Im Schatten der Energetiker

K. Kroy

Leipzig, 2015*

---

*The script is not meant to be a substitute for reading proper textbooks nor for dissemination. Comments and suggestions are highly welcome.

Contents

> The ability to coarse-grain has been our most important tool to bring an infinitely complicated world within reach of analysis and comprehension
>
> Eric Smith

What it is all about

The main subject of the lecture are systems with too many microscopic degrees of freedom to treat them all in detail. The systems of interest typically display a scale separation (in space and/or in time) between those many uninteresting degrees of freedom and a few interesting macroscopic variables, though. This property is not limited to physical systems, which explains the extraordinary interdisciplinary interest in the concepts of Statistical Mechanics. Essentially three reduced ("coarse grained") levels of description are commonly used and will be introduced here: *Thermostatics*, which tries to establish a deterministic theory, much like classical continuum mechanics, on the basis of the static equilibrium of a few macroscopic variables alone; *Thermo-/Hydrodynamics*, which considers that these variables may slowly vary in space and time, thereby causing irreversible fluxes that increase the total entropy and restore the global static equilibrium; finally *Stochastic Thermodynamics*, which deals explicitly not only with deterministic variables, but also with some of the fluctuations arising through their coupling to the many uninteresting microscopic degrees of freedom. Broadly speaking, *Statistical Mechanics* comprises all attempts to address the question how simple deterministic (and stochastic) equations for a few variables and their fluctuations emerge from the underlying microscopic mess of a high dimensional phase space. In fact, it includes also those quite ambitious studies trying to establish how a system becomes messy (apparently stochastic) in the first place, starting from the deterministic Liouville equation for the microscopic motion or its quantum version, the von-Neumann equation. This interesting question will be given little attention, here. The question pursued instead is: what can be gained by postulating that the motion of the microscopic degrees of freedom of a system in thermodynamic equilibrium can be considered as structureless noise as far as the construction of coarse-grained descriptions is concerned? Boltzmann, Maxwell and Gibbs were among the first to realize the power of this postulate in deriving the highly universal laws governing the macroscopic order that emerges by virtue of the high dimensionality (the law of large numbers): simple laws that become exact statements about strongly interacting many particle systems in the thermodynamic (hydrodynamic) limit.

Part I
Thermodynamics

> Thermodynamics is a funny subject: The first time you go through it, you don't understand it at all. The second time you go through it, you think you now understand it, except for one or two little details. The third time you go through it, you realize you understand close to nothing, but by that time you are so used to it that it doesn't bother you any more.
>
> attributed to Arnold Sommerfeld

1 The Birth of Thermostatics: Out of the Spirit of Mechanics

The program is to develop Thermostatics following the paradigm of continuum mechanics, but extending the concepts used there by allowing for the reversible[1] conversion of mechanical energy into heat and vice versa. Beyond its historical relevance for the mathematical description of heat engines, this development gives a glimpse of some basic concepts of great generality, from which a basic understanding of the conditions under which viable coarse-grained theories (which comprise all known physical theories) can be constructed, in, near, and far from equilibrium, is emerging.

1.1 Basic notions

The notions of *equilibrium, work, energy, potentials, extensive and intensive* (proportional/independent of system size) *state variables, state space* and *equation of state* are familiar from mechanics. The thermodynamic state space contains apart from a relevant set of a few mechanical (electrical, magnetic, whatever) state variables also a variable called *entropy* S. This will be sort of established

---

[1] Irreversible processes that increase the entropy beyond what is provided by heat influx are the subject of irreversible thermodynamics or generalized hydrodynamics, which explicitly deals with spatio-temporal gradients in the thermostatic variables.

after taking the reader's everyday working knowledge about heat transfer, heat isolation, and temperature for granted (i.e., there is no attempt to achieve axiomatic rigor[2]).

Fundamental relations, equations of state, integrability

As for any mechanostatic, hydrostatic, electrostatic, magnetostatic, etc. system, the equilibrium state can be specified by a so-called fundamental relation that describes a hypersurface $Y(\{X_i\})$ in a state space spanned by $m+1$ macroscopic state variables $\{Y, X_i\}$. This fundamental relation is not itself provided by the macroscopic theory, but is a material or constitutive relation characterizing a particular system, usually in the form of an energy function. It is a major task of statistical mechanics and condensed matter physics to provide and classify such relations. The special new thing in thermostatics is that the notion of heat is added to the other forms of energy mentioned above. The differential of the fundamental relation[3]

$$dY = \sum_i \xi_i \mathrm{d}X_i \ , \qquad \xi_i(\{X_i\}) \equiv (\partial_{X_i} Y)_{X_{j\neq i}} \tag{1}$$

is itself fully specified by the m equations of state $\xi_i(\{X_i\})$. Taking another derivative leads to m^2 (specific) susceptibilities χ_{ij} or response coefficients

$$\chi_{ij}^{-1} = X_i \partial_{X_i} \xi_j = X_i \partial_{X_i} \partial_{X_j} Y \ . \tag{2}$$

In practice, it is the susceptibilities that are measured and from which one infers the equations of state, while integrating the latter up to obtain a full fundamental relation is rarely achieved outside the realm of theoretical toy models. Note that neither the equations of state nor the susceptibilities are independent of each other. For $\mathrm{d}Y$ to be a total differential, one needs its mixed second derivatives to be equal, namely

$$\partial_{X_j} \xi_i = \partial_{X_i} \xi_j \qquad \text{Maxwell integrability relations} \tag{3}$$

To give a standard example, consider the fundamental relation $\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V$,

$$(\partial_V T)_S = \partial_V \partial_S U = -(\partial_S p)_V \ . \tag{4}$$

Equivalently, you could start from $\mathrm{d}S$ or other potentials obtained from S or U by Legendre transforms (see below), which considerably constrains the possible forms of response coefficients and equations of state.

---

[2] See recent work by Lieb and Yngvason, if that is what you are after.

[3] The $X_{j\neq i}$ are constant when taking the partial derivative w.r.t. X_i.

Examples

An ideal elastic spring of rest length L, spring stiffness k has two state variables x (extension, extensive in L), f (tension, intensive), and the equation of state $f = kx$. In contrast to k and its inverse, the extensive susceptibility $k^{-1} \propto L$, the *specific* spring stiffness $K = kL$ is an *intensive response coefficient*, independent of L, which characterizes the material and basic design of the spring. If the spring is extended gently, so that the internal tension f is always spatially constant throughout the spring and equal to the external force f_{ext}, the work $\delta W = f_{\text{ext}}\mathrm{d}x$ can be integrated to give a fundamental relation, namely the (extensive) state function known as internal elastic energy $E_e = kx^2/2$ as a function of x. In this case the work increment is the total differential of the elastic energy $\mathrm{d}W = \mathrm{d}E_e = kx\mathrm{d}x$. Otherwise one has a nonequilibrium situation and this is not the case.

A real spring dissipates some (positive) part δW_{diss} of the work δW it receives as irretrievable heat into its environment,

$$\delta W = f_{\text{ext}}\mathrm{d}x = \mathrm{d}E_e + \delta W_{\text{diss}} = kx\mathrm{d}x + \delta W_{\text{diss}} \tag{5}$$

and the total work $\delta W > \mathrm{d}E_e$ is, in contrast to the elastic energy, not a total differential. The same happens for the reverse operation (spring contraction): some of the elastic energy is dissipated as heat and an even smaller part is provided as work back to the external world.

A rubber band seems superficially very similar, but it heats and cools considerably upon (fast) extension and release. Hence a very substantial amount of the external work δW (if not all of it) is converted into heat δQ and vice versa,

$$\delta W = f_{\text{ext}}\mathrm{d}x = \mathrm{d}E_e - \delta Q\ . \tag{6}$$

We still know what the external work is but are *a priori* uncertain how it is distributed between heat and elastic energy inside the rubber, and whether or not it is a total differential. The observed cooling suggests that, in contrast to the spring, heat once released can later help in the contraction, so that despite of the substantial heat flows we might reversibly gain all of our mechanical work back, which would then again be expressed as a total differential. (In this case it would not be appropriate to call the heat "dissipated".)

An ideal rubber band is a such a perfect rubber, in which the total response is attributed to the conformational change of some idealized flexible polymers. The filler, in which these polymers are embedded, is assumed to be mechanically inactive but to act as an ideal heat reservoir. The external work is completely transformed into an *entropy reduction* in the mechanically active polymers (microscopically: stretching restricts the available microstates for their internal degrees of freedom), not into an elastic energy increase. The internal energy U of the polymers does not change, $\mathrm{d}U = \mathrm{d}E_e = 0$, and all work is transformed into heat that flows from the polymer degrees of freedom into the filler material. For reversible isothermal ($\mathrm{d}T = 0$) state changes the complete heat $\delta Q = T\mathrm{d}S$ can be recovered as work

$$f_{\mathrm{ext}}\mathrm{d}x = \delta W = \mathrm{d}U - \delta Q = \mathrm{d}F \tag{7}$$

which is in this case given by the total differential of a potential called the free energy $F = U - TS$. This is the part of the energy that can be recovered as work under isothermal equilibrium conditions and takes over the role played by the potential energy in mechanics. Note that (in contrast to the dissipation in the real spring) the heat in Eq. (7) is not lost. It flows back into the polymers when the rubber band contracts again. There is "free" energy available to perform work, although there is no potential energy stored.

Solids, Fluids, Magnets can be discussed along similar lines. For a solid, $\delta W = \int \mathrm{d}V\, S_{ij}\delta e_{ij}$ with S_{ij} and e_{ij} the elements of the second Piola-Kirchhoff stress tensor and the Green-Lagrange strain tensor, respectively. Both are related linearly in a popular approximation of the equation of state known as Hooke's law, valid for small deformations, via a fourth rank elasticity tensor E_{ijkl} with (depending on the degree of symmetry) 2 to 21 independent material constants. That helps you understand why you should love fluids, gases and liquids, for which the work is simply $\delta W = -p\mathrm{d}V$ (p pressure, V volume) and magnets with $\delta W = H\,\mathrm{d}M$. For the latter, H is the somewhat misleading but common notation for the external magnetic field B_{ext} (see Brenig's book for a careful discussion), M the magnetic moment (often called magnetization), and $\mathcal{M} \equiv M/V$ the magnetization (referred to as the magnetization density by those otherwise running into notational trouble).

Particle exchange is not limited to but common in chemical and nuclear reactions. The chemical work is $\delta W = \mu_i \mathrm{d}N_i$ for N_i particles of species i and chemical potential μ_i. So particles do not come and go for free. Moreover, they usually carry heat ("convection").

A little dictionary

Equilibrium is a state not changing in (any reasonable) time. Following the mechanical paradigm, the notion of thermostatic equilibrium implies that the intensive parameters (like the spring tension) are constant throughout.

Temperature T is the intensive parameter regulating the exchange of heat much as tension regulates the exchange of extension in a spring.

Isolated systems do not exchange anything (no particles, work or heat) with their surroundings, so that their internal energy U stays constant, $\mathrm{d}U = 0$.

Closed systems allow energy to be exchanged (heat and work) but not particles.

Thermostats or heat reservoirs take up or give away finite amounts of heat without changing their temperature.

Quasi-static processes are required to be locally in equilibrium, i.e. to be locally reversible processes.

Reversible processes are quasi-static processes that proceed so gently that the system remains globally in equilibrium with itself and exchange partners (intensive variables are well-defined and constant throughout the system and its partners at any time). Reverting the protocol takes you back to start in the thermodynamic state space.

Isothermal processes live on manifolds called isotherms in state space, where the system stays in contact with a heat reservoir that keeps the temperature constant and equal to the reservoir temperature, i.e. $\mathrm{d}T = 0$.

Adiabatic processes live on adiabatic manifolds where they do not exchange heat, $\delta Q = 0$ if no particle exchange occurs (while the name convention in mechanics rather refers to slow processes conserving the energy, the action, the phase space volume, or the like).

The Carnot process is a reversible cycle with two adiabatic and two isothermal sub-processes that can be used to convert some heat into work or vice versa.

The efficiency of a thermodynamic machine that turns heat into (mechanical) work is defined as

$$\eta \equiv |\Delta W| / |\Delta Q_+| \;, \tag{8}$$

the fraction of the net work $-\Delta W$ provided to the outside world to the heat ΔQ_+ supplied to the process through the hot reservoir.

Response coefficients (or "functions", once you realize they are actually not constant if measured over a sufficiently wide parameter range) control how a system's state variables respond to external perturbations. These are obviously nice observables to characterize a system. Some examples are the isothermal/adiabatic *compressibility* $\kappa_X \equiv -(\partial_p V)_X / V$ $(X = T,\ S)$, the thermal *expansion coefficient* $\alpha \equiv (\partial_T V)_p/V$, or the magnetic susceptibility $\chi \equiv \partial_H \mathcal{M}$. If normalized to be intensive, they are called "specific". A purely thermal response coefficient is the *heat capacity*

$$C_X \equiv (\delta Q/\delta T)_X \tag{9}$$

or, after division with mass, volume or particle number, the *specific heat* c_X. It measures how much the temperature changes upon heat exchange — a large specific heat meaning that much heat must be added to achieve a small temperature increment, or much heat can be taken out before the temperature falls considerably, respectively (as it should be for a good heat storage device). The specific heat clearly depends on what the system is doing (e.g. working or resting etc.) during the heat exchange, as notified by an index X for each variable held constant.

Sign convention: things added to a system are counted as positive; some people insist that it is positive if the system performs some work, instead.

The fundamental laws of thermodynamics are energy conservation (first law) and entropy para-conservation (second law), but people with a strong passion for complete collections make you feel ashamed if you do not know at least four laws (including a zeroth law saying that temperature exists and a third law saying that things tend to behave increasingly orderly if frozen).

1.2 The laws of thermostatics

The zeroth law ("thermal equilibrium exists")

postulates that thermal equilibrium is transitive[4] and mediated by an intensive state variable, the empirical temperature. If equilibrium exists and is transitive, one can bring three systems into mutual equilibrium, one of which acts as a thermometer for the other two. The mechanical paradigm for temperature is the tension in a spring. This intensive quantity has to be constant throughout and at mechanical contacts with other bodies, in mechanical equilibrium.

The first law ("heat is energy")

says that for any system *a state function U called internal energy can be defined, with its change $\mathrm{d}U$ in an infinitesimal state change being a total differential equal to the net energy influx (counting all forms of energy including heat).*

For a closed system (no matter exchanged) the energy influx is divided into the absorbed heat δQ and work δW,

$$\mathrm{d}U = \delta Q + \delta W\,. \tag{10}$$

---

[4] If A is in equilibrium with B, and B with C, then A and C are in equilibrium, too.

For $\delta Q = 0$, the process is said to be adiabatic and the work can be expressed as a total differential $\mathrm{d}U$. Note that the first law holds generally but that only for reversible processes something like $\delta W = -p\mathrm{d}V$ holds, where p is "the" pressure in the fluid. (Generally, we only might know that $\delta W = -p_{\mathrm{ext}}\mathrm{d}V$.) Also note that the separation into heat and work is not an entirely objective concept. One person's heat might be another (more sophisticated) person's work (similarly as one person's noise might be another person's music). Speaking from the molecular perspective of statistical mechanics, it matters which degrees of freedom of the system can actually be accessed to systematically extract "useful work".

Historically, heat was long thought to be a substance ("chaleur", "calorie"), which is maybe not entirely incomprehensible if you think of convection and of how heat is provided to our bodies. Count Rumford, who is also remembered for the Rumford soup and the English garden in Munich, demonstrated at the end of the 18^{th} century that work can be transformed into heat. Canon drilling horses could make the cooling water cook as much as one liked without having any direct access to the water. Rumford concluded that heat must be a form of motion of materials but could not provide a consistent formalism to replace the conventional theory.

The first law may be used to express the Carnot efficiency Eq. (8) in terms of the exchanged heat alone. Integrating Eq. (10) over the full cycle gives

$$\oint \mathrm{d}U = 0 \qquad \Rightarrow \qquad -\oint \delta W \equiv -\Delta W = \Delta Q \equiv Q_+ + Q_- = |Q_+| - |Q_-| \quad (11)$$

with $Q_\pm$ the heat absorbed with the hot/cold thermostat, hence $\eta = 1 - |Q_-|/|Q_+|$.

The second law ("heat is inferior energy")

can be formulated similarly as the first by saying that for any system *a state function S, called entropy, can be defined, with its change $\mathrm{d}S$ in an infinitesimal state change being a total differential equal to the net entropy influx (equal to the net heat influx divided by the reservoir temperature) plus some non-negative amount $\delta S_{\mathrm{irr}} > 0$ spontaneously created if the process is not reversible.*

The historical formulations of the second law, developed by Carnot, Mayer, Joule, and Clausius in the early 19^{th} century are: *no periodically working process can have the sole effect of transferring heat from the cold to the warm (Clausius) or of turning heat completely into work (Kelvin).*

Universal absolute temperature

Exploiting the Carnot cycle, a universal (up to a scale factor) temperature scale can be established on this basis, which then implies that the Carnot efficiency can be expressed by the two reservoir temperatures T_+ and T_- alone. To prove this, suppose the opposite, namely that the exchanged heat was not tied uniquely and universally to temperature. Then one could have two *ideal reversible* Carnot cycles operating in opposite sense between the same heat reservoirs, the heat pump providing the heat Q_+ consumed by the heat engine, the heat engine providing the work W consumed by the heat pump. Now, the second law in Kelvin's formulation could be violated if the heat engine could be chosen such as to provide less heat to the lower reservoir and more work than consumed by the heat pump. Hence, *at least the fraction Q_-/Q_+ entering the optimum*[5] *efficiency must be a universal function of the corresponding temperatures*[6], say $f(T_-, T_+)$. But then, replace one of the Carnot engines by two, one working between T_+ and T_i and the other between T_i and T_-, so that

$$|Q_+|/|Q_i| = f(T_+, T_i), \quad |Q_i|/|Q_-| = f(T_i, T_-), \quad |Q_+|/|Q_-| = f(T_+, T_-)\,, \tag{12}$$

and, after multiplying the first two equations,

$$f(T_+, T_-) = f(T_+, T_i) f(T_i, T_-)\ . \tag{13}$$

One concludes that $f(T_1, T_2) = \varphi(T_1)/\varphi(T_2)$, so that there is up to an overall constant (setting the temperature unit) a unique universal temperature scale $\varphi(T)$ with a unique absolute zero point. For simplicity, it shall still be called T. From this, we conclude:

The universal Carnot efficiency

The Carnot efficiency is a unique and universal function of its working temperatures,

$$\eta = 1 - |Q_-|/|Q_+| = 1 - T_-/T_+\ . \tag{14}$$

This statement can be understood as a condensed form of the first and second law for reversible processes. In fact, we can rewrite it as $Q_+/T_+ + Q_-/T_- = 0$. Now, exploiting that every arbitrary thermodynamic cycle can be approximated by little Carnot cycles (sketch), we have *for all ideal reversible cycles*

Carnot's theorem

$$\oint \delta Q/T = 0\,, \qquad \text{hence} \qquad \mathrm{d}S \equiv \delta Q/T \tag{15}$$

defines a total differential of a state function S called **entropy**. The inverse temperature $1/T$ acts as a so-called integrating factor that turns an incomplete differential δQ into a total differential $\mathrm{d}S$.

---

[5] It is of course always allowed to operate a device at lower efficiency, the second law only limits the optimum efficiency reached in a completely reversibly operating ideal process.

[6] The zeroth law provides a sufficient empirical basis for speaking of temperatures.

The entropy is the new potential that extends the mechanical (etc.) state space for thermodynamic purposes and allows for a self-contained reduced description that accounts for work being transformed into heat (that flows into unresolved degrees of freedom). Reversible state changes live on hyper-surfaces defined by the *fundamental relation* $S(\{X_i\})$ in the state space spanned by $\{S, X_i\}$. In particular, for reversible state changes of a simple $p - V$ system $X_1 = U$, $X_2 = V$, and the combined first and second law can be written as

$$\mathrm{d}S(U,V) = T^{-1}\mathrm{d}U + pT^{-1}\mathrm{d}V \ . \tag{16}$$

or, in engineering style (emphasizing energy),

$$\mathrm{d}U(S,V) = T\mathrm{d}S - p\mathrm{d}V \ . \tag{17}$$

It is essential to understand that the whole discussion of thermodynamic processes so far is completely independent of the specific technical realization (e.g. the materials used, etc) of the hypothetical heat engines. It requires *perfectly reversible behavior*, however, i.e., in a sense it describes an upper bound of what is possible in real life. It is certainly of great interest to additionally know how the imperfections of the real world will change the picture. According to the everyday experience echoed by the second law, all that can happen is that some work is wasted into heat, such that the efficiency decreases, as stated by

Clausius' theorem

$$\oint \delta Q/T \le 0 \ . \tag{18}$$

Note that for irreversible processes the reservoir temperature T appearing in this formula can no longer be identified with the system temperature, which is in general not well-defined. Only the temperature of the bath is well-defined (though not generally constant) throughout the process. To prove Eq. (18), it is enough to realize that the Carnot cycle may always be operated at a less than optimum efficiency, so that

$$\eta = 1 - |Q_-|/|Q_+| \le 1 - T_-/T_+ \ , \tag{19}$$

which says that less heat is turned into work than would be allowed by the second law. Accordingly, one has $|Q_-|/T_- \ge |Q_+|/T_+$ or

$$Q_+/T_+ + Q_-/T_- \le 0 \quad \Box \tag{20}$$

For a more general formulation consider any two states (T_1, V_1) and (T_2, V_2), say, connected by paths Γ in the $T - V$ plane

$$S(T_2,V_2) - S(T_1,V_1) \ge \int_\Gamma \frac{\delta Q}{T} \qquad \text{hence} \qquad dS \ge \delta Q/T \ . \tag{21}$$

Some remarks

- Isolated systems: Thermodynamic state changes of an isolated system are necessarily irreversible (which thermodynamic function should change in a reversible state change?), increasing the entropy until equilibrium is reached. From Eq. (21), $dS \geq 0$ for $\delta Q = 0$. *The entropy of an isolated system attains a maximum in equilibrium.* (The 2^{nd} law for an isolated system.)

- $\eta \neq 1$, since reversible heating at $T = 0$ is impossible due to $T\mathrm{d}S = 0$.

- The mismatch between the entropy $\delta S_{\text{ext}} = \delta Q/T_R$ released from the thermostat at temperature T_R and the system's total entropy change $\mathrm{d}S$ can of course always be given a name, δS_{irr} ("entropy production"), so that the second law takes the form

$$\mathrm{d}S = \delta S_{\text{ext}} + \delta S_{\text{irr}} = \delta Q/T_R + \delta S_{\text{irr}} \ , \qquad \delta S_{\text{irr}} \geq 0 \ . \tag{22}$$

  For weakly irreversible quasi-static state changes with local equilibrium, δS_{irr} can be quantified by including additional internal variables into the thermodynamic state space, thereby turning the Clausius inequality back into an equality and giving the symbol T, at least locally, a meaning for the system itself. Generally, an inequality indicates that some of the sucking internal degrees of freedom are not explicitly resolved.

- The so called *fluctuation theorems* quantify how violations of the second law vanish with system size, making it a strict law in the thermodynamic limit (special care is needed for phase separating and non-extensive, e.g. gravity-dominated, systems).

- Our everyday experience with entropy is only partially related to the second law of thermostatics as formulated above. The rich structures and processes on earth including all forms of life are due to an entropy gradient established by a persistent enormous net negative entropy influx from the sun (less hot photons in than cold photons out — as dictated by energy conservation and surface temperature), which keeps the surface of the earth far from equilibrium. The second law as a statement about the stability of equilibrium can only partially or locally be brought to bear onto this experience.

The third law ("quantum nature of absolute zero")

(postulated on empirical grounds by Nernst 1905) says that *the entropy at absolute zero vanishes*[7], $S(T = 0) = 0$.

---

[7] It suffices that it is a non-extensive universal constant that may be set to zero.

This is a macroscopic manifestation of quantum statistics, which is best understood by studying the specific heat of quantum mechanical model systems. The third law requires that it (as well as some other thermal response coefficients) has to vanish sufficiently quickly that $S(T,X) = \lim_{T_0\to 0}\int_{T_0}^{T} dT'\, C_X/T'$ vanishes for $T \to 0$ — independently of X. The latter condition implies that absolute zero cannot be reached by a finite number of state changes, since one would have to cycle the system between $S(T, X = X_1)$ and $S(T, X = X_2)$ for some given variable X (see sketch!). Or, as noticed above, there is no reversible heating from (and thus no reversible cooling to) $T = 0$, where $T\mathrm{d}S = 0$.

Many practically (sometimes even at very low temperature) very useful classical models for fluids, solids, magnets etc. violate the third law, but in principle any system should always be expected to settle at very low temperature into some coherent wave function that has no (or at least non-extensive) entropy or ground-state degeneracy.

Summary of the laws

0) You have to play the game
1) You can't win
2) You can't break even, except on a very cold day
3) It doesn't get that cold

1.3 Stability, Homogeneity

Thermodynamic stability

The second law expresses the stability of thermal equilibrium. This can be made explicit in a differential form and related to the sign of the response coefficients. The essential trick, known as d'Alembert's principle, can be borrowed from mechanostatics. It consists in artificially dividing a homogeneous system into two subsystems 1 and 2 and considering a repartition of the extensive quantities, i.e. taking some small amount $\delta X_i \equiv \delta X_i^{(1)} = -\delta X_i^{(2)}$ from one to the other subsystem. In equilibrium this should leave the relevant effective energy function Y (which for an isolated system happens to be the negative entropy $-S$, for an isothermal system the free energy F, and so on) unchanged to leading order, which is indeed the case for homogeneous intensive parameters. In general notation, with $\xi_i^{(k)} = \partial_{X_i} Y^{(k)}(\{X_i^{(k)}\})$,

$$0 = \mathrm{d}Y = \mathrm{d}Y^{(1)} + \mathrm{d}Y^{(2)} = \sum_i (\xi_i^{(1)} - \xi_i^{(2)})\delta X_i \quad \Leftrightarrow \quad \xi_i^{(1)} = \xi_i^{(2)} \,. \tag{23}$$

To second order in the perturbation, the repartition must cause a penalty for the equilibrium to be stable, which corresponds to the *convexity* or positive-definite matrix of second derivatives of the relevant energy function, so that the response acts back against its cause. This is of course nothing but the formalization of the intuitive picture of a system living for all possible boundary conditions (isolated, closed, open,…) near the bottom of an appropriately defined potential valley. Think of the ideal spring, where $Y = E$, $X = x$, $\xi = -f$, $\chi^{-1} = -\partial_x f = k$, or an isentropic gas with $Y = U$, $X = V$, $\xi = -p$, $\chi = V\kappa_T = -(\partial_p V)_T$ as purely mechanical examples. Generally, in one dimension,

$$\begin{aligned} \mathrm{d}Y &= \mathrm{d}Y^{(1)} + \mathrm{d}Y^{(2)} \\ &= (\partial_X Y^{(1)})\delta X + (\partial_X Y^{(2)})(-\delta X) + (\partial_X^2 Y^{(1)} + \partial_X^2 Y^{(2)})(\delta X)^2/2 = \\ &= \underbrace{(\xi^{(1)} - \xi^{(2)})}_{=0 \text{ in equil.}} \delta X + \underbrace{(\partial_X \xi^{(1)} + \partial_X \xi^{(2)})}_{\partial_X \xi^{(1)} = \chi^{-1} = \partial_X \xi^{(2)}} (\delta X)^2/2 = \underbrace{(\partial_X^2 Y)}_{=\chi^{-1}} (\delta X)^2 > 0 \end{aligned} \tag{24}$$

In the final step the irrelevant reference to the arbitrary subdivision of the system was omitted. Spontaneous repartitioning of the extensive variables is thus suppressed in equilibrium if the response coefficients have the right sign. If imposed externally, a repartitioning relaxes spontaneously by decreasing an appropriate free energy (or the negative entropy). The case of a relaxing isothermal system lowering its free energy corresponds to our most common everyday experience, and the related increase in entropy links it to the above formulation of the second law.

To discuss a higher-dimensional example that includes a mechanical and a thermal sector, take an isolated $p - V$ system ruled by Eq. (16) and divide it artificially into two subsystems $i = 1, 2$. The relevant potential valley is now the negative entropy, $Y = -S$. To leading order, it is insensitive to an internal reshuffling of the extensive variables U and V. To provide stability, it has to increase to second order,

$$-(\mathrm{d}S)_{U,V} = \frac{1}{2}(\delta U, \delta V) \begin{pmatrix} -(\partial_U^2 S)_V & -\partial_U \partial_V S \\ -\partial_V \partial_U S & -(\partial_V^2 S)_U \end{pmatrix} \begin{pmatrix} \delta U \\ \delta V \end{pmatrix} > 0 \, . \tag{25}$$

So the Hesse matrix (of $-S$) must be positive definite, which requires all subdeterminants to be positive definite, i.e.

$$0 < -(\partial_U^2 S)_V = \frac{1}{T^2 C_V} \quad \Leftrightarrow \quad C_V > 0 \tag{26}$$

and[8]

$$0 < (\partial_U^2 S)_V (\partial_V^2 S)_U - (\partial_U \partial_V S)^2 = \frac{1}{T^3 V \kappa_T C_V} \quad \Leftrightarrow \quad \kappa_T > 0 \tag{27}$$

Note that the identification of stability with positive response coefficients requires the parameter T^{-1} that controls the repartition of the internal energy U to be

---

[8] Apply the Jacobi-determinant chain rule to $\partial(1/T, p/T)/\partial(U, V)$ to get the second equality.

positive. While this is usually the case, it may sometimes (though not for thermodynamic efficiency considerations) make sense to characterize a system under extreme nonequilibrium conditions by a negative temperature $T < 0$. A simple example is provided by a paramagnet, a collection of (independent) magnetic moments which are strongly oriented by an external field that is suddenly reversed. The state space explored by the microstate of the system, and hence its entropy, clearly increases[9] when the moments start to de-correlate to lower their energy. So for a certain time before the moments all settle into the new direction of the field, the derivative $\partial_S U$ of energy with respect to entropy is negative, hence $T < 0$. The system is thus temporarily (magnetically) unstable although the response coefficients are positive.

Homogeneity, extensivity

As familiar from continuum mechanics, hydrodynamics, etc., it turns out that there is a large class of interesting systems that are homogeneous or extensive, with their energy and entropy scaling linearly in N, which is usually[10] equivalent to saying that they are trivially decomposable into additive subsystems. For example, half of the content of a bottle of water contains not only half of the atoms but also half of the energy, entropy, etc. The same cannot be said about a neutron star, an atomic nucleus, or a protein that you cut into halves, though, because in these system the interactions are not short ranged. Many of the standard tricks you are commonly supposed to learn in this lecture either need special care or do not work at all for these systems, as they rely crucially on dividing a system into weakly coupled subsystems. This is why so-called "Small" systems are treated as exceptional cases in the common folklore, including this text[11].

Now, given the restriction that *all of the systems to be discussed in the following are extensive/homogeneous*, it is interesting to analyze the general thermodynamic consequences. Mathematically, the homogeneity of the system is formulated by requiring for a fundamental relation $Y(\{X_i\})$ written in terms of *extensive* quantities $\{Y, X_i\}$ that

$$\lambda Y(\{X_i\}) = Y(\{\lambda X_i\}) \,. \tag{28}$$

Taking the derivative with respect to λ at $\lambda = 1$ gives the *Euler relation*

$$Y = \sum_i (\partial_{X_i} Y) X_i \equiv \sum_i \xi_i X_i \,, \tag{29}$$

---

[9] This explanation appeals to a microscopic interpretation that anticipates major insights from statistical mechanics. For the time being, just swallow it.

[10] See H. Touchette, Physica A, 2002 vol. 305 pp. 84-88 for a more careful discussion.

[11] But see the book by Dieter Gross on *Microcanonical Thermodynamics*.

which tells us that the equations of state $\xi(\{X_i\})$ are equations for the *intensive* variables. From the last equation we have

$$\mathrm{d}Y = \sum_i \xi_i \,\mathrm{d}X_i + \sum_i X_i \,\mathrm{d}\xi_i \; . \tag{30}$$

Comparison with the total differential

$$\mathrm{d}Y(\{X_i\}) = \sum_i \xi_i \mathrm{d}X_i \tag{31}$$

of the fundamental relation reveals that the second sum in Eq. (30) has to vanish,

$$\sum_i X_i \,\mathrm{d}\xi_i = 0 \; . \tag{32}$$

This is known as the *Gibbs–Duhem* relation. For illustration, consider the case of a simple fluid. Starting from the differential form of the fundamental relation

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N \tag{33}$$

the *Gibbs-Duhem* and *Euler* relations take the forms

$$S\mathrm{d}T - V\mathrm{d}p + N\mathrm{d}\mu = 0 \qquad \text{or} \qquad \mathrm{d}\mu = -s\mathrm{d}T + v\mathrm{d}p \; , \tag{34}$$

$$U = TS - pV + \mu N \qquad \text{or} \qquad \mu = u - Ts + pv \tag{35}$$

with

$$s \equiv S/N \qquad v \equiv 1/n \equiv V/N \qquad u \equiv U/N \; . \tag{36}$$

As a consequence, of the m equations of state in a $m+1$ dimensional thermodynamical state space only $m-1$ can be independent, even before invoking the Maxwell integrability conditions, which provide further, more general constraints. Moreover, the response coefficients are seen to be interrelated. For example, if p is interpreted as a function of T and n, so that $\mathrm{d}p = (\partial_T p)_n \mathrm{d}T + (\partial_n p)_T \mathrm{d}n$, Eq. (34) is readily seen to imply,

$$\kappa_T \equiv n^{-1}(\partial_p n)_T = n^{-2}(\partial_\mu n)_T \tag{37}$$

for the isothermal compressibility.

1.4 Thermodynamic potentials

It was mentioned during the foregoing discussion that there are a number of equivalent formulations of the fundamental relation, their usefulness depending on the boundary conditions applied to the system. To each of the possible external conditions (fixed energy, fixed temperature, fixed volume, fixed pressure, fixed magnetization, fixed external field, etc.) [sketch] suitable generalized energy functions can be defined that all reflect, *for their extensive variables*, our intuition of an energy valley, in which the system slides down to the minimum.

In classical mechanics, say, a system that is not at the energy minimum would have to oscillate indefinitely. Only by introducing dissipation, the minimum becomes attractive. But then one automatically deals with a (generalized) free energy, which is actually the more accurate representation of our everyday experience of energy valleys. All these generalized free energies inherit their stability properties from the fact that the entropy of an isolated system attains its maximum in equilibrium, i.e. from the second law. The common notion of stability of energy valleys, even if familiar from classical mechanics, can thus not simply be taken over from there, because its use relies on a genuinely thermodynamic justification. One may say that the fox is put in charge of the henhouse, since, somewhat paradoxically, dissipation (or "energy loss") is the ordering principle to which the energy picture owes its stabilizing power.

Mathematically, the conversion between the different potentials is performed with the help of a *Legendre transformation*, which exchanges the roles of "slope" ξ and "variable" X in the differential $\mathrm{d}Y = \xi \mathrm{d}X$ of a convex function $Y(X)$. Namely,

$$Y(X) \to Z(\xi) \equiv Y - \xi X \quad \text{with } \xi \equiv \partial_X Y \quad \text{so that } \mathrm{d}Z = -X\mathrm{d}\xi \tag{38}$$

where all other variables of Y are kept fixed in taking the derivative with respect X. The Legendre transform thus represents a one-to-one recoding of thermodynamic potentials, which are *convex*[12] *functions*, in terms of their tangents. It thereby exchanges extensive and intensive variables. Each transformation replaces one convex extensive sector by a concave intensive one, indicating, so to say, that the responsibility for the stability with respect to this variable has been delegated to a larger system.

As a simple mechanical example, consider the case of the ideal spring with a mass m attached under gravity, which exerts a force $-mg$. The equilibrium position x_0 is now not any more given by the minimum of the internal elastic energy $U(x) = kx^2/2$ (namely $x_0 = 0$), but by that of the enthalpy $H(mg) \equiv U - mgx$ (namely $x_0 = mg/k$), which is the sum of the elastic energy and the gravitational potential energy. Also observe that the equilibrium enthalpy $-mg^2/(2k)$ happens to be concave with respect to the intensive variable mg. As emphasized above, this is not a reason to panic, since it is the convexity with respect to extensive quantities that matters for thermodynamic stability.

Following the engineers' preference for energies, one usually starts the discussion of thermodynamic potentials—somewhat inappropriately—by expressing energy and enthalpy for gases and liquids in the canonical form, as follows

---

[12] Consider $-S$ instead of S.

Internal energy

$$U(S,V,N) = TS - pV + \mu N \qquad \mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N\ . \tag{39}$$

In contrast to the entropy S, U does not *per se* obey an extremum principle, as it is conserved for an isolated system. Alternatively, people who are used to cooking with open pots exposed to ambient pressure (e.g. chemists) prefer to start from

Enthalpy

$$H(S,p,N) \equiv U + pV = TS + \mu N \quad \Rightarrow \quad \mathrm{d}H = T\mathrm{d}S + V\mathrm{d}p + \mu\mathrm{d}N \tag{40}$$

Now, the really interesting physics (the second law) comes in with

Entropy

$$S(U,V,N) = \frac{1}{T}U + \frac{p}{T}V - \frac{\mu}{T}N \qquad \mathrm{d}S = \frac{1}{T}\mathrm{d}U + \frac{p}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N\ , \tag{41}$$

and in principle one should of course start the whole discussion from here. It is the entropy, which for an isolated system ($\mathrm{d}U = \mathrm{d}V = \mathrm{d}N = 0$) takes its maximum in equilibrium and embodies thermodynamic stability by $-S$ being convex in its extensive variables (which is the pertinent formulation of the 2^{nd} law for isolated systems). By discussing Legendre transforms of S, one would then obtain a series of potentials known as *Massieu potentials*, which clearly reveal the underlying structure of thermodynamics. However, the conventional practice reflects the Energetikers' preference for energies, which most people find more intuitive, and starts from U and H, instead.

From U, the first Legendre transformation then takes one to the

Helmholtz free energy

$$F(T,V,N) \equiv U - TS = \mu N - pV \quad \Rightarrow \quad \mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \mu\mathrm{d}N \tag{42}$$

It is *minimal for an isothermal system of fixed volume and particle number and convex in its extensive variables* — precisely what our intuition requires from an energy valley — which provides the formulation of the 2^{nd} law pertinent to isothermal systems. So the free energy is the legitimate heir[13] of entropy for an isothermal system ($\mathrm{d}T = 0$) and at the same time satisfies the engineers' preference for energies, which makes it the most popular kid in town.

Analogously, starting from H one arrives at the

---

[13] While starting from U did conceal the underlying logic, F did nicely inherit its convexity from $-S$, just as it should be.

Gibbs free energy (free enthalpy)

$$G(T,p,N) \equiv U - TS + pV = \mu N \quad \Rightarrow \quad \mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \mu\mathrm{d}N \tag{43}$$

which takes the role of F at constant pressure instead of constant volume. So G is minimal for a chemically closed isothermal isobaric system in equilibrium[14].

Finally, if the volume is kept fixed for an isothermal system but matter can be exchanged at a constant chemical potential, the

Grand canonical potential

$$J(T,V,\mu) \equiv U - TS - \mu N = -pV \quad \Rightarrow \quad \mathrm{d}J = -S\mathrm{d}T - p\mathrm{d}V - N\mathrm{d}\mu \tag{44}$$

attains a minimum in equilibrium. While particle exchange at fixed volume may sound a bit odd, at first, it is just the perfect starting point for formulating hydrodynamics. Space is virtually divided up into fixed little volume elements, and matter can flow freely between them. The volume elements serve to normalize all thermodynamic quantities to densities, which results in a very natural and commonplace formalism.

What, if the isothermal pot is open at constant pressure and matter can be exchanged at a fixed chemical potential? In other words, can we iterate the trick once more? Well, yes we can. But subtracting the last extensive sector takes us to end of the road. It renders the remaining potential useless as a means to express thermodynamic stability. Yet, the potential obtained from the internal energy in an $m+1$ dimensional state space after m Legendre transformations, which only has intensive variables, is sometimes referred to as the generalized grand canonical potential $K = 0$, $\mathrm{d}K = 0$. It comes in handy for those who have already forgotten the Gibbs–Duhem and Euler relations from the preceding section, since that is exactly what these equations mean. How nice and friendly thermodynamics is to us!

The above list of potentials (together with their possible reparametrizations) certainly gives you enough Maxwell relations between their mixed second derivatives to keep you happy for quite a while — or the characteristic thermodynamic headache, as some call it, if you happen to need to transform some odd thermodynamic expression into an equivalent but different form. The implicit function calculus revisited in the exercises is always good for some additional entertainment in such situations.

The second law for isothermal systems

To illustrate the somewhat more intuitive and practical formulation of the second law as a minimum condition for the free energy of an isothermal system (instead of a maximum principle for the entropy of an isolated system) consider

---

[14] Note that G/N is simply the chemical potential μ; so it is not really a big surprise that chemists (including those calling themselves semiconductor- or particle physicists) love it.

an isothermal system on which some work is performed by external means. If the system transfers the amount of heat $|\delta Q| = -\delta Q$ into a surrounding isothermal medium that acts as a heat reservoir R at temperature T_R, one has

$$\delta W = \mathrm{d}U - \delta Q = \mathrm{d}U + T_R \mathrm{d}S_R \; . \tag{45}$$

This is the first law, where, as usual, the internal energy change $\mathrm{d}U$ is meant to comprise all forms of energy (elastic, electric, etc., *and* heat) stored internally in the system. In equilibrium, i.e., for reversible transformations with $T = T_R$, one could write the heat $-\delta Q$ given by the system to the reservoir in the form $-\delta Q_{\mathrm{rev}} = -T_R \mathrm{d}S$, and only then the work δW could be written as a total differential $\delta W_{\mathrm{rev}} = \mathrm{d}F = \mathrm{d}U - T_R \mathrm{d}S$ of the equilibrium free energy. But the above expression is more general. The reservoir is per definition always in equilibrium, so its temperature T_R is always defined, even if the embedded system, and thus also the "universe" Σ comprising the system *and* its reservoir R, are out of equilibrium. The heat flowing from the system to the reservoir can therefore be written in the form $-\delta Q = \delta Q_R = T_R \mathrm{d}S_R$. The entropy change $\mathrm{d}S_\Sigma$ for the isolated "universe" is non-negative, according to Clausius' theorem. Using its additivity,

$$\mathrm{d}S_\Sigma \equiv \mathrm{d}S + \mathrm{d}S_R \geq 0 \; , \tag{46}$$

one finds that the so-called *dissipated work* (or "dissipation"), defined as the difference between the actual work performed and the equilibrium work $\mathrm{d}F$ needed to achieve the same state change under reversible isothermal conditions, is always non-negative:

$$\delta W_{\mathrm{diss}} \equiv \delta W - \mathrm{d}F = \mathrm{d}U + T_R \mathrm{d}S_R - (\mathrm{d}U - T_R \mathrm{d}S) = T_R \mathrm{d}S_\Sigma \geq 0 \; . \tag{47}$$

This is yet another formulation of the 2nd law for isothermal systems, which is equivalent to saying that *the free energy happens to be the minimum work required to realize a certain state change.* Since the reservoir is in equilibrium, the system alone must be blamed for the irreversibility quantified by the dissipated work. So the total entropy production $\mathrm{d}S_\Sigma$ in the universe equals the extra entropy production of the system, as introduced in Eq. (22), i.e., $\mathrm{d}S_\Sigma = \delta S_{\mathrm{irr}} \geq 0$, and the dissipated work can also be written as $\delta W_{\mathrm{diss}} = T_R \delta S_{\mathrm{irr}}$. From

$$-\delta Q = T_R \mathrm{d}S_R = T_R(\delta S_{\mathrm{irr}} - \mathrm{d}S) = \delta W_{\mathrm{diss}} - \delta Q_{\mathrm{rev}} \tag{48}$$

it follows that δW_{diss} is the extra (non-reversible bit of) heat flowing from the system to the reservoir under non-equilibrium conditions, which is therefore often also called the dissipated heat $-\delta Q_{\mathrm{diss}} = T_R \delta S_{\mathrm{irr}} \geq 0$ "lost" in the process.

1.5 Example: the classical ideal gas

As a paradigmatic example for a fundamental relation consider the classical ideal gas. A common situation is that someone has measured, say, the pressure equation of state[15],

$$pV = Nk_BT \ , \tag{49}$$

and[16] wants to know the full fundamental relation. Surely, Eq. (49) alone, which simply says that each particle independently contributes k_BT/V to the pressure, is not enough to fully define the model. Such an equation $p(T,V,N)$ can only determine the mechanical sector of the fundamental relation for the free energy $F(T,V,N)$. However, as pointed out above, the missing information is constrained by the requirements of extensivity and thermal equilibrium. Because of the former, a formulation with densities of extensive quantities suffices, so that a *single* additional equation of state will uniquely specify the model. One can, for example, restrict the discussion to the work and heat sectors by considering a constant particle number N ($\mathrm{d}N = 0$) throughout, to rewrite everything "per particle" ($u \equiv U/N$, $s \equiv S/N$, $f \equiv F/N$, $v \equiv V/N \equiv n^{-1}$). The $N-$dependence of U, F, or S can always be recovered by multiplying N back in (e.g. $f \to F = Nf$). The equations of state and the fundamental relation are therefore more economically expressed in a manifestly intensive form, e.g.

$$\mathrm{d}f = -s(T,v)\mathrm{d}T - p(T,v)\mathrm{d}v \ . \tag{50}$$

But even the choice of the second equation of state $s(T,v)$ is not arbitrary once $p(T,v) = k_BT/v$ is given. The integrability of f requires $T\partial_v s)_T = T\partial_T p)_v = p$. Expanding the fundamental relation $T\mathrm{d}s(u,v)$ for the entropy in T and v

$$T\mathrm{d}s(T,v) = \mathrm{d}u(T,v) + p\mathrm{d}v = (\partial_T u)_v \mathrm{d}T + \underbrace{[p + (\partial_v u)_T]}_{\partial_v s)_T} \mathrm{d}v \ , \tag{51}$$

then immediately reveals $(\partial_v u)_T = 0$. Hence $u(T,v)$ has to be independent of v, which leaves little freedom, given the physical units, namely $u(T) \propto k_BT$ and $c_V = \partial_T u)_V \propto k_B$. All that remains to be specified about the second equation of state, which is sometimes called caloric or energy equation, is the proportionality constant that can be measured by calorimetry. It depends on the dispersion relation of the gas particles (e.g. relativistic or non-relativistic) and is material specific (e.g. dependent on whether the particles can store energy in rotational or vibrational modes, etc). Hence, thermodynamics allows for more than one ideal gas. Deviations from a strict proportionality are possible in temperature ranges where new energy scales (e.g. the particles' rest mass or internal excitation energies), or quantum effects ($\hbar$) enter the stage. Classically, the prefactor in Eq. (52) is fixed by *equipartition*, saying that each equilibrated degree of freedom stores $k_BT/2$ of thermal energy, so that

$$c_V = 3k_B/2 \qquad \text{and} \qquad u = 3k_BT/2 \tag{52}$$

---

[15] using atomistic in place of the proper macroscopic notation with gas constant and moles.

[16] like Faust's famulus saying "zwar weiß ich viel, doch möcht' ich alles wissen"

for a non-relativistic mono-atomic classical gas[17].

With this information, the differential fundamental relation for s is straightforwardly integrated, up to an unknown additive constant that is tacitly suppressed:

$$s = k_B \ln \left(v u^{3/2} \right) = k_B \ln \left(v T^{3/2} \right) \ . \tag{53}$$

Importantly, only the first form, expressing the entropy in its natural variables, is the desired fundamental relation $S(U, V, N) = N s(U/N, V/N)$, which contains the full thermodynamic information. In contrast, the second one merely is an equation of state for the fundamental relation of the free energy per particle in Eq. (50). In terms of its natural variables T, $v = 1/n$,

$$f = u - Ts \cong -Ts = -k_B T \ln \left(v T^{3/2} \right) = k_B T \ln \left(n T^{-3/2} \right) \tag{54}$$

also contains the complete thermodynamic information about the classical ideal gas in a slightly different form, more appropriate for isothermal systems. Again, irrelevant terms were omitted, which, because of $s = -(\partial_T f)_v$, are now proportional to $k_B T$. Since $\mu = f + pv = f + k_B T$, the free energy per particle and the chemical potential μ can, up to irrelevant terms, both be identified with $-Ts$ so that all potentials are nicely convex in their extensive variables, as it should be.

Incidentally, $Ts = -\mu$ implies that μ must be negative, if the entropy of a dilute gas increases upon adding more particles, as one would expect. And also that the entropy of the classical ideal gas is (up to a thermodynamically irrelevant gauge constant) "all chemical". In return, the chemistry of the ideal gas is "all entropic", of course. You may notice something mildly odd about the notion of "chemistry" in the context of an ideal gas that we imagine to be made of non-interacting identical particles. Take m different ideal (i.e. non-interacting) gases at ambient temperature in m identical bottles. Together they have m times the entropy of each single one. Putting them all into the same bottle should not change this, since the gases have no way of distinguishing in which bottle they are, and with whom, as there are no mutual (chemical) potentials. They also will not notice that the pressure in the bottle grows by a factor of m upon mixing them all together, since they share the burden democratically. But if the bottles initially all contained the same sort of gas, mixing their contents together into one bottle amounts to an $m-$fold compression of the gas. The bottle then contains mN *identical* particles at the m-fold density. The entropy decreases accordingly, as the chemical potential becomes less negative. Conversely, pulling a separating wall out of a bottle containing the same gas at the same conditions on both sides of the wall, changes nothing (this is the definition of homogeneity). Instead, if the compartments initially held two different (ideal) gases, the entropy will of course increase, because these gases expand. Assuming that the

---

[17] If you believe in atoms, it is easy to show that, for purely kinematic reasons, $p = \tau u/d$, with d being the space dimension and τ characterizing the dispersion relation (e.g. quadratic or linear), which also allows Eq. (52) entirely to be inferred from Eq. (49) (or *vice versa*).

initial thermodynamic conditions were again homogeneous except for the type of atoms in the compartments, this changes nothing in the thermodynamic relation with the outside world. Yet, the internal expansion could be harnessed, e.g. using semipermeable membranes, to do some useful work, which is witnessed by the entropy. In summary, ideal gases have a (purely entropic) chemical self-interaction but no mutual interaction. A better understanding of the origin of this somewhat counter-intuitive self-interaction of gases (sometimes referred to as the Gibbs paradox of mixing) emerges in statistical mechanics.

Clearly, all of the above expressions lack a proper normalization to render the arguments of the logarithms dimensionless. Thermodynamics is insensitive to this normalization, which may be assumed to be hidden in the omitted terms. But the physical origin and nature of such a (mathematically necessary) normalization remains a mystery in thermodynamics, as does the strange behavior for $T \to 0$.

Finally, as an illustration of the basic elements introduced so far, and before entering more elevated terrain, you are encouraged to consider the following example: an ideal gas in a closed container is heated/cooled from T_1 to T_2 by bringing it into contact with a heat reservoir either reversibly — the reservoir temperature is raised/lowered slowly from T_1 to T_2 — or irreversibly — the reservoir temperature is already T_2 when the contact is established. Calculate in each case the heat and entropy exchanged and the entropy of the gas before and after heating/cooling. Verify the second law.

2 Phase Equilibria and Phase Transitions

2.1 Natura salta

The fundamental laws of thermodynamics introduced above might seem nice to know from a historical perspective, but somewhat redundant in practice. All the practitioner really needs should be a list of fundamental relations for all possible materials, after all. However, the multitude of potentially interesting many-body problems is virtually infinite, and even for quite mundane substances like H_2O, say, the fundamental relations turn out to be very complicated functions that are only incompletely known[18]. Most interestingly, they usually display non-analytic behavior indicating phase changes. A thorough first-principles understanding of the mathematical origin of kinks and other anomalies in the fundamental relations only emerged in the late 20th century. As familiar from mountain skylines, non-analytic points such as kinks naturally arise from smooth functions upon projection from a higher dimensional space (the many-particle phase space) onto a lower dimensional space (the thermodynamic state space). The concept of a phase transition has become very popular beyond the realm of physics, since it describes quite generally how a qualitatively new collective behavior emerges in

---

[18] The above example of the ideal gas is a completely atypical and rare exception to this rule.

a complex system in response to a gradual parameter change. Phase transitions are intimately related to so-called castastrophees, bifurcations, emergence, and spontaneous symmetry breaking. One does not easily find a scientific notion that comes closer to being synonymous with “interesting”.

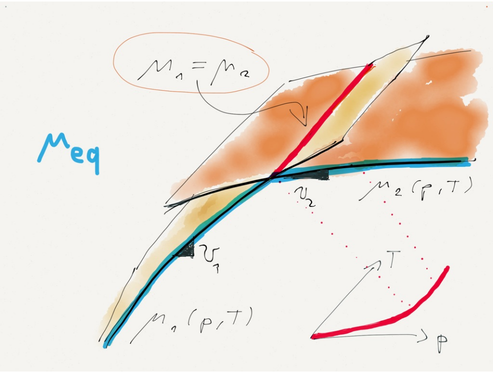
Continuous and discontinuous phase changes

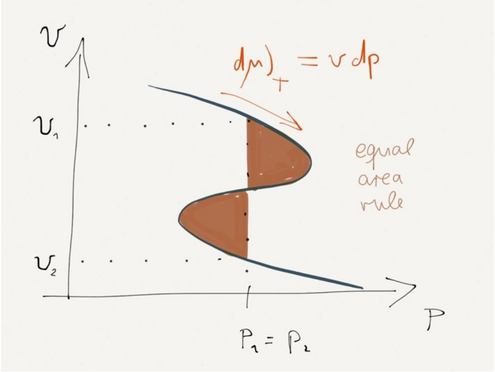
Familiar examples of phase transitions are cooking water and melting ice. Both processes are reversible and occur at well defined temperatures and pressures, i.e. they are equilibrium phenomena. The constant transition temperature is of great help in cooking, its sensitivity to pressure in pressure cookers, ice skating, etc. Phase changes are usually accompanied by the vanishing or the emergence of collective qualities such as surface tension or elasticity, and often by abrupt density changes or jumps upon gradual parameter changes. For example, the densities of fluid and solid phase at freezing and melting, or of the vapor and liquid phase at condensation and evaporation, can be pretty different, although both phases coexist. Accordingly, the isothermal compressibility of the whole system has to diverge at coexistence (compression and expansion can be achieved without a force by simply transforming vapor into liquid or *vice versa*).

Phase changes involving jumps in (densities of) the extensive variables are classified as *discontinuous (or first order) transitions*, “first order” referring to discontinuities and singularities in the first derivatives of the thermodynamic potentials with respect to the gradually changing intensive control parameters. The generic example is the (specific) volume (the reciprocal density), which is the first derivative of the chemical potential with respect to the ambient pressure, at a phase coexistence line. Further, there are distinguished points in a phase diagram, called *critical points*, where a phase coexistence line terminates. (Following the coexistence line to its end is also the only realistic way to find such a needle in the haystack.) This corresponds to a kink in the mountain skyline turning into a smooth pass after shifting the point of view, so to speak. There are no jumps in the extensive thermodynamic variables and their densities at critical points. This is why such critical phase transitions are classified as *continuous (or second order) transitions*, “second order” referring to the appearance of discontinuities in the higher derivatives of the thermodynamic potentials, *viz.* the response coefficients, such as the compressibility or the specific heat.

Phase boundaries, Gibbs phase rule, Clausius-Clapeyron

A very natural representation of phase equilibria employs $p - T$ diagrams (or their equivalents for, say, magnets, etc.). Pressure and temperature are usually the ambient control parameters. Moreover, mechanical and thermal equilibrium requires them to be the same in the coexisting phases, which means that these are neighbors that share coexistence lines in the $p-T$ plane. In contrast, in diagrams





with at least one extensive degree of freedom, they are generally separated by finite gaps. With p and T controlled, the appropriate potential is the Gibbs free energy (per particle) or the chemical potential $\mu = G/N$, with the total differential

$$\mathrm{d}\mu = -s\mathrm{d}T + v\mathrm{d}p\ . \tag{55}$$

The function $\mu(T,p)$ generally has a different functional form in each of the phases; but both must coincide at coexistence to ensure chemical stability. As with mechanical and thermal equilibrium, two phases are said to be in chemical equilibrium with respect to particle exchange, or to coexist, if their responsible intensive parameters, namely their chemical potentials, are equal. Coexistence does not restrain the (densities of the) extensive variables, though. In the examples of melting and evaporation, the densities $n_{1,2} = v_{1,2}^{-1}$ and the compressibilities $\kappa_T^{1,2}$ are quite different. Ice floats on water, vapor is much easier compressed than water. Therefore, for a fixed given T, the monotonous concave curves corresponding to the two "branches" $\mu_1(p)$ and $\mu_2(p)$ of the chemical potential for the two phases differ in their slopes $(\partial_p\mu)_T = v > 0$ and curvatures $(\partial_p^2\mu)_T = (\partial_p v)_T = -v\kappa_T < 0$ such that they can only intersect in one point, the coexistence point (sketch). Allowing T to vary, this point expands into a coexistence line, then. Outside coexistence, only the lower of the two branches $\mu_{1,2}(p)$ is thermodynamically stable, since the relevant thermodynamic potential $\mu = G/N$ must be at its minimum in equilibrium. Consequently, at coexistence, $\mu(p)$ has a kink and the compressibility κ_T diverges. This is analogous to the above mentioned mountain hiker experience: we only can see the (visually) more elevated of two smooth contours of mountains at different distances, which gives rise to the optical illusion of kinked valleys in the skyline, where a virtual mountain biker would crash (experience a diverging acceleration).

So, while two phases are separated by a *coexistence line* in the $p-T$ plane, three phases (e.g. vapor, liquid, solid) can only coexist in a single point, the *triple point*. Just as you exploit the evaporation transition as a simple thermostat whenever you cook food, three phase coexistence of a (very) pure substance can be used as a pressure and temperature gauge, as in some early space missions.

The general situation is neatly summarized in the *Gibbs phase rule*: in a mixture of $\mathcal{N}$ substances $\mathcal{P}$ phases can be in equilibrium on a manifold of

$$2 + \mathcal{N} - \mathcal{P} \geq 0 \tag{56}$$

dimensions spanned by intensive thermodynamic degrees of freedom. Whenever two phases are required to be in equilibrium, this constrains one intensive variable, thereby reducing the degrees of freedom of the system by one. In a pure substance three-phase coexistence is constrained to a point, the triple point, since $2+1-3 = 0$. Note that the Gibbs phase rule refers to a phase space characterized by *intensive control parameters*. The triple point, quite in contrast to the critical point, expands to a line or an area, respectively, if you exchange one or two of the intensive parameters by extensive ones (or their densities).

The matching condition for the chemical potentials of coexisting phases can be exploited to derive the shape of the coexistence line $p(T)$ in the $p-T$ diagram. On the coexistence line,

$$0 = d\mu_1 - d\mu_2 = -(s_1 - s_2)dT + (v_1 - v_2)dp \qquad \Rightarrow \qquad \frac{dp}{dT} = \frac{\Delta s}{\Delta v}\,. \tag{57}$$

This is known as the *Clausius-Clapeyron* equation. The slope of the coexistence line is controlled by the ratio of the specific volume change $\Delta v \equiv v_1 - v_2$ and the *latent heat* per particle $T\Delta s \equiv T(s_1 - s_2)$. At the critical point both Δv and $T\Delta s$, which are both density differences of extensive quantities, have to vanish, so they can both serve as so-called *order parameters* that help to quantify the newly emerging quality (related to the emerging "kinkiness"). Typically, one expects the hotter phase to be more disordered and less dense, which suggests that the coexistence lines should have positive slopes. A notable exception is the liquid-solid coexistence line of water, where the density increases upon freezing due to H-bond network formation; another one is the Pomerantschuk effect in He^3, where the localization of atoms in the solid phase effectively disentangles the wave functions of adjacent electrons. This liberates the otherwise Pauli-correlated spins, and the entropy increases accordingly upon cooling.

2.2 Van der Waals' model for phase transitions

The middle school teacher Johannes Diderik van der Waals managed to delegate a substantial part of his professional duties to his wife, so that he could reserve some time for taking math and physics classes at the university. At the age of 44 he handed in his PhD thesis "Over de Continuiteit van den Gas- en Vloeistoftoestand", which contained an equation of state with a first order liquid-gas transition that vanishes in a critical point. For this and some other achievements he was honored with the Nobel prize in 1910. Till today, the van der Waals gas has remained a favorite tutorial example to gain some intuition about phase transitions and a starting point for improved theories, although this turns into a somewhat tortuous bricolage.

Statistical Mechanics Kindergarden

Motivated by his believe in atoms, van der Waals extended the ideal gas model, which he understood as a model of N non-interacting tiny balls, in two directions, namely by introducing *excluded volume* and *mutual attractions* between the gas molecules in a schematic manner. Subtracting some excluded volume b from the total available specific volume $v = V/N$ makes the pressure $p = k_B T/(v - b)$ at $v = b$ rather than at $v = 0$. So b could be interpreted as a rough estimate of the volume per molecule in a densely packed solid or liquid. Alternatively, arguing that the way a dilute gas realizes the presence of excluded volume is via rare pair collisions, one can also consider the volume $4\pi\sigma^3/3$ around the center of mass of a sphere of diameter σ, from which the center of mass of a second sphere is excluded. Since there are $N/2$ pairs, the specific volume v should thus be reduced by $b = 2\pi\sigma^3/3$. As you can quickly convince yourself by considering three spheres in close proximity (sketch), the excluded volume is non-additive, which in fact poses a major problem to any liquid state theory trying to seriously address condensation phenomena. So the trick with the parameter b cannot really work, not even for hard spheres, for which the phase diagram is sketched in Fig. 1.

As suggested by the title of van der Waals' thesis, the emphasis of his work was not so much on the excluded volume but rather on the liquid-gas transition, which is solely due to the attractive part of the molecular interactions. It is not difficult to imagine that any attempt to realistically model the attractions encounters comparable problems as for the repulsions. The problem can however in this case be circumvented by choosing infinitely long ranged attractions, which essentially "smears out" the nasty packing problem. If all pairs of molecules attract each other with the same strength, independent of their distance, the packing structure does not matter for the attraction. As a simple realization think of a square well potential of range $\propto V^{1/3}$ and infinitesimal strength $\propto V^{-1}$ such that the overall reduction of the energy density $u = U/N = N(N-1)/(2VN)$ remains finite in the thermodynamic limit ($V,\ N \to \infty$ at fixed $n = 1/v$) and can simply be written as a term $-a\,n = -a/v$, proportional to the number density n of particles. This prescription is clearly an attempt to dress a manifestly non-extensive model in an extensive garment. The apparent extensivity is purchased at the cost of a strange system-size dependent microscopic interaction, which is not literally what is usually meant with extensivity, of course, and the reckless ruse will be seen to entail some unphysical artifacts in the predictions of the model.

Before analyzing predictions, the model should first be completely formulated, though. The repulsive and attractive corrections in the form $p = k_B T/(v - b)$ and $u = 3k_B T/2 - a/v$, as motivated above, are not yet thermodynamically consistent. This is betrayed by the argument used above, around Eq. (51), to demonstrate that u is independent of v for the ideal gas. It still goes through with the new expression for the pressure, in contradiction to the new expression for the internal energy per particle, for which $\partial_v u)_T \neq 0$. The obvious cure is to correct the pressure equation such that

$$T\partial_v s)_T = T\partial_T p)_v = p + \partial_v u)_T = p + a/v^2 \tag{58}$$

is consistent with the integrability condition for the entropy $s(T, v)$ in Eq. (51):

$$p(T, v) = k_B T/(v - b) - a/v^2 \ , \qquad u(T, v) = 3k_B T/2 - a/v \ , \tag{59}$$

That the pressure feels the attractions between the gas particles is physically reasonable and could have been anticipated. Using Eqs. (58), (59) to integrate $\mathrm{d}s(T, v)$ and to rewrite the result in the appropriate form $s(u, v)$ of a fundamental relation then yields (up to an irrelevant constant)

$$s = k_B \ln[(u + a/v)^{3/2}(v - b)] \ . \tag{60}$$

The free energy per particle follows as

$$f(T, v) = u - Ts = -k_B T \ln[T^{3/2}(v - b)] - a/v \ . \tag{61}$$

Since the entropy per particle $s = -\partial_T f$ is only known up to a numerical constant, f and $\mu = f + pv$ are only determined up to irrelevant terms $\mathcal{O}(k_B T)$, as already pointed out for the ideal gas. For the most commonly studied volume density $\mathfrak{f}(T, n) \equiv F/V = nf$ of the free energy, the gauge freedom even corresponds to a “tilting” of the free energy landscape by a term $\mathcal{O}(nk_B T)$.

Academic first aid: equal area & common tangent constructions

Maxwell realized that van der Waals’ model did not respect the rules following from homogeneity and thermodynamic stability. At low temperature it displays regions of negative compressibility corresponding to sign changes in the curvature of the free energy, and even negative pressures. The origin of the problems lies of course in the artificial infinite range interactions, which make the van der Waals fluid similar to a system dominated by gravity. They prevent the gas from becoming heterogeneous, since interfaces between coexisting phases are extensive and therefore energetically too costly. The “unphysical” loops in the van der Waals isotherms thus indicate where phase coexistence should occur in a gas with short-ranged interactions (not dominated by gravity). Rather than suggesting to dump the model into the dust bin, Maxwell proposed to cure the flaw by applying some plastic surgery to the symptoms. His quick fix consists in correcting the model predictions, wherever they violate thermodynamic stability, by hand. It can be condensed into two practical rules known as *equal area rule* and *common tangent construction*, respectively. For the phases to be in thermal, mechanical, and chemical equilibrium, they need to have identical intensive parameters T, p, μ. Integrating the differential $\mathrm{d}\mu|_T = v\mathrm{d}p$ of the chemical potential along a van der Waals isotherm from p_1 to $p_2 = p_1$ (sketch),

$$0 = \mu_2 - \mu_1 = \int_{\mu_1}^{\mu_2} \mathrm{d}\mu = \int_{p_1}^{p_2 = p_1} v \, dp \tag{62}$$

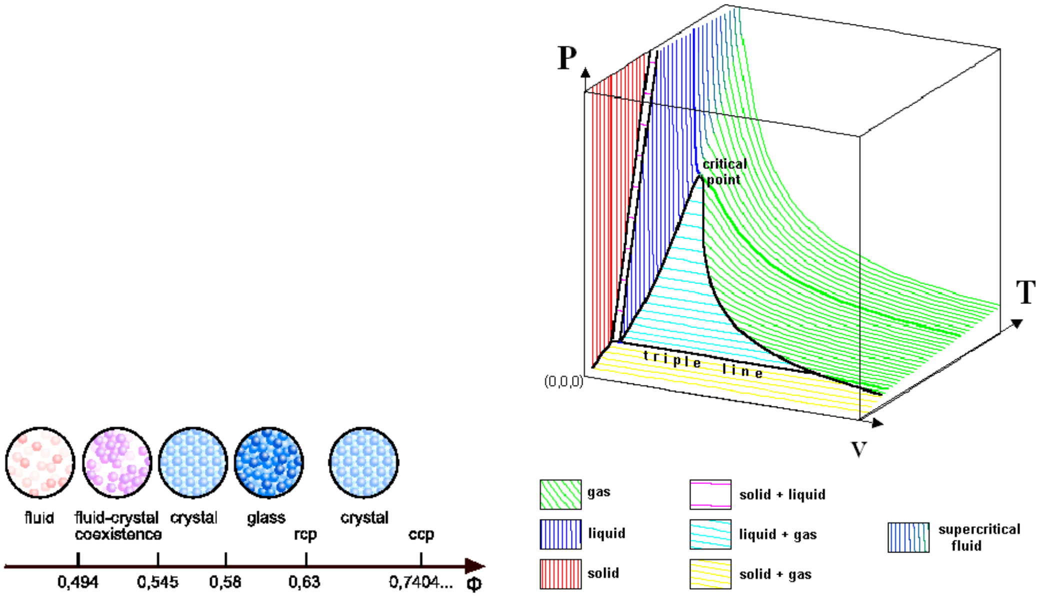


Figure 1: *Left:* Phase diagram of hard spheres as a function of volume fraction $\phi \equiv \pi n \sigma^3/6$. The equilibrium phase behavior comprises fluid, fluid-crystal coexistence and crystal. The transition is driven by the free volume v per particle (and the corresponding configurational entropy), which is higher in an ordered crystal compared to a disordered fluid. The glass is a disordered nonequilibrium solid that interferes with crystallization between $\phi \approx 0.58$ and rcp = random close packing. The closest possible packing is ccp = crystalline close packing (corresponding to hcp/fcc) as conjectured by Johannes Kepler in the early 17th century and proved by Thomas Hales in the late 20th century. *Right:* Adding attractions to hard spheres makes temperature an additional interesting control parameter of the phase diagram. If the attractions are sufficiently long ranged compared to the hard core diameter, a liquid-vapor phase coexistence appears at low temperatures and intermediate volume fractions. In the high temperature limit, the phase behavior of hard spheres is recovered. (Plot from Wikipedia)

one finds the following geometric interpretation in a $p - v-$diagram: the true isobaric connections between the points on the isotherm that represent the onset of evaporation and condensation at densities $n_{1,2} = v_{1,2}^{-1}$, respectively, divide the "unphysical" loops in the isotherm into regions of equal area. Physically, the straight connections indicate two-phase coexistence, the relative amount of either phase being controlled by the position along the isobaric *tie-line.* For the graph of the free energy density

$$F/V \equiv \mathfrak{f}(n) = -p + \mu n \;, \tag{63}$$

this construction amounts to the convex hull of the non-convex coexistence region, a straight line (of constant slope $\partial_n \mathfrak{f}|_{T,V} = \mu_1 = \mu_2$) tangential to the graph of $\mathfrak{f}(n)$ (sketch).

After these corrections, the phase behavior implied by the van der Waals model is qualitatively similar to that observed for real molecular gases, as exemplified by the landscape in Fig. 1. Note the steep cliffs of crystallization and the more shallow flank of liquid-gas coexistence (light blue) enclosed by the *binodal* line, which delimits the region where a single phase is thermodynamically unstable and separates into the two phases delineated by the binodal. It contains the similarly shaped but narrower *spinodal* region, where the van der Waals isotherms predict negative compressibility $\kappa_T = -n\partial_p v \leq 0$, indicating an even more severe, namely mechanical, instability. Inside the spinodal, phase separation is not only driven thermodynamically, but also by a spontaneous mechanical collapse.

Are all gases the same? ("law of corresponding states")

Each substance that obeys the predictions of the van der Waals model *qualitatively* has a *unique liquid-gas critical point* at the end of the coexistence region, which is given by $\partial_v p = \partial_v^2 p = 0$. One can therefore get rid of the material parameters with a normalization that renders the van der Waals equations dimensionless by normalizing density, volume, temperature, and pressure by their respective values at the critical point, notified by an index c,

$$\tilde{v} \equiv v/v_c \,, \qquad \tilde{n} \equiv n/n_c \,, \qquad \tilde{T} \equiv T/T_c \,, \qquad \tilde{p} \equiv p/p_c \tag{64}$$

with

$$v_c = 1/n_c = 3b \,, \qquad p_c = a/27b^2 \,, \qquad k_B T_c = 8a/27b \,. \tag{65}$$

The corresponding reduced thermal equation of state reads

$$\tilde{p} = 8\tilde{T}/(3\tilde{v} - 1) - 3/\tilde{v}^2 = 8\tilde{n}\tilde{T}/(3 - \tilde{n}) - 3\tilde{n}^2 \;, \tag{66}$$

Similarly, on has the reduced free energy density

$$\tilde{\mathfrak{f}} \equiv \mathfrak{f}/n_c k_B T_c = \tilde{n}\tilde{T} \ln \left[\tilde{T}^{-3/2}\tilde{n}/(3 - \tilde{n})\right] - 9\tilde{n}^2/8 + \mathcal{O}(nT) \tag{67}$$

and the reduced chemical potential

$$\tilde{\mu} \equiv \mu/k_B T_c = 3\tilde{T}/(3-\tilde{n}) + \tilde{T}\ln\left[\tilde{T}^{-3/2}\tilde{n}/(3-\tilde{n})\right] - 9\tilde{n}/4 + \mathcal{O}(T) \; . \tag{68}$$

These reduced gas equations and in particular the combination $\tilde{p}\tilde{v}/\tilde{T}$ (which is predicted to converge to the constant $n_c k_B T_c/p_c = 8/3$ in the ideal gas limit) should according to the model be independent of the two material parameters. Experimental data for a number of molecular gases provides reasonable evidence for the existence of such a universal "law of corresponding states", although it exhibits some deviations from the specific van-der-Waals form. (*Exercise*: What can be deduced from this observation about the microscopic interactions?)

2.3 Critical phenomena

Critical exponents

An even greater universality is observed for so-called equilibrium critical behavior, which refers to various divergencies encountered upon approaching a critical point. Critical points can be understood as organizers of the surrounding thermodynamic state space, and thermodynamic behavior may be classified into universality classes, accordingly. It turns out that systems that are microscopically and even phenomenologically as different as magnets and fluids may fall into the same universality class with respect to their critical behavior. Quite generally one summarizes critical behavior by a number of critical exponents characterizing the singular dependencies of various thermodynamic quantities of interest at small $t \equiv T/T_c - 1$ (or $t \equiv 1 - T_c/T$),

$$c_V \propto |t|^{-\alpha} \qquad\qquad \text{specific heat} \tag{69}$$

$$\psi \propto |t|^{\beta} \qquad (t<0) \qquad \text{order parameter} \tag{70}$$

$$\chi_T \propto |t|^{-\gamma} \qquad (\psi_> = 0; \psi_< = \psi_{1,2}) \qquad \text{susceptibility} \tag{71}$$

$$h \propto |\psi|^{\delta} \qquad (t=0) \qquad \text{critical isotherm} \tag{72}$$

In the exercises, you can convince yourself that these exponents are mutually related as a consequence of general thermodynamic constraints, a point that will come up again further below. Broadly speaking, α characterizes the thermal effects such as those associated with the appearance of a latent heat, and δ the mechanical, magnetic, etc. instability at the transition, as encoded in the shape of the critical isotherm. For the liquid-gas critical point $\chi_T = \kappa_T$, so that γ describes the divergence of the compressibility. The notation $|t|$ indicates that both positive and negative t are to be considered and impies that the singular behavior above and below the transition is quite symmetric (except for prefactors). For the van der Waals gas and similar models, one has to remember to use the Maxwell construction to correct for model failures at $t < 0$ as discussed above.

Most interestingly, a quantity called the order parameter can usually be defined, here denoted by ψ, which vanishes above T_c and takes on non-zero values below T_c. It serves to characterize the emergence of qualitatively new thermodynamic behavior, e.g. the density difference between the two new phases emerging at the liquid-gas critical point, the spontaneous magnetization of a ferromagnet cooled below its Curie temperature, the condensate fraction in Bose–Einstein condensation, the Higgs field, etc. Its growth is quantified by β, defined only for $t < 0$. As the mentioned examples suggest, a *spontaneous symmetry breaking*[19] is often associated with the transition, as e.g. the spontaneously emerging magnetization clearly has to point in some definite but *a priori* arbitrary direction.

To illustrate the abstract concepts, the van der Waals gas or the very similar Weiss model for a ferromagnet may serve as simple tutorial examples (see exercises).

Landau theory and universality

To rationalize the amazing universality of critical phenomena, Lev Landau suggested that a Taylor expansion of the relevant free energy density as a function of small deviations from the critical point should be expected to look very similar for a large variety of phenomenologically very different systems. The common form of the *Landau free energy density* $\mathcal{L}$ is that of a standard hydrodynamic description, namely a free energy per volume controlled by intensive parameters. For the example of the van der Waals model, the Landau free energy is given by a volume integral over the density of the grand canonical potential, which, for prescribed T and μ, takes its minimum in equilibrium:

$$\mathcal{L} \equiv J(T,\mu)/(n_c k_B T_c V) = -3\tilde{p}/8 = \tilde{\mathfrak{f}}(\tilde{T},\tilde{n}) - \tilde{\mu}\tilde{n} \; . \tag{73}$$

With the short-hand notation $c_1 \equiv 9/4$, $c_2 \equiv 9/16$,

$$\psi \equiv \tilde{n} - 1 \; , \qquad t \equiv \tilde{T} - 1 \; , \qquad h \equiv \tilde{\mu} - \tilde{\mu}_c \; , \tag{74}$$

and expanding $\tilde{\mathfrak{f}}$ to leading orders in t and in the order parameter ψ, one finds

$$\mathcal{L}(t,h) = \frac{c_1 t}{2}\psi^2 + \frac{c_2}{4}\psi^4 - h\psi \tag{75}$$

A proper thermodynamic free energy ought to be convex in ψ to reflect thermodynamic stability. Instead, $\mathcal{L}$ is easily seen to develop a double-well structure for $t < 0$. This problem, which is inherited from the canonical free energy $\mathfrak{f}(t,\psi)$, must again be corrected by the Maxwell construction. For vanishing "external field" h, the latter boils down to simply joining the two local minima, the positions of which give the possible values of the order parameter, by a horizontal straight line. The critical exponents are straightforwardly calculated as $\beta = 1/2$, $\gamma = 1$, $\delta = 3$; the specific heat has a jump but no divergence, i.e. $\alpha = 0$ (exercise!).

---

[19] …plus the associated Goldstone modes, such as phonons, photons, magnons, etc., hydrodynamic excitations that aim to restore the broken symmetry.

Scaling hypothesis and exponent relations

Landau's approach makes it plausible that the above predictions for the critical exponents are almost inevitable in any simple model and quite insensitive to the underlying microscopic interactions. Some variation may be expected due to different order parameter dimensions (scalar, vector, tensor) or special symmetries that require certain orders in the expansion to be absent. But apart from that there seems little scope for surprises. While the implication concerning the universality of the critical exponents is essentially borne out by experiments, Landau's (or van der Waals') numerical values turn out to be almost always inappropriate. This is not merely a quantitative failure, which could be overcome by a slightly improved model to be worked out by some diligent chemists, but a failure in principle. It is the assumption that the free energy can be expanded in a Taylor series, which is at stakes, since the measured exponents are not simple rational numbers, indicating a severe non-analyticity. A full understanding of the origin of this odd but common behavior on the basis of Statistical Mechanics only emerged in the 1970's and was rewarded with the Nobel prize for K. G. Wilson (see the link on the homepage).

Phenomenologically, it can be rationalized by the postulate of a *generalized homogeneity* relation for the Landau free energy $\mathcal{L}$, or rather its *singular part* $\mathcal{L}_s$, in its natural intensive variables (the control parameters T and h). This postulate is known as the scaling hypothesis:

$$\lambda \mathcal{L}_s(t,h) = \mathcal{L}_s(\lambda^{a_t} t, \lambda^{a_h} h) \ , \tag{76}$$

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. (76) reduces to

$$\mathcal{L}_s(t,h) = |t|^{1/a_t} \mathcal{L}_s(\pm 1, h/|t|^{a_h/a_t}) \equiv |t|^{1/a_t} \mathcal{L}_\pm(h/|t|^{a_h/a_t}) \ . \tag{77}$$

Landau's analyticity assumption for $\mathcal{L}$ is replaced by regularity requirements for the *scaling function* $\mathcal{L}_\pm$ and its first and second derivatives at the origin, complemented by the condition $\mathcal{L}'_\pm \equiv \partial_x \mathcal{L}_\pm \overset{x\to\infty}{\sim} -x^{1/\delta}$ to meet Eq. (72). Then Eq. (77) immediately produces the desired power-law behavior for small t and h:

$$\begin{aligned}
C_V &\propto -T\partial_t^2 \mathcal{L}_s(t, h=0) \propto |t|^{1/a_t - 2} \mathcal{L}_\pm(0) && (78)\\
\psi &\propto -\left.\partial_h \mathcal{L}_s\right|_{h=0} \propto |t|^{(1-a_h)/a_t} \mathcal{L}'_\pm(0) && (79)\\
\chi_T &\propto -\left.\partial_h^2 \mathcal{L}_s\right|_{h=0} \propto |t|^{(1-2a_h)/a_t} \mathcal{L}''_\pm(0) && (80)\\
\psi(h) &\propto -\partial_h \mathcal{L}_s \sim -|t|^{(1-a_h)/a_t} \mathcal{L}'_\pm(h/|t|^{a_h/a_t}) && (81)\\
&\propto |t|^{1/a_t - a_h/a_t(1+1/\delta)} h^{1/\delta} \qquad (t \to 0) \ .
\end{aligned}$$

Comparison with the definitions in Eqs. (69)-(72) altogether yields

$$\alpha = 2 - 1/a_t \tag{82}$$

$$\beta = (1 - a_h)/a_t \tag{83}$$

$$\gamma = (2a_h - 1)/a_t \tag{84}$$

$$\delta = a_h/(1 - a_h)\ . \tag{85}$$

The fact that two independent numbers generate (at least) four critical exponents immediately explains the existence of exponent relations. To familiarize yourself with these relations, try to re-derive the relation $2 - \alpha = 2\beta + \gamma$ obtained from thermodynamic considerations in the exercises. The odd numerical values of the exponents have thus been traced back to two singular limits of the scaling function for the free energy, which, of course, cannot itself be derived or computed without the help of Statistical Mechanics.

Summary and outlook

It may seem strange that one introduces a model that is criticized as somewhat unphysical, and then tries to repair its predictions *a posteriori*, instead of introducing an improved model. The reason is that the unphysical behavior of the model is due to the neglect of fluctuations and will therefore never be entirely repaired within thermodynamics. In other words, its failure is paradigmatic for a principle flaw of the whole thermodynamic approach, i.e. the attempt of finding a simple self-contained coarse-grained description.

The long-ranged atomic interactions implicitly assumed in the van der Waals model helped circumvent the complicated problem of packing atoms and molecules in space. But it achieved this by preventing fluctuations and heterogeneities by imposing extensive energy penalties on heterogeneities. Since the model is primarily not meant to describe an ensemble of gravitating particles but an ordinary bottle of gas at room temperature, this must be considered an artifact. For the purpose of describing phase equilibria, Maxwell's common-tangent construction "solved" the problem. While this is clearly a case of locking the stable door after the horse has bolted, it can be justified by taking the thermodynamic limit in van-der-Waals type models more carefully[20]. In any case, the "solution" can be criticized for being too radical, as it throws the baby out with the bathwater. By ironing out the van-der-Waals loops, one effectively neglects surface tension effects, altogether. These play a crucial role in phase separation phenomena, however. They dominate the initial nucleation of new phases, which relies on the spontaneous formation of *finite* supercritical droplets that subsequently grow algebraically slowly until a complete separation into the phases is achieved. Surface

---

[20] For a pedagogic summary of work by Kac, Uhlenbeck, Hemmer, van Kampen, Lebowitz and O. Penrose in the 1960's see Colin Thompson's book *Classical Equilibrium Thermodynamics.*

energy is also essential for taming critical fluctuations. Its vanishing at the critical point gives rise to the long-range correlations and collective fluctuations that are behind the mysterious odd values for the critical exponents and the super-universality of critical phenomena that the thermodynamic scaling hypothesis can only describe but not explain. The unphysical van-der-Waals loops are therefore not entirely unphysical, after all, only much too pronounced for systems with short-ranged interactions.

In summary, the shortcomings of the van-der-Waals model and similar models cannot be cured entirely within thermodynamics but hint at the rich and interesting physics of fluctuations. It is the suppression of fluctuations in simple coarse-grained theories that is responsible for their failure to properly describe the odd behavior at critical points. The quantitative characterization of fluctuations is, next to the recovery of the laws of thermodynamics and the calculation of (approximate) fundamental relations and the position and form of their non-analyticities, a main task in statistical mechanics.

Part II
Statistical Mechanics

$$\langle \hat{A} \rangle = \frac{\sum_\nu \langle \nu | \hat{A} | \nu \rangle e^{-E_\nu / k_B T}}{\sum_\nu e^{-E_\nu / k_B T}} \tag{86}$$

> This fundamental law is the summit of statistical mechanics, and the entire subject is either the slide-down from this summit, as the principle is applied gto various cases, or the climb-up to where the fundamental law is derived and the concepts of thermal equilibrium and temperature T are clarified.
>
> R. Feynman

3 The tedious climb-up

Statistical Mechanics has become the common name for attempts to approach the general principles that help to predict the salient characteristics of macroscopic (physical) systems on a microscopic basis. These attempts usually either try to explain or rely on the following two observations. First, microscopically, interacting many-particle systems exhibit complex dynamical behavior that is best characterized in terms of trajectory densities, borrowing notions from probability theory. Secondly, amazingly simple deterministic laws emerge from this mess on a macroscopic scale. In particular, Equilibrium Statistical Mechanics as developed by Boltzmann, Maxwell, Gibbs, Einstein, and others, takes for granted that accurate predictions of macroscopic deterministic behavior can be obtained based on the postulate of microscopic chaos in one or another form. Predictions are made without actually deriving specific solutions to the complicated dynamic equations, by merely estimating their weights in the space of all possible solutions. Accordingly, statistical mechanics naturally puts some emphasis on universal aspects, and many physicists think that this is a good thing to do, anyway. The study of simple but paradigmatic model systems provides an understanding of how, in principle, simplicity, universality and higher-level collective qualities emerge from the complicated micro-scale physics. In proportion to the degree of specialization, the subjects of statistical mechanics are delegated to condensed matter physics, chemical engineering, econo-physics, or some other specialized disciplines that strongly rely on the overarching framework of statistical mechanics and therefore share major fundamental principles and methods.

3.1 Stairway to heaven

Phase space density, density matrix, ensembles

The most convenient formal starting point for a microscopic description of many-body systems is the phase space density or density operator formalism, respectively. To describe a specific N-particle system corresponding to a pure quantum state $|\mu(t)\rangle$ in a Hilbert space, or, alternatively, the phase space density corresponding to a single phase space trajectory, one writes

$$\hat{\rho}_\mu(t) \equiv |\mu(t)\rangle\langle\mu(t)| \;, \qquad \hat{\rho}_\mu(\{\mathbf{q}_i, \mathbf{p}_i\}, t) \equiv \prod_i^N \delta[\mathbf{q}_i - \mathbf{q}_i^{(\mu)}(t)]\delta[\mathbf{p}_i - \mathbf{p}_i^{(\mu)}(t)] \;. \quad (87)$$

Observables are represented as self-adjoined operators $\hat{A}$ or phase space functions $A(\{\mathbf{q}_i, \mathbf{p}_i\})$, respectively, which are evaluated as

$$A_\mu(t) \equiv \mathrm{tr}\big(\hat{\rho}_\mu(t)\hat{A}\big) \;, \qquad A_\mu(t) = \int \prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \, \hat{\rho}_\mu(\{\mathbf{q}_i, \mathbf{p}_i\}, t) A(\{\mathbf{q}_i, \mathbf{p}_i\}) \;. \quad (88)$$

In the classical case, this merely amounts to a cumbersome formalism for saying that one evaluates a phase space function on a trajectory. The quantum mechanical formulation looks less artificial, as one is accustomed to probabilistic elements involved in calculating the expectation value for the outcome of a future measurement from the knowledge of the state $|\mu(t)\rangle$ in the past. In both cases there are good reasons to generalize the formalism to ensembles of pure states, sometimes called "mixed states".

In fact, in quantum mechanics, partial traces of density matrices generally result in reduced density operators corresponding to a statistical ensemble of states even if the total system including the environment (sometimes called the "universe") is in a pure state. So, whenever one considers a subsystem entangled with an environment, one automatically has to deal with a mixed state of the subsystem, alone. In classical physics, a comparable situation is encountered when the environment acts as a weakly coupled "bath" that communicates with the system of interest and constantly changes its precise *microstate* while leaving the so-called *macrostate*, which is defined by some appropriate reduced set of macroscopic observables, invariant.

So in quantum and classical mechanics it makes sense to take a step back from the specific current microstate of the system at hand and consider an *ensemble* of microscopic states or trajectories that belong to an equivalence class in some reduced description. A gas characterized by its particle number, temperature and pressure is a canonical example. In practice, the precise microstate (comprising the positions and momenta of all individual gas particles or the precise spectrum and occupation numbers) is usually neither accessible nor of particular interest; much more so the question, whether a sensible self-contained reduced description in terms of equivalence classes corresponding to (essentially) the same values of a few thermodynamic variables can be established on microscopic grounds.

To address this question, it is practical to generalize the definitions in Eq. (87):

$$\hat{\rho}(t) = \sum_{\nu} \rho_\nu |\nu(t)\rangle\langle\nu(t)| \ , \quad \hat{\rho}(\{\mathbf{q}_i,\mathbf{p}_i\},t) = \int \mathrm{d}\nu \, \rho_\nu \prod_i^N \delta[\mathbf{q}_i - \mathbf{q}_i^{(\nu)}(t)]\delta[\mathbf{p}_i - \mathbf{p}_i^{(\nu)}(t)] \tag{89}$$

The normalized weights ρ_ν quantify how much the pure states or trajectories (are thought to) contribute to the macrostate of the system characterized by $\hat{\rho}$. Normalization for both the weights ρ_ν and for the individual microstates or trajectories reads

$$\mathrm{tr}\hat{\rho} = \sum_{\mu,\nu} \rho_\nu \langle\mu|\nu\rangle\langle\nu|\mu\rangle = \sum_{\mu,\nu} \rho_\nu \langle\nu|\mu\rangle\langle\mu|\nu\rangle = \sum_{\nu} \rho_\nu = 1 \ , \tag{90}$$

$$\int \prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \mathrm{d}\nu \rho_\nu \, \delta[\mathbf{q}_i - \mathbf{q}_i^{(\nu)}(t)]\delta[\mathbf{p}_i - \mathbf{p}_i^{(\nu)}(t)] = \int \mathrm{d}\nu \, \rho_\nu = 1 \ . \tag{91}$$

The predictions for measurements are now formulated in terms of averages over the ensemble, in the simplest case in the form of expectation values,

$$\langle \hat{A}(t)\rangle \equiv \mathrm{tr}\big(\hat{\rho}(t)\hat{A}\big) \ , \qquad \langle A(t)\rangle \equiv \int \prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \, \hat{\rho}(t) A(\{\mathbf{q}_i,\mathbf{p}_i\}) \ , \tag{92}$$

but variances and correlation functions (averages over multivariate expressions) will also be of interest. *Examples:* $\langle\hat{1}\rangle = 1$ (normalization); $\langle\hat{\rho}\rangle = \mathrm{tr}\hat{\rho}^2 \leq 1$ (“=” for a pure state).

Note the dramatic difference between Eqs. (87) and Eqs. (89). The first two describe a lone single trajectory (or pure state) in an otherwise completely empty, high dimensional state space. In contrast, the second set of equations describes an ensemble comprising an overwhelmingly large number of possible microstates, i.e. a collection of an extremely dense tangle of trajectories or states compatible with some coarse macroscopic criteria (e.g., a given energy of temperature, etc.). In fact, there are more than astronomically many microstates (typically on the order of e^N), against which the “rather few” ($N \simeq 10^{23}$) dimensions of phase space pale to nothing. Since for almost anything any atom does at any moment, the opposite action is also part of the ensemble, it is plausible (though practically almost always impossible to prove) that this tangle of states or trajectories basically fills the available state space densely and more or less uniformly, hence cannot evolve anywhere else in time, and the system is, broadly construed, said to be in equilibrium. Usually, much weaker requirements suffice, e.g. that the projection of the phase space tangle into a low-dimensional state space of very few (local) thermodynamic variables is well enough represented by a uniform density for the purpose of calculating some thermodynamic averages. To understand how, and under precisely which circumstances, this intuition fails (spontaneously or due to external constraints), which it may do in a most perplexing and intricate way, is the grand unsolved puzzle investigated in *non-equilibrium* statistical mechanics.

No matter whether the system is in or out of equilibrium, the dynamics of the (bare and ensemble averaged) densities is always ruled by the equations of *von Neumann*

$$\frac{\hbar}{i}\partial_t\hat{\rho} = [\hat{\rho}, \hat{H}] \equiv \hat{\rho}\hat{H} - \hat{H}\hat{\rho}, \tag{93}$$

and *Liouville*,

$$\partial_t\hat{\rho} = \{\hat{\rho}, H\} \equiv -\mathcal{L}\hat{\rho}\ , \qquad \mathcal{L} \equiv \sum_i (\partial_{\mathbf{p}_i} H)\partial_{\mathbf{q}_i} - (\partial_{\mathbf{q}_i} H)\partial_{\mathbf{p}_i} \tag{94}$$

with commutator $[,]$, Poisson-bracket $\{,\}$ and Liouville-operator $\mathcal{L}$. They are related by the correspondence principle $i[,]/\hbar \leftrightarrow \{,\}$. In a (macroscopically) stationary[21] state, and in particular in equilibrium, $\hat{H}$ and $\hat{\rho}$ commute,

$$\partial_t\hat{\rho} = 0 \quad \Leftrightarrow \quad [\hat{H}, \hat{\rho}] = 0\ , \quad \{H, \hat{\rho}\} = 0\ . \tag{95}$$

An ensemble density $\hat{\rho}$ that describes equilibrium and therefore does not change in time can thus be treated as a function that depends on the microscopic degrees of freedom only indirectly via the Hamiltonian. Other conserved quantities are rarely directly accessible, in practice. Below, this is used to motivate the basic postulate of statistical mechanics, namely to approximate the ensemble density of an isolated system by the particularly simple solution $\hat{\rho}(\{\mathbf{q}_i, \mathbf{p}_i\}) = \hat{\rho}\big(H(\{\mathbf{q}_i, \mathbf{p}_i\})\big)$ of this equation, which happens to be constant on "energy shells" ($H = E =$const.) in phase space for a given set of thermodynamic parameters.

The next two subsections introduce two important general ideas that play a crucial role for the construction of statistical mechanics as a general theory of many identical particles with short-ranged interactions. (With some care, basic elements of the discussion will be generalizable to not so many particles or long-ranged interactions.)

Indistinguishable particles and μ-space

In practice, for a given many-particle system, one is often interested in the particle density $n(\mathbf{r}, t)$ at a certain position $\mathbf{r}$ in 3-dimensional position space or rather in a 6-dimensional one-particle phase space consisting of positions (then usually denoted by $\mathbf{q}$) and momenta $\mathbf{p}$. One thus asks for the chance to find any particles

---

[21] Stationarity of the ensemble averaged density does not imply that nothing moves—the microstates may still fluctuate wildly within the constraints defined by the ensemble.

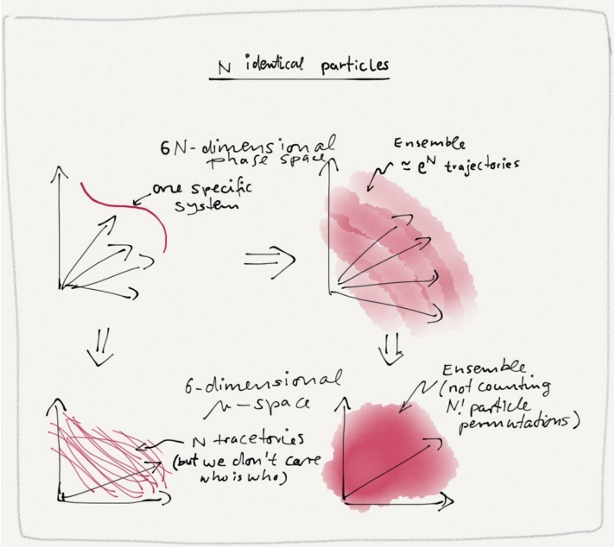


Figure 2: Trajectories and ensembles in phase space and $\mu-$space.

(not a certain particle) at some point in a $2d$-dimensional projection of the $2dN$-dimensional phase space, called 1-particle phase space or μ-space (to avoid the misunderstanding that it only contains one particle). Accordingly, it is useful to define the *1-body density*

$$f_1(\mathbf{q},\mathbf{p},t) \equiv \sum_i \langle \delta[\mathbf{q}-\mathbf{q}_i(t)]\delta[\mathbf{p}-\mathbf{p}_i(t)]\rangle = N\int \prod_{i>1} \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i\, \hat{\rho} \tag{96}$$

to find any particle in the state $\{\mathbf{q},\mathbf{p}\}$ in $\mu-$ *space*. More ambitiously, one may want to know the weight to find any M particles in a certain correlated state (e.g., for $M=2$, to find any two particles with certain velocities at a certain distance, etc.), no matter which particles. This information is then encoded in the *M-body densities*

$$f_M(\mathbf{q}_1,\mathbf{p}_1,\ldots \mathbf{q}_M,\mathbf{p}_M,t) \equiv \frac{N!}{(N-M)!}\int \prod_{i>M} \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_{\mathrm{i}}\, \hat{\rho}\;, \tag{97}$$

which have the form of joint probabilities (or correlation functions) to find any (no matter which) M particles with the given positions and momenta.

In particular, the N-body density f_N is formally identical to the phase space density $\hat{\rho}$ of Eq. (89) up to the factor $N!$ indicating that it does not care about which particle is where. Like all the M-body densities, it however has a different interpretation with respect to the space on which it is defined. The $N!$ different phase space trajectories corresponding to all of the (otherwise identical) N particle-systems obtained by all possible permutations of the particles map to a single trajectory in a projective space, in which the permutations are not distinguished. Going one step further, namely to a description in μ-space, they are represented by N trajectories (reflecting the presence of N particles) in this $2d$-dimensional space. The one-body density f_1 is simply the trajectory density in μ-space and is normalized to the particle number N. Similarly, by virtue of their normalization, the M-body densities f_M are or joint densities for finding M correlated trajectories in μ-space, i.e. $M-$point correlation functions.

In fact, the admittedly somewhat perverse mapping from Liouville's phase space to μ-space is argued to better represent the proper phase space structure of a system of N *practically indistinguishable particles occupying the same configuration space*, for which permutations should rather not be counted as new states if the notion of indistinguishability is taken seriously. The difference between Liouville- and μ-space may seem of minor relevance at first sight but matters as soon as a full thermodynamic fundamental relation has to be computed from statistical mechanics, which amounts to counting the total number of available states or to measuring the accessible phase space volume, respectively. Below, this aspect will be illustrated by a calculation of the entropy of an ideal gas of N (indistinguishable) particles, which fails to be extensive if the reduction of phase space volume implied by the mapping to μ-space is overlooked.

Law of large numbers for homogeneous systems

The last remark alludes to a general crucial issue in equilibrium statistical mechanics, namely the recovery of extensive thermodynamics in the limit of large system size for many-body systems with short-ranged interactions. In particular, one expects large equilibrium systems to behave deterministically with vanishing fluctuations of the thermodynamic quantities, "typical" and "average" being the same, which then implies a most crucial property of such systems: *that a measurement performed on one particular real-world realization of the members of the ensemble at one instance practically always corresponds to a measurement over the complete ensemble.*

All this can indeed be demonstrated for large homogenous systems. Due to the short range of the microscopic interactions, they can be subdivided at essentially no cost (neglecting only sub-leading surface terms in the energy) into many small, yet macroscopic (compared to the molecular scale) volume elements. The volume elements may then be used to define local densities a of extensive quantities $\mathcal{A}$. Fluctuations $\Delta\mathcal{A} \equiv \mathcal{A} - \langle\mathcal{A}\rangle$ of an extensive quantity $\mathcal{A}$ around its average $A \equiv \langle\mathcal{A}\rangle = V\langle a\rangle$ may then be expressed in terms of these densities a, i.e. $\Delta a = a - \langle a\rangle$. Since the volume elements are only weakly coupled in the complete system, local averages $\langle a(\mathbf{r})\rangle$ decorrelate ("quickly loose their mutual dependence") for distances $|\mathbf{r}-\mathbf{r}'|$ longer than the range ℓ of interactions. Accordingly, the average mean-square fluctuations of an observable can be rewritten in the integral form

$$\Delta A^2 \equiv \langle \Delta \mathcal{A}^2 \rangle = \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \, \langle \Delta a(\mathbf{r}) \Delta a(\mathbf{r}') \rangle \simeq \langle \Delta a^2 \rangle V \ell^3 \, . \tag{98}$$

The final estimate follows from the form of the (spatial) 2-point density

$$\langle \Delta a(\mathbf{r}) \Delta a(\mathbf{r}') \rangle \sim \begin{cases} \langle \Delta a^2(\mathbf{r}) \rangle = \langle \Delta a^2(\mathbf{r}') \rangle = \langle \Delta a^2 \rangle & (|\mathbf{r} - \mathbf{r}'| \ll \ell) \\ \langle \Delta a(\mathbf{r}) \rangle \langle \Delta a(\mathbf{r}') \rangle = 0 & (|\mathbf{r} - \mathbf{r}'| \gg \ell) \end{cases} .$$

Take for example a chain of people who have a (weak) tendency to align their vote $s = \pm 1$ with that of their neighbors in a range ℓ, but no systematic preference for either vote. The spatial average over $s(\mathbf{r})$ vanishes because of the randomly changing sign of $s(\mathbf{r})$ along the chain. Similarly, $s(\mathbf{r})s(\mathbf{r} + \Delta\mathbf{r})$ can take either of two values ± 1. But positive correlation $(+)$ is more likely than negative correlation or anticorrelation $(-)$ for $|\Delta\mathbf{r}| \ll \ell$, independent of $\mathbf{r}$.

The estimate in Eq. (98) suggests that *large additive extensive systems will effectively behave deterministically*, since $\Delta A/A \propto V^{-1/2}$. Moreover, *typical values and average values coincide* to increasing relative precision for $V \to \infty$, which allows to regard a particular macroscopic system at a particular time as a representative sample of the ensemble. A macroscopic amount of gas or liquid are the standard examples for this general expectation. Extensive variables in such systems are spatially self-averaging. The measurement of an extensive quantity of a single system amounts to the measurement of a density, summed up over many practically independent identical sub-systems to yield a virtually deterministic outcome. In other words, for additive systems "one is many", effectively, and average is typical, so that "any is average". (The probabilistic version of the central limit theorem, which is discussed in the exercises, indeed asserts that the probability distribution for extensive variables that can be represented as sums of effectively independent and identically distributed stochastic variables of finite variance quickly approaches an incredibly sharp Gaussian with this property, if the number of terms in the sum is not very small.) For such systems, there is therefore no need to worry about ensembles in everyday life, as a single system behaves like an ensemble-averaged system, already. Introducing ensembles is therefore sometimes an unnecessary, merely theoretically motivated exercise, and it can indeed be useful to consider one concrete system, instead. Special care is generally needed near critical points, though, where the correlation length diverges, correlations decay like power-laws, and certain limits, such as $t \to \infty$, $T \to T_c$, and $V \to \infty$, may fail to interchange.

It should be said that it even often makes sense to apply the usual statistical mechanics formalism to systems that fail to meet the prerequisites for applying the law of large numbers, since they are small in the sense that they only have few degrees of freedom. The above discussion does not apply, and they are not self-averaging by themselves, but this can be remedied if the small system is coupled to a bath. One can then measure the time average of an observable for the small system, which will usually provide an excellent match to an ensemble average. A system for which time and ensemble averages coincide, i.e., which explores the allowed phase space of the ensemble on a reasonable time scale, is called *ergodic*. Tossing a single coin is an ergodic process. If repeated N times, it leads to the same statistics as tossing N coins simultaneously (whence the saying "time is money"). Systems that are self-averaging in the above sense, are ergodic. However, you clearly would not like your portfolio to be ergodic in this sense, although you would like it to be self-averaging to minimize volatility.

As the last remarks suggests, there are serious counterexamples to and exceptions from the law of large numbers. These are naturally suspected among the systems that are "small" in the sense that they are non-extensive and not trivially decomposable into additive compartments, such as proteins, neutron stars, or atomic nuclei. Their number of available microstates per macrostate is highly constrained due to strong internal correlations. But also "large" systems, for which the number of possible microstates grows exponentially with system size, can pose problems, since one might well need an exponentially large number of weakly coupled compartments (hence an exponentially large volume) to exhaust the ensemble by a spatial average. Think of the little volume elements being exposed to so-called quenched disorder, e.g. an external field that takes different but fixed values for each of them, with extreme values being very rare but relevant for the mean. Even worse, a window glass may seem to be an equilibrium solid for all practical purposes, if one adopts Feynman's definition of equilibrium as the state when all the fast things have happened but the slow things have not. Yet, philosophically, it is possibly better characterized as a highly viscous undercooled liquid evolving (and aging) extremely slowly in time. Finally, if you ask people how much money they have, or how many friends or sex partners, you will discover that these numbers are power-law distributed, so that the average value can be very different from the most probable value, if it exists at all[22]. The "Petersburg paradox", a drastic example for such unruly behavior is discussed in the exercises. It gives a hint why tax payers always end up paying the free lunch of smart bankers that have everything under control except for some rare events (and, sure enough, they do not eat at the university canteen). The unintuitive behavior of all such strange systems indicates that they are far from equilibrium. However, it is not generally possible to tell whether or when a system is ergodic or will fall out of equilibrium by looking at either the system itself or its Hamiltonian (think of the glass transition in hard sphere suspensions). But it is quite easy to write down innocent-looking toy models that can keep larger parts of the statistical-physics community busy for decades.

---

[22] A true power-law distribution over the (positive) real numbers is of course an idealization, which will usually need to be cut off somewhere in practical applications (e.g., the probability to find someone with 10^{23} sex partners is exactly zero and not a very small positive number).

The infinite staircase

With these general considerations in mind, one may try and see what can be learned directly from the microscopic formulation. Consider N identical interacting classical particles with the pair-Hamiltonian

$$H = \sum_i \frac{\mathbf{p}^2}{2m} + \sum_i \mathcal{U}(\mathbf{q}_i) + \sum_{i<j} \mathcal{V}(|\mathbf{q}_i - \mathbf{q}_j|) \ . \tag{99}$$

With the abbreviations

$$\mathbf{F}_i \equiv -\partial_{\mathbf{q}_i} \mathcal{U}(\mathbf{q}_i) \qquad \text{and} \qquad \mathbf{K}_{ij} \equiv -\partial_{\mathbf{q}_i} \mathcal{V}(|\mathbf{q}_i - \mathbf{q}_j|) = -\mathbf{K}_{ji} \tag{100}$$

for the external and internal forces, respectively, the Liouville equation takes the form

$$[\partial_t + \mathcal{L}]\, \hat{\rho} = 0 \ , \qquad \mathcal{L} \equiv \sum_i \left(\frac{\mathbf{p}_i}{m} \cdot \partial_{\mathbf{q}_i} + \mathbf{F}_i \cdot \partial_{\mathbf{p}_i} \right) + \frac{1}{2} \sum_{i,j} \mathbf{K}_{ij} \cdot \left(\partial_{\mathbf{p}_i} - \partial_{\mathbf{p}_j} \right) \ . \tag{101}$$

Here $\hat{\rho}$ is the statistical phase space density from Eq. (89) for which there is no explicit expression so far. In a straightforward (but ambitious) approach to Statistical Mechanics you might want to solve Eq. (101) and then take the long-time limit $t \to \infty$, where the dynamics hopefully comes to rest, to find the $\hat{\rho}$ corresponding to the equilibrium distribution.

As a more modest starting point, one may investigate the consequences of Eq. (101) for the one-body density in $\mu-$space,

$$\partial_t f_1 = -N \int \prod_{i>1} \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \, \mathcal{L} \hat{\rho} \ . \tag{102}$$

The terms resulting from operators with indices i, $j > 1$ in $\mathcal{L}$ are total derivatives (their prefactors are independent of the variables in the derivatives) that vanish upon integration and evaluation at the borders, where ρ is assumed to vanish:

$$\partial_t f_1 = -N \int \prod_{i>1} \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \left[\frac{\mathbf{p}_1}{m} \cdot \partial_{\mathbf{q}_1} + \mathbf{F}_1 \cdot \partial_{\mathbf{p}_1} + \sum_{j>1} \mathbf{K}_{1j} \cdot \partial_{\mathbf{p}_1} \right] \hat{\rho} \ . \tag{103}$$

The first two operator terms on the right can be taken out of the integral, so that with Eq. (96) the corresponding contributions can be transferred to the left. Altogether, one thus has

$$\begin{aligned} \left[\partial_t + \frac{\mathbf{p}_1}{m} \cdot \partial_{\mathbf{q}_1} + \mathbf{F}_1 \cdot \partial_{\mathbf{p}_1} \right] f_1 &= -\, N \int \prod_{i>1} \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \sum_{j>1} \mathbf{K}_{1j} \cdot \partial_{\mathbf{p}_1} \hat{\rho} \\ &= -\, N(N-1) \int \mathrm{d}\mathbf{q}_2 \mathrm{d}\mathbf{p}_2 \, \mathbf{K}_{12} \cdot \partial_{\mathbf{p}_1} \int \prod_{i>2} \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i \, \hat{\rho} \\ &= -\int \mathrm{d}\mathbf{q}_2 \mathrm{d}\mathbf{p}_2 \, \mathbf{K}_{12} \cdot \partial_{\mathbf{p}_1} f_2(\mathbf{q}_1, \mathbf{p}_1, \mathbf{q}_2, \mathbf{p}_2, t) \ . \end{aligned} \tag{104}$$

While the left hand side characterizes the free streaming of the density f_1 in $\mu-$space, the right hand side is identified as its change due to two-body interactions that depend on the two-body density f_2. The equation is thus not closed. An additional equation specifying the time evolution of f_2 is of course readily derived along the same lines, but it turns out to depend on f_3 and so on. Such an infinite hierarchy of equations clearly needs to be truncated to make it practically useful. Boltzmann's famous kinetic equation results if one estimates the contributions from f_2 by the factorization approximation $f_2 \propto f_1 f_1$, which inflicts only small errors for dilute gases. Hence, one can get useful results if one takes the interacting pairs to be uncorrelated prior to their encounter, which corresponds to the famous assumption of "molecular chaos". (The chaos should be chaotic enough that there is no need to worry about situations where the two particles have recently already met or where the first particle collided with a third particle that hit a fourth particle that had an encounter with the second, etc.). Boltzmann's strategy to derive statistical averages and correlation functions using the assumption of an underlying molecular chaos has turned out to be a key idea of Statistical Mechanics. In contrast, Boltzmann's equation and the whole approach based on this BBGKY[23] hierarchy relies on truncating the hierarchy, and is therefore limited to dilute gases. And no matter how well you treat a dilute gas, it will not become a dense fluid at the end of the day. This is a good excuse to abandon the discussion of Boltzmann's approach via the infinite staircase and look for the elevator to the summit. After all the exhausting climbing it is worthwhile, though, to pause for a little intermezzo and consider at least the trivial case of an ideal (non-interacting) gas, explicitly, if only to anticipate the whole discussion of the remainder of this section in a nutshell.

3.2 First little slide-down with the ideal gas

A special solution

For non-interacting particles the infinite staircase collapses and the stationary one-body density f_1 can be obtained from the much simpler equation

$$\left[\frac{\mathbf{p}}{m} \cdot \partial_{\mathbf{q}} + \mathbf{F} \cdot \partial_{\mathbf{p}}\right] f_1 = 0 \; . \tag{105}$$

It is obviously solved by any differentiable function $f_1(H)$. With a separation ansatz in $\mathbf{q}$ and $\mathbf{p}$ follows the (special) solution

$$f_1(\mathbf{q},\mathbf{p}) = N e^{-\beta H_1(\mathbf{q},\mathbf{p})}/C \; , \quad H_1(\mathbf{q},\mathbf{p}) \equiv \frac{\mathbf{p}^2}{2m} + \mathcal{U}(\mathbf{q}) \; , \tag{106}$$

$$C \equiv \int \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}\, e^{-\beta H_1} = (2\pi m/\beta)^{3/2} \int \mathrm{d}\mathbf{q}\, e^{-\beta \mathcal{U}} \; . \tag{107}$$

with an arbitrary separation constant β.

---

[23] Named after Bogoliubov, Born, Green, Kirkwood, Yvon.

Connection to thermodynamics

To get an idea of the physical meaning of β, consider the average energy of the gas without external field ($\mathcal{U} \equiv 0$)

$$U \equiv \langle H_1 \rangle_{f_1} = \frac{\langle \mathbf{p}^2 \rangle_{f_1}}{2m} = \frac{1}{2m} \int \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}\, \mathbf{p}^2 \, f_1 = \frac{3N}{2\beta} \; . \tag{108}$$

The average is taken with the one-body phase-space density f_1 that is normalized to N. Comparison with the caloric equation of state of an ideal gas as introduced in thermodynamics suggests

$$\beta^{-1} = k_B T \; . \tag{109}$$

Incidentally, this also proofs the so-called *equipartition theorem*: Eq. (108) shows that the thermal energy stored on average in a velocity-degree of freedom (which is penalized with a quadratic term in the Boltzmann factor) is $k_B T/2$. It is clear that this result still holds in presence of potential interactions that do not depend on the momenta, and that the practical rule generalizes to any term in the Hamiltonian that is quadratic in whatever degree of freedom. So, each equilibrated quadratic degree of freedom will carry the thermal energy $k_B T/2$.

An elementary argument (due to Bernoulli) allows to connect the atomistic picture also to macroscopic gas *mechanics*. The pressure onto the container walls is calculated as the force per unit area. The momentum transfer in a single collision with the x-wall is $2p_x$. There are np_x/m collisions per unit area and time, with n the particle density, exerting a pressure (only particles moving in the right direction contribute)

$$P = \frac{2n}{m} \int_0^\infty \mathrm{d}p_x \, \frac{p_x^2 e^{-\beta p_x^2/2m}}{(2\pi m/\beta)^{1/2}} = \frac{\langle p_x^2 \rangle_{f_1}}{Vm} = \frac{\langle \mathbf{p}^2 \rangle_{f_1}}{3mV} = \frac{2}{3}\frac{U}{V} = \frac{n}{\beta} \tag{110}$$

on the walls. This can indeed again be reconciled with the thermodynamic equation of state using Eq. (109). Equations (108), (109), and (110) link the microscopic description to macroscopic thermostatics for the ideal gas.

In the exercises you moreover verify that Boltzmann's interpretation of

$$\mathfrak{H} = \langle \ln f_1 \rangle = \int \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}\, f_1 \ln f_1 \tag{111}$$

as the negative entropy (in units of k_B) does yield the proper thermodynamic expression. Notice that the normalization of f_1 to N (rather than to one) both in front of and inside the logarithm, renders Boltzmann's $\mathfrak{H}-$function extensive.

The foregoing discussion may seem slightly odd, as it appears to be formulated for only one single particle rather than an ensemble of N particles. The particle number only enters via the normalization of the one-particle distribution function. This perspective may resonate well with the dictum by Wheeler and Feynman that there actually is "only one electron in the universe" (i.e. one electron-positron field with N excitations that are totally antisymmetrized in space and can move forward and backward in time). Yet, a more comprehensive description of an ensemble with the functions f_M that can encode correlations induced by interactions is often highly desirable. For the interaction-free ideal gas, all M-body densities, which have the form of joint probabilities, factorize into products of single-particle densities. In particular, the N-body density f_N reads

$$f_N(\{\mathbf{q}_i, \mathbf{p}_i\}) = N! \prod_{i=1}^{N} e^{-\beta H_1(\mathbf{q}_i,\mathbf{p}_i)}/C = N! e^{-\beta \sum_i H_1(\mathbf{q}_i,\mathbf{p}_i)}/C^N \ . \tag{112}$$

The factor $N!$ is a reminder that all permutations of the particles are considered as equivalent and not counted as independent states. While the original density $\hat{\rho}$ was normalized to a phase space integration, its appropriate replacement in the case of N indistinguishable particles, namely f_N, is normalized with respect to an integration over the projective phase space, in which all $N!$ different trajectories obtained upon particle permutations are identified. Because of the complicated topology of such a space, it is of course much more convenient to integrate over the original phase space and then divide by the number of permutations, i.e. to use $\prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i/N!$ as the phase space measure[24] for f_N, which is then indeed normalized to one rather than to $N!$. In the exercises you show that, with this convention,

$$S/k_B = -\langle \ln f_N \rangle = \int \frac{\prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i}{N!} f_N \ln f_N \tag{113}$$

is another way of writing the entropy of an ideal gas of N particles. Notice again how extensivity is realized, this time by the factor $N! \approx N^N$ inside the logarithm.

To summarize, N indistinguishable non-interacting particles occupying the same configuration space are described by the phase space density,

$$\rho \equiv h^{3N} f_N = \frac{e^{-\beta H}}{Z} \ , \quad H(\{\mathbf{q}_i, \mathbf{p}_i\}) = \sum_i \frac{\mathbf{p}_i^2}{2m} + \mathcal{U}(\mathbf{q}_i) \ , \tag{114}$$

with normalization, phase space integral, and averages, respectively, defined by

$$Z \equiv \int \mathrm{d}\Gamma \, e^{-\beta H} \qquad \mathrm{d}\Gamma \equiv \frac{\prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i}{N! h^{dN}} \ , \qquad \langle A \rangle \equiv \int \mathrm{d}\Gamma \, A \rho \ . \tag{115}$$

Below, this form is generalized to allow for interactions of the particles, simply by extending the free Hamiltonian to an interacting Hamiltonian. The supposition

---

[24] This may seem strange. After a lengthy discussion of why one should use $f_N = N!\rho$ as appropriate replacement for the original phase space weight ρ, the factor $N!$ is divided out again, in the measure. It is only the latter, though, that remains visible in practical calculations, while its opponent remains hidden inside the new definition of ρ.

that this simple rule works is a formulation of the fundamental postulate of Equilibrium Statistical Mechanics, namely the one appropriate for isothermal systems. Note that Planck's constant h has been smuggled into Eqs. (114) & (115) as an elementary phase space volume to render ρ and the integral measure both dimensionless. Thanks to this slight amendment the interpretation of the phase space volume Γ (and the normalization Z) is no longer that of an abstract measure of the (relevant) *continuous* phase space volume but something more mundane: a ridiculously large but in principle—not in practice—*countable number of discrete states*. An explicit demonstration how the classical phase space measure emerges from a quantum statistical description can be found below in Sec. 4.3, where the classical limit is more closely examined.

For the ideal gas, explicit expressions can be computed for the so-called *canonical partition sum*

$$Z(\beta, V, N) = \int \mathrm{d}\Gamma\, e^{-\beta H} = \int^{\infty} \mathrm{d}\varepsilon\, g(\varepsilon) e^{-\beta\varepsilon} = \frac{V^N (2\pi m/\beta)^{3N/2}}{N! h^{3N}} \,, \tag{116}$$

the *phase space volume*

$$\Gamma(E, V, N) \equiv \int \mathrm{d}\Gamma\, \theta(E - H) = \int^{E} \mathrm{d}\varepsilon\, g(\varepsilon) = \frac{V^N (2\pi m E)^{3N/2}}{N!(3N/2)! h^{3N}} \,, \tag{117}$$

and its derivative at energy E, *the density of states*

$$g(E) \equiv \frac{\mathrm{d}\Gamma}{\mathrm{d}E} = \frac{V^N 2\pi m (2\pi m E)^{3N/2-1}}{N!(3N/2-1)! h^{3N}} = \frac{3N\Gamma}{2E} \,. \tag{118}$$

Approximating the factorials by $x! \approx x^x$ helps to establish the close relations between $g(E)$, $\Gamma(E)$ and $Z(\beta)$. If integrated over a small energy interval (or multiplied by the energy E/N per particle), the density of states $g(E)$ basically gives the volume of the energy shell at energy E, which essentially contains the total phase space volume $\Gamma(E)$, in a high dimensional space (i.e. for large particle numbers N), as further discussed in the exercises. Also, the partition sum Z is seen to boil down to the phase space volume Γ if one identifies the energy per particle with the thermal energy $3k_BT/2$ and uses $N! \approx N^N$. So, up to minor technical details, the phase space measures Γ, g, and Z can all basically be identified for large N.

Fluctuations

Beyond the calculation of thermodynamic quantities, the framework of statistical mechanics enables us to analyze fluctuations around the average behavior. For instance, the mean-square fluctuations

$$\Delta U^2 \equiv \langle (H - \langle H \rangle)^2 \rangle = \langle H^2 \rangle - \langle H \rangle^2 \,, \tag{119}$$

of the total internal energy U of the ideal gas are explicitly given by

$$\Delta U^2 = \frac{1}{4m^2}\langle\big(\sum_i p_i^2\big)^2\rangle - \langle H\rangle^2 = \frac{N}{4m^2}\left(\langle p^4\rangle - \langle p^2\rangle^2\right) = \frac{3N}{2\beta^2}\ . \tag{120}$$

For the squared sums $(\sum_i p_i^2)^2 = \sum_i (p_i^2)^2 + \sum_{i\neq j} p_i^2 p_j^2$ was used, and that the terms in each sum are all equal upon averaging, the sum over the mixed terms giving $N(N-1)\langle p^2\rangle^2$. In particular, one observes that the ideal gas supports the general expectation for all many-body systems with short-range interactions that the absolute strength of fluctuations scales like the square-root in the system size, hence

$$\Delta U/U \simeq N^{-1/2}\ . \tag{121}$$

A more comprehensive characterization of the energy fluctuations is provided by the *energy distribution*

$$\begin{aligned} w(E) \equiv& \langle\delta(E-H)\rangle \\ =& \int \mathrm{d}\Gamma\, \frac{e^{-\beta H}}{Z}\delta(E-H) = \int \mathrm{d}\varepsilon\, g(\varepsilon)\,\frac{e^{-\beta\varepsilon}}{Z}\delta(E-\varepsilon) = \frac{g(E)}{Z}e^{-\beta E}\ . \end{aligned} \tag{122}$$

Note that the essential specific information is again contained in the density of states $g(E)$. In the exercises, you show that the resulting energy distribution for the ideal gas takes a Gaussian form in the thermodynamic limit,

$$w(E) = \frac{\beta(\beta E)^{3N/2-1}e^{-\beta E}}{(3N/2-1)!} \overset{N\to\infty}{\sim} \frac{1}{\sqrt{2\pi}\Delta U}\exp\left[-\frac{(E-U)^2}{2(\Delta U)^2}\right]\ . \tag{123}$$

So in this particularly simple case, the information gained by calculating the distribution rather than its mean and variance, which was already achieved in Eqs. (108)) and (120), respectively, does not seem too precious. A Gaussian is fully determined by its mean and variance, after all. But in more complicated situations, the shape of the distribution could be much more interesting (and, unfortunately, much harder to calculate). In the following, the strategy is to postulate that the basic properties (not the detailed numerical form) of the functions calculated above for the ideal gas are generic beyond the realm of gases, for all many-body systems with short-range interactions — if they are in equilibrium.

3.3 The elevator to the summit

The general strategy is to argue or to hope that the structure discovered for the isothermal ideal gas carries over to interacting many-particle systems with short-range interactions. The notions motivated by the example of the ideal gas in the previous section are then often understood as a derived version of a “simpler” postulate; namely, that the equilibrium statistical mechanics of an *isolated* interacting many-body system may be obtained from a (fake) phase space weight $\rho(\{\mathbf{q}_i, \mathbf{p}_i\}) = \rho\big(H(\{\mathbf{q}_i, \mathbf{p}_i\})\big)$ that is uniform on the energy shell $H(\{\mathbf{q}_i, \mathbf{p}_i\}) = E$. This is called the micro-canonical ensemble. In the following, the classical so-called "Gibbs ensembles" are all introduced as independent postulates, though further below some remarks concerning their mutual dependence and the limitations of their applicability shall be made.

Micro-canonical Ensemble

The equilibrium weight (that reproduces the macrostates) of an isolated conservative system is the uniform distribution on the energy shell.

$$\rho(H; E, V, N) = \frac{\delta(H - E)}{g(E, V, N)}\,, \qquad g(E, V, N) = \int \mathrm{d}\Gamma\, \delta(H - E) = \frac{\mathrm{d}\Gamma}{\mathrm{d}E}\,. \tag{124}$$

This is the (unproven) fundamental postulate of Equilibrium Statistical Mechanics. The uniform distribution corresponds to a maximally chaotic or "featureless" distribution of trajectories in phase space. The true microscopic dynamics will generally not correspond to equal weights everywhere on the energy shell — but the hope is that pretending it does so causes no harm for all common purposes.

For the quantum mechanical prescription one considers $\hat{H}$ and its spectrum $\{E_\nu\}$, the trace tr replaces the phase space integration $\mathrm{d}\Gamma \equiv (\prod_i \mathrm{d}\mathbf{q}_i \mathrm{d}\mathbf{p}_i)/(N! h^{3N})$. The density of states and the total number of states with energies up to E read

$$g(E) = \sum_\nu \delta(E - E_\nu)\,, \qquad \Gamma = \int \mathrm{d}E\, g(E)\,. \tag{125}$$

Do not take any of the δ-functions too literally. As always in physics, they idealize a suitable narrowly peaked function. The δ-function in Eq. (124) is supposed to have width ϵ_0 (independent of system size, e.g. something about the typical energy of one particle). For a generic interacting system, it will usually be assumed that so many energy levels (i.e. δ-functions) are contained in this narrow peak region that a continuum approximation of $g(E)$ is appropriate[25].

Recall that for an ideal classical gas the phase space volume takes the explicit form

$$\Gamma(E, V, N) \equiv \frac{V^N (2\pi m E)^{3N/2}}{N!(3N/2)! h^{3N}}\,. \tag{126}$$

Upon introducing the de Broglie wavelength

$$\lambda_E \equiv \frac{h}{\sqrt{2\pi m (2E/3N)}} \tag{127}$$

---

[25] The generic expectation for an extensive system with short-range interactions is that $g(E)$ grows like some number to the power of $E/\epsilon_0 \simeq N$. For macroscopic systems, there may thus easily be many more accessible states than particles in the known universe, and a continuum description is not only well justified but mandatory. As a peculiarity, this generic structure is only recovered somewhat indirectly in the classical limit of an ideal quantum gas (see below).

of a particle of mass m with a kinetic energy of about E and approximating $N! \approx N^N$, one observes

$$\Gamma(E,V,N) \approx (V/N)^N/\lambda_E^{3N} \equiv \left(v/\lambda_E^3\right)^N . \tag{128}$$

This suggests the interpretation that the uncertainty relation provides a discretization of the configuration space of the gas into units of linear dimension λ_E. Moreover, with the interpretation of Γ as the number of available states, each of N non-interacting indistinguishable particles in a volume V is seen to have only $v/\lambda_E^3 = (n\lambda_E^3)^{-1}$ states available (with $v \equiv n^{-1} \equiv V/N$ the specific volume) — rather than V/λ_E^3, as one might naively have thought. This is clearly a direct consequence of the identification of states which only differ by particle permutations. It turns out to be the single most important insight for understanding some fundamental properties of condensed matter such as the temperature dependence of their specific heat, their electrical conductivity, and even transitions to macroscopically coherent quantum states like Bose-Einstein condensation, which can all be understood in terms of (effective) gases of elementary excitations.

Canonical Ensemble

In the limit of an infinite heat reservoir, the equilibrium weight function for an isothermal system is

$$\begin{aligned} \rho(H;\beta,V,N) &= \frac{\exp(-\beta H)}{Z(\beta,V,N)} , \\ Z(\beta,V,N) &= \int \mathrm{d}\Gamma \, \exp(-\beta H) = \int \mathrm{d}\epsilon \, g(\epsilon) \exp(-\beta\epsilon) . \end{aligned} \tag{129}$$

The translation rules to switch to a quantum mechanical description are as above.

As one usually considers Hamiltonians that are a quadratic form in the momenta, the momentum integration can be performed once and for all, which reduces the task of calculating the canonical partition sum to the so-called *configuration integral*

$$Z = \int \mathrm{d}\Gamma \, e^{-\beta[\sum_i \mathbf{p}_i^2/2m + \mathcal{V}(\{\mathbf{r}_j\})]} = \frac{1}{N!\lambda_T^{3N}} \int \mathrm{d}^N\mathbf{r} \, e^{-\beta\mathcal{V}(\{\mathbf{r}_i\})} . \tag{130}$$

Here, the *thermal wavelength*

$$\lambda_T \equiv \frac{h}{\sqrt{2\pi m k_B T}} \tag{131}$$

was introduced, which is easily recognized as the de Broglie wavelength of a particle of mass m with a thermal kinetic energy of about k_BT. To connect to the motto preceding this part of the lecture, one may pinpoint the computation of the configuration integral as the central practical task in statistical mechanics.

Needless to say, that this task can only in a few exceptional cases be performed analytically. For the ideal classical gas,

$$Z(T,V,N) = \frac{1}{N!} Z(T,V,1)^N = \frac{(V/\lambda_T^3)^N}{N!} \approx (v/\lambda_T^3)^N \; . \tag{132}$$

As for the micro-canonical case, interpreting Z as the number of available states and using $N! \approx N^N$, one recovers the observation that for each particle in a gas of N non-interacting indistinguishable particles in a volume V there are only v/λ_T^3 (instead of V/λ_T^3) states available.

Grand-canonical ensemble

In the limit of an infinite heat and particle reservoir, the weight function for a system in thermal and chemical equilibrium is

$$\begin{aligned} \rho(H,N;\beta,V,\mu) &= \frac{\exp[-\beta(H-\mu N)]}{Z_G(\beta,V,\mu)} \\ Z_G(\beta,V,\mu) &= \sum_{N=0}^{\infty} \int \mathrm{d}\Gamma \, \exp[-\beta(H-\mu N)] \; . \end{aligned} \tag{133}$$

The translation rules to switch to a quantum mechanical description are as above, N becomes the number operator $\hat{N}$. To emphasize that compared to other players in the game $N \simeq 10^{23}$ is not such a large number after all, the discrete notation is often preferred. The thermodynamic conjugate to the particle number N is the *chemical potential* (in units of $k_B T$). Its Boltzmann factor $z \equiv e^{\beta\mu}$ is called the *fugacity.* Note that

$$Z_G(\beta,V,z) = \sum_N z^N Z(\beta,V,N) \quad \Rightarrow \quad Z(\beta,V,N) = \frac{1}{N!} \partial_z^N Z_G|_{z=0} \; , \tag{134}$$

which is why Z_G is sometimes called the generating function of Z.

The fluctuating particle number often simplifies the task of summing over all states, since it allows finite sums to be extended to infinity. In return, the task of relating the results thus obtained back to a system of fixed known particle number (but unknown μ) involves the identification of that particle number with

$$\langle N \rangle = z \partial_z \ln Z_G = \partial_{\beta\mu} \ln Z_G \tag{135}$$

and the elimination of z with help of this equation. For an isothermal gas or fluid in a closed box, the canonical partition sum Z pertains to the whole box, whereas the grand canonical sum Z_G is useful to address local properties in the bulk of the material, such as local density fluctuations.

Consider again the ideal classical gas as the standard example, for which Eq. (132) implies

$$Z_G(T,V,z) = \sum_N z^N Z(T,V,N) = \exp[zZ(T,V,1)] = e^{zV/\lambda_T^3} = e^{\langle N\rangle} \,. \tag{136}$$

In the last step Eq. (135) was used to eliminate the fugacity z in favor of the average particle number $\langle N\rangle = Vn = Vz/\lambda_T^3$ (or density n). The discussion in the preceding paragraph identified the inverse of $n\lambda_T^3$ as the number of available states per particle. This is now seen to imply that the fugacity is the crucial dimensionless parameter controlling the importance of quantum effects in ideal gases. A more thorough discussion of ideal quantum gases (see Sec. 4.2) indeed reveals that gases condense into a single macroscopically occupied quantum state for $n\lambda_T^3 \approx 1$.

"Deducing" open ensembles

Instead of postulating the canonical, grand-canonical and other generalized ensembles constructed along the same lines, one may derive the equilibrium distributions for these ensembles from the uniform distribution of the micro-canonical ensemble. The basic strategy is to divide a micro-canonical system Σ into a system S with Hamiltonian H and a much larger reservoir R, with a weak (short-range) coupling term in their Hamiltonians that needs not to be specified explicitly, since it is dropped in the energy balance $H_\Sigma = H + H_R$ anyway (imagine it as something non-extensive, like a surface term). For long-range interactions between system and reservoir this strategy obviously does not work, which hints at some limitations. It moreover requires strong assumptions about the part of the micro-canonical system treated as the reservoir. Basically, one wants the bath to be a classical ideal gas, which clearly comes pretty close to postulating the result from the outset. One generates the distribution of the system S from that of Σ by integrating out (averaging over) the degrees of freedom of the reservoir R,

$$\rho = \int \mathrm{d}\Gamma_R \, \frac{\delta(H + H_R - E_\Sigma)}{g_\Sigma(E_\Sigma)} = \frac{g_R(E_\Sigma - H)}{g_\Sigma(E_\Sigma)} = \frac{g_R(E_\Sigma - H)}{g_R(E_\Sigma)} \frac{g_R(E_\Sigma)}{g_\Sigma(E_\Sigma)} \,. \tag{137}$$

The purpose of the last step is to facilitate entering the strong assumptions about the reservoir. Using the classical ideal gas form Gl. (118) for the density of states of the reservoir and taking the thermodynamic limit ($N_R \approx N_\Sigma \to \infty$), so that $(1 + x/N)^N \sim e^x$, one finds

$$\frac{g_R(E_\Sigma - H)}{g_R(E_\Sigma)} \to \left(1 - \frac{H}{E_\Sigma}\right)^{\frac{3}{2}N_R} \to \exp(-\beta H) \,, \qquad \beta = \frac{3N_R}{2E_\Sigma} \to \frac{3N_{\Sigma,R}}{2E_{\Sigma,R}} \,. \tag{138}$$

More generally, if g_R has another form (that should however exhibit a comparably dramatic growth as for the ideal gas[26]) one expands $\ln g_R$ with respect to $H \ll E_\Sigma$,

$$\ln\big[g_R(E_\Sigma - H)/g_R(E_\Sigma)\big] = -\beta H \ldots , \qquad \beta = \partial_{E_{\Sigma,R}} \ln g_R(E_{\Sigma,R}) , \tag{139}$$

which yields a more general definition of the intensive constant β that controls the energy exchange in the canonical ensemble. The thermodynamic limit is taken such that $R \sim \Sigma \to \infty$ for a fixed system size; i.e., there is a strong scale separation between the reservoir and the system. If the system itself is large, there is a double scale separation with respect to the atomic scale, but the subsystem may in principle be chosen arbitrarily small, provided that one does not mind fluctuations of a comparable order of magnitude as the absolute values. The observation that weakly coupled subsystems of homogeneous macroscopic systems are "thermal" in the sense that they obey a canonical distribution is very general, and the applicability of the canonical ensemble therefore extends far beyond the limitations of the above derivation (and maybe even its underlying assumption of a microcanonical ensemble for the total system).

3.4 Recovering thermodynamics

Generally, since the different ensembles correspond to different degrees of isolation of a thermodynamic system, the precise identifications of the formalisms of statistical mechanics and thermostatics may be expected to exhibit slight differences for the different ensembles.

Fundamental relations and the first law

The discussions in Sec. 3.2 (isothermal ideal gas) and in the preceding section (isothermal interacting many-body system) suggested the identification

$$\beta = 1/k_B T = \partial_E \ln g(E) \tag{140}$$

for the parameter controlling the energy partition in the canonical ensemble. Comparison with the thermodynamic relation

$$(\partial_E S)_V = 1/T \tag{141}$$

then suggests that the identifications $T = 1/k_B\beta$ and $S = k_B \ln g(E)$ hold beyond the special case of the ideal gas. To render the argument of the logarithm dimensionless one re-introduces the small energy width ϵ_0 of the energy shell, which can be taken on the order of the energy of a single particle. Writing the number of states in the energy shell as

$$\mathsf{W} \equiv \epsilon_0 g(E) \tag{142}$$

---

[26] Hence, one should refrain from expanding $g_R(E-H)/g_R(E)$ directly.

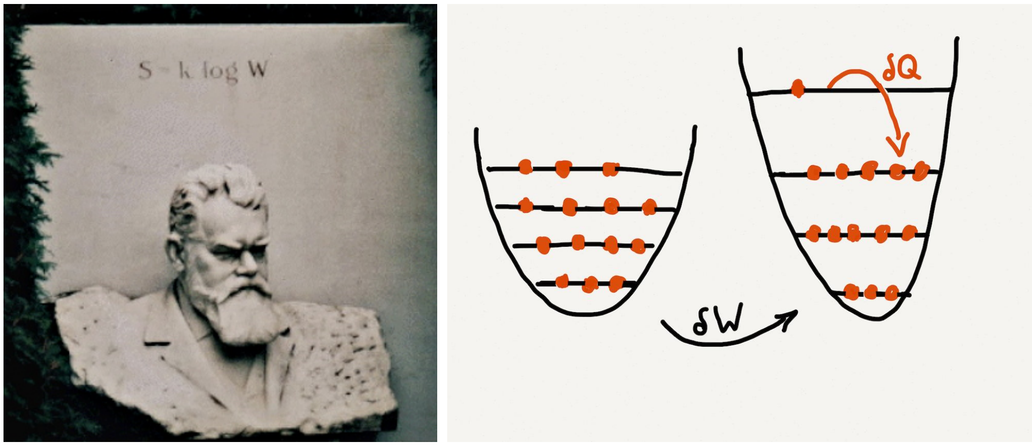


Figure 3: Boltzmann's epitaph with his ground-breaking insight in Planck's modernized formulation (left) and visualization of work and heat as (adiabatic) shifting of energy levels and redistribution of occupancies, respectively (right).

harmonizes the notation with that on Boltzmann's gravestone[27] (Fig. 3)

$$\boxed{S = k_B \ln \mathsf{W}} \tag{143}$$

Recall that $g(E, V, N)$ comprises the full statistical information about an isolated system. Accordingly, the equation on the gravestone represents a full thermodynamic fundamental relation $S(E, V, N)$ once the density of states is known as a function of E, V and N. This is why Eq. (143) is often regarded as the central equation of statistical mechanics. Note that for S to be an extensive quantity, g and W must obviously exhibit an exponential dependence on the particle number, i.e., they must be of the form $(\text{something intensive})^N$. In other words, the phase space volume (or number of states) available to a classical gas in a bottle grows by a fantastic factor of about $10^{10^{20}}$ if one allows the gas to extend its volume by a factor of two. Even bigger is your surplus in phase space over your neighbor's if you consume a proper beer rather than his half beer (which amounts to taking the square). Also note that, in line with the identification of heat as the energy stored in macroscopically unresolved degrees of freedom, this observation suggests to interpret entropy (up to the conventional factor k_B) as the *number of the unresolved degrees of freedom* (or the missing information compared to whatever might be considered a complete microscopic description).

The above motivation of the relation between the formalisms of statistical mechanics and thermostatics was still quite heuristic and formal. It becomes more compelling if the analogy is first established for the work sector, which is

---

[27] "The range of validity and the proper interpretation of this relation are unclear to me" (E.G.D. Cohen in *Dynamics of Dissipation* 2002)

independent of the notions of temperature and entropy. This is achieved in a most simple and elegant way in the canonical ensemble. The parametric change $\delta\hat{H}$ of the Hamiltonian (in a form like $\delta\hat{H} = \hat{X}\delta f$, $\hat{M}\delta h$ etc., as a product of conjugate generalized "position" variables and "force" parameters) is identified as the microscopic work increment[28]. It changes the kinematics of the system and thereby its energy levels. According to the first law, the remaining contribution to the total energy, which is due to the reshuffling of the occupancies of the energy levels (i.e. the statistics rather than the kinematics, cf. Fig. 3), is identified as heat. For brevity, the quantum mechanical notation is employed:

$$\mathrm{d}U = \mathrm{d}\langle\hat{H}\rangle = \langle\delta\hat{H}\rangle + \mathrm{tr}(\hat{H}\delta\hat{\rho}) = \delta W + \delta Q \ . \tag{144}$$

Now, note that $\mathrm{tr}(\delta\hat{\rho}) = 0$ (normalization). Hence, using $\hat{\rho} = e^{-\beta\hat{H}}/Z$, one finds

$$\mathrm{d}\langle\ln\hat{\rho}\rangle = \mathrm{tr}(\ln\hat{\rho}\ \delta\hat{\rho}) = -\beta\mathrm{tr}(\hat{H}\delta\hat{\rho}) = -\beta\delta Q \ , \tag{145}$$

which establishes β as an integrating factor for the heat increment δQ. A comparison with the Carnot theorem then establishes $\beta = 1/k_B T$ and the corresponding expressions for the canonical entropy

$$S = -k_B\langle\ln\hat{\rho}\rangle = \langle\hat{H}\rangle/T + k_B\ln Z(T,V,N) \ , \tag{146}$$

and, consequently, the canonical free energy

$$U - TS = F(T,V,N) = -k_B T\ln Z(T,V,N) \ . \tag{147}$$

In complete analogy, one shows the relation

$$pV = -J(T,V,\mu) = k_B T\ln Z_G(T,V,\mu) \tag{148}$$

between the *grand canonical* potential J and partition sum Z_G. Clearly, both Eq. (147) and Eq. (148) represent — like the formula on Boltzmann's gravestone — thermodynamic fundamental relations that contain the complete thermodynamic information.

The second law

Given the differences between the ensembles and the different expressions for the entropy, one may well be worried whether the differently defined entropies do indeed take a maximum in their respective ensembles, as required by the second

---

[28] There is a slight subtlety hidden in this argument, as already noted in the discussion of thermostatics, namely that the work a system can perform typically depends on our practical ability to access certain (possibly internal) system variables. Therefore, one person's heat may to some extent be another (more sophisticated) person's work. Also note that there is no hermitian operator for work, hence it is not an observable in the usual sense.

law. This is, in fact, automatically guaranteed by the sufficiently intelligent construction of the Gibbs ensembles. One can show this, starting from a definition of the entropy in terms of the density operator as in Eq. (146). Under various thermodynamic constraints, the Gibbs expressions for the density operators then follow from the requirement that they maximize this so-called "Gibbs" or "information" entropy in equilibrium, as demanded by the second law. Take as an example an *isolated system* and assume that it can under certain circumstances be described by a general normalized phase space density operator $\hat{\tilde{\rho}}$ with normalized eigenstates $|\tilde{\nu}\rangle$ and eigenvalues $\tilde{\rho}_{\tilde{\nu}}$, which is required to give the right thermodynamic averages. For the corresponding Gibbs (information) entropy

$$\tilde{S} \equiv -k_B \mathrm{tr} \hat{\tilde{\rho}} \ln \hat{\tilde{\rho}} \tag{149}$$

one can show that it is never larger than the Boltzmann entropy

$$S = k_B \ln \mathsf{W} = k_B \langle \ln \mathsf{W} \rangle \ , \tag{150}$$

for which the density used in the calculation of the average clearly does not matter. The information entropy is at best equal to the Boltzmann entropy, when $\hat{\tilde{\rho}}$ is equal to the uniform micro-canonical density (i.e. $\tilde{\rho} = \rho = 1/\mathsf{W}$ on the energy shell). To formally demonstrate the inequality between the two entropies, one introduces an operator-1 constructed from the eigenstates $|\nu\rangle$ of the microcanonical density operator $\hat{\rho}$ (with eigenvalues ρ_ν) into Eqs. (149), (150), while formulating the trace using the eigenstates $|\tilde{\nu}\rangle$,

$$\frac{\tilde{S} - S}{k_B} = \sum_{\nu, \tilde{\nu}} |\langle \nu | \tilde{\nu} \rangle|^2 \tilde{\rho}_{\tilde{\nu}} \ln \frac{\rho_\nu}{\tilde{\rho}_{\tilde{\nu}}} \leq \sum_{\nu, \tilde{\nu}} |\langle \nu | \tilde{\nu} \rangle|^2 \tilde{\rho}_{\tilde{\nu}} \left(\frac{\rho_\nu}{\tilde{\rho}_{\tilde{\nu}}} - 1 \right) = \mathrm{tr}\hat{\rho} - \mathrm{tr}\hat{\tilde{\rho}} = 0 \ . \tag{151}$$

(For the final equality the initial steps of the calculations were reverted.) Starting from the definition Eq. (149) one has thus "derived" the microcanonical equilibrium form of the density and the entropy by insisting on the maximum principle embodied in the second law: *the Gibbs entropy $\tilde{S}$ of an isolated conservative system takes its maximum $S = k_B \ln \mathsf{W}$ (the Boltzmann entropy) for a uniform phase space density on the energy shell, i.e. if the system is in equilibrium as postulated by the micro-canonical ensemble.* Note that the only essential ingredients in the derivation of this result are the independence of the Boltzmann entropy of the density used for averaging in Eq. (150), and the convexity of the logarithm.

The entropy $\tilde{S}$ is called information entropy, because it is used in this form, except for the different base of the logarithm and the conventional factor k_B, in computer science and information theory. There it serves as a measure for our ignorance, which is defined as the complement to our bias or *a priori* information about the state of a system that is statistically characterized by $\tilde{\rho}$. Equation (151) then has the interpretation that our bias about the true microscopic distribution on the energy shell is minimal for the micro-canonical density. This does of course nothing to prove that it is the "right" or "best" choice to mimic any actual microscopic distribution. Yet it is clearly helpful to think of entropy in terms of ignorance or "missing information[29]" about the microstate of a given thermodynamic system.

A similar procedure as above can be applied to an isothermal system to derive the canonical distribution (exercises). One finds that the Gibbs entropy $\tilde{S}$ of a system constrained to have an average energy $\langle H\rangle = U$ takes its maximum $S = (U-F)/T$ (the Legendre transform of the free energy) for a canonical form of the phase space density, i.e. if the system is in equilibrium as defined by the canonical ensemble. This is of course nothing but the familiar *minimum principle for the free energy* F (the second law for an isothermal system) turned around.

This essentially completes the recovery of thermodynamics, although it should be noted, as an outlook, that the maximum entropy or minimum free energy principle can be generalized to (and even understood as a consequence of) the simple equality

$$\langle e^{-\beta W_{\text{diss}}}\rangle = 1 \tag{152}$$

for the dissipated work $W_{\text{diss}} \equiv W - F$. The latter is here understood as a fluctuating quantity, since the system at hand is assumed not to be in the thermodynamic limit. The average refers to an ensemble of realizations of the same isothermal system subjected to the same prescribed external work protocol that may take the system far from equilibrium. This and similar so-called *fluctuation theorems* can be shown to hold on very general grounds, as discussed in Part III of the lecture. The second law in the form of Eq. (47), which says that the free energy is the minimum average work needed to achieve a certain state change of an isothermal system (or that the dissipation cannot be negative on average), follows from Eq. (152) by applying the inequality $e^{-x} \geq 1 - x$.

Also note that along the same lines as in Eq. (151), replacing $\hat{\rho}$ by a time-dependent quasi-static density operator $\hat{\rho}(t)$ and $\hat{\tilde{\rho}}$ by $\hat{\rho}(t=0)$, one can easily show that the entropy $-\langle k_B \ln \rho(t)\rangle$ averaged with $\hat{\tilde{\rho}} = \hat{\rho}(0)$ increases with time[30]. Physically, it means that if one generates a phase space weight $\rho(0)$ according to the maximum entropy (minimum free-energy) rule such that it properly predicts some thermodynamic data at $t = 0$, and then lets the system — and its distribution $\hat{\rho}(t)$ — evolve in time under the same conditions, its entropy will grow except if $\hat{\rho}(t)$ does not evolve, in which case the system happened to be in equilibrium at $t = 0$, already. That such an increase in entropy is only observed in one time direction (called future) in everyday life and not in both, as one would expect from such a formalism, is an indication that we (and in fact the whole universe) experience conditions very far from equilibrium[31].

---

[29] Arieh Ben-Naim, A Farewell to Entropy: Statistical Thermodynamics based on Information

[30] Wilhelm Brenig, Statistical Theory of Heat. Nonequilibrium Phenomena; Springer 1989.

[31] J. L. Lebowitz, Boltzmann's Entropy and Time's Arrow, Physics Today, 9/1993, 32; D. Wallace, Gravity, Entropy, & Cosmology: in Search of Clarity, Br. J. Phil. Sci. 61 (2010) 513

Partial traces and free energies

Note that Eq. (147) for the free energy may be rephrased as

$$Z(T,V,N) = \operatorname{tr} e^{-\beta \hat{H}} = e^{-\beta F(T,V,N)} \ . \tag{153}$$

The final form emphasizes that the Hamiltonian turns into a free energy upon integrating/tracing out the microscopic degrees of freedom. This statement generalizes to partial traces. The calculation of a (partial) partition sum can be understood as a formal prescription to coarse-grain a more microscopic description into a more macroscopic one. In this sense, a free energy is nothing but a coarse grained Hamiltonian, and the distinction between the two gets blurred. An example is provided by the Landau free energy discussed in the Thermodynamics Part of the lecture, where the order parameter is kept constant upon taking the trace,

$$\operatorname{tr}|_{\psi}\, e^{-\beta \hat{H}} = e^{-\beta L(\psi;T,V,h)} \ , \tag{154}$$

or in fact by any of the free energies of the restrained ensembles. Stretching the analogy between Hamiltonian and free energy to its extreme, one may take the trace over fluctuations on successively coarser length scales iteratively and attempt to absorb the coarse-graining effect in a renormalization of the phenomenological parameters of the Hamiltonian. This produces a flow of Hamiltonians in parameter space, which is known as the renormalization group flow. Critical points, which entail self-similar (fractal) fluctuations on all scales, correspond to fixed points of this flow, an insight rewarded with the Nobel prize in 1982.

Mathematical relation between the ensembles

The various ensembles are related in a very systematic way. Starting point is an isolated conservative system described by the micro-canonical ensemble, which is completely specified by its density of states $g(E,V,N)$ as a function of energy, volume and particle number. The density of states (or, more precisely, $\mathsf{W} = \epsilon_0 g$) is the micro-canonical partition sum. If the system is brought into contact with reservoirs with which it is allowed to exchange some extensive quantities, the new fundamental relation (or partition sum) for the appropriate larger ensemble is obtained by Laplace transformations with respect to the liberated quantities. In the thermodynamic limit, the Laplace transformation boils down to the (much simpler) Legendre transformation. As a scheme:

$$\begin{aligned} g(E,V,N) &\xrightarrow{\text{Laplace trafo}} Z(T,V,N) \xrightarrow{\text{Laplace trafo}} Z_G(T,V,\mu) \\ S(E,V,N) &\xrightarrow{\text{Legendre trafo}} F(T,V,N) \xrightarrow{\text{Legendre trafo}} J(T,V,\mu) \end{aligned} \tag{155}$$

The reason for the drastic simplification from a Laplace to a Legendre transformation is the thermodynamic limit $N,\ E,\ V \to \infty$, which allows to evaluate the

Laplace integral in a saddle-point approximation. To leading order, this amounts to the approximation of the integral by the maximum of its integrand (see below). As demonstrated above for the example of the ideal gas, the integrand is (up to a normalization constant) nothing but the distribution function of the liberated quantity (in the example, the energy E), which is fixed in the restrained ensemble but allowed to fluctuate in the open one. Only in the thermodynamic limit, when this distribution becomes infinitely sharp and the relative magnitude of the fluctuations decays to zero, all the ensembles and the corresponding thermodynamic potentials are essentially[32] equivalent. The following section dwells on this a bit more and makes this statement more precise.

3.5 Fluctuations

Beyond recovering thermodynamics and gaining a method to compute (at least in principle) constitutive laws from microscopic Hamiltonians, another benefit of statistical mechanics is that it provides information about thermal fluctuations and their spatial and temporal correlations. The study of fluctuations moreover provides additional insight into the role of the various definitions of the thermodynamic potentials in the various ensembles and their mutual relations.

Liberté, Egalité, Stabilité

Recall that the Laplace transform that takes you from the micro-canonical to the canonical ensemble and liberates energy from the micro-canonical energy shell is simply the partition integral; and that the partition function $Z(\beta)$ is nothing but the normalization of the canonical energy distribution function $w(E)$,

$$Z(\beta) = \int \mathrm{d}E\, g(E) e^{-\beta E} \;, \qquad g(E)e^{-\beta E} \propto w(E) \equiv \langle \delta(H-E) \rangle \;. \tag{156}$$

Here $\langle \dots \rangle$ refers to the canonical average. The integral may be rewritten in the form

$$Z(\beta) = \int \frac{\mathrm{d}E}{\epsilon_0}\, e^{\ln \mathsf{W}(E) - \beta E} = \int \frac{\mathrm{d}E}{\epsilon_0}\, e^{-\beta A(E,\beta)} \;, \tag{157}$$

with the micro-canonical free energy

$$A(E,\beta) \equiv E - \beta^{-1} \ln \mathsf{W}(E) \;. \tag{158}$$

It is taken for granted that $N^{-1} \ln \mathsf{W}(E)$ approaches an intensive function of the energy in the thermodynamic limit, so that the integral can be evaluated in a saddle point approximation — which requires that a diverging multiplicative factor N can be isolated in the exponent. The position of the maximum of the energy distribution and the minimum of $A(E,\beta)$ is the energy giving the

---

[32] One may still worry about situations with diverging fluctuations, e.g. phase transitions.

dominant contribution to the integral. This is found to be the (canonical) average energy $U(\beta)$. Beyond the leading order, the saddle-point approximation consists in expanding $A(E,\beta)$ to second order in E around this minimum and performing the Gaussian integral. With the abbreviation

$$A_0(U,\beta) \equiv U - \beta^{-1} \ln \mathsf{W}(U)$$

for the minimum of the micro-canonical free energy, the expansion reads

$$A(E,\beta) \approx A_0(U,\beta) + \left(1 - \beta^{-1}\partial_E \ln \mathsf{W}|_U\right)(E-U) - (2\beta)^{-1}\partial_E^2 \ln \mathsf{W}|_U (E-U)^2$$

and one recovers the micro-canonical definition of temperature,

$$(k_BT)^{-1} = \beta = \partial_E \ln \mathsf{W}(E)|_U \ , \tag{159}$$

since the linear term must vanish at the minimum. The coefficient of the quadratic term is thereby identified as $(2\beta)^{-1}\partial_U\beta = -(2T\partial_T U)^{-1} = -(2TC_V)^{-1}$, hence

$$A(E,T) \approx A_0(U,T) + (E-U)^2/(2TC_V) \ . \tag{160}$$

Finally, the (to this order) Gaussian integration in Eq. (157) yields the asymptotically exact result

$$Z(T) \sim e^{-A_0(U,T)/k_BT} \int \frac{\mathrm{d}E}{\epsilon_0}\, e^{-\frac{(E-U)^2}{2k_BT^2C_V}} = \sqrt{2\pi C_V/k_B}\,\frac{k_BT}{\epsilon_0} e^{-A_0(U,T)/k_BT} \ . \tag{161}$$

Note that for *small systems*, the equivalence of the ensembles is troubled by non-extensive terms, such as the prefactor of the exponential in the last term, which is of order $\mathcal{O}(N^{1/2})$, corresponding to a positive logarithmic $\mathcal{O}(\ln N)$ term by which A_0 exceeds the canonical free energy $F(T) = -k_BT \ln Z(T)$. So, the canonical expression is seen to be superior to the microcanonical one under isothermal conditions, and the central thermodynamic minimum condition is seen to emerge from the saddle-point condition. Only *in the thermodynamic limit* the non-extensive corrections can be dropped, so that the micro-canonical and canonical definitions of temperature, energy, entropy, and free energy (average or not) all coincide, demonstrating the *asymptotic equivalence of the ensembles*:

$$F(T) = -k_BT \ln Z(T) \sim A_0(U,T) = U - k_BT \ln \mathsf{W}[U(T)] \sim U - TS \ . \tag{162}$$

Fluctuations and response coefficients in open ensembles

From Eq. (160), the energy fluctuations are seen to be controlled by the square of the temperature and the heat capacity C_V, the response coefficient linking energy to temperature. Its inverse provides the restoring force that restrains the energy fluctuations. For positive C_V, the variance of the mean-square fluctuations is proportional to C_V, hence extensive, and their square root $\Delta U \propto C_V^{1/2} \propto N^{1/2}$

(the standard deviation) is therefore non-extensive. The relative importance of fluctuations vanishes as $\Delta U/U \propto N^{-1/2}$ in the system size. For finite C_V, the positivity condition $C_V > 0$ thus guarantees *thermodynamic stability* against fluctuations (even at formally negative temperatures). In fact, $C_V > 0$ is always satisfied in the canonical ensemble:

$$k_B T^2 C_V = -\partial_\beta \langle H \rangle = \partial_\beta^2 \ln Z = Z^{-1} \partial_\beta^2 Z - (\partial_\beta \ln Z)^2 = \langle \Delta H^2 \rangle \geq 0 \,. \tag{163}$$

The observation can be generalized. It is a special case of the fluctuation-dissipation theorem, saying that the strength of the thermal fluctuations of thermodynamic quantities are controlled by response coefficients, more precisely, by the response coefficient that quantifies the response of the fluctuating variable to its conjugate intensive parameter — the one that controls its repartition throughout the system. When going from the canonical to the grand canonical ensemble, the roles of energy, temperature and heat capacity in Eq. (163), and in the foregoing discussion, are taken by particle number, chemical potential, and the (isothermal) compressibility $\kappa_T = v\partial_p n)_T = v^2 \partial_\mu n)_T$, respectively. The corresponding fluctuation-response relation reads

$$n k_B T \kappa_T = \langle \Delta N^2 \rangle / \langle N \rangle \geq 0 \,. \tag{164}$$

The divergence of the compressibility at phase coexistence entails a divergence of the number (density) fluctuations.

Phase coexistence in open and closed ensembles

Two immediate general insights gleaned from Eqs. (163), (164) are that *fluctuations grow with the square-root of the system size* and *response coefficients cannot become negative in open ensembles.* Response coefficients vanish, susceptibilities diverge, and fluctuations become extensive at phase transitions, however. Moreover, the same response coefficients then turn negative but stay finite (except at critical points), in the more restrained ensembles, where fluctuations are suppressed. They rightly denounce phase transitions as thermodynamical (mechanical) instabilities. Consider the example of a grand canonical density distribution

$$w(N) \equiv \langle \delta_{N,M} \rangle \propto z^N Z(T,V,N) = e^{-\beta V [\mathfrak{f}(T,n) - \mu n]} \,, \tag{165}$$

In complete analogy to Eqs. (156), (157) the (left out) normalization of $w(N)$ is

$$Z_G(T,V,\mu) = V \int \mathrm{d}n \, e^{-\beta V [\mathfrak{f}(T,n) - \mu n]} \,, \tag{166}$$

where the sum over N has been turned into an integral over n, for convenience. In case of two coexisting phases (vapor and liquid) the canonical (Landau) free energy $F(T,n) = V\mathfrak{f}(T,n)$ develops a double-well form with minima at the densities

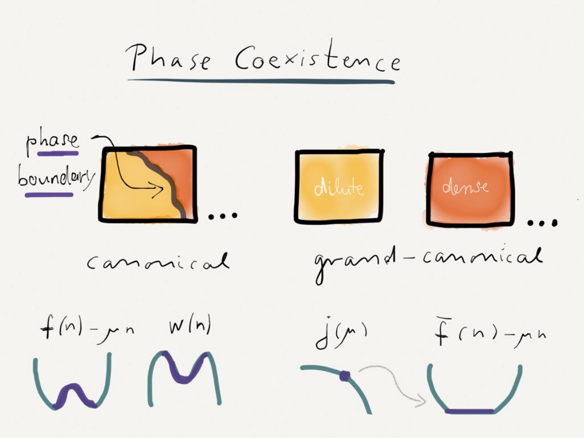


Figure 4: Phase coexistence in the canonical and grand-canonical ensemble. The hump/valley in the canonical free energy $\mathfrak{f}(n) - \mu n$ and density distribution $w(n)$ indicates the thermodynamic instability of the homogeneous system with respect to phase separation at the cost of creating phase boundaries. It is ironed out by the saddle-point approximation involved in the calculation of $j(\mu) \equiv J(\mu)/V$ via a Laplace transform, which only sees the largest saddle point (the deepest free energy minimum) and thus amounts to a Maxwell construction $\mathfrak{f}(n) \to \bar{\mathfrak{f}}(n)$.

$n_<$ and $n_>$. The energy distribution $w(N)$, or $w(n)$, accordingly has a bimodal form with two peaks (Fig. 4). In a finite container of fixed volume, phase coexistence implies phase boundaries. They cost some non-extensive[33] surface free energy showing up as a non-extensive hump between the free-energy minima and the corresponding valley between the two peaks of $w(n)$, respectively.

In contrast, in an open container that allows particles to be exchanged with a bath, if the average particle number is raised by adjusting the chemical potential, the content of the container can switch between the dense all-liquid and the dilute all-vapor state from one ensemble element to the next (Fig. 4). Thereby, the costly phase boundary is avoided at the price of large ensemble fluctuations, thermodynamically reflected by a *kinked form of the grand canonical potential* $J(\mu) = -Vp(\mu) = -k_BT \ln Z_G$, which is nothing but the flipped-over version of the function $\mu(p)$ discussed in the Thermodynamics Part of the lecture. The kink is now understood as a consequence of the saddle-point approximation of the integral in Eq. (166) and arises when the largest saddle point jumps from one of the free energy minima of the double-welled free energy $\mathfrak{f}(n) - \mu n$ to the other, upon tilting the free-energy landscape by varying μ. Now, going backwards from this kinked potential $J(\mu)$ to a density-dependent free energy yields a different form $\bar{\mathfrak{f}}(n)$ that is no longer double-welled but has a flat region (a Maxwell construction!), where the compressibility diverges — and hence the fluctuations as

---

[33] …unless the interactions are long-ranged, as e.g. for the van der Waals gas.

well. The function $\bar{\mathfrak{f}}(n)$ is the grand canonical version of the free energy (Fig. 4). The corresponding density distribution $\bar{w}(n)$ then also has a flat region, i.e., a bridge over the valley between the two peaks of the canonical $w(n)$. In summary,

- For many-body systems with short-ranged interactions the ensembles are thermodynamically equivalent in the thermodynamic limit, up to non-extensive fluctuations ($\propto N^{1/2}$) of the liberated extensive variables.
- In open ensembles, response coefficients of liberated variables are proportional to their mean-square fluctuations. In contrast to the restrained ensembles, they thus remain non-negative (but diverge) at phase transitions.
- The saddle-point approximation of the integral over fluctuations involves the minimum condition for the free energies. It turns the free energies of the restrained ensembles, which become non-convex at a first-order phase transition, into a kinked convex potential for the open ensemble. Going backwards from the kinked open-ensemble to the restrained-ensemble free energy, divergent response coefficients imply flat (but marginally convex) free energies, thereby generating a Maxwell construction.

Chapter summary

Finally, standing on the summit of statistical mechanics and enjoying the impressive panorama, contemplate for a moment on what has been achieved. Thermodynamics has *not* been explained on the basis of the microscopic dynamics, as one might have hoped. Instead, the true dynamics was replaced by an assumption about its maximally chaotic character, at least with regard to the usually considered low dimensional (thermodynamic) projections. This helped to short-cut the intractable equations of motion and to propose a computational scheme how to predict thermodynamic quantities and their fluctuations for systems with short-ranged interactions on the basis of a crude caricature of the overall phase space structure *directly from the microscopic Hamiltonian.* A systematic recipe for this coarse-graining procedure was given, namely the computation of a (partial) partition sum that is closely related to the density of states or the phase space volume. This yields the wanted thermodynamic fundamental relations in the form of free energies. The postulated weights for the phase-space density were seen to take different forms, depending on the thermodynamic boundary conditions — and the associated thermodynamic potentials accordingly. But the various ensembles were found to be thermodynamically equivalent in the thermodynamic limit, up to some (usually subdominant) fluctuations. A one-to-one relation was discovered between fluctuations and response coefficients, which constrains the former and suggests non-invasive (scattering, microscopy, . . .) techniques to measure the latter. What remains to be done is to slide down from the summit and plunge into the rich world of applications of the established formalism.

4 The merry slide-down

The remaining task is basically the calculation of paradigmatic partition sums. Due to the tremendous reduction of information from the microscopic Hamiltonian to the thermodynamic fundamental relations or correlation functions, it is plausible to expect a highly redundant mapping from *huge classes* of possible microscopic systems to only a few characteristic types of thermodynamic behavior. (Indeed, we have innumerable different atomic and molecular constituents to make up fluids but only one Navier–Stokes equation to describe fluid flow.) An economic approach will therefore try to single out a minimalistic Hamiltonian of ultimate simplicity for each class and try to figure out the corresponding thermodynamics.

As the partition sum of non-interacting classical systems disintegrates into a product of single-particle contributions in the open ensembles, solving a single element obviously amounts to solving the whole. However, they give no clue as to the rich complexity and the phase transitions possible in interacting many-body systems. The remainder is therefore mostly devoted to two easily analyzable model systems that are a bit more complex: the Ising spin chain with nearest neighbor interactions, as an example for an interacting classical system; and ideal quantum gases. Somewhat paradoxically, the latter have a phase transition but not the former (at least no "proper" one with finite transition temperature. The reason is that, on the one hand, the exchange symmetries of indistinguishable particles give rise to effective exchange interactions of range λ_T that persist even if the potential interactions between the particles are completely neglected, while, on the other hand, a one-dimensional Ising chain with finite-range interactions corresponds to an interacting system at its lower critical dimension. Despite being somewhat pathological examples for generic (interacting) many-body systems, the studied model systems turn out to be remarkably helpful. The solutions for the mean-field version of the Ising model (corresponding to an infinite dimensional space or infinite-range interactions, respectively) and for the one-dimensional Ising chain provide a nice bracket for what really happens in interacting many-particle systems with finite-range interactions in spaces of more interesting dimensionality. Ideal quantum gases are the template for the zoo of elementary excitations or quasi-particles that capture the physics of solids and fluids in the quantum regime.

4.1 Interacting many-body systems

4.1.1 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 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

$$H(\{s_i\}) = -J\sum_{\mathrm{ip}} s_i s_j - h\sum_i s_i \tag{167}$$

Here ip stands for a sum over all N_{ip} interacting pairs. Two very useful idealizations are provided by the extreme choices that the pair interactions are either (1) *nearest-neighbor interactions* ($\sum_{\mathrm{ip}} \to \sum_{\langle ij\rangle}$, $N_{\mathrm{ip}} = Nq/2$, with q being the coordination number, i.e. the number of nearest neighbors per lattice site), or (2) van-der-Waals type *infinite-range interactions* ($\sum_{\mathrm{ip}} \to \sum_{i<j}$, $N_{\mathrm{ip}} = N(N-1)/2$). 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 model.

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 Fischer zeros finally elucidated the general mathematical mechanism how a singular free energy may emerge from a sum of infinitely many positive analytical Boltzmann factors.

4.1.2 Mean-field theory

In infinite dimensions or for infinite-range interactions, i.e. whenever the number of neighbors of each site is extensive, the interactions with the neighbors create an essentially deterministic field at that site. Solving the model then boils down to the evaluation of the partition sum of a single spin in this self-generated field. More precisely, writing $J = I/q$ for the interaction strength of the q interacting spin pairs in aribtrary space dimension, and setting $q = N$, similar as done for the molecular attractions in the van der Waals gas, one can rewrite the Boltzmann factor (dropping a non-extensive term $\beta I/2$) as

$$\begin{aligned} \exp(-\beta H) &= \exp\Big[\frac{\beta I}{2N}\sum_{i\neq j} s_i s_j + \beta h\sum_i s_i\Big] \\ &= \Big(\frac{\beta I N}{2\pi}\Big)^{1/2}\int_{-\infty}^{+\infty} \mathrm{d}\zeta\, \exp\Big[-\frac{1}{2}\beta I N\zeta^2 + \beta(I\zeta + h)\sum_i s_i\Big]\,, \end{aligned} \tag{168}$$

Up to the ζ-integration, the calculation of the partition sum Z of the ferromagnet is thereby reduced to that for a *paramagnet*[34] *in an effective field* $I\zeta + h$:

$$\begin{aligned}
Z =& \sum_{\{s_i\}} \exp(-\beta H) \\
=& \Big(\frac{\beta I N}{2\pi}\Big)^{1/2} \int_{-\infty}^{+\infty} \mathrm{d}\zeta \, e^{-\beta I N \zeta^2/2} \sum_{\{s_i\}} e^{\beta(I\zeta+h)\sum_i s_i} \\
=& \Big(\frac{\beta I N}{2\pi}\Big)^{1/2} \int_{-\infty}^{+\infty} \mathrm{d}\zeta \, e^{-\beta I N \zeta^2/2} \prod_i \sum_{s_i} e^{\beta(I\zeta+h)s_i} \\
=& \Big(\frac{\beta I N}{2\pi}\Big)^{1/2} \int_{-\infty}^{+\infty} \mathrm{d}\zeta \, e^{-\beta I N \zeta^2/2} [2\cosh(\beta I\zeta + \beta h)]^N \\
=& \Big(\frac{\beta I N}{2\pi}\Big)^{1/2} \int_{-\infty}^{+\infty} \mathrm{d}\zeta \, e^{-\beta N a(\zeta,\beta,h)} \ , \\
\text{with} \quad & a(\zeta,\beta,h) \equiv I\zeta^2/2 - \beta^{-1} \ln[2\cosh(\beta I\zeta + \beta h)] \ .
\end{aligned} \tag{169}$$

The limit $N \to \infty$ calls for the saddle-point method, which allows to replace the integral, up to sub-dominant contributions, by the maximum of the integrand $\exp[-\beta \min_\zeta a(\zeta)]$. The position ζ_0 of the saddle point follows from the minimization of $a(\zeta)$ via

$$0 = \partial_\zeta a(\zeta) = I\zeta - I \tanh(\beta I\zeta + \beta h) \ . \tag{170}$$

Note from Eq. (169) that this minimum $a(\zeta_0,\beta,h)$ of $a(\zeta,\beta,h)$ then represents the thermodynamic free energy (per spin) $F(\beta,h)/N$. In fact, since it depends on the external field h, it should actually rather be called a Gibbs free energy or grand potential. According to the discussion in Sec. 3.5, the integral in Eq. (169) should thus run over the distribution of the extensive quantity conjugate to h, i.e. the (site) magnetization

$$\sigma \equiv \frac{1}{N} \sum_i \langle s_i \rangle \tag{171}$$

If the latter is written as the derivative of the free energy,

$$\begin{aligned}
\sigma =& - \lim_{N\to\infty} \partial_h F/N \\
=& - \partial_h|_{\zeta_0} a(\zeta_0) - \partial_{\zeta_0}|_h a(\zeta_0) \partial_h \zeta_0 \\
=& \tanh(\beta I\zeta_0 + \beta h) = \zeta_0 \ ,
\end{aligned} \tag{172}$$

one indeed recovers the above saddle point equation. The magnetization is therefore obtained as the solution of the implicit transcendental equation

$$\sigma = \tanh(\beta I\sigma + \beta h) \ . \tag{173}$$

---

[34] Observe from the second and third line in Eq. (169) that the partition sum of N independent spins in a field h is $(2\cosh\beta h)^N$.

This self-consistency equation for the order parameter is the major result of this paragraph, and it is the typical result of any mean-field theory for whatever underlying microscopic model. In the exercises you discuss this equation further and show that it implies the emergence of a spontaneous magnetization below a critical temperature, the Curie temperature T_c, given by

$$k_B T_c = I \; . \tag{174}$$

Also note that $Na(\zeta)$, the logarithm of the magnetization distribution, is the Landau free energy of the model. The familiar ψ^4-structure emerges if $a(\zeta)/I \equiv (V/N)\mathcal{L}(\zeta)$ is expanded for small ζ and βh,

$$\frac{a(\zeta)}{I} \approx (1-\beta I)\frac{\zeta^2}{2} + (\beta I)^3\frac{\zeta^4}{12} - \frac{h}{I}\zeta \propto \mathcal{L} \; , \tag{175}$$

which incidentally corroborates Eq. (174). For the Ising model with infinite-range interactions, we thus find Landau theory to be an exact theory in the vicinity of the critical point.

In summary, by allowing for mutual interactions between an extensive number of elements (degrees of freedom), one obtains so-called mean-field behavior: the interacting microscopic elements become effecetively decoupled and their mutual interaction is replaced by the interaction with a self-generated homogeneous effective field σ, which, for the Ising model, turns out to be the magnetization. If the actual microscopic interactions have only a *finite range*, the mean-field theory is still exact in *infinite dimensions*. The higher the dimensionality, the more nearest neighbors there are, the more stable is a bulky arrangement of atoms, and the less destructive is the effect of thermal fluctuations. Ultimately, in an infinite dimensional bulk material, each atom is surrounded by infinitely many neighbors, leading to a total slaving, even with short-ranged interactions. In this limit, there is no room left for critical fluctuations, just as in the case of infinitely long-ranged interactions in finite dimensions, and the critical behavior becomes mean-field like, as one says, i.e. of the van-der-Waals or Curie-Weiss type. The interesting question is whether there is a finite upper critical dimension d_u above which this formidable simplification occurs. As it turns out, the answer is positive. Moreover, mean-field theory even provides a decent approximation for a finite window of dimensions between an upper and a lower critical dimension $d_\ell < d < d_\mathrm{u}$, as discussed further below. It then corresponds to the approximation of replacing an (anti-)ferromagnet by a paramagnet in a self-generated field σ or, equivalently, dropping the correlations of fluctuations in the Hamiltonian:

$$s_i s_j = -\sigma^2 + \sigma(s_i + s_j) + (s_i - \sigma)(s_j - \sigma) \approx -\sigma^2 + \sigma(s_i + s_j) \; . \tag{176}$$

4.1.3 Hydrodynamics of fluctuations

Some of the neglected fluctuations can be put back into the mean-field model, *a posteriori*. A generic method to do this is hydrodynamics. Assuming that the thermal fluctuations of the order parameter vary only slowly with the spatial coordinate (which may not sound like a great idea for Ising spins in low spatial dimensions but much more so for systems with continuous-range order parameters), one can take them approximately into account by a square-gradient term in the free energy. This results in a popular density functional theory that goes under the names of Landau–Ginzburg or Ornstein–Zernike theory. The (normalized) Landau–Ginzburg free energy L_G is given by a spatial integral over the Landau–Ginzburg density functional $\mathcal{L}_G$,

$$L_G = n_c k_B T_c \int_V \mathrm{d}\mathbf{r}\, \mathcal{L}_G \,, \quad \mathcal{L}_G \equiv \frac{\ell^2}{2}(\nabla\psi)^2 + \frac{t}{2}\,\psi^2 + \frac{g}{4}\,\psi^4 - h\psi \,. \tag{177}$$

Here, the position-dependent magnetization $\sigma(\mathbf{x})$ was replaced by the general notation $\psi(\mathbf{x})$ introduced for the order parameter in the Thermodynamics Part of the lecture. From there, also the notation $t \sim 1 - T_c/T \sim T/T_c - 1 \ll 1$ for the small dimensionless temperature deviation from the critical point, and g for a phenomenological numerical coefficient was adapted. Finally, ℓ denotes a characteristic microscopic length scale (the lattice constant of the Ising model or the size of an atom, etc.).

Many sophisticated theories have grown out of the same spirit over the years to deal with a wide range of phenomena in heterogeneous gases, liquids, and solids. Classical applications of the Ginzburg–Landau theory comprise phase boundaries and nucleation phenomena in phase separation. Both phenomena involve surface energy in a crucial way. Hence, it really pays to explicitly include it in the discussion, although surfaces are of zero measure in the thermodynamic limit, and their contribution to the free energy is ultimately negligible compared to the bulk contributions in an infinite system.

Phase boundaries

As the simplest example for an application of the theory consider an isolated plane phase boundary at $T < T_c$ within Landau–Ginzburg theory. The two phases are characterized by the two possible values $\psi = \psi_{1,2} = \pm\psi_1$ corresponding to the two degenerate free energy minima, respectively. Upon imposing conflicting boundary conditions $\psi(z \to \pm\infty) \to \psi_{1,2}$ the variational minimization of the Ginzburg functional

$$\frac{\delta L_G}{\delta\psi} \equiv \nabla\frac{\partial\mathcal{L}_G}{\partial\nabla\psi} - \frac{\partial\mathcal{L}_G}{\partial\psi} = \ell^2\partial_z^2\psi - t\psi - g\psi^3 = 0 \tag{178}$$

yields the sigmoidal order parameter profile

$$\psi(z) = \psi_1 \tanh(z/\sqrt{2}\xi) \tag{179}$$

perpendicular to the phase boundary (exercises), and the values $\beta = \nu = 1/2$ for the critical exponents characterizing the emergence of the order parameter and the divergence of the characteristic thickness of the phase boundary, respectively:

$$\psi_{1,2} = \pm\sqrt{-t/g} \propto |t|^{\beta} , \quad \xi \equiv \ell\sqrt{-1/t} \propto |t|^{-\nu} . \tag{180}$$

Generally speaking, ξ represents a *correlation length* or persistence length for the fluctuations in the order parameter profile. Its divergence at the critical point indicates that critical phenomena are characterized by strong (long-ranged) spatial correlations, or "infinitely extended phase boundaries".

For phase boundaries to diverge, the energy penalty associated with them has to vanish at the critical point. Within in the hydrodynamic theory, the *surface tension* Σ can be calculated as the surface contribution of the free energy per unit area. One simply has to plug in the solution for $\psi(z)$ from Eq. (179) into L_G and integrate it through the phase boundary:

$$\begin{aligned}\Sigma &= n_c k_B T_c \int_{-\infty}^{\infty} dz \left[\mathcal{L}_G\left(\psi(z)\right) - \mathcal{L}_G\left(\psi_1\right)\right] \\ &= n_c k_B T_c \int_{-\infty}^{\infty} dz \left[\underbrace{\mathcal{L}\left(\psi(z)\right) - \mathcal{L}\left(\psi_1\right)}_{=\ell^2(\partial_z\psi)^2/2} + \frac{\ell^2}{2}\left(\partial_z\psi\right)^2\right] \\ &\simeq n_c k_B T_c \ell^2 \psi_1^2/\xi \propto |t|^{2\beta+\nu} .\end{aligned} \tag{181}$$

(The second step makes use of an identity you discover when solving the Euler–Lagrange equation for the order parameter profile.) Notice that, up to a prefactor, the result could have been anticipated without calculation, from dimensional analysis, namely $\nabla \to \xi^{-1}$ and $\int \mathrm{d}z \to \xi$:

$$\Sigma \simeq n_c k_B T_c \int \mathrm{d}z\, \ell^2 (\nabla\psi)^2 \simeq n_c k_B T_c \xi \ell^2 (\psi_1/\xi)^2 . \tag{182}$$

So, one observes that surface tension does in indeed vanish very dramatically at the critical point, with the relatively large exponent $2\beta + \nu = 3/2$. The effect can thus be attributed to the vanishing of the order parameter (the difference between the two symmetry-broken phases), quantified by β, and to the increasing spatial correlations, quantified by ν, respectively. Experimentally, the consequences can be seen in transparent fluids by the bare eye as *critical opalescence*. The sample turns turbid when light is strongly scattered from the density fluctuations, as soon as their correlation length becomes comparable to its wavelength. This phenomenon already caught the attention of Einstein, who understood this connection quite well. It indicates that order parameter fluctuations not only become long-ranged but also strong at the critical point. In other words, they become large in more than one sense. This should of course be expected, because the susceptibility corresponding to the order parameter (i.e. the compressibility in a fluid, the magnetic susceptibility in a magnet) diverges at the critical point[35].

---

[35] One can indeed show the exponent relation $\nu = \frac{\gamma}{d}\frac{\delta+1}{\delta-1}$ recovering the mean-field estimate for ν at the upper critical dimension $d_u = 4$ and relating ν to the softening encoded in γ, δ.

And it is a clear demonstration that the homogeneity of the order parameter profile predicted by mean-field theory is not trustworthy for common materials in physical dimensions, and may even make one worry about the reliability of the hydrodynamic approach of the Ginzburg theory.

Breakdown of mean-field theory: upper and lower critical dimension

Luckily, a slight variation of the dimensional argument put forward in the previous paragraph provides a simple derivation of a criterion for the reliability of the predictions of the theories by Landau and Ginzburg. The idea is that at the critical point, where the quadratic term $\propto \psi^2$ in the Landau free energy vanishes, the task of confining the order parameter fluctuations is largely left to the gradient term. The relevant integration range and the ∇ are again estimatd by the characteristic correlation length ξ, i.e. $\int d\mathbf{r} \to \xi^d$, $\nabla \to \xi^{-1}$. Invoking equipartition (which states that each quadratic classical degree of freedom carries $k_BT/2$ of thermal energy in its fluctuations) for correlation volumes ξ^d, which correspond to the effective (collective) degrees of freedom, one has

$$n_c k_B T_c \int_V d\mathbf{r}\, \ell^2 (\nabla\psi)^2 \simeq n_c k_B T_c \xi^{d-2} \ell^2 \delta\psi^2 \simeq k_B T_c \tag{183}$$

The overall size of the fluctuations of the order parameter is then found to be

$$\delta\psi^2 \simeq \ell^{-d} n_c^{-1} (\xi/\ell)^{2-d} = \ell^{-d} n_c^{-1} |t|^{(d-2)/2} \,. \tag{184}$$

It is now compared with the absolute value $\psi_{1,2}^2 = |t|/g$ predicted by Landau theory below T_c. Both $\delta\psi$ and $\psi_{1,2}$ are changing as a function of the critical parameter t, and one would expect the Ginzburg theory to be trustworthy in a range of t where the fluctuations $\delta\psi$ do not exceed $\psi_{1,2}$ in magnitude, i.e.

$$\delta\psi^2/\psi_1^2 \simeq \ell^{-d} n_c^{-1} g |t|^{(d-4)/2} \lesssim 1 \quad \Rightarrow \quad |t| \gtrsim (g/\ell^d n_c)^{2/(4-d)} \,. \tag{185}$$

This is the so-called *Ginzburg criterion.* It shows that $d = d_u = 4$ plays the role of an upper critical dimension above which fluctuations vanish for $t \to 0$ and Landau theory should hold. For $d < 4$, it defines a critical region $|t| \lesssim (g/\ell^d n_c)^{2/(4-d)}$ 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 obey the mean-field prediction very well since the critical region is very small due to the large off-critical correlation length $\ell \gg n_c^{-1/3}$ of the strongly delocalized Cooper pairs.

The above estimate of the absolute order parameter fluctuations $\delta\psi$ in Eq. (184) moreover suggests that the order-parameter fluctuations become large in absolute terms, not only relatively to $\psi_{1,2}$, 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, using the Fourier modes

$$\psi_{\mathbf{q}} \equiv \frac{1}{\sqrt{V}} \int_V \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \qquad \psi(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} \psi_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} \, , \tag{186}$$

in the harmonic approximation (exercises)

$$L_G = n_c k_B T_c \frac{1}{2} \sum_{\mathbf{q}} (t + q^2 \ell^2) |\psi_{\mathbf{q}}|^2 \, . \tag{187}$$

For simplicity, $T \gtrsim T_c$ ($t > 0$) and $h = 0$ shall be required. Now the equipartition theorem is again employed; this time to the eigenmodes $\psi_{\mathbf{q}}$, which looks less heuristic than the above application to effective correlation volumes ξ^d, but essentially reproduces the same result. The average strength of the order parameter fluctuations is then seen to be given by

$$n_c k_B T_c (t + q^2 \ell^2) \langle |\psi_{\mathbf{q}}|^2 \rangle = k_B T_c \, . \tag{188}$$

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

$$n_c \ell^2 \langle |\psi_{\mathbf{q}}|^2 \rangle = \frac{1}{\xi^{-2} + q^2} \, . \tag{189}$$

Here the correlation length ξ of the order parameter fluctuations shows up again. It controls the spatial decay of correlations, as becomes more evident from the Fourier back transform, the order parameter *correlation function* in real space,

$$\langle \psi(\mathbf{r}) \psi(0) \rangle = \mathrm{FT}^{-1} \frac{\langle |\psi_{\mathbf{q}}|^2 \rangle}{\sqrt{V}} \propto \begin{cases} \dfrac{e^{-r/\xi}}{r^{(d-1)/2}} & (r \to \infty) \\ 1/r^{(d-2)} & (\xi \to \infty) \, . \end{cases} \tag{190}$$

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.

From the last expression in Eq. (190), one can again read of the lower critical space dimension $d_l = 2$. 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[36]. It is not an artifact of the theory but due to a real physical effect. Due to the reduced number of constraining neighbors, long-wavelength hydrodynamic excitations (also called Goldstone

---

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

modes) become all pervasive in and below the lower critical dimension $d_\mathrm{l} = 2$ at any $T > 0$, which prevents the phase ordering phenomena observed in higher dimensions (although not phase separation as such). In particular, it has been proved in great generality that thermal fluctuations of a continuous order parameter prevent long-range order in $d \leq 2$ if the interactions are not long-ranged ("Mermin–Wagner theorem"). In other words, it says that there are no proper crystals in one or two dimensions, thereby hinting at the fact that the physics of membranes and polymers is dominated by fluctuations. Phase transitions exactly at $d = d_\mathrm{l}$ (first 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. (190). Note that it makes a prediction for the behavior inside the critical region, where one should actually not trust the prediction of the Ginzburg theory. As precise measurements have indeed revealed, there are small systematic deviations from the Ginzburg form in $d < 4$ space dimensions. They have (first phenomenologically) been rationalized by introducing a so-called "anomalous dimension" η,

$$\langle |\psi_\mathbf{q}|^2 \rangle \propto q^{-2+\eta} \,, \qquad \langle \psi(\mathbf{r})\psi(0) \rangle \propto 1/r^{d-2+\eta} \,, \qquad \xi \to \infty \,. \tag{191}$$

The name derives from the perplexing observation that this relation apparently violates ordinary dimensional analysis. It implies that there must be a subtle hidden dependence on the microscopic length scale ℓ, even — or rather in particular — at the critical point, where $\ell/\xi \to 0$. As a side remark, Eq. (191) implies the exponent relation $\gamma = \nu(2 - \eta)$. It follows by analogy with the compressibility equation Eq. (164) relating the order-parameter fluctuations to the corresponding response coefficient of a gas. For the magnet, we simply have to replace particle number fluctuations $\langle N^2 \rangle$ by spin number (or magnetization) fluctuations $\langle \psi^2 \rangle$ and the isothermal compressibility κ_T by the magnetic susceptibility χ_T.

4.1.4 Defects and low-temperature expansion

Defects as localized fluctuations

The hydrodynamic theory of fluctuations suggests $d_\ell = 2$ for the lower critical dimension, as is indeed widely observed. But the Ising model is a bit exceptional because of its discrete state space. In contrast to ordinary classical variables the binary spin variables do not have a continuous range and thus cannot change smoothly in space and time. As a result, the low temperature excitations from the ground state are not hydrodynamic waves (such as spin waves, sound waves, etc., collectively known as Goldstone modes) as they would occur in systems with a continuously varying order parameter amenable to a hydrodynamic treatment. Instead, they are local defects of microscopic width. Consider for simplicity first the 1-dimensional Ising spin chain consisting of N spins (or bonds). In the ground state all spins are aligned. But what does temperature do to the chain?

It provides thermal energy to topple some spins. A "broken bond" between a pair of misaligned spins is called a *defect*, because it is regarded as an annoying perturbation of the uniform ground state. Its creation costs an energy $2J$. Since an individual bond between spins can either be satisfied or broken, corresponding to the Boltzmann factors $e^{\beta J}$ and $e^{-\beta J}$, and the partition sum $Z = e^{\beta J} + e^{-\beta J}$, the fraction of broken bonds in equilibrium is simply[37] $(1 + e^{2\beta J})^{-1}$.

Another way of looking at this is the following. Single defects can wander about along the chain at no cost like gas particles. At low temperatures, $\beta J \gg 1$, only a small number N_b of all N bonds is broken, corresponding to a defect fraction $\phi \equiv N_\mathrm{b}/N \ll 1$. Hence the defects represent a very dilute gas and the probability ϕ^2 that two defects meet is very small, so that they can be treated as an ideal gas. They then supply the entropy $S \simeq k_B N_\mathrm{b} \ln(T^{1/2}/\phi)$, and the defect free energy is

$$F = U - TS = N_\mathrm{b} \left[2J - k_B T \ln(T^{1/2}/\phi)\right] \ . \tag{192}$$

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 down), the chain would spontaneously split up into up- and down-domains by a spontaneous creation of defects. 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. Note that the limits $T \to 0$ 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), indicative of a non-analytic structure at $T = 0$, $N = \infty$, which is studied in detail, further below.

Low-temperature expansion and Peierls' argument

In one dimension, point defects thus proliferate at any finite temperature. What about higher dimensions, where defect lines and surfaces etc. take over the role of the point defects in the Ising chain? When does the mean-field prediction of a spontaneous magnetization apply? Defect hypersurfaces of dimension $d - 1$ are phase boundaries, since they separate regions of spins with opposite orientation. They have the tendency to freeze out at low temperatures, and this tendency increases with the number of intolerant neighbors, which in turn increases with the space dimension. Upon freezing, they will contract the enclosed volume of "wrong" spins such that, below a critical temperature $T_c > 0$, spins of one orientation will dominate over those with the other, giving rise to a spontaneous magnetization. For a two-dimensional Ising lattice this effect is easily demon-

---

[37] This assumes open boundary conditions. In the case of periodic boundary conditions, defects can only be created pairwise, hence their fraction is $2(1 + e^{2\beta J})^{-2}$.

strated, as first shown by R. Peierls in 1935, which proves that there is a non-zero transition temperature in the Ising model with nearest-neighbor interactions in two dimensions. This establishes $d_\ell^{\rm Ising} = 1$ as the lower critical dimension, in contrast to what is found for models with continuous order parameters.

The starting point of Peierls' discussion is again the observations that for each pair of unsatisfied neighboring spins an energy $2J$ has to be paid. At low temperatures, few such excitations are present, and it makes sense to order the terms in the partition sum according to the excitation energy. For the above example of the one dimensional spin chain with N bonds:

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

The leading term is the Boltzmann factor of the twofold degenerate ground state, the other terms correspond to one defect, two defects, etc., which can occur anywhere along the chain, i.e. at m out of N places. This form of the partition sum as sum over configurations with increasing numbers of defects is particularly useful at low temperatures, where accurate approximations are obtained by truncating the sum after a finite number of terms.

The analogous expression in two dimensions reads

$$Z^{(2{\rm d})} = 2e^{\beta JdN}\sum_{N_b} g_{N_b}e^{-2\beta JN_b}\ . \tag{193}$$

The leading term outside the sum again corresponds to the ground-state free energy, which is attained if all dN bonds are satisfied, and there remains but the binary up/down degeneracy of the total magnetization for the entropy. The partition sum for the defects runs over all g_{N_b} droplets with a circumference corresponding to N_b broken bonds. The entropic contributions to the droplet free energy are contained in the degeneracy factors g_{N_b}, which is of course difficult to calculate in all generality. But for the task of demonstrating the existence of a finite transition temperature only an upper bound for the average number of "wrong" spins inside the droplets, say the $\downarrow$-spins, is needed. So finding an upper bound for g_{N_b} suffices, and one may simply say that there are $2d = 4$ directions the defect line can chose to go to from an arbitrarily chosen broken bond to the next (neglecting the possibility that at least one of these other bonds might already be broken). Lines of broken bonds ending on the boundaries of the square lattice are relatively rare of order $1/\sqrt{N}$ relative to closed loops that may start (and end) on any of the N bonds in the lattice, so they are dismissed. Finally, there clearly cannot be boundaries of length shorter than $2d$, but including them in the sum should not cause much trouble, nor does extending the sums to infinity. Hence, one sloppily bounds the average number $N_\downarrow$ of spins in droplets with closed boundaries by

$$N_\downarrow \leq \frac{\langle N_b^2 \rangle}{4^2} \lesssim \frac{\sum_{N_b=0}^{\infty} N_b^2 N (4e^{-2\beta J})^{N_b}}{4^2 \sum_{N_b=0}^{\infty} N (4e^{-2\beta J})^{N_b}} = \frac{(4 + e^{2\beta J})}{4(e^{2\beta J} - 4)^2} \tag{194}$$

in $d = 2$ space dimensions. This function is quickly decreasing with βJ for not too small arguments, so that the number $N_\downarrow$ of spins with the wrong orientation drops below $N/2$ at $k_B T \approx J$, where a finite spontaneous magnetization must therefore develop[38].

Remark: Similarly to the low temperature expansion, one can write down a high temperature expansion, which has a similar diagrammatic interpretation, and which turns out to be "dual" (closely related) to the low temperature expansion by virtue of a self-duality symmetry of the two-dimensional Ising model.

4.1.5 Thermostatics of the Ising chain

Transfer matrix

The one dimensional Ising chain with nearest neighbor interactions has been discussed in the exercises for open boundaries. Here, it shall be reconsidered with periodic boundary conditions to introduce the popular method of the transfer matrix. Note that the partition sum

$$Z = \sum_{\{s_i = \pm 1\}} e^{-\beta H} = \sum_{\{s_i = \pm 1\}} e^{\beta \sum_i \left[J s_i s_{i+1} + h s_i \right]} \ . \tag{195}$$

decays into a product of structurally identical terms

$$Z = \sum_{\{s_i = \pm 1\}} \prod_i e^{\beta J s_i s_{i+1} + \beta h (s_i + s_{i+1})/2} \ . \tag{196}$$

This calls for introducing an operator $\hat{T}$ with matrix elements

$$\langle s_i | \hat{T} | s_{i+1} \rangle = \exp[\beta J s_i s_{i+1} + \beta h (s_i + s_{i+1})/2] \tag{197}$$

and matrix representation

$$\begin{pmatrix} \exp[\beta(J+h)] & \exp[-\beta J] \\ \exp[-\beta J] & \exp[\beta(J-h)] \end{pmatrix} \qquad (\text{"transfer matrix"}) \tag{198}$$

in the 2-dimensional spin space with basis $\{|+\rangle, |-\rangle\}$ for $s_i = \pm 1$. With $\hat{T}$ one can write

$$Z = \sum_{\{s_i = \pm 1\}} \langle s_1 | \hat{T} | s_2 \rangle \langle s_2 | \hat{T} | s_3 \rangle \ldots \langle s_N | \hat{T} | s_1 \rangle \ . \tag{199}$$

---

[38] Onsager's exact solution says $k_B T_c \approx 2.27 J$.

Using the completeness of the spin basis, this reduces to

$$Z = \sum_{s_1=\pm 1} \langle s_1|\hat{T}^N|s_1\rangle = \mathrm{tr}\hat{T}^N = \lambda_+^N + \lambda_-^N \,, \tag{200}$$

where λ_+ and λ_- denote the transfer matrix eigenvalues

$$e^{\beta J}\big[\cosh\beta h \pm (e^{-4\beta J} + \sinh^2\beta h)^{1/2}\big] \sim \begin{cases} e^{\beta J} \pm e^{-\beta J} + \frac{e^{\beta J} \pm e^{3\beta J}}{2}(\beta h)^2 & (\beta h \to 0) \\ 2\cosh\beta h \,,\, 0 & (\beta J = 0) \\ e^{\beta J \pm \beta |h|} & (\beta J \to \infty) \end{cases} . \tag{201}$$

Emergence of singularities for $T \to 0$, $N \to \infty$

Observe that for finite βJ and βh there is a non-degenerate largest eigenvalue $\lambda_+ > \lambda_-$, which is an analytical function of βJ and βh. (This is the Frobenius–Perron theorem for matrices with finite positive matrix elements). So for finite matrix entries (finite βJ and βh), the larger eigenvalue λ_+ will dominate the partition sum and the free energy per spin

$$\lim_{N\to\infty} \beta F/N \sim -\ln\lambda_+ = -\beta J - \ln[\cosh\beta h + (e^{-4\beta J} + \sinh^2\beta h)^{1/2}] \tag{202}$$

in the thermodynamic limit ($N \to \infty$). Asymptotically, this simplifies to

$$\begin{aligned} -\ln[2\cosh\beta J] - (\beta h e^{\beta J})^2/2 \qquad & (\beta h \to 0) \\ -\ln[2\cosh\beta h] \qquad & (\beta J = 0) \\ -\beta J - |\beta h| \qquad & (\beta J \to \infty) \end{aligned} \tag{203}$$

Also note that, in the high-temperature limit $T \to \infty$, the free energy F reduces to $-TS$ with the entropy $S = k_B N \ln 2$.

All thermodynamic quantities of interest can now be obtained as derivatives of the free energy density $f = F/N$. For example, the specific heat per particle $c_h = -T\partial_T^2 f$ is found as

$$\frac{c_h}{k_B} = \frac{(\beta J)^2}{\cosh^2\beta J} \qquad (\beta h \to 0)\,, \qquad \frac{c_h}{k_B} = \frac{(\beta h)^2}{\cosh^2\beta h} \qquad (\beta J \to 0)\,. \tag{204}$$

The result indicates that the thermodynamics of spins is somewhat untypical and counter-intuitive, due to their odd state space. Namely, instead of saturating, the specific heat vanishes for high temperatures due to the limited energy that a spin degree of freedom (in contrast to particles in ordinary fluids or solids, say) can take up. The dramatic non-analytic decrease at low temperatures is characteristic of the *energy gap* $2J$ that has to be overcome to achieve any excitation at all, namely to introduce a single defect and thereby put some heat into the system.

In practice, such behavior is therefore characteristic of semiconductors and superconductors — rather than common magnets. It is called a Schottky-anomaly to emphasize that it is rather unconventional compared to a dense (gap-less) spectrum of energy levels that can gradually be populated if heat is added.

Similarly, taking derivatives of f with respect to βh, one arrives at the site magnetization

$$\partial_{\beta h}\beta f = \langle\sigma\rangle = \frac{\sinh\beta h}{\sqrt{\sinh^2\beta h + e^{-4\beta J}}} \sim \begin{cases} \beta h e^{2\beta J} & (\beta h \to 0) \\ \tanh\beta h & (\beta J \to 0) \\ \mathrm{sgn}(\beta h) & (\beta J \to \infty) \end{cases} \tag{205}$$

and the susceptibility

$$\chi = \partial_h \langle\sigma\rangle|_{h=0} = \beta e^{2\beta J} \ . \tag{206}$$

This reduces to the well-known Curie law $\chi \simeq \beta$ for a paramagnet in the limit $\beta J \to 0$. Because of the analytic parameter dependence, there cannot be a phase transition for finite parameters. However, in the limit $\beta J \to \infty$, Eq. (205) produces a discontinuous magnetization as characteristic of a ferromagnet. In this limit, the gap between the eigenvalues diminishes if $h = 0$, so that a singularity may appear. For this reason, it is advisable to try and see what happens if the order of the limits is reversed and the contribution from the second eigenvalue λ_- is kept, if one is interested in the low temperature or strong interaction behavior $\beta J \to \infty$ of the model. (Recall that for any finite $J \neq 0$ any finite chain will order at $T = 0$ while an infinite chain will always be disordered at any $T > 0$, which hints at the nature of the singularity at $T = 0$.)

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. (200) & (201) give for small βh

$$Z = e^{N\beta J}\left[(1 + |\beta h|)^N + (1 - |\beta h|)^N\right] = e^{N\beta J} 2\cosh(N\beta h) \ , \tag{207}$$

hence

$$f \equiv \frac{F}{N} = -J - \frac{1}{\beta N}\ln[2\cosh(N\beta h)] \overset{N\to\infty}{\sim} -J - |h| \ . \tag{208}$$

While the ultimate result in the double limit βJ, $N \to \infty$ is the same as in Eq. (202) above, 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

$$\langle\sigma\rangle = \tanh N\beta h \sim \mathrm{sgn}(\beta h) \, , \tag{209}$$

and in the susceptibility

$$\chi = \frac{N\beta}{\cosh^2 N\beta h} \sim 2\beta\delta(\beta h) \ . \tag{210}$$

They both clearly display how the singularity develops 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 $\chi \sim \beta N$, 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 consequence of the formation of a giant spin can also be seen in the behavior of the specific heat, if calculated for finite N and diverging βJ. Starting from Eq. (208) instead of Eq. (202), which corresponds to the reverse order of the limits $N \to \infty$ and $\beta J \to \infty$, reveals a tiny specific heat

$$\frac{c_h}{k_B} = \frac{N(\beta h)^2}{\cosh^2 N\beta h} \sim \frac{p}{N}[\delta^+_{\beta h} + \delta^-_{\beta h}]\ . \tag{211}$$

The symbolic notation $\delta^\pm$ refers to a Kronecker$-\delta$ centered around slightly positive or negative arguments, respectively, and $p \approx 0.44$. 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. (208) 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, so that the specific heat vanishes in the thermodynamic limit. Yet, the limit in Eq. (211) 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 switched off ($h \to 0$) 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.

4.1.6 Critical scaling and Yang–Lee zeros

Recovering the scaling hypothesis and finite-size scaling

The non-analytic strong coupling behavior revealed by the above discussion is indicative of a phase transition at $T = 0$. It is worthwhile checking whether the general framework of the scaling hypothesis introduced in Part I is still applicable here. Recall that the singular part of the free energy density $\mathfrak{f}_s$ was supposed to diverge as $t^{1/a_t} = t^{2-\alpha}$ in absence of an external field ($h = 0$), where by convention $t = |T/T_c - 1|$ (or better $|1 - T_c/T|$) is used to measure the dimensionless distance from the critical point. This was the scaling hypothesis, which provided a neat phenomenological explanation of the observed non-trivial divergencies at the critical point $T = T_c$. The existence of a single characteristic diverging length scale $\xi \propto t^{-\nu}$ near the critical point together with dimensional analysis moreover suggests that $\mathfrak{f}_s$, as a spatial density, should obey $\mathfrak{f}_s \propto \xi^{-d} \propto t^{d\nu}$, together with the above $\mathfrak{f}_s \propto t^{2-\alpha}$, hence the exponent relation

$$d\nu = 2 - \alpha \tag{212}$$

known as the *hyperscaling relation*. Moreover, for transfer matrices, one can show that the dimensionless correlation length $\hat{\xi}$ (in units of the lattice constant) generally obeys

$$\hat{\xi}^{-1} = -\lim_{r\to\infty} \frac{1}{r} \ln\langle s_i s_{i+r}\rangle = \ln(\lambda_1/\lambda_2) \ , \tag{213}$$

showing that its divergence at the critical point is related to the decreasing distance between the eigenvalues. In the exercises it was indeed shown by direct calculations that the spin correlations in the open Ising chain with nearest neighbor interactions decay exponentially

$$h(n-m) \equiv \langle s_n s_m\rangle = [\tanh\beta J]^{|n-m|} = e^{-|n-m|/\hat{\xi}} \ . \tag{214}$$

with the *correlation length*

$$\hat{\xi}^{-1} = -\ln\tanh\beta J \sim 2e^{-2\beta J} \qquad (\beta J \to \infty) \ . \tag{215}$$

Physically it is clear, that for $|n-m| \ll N$ this will also hold in a closed chain. To match this prediction to the generic power-law singularities discussed in Part I of the lecture, one apparently should *identify* $e^{-\beta J}$ *with the critical parameter* t, so that $\mathrm{d}\nu = \nu = 2$ and $\alpha = 0$ as required by the absence of a divergence in the specific heat. The unusual critical parameter is obviously a consequence of the Schottky-anomaly (or of the system being at its lower critical dimension). From Eq. (206) $\gamma = 2$ is read off, which together with the exponent relation $2-\alpha = 2\beta + \gamma$ implies $\beta = 0$, in accord with the absence of a spontaneous magnetization for any finite coupling strength. Finally, the anomalous dimension $\eta = 1$ follows from the exponent relation deduced from Eq. (191), namely $\gamma = \nu(2-\eta)$. One may also check the scaling ansatz for the singular part of the free energy density, directly. Keeping only the leading order terms in the critical parameters $e^{-\beta J}$ and βh for large βJ, one finds from Eq. (202)

$$\beta f \sim -\beta J - e^{-2\beta J}\sqrt{1 + (\beta h e^{2\beta J})^2} \ . \tag{216}$$

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

$$\beta f_s = \beta f + \beta J \sim -\hat{\xi}^{-1}\varphi(\beta h\hat{\xi}) = -t^{\nu}\varphi(\beta h/t^{\nu}) = -t^{2-\alpha}\varphi(\beta h/t^{\gamma+\beta}) \tag{217}$$

where the last expression corresponds to the general scaling hypothesis introduced in Part I. Similarly, for large (but finite) N, βJ and $\beta h = 0$

$$\begin{aligned}
Z =&(e^{\beta J}+e^{-\beta J})^N + (e^{\beta J}-e^{-\beta J})^N = e^{N\beta J}[(1+e^{-2\beta J})^N + (1-e^{-2\beta J})^N] \\
\sim& 2e^{N\beta J}\cosh(Ne^{-2\beta J}) \qquad \Rightarrow \\
\beta f_s \sim& -\frac{1}{N}\ln[(2\cosh(Ne^{-2\beta J})]) \\
\sim& -e^{-2\beta J}\big[Ne^{-2\beta J} - N^3 e^{-6\beta J}/6 + \dots\big]/2 \\
\sim& -\hat{\xi}^{-1}\phi(N/\hat{\xi}) \qquad\qquad (\beta J \to \infty\,,\ N \to \infty)
\end{aligned} \tag{218}$$

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. The *finite size scaling* relation, Eq. (218), clarifies 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.

Yang-Lee zeros

The above calculations have shown explicitly how thermodynamic quantities develop non-analytic behavior as a function of the coupling strength, the external field, and the particle number. Another way of looking at what is going on here focuses on the zeros of the partition sum in the complex field or temperature plane for large βJ and N. Starting from the partition sum for finite N and $\beta J \to \infty$, corresponding to the free energy in Eq. (208), one has to find the complex solutions of

$$Z = 2e^{N\beta J}\cosh N\beta h = 0 \quad \Rightarrow \quad \beta h = \pm i\pi(n-1/2)/N \quad n \in \mathbb{N}\,. \tag{219}$$

While these lie on the imaginary h-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 again $(1+x/N)^N \sim e^x$, one finds the zeros in the complex "temperature" plane

$$Z = 2e^{N\beta J}\cosh[Ne^{-2\beta J}] = 0 \quad \Rightarrow \quad t^2 \equiv e^{-2\beta J} = \pm i\pi(n-1/2)/N \quad n \in \mathbb{N}\,. \tag{220}$$

Again, there is an accumulation of zeros that cut the complex plane into pieces at the origin for $N \to \infty$. They approach on the imaginary axis of the complex t^2-plance, or along the diagonals in the complex t-plane, respectively. Somewhat more impressive sets of Yang-Lee zeros can e.g. be found in the book "The Beauty of Fractals" by H.-O. Peitgen and P. H. Richter.

4.2 Ideal quantum gases

Ideal quantum gases are of similar complexity as interacting classical gases. This is so because of effective exchange interactions, which are a consequence of identifying states that differ only in the exchange of indistinguishable particles. This condition is usually implemented by imposing appropriate symmetries on the wave function (rather than by formulating the wave function in a projective configuration space). Identical bosons, with symmetric many-body wave functions, then effectively attract each other. And identical fermions, with asymmetric many-body wave functions, effectively repel each other. While an approximate description of an ideal quantum gas as an interacting classical gas is therefore possible, as discussed further below, an exact quantum mechanical description is more appropriate and not more complicated. It turns out to be surprisingly helpful in the understanding of a broad range of phenomena. In fact, it is the passkey to a quantum theory of the (sometimes not so) low-temperature thermodynamics of condensed matter. The reason is that the *elementary collective excitations* responsible for all types of transport in cold solids and liquids can after quantization be interpreted as weakly interacting *quasi-particles* of either bosonic or fermionic nature[39], at sufficiently low temperatures. Familiar examples comprise electronic excitations (so-called "electrons" and "holes") in metals and semiconductors, sound and spin waves (phonons, magnons) and hybrid excitations made e.g. from light and matter (polarons, polaritons, etc) or from other quasiparticles (Cooper pairs). The low temperature physics of all matter can therefore be understood on the basis of the following discussion of ideal quantum gases. Here "low temperature" can mean several ten-thousand Kelvin for some systems (conduction electrons in normal metals) but the lowest temperatures in the universe for others (ordinary gases). The remoteness of the Planck scale makes it quite plausible[40] that even the so-called elementary particles of the standard model are but another example of such emergent low-temperature elementary excitations that hardly give a clue at whatever theory may describe the constituents of the underlying "vacuum". In any case, the densities and temperatures where the classical gas description breaks down are defined by the condition $n\lambda_T^3 \simeq 1$.

---

[39] This statement holds in a three-dimensional world. Due to the different topology (J.M. Leinaas, J. Myrheim, *Il Nuovo Cimento* 37 B (1977), p.1), *anyons*, which defy this classification, should prevail in a two dimensional space (F. Wilczek, *Physics World* 19/1 2006, p.22).

[40] Grigory E. Volovik: The universe in a helium droplet; Oxford 2003.

4.2.1 Kinematics and statistics

Kinematics

The basis for the quantum mechanical description of ideal gases is the free Hamiltonian

$$\hat{H} = \sum_i \frac{\hat{\mathbf{p}}_i^2}{2m} \tag{221}$$

which is a sum over non-interacting particles. Each sector of the Hamiltonian can thus be treated individually and is diagonalized by the eigenfunctions

$$\langle \mathbf{r} | \mathbf{p} \rangle = e^{i\mathbf{p}\cdot\mathbf{r}/\hbar} \,. \tag{222}$$

Usually one imagines the particles to sit in a cubic box of volume V (to be taken to infinity, in the end) — in which case the eigenfunctions will be sines instead of exponentials — so that

$$\varepsilon_{\mathbf{k}} = \frac{\mathbf{p}^2}{2m} \qquad \mathbf{p} = \frac{\hbar\pi}{V^{1/d}}(k_1, \ldots, k_d) \qquad k_i \in \mathbb{N} \,. \tag{223}$$

The *one-particle density of states* (or level density)

$$g(\varepsilon) \equiv 2\pi (2m)^{d/2} h^{-d} \varepsilon^{d/2-1} \tag{224}$$

will always enter the discussion if a sum over eigenstates is turned into an energy integral,

$$\sum_{\text{states}} = \sum_{\{\mathbf{p}\}} = V \sum_{\{\mathbf{p}\}} \left[\frac{\Delta p}{\hbar\pi} \right]^d \to V \int_{-\infty}^{\infty} \frac{\mathrm{d}^d p}{h^d} = V \int_{\varepsilon_0}^{\infty} \mathrm{d}\varepsilon \, g(\varepsilon) \,, \tag{225}$$

Note that the sum over $k_i \in \mathbb{N}$ turns into an integral over the positive momentum sector — $1/2^d$ of the full momentum space — and also that the volume V is not included in $g(\varepsilon)$. This slight variation of the convention used so far will be more convenient when dealing with densities, as opposed to extensive values, below. To include slightly more general cases involving quasi-particles such as phonons, photons, magnons, etc. which may have different dispersion relations, the density of states is in the following assumed to be of the form

$$g(\varepsilon) = \mathfrak{g} \, g_0 \, \varepsilon^{d/\tau - 1} \,. \tag{226}$$

The factor g_0 comprises some numerical constants, $\mathfrak{g}$ counts the spin degeneracy (or polarizations for photons), and the effective mass m or velocity c characterize, together with the exponent τ, the dispersion relation[41]

$$\mathfrak{g} = \begin{cases} 2 & \text{(photons)} \\ d & \text{(phonons)} \\ 2 & \text{(electrons)} \end{cases} \qquad g_0 = \begin{cases} \dfrac{2\beta^{d/2}}{\lambda_T^d \Gamma(d/2)} \\ \dfrac{4\pi}{(hc)^d} \end{cases} \qquad \tau = \begin{cases} 2 & \text{(massive)} \\ 1 & \text{(massless)} \end{cases} \,. \tag{227}$$

---

[41] Here, $\Gamma(d/2)$ ($= 1$ in $d = 2$ and $\sqrt{\pi}/2$ in $d = 3$ dimensions) denotes the gamma function.

It is often o.k. to think of massive modes as "particles" and of massless modes as "waves". Note that everything up to now is *pure kinematics*, independent of statistics, and will stay fixed for the rest of the discussion.

Statistics

While the kinematics of noninteracting gas particles in a box is invariably fixed by the above expressions that only depend on the properties of a single particle, the "exchange behavior" determines how the single-particle energy spectrum is populated, if a given volume is occupied by many gas particles. The cases

$$\delta \to \begin{cases} -1 & \text{Bose statistics} \\ 1 & \text{Fermi statistics} \\ 0 & \text{Maxwell–Boltzmann statistics} \end{cases} \tag{228}$$

can efficiently be treated in parallel. For N indistinguishable particles, the projective nature of their phase space may be taken care of on the level of the wave function. If it is written as a *totally (anti-)symmetrized product wave function*

$$|\{\mathbf{p}_i\}\rangle \equiv \frac{1}{\sqrt{N!}} \sum_{\mathcal{P}} (-\delta)^P |\mathbf{p}_{\mathcal{P}1}\rangle|\mathbf{p}_{\mathcal{P}2}\rangle \dots |\mathbf{p}_{\mathcal{P}N}\rangle \tag{229}$$

constructed from N ortho-normalized one-particle wave functions $|\mathbf{p}\rangle$, one can think of it as a function living on the naive (non-projective) phase space. The short-hand notation for permutations of the single particle wave functions employs the permutation operator $\mathcal{P}$ and the length of the permutation P (i.e. the number of pairs exchanged by $\mathcal{P}$). Note that the total anti-symmetrical construction for fermions, called a "slater determinant", suppresses multiple occupancies of the same state. This is the *Pauli principle.* In contrast, occupancies of arbitrary multiplicity are allowed for bosons[42].

For the purpose of calculating partition sums, all that Eq. (229) adds to the above discussion of the single particle spectrum, is the constraint that the many-body states have to conform with (anti-)symmetrization. This means that two or more identical fermions cannot occupy the same state (spin degeneracy is counted separately in $\mathfrak{g}$), while up to N identical bosons can occupy the same state. One can take advantage of this situation by a clever way of counting states, which consists in counting occupation numbers within the preset one-particle energy spectrum. Formally, one may say that one can uniquely represent the actual $N-$particle states in an *occupation number representation*

$$|\{n_{\mathbf{k}}\}\rangle \tag{230}$$

---

[42] Possible worries about the proper normalization in the case of multiple occupancies are best vetted later, when taking the trace.

with $n_{\mathbf{k}}$ the occupancy of the one-particle eigenstates from Eq. (223). As a consequence of this smart way of counting degrees of the freedom, the N-body density of states, which was the central object of the classical treatment in Sec. 3.2, now never enters the discussion. All one ever needs is the one-particle level density from Eq. (226).

If the total number of particles N is fixed, i.e. for true non-relativistic particles that cannot arbitrarily be created and destroyed, the convenience of the number representation is somewhat spoiled by the auxiliary condition

$$N = \sum_{\mathbf{k}} n_{\mathbf{k}} \ . \tag{231}$$

This is a good reason to turn to the grand-canonical ensemble[43], in which a chemical potential will take care of fixing N on average without one having to bother about Eq. (231). The grand canonical partition sum reads

$$Z_G = \sum_{N=0}^{\infty} \mathrm{tr}_N \exp[-\beta(\hat{H} - \mu\hat{N})] \ . \tag{232}$$

The index N indicates that the canonical trace has to respect Eq. (231). But, luckily, the additional summation over all N resolves the inherent difficulty: summing traces respecting Eq. (231) over all N is the same as taking the unrestricted trace over all possible states of arbitrary N, right away. The index N at the trace and the N-summation thus effectively cancel each other, and one is left with the much simpler task of computing the infinite sum

$$Z_G = \sum_{\{n_{\mathbf{k}}\}} \exp\Big[-\beta \sum_{\mathbf{k}} n_{\mathbf{k}}(\varepsilon_{\mathbf{k}} - \mu)\Big] = \prod_{\mathbf{k}} \sum_{n_{\mathbf{k}}} \exp[-\beta n_{\mathbf{k}}(\varepsilon_{\mathbf{k}} - \mu)] \ . \tag{233}$$

Taking into account the different occupancies allowed for fermions and bosons, this yields the neat result

$$Z_G = \prod_{\mathbf{k}} [1 + \delta e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}]^{1/\delta} \ . \tag{234}$$

For fermions (relegating the issue of spin degeneracies to the density of states) the sum has only two terms, $n_{\mathbf{k}} = 0, 1$, because of the Pauli principle. For bosons the sum is infinite, $n_{\mathbf{k}} = 0, 1, 2, 3, \dots$, and one uses the Euler formula for geometric series. Its convergence is only guaranteed for $\beta(\varepsilon_{\mathbf{k}} - \mu) > 0$, though. Given that the ground state energy in the box is $\varepsilon_0 \simeq \mathcal{O}(V^{-\tau/d}) \to 0$, the chemical potential is therefore required to be negative for bosons, as for classical gases, for which

---

[43] One should of course check *a posteriori* that the results do indeed apply to experimental situations with constant particle number. For the thermodynamic averages computed below, this does not cause problems. Occupancies of individual states are a more subtle issue.

$\mu \simeq -s < 0$, up to irrelevant terms. The marginal case $\mu \to 0$ is analyzed further below. The classical limiting case (Boltzmann statistics) is obtained for $\delta \to 0$, where the characteristic exponential scaling with the particle number, Eq. (136), is recovered by the effective exponentiation: $(1 + \delta x)^{1/\delta} \sim e^x$.

According to Eq. (148) the logarithm of the grand canonical partition sum is the grand canonical potential, and its density is the pressure,

$$pV = -J(T, V, \mu) = k_B T \ln Z_G = \frac{k_B T}{\delta} \sum_{\mathbf{k}} \ln[1 + \delta e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}] \tag{235}$$

This represents the thermodynamic *fundamental relation of an ideal quantum gas*, i.e. it contains the complete thermodynamic information. Yet, some work remains to be done to squeeze out the relevant physics from it. Some easy-to-derive major implications are discussed in the following.

4.2.2 Thermodynamics of ideal quantum gases

Practically speaking, for systems with a prescribed number of particles N, the chemical potential μ, on which the thermodynamic properties apparently depend quite sensitively, is *a priori* not known. It has to be inferred from Eq. (231) to make Eq. (235) useful. In other words,

$$\partial_{\beta\mu} \ln Z_G = \langle N \rangle = \sum_{\mathbf{k}} \langle n_{\mathbf{k}} \rangle \;, \tag{236}$$

must be solved for $\mu(\langle N \rangle)$, which can only be achieved perturbatively. In the following, the strategy is to avoid this laborious task as long as possible, since a sufficient amount of interesting results can already be obtained without it.

The crucial quantity that appears in Eq. (236) and in many other tasks is the *number distribution* (sketch!)

$$n(\varepsilon_{\mathbf{k}}) \equiv \langle n_{\mathbf{k}} \rangle = \partial_{\varepsilon_{\mathbf{k}}} J = \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} + \delta} \;. \tag{237}$$

It generalizes the classical Boltzmann factor, to which it reduces for $\delta = 0$. The thermodynamic properties of ideal quantum gases are obtained from this central quantity (and its derivatives). The Fermi- or Bose distribution $n(\varepsilon)$, as it is called for $\delta = \pm 1$, respectively, contains the *essential statistical information* as opposed to the kinematic information encoded in the density of states $g(\varepsilon)$. The product $w(\varepsilon) = g(\varepsilon) n(\varepsilon)$ is the energy distribution, which determines the relative amount of particles with a given energy ε at a given temperature and chemical potential.

The task of computing the pressure from Eq. (235) (as most others) leads one directly to the number distribution and to the energy distribution. In the integral approximation, Eq. (225), which is admissible as long as the individual level contributions to the total sum are negligibly small,

$$p = \frac{k_B T}{\delta} \int_0^\infty \mathrm{d}\varepsilon \, g(\varepsilon) \ln[1 + \delta e^{-\beta(\varepsilon - \mu)}] \;. \tag{238}$$

In the thermodynamic limit the ground state energy vanishes, so that the lower bound of the integral can be set to zero, where the integrand vanishes (at least in three space dimensions). Using Eq. (226), i.e., $g(\varepsilon) \propto \varepsilon^{d/\tau-1}$ so that $\int \mathrm{d}\varepsilon\, g(\varepsilon) = \varepsilon\tau g(\varepsilon)/d$, an integration by parts relates the pressure

$$p = \frac{\tau}{d}\int_0^\infty \mathrm{d}\varepsilon\, g(\varepsilon) n(\varepsilon)\varepsilon = \frac{\tau\mathfrak{u}}{d} = \begin{cases} 2\mathfrak{u}/d & \text{(massive)} \\ \mathfrak{u}/d & \text{(massless)} \end{cases} \tag{239}$$

to the energy density

$$\mathfrak{u} \equiv \frac{\langle H\rangle}{V} = -\frac{1}{V}\partial_\beta|_z \ln Z_G = \sum_{\mathbf{k}} \langle n_{\mathbf{k}}\rangle \varepsilon_{\mathbf{k}} = \int_0^\infty \mathrm{d}\varepsilon\, g(\varepsilon) n(\varepsilon)\varepsilon = \int_0^\infty \mathrm{d}\varepsilon\, w(\varepsilon)\varepsilon \; . \tag{240}$$

The purely kinematic relation between p and $\mathfrak{u}$ was already derived in Eq. (110) from elementary considerations for the classical ideal gas. Alternatively, it may be obtained from the virial theorem (exercises). Both alternative derivations underscore that it holds for arbitrary statistics (arbitrary δ), and merely depends on the space dimension and the dispersion relation encoded in the density of states of the ideal gas. It is exactly this part of the classical ideal gas equations of state that survives at low temperatures, whereas the dependence of the energy density and pressure on temperature and number density are modified by exchange effects.

So far, the computational task has only been shifted from the pressure to the internal energy. In calculating thermodynamic quantities one will typically end up with integrals of the above type, which have been tabulated as functions of the fugacity $z = e^{\beta\mu}$, under the name of polylogarithms or Fermi–Dirac ($\delta = 1$) and Bose–Einstein ($\delta = -1$) integrals, respectively[44]

$$\frac{1}{\Gamma(\nu)}\int_0^\infty \mathrm{d}x\, \frac{x^{\nu-1}}{z^{-1}e^x \pm 1} = \begin{cases} f_\nu(z) \\ g_\nu(z) \end{cases} \; . \tag{241}$$

Their graphs are not very exciting. They lie close to z if ν is large, curving somewhat upwards and eventually diverging at $z = 1$ (bosons) or downwards like $\log^\nu(z)$ for large z (fermions), for the more relevant small values of ν. For $z = 1$, corresponding to the marginal case $\mu = 0$, the Bose integrals $g_\nu(1)$ reduce to the Riemann zeta-function[45]

$$g_\nu(1) = \zeta(\nu) \equiv \sum_{k=1}^\infty k^{-\nu} \; . \tag{242}$$

---

[44] The Bose–Einstein function, which always comes with an index and a dimensionless argument, should of course not be confused with the density of states.

[45] Use Euler's geometric sum formula, $(e^x - 1)^{-1} = e^{-x}/(1 - e^{-x}) = \sum_{k=1}^\infty e^{-kx}$ and the definition of the Γ-function to confirm this.

With the new notation, the pressure $p = \tau \mathfrak{u}/d$ and the energy density of some important ideal quantum gases in three dimensions ($d = 3$) are expressed in the form

$$p = \frac{\tau}{3}\mathfrak{u} = \begin{cases} \mathfrak{g}\frac{f_{5/2}(z)}{\beta\lambda_T^3} & \text{(massive fermions)} \\ \mathfrak{g}\frac{g_{5/2}(z)}{\beta\lambda_T^3} & \text{(massive bosons)} \\ \mathfrak{g}\frac{\pi^2}{45}\frac{(k_BT)^4}{(\hbar c)^3} & \text{(massless non-conserved bosons)} \end{cases} \tag{243}$$

The number of **massless bosons** like photons and phonons is usually not *per se* conserved[46] so that their chemical potential vanishes in equilibrium. This allows for the explicit evaluation of the energy density using $\zeta(4) = \pi^4/90$ for $z = 1$ and moreover implies

$$\mu = 0 \quad \Longrightarrow \quad \mathfrak{s} = (\mathfrak{u} + p)/T = (1 + \tau/3)\mathfrak{u}/T = (1 + 3/\tau)p/T \tag{244}$$

After considering the emission geometry from a hole in a hot oven, this equation yields the Stefan–Boltzmann law of black-body radiation. Moreover, the integrand in Eq. (240) can for massless bosons be rewritten as a function of frequency ω in the form

$$\hbar\omega w(\omega) = \hbar\omega g(\omega) n(\omega) = \frac{\mathfrak{g}\hbar\omega^3}{2\pi^2 c^3}\frac{1}{e^{\beta\hbar\omega} - 1} \,. \tag{245}$$

Up to factors considering the radiation geometry this should be recognized as Planck's famous *black-body spectrum*. It is found to be very accurately realized in the cosmic microwave background (CMB) with $T \approx 2.7K$, which is nowadays measured with ever increasing precision. Alternatively, if c is interpreted as the sound velocity instead of the velocity of light (and considering $\mathfrak{g} = 3$ instead of $\mathfrak{g} = 2$), it can be interpreted as the *Debye phonon spectrum* in a solid at frequencies lower than a cutoff frequency $c\,v^{-1/3}$ corresponding to a minimum wavelength $v^{1/3}$ (on the order of the size of the unit cell) and a maximum phonon energy $k_B T_\mathrm{D} \simeq \hbar\, c\, v^{-1/3}$, with T_D the Debye temperature (roughly around 10^3 K for many crystals).

In summary, thermodynamic observables for ideal quantum gases will generally have the form

$$\langle \hat{A} \rangle = \int_0^\infty \mathrm{d}\varepsilon\, w(\varepsilon) A(\varepsilon) \quad \text{with} \quad w(\varepsilon) = \underbrace{g(\varepsilon)}_{\text{kinem.}}\ \underbrace{n(\varepsilon)}_{\text{statist.}} \,. \tag{246}$$

At first sight, the expression for the energy distribution $w(\varepsilon)$ looks very much like the one for a classical ideal gas, Eq. (122). But note that, for $\delta \neq 0$, the good old Boltzmann weight is replaced by the Fermi- and Bose distributions $n(\varepsilon)$

---

[46] Massless bosonic excitations (quantized classical waves) usually have weak self-interactions and are thermalized "externally" by mechanisms that do not conserve their numbers.

(which, in general, still depends on μ) and the $N-$particle density of states is replaced by the level density of a *single* particle in a box. The peculiarity that an ideal gas of N particles admits a description with a single-particle formalism, is already familiar from the discussion in Sec. 3.2. If $A(\varepsilon)$ is of polynomial form, the integrals can be expressed by polylogarithms, as exemplified in Eq. (243).

The case of non-conserved massless bosons is special, since $\mu = 0$ means $z = 1$, where the integrals can then be evaluated in closed form using the zeta function. For fermions and conserved bosons, the interpretation of the results in Eqs. (243), (246) is less obvious. They still depend on the (generally unknown) chemical potential μ via the fugacity $z = e^{\beta\mu}$, the dimensionless argument of the Bose- and Fermi-functions. The following paragraphs show that a similar simplification as for non-conserved bosons occurs for conserved bosons if a condensate is present, and how one otherwise can calculate $\mu(\langle N\rangle)$ and the normalization of $n(\varepsilon)$, perturbatively.

4.2.3 Bose–Einstein condensation

In this subsection the low temperature behavior of Bose gases is discussed in greater detail. For concreteness, consider massive bosons with spin 0 in three space dimensions

$$\delta = -1\ ,\ \tau = 2\ ,\ \mathfrak{g} = 1\ ,\ d = 3\ . \tag{247}$$

Note that there appears to be a potential problem with Eq. (236) for the average particle number if the integral approximation Eq. (225) is used,

$$\frac{\langle N\rangle}{V} = \frac{1}{V}\sum_{\mathbf{k}}\langle n_{\mathbf{k}}\rangle \simeq \int_0^\infty \mathrm{d}\varepsilon\, g(\varepsilon)n(\varepsilon) \overset{\delta=-1}{=} g_{3/2}(z)/\lambda_T^3 \tag{248}$$

The Bose function $g_{3/2}(z)$ is a monotonically increasing function of z which takes its maximum for $z = 1$:

$$n\lambda_T^3 = g_{3/2}(z) \le g_{3/2}(1) = \zeta(3/2) \approx 2.612\ . \tag{249}$$

This equation clearly fails to hold for high densities $\langle N\rangle/V \equiv n > n_c \equiv \zeta(3/2)\lambda_T^{-3}$ (which defines a critical density n_c for prescribed temperature) and below a critical temperature $T_c =$ given by $\lambda_{T_c}^3 \equiv \zeta(3/2)/n$ (for prescribed density n). The intuitive interpretation of the condition Eq. (249) is that the particles' wave packets start to overlap or that their thermal energy would be overwhelmed by their kinetic rest energy due to uncertainty, if you tried to prevent this.

One might think that the breakdown of Eq. (248) detected in Eq. (249) could be indicative of a maximum density, which would make sense for fermions but sounds implausible for bosons. In fact, the limit $z \to 1$ was noticed to suffer from convergence problems, above, and it is now high time to have a closer look at these issues. For $\mu \to 0$ or $z = e^{\beta\mu} \to 1$ the ground state contribution (for $\varepsilon = \varepsilon_0 \to 0$)

to the number distribution $n(\varepsilon_{\mathbf{k}})$ in Eq. (237) diverges in the thermodynamic limit. This comes in quite handy now. The divergence that is required to rescue Eq. (236) corresponds to an *extensive ground state occupancy*,

$$n(\varepsilon_0) = \frac{1}{e^{\beta(\varepsilon_0-\mu)}-1} \sim \frac{k_B T}{\varepsilon_0-\mu} = \mathcal{O}(N) = \mathcal{O}(V) \ . \tag{250}$$

One can convince oneself that the ground state alone absorbs the problem and excited states yield only negligible contributions in the thermodynamic limit[47]. Accordingly, for $T \leq T_c$ (or $n \geq n_c$) and in the thermodynamic limit, the appropriate way to write Eq. (248) with the integral approximation is

$$\begin{aligned} n = \frac{\langle N\rangle}{V} &= \int_0^\infty \mathrm{d}\varepsilon\, g(\varepsilon) n(\varepsilon) + \frac{n(\varepsilon_0)}{V} \\ \Rightarrow \qquad n\lambda_T^3 &= \zeta(3/2) + \frac{\lambda_T^3}{V} n(\varepsilon_0) \ . \qquad (T \leq T_c) \end{aligned} \tag{251}$$

The integral approximation is now only used for terms yielding negligible individual contributions to the total sum, as it should be. The fact that the additional term $n(\varepsilon_0)/V$ is of order one in Eq. (251) explains why it could be neglected in the discussion of u and p, above, where its contribution to the integral would have been suppressed by the multiplicative term ε in the integrand, anyway.

The sudden onset of an extensive population of the ground state is called *Bose–Einstein condensation* (BEC). It manifests itself most dramatically in the velocity distribution of the gas that develops a central peak at the origin for $T < T_c$ (figure!). BEC is the paradigm for transitions into a state with a macroscopic occupancy of a pure quantum state, for which suprafluids and superconductors provide spectacular (but more complicated) examples. The first realizations of BEC with atoms were achieved in the 1990s, 70 years after it was theoretically predicted, by rapid cooling of relatively dilute (hence weakly interacting) atomic gases into the μK range using elaborate technical equipment. It took another decade to achieve the same for gases of fermionic atoms, since fermions have first to pair up into bosonic quasi-particles that can then themselves undergo BEC. In this form BEC is very analogous to superconductivity. Nowadays BECs are studied in free fall to reduce gravity, which, amazingly, counts among the limiting disturbances of the homogeneity of the temperature in a BEC. BECs of a variety of quasi-particles such as exciton-polaritons, and even of bare photons, have been produced.

The thermodynamics of a massive Bose gas in presence of a condensate, i.e. below T_c, is relatively easy to analyze due to the free exchange of particles between the condensate and the excited states, which pins the chemical potential

---

[47] Here is a hint: from Eq. (223), the occupancy of the first excited state will at most be on the order of $k_BT/(\hbar^2\pi^2/2mV^{2/3}+\varepsilon_0-\mu) \propto V^{2/3}$ — hence still large but down by some factor $V^{-1/3}$ and therefore non-extensive in the thermodynamic limit.

(essentially) at $\mu = 0$. To obtain the "normal properties" just above the transition is slightly less straightforward, since they depend on a negative chemical potential $\mu < 0$, which has to be calculated from Eq. (231). An example of a perturbative calculation of μ is given for the Fermi gas in the next subsection, and a calculation of the limiting asymptotic approach to the classical gas behavior can be found in the subsequent paragraph about the classical limit. This is taken as an excuse for restricting the present discussion to the case with condensate, where $\mu = 0$, $z = 1$.

Then, Eq. (251) can be read as an equation for the redistribution of particles between the ground state and the excited states. The *condensate fraction*

$$\psi \equiv \frac{N_0}{\langle N \rangle} = \frac{n(\varepsilon_0)}{nV} \overset{T \leq T_c}{=} 1 - \frac{\zeta(3/2)}{n\lambda_T^3} \tag{252}$$

plays the role of an *order parameter*: it vanishes above the transition and continuously grows towards one below T_c (above n_c) (sketch!),

$$\psi = \begin{cases} 0 & (T \geq T_c,\ n \leq n_c) \\ 1 - (T/T_c)^{3/2} = 1 - n_c/n & (T \leq T_c,\ n \geq n_c) \end{cases} . \tag{253}$$

BEC is thus a continuous transition. As already noticed above, the expressions derived for the thermodynamic quantities such as the pressure p, the energy density $\mathfrak{u}$, the entropy density $\mathfrak{s} = (\mathfrak{u}+p)/T = 5\mathfrak{u}/3T$, etc., do not acquire corrections below T_c. But with Eq. (253) they can now be given a new intuitive physical interpretation by rewriting them in a form that emphasizes their proportionality to the normal fraction $1 - \psi$ of the gas. So by combination of Eqs. (243), Eq. (249), (253), one finds

$$p = 2\mathfrak{u}/3 = 2\mathfrak{s}T/5 = \frac{\zeta(5/2)}{\zeta(3/2)} n k_B T (T/T_c)^{3/2} \approx 0.51 k_B T n_c \,, \tag{254}$$

$$n c_V = T\partial_T \mathfrak{s} = \partial_T \mathfrak{u} = \frac{15\zeta(5/2)}{4\zeta(3/2)} n (T/T_c)^{3/2} \approx 1.9 k_B n_c \,, \tag{255}$$

with the density $n_c(T) = n(1-\psi) = n(T/T_c)^{3/2}$ of normal (not condensed) particles. Plainly, the condensate, as a pure quantum state, neither carries entropy nor heat, nor does it exert pressure to the surrounding normal fraction[48]. Therefore, the thermodynamic properties of the gas are entirely determined by the normal fraction. A rough but essentially correct characterization of the thermodynamics of the Bose gas in presence of the condensate is thus the following: *it is all due to the particles of the normal fraction* $1 - \psi$*, which behave more or less like a classical ideal gas.*

---

[48] It exerts a quantum mechanical ground-state pressure on the walls of the box, though.

From Eq. (254), the isotherms in a $p - v$ or $p - n$ diagram are found to be horizontal (independent of n), which is reminiscent of a liquid-vapour transitions (sketch!). The coexistence pressure grows with temperature according to

$$p_c(T) = p(n = n_c) = \zeta(5/2)(2\pi m/h^2)^{3/2}(k_B T)^{5/2} \; . \tag{256}$$

This also defines a line of maximum pressure that delimits the physically accessible region, hence a critical pressure $p_c = p(n \geq n_c)$ in a $p - T$ diagram. Using Eq. (249), it may alternatively be expressed as a function of the specific volume v at $T = T_c$,

$$p_c(v) = p(T = T_c) = \frac{h^2 \zeta(5/2)}{2\pi m \zeta(3/2)^{5/3}} v^{-5/3} \; , \tag{257}$$

defining a line akin to the liquid-vapour binodal in the $p-v$ diagram of a classical liquid-gas transition. As familiar from liquid-vapor coexistence, any attempt to compress the gas beyond that limit merely increases the condensate fraction ψ. (In reality, if one crosses this line, one compresses a pure condensate that eventually turns into a solid or liquid, which is clearly out of reach for the present discussion based on non-interacting particles.)

The thermodynamic properties in presence of a BEC do not attain their classical limiting form right at T_c. To study the convergence to the classical limit, $\mu(T > T_c)$ has to be calculated perturbatively, as already mentioned. (Such a calculation is exemplarily performed for the Fermi gas, below.) One then finds, for example, that the specific heat decreases continuously from its critical value $c_V(T_c) \approx 1.9 k_B$ to its classical value $c_V(T \to \infty) = 3k_B/2$, with a kink but no discontinuity at T_c (sketch!).

4.2.4 Almost degenerate Fermi gas and Sommerfeld expansion

Consider, as in the preceding paragraph, an ideal gas of massive particles in three-dimensional space, but now let them be fermions instead of bosons, i.e.

$$\delta = 1 \; , \; \tau = 2 \; , \; d = 3 \; . \tag{258}$$

What is the difference? It turns out that, at low temperatures and high densities, the behavior of an ideal Fermi gas could hardly be less similar to that of a Bose gas. Maybe the most striking difference is the different sign of the chemical potential, which results from the self-avoidance due to the different exchange properties.

The chemical potential μ is implicitly given by the number conservation equation. In the integral approximation, it reads

$$n = \int \mathrm{d}\varepsilon \, g(\varepsilon) n(\varepsilon) = \mathfrak{g} f_{3/2}(z)/\lambda_T^3 \; . \tag{259}$$

At low temperature, the Fermi distribution degenerates to a Heavyside step function with the step marking the value of the chemical potential at $T = 0$,

$$n(\varepsilon) = \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} \xrightarrow{\beta\to\infty} \theta(\varepsilon_F - \varepsilon) \,, \tag{260}$$

so that Eq. (259) simplifies to

$$n = \Gamma(\varepsilon_F) = \int_0^{\varepsilon_F} \mathrm{d}\varepsilon \, g(\varepsilon) = \frac{\tau}{d} \mathfrak{g} g_0 \varepsilon_F^{d/\tau} \,, \tag{261}$$

with the single-particle momentum space volume $\Gamma(\varepsilon)$ and the Fermi energy

$$\varepsilon_F \equiv \mu(T=0) = [d\,n/(\mathfrak{g} g_0 \tau)]^{\tau/d} \,. \tag{262}$$

Here Γ is not the gamma function, and there is the same abuse of the earlier notation for the total phase space volume as for $g(\varepsilon)$ in Eq. (225). Equation. (261) says that the number of accessible quantum states (or the phase space volume in unit of λ_T^3) $\Gamma(\varepsilon_F)V$ is equal to the number of particles N (rather than on the order of $e^{\langle N\rangle}$). The $N/\mathfrak{g}$ lowest levels are occupied, each $\mathfrak{g}$ times, and all higher levels are empty. This is just as in the ground state of an atom, which certainly justifies the notion of a "degenerate" gas. Writing $\varepsilon_F \equiv k_B T_F \equiv p_F^2/2m$, one defines the Fermi temperature and Fermi momentum, accordingly. With Eq. (258), Eq. (262) can then alternatively be written in the forms

$$p_F = \hbar(6\pi^2 n/\mathfrak{g})^{1/3} \quad \text{and} \quad n\lambda_{T_F}^3 = \frac{4\mathfrak{g}}{3\sqrt{\pi}} \,. \tag{263}$$

Given that the density of (unbound) conduction electrons in a metal is very roughly of the same order of magnitude as the electron density in an atom, one estimates $n \simeq \text{Å}^{-3}$ and $\varepsilon_F \simeq \text{Ry}$, which implies $T_F \simeq 10^5$ K. Hence, at room temperature the conduction electrons in metals correspond very well to a degenerate Fermi gas, if their mutual Coulomb interactions are neglected. Due to their long range, this is indeed a very reasonable approximation.

In sharp contrast to the Bose gas, the average energy and pressure do not vanish in a Fermi gas for $T \to 0$,

$$p_0 = \frac{\tau}{d}\mathfrak{u}_0 = \frac{\tau}{d}\int_0^{\varepsilon_F} \mathrm{d}\varepsilon \, g(\varepsilon)\varepsilon = \frac{n\varepsilon_F}{1 + d/\tau} \propto n^{1+\tau/d} \,. \tag{264}$$

Superficially, the result $p_0 \propto n^{5/3} \propto v^{-5/3}$ is reminiscent of the relation for the critical (maximum) pressure of the Bose gas, but while Bose condensation bounds the pressure of the Bose gas from above, the Pauli principle bounds the pressure of a Fermi gas from below.[49] This means that a Fermi gas needs to be confined

---

[49] This should intuitively be clear. The formal justification follows below.

by some external means that counterbalance the Fermi pressure. (For the gas of conduction electrons in a metal the confinement is due to the potential energy that holds the material together.) Similar as for the Bose condensate, the entropy $\mathfrak{s}_0$ of the "Fermi condensate" (the degenerate Fermi gas) vanishes, though. To see this, note that, with $\mathfrak{u}_0 = 3n\varepsilon_F/5$ and $p_0 = 2n\varepsilon_F/5$, $T\mathfrak{s}_0 = \mathfrak{u}_0 + p_0 - n\varepsilon_F = 0$. It turns out that the finite-temperature corrections to $\mathfrak{u}_0$, $\mu_0 \equiv \varepsilon_F$, and p_0 are of order $\mathcal{O}(T^2)$, hence $\mathfrak{s}(T \to 0) = \mathcal{O}(T)$, as required by the third law.

To formally demonstrate the last statements, $\mu(T > 0)$ has to be calculated perturbatively. One can exploit that the derivative of $n(\varepsilon)$ is almost a Dirac δ-function for the almost degenerate Fermi gas to arrive at a low-temperature expansion known as the "Sommerfeld expansion". The idea is to rewrite the integrands in the form $\partial_\varepsilon n(\varepsilon)$ (the δ-function like peak) times a rest that can be expanded around the peak position (sketch). At vanishing temperature the peak sits at $\varepsilon = \varepsilon_F$, but at finite temperatures $T > 0$ at $\mu(T) \neq \varepsilon_F$, which must be calculated perturbatively, itself, using the number conservation equation:

$$n = \int_0^\infty \mathrm{d}\varepsilon\, g(\varepsilon) n(\varepsilon) = \underbrace{\Gamma(\varepsilon) n(\varepsilon)\big|_0^\infty}_{0 \text{ in } d=3} - \int_0^\infty \mathrm{d}\varepsilon\, \Gamma(\varepsilon) \partial_\varepsilon n(\varepsilon)\ . \tag{265}$$

Now, expanding $\Gamma(\varepsilon)$ to second order around the yet unknown position μ of the peak yields

$$n = \underbrace{\Gamma(\mu)}_{\mathfrak{g} g_0 \frac{\tau}{d} \mu^{\frac{d}{\tau}}}\ \underbrace{n(0)}_{1} - \frac{1}{2!} \underbrace{\partial_\varepsilon g(\varepsilon)\big|_{\varepsilon=\mu}}_{\mathfrak{g} g_0 (\frac{d}{\tau} - 1) \mu^{\frac{d}{\tau} - 2}} \int_0^\infty \mathrm{d}\varepsilon\, (\varepsilon - \mu)^2 \partial_\varepsilon n(\varepsilon)\ . \tag{266}$$

The lower bound of the integral can be shifted to negative infinity without incurring serious errors, and, due to the symmetry of $\partial_\varepsilon n(\varepsilon)$, the integral over the first order term vanishes. Hence, the calculation of corrections to the fully degenerate Fermi gas in Eq. (264) boils down to the evaluation of integrals of the form

$$\int_{-\infty}^\infty \mathrm{d}\varepsilon\, (\varepsilon - \mu)^\gamma \partial_\varepsilon n(\varepsilon) = \beta^{-\gamma} \int_{-\infty}^\infty \mathrm{d}x\, \frac{x^\gamma e^x}{(e^x + 1)^2} \xrightarrow{\gamma=2} \frac{\pi^2}{3\beta^2} \tag{267}$$

with $x \equiv \beta(\varepsilon - \mu)$. The first non-vanishing correction to the chemical potential is obtained for $\gamma = 2$ after eliminating n from Eq. (266) by use of Eqs. (261), (262),

$$1 = \Big(\frac{\mu}{\varepsilon_F}\Big)^{d/\tau} \Big[1 + \frac{\pi^2}{6\beta^2} \frac{d}{\tau} \Big(\frac{d}{\tau} - 1\Big) \mu^{-2}\Big]\ . \tag{268}$$

Perturbative solution for μ (and completely analogous calculations for $\mathfrak{u}$) yield for $d = 3$, $\tau = 2$

$$\mu \sim \varepsilon_F \Big[1 - \frac{\pi^2}{12} \Big(\frac{T}{T_F}\Big)^2\Big] \qquad \mathfrak{u} = \frac{3}{2} p \sim \frac{3}{5} n \varepsilon_F \Big[1 + \frac{5\pi^2}{12} \Big(\frac{T}{T_F}\Big)^2\Big]\ . \tag{269}$$

The obvious interpretation goes much along the lines of what was said for the cold Bose gas, above. The correction to the ground state energy $\mathfrak{u}_0$ and pressure p_0 of the frozen fermion condensate is given by a normal fraction $1-\psi = T/T_F$ of particles which exhibit essentially classical ideal gas behavior:

$$p - p_0 = \frac{2}{3}(\mathfrak{u} - \mathfrak{u}_0) = \frac{\pi^2}{6} n_c k_B T \quad \text{where} \quad n_c \equiv n(1-\psi) = nT/T_F \tag{270}$$

again denotes the normal (not condensed) fraction of the gas.

4.2.5 Low-temperature specific heats and the third law

Particularly wide reaching consequences that follow immediately from the above results concern the low-temperature specific heat not only of gases but also of fluids and solids. Thermal transport in the latter is dominated by *elementary excitations*, which can in a reasonable approximation often be represented as *weakly interacting or ideal quantum gases of quasi-particles.* These comprise sound waves ("phonons") and excitations of the electronic degrees of freedom (so-called "electrons" or "holes") in conductors or semiconductors, spin waves ("magnons") in magnets, and so on. The derivation of the specific heat of ideal quantum gases goes a long way towards explaining the microscopic origin of the third law.

The contributions by the conduction electrons to heat and charge transport in insulators and semiconductors freeze out with a Schottky anomaly (i.e. exponentially, as for the Ising chain), because of their band gap. Therefore, sound waves, which are bosons, dominate their low-temperature specific heat and heat transport. In normal conductors, the conduction electrons, which are reasonably well described as an almost degenerate Fermi gas, also make a relevant contribution to the specific heat at low temperatures.

More specifically, for non-conserved bosons or conserved (massive) bosons at low temperatures, assuming a vanishing chemical potential $\beta\mu = 0$ such that $z \to 1$ and $g_{d/\tau+1}(z) \to \zeta(d/\tau + 1)$, Eq. (243) gives

$$nc_V = (\partial_T \mathfrak{u})_V \qquad \Rightarrow \qquad c_V/k_B \propto T^{d/\tau} = \begin{cases} T^{d/2} & \text{(massive bosons)} \\ T^{d} & \text{(massless bosons)} \end{cases} . \tag{271}$$

which should be contrasted with the classical predictions $c_V/k_B = d/\tau =$ constant. Sound waves in solids, i.e. phonons with $c_V/k_B = (2\pi^2/5)(T/T_D)^3$, and spin waves in antiferromagnets behave as massless bosons, while the massive case pertains to ferromagnetic spin waves and bosonic atoms. For all except the last example, the temperature does not have to be terribly low to observe the low-temperature behavior.

In contrast, the specific heat of an ideal Fermi gas at low temperatures is

$$c_V/k_B = \frac{\pi^2}{2}\frac{T}{T_F} \qquad \text{(fermions)} \tag{272}$$

Formally, this is a consequence of $\mathfrak{u}(T \ll T_F) \simeq n[k_B T_F + \mathcal{O}(T/T_F)^2]$. Physically, it can be rationalized by observing that the Fermi gas freezes if $T \ll T_F$. The Fermi distribution near ε_F is softened within a range $k_B T$ compared to the limiting step function, so T/T_F is the fraction of particles excited from the region $\varepsilon < \varepsilon_F$ to the region $\varepsilon > \varepsilon_F$. In other words, since the fraction ψ of particles frozen into their ground state does not contribute to transport (except for the Fermi pressure resulting from the Pauli principle), only the normal fraction $1-\psi$ of about T/T_F behaves gas-like. Similarly as discussed for bosons with a BEC, it is only the normal fraction that contributes to heat and charge transport and carries entropy, so that the corresponding potentials and transport coefficients are reduced compared to the classical laws at $T \ll T_F$ by this factor.

An important aspect of the foregoing discussion is that the notion of "low temperatures", which always means $n\lambda_T^3 \gtrsim 1$ for an ideal gas, can refer to utterly different temperature values for typical fermi and bose gases. Remembering that $n\lambda_T = z = e^{\beta\mu}$ for a classical ideal gas, one expects the interesting things to happen at $k_B T \simeq |\mu|$. But the chemical potentials of bosons and fermions have different signs and magnitudes. Bosons can all settle into the ground state at $T \to 0$, which can give or take particles at no cost, so that $\mu(T \to 0)$ vanishes upon cooling, and the interesting things happen at very low temperatures, accordingly. In contrast, due to the Pauli principle, N fermions will condense into the lowest N states, instead. For conduction electrons in a metal, this implies that the chemical potential is on the order of the Rydberg energy, even for $T \to 0$, i.e. $\mu(T \to 0) = k_B T_F$ with a Fermi temperature T_F on the order of tens of thousands of Kelvin, so that conduction electrons are deeply frozen at room temperature.

4.3 Classical limit and exchange interactions

As demonstrated in the preceding subsections, *non-interacting but indistinguishable particles* exhibit strongly correlated behavior (dependent on their exchange statistics encoded in the parameter δ) at high densities and low temperatures, where a quantum mechanical description applies. In the case of bosons, even a phase transition occurs. In brief, they share certain features with, and are of comparable complexity as, interacting classical systems. In fact, requiring N particles to be indistinguishable imposes a type of N-body interaction that is *not decomposable into pair interactions* (which would act between any pair of particles, irrespective of what all the other particles are meanwhile doing). A simple, convenient pair description only emerges in the "classical limit"

$$n\lambda_T^3 \to 0 \tag{273}$$

(in $d = 3$ dimensional space) corresponding to high temperatures and low densities, such that the available energy states are sparsely occupied ($\Gamma \gg N$) and can be approximated as a continuum, and multiple occupations have little chance to occur. The different exchange properties of fermions and bosons turn out to still make a sizable effect through the (anti-)symmetrized wave functions, giving rise to effective entropic interactions. By exploring this limit quantitatively, one can therefore get some idea of how interacting many-body systems might be approached.

For simplicity, spin degeneracies are in the following neglected ($\mathfrak{g} = 1$). With the (anti-)symmetric product eigenstates $|\{\mathbf{p}_i\}\rangle$ of Eq. (229) and (222) the canonical partition sum then reads

$$Z = \mathrm{tr}\exp\bigl[-\beta\hat{H}\bigr] = \sum_{\{\mathbf{p}_i\}} \langle\{\mathbf{p}_i\}|\exp\bigl[-\beta\textstyle\sum_i \frac{\mathbf{p}_i^2}{2m}\bigr]|\{\mathbf{p}_i\}\rangle \ . \tag{274}$$

The trace amounts to a summation "$\sum_{\{\mathbf{p}_i\}}$" extending over *all (anti-)symmetrized N-particle states* enumerated by the sets $\{\mathbf{p}_i\}$. This counts all possible combinations of single particle energy levels. It therefore should not be interpreted as $\sum_{\mathbf{p}_1}\cdots\sum_{\mathbf{p}_N}$, as this would generate $N!$ redundant terms[50]. With Eq. (229) the bracket is written in the form

$$\langle\{\mathbf{p}_i\}|\{\mathbf{p}_i\}\rangle = \frac{1}{N!}\sum_{\mathcal{P},\,\mathcal{P}'}(-\delta)^{P+P'}\langle\mathbf{p}_{\mathcal{P}1}|\mathbf{p}_{\mathcal{P}'1}\rangle\langle\mathbf{p}_{\mathcal{P}2}|\mathbf{p}_{\mathcal{P}'2}\rangle\ldots\langle\mathbf{p}_{\mathcal{P}N}|\mathbf{p}_{\mathcal{P}'N}\rangle \tag{275}$$

Here, one of the redundant sets of permutations can be canceled against the factor $1/N!$, i.e. all occurrences of $\mathcal{P}'$ and P' can be crossed out together with the factorial factor. Now, insert N unity operators $\int\!\mathrm{d}\mathbf{r}_i\,|\mathbf{r}_i\rangle\langle\mathbf{r}_i|/V$ constructed from one-particle position eigenstates $|\mathbf{r}\rangle$ into the one-particle momentum-state brackets. After rearranging the factors in order to gather terms with the same momentum, and using Eq. (222), one has

$$\langle\{\mathbf{p}_i\}|\{\mathbf{p}_i\}\rangle = \int\!\frac{\mathrm{d}^N\mathbf{r}}{V^N}\,\sum_{\mathcal{P}}(-\delta)^P e^{-i\mathbf{p}_1(\mathbf{r}_1-\mathbf{r}_{\mathcal{P}1})/\hbar}\ldots e^{-i\mathbf{p}_N(\mathbf{r}_N-\mathbf{r}_{\mathcal{P}N})/\hbar} \ .$$

Upon rearranging factors, the permutation symbol has moved from the momentum to the particle coordinates. The next step involves going from the level sum in Eq. (274) to a sum over all positive one-particle momenta (which is then straightforwardly turned into an integration over the full momentum space divided by the factor 2^d),

$$\sum_{\{\mathbf{p}_i\}} = \frac{1}{N!}\sum_{\mathbf{p}_1}\cdots\sum_{\mathbf{p}_N} \ \to\ V^N\!\int\!\frac{\mathrm{d}^N\mathbf{p}}{h^{3N}N!} \ . \tag{276}$$

This is how the classical phase space integration measure evolves from the sum over (anti-)symmetrized many-particle quantum states, thereby justifying *a posteriori* the normalization of the phase space volume by Plack's constant, as introduced in Part II of the lecture. Altogether, the partition sum in the continuum

---

[50] One cannot prescribe to the (anti-)symmetrized wave function which particle has which momentum — all combinations are automatically included.

limit reads

$$\begin{aligned} Z = \int \frac{\mathrm{d}^N\mathbf{p}\,\mathrm{d}^N\mathbf{r}}{N!h^{3N}} \sum_{\mathcal{P}} (-\delta)^P \exp\bigl[-\beta\mathbf{p}_1^2/2m - i\mathbf{p}_1\cdot(\mathbf{r}_1 - \mathbf{r}_{\mathcal{P}1})/\hbar\bigr]\cdots \\ \cdots\exp\bigl[-\beta\mathbf{p}_N^2/2m - i\mathbf{p}_N\cdot(\mathbf{r}_N - \mathbf{r}_{\mathcal{P}N})/\hbar\bigr] \; . \end{aligned} \tag{277}$$

As usual, the momentum integrals can be performed immediately (first complete the square in the exponentials), and using the definition

$$f(\mathbf{r}) \equiv \exp[-\pi\mathbf{r}^2/\lambda_T^2] \sim \begin{cases} 1 & (|\mathbf{r}| \ll \lambda_T) \\ 0 & (|\mathbf{r}| \gg \lambda_T) \end{cases} \tag{278}$$

one then is left with the configuration integral (Z_{ig} denotes the partition sum for a classical ideal gas)

$$Z = Z_{\text{ig}} \int \frac{\mathrm{d}^N\mathbf{r}}{V^N} \sum_{\mathcal{P}} (-\delta)^P f(\mathbf{r}_1 - \mathbf{r}_{\mathcal{P}1})\cdots f(\mathbf{r}_N - \mathbf{r}_{\mathcal{P}N}) \; . \tag{279}$$

The length of the permutation P (i.e. the number of exchanged particle pairs) counts the factors that differ from unity in each term of the sum. The unit permutation ($P = 0$) simply produces a one in the form $f(0)^N$, which, of course, corresponds to the classical ideal gas of indistinguishable particles. Since the function $f(\mathbf{r})$ vanishes quickly for distances $r \gg \lambda_T$, permutations are likely to produce vanishing contributions to the sum. Non-vanishing contributions only result if the exchanged particles happen to reside in the same volume element of size λ_T^3, which occurs rarely in the classical limit $n\lambda_T^3 \to 0$. The estimated contributions of the products of f's thus decrease like $(n\lambda_T^3)^P$ with the length P of the permutation, which can therefore be used as an ordering scheme for the series in Eq. (279). To extract the dominant correction to the classical ideal gas in a low-density expansion, it suffices to truncate the series after the second order, with only two non-trivial factors of f (for indices ij and ji, hence the $\sum_{i<j}$)

$$\sum_{\mathcal{P}} (-\delta)^P f(\mathbf{r}_1 - \mathbf{r}_{\mathcal{P}1})\cdots f(\mathbf{r}_N - \mathbf{r}_{\mathcal{P}N}) = 1 - \delta\sum_{i<j} f^2(\mathbf{r}_i - \mathbf{r}_j) + \mathcal{O}(f^3) \; . \tag{280}$$

The leading correction to the classical ideal gas (of indistinguishable particles) can be interpreted as consequence of a fictitious pair interaction potential (sketch!)

$$V_{\text{eff}}(\{\mathbf{r}_i\}) \equiv \sum_{i<j} \nu_{\text{eff}}(\mathbf{r}_i - \mathbf{r}_j) \, , \quad \nu_{\text{eff}}(\mathbf{r}) \equiv -k_B T \ln\bigl[1 - \delta\exp\bigl(-2\pi\mathbf{r}^2/\lambda_T^2\bigr)\bigr] \; . \tag{281}$$

via the identification of the integrand in Eq. (279) with a Boltzmann factor,

$$1 - \delta\sum_{i<j} f^2(\mathbf{r}_i - \mathbf{r}_j) \sim \prod_{i<j}[1 - \delta f^2(\mathbf{r}_i - \mathbf{r}_j)] = \exp\bigl[-\beta V_{\text{eff}}(\mathbf{r}_i - \mathbf{r}_j)\bigr] \; . \tag{282}$$

The purely entropic origin of this "statistical potential", which does not involve a real interaction energy, is betrayed by the prefactor k_BT of ν_{eff}. It arises solely from the (anti-)symmetrization of the wave functions in Eq. (229). The corresponding positional correlations between particles sharing the same volume can thus be represented by an effective pair interaction ν_{eff}, in the dilute limit. It is as if identical non-interacting bosons and fermions attract and repel each other, respectively, a point of view that is clearly limited to the classical limit.

If naively extrapolated to higher densities, the interaction representation of the many-body exchange correlations suggests a classical condensation in real space, whereas the real BEC is of course a *condensation in momentum space.* Yet, Eq. (281) can be useful for many practical purposes. Namely, using the fictitious pair potential, which has $N(N-1)/2$ equivalent terms, corrections to the fundamental relation $F_{\text{ig}}(T,V,N)$ of a classical ideal gas are readily estimated (exercises). You simply have to revert the above arguments to get the leading correction to the fundamental relation $F(T,V,N)$ for the classical ideal gas from

$$\frac{Z}{Z_{\text{ig}}} = \int \frac{\mathrm{d}^N \mathbf{r}}{V^N} \exp\bigl[-\beta V_{\text{eff}}(\{\mathbf{r}_i\})\bigr] \sim 1 - \delta \frac{N(N-1)}{2} \int \frac{\mathrm{d}\mathbf{r}}{V}\, e^{-2\pi \mathbf{r}^2/\lambda_T^2} \,,$$

which amounts to the interaction or "excess" free energy (per particle)

$$\beta(F - F_{\text{ig}})/N = Bn \,, \quad \text{with the second virial coefficient} \quad B = \delta \lambda_T^3/2^{5/2} \,. \qquad (283)$$

From this, the proper asymptotic behavior of the specific heat c_V above the BEC transition (a decay $\propto T^{-3/2}$ to the classical value $3k_B/2$) follows, and similarly good estimates for the pressure, $p = nk_BT(1+Bn)$. The maximum errors — with respect to the correct limits at $T = T_c$ obtained from Eqs. (254), (255), (270), respectively — are only a few percents. The exchange interactions are thus recognized as the culprits for the slight deviations of the low-temperature thermodynamics of the normal fraction from the classical ideal-gas laws, and it often suffices to deal with them on the level of the second virial coefficient.

Chapter summary

Ideal quantum gases, though slightly artificial at first sight, are the basis of our understanding of quantum effects in condensed matter. At low temperature the latter are often governed by dilute gases of elementary excitations, the so-called quasi-particles that arise upon quantizing low-energy collective excitations such as sound waves, spin waves, etc., or representing the effective behavior of an electron or an electron vacancy, say, in a solid. Apart from the most spectacular macroscopic quantum properties such as suprafluidity, superconductance, quantum-Hall resistance, and Bose–Einstein condensation, the consequences of this observation also concern such apparently mundane properties as the normal metal conductance or the specific heat of solids, which could impossibly be understood without understanding the different statistics of bosons and fermions. The effective exchange interactions completely alter the behavior of the gas compared to the classical ideal gas whenever $n\lambda_T^3$ is non-negligible. The consequences are very different for the gas of conductance electrons in a metal as opposed to dilute atomic Bose gases, say, due to their very different chemical potentials μ at low temperatures. Exchange effects start to dominate over conventional classical ideal gas behavior at temperatures as high as many thousands Kelvin in the first case, but not before reaching the lowest temperatures in the universe, in the second. Moreover, the possibility to form a condensate bounds the pressure of Bose gases from above and makes it vanish at low temperatures, whereas the Pauli principle keeps the pressure of Fermi gases finite and thus prevents a collapse of the gas, down to zero temperature. Together with what was said about quasi-particles and elementary excitations, the fact that the specific heat for ideal Fermi and Bose gases vanishes according to a power-law at low temperatures goes a long way in explaining the third law of thermodynamics.

Fermata

Statistical Mechanics has traced back the great generality and simplicity of the laws of Thermodynamics to the assumption of a uniform phase space density on the energy shell. Technically, obtaining thermodynamic fundamental relations for a given Hamiltonian was reduced to the calculation of a phase space volume or a sum over quantum states. Under appropriate additivity assumptions, the fundamental law called the “Summit of Statistical Mechanics” by Feynman was motivated, and quite a number of interesting predictions for the thermodynamics and the fluctuations of gases and interacting spin systems were derived from it. If combined with the concept of quantized elementary excitations, or quasi-particles, and the dissimilar exchange rules for bosons and fermions, far reaching general consequences for the material behavior of gases, fluids, and solids can be deduced. The stunning success of this simple recipe amazes me.

> This brings us to the essential mystery of statistical mechanics, whether equilibrium or nonequilibrium — why do such models work in the first place?
>
> R. Zwanzig