# Statistical Physics

G. Falkovich

http://www.weizmann.ac.il/home/fnfal/papers/statphys13-14.pdf

More is different (P.W. Anderson)

## Contents

This is a graduate one-semester course. Chapters 1-2 briefly remind what is supposed to be known from the undergraduate courses, using a bit more sophisticated language. Chapter 3 treats ideal gases on a level a bit higher than undergraduate, chapters 4-8 is a graduate material.

# 1 Thermodynamics (brief reminder)

Physics is an experimental science, and laws appear usually by induction: from particular cases to a general law and from processes to state functions. The latter step requires integration (to pass, for instance, from Newton equation of mechanics to Hamiltonian or from thermodynamic equations of state to thermodynamic potentials). Generally, it is much easier to differentiate then to integrate and so deduction (or postulation approach) is usually much more simple and elegant. It also provides a good vantage point for further applications and generalizations. In such an approach, one starts from postulating some function of the state of the system and deducing from it the laws that govern changes when one passes from state to state. Here such a deduction is presented for thermodynamics following the book H. B. Callen, *Thermodynamics* (John Wiley & Sons, NYC 1965).

## 1.1 Basic notions

We use macroscopic description so that some degrees of freedom remain hidden. In mechanics, electricity and magnetism we dealt with the explicitly known macroscopic degrees of freedom but in thermodynamics we deal with macroscopic manifestations of the hidden degrees of freedom. When detailed knowledge is unavailable, physicists use symmetries or conservation laws. Thermodynamics studies restrictions on the possible properties of macroscopic matter that follow from the symmetries of the fundamental laws. Therefore, thermodynamics does not predict numerical values but rather sets inequalities and establishes relations among different properties.

The basic symmetry is invariance with respect to time shifts which gives energy conservation[1]. That allows one to introduce the internal energy $E$.

---

[1] Be careful trying to build thermodynamic description for biological or social-economic systems, since generally they are not time-invariant. For instance, living beings age and the amount of money is not always conserved.

Energy change generally consists of two parts: the energy change of macroscopic degrees of freedom (which we shall call work) and the energy change of hidden degrees of freedom (which we shall call heat). To be able to measure energy changes in principle, we need adiabatic processes where there is no heat exchange. We wish to establish the energy of a given system in states independent of the way they are prepared. We call such states equilibrium, they are those that can be completely characterized by the *static* values of extensive parameters like energy $E$, volume $V$ and mole number $N$ (number of particles divided by the Avogadro number $6.02 \times 10^{23}$). Other extensive quantities may include numbers of different sorts of particles, electric and magnetic moments etc i.e. everything which value for a composite system is a direct sum of the values for the components. For a given system, any two equilibrium states A and B can be related by an adiabatic process either $A \to B$ or $B \to A$, which allows to measure the difference in the internal energy by the work $W$ done by the system. Now, if we encounter a process where the energy change is not equal to minus the work done by the system, we call the difference the heat flux into the system:

$$dE = \delta Q - \delta W \ . \tag{1}$$

This statement is known as the first law of thermodynamics. The energy is a function of state so we use differential, but we use $\delta$ for heat and work, which aren't differentials of any function as they refer to particular forms of energy transfer (not energy content).

**The basic problem** of thermodynamics is the determination of the equilibrium state that eventually results after all internal constraints are removed in a closed composite system. The problem is solved with the help of extremum principle: there exists an extensive quantity $S$ called entropy which is a function of the extensive parameters of any composite system. The values assumed by the extensive parameters in the absence of an internal constraint maximize the entropy over the manifold of constrained equilibrium states. Since the entropy is extensive it is a homogeneous first-order function of the extensive parameters: $S(\lambda E, \lambda V, \ldots) = \lambda S(E, V, \ldots)$. The entropy is a continuous differentiable function of its variables. This function (called also fundamental relation) is *everything* one needs to know to solve the basic problem (and other problems in thermodynamics as well).

Since the entropy is generally a monotonic function of energy[2] then $S =$

---

[2] This is not always so, we shall see in the second Chapter that the two-level system

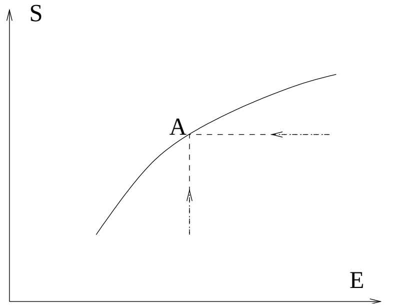
$S(E, V, \ldots)$ can be solved uniquely for $E(S, V, \ldots)$ which is an equivalent fundamental relation. Indeed, assume $(\partial E/\partial S)_X > 0$ and consider $S(E, X)$ and $E(S, X)$. Then[3]

$$\left(\frac{\partial S}{\partial X}\right)_E = 0 \Rightarrow \left(\frac{\partial E}{\partial X}\right)_S = -\frac{\partial(ES)}{\partial(XS)}\frac{\partial(EX)}{\partial(EX)} = -\left(\frac{\partial S}{\partial X}\right)_E \left(\frac{\partial E}{\partial S}\right)_X = 0 \ .$$

Differentiating the last relation one more time we get

$$(\partial^2 E/\partial X^2)_S = -(\partial^2 S/\partial X^2)_E (\partial E/\partial S)_X \ ,$$

since the derivative of the second factor is zero as it is at constant $X$. We thus see that the equilibrium is defined by the energy minimum instead of the entropy maximum (very much like circle can be defined as the figure of either maximal area for a given perimeter or of minimal perimeter for a given area). On the figure, unconstrained equilibrium states lie on the curve while all other states lie below. One can reach the state A either maximizing entropy at a given energy or minimizing energy at a given entropy:



One can work either in energy or entropy representation but ought to be careful not to mix the two.

Experimentally, one usually measures *changes* thus finding derivatives (called equations of state). The partial derivatives of an extensive variable with respect to its arguments (also extensive parameters) are intensive parameters[4]. For example, for the energy one writes

$$\frac{\partial E}{\partial S} \equiv T(S, V, N) \ , \quad \frac{\partial E}{\partial V} \equiv -P(S, V, N) \quad \frac{\partial E}{\partial N} \equiv \mu(S, V, N) \ , \ldots \tag{2}$$

---

gives a counter-example as well as other systems with a finite phase space.

[3] An efficient way to treat partial derivatives is to use jacobians $\partial(u, v)/\partial(x, y) = (\partial u/\partial x)(\partial v/\partial y) - (\partial v/\partial x)(\partial u/\partial y)$ and the identity $(\partial u/\partial x)_y = \partial(u, y)/\partial(x, y)$.

[4] In thermodynamics we have only extensive and intensive variables (and not, say, surface-dependent terms $\propto N^{2/3}$) because we take thermodynamic limit $N \to \infty$, $V \to \infty$ keeping $N/V$ finite.

These relations are called the *equations of state* and they serve as *definitions* for temperature $T$, pressure $P$ and chemical potential $\mu$ while the respective extensive variables are $S, V, N$. From (2) we write

$$dE = \delta Q - \delta W = TdS - PdV + \mu dN \ . \tag{3}$$

Entropy is thus responsible for hidden degrees of freedom (i.e. heat) while other extensive parameters describe macroscopic degrees of freedom. The derivatives (2) are defined only in equilibrium. Therefore, $\delta Q = TdS$ and $\delta W = PdV - \mu dN$ for *quasi-static processes* i.e such that the system is close to equilibrium at every point of the process. A process can be considered quasi-static if its typical time of change is larger than the relaxation times (which for pressure can be estimates as $L/c$, for temperature as $L^2/\kappa$, where $L$ is a system size, $c$ - sound velocity and $\kappa$ thermal conductivity). Finite deviations from equilibrium make $dS > \delta Q/T$ because entropy can increase without heat transfer.

Let us give an example how the entropy maximum principle solves the basic problem. Consider two simple systems separated by a rigid wall which is impermeable for anything but heat. The whole composite system is closed that is $E_1 + E_2$ =const. The entropy change under the energy exchange,

$$dS = \frac{\partial S_1}{\partial E_1} dE_1 + \frac{\partial S_2}{\partial E_2} dE_2 = \frac{dE_1}{T_1} + \frac{dE_2}{T_2} = \left(\frac{1}{T_1} - \frac{1}{T_2}\right) dE_1 \, ,$$

must be positive which means that energy flows from the hot subsystem to the cold one ($T_1 > T_2 \Rightarrow \Delta E_1 < 0$). We see that our definition (2) is in agreement with our intuitive notion of temperature. When equilibrium is reached, $dS = 0$ which requires $T_1 = T_2$. If fundamental relation is known, then so is the function $T(E, V)$. Two equations, $T(E_1, V_1) = T(E_2, V_2)$ and $E_1 + E_2$ =const completely determine $E_1$ and $E_2$. In the same way one can consider movable wall and get $P_1 = P_2$ in equilibrium. If the wall allows for particle penetration we get $\mu_1 = \mu_2$ in equilibrium.

Both energy and entropy are homogeneous first-order functions of its variables: $S(\lambda E, \lambda V, \lambda N) = \lambda S(E, V, N)$ and $E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N)$ (here $V$ and $N$ stand for the whole set of extensive macroscopic parameters). Differentiating the second identity with respect to $\lambda$ and taking it at $\lambda = 1$ one gets the Euler equation

$$E = TS - PV + \mu N \ . \tag{4}$$

Let us show that there are only two independent parameters for a simple one-component system, so that chemical potential $\mu$, for instance, can be found as a function of $T$ and $P$. Indeed, differentiating (4) and comparing with (3) one gets the so-called Gibbs-Duhem relation (in the energy representation) $Nd\mu = -SdT + VdP$ or for quantities per mole, $s = S/N$ and $v = V/N$: $d\mu = -sdT + vdP$. In other words, one can choose $\lambda = 1/N$ and use first-order homogeneity to get rid of $N$ variable, for instance, $E(S, V, N) = NE(s, v, 1) = Ne(s, v)$. In the entropy representation,

$$S = E\frac{1}{T} + V\frac{P}{T} - N\frac{\mu}{T} ,$$

the Gibbs-Duhem relation is again states that because $dS = (dE + PdV - \mu dN)/T$ then the sum of products of the extensive parameters and the differentials of the corresponding intensive parameters vanish:

$$Ed(1/T) + Vd(P/T) - Nd(\mu/T) = 0 \; . \tag{5}$$

One uses $\mu(P, T)$, for instance, when considering systems in the external field. One then adds the potential energy (per particle) $u(\mathbf{r})$ to the chemical potential so that the equilibrium condition is $\mu(P, T) + u(\mathbf{r})$ =const. Particularly, in the gravity field $u(\mathbf{r}) = mgz$ and differentiating $\mu(P, T)$ under $T = const$ one gets $vdP = -mgdz$. Introducing density $\rho = m/v$ one gets the well-known hydrostatic formula $P = P_0 - \rho gz$. For composite systems, the number of independent intensive parameters (thermodynamic degrees of freedom) is the number of components plus one. For example, for a mixture of gases, we need to specify the concentration of every gas plus temperature, which is common for all.

**Processes**. While thermodynamics is fundamentally about states it is also used for describing processes that connect states. Particularly important questions concern performance of engines and heaters/coolers. Heat engine works by delivering heat from a reservoir with some higher $T_1$ via some system to another reservoir with $T_2$ doing some work in the process[5]. If the entropy of the hot reservoir decreases by some $\Delta S_1$ then the entropy of the cold one must increase by some $\Delta S_2 \geq \Delta S_1$. The work $\Delta W$ is the difference between the heat given by the hot reservoir $\Delta Q_1 = T_1 \Delta S_1$ and the heat absorbed by

---

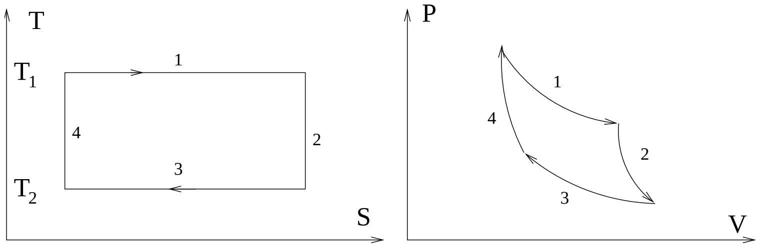
[5] Think about how any real internal combustion engine works to appreciate the level of idealization achieved in distillation of that definition

the cold one $\Delta Q_2 = T_2 \Delta S_2$ (assuming both processes quasi-static). Engine efficiency is the fraction of heat used for work that is

$$\frac{\Delta W}{\Delta Q_1} = \frac{\Delta Q_1 - \Delta Q_2}{\Delta Q_1} = 1 - \frac{T_2 \Delta S_2}{T_1 \Delta S_1} \leq 1 - \frac{T_2}{T_1} \ .$$

It is clear that maximal work is achieved for minimal entropy change $\Delta S_2 = \Delta S_1$, which happens for reversible (quasi-static) processes — if, for instance, the system is a gas which works by moving a piston then the pressure of the gas and the work are less for a fast-moving piston than in equilibrium. Similarly, refrigerator/heater is something that does work to transfer heat from cold to hot systems. The performance is characterized by the ratio of transferred heat to the work done. For the cooler, the efficiency is $\Delta Q_2/\Delta W \leq T_2/(T_1 - T_2)$, for the heater it is $\Delta Q_1/\Delta W \leq T_1/(T_1 - T_2)$. When the temperatures are close, the efficiency is large, as it requires almost no work to transfer heat.

A specific procedure to accomplish reversible heat and work transfer is to use an auxiliary system which undergoes so-called Carnot cycle, where heat exchanges take place only at two temperatures. Engine goes through: 1) isothermal expansion at $T_1$, 2) adiabatic expansion until temperature falls to $T_2$, 3) isothermal compression until the entropy returns to its initial value, 4) adiabatic compression until the temperature reaches $T_1$. The auxiliary system is connected to the reservoirs during isothermal stages: to the first reservoir during 1 and to the second reservoir during 3. During all the time it is connected to our system on which it does work during 1 and 2, increasing the energy of our system, which then decreases its energy by working on the auxiliary system during 3 and 4. The total work is the area of the rectangle between the lines 1,3, the heat $\Delta Q_1$ is the area below the line 1. For heat transfer, one reverses the direction.



Carnot cycle in T-S and P-V variables

Carnot cycle provides one with an operational method to measure the ratio of two temperatures by measuring the engine efficiency[6].

**Summary of formal structure**. The fundamental relation (in energy representation) $E = E(S, V, N)$ is equivalent to the three equations of state (2). If only two equations of state are given then Gibbs-Duhem relation may be integrated to obtain the third up to an integration constant; alternatively one may integrate molar relation $de = Tds - Pdv$ to get $e(s, v)$ again with an undetermined constant of integration.

Example: consider an ideal monatomic gas characterized by two equations of state (found, say, experimentally with $R \simeq 8.3\,\mathrm{J/mole\,K} \simeq 2\,\mathrm{cal/mole\,K}$ ):

$$PV = NRT \ , \qquad E = 3NRT/2 \ . \tag{6}$$

The extensive parameters here are $E, V, N$ so we want to find the fundamental equation in the entropy representation, $S(E, V, N)$. We write (4) in the form

$$S = E\frac{1}{T} + V\frac{P}{T} - N\frac{\mu}{T} \ . \tag{7}$$

Here we need to express intensive variables $1/T, P/T, \mu/T$ via extensive variables. The equations of state (6) give us two of them:

$$\frac{P}{T} = \frac{NR}{V} = \frac{R}{v} \ , \qquad \frac{1}{T} = \frac{3NR}{2E} = \frac{3R}{e} \ . \tag{8}$$

Now we need to find $\mu/T$ as a function of $e, v$ using Gibbs-Duhem relation in the entropy representation (5). Using the expression of intensive via extensive variables in the equations of state (8), we compute $d(1/T) = -3Rde/2e^2$ and $d(P/T) = -Rdv/v^2$, and substitute into (5):

$$\begin{aligned} d\left(\frac{\mu}{T}\right) &= -\frac{3}{2}\frac{R}{e}de - \frac{R}{v}dv \, , \quad \frac{\mu}{T} = C - \frac{3R}{2}\ln e - R\ln v \, , \\ s &= \frac{1}{T}e + \frac{P}{T}v - \frac{\mu}{T} = s_0 + \frac{3R}{2}\ln\frac{e}{e_0} + R\ln\frac{v}{v_0} \ . \end{aligned} \tag{9}$$

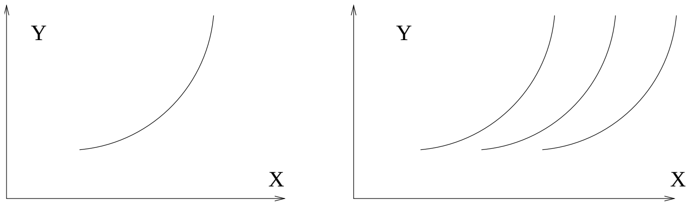
Here $e_0, v_0$ are parameters of the state of zero internal energy used to determine the temperature units, and $s_0$ is the constant of integration.

---

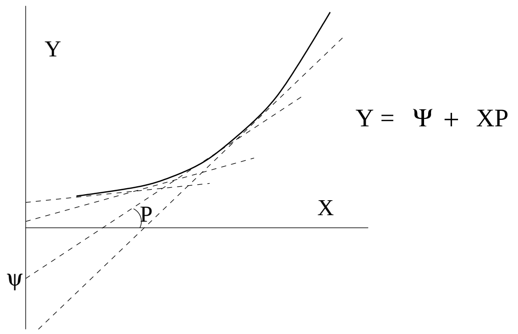
[6] Practical needs to estimate the engine efficiency during the industrial revolution led to the development of such abstract concepts as entropy

## 1.2 Legendre transform

Let us emphasize that the fundamental relation always relates extensive quantities. Therefore, even though it is always possible to eliminate, say, $S$ from $E = E(S, V, N)$ and $T = T(S, V, N)$ getting $E = E(T, V, N)$, this *is not* a fundamental relation and it does not contain all the information. Indeed, $E = E(T, V, N)$ is actually a partial differential equation (because $T = \partial E/\partial S$) and even if it can be integrated the result would contain undetermined function. Still, it is easier to measure, say, temperature than entropy so it is convenient to have a complete formalism with intensive parameters as operationally independent variables and extensive parameters as derived quantities. This is achieved by the Legendre transform: To pass from the relation $Y = Y(X)$ to that in terms of $P = \partial Y/\partial X$ it is not enough to eliminate $X$ and consider the function $Y = Y(P)$, which determines the curve $Y = Y(X)$ only up to a shift along $X$:



For example, the same $Y = P^2/4$ correspond to the family of functions $Y = (X + C)^2$ for arbitrary $C$. To fix the shift one may consider the curve as the envelope of the family of the tangent lines characterized by the slope $P$ and the position $\psi$ of intercept of the $Y$-axis. The function $\psi(P) = Y[X(P)] - PX(P)$ completely defines the curve; here one substitutes $X(P)$ found from $P = \partial Y(X)/\partial X$ (which is possible only when $\partial P/\partial X = \partial^2 Y/\partial X^2 \neq 0$). The function $\psi(P)$ is referred to as a Legendre transform of $Y(X)$. From $d\psi = -PdX - XdP + dY = -XdP$ one gets $-X = \partial\psi/\partial P$ i.e. the inverse transform is the same up to a sign: $Y = \psi + XP$. In mechanics, we use the Legendre transform to pass from Lagrangian to Hamiltonian description.



**Different thermodynamics potentials** suitable for different physical situations are obtained replacing different extensive parameters by the respective intensive parameters.

Free energy $F = E - TS$ (also called Helmholtz potential) is that partial Legendre transform of $E$ which replaces the entropy by the temperature as an independent variable: $dF(T, V, N, \ldots) = -SdT - PdV + \mu dN + \ldots$. It is particularly convenient for the description of a system in a thermal contact with a heat reservoir because then the temperature is fixed and we have one variable less to care about. The maximal work that can be done under a constant temperature (equal to that of the reservoir) is minus the differential of the free energy. Indeed, this is the work done *by the system and the thermal reservoir*. That work is equal to the change of the total energy

$$d(E + E_r) = dE + T_r dS_r = dE - T_r dS = d(E - T_r S) = d(E - TS) = dF \ .$$

In other words, the free energy $F = E - TS$ is that part of the internal energy which is *free* to turn into work, the rest of the energy $TS$ we must keep to sustain a constant temperature. The equilibrium state minimizes $F$, not absolutely, but over the manifold of states with the temperature equal to that of the reservoir. Indeed, consider $F(T, X) = E[S(T, X), X] - TS(T, X)$, then $(\partial E/\partial X)_S = (\partial F/\partial X)_T$ that is they turn into zero simultaneously. Also, in the point of extremum, one gets $(\partial^2 E/\partial X^2)_S = (\partial^2 F/\partial X^2)_T$ i.e. both $E$ and $F$ are minimal in equilibrium. Monatomic gas at fixed $T, N$ has $F(V) = E - TS(V) = -NRT \ln V +$const. If a piston separates equal amounts then the work done in changing the volume of a subsystem from $V_1$ to $V_2$ is $\Delta F = NRT \ln[V_2(V - V_2)/V_1(V - V_1)]$.

Enthalpy $H = E + PV$ is that partial Legendre transform of $E$ which replaces the volume by the pressure $dH(S, P, N, \ldots) = TdS + VdP + \mu dN + \ldots$. It is particularly convenient for situation in which the pressure is maintained constant by a pressure reservoir (say, when the vessel is open into atmosphere). Just as the energy acts as a potential at constant entropy and the free energy as potential at constant temperature, so the enthalpy is a potential for the work done *by the system and the pressure reservoir* at constant pressure. Indeed, now the reservoir delivers pressure which can change the volume so that the differential of the total energy is

$$d(E+E_r)=dE-P_rdV_r=dE+P_rdV=d(E+P_rV)=d(E+PV)=dH\ .$$

Equilibrium minimizes $H$ under the constant pressure. On the other hand, the heat received by the system at constant pressure (and N) is the enthalpy change: $\delta Q=dQ=TdS=dH$. Compare it with the fact that the heat received by the system at constant volume (and N) is the energy change since the work is zero.

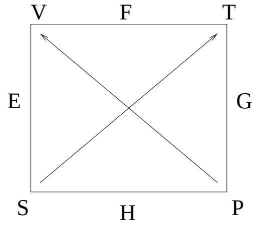
One can replace both entropy and volume obtaining (Gibbs) thermodynamics potential $G=E-TS+PV$ which has $dG(T,P,N,\ldots)=-SdT+VdP+\mu dN+\ldots$ and is minimal in equilibrium at constant temperature and pressure. From (4) we get (remember, they all are functions of different variables):

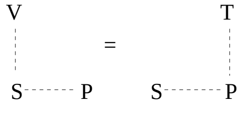
$$F=-P(T,V)V+\mu(T,V)N\ ,\quad H=TS+\mu N\ ,\quad G=\mu(T,P)N\ . \tag{10}$$

When there is a possibility of change in the number of particles (because our system is in contact with some particle source having a fixed chemical potential) then it is convenient to use the grand canonical potential $\Omega(T,V,\mu)=E-TS-\mu N$ which has $d\Omega=-SdT-PdV-Nd\mu$. The grand canonical potential reaches its minimum under the constant temperature and chemical potential.

Since the Legendre transform is invertible, all potentials are equivalent and contain the same information. The choice of the potential for a given physical situation is that of convenience: we usually take what is fixed as a variable to diminish the number of effective variables.

**Maxwell relations**. Changing order of taking mixed second derivatives of a potential creates a class of identities known as Maxwell relations. For example, $\partial^2E/\partial S\partial V=\partial^2E/\partial V\partial S$ gives $(\partial P/\partial S)_V=-(\partial T/\partial V)_S$. That can be done for all three combinations $(SV,SN,VN)$ possible for a simple single-component system and also for every other potential $(F,H,G)$. Maxwell relations for constant $N$ can be remembered with the help of the mnemonic diagram with the sides labelled by the four common potentials flanked by their respective natural independent variables. In the differential expression for each potential in terms of the natural variables arrow pointing away from the variable implies a positive sign while pointing towards the variable implies negative sign like in $dE = TdS - PdV$:





Maxwell relations are given by the corners of the diagram, for example, $(\partial V/\partial S)_P = (\partial T/\partial P)_S$ etc. If we consider constant $N$ then any fundamental relation of a single-component system is a function of only two variables and therefore have only three independent second derivatives. Traditionally, all derivatives are expressed via the three basic ones (those of Gibbs potential), the specific heat and the coefficient of thermal expansion, both at a constant pressure, and isothermal compressibility:

$$c_P = T\left(\frac{\partial S}{\partial T}\right)_P = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_P , \ \alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P , \ \kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T .$$

In particular, the specific heat at constant volume is as follows:

$$c_V = T\left(\frac{\partial S}{\partial T}\right)_V = c_P - \frac{TV\alpha^2}{N\kappa_T} . \tag{11}$$

That and similar formulas form a technical core of thermodynamics and the art of deriving them ought to be mastered. It involves few simple rules in treating partial derivatives:

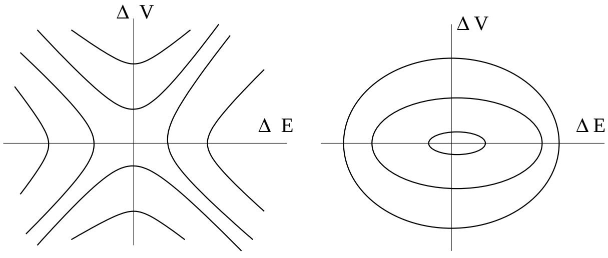
$$\left(\frac{\partial X}{\partial Y}\right)_Z = \left(\frac{\partial Y}{\partial X}\right)_Z^{-1}, \ \left(\frac{\partial X}{\partial Y}\right)_Z\left(\frac{\partial Y}{\partial W}\right)_Z = \left(\frac{\partial X}{\partial W}\right)_Z, \ \left(\frac{\partial X}{\partial Y}\right)_Z\left(\frac{\partial Y}{\partial Z}\right)_X\left(\frac{\partial Z}{\partial X}\right)_Y = -1.$$

An alternative (and more general) way to manipulate thermodynamic derivatives is to use jacobians and identity $\partial(T,S)/\partial(P,V) = 1$. Taking, say, $S, V$ as independent variables,

$$\frac{\partial(T,S)}{\partial(P,V)} = \frac{\partial(T,S)}{\partial(S,V)}\frac{\partial(S,V)}{\partial(P,V)} = -\frac{(\partial T/\partial V)_S}{(\partial P/\partial S)_V} = \frac{E_{SV}}{E_{VS}} = 1 .$$

## 1.3 Stability of thermodynamic systems

Consider entropy representation. Stationarity of equilibrium requires $dS = 0$ while stability requires $d^2S < 0$. In particular, that means concavity of $S(E, X)$. Indeed, for all $\Delta E$ one must have $S(E+\Delta E, X)+S(E-\Delta E, X) \leq 2S(E, X)$ otherwise our system can break into two halves with the energies $E \pm \Delta E$ thus increasing total entropy. For $\Delta E \to 0$ the stability requirement means $(\partial^2 S/\partial E^2)_X \leq 0 \Rightarrow (\partial T/\partial E)_X \geq 0$ — increase of the energy must increase temperature. For the case $X = V$ this can be also recast into $(\partial T/\partial E)_V = [\partial(TV)/\partial(EV)][\partial(SV)/\partial(SV)] = T^{-1}(\partial T/\partial S)_V = 1/c_v \geq 0$ (adding heat to a stable system increases temperature). The same concavity requirement is true with respect to changes in other parameters $X$, in particular, $(\partial^2 S/\partial V^2)_E \leq 0 \Rightarrow (\partial P/\partial V)_T \leq 0$ that is isothermal expansion must reduce pressure for the stable system. Considering both changes together we must require $S_{EE}(\Delta E)^2 + 2S_{EV}\Delta E\Delta V + S_{VV}(\Delta V)^2 \leq 0$. This quadratic form, $S_{EE}(\Delta E)^2 + 2S_{EV}\Delta E\Delta V + S_{VV}(\Delta V)^2 = S_{EE}^{-1}(S_{EE}\Delta E + S_{EV}\Delta V)^2 + (S_{VV} - S_{EV}^2 S_{EE}^{-1})(\Delta V)^2$, has a definite sign if the determinant is positive: $S_{EE}S_{VV} - S_{EV}^2 \geq 0$. Manipulating derivatives one can show that this is equivalent to $(\partial P/\partial V)_S \leq 0$. Alternatively, one may consider the energy representation, here stability requires the energy minimum which gives $E_{SS} = T/c_v \geq 0$, $E_{VV} = -(\partial P/\partial V)_S \geq 0$. Considering both variations one can diagonalize $d^2E = E_{SS}(dS)^2 + E_{VV}(dV)^2 + 2E_{SV}dSdV$ by introducing the temperature differential $dT = E_{SS}dS + E_{SV}dV$ so that $2d^2E = E_{SS}^{-1}(dT)^2 + (E_{VV} - E_{SV}^2 E_{SS}^{-1})(dV)^2$. It is thus clear that $E_{VV} - E_{SV}^2 E_{SS}^{-1} = (\partial^2 E/\partial V^2)_T = -(\partial P/\partial V)_T$ and we recover all the same inequalities. Note that the pressure must decrease under both isothermal and adiabatic expansion.

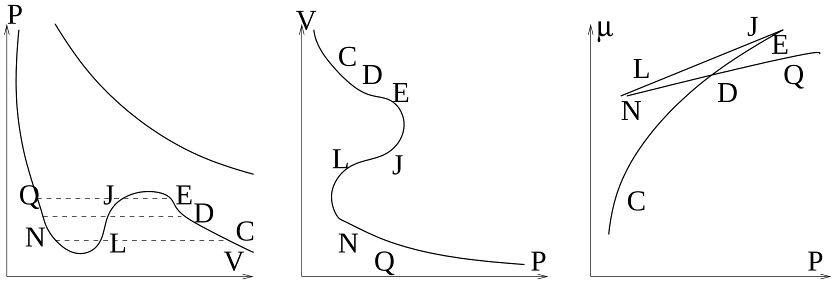


Lines of constant entropy in unstable and stable cases

The physical content of those stability criteria is known as Le Chätelier's principle: if some perturbation deviates the system from a stable equilibrium that induces spontaneous processes that reduce the perturbation.

**The third law of thermodynamics (Nernst theorem)**. It claims that $S \to 0$ as $T \to 0$. A standard argument is that since stability requires the positivity of the specific heat $c_v$ then the energy must monotonously increase with the temperature and zero temperature corresponds to the ground state. If the ground state is non-degenerate (unique) then $S = 0$. The ground can be degenerate yet generally that degeneracy grows slower than exponentially with $N$, then the entropy per particle is zero in the thermodynamic limit. While this argument is correct it is relevant only for temperatures less than the energy difference between the first excited state and the ground state. As such, it has nothing to do with the third law established generally for much higher temperatures and related to the density of states as function of energy (as we shall see in Chapter 3, entropy goes to zero as $T, T^{3/2}, T^3$ for fermions, massive and massless bosons respectively). See Huang (Section 9.4) and L&L (Section 45) for more details.

**Phase transitions** happen when some stability condition is not satisfied like in the region with $(\partial P/\partial V)_T > 0$ as at the lowest isotherm in the below figure. When the pressure corresponds to the level NLC, it is clear that L is an unstable point and cannot be realized. But which stable point is realized, N or C? To get the answer, one must minimize the Gibbs potential since we have $T$ and $P$ fixed. For one mole, it is the chemical potential which can be found integrating the Gibbs-Duhem relation, $d\mu(T,P) = -sdT + vdP$, under the constant temperature: $G = \mu = \int v(P)dP$. The chemical potential increases up to the point (after E) with infinite $dV/dP$. After that we move along the isotherm back having $dP < 0$ so that the integral decreases and then passes through another point of infinite derivative and start to increase again. In other words, the third graph below represents three branches of the function $\mu(P)$ that has its derivative the function $v(P)$ shown in the second graph. It is clear that to the intersection point D correspond to equal areas below and above the horizontal line on the first graph.The pressure that corresponds to this point separates the absolute minimum at the left branch marked Q (solid-like) from that on the right one marked C (liquid-like). The dependence of volume on pressure is discontinuous along the isotherm.


P
Q
J
E
D
N
L
C
V
V
C
D
E
L
J
N
Q
P
μ
J
E
L
Q
D
N
C
P

# 2 Basic statistical physics (brief reminder)

Here we introduce microscopic statistical description in the phase space and describe three principal ways (microcanonical, canonical and grand canonical) to derive thermodynamics from statistical mechanics.

## 2.1 Distribution in the phase space

We consider macroscopic bodies, systems and subsystems. We define probability for a subsystem to be in some $\Delta p \Delta q$ region of the phase space as the fraction of time it spends there: $w = \lim_{T\to\infty} \Delta t/T$. Assuming that the probability to find the subsystem within the volume $dpdq$ is proportional to this volume, we introduce the statistical distribution in the phase space as density: $dw = \rho(p,q)dpdq$. By definition, the average with the statistical distribution is equivalent to the time average:

$$\bar{f} = \int f(p,q)\rho(p,q)dpdq = \lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)dt \ . \tag{12}$$

The main idea is that $\rho(p,q)$ for a subsystem does not depend on the initial states of this and other subsystems so it can be found without actually solving equations of motion. We define statistical equilibrium as a state where macroscopic quantities equal to the mean values. Assuming short-range forces we conclude that different macroscopic subsystems interact weakly and are statistically independent so that the distribution for a composite system $\rho_{12}$ is factorized: $\rho_{12} = \rho_1\rho_2$.

Now, we take the ensemble of identical systems starting from different points in phase space. In a flow with the velocity $\mathbf{v} = (\dot{p}, \dot{q})$ the density changes according to the continuity equation: $\partial\rho/\partial t + div\,(\rho\mathbf{v}) = 0$. If the motion is considered for not very large time it is conservative and can be described by the Hamiltonian dynamics: $\dot{q}_i = \partial\mathcal{H}/\partial p_i$ and $\dot{p}_i = -\partial\mathcal{H}/\partial q_i$. Here the Hamiltonian generally depends on the momenta and coordinates of the given subsystem and its neighbors. Hamiltonian flow in the phase space is incompressible, it conserves area in each plane $p_i, q_i$ and the total volume: $div\,\mathbf{v} = \partial\dot{q}_i/\partial q_i + \partial\dot{p}_i/\partial p_i = 0$. That gives the Liouville theorem: $d\rho/dt = \partial\rho/\partial t + (\mathbf{v}\cdot\nabla)\rho = -\rho div\,\mathbf{v} = 0$. The statistical distribution is thus conserved along the phase trajectories of any subsystem. As a result, equilibrium $\rho$ is an integral of motion and it must be expressed solely via the integrals of motion. Since $\ln\rho$ is an additive quantity then it must be expressed linearly via the additive integrals of motions which for a general mechanical system are energy $E(p,q)$, momentum $\mathbf{P}(p,q)$ and the momentum of momentum $\mathbf{M}(p,q)$:

$$\ln\rho_a = \alpha_a + \beta E_a(p,q) + \mathbf{c}\cdot\mathbf{P}_a(p,q) + \mathbf{d}\cdot\mathbf{M}(p,q)\ . \tag{13}$$

Here $\alpha_a$ is the normalization constant for a given subsystem while the seven constants $\beta, \mathbf{c}, \mathbf{d}$ are the same for all subsystems (to ensure additivity) and are determined by the values of the seven integrals of motion for the whole system. We thus conclude that the additive integrals of motion is all we need to get the statistical distribution of a closed system (and any subsystem), those integrals replace all the enormous microscopic information. Considering system which neither moves nor rotates we are down to the single integral, energy. For any subsystem (or any system in the contact with thermostat) we get Gibbs' *canonical distribution*

$$\rho(p,q) = A\exp[-\beta E(p,q)]\ . \tag{14}$$

See Landau & Lifshitz, Sects 1-4.

## 2.2 Microcanonical distribution

Consider now a *closed* system with the energy $E_0$. Boltzmann *assumed* that all microstates with the same energy have equal probability (ergodic hypothesis) which gives the *microcanonical distribution*:

$$\rho(p,q) = A\delta[E(p,q) - E_0]\ . \tag{15}$$

Usually one considers the energy fixed with the accuracy $\Delta$ so that the microcanonical distribution is

$$\rho = \begin{cases} 1/\Gamma & \text{for } E \in (E_0, E_0+\Delta) \\ 0 & \text{for } E \notin (E_0, E_0+\Delta)\,, \end{cases} \tag{16}$$

where $\Gamma$ is the volume of the phase space occupied by the system

$$\Gamma(E,V,N,\Delta) = \int_{E<\mathcal{H}<E+\Delta} d^{3N}p\,d^{3N}q\ . \tag{17}$$

For example, for $N$ noninteracting particles (ideal gas) the states with the energy $E = \sum p^2/2m$ are in the $\mathbf{p}$-space near the hyper-sphere with the radius $\sqrt{2mE}$. Remind that the surface area of the hyper-sphere with the radius $R$ in $3N$-dimensional space is $2\pi^{3N/2}R^{3N-1}/(3N/2-1)!$ and we have

$$\Gamma(E,V,N,\Delta) \propto E^{3N/2-1}V^N\Delta/(3N/2-1)! \approx (E/N)^{3N/2}V^N\Delta \ . \tag{18}$$

To link statistical physics with thermodynamics one must define the fundamental relation i.e. a thermodynamic potential as a function of respective variables. It can be done using either canonical or microcanonical distribution. We start from the latter and introduce the entropy as

$$S(E,V,N) = \ln\Gamma(E,V,N) \ . \tag{19}$$

This is one of the most important formulas in physics[7] (on a par with $F = ma\,, E = mc^2$ and $E = \hbar\omega$).

Noninteracting subsystems are statistically independent. That means that the statistical weight of the composite system is a product - indeed, for every state of one subsystem we have all the states of another. If the weight is a product then the entropy is a sum. For interacting subsystems, this is true only for short-range forces in the thermodynamic limit $N \to \infty$. Consider two subsystems, 1 and 2, that can exchange energy. Assume that the indeterminacy in the energy of any subsystem, $\Delta$, is much less than the total energy $E$. Then

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E-E_i) \ . \tag{20}$$

We denote $\bar{E}_1, \bar{E}_2 = E - \bar{E}_1$ the values that correspond to the maximal term in the sum (20). The derivative of it is proportional to $(\partial\Gamma_1/\partial E_i)\Gamma_2 + (\partial\Gamma_2/\partial E_i)\Gamma_1 = (\Gamma_1\Gamma_2)^{-1}[(\partial S_1/\partial E_1)_{\bar{E}_1} - (\partial S_2/\partial E_2)_{\bar{E}_2}]$. Then the extremum condition is evidently $(\partial S_1/\partial E_1)_{\bar{E}_1} = (\partial S_2/\partial E_2)_{\bar{E}_2}$, that is the extremum corresponds to the thermal equilibrium where the temperatures of the subsystems are equal. The equilibrium is thus where the maximum of probability is. It is obvious that $\Gamma(\bar{E}_1)\Gamma(\bar{E}_2) \le \Gamma(E) \le \Gamma(\bar{E}_1)\Gamma(\bar{E}_2)E/\Delta$. If the system consists of $N$ particles and $N_1, N_2 \to \infty$ then $S(E) = S_1(\bar{E}_1) + S_2(\bar{E}_2) + O(logN)$ where the last term is negligible in the thermodynamic limit.

Identification with the thermodynamic entropy can be done considering any system, for instance, an ideal gas. The problem is that the logarithm

---

[7] It is inscribed on the Boltzmann's gravestone.

of (18) contains non-extensive term $N\ln V$. The resolution of this controversy is that we need to treat the particles as indistinguishable, otherwise we need to account for the entropy of mixing different species. We however implicitly assume that mixing different parts of the same gas is a reversible process which presumes that the particles are identical. For identical particles, one needs to divide $\Gamma$ (18) by the number of transmutations $N!$ which makes the resulting entropy of the ideal gas extensive in agreement with (9): $S(E,V,N)=(3N/2)\ln E/N+N\ln eV/N+$const. Note that quantum particles (atoms and molecules) are indeed indistinguishable, which is expressed by a proper symmetrization of the wave function. One can only wonder at the genius of Gibbs who introduced $N!$ long before quantum mechanics (see, L&L 40 or Pathria 1.5 and 6.1). Defining temperature in a usual way, $T^{-1}=\partial S/\partial E=3N/2E$, we get the correct expression $E=3NT/2$. We express here temperature in the energy units. To pass to Kelvin degrees, one transforms $T\to kT$ and $S\to kS$ where the Boltzmann constant $k=1.38\cdot 10^{23}\,J/K$. The value of classical entropy (19) depends on the units. Proper quantitative definition comes from quantum physics with $\Gamma$ being the number of microstates that correspond to a given value of macroscopic parameters. In the quasi-classical limit the number of states is obtained by dividing the phase space into units with $\Delta p\Delta q=2\pi\hbar$.

The same definition (entropy as a logarithm of the number of states) is true for any system with a discrete set of states. For example, consider the set of $N$ two-level systems with levels 0 and $\epsilon$. If energy of the set is $E$ then there are $L=E/\epsilon$ upper levels occupied. The statistical weight is determined by the number of ways one can choose $L$ out of $N$: $\Gamma(N,L)=C_N^L=N!/L!(N-L)!$. We can now define entropy (i.e. find the fundamental relation): $S(E,N)=\ln\Gamma$. Considering $N\gg 1$ and $L\gg 1$ we can use the Stirling formula in the form $d\ln L!/dL=\ln L$ and derive the equation of state (temperature-energy relation),

$$T^{-1}=\partial S/\partial E=\epsilon^{-1}\frac{\partial}{\partial L}\ln\frac{N!}{L!(N-L)!}=\epsilon^{-1}\ln\frac{N-L}{L}\ ,$$

and specific heat $C=dE/dT=N(\epsilon/T)^2 2\cosh^{-2}(\epsilon/T)$. Note that the ratio of the number of particles on the upper level to those on the lower level is $L/(N-L)=\exp(-\epsilon/T)$ (Boltzmann relation).

The derivation of thermodynamic fundamental relation $S(E,\ldots)$ in the microcanonical ensemble is thus via the number of states or phase volume.

## 2.3 Canonical distribution

Let us re-derive the canonical distribution from the microcanonical one which allows us to specify $\beta = 1/T$ in (13,14). Consider a small subsystem or a system in a contact with the thermostat (which can be thought of as consisting of infinitely many copies of our system — this is so-called canonical ensemble, characterized by $N, V, T$). Here our system can have any energy and the question arises what is the probability $W(E)$. Let us find first the probability of the system to be in a given microstate $a$ with the energy $E$. Since all the states of the thermostat are equally likely to occur, then the probability should be directly proportional to the statistical weight of the thermostat $\Gamma_0(E_0 - E)$, where we assume $E \ll E_0$, expand $\Gamma_0(E_0 - E) = \exp[S_0(E_0 - E)] \approx \exp[S_0(E_0) - E/T)]$ and obtain

$$w_a(E) = Z^{-1} \exp(-E/T) \ , \tag{21}$$

$$Z = \sum_a \exp(-E_a/T) \ . \tag{22}$$

Note that there is no trace of the thermostat left except for the temperature. The normalization factor $Z(T, V, N)$ is a sum over all states accessible to the system and is called the partition function.

The probability to have a given energy is the probability of the state (21) times the number of states i.e. the statistical weight of the *subsystem*:

$$W(E) = \Gamma(E) w_a(E) = \Gamma(E) Z^{-1} \exp(-E/T) \ . \tag{23}$$

Here the weight $\Gamma(E)$ grows with $E$ very fast for large $N$. But as $E \to \infty$ the exponent $\exp(-E/T)$ decays faster than any power. As a result, $W(E)$ is concentrated in a very narrow peak and the energy fluctuations around $\bar{E}$ are very small (see Sect. 2.4 below for more details). For example, for an ideal gas $W(E) \propto E^{3N/2} \exp(-E/T)$. Let us stress again that the Gibbs canonical distribution (21) tells that the probability of a given microstate exponentially decays with the energy of the state while (23) tells that the probability of a given energy has a peak.

An alternative and straightforward way to derive the canonical distribution is to use consistently the Gibbs idea of the canonical ensemble as a virtual set, of which the single member is the system under consideration and the energy of the total set is fixed. The probability to have our chosen system in the state $a$ with the energy $E_a$ is then given by the average number of systems $\bar{n}_a$ in this state divided by the total number of systems $N$. Any set of occupation numbers $\{n_a\} = (n_0, n_1, n_2 \ldots)$ satisfies obvious conditions

$$\sum_a n_a = N \ , \qquad \sum_a E_a n_a = E = \epsilon N \ . \tag{24}$$

Any given set is realized in $W\{n_a\} = N!/n_0!n_1!n_2! \ldots$ number of ways and the probability to realize the set is proportional to the respective $W$:

$$\bar{n}_a = \frac{\sum n_a W\{n_a\}}{\sum W\{n_a\}} \ , \tag{25}$$

where summation goes over all the sets that satisfy (24). We assume that in the limit when $N, n_a \to \infty$ the main contribution into (25) is given by the most probable distribution that is maximum of $W$ (we actually look at the maximum of $\ln W$ which is the same yet technically simpler) under the constraints (24). Using the method of Lagrangian multipliers we look for the extremum of $\ln W - \alpha \sum_a n_a - \beta \sum_a E_a n_a$. Using the Stirling formula $\ln n! = n \ln n - n$ we write $\ln W = N \ln N - \sum_a n_a \ln n_a$. We thus need to find the value $n_a^*$ which corresponds to the extremum of $\sum_a n_a \ln n_a - \alpha \sum_a n_a - \beta \sum_a E_a n_a$. Differentiating we obtain: $\ln n_a^* = -\alpha - 1 - \beta E_a$ which gives

$$\frac{n_a^*}{N} = \frac{\exp(-\beta E_a)}{\sum_a \exp(-\beta E_a)} \ . \tag{26}$$

The parameter $\beta$ is given implicitly by the relation

$$\frac{E}{N} = \epsilon = \frac{\sum_a E_a \exp(-\beta E_a)}{\sum_a \exp(-\beta E_a)} \ . \tag{27}$$

Of course, physically $\epsilon(\beta)$ is usually more relevant than $\beta(\epsilon)$. See Pathria, Sect 3.2.

To get thermodynamics from the Gibbs distribution one needs to define the free energy because we are under a constant temperature. This is done via the partition function $Z$ (which is of central importance since macroscopic quantities are generally expressed via the derivatives of it):

$$F(T, V, N) = -T \ln Z(T, V, N) \ . \tag{28}$$

To prove that, differentiate the identity $Z = \exp(-F/T) = \sum_a \exp(-E_a/T)$ with respect to temperature, which gives

$$F = \bar{E} + T \left( \frac{\partial F}{\partial T} \right)_V \ ,$$

equivalent to $F = E - TS$ in thermodynamics.

One can also come to this by defining entropy. Remind that for a closed system we defined $S = \ln \Gamma$ while the probability of state was $w_a = 1/\Gamma$. In other words, the entropy was minus the log of probability. Let us generalize this defining entropy as minus the mean log of probability. For a system in a contact with a thermostat that has a Gibbs distribution we have $\ln w_a$ linear in $E$, so that

$$\begin{aligned} S(\bar{E}) &= - \quad \ln w_a(\bar{E}) = -\langle \ln w_a \rangle = -\sum w_a \ln w_a \qquad (29) \\ &= \quad \sum w_a (E_a/T + \ln Z) = E/T + \ln Z = (E - F)/T \ . \end{aligned}$$

Even though we derived the formula for entropy, $S = -\sum w_a \ln w_a$, for an equilibrium, this definition can be used for any set of probabilities $w_a$, since it provides a useful measure of our ignorance about the system, as we shall see later.

See Landau & Lifshitz (Sects 31,36).

## 2.4 Grand canonical ensemble

Let us now repeat the derivation we did in Sect. 2.3 but in more detail and considering also the fluctuations in the particle number $N$. Consider a subsystem in contact with a particle-energy reservoir. The probability for a subsystem to have $N$ particles and to be in a state $E_{aN}$ can be obtained by expanding the entropy of the reservoir. Let us first do the expansion up to the first-order terms as in (21,22)

$$\begin{aligned} w_{aN} &= A \exp[S(E_0 - E_{aN}, N_0 - N)] = A \exp[S(E_0, N_0) + (\mu N - E_{aN})/T] \\ &= \exp[(\Omega + \mu N - E_{aN})/T] \ . \qquad (30) \end{aligned}$$

Here we used $\partial S/\partial E = 1/T$, $\partial S/\partial N = -\mu/T$ and introduced the grand canonical potential which can be expressed through the grand partition function

$$\Omega(T, V, \mu) = -T \ln \sum_N \exp(\mu N/T) \sum_a \exp(-E_{aN})/T) \ . \qquad (31)$$

That this is equivalent to the thermodynamic definition, $\Omega = \bar{E} - T\bar{S} - \mu\bar{N}$ can be seen calculating the mean entropy similar to (29):

$$\bar{S} = -\sum_{a,N} w_{aN} \ln w_{aN} = (\mu\bar{N} + \Omega - \bar{E})/T \ . \qquad (32)$$

The grand canonical distribution must be equivalent to canonical if one neglects the fluctuations in particle numbers. Indeed, when we put $N = \bar{N}$ the thermodynamic relation gives $\Omega + \mu\bar{N} = F$ so that (30) coincides with the canonical distribution $w_a = \exp[(F - E_a)/T]$.

Generally, there is a natural hierarchy: microcanonical distribution neglects fluctuations in energy and number of particles, canonical distribution neglects fluctuations in $N$ but accounts for fluctuations in $E$, and eventually grand canonical distribution accounts for fluctuations both in $E$ and $N$. The distributions are equivalent only when fluctuations are small. In describing thermodynamics, i.e. mean values, the distributions are equivalent, they just produce different fundamental relations, $S(E, N)$ for microcanonical, $F(T, N)$ for canonical, $\Omega(T, \mu)$ for grand canonical, which are related by the Legendre transform. How operationally one checks, for instance, the equivalence of of canonical and microcanonical energies? One takes an isolated system at a given energy $E$, measures the derivative $\partial E/\partial S$, then puts it into the thermostat with the temperature equal to that $\partial E/\partial S$; the energy now fluctuates but the *mean* energy must be equal to $E$ (as long as system is macroscopic and all the interactions are short-range).

To describe fluctuations one needs to expand the respective thermodynamic potential around the mean value, using the second derivatives $\partial^2 S/\partial E^2$ and $\partial^2 S/\partial N^2$ (which must be negative for stability). That will give Gaussian distributions of $E - \bar{E}$ and $N - \bar{N}$. A straightforward way to find the energy variance $\overline{(E - \bar{E})^2}$ is to differentiate with respect to $\beta$ the identity $\overline{E - \bar{E}} = 0$. For this purpose one can use canonical distribution and get

$$\frac{\partial}{\partial\beta}\sum_a (E_a - \bar{E})e^{\beta(F-E_a)} = \sum_a (E_a - \bar{E})\Big(F + \beta\frac{\partial F}{\partial\beta} - E_a\Big)e^{\beta(F-E_a)} - \frac{\partial\bar{E}}{\partial\beta} = 0\,,$$

$$\overline{(E - \bar{E})^2} = -\frac{\partial\bar{E}}{\partial\beta} = T^2 C_V\ . \qquad (33)$$

Magnitude of fluctuations is determined by the *second* derivative of the respective thermodynamic potential:

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\frac{1}{T} = -\frac{1}{T^2}\frac{\partial T}{\partial E} = -\frac{1}{T^2 C_V}\ .$$

This is natural: the sharper the extremum (the higher the second derivative) the better system parameters are confined to the mean values. Since both $\bar{E}$ and $C_V$ are proportional to $N$ then the relative fluctuations are small indeed:

$\overline{(E-\bar{E})^2}/\bar{E}^2 \propto N^{-1}$. Note that any extensive quantity $f = \sum_{i=1}^{N} f_i$ which is a sum over independent subsystems (i.e. $\overline{f_i f_k} = \bar{f}_i \bar{f}_k$) have a small relative fluctuation:

$$\frac{(\overline{f^2} - \bar{f}^2)}{\bar{f}^2} = \frac{\sum(\overline{f_i^2} - \bar{f}_i^2)}{(\sum f_i)^2} \propto \frac{1}{N} \ .$$

Let us now discuss the fluctuations of particle number. One gets the probability to have $N$ particles by summing (30) over $a$:

$$W(N) \propto \exp\{\beta[\mu(T,V)N - F(T,V,N)]\}$$

where $F(T,V,N)$ is the free energy calculated from the canonical distribution for $N$ particles in volume $V$ and temperature $T$. The mean value $\bar{N}$ is determined by the extremum of probability: $(\partial F/\partial N)_{\bar{N}} = \mu$. The second derivative determines the width of the distribution over $N$ that is the variance:

$$\overline{(N-\bar{N})^2} = 2T\left(\frac{\partial^2 F}{\partial N^2}\right)^{-1} = -2TNv^{-2}\left(\frac{\partial P}{\partial v}\right)^{-1} \propto N \ . \qquad (34)$$

Here we used the fact that $F(T,V,N) = Nf(T,v)$ with $v = V/N$ so that $P = (\partial F/\partial V)_N = \partial f/\partial v$, and substituted the derivatives calculated at fixed $V$: $(\partial F/\partial N)_V = f(v) - v\partial f/\partial v$ and $(\partial^2 F/\partial N^2)_V = N^{-1}v^2\partial^2 f/\partial v^2 = -N^{-1}v^2\partial P(v)/\partial v$. As we discussed in Thermodynamics, $\partial P(v)/\partial v < 0$ for stability. We see that generally the fluctuations are small unless the isothermal compressibility is close to zero which happens at the first-order phase transitions. Particle number (and density) strongly fluctuate in such systems which contain different phases of different densities. This is why one uses grand canonical ensemble in such cases.

Let us repeat this important distinction: all thermodynamics potential are equivalent but respective statistical distributions are not. System that can exchange energy and particles with a thermostat has its extensive parameters $E$ and $N$ fluctuating and the grand canonical distribution describes those fluctuations. The choice of description is dictated only by convenience in thermodynamics. But in statistical physics, if we want to describe the whole statistics of the system in thermostat, we need to use canonical distribution, not the micro-canonical one. That does not mean that one cannot learn everything about the system by considering it isolated (micro-canonically). Indeed, we can determine $C_V$ (and other second derivatives) for an isolated system and then will know the mean squared fluctuation of energy when we bring the system into a contact with a thermostat.

See also Landau & Lifshitz 35 and Huang 8.3-5.

## 2.5 Two simple examples

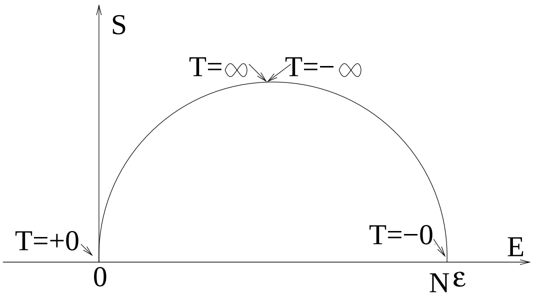
We have seen that the central element of statistical physics is counting the states. Here we consider two examples with the simplest structures of energy levels to illustrate the use of microcanonical and canonical distributions.

### 2.5.1 Two-level system

Assume levels 0 and $\epsilon$. Remind that in Sect. 2.2 we already considered two-level system in the microcanonical approach calculating the number of ways one can distribute $L = E/\epsilon$ portions of energy between $N$ particles and obtaining $S(E,N) = \ln C_N^L = \ln[N!/L!(N-L)!] \approx N\ln[N/(N-L)] + L\ln[(N-L)/L]$. The temperature in the microcanonical approach is as follows:

$$T^{-1} = \frac{\partial S}{\partial E} = \epsilon^{-1}(\partial/\partial L)\ln[N!/L!(N-L)!] = \epsilon^{-1}\ln(N-L)/L\ . \tag{35}$$

The entropy as a function of energy is drawn on the Figure:



Indeed, entropy is zero at $E = 0, N\epsilon$ when all the particles are in the same state. The entropy is symmetric about $E = N\epsilon/2$. We see that when $E > N\epsilon/2$ then the population of the higher level is larger than of the lower one (inverse population as in a laser) and the temperature is negative. Negative temperature may happen only in systems with the upper limit of energy levels and simply means that by adding energy beyond some level we actually decrease the entropy i.e. the number of accessible states. That example with negative temperature is to help you to disengage from the everyday notion of temperature and to get used to the physicist idea of temperature as the derivative of energy with respect to entropy.

Available (non-equilibrium) states lie below the $S(E)$ plot, notice that the entropy maximum corresponds to the energy minimum for positive temperatures and to the energy maximum for the negative temperatures part. A glance on the figure also shows that when the system with a negative temperature is brought into contact with the thermostat (having positive temperature) then our system gives away energy (a laser generates and emits light) decreasing the temperature further until it passes through infinity to positive values and eventually reaches the temperature of the thermostat. That is negative temperatures are actually "hotter" than positive. By itself though the system is stable since $\partial^2 S/\partial E^2 = -N/L(N-L)\epsilon^2 < 0$.

Let us stress that there is no volume in $S(E,N)$ that is we consider only subsystem or only part of the degrees of freedom. Indeed, real particles have kinetic energy unbounded from above and can correspond only to positive temperatures [negative temperature and infinite energy give infinite Gibbs factor $\exp(-E/T)$].

Apart from laser, an example of a two-level system is spin 1/2 in the magnetic field $H$. Because the interaction between the spins and atom motions (spin-lattice relaxation) is weak then the spin system for a long time (tens of minutes) keeps its separate temperature and can be considered separately.

External fields are parameters (like volume and chemical potential) that determine the energy levels of the system. They are sometimes called generalized thermodynamic coordinates, and the derivatives of the energy with respect to them are called respective forces. Let us derive the generalized force $M$ that corresponds to the magnetic field and determines the work done under the change of magnetic field: $dE(S,H) = TdS - MdH$. Since the projection of every magnetic moment on the direction of the field can take two values $\pm\mu$ then the magnetic energy of the particle is $\mp\mu H$ and $E = -\mu(N_+ - N_-)H$. The force (the partial derivative of the energy with respect to the field at a fixed entropy) is called magnetization or magnetic moment of the system:

$$M = -\left(\frac{\partial E}{\partial H}\right)_S = \mu(N_+ - N_-) = N\mu\frac{\exp(\mu H/T) - \exp(-\mu H/T)}{\exp(\mu H/T) + \exp(-\mu H/T)} \ . \tag{36}$$

The derivative was taken at constant entropy that is at constant populations $N_+$ and $N_-$. Note that negative temperature for the spin system corresponds to the magnetic moment opposite in the direction to the applied magnetic field. Such states are experimentally prepared by a fast reversal of the magnetic field. We can also define magnetic susceptibility: $\chi(T) = (\partial M/\partial H)_{H=0} = N\mu^2/T$, yet another second derivative that determines the response and fluctuations and will feature prominently in what follows.

At weak fields and positive temperature, $\mu H \ll T$, (36) gives the formula for the so-called Pauli paramagnetism

$$\frac{M}{N\mu} = \frac{\mu H}{T} \ . \tag{37}$$

*Para* means that the majority of moments point in the direction of the external field. This formula shows in particular a remarkable property of the spin system: adiabatic change of magnetic field (which keeps constant $N_+, N_-$ and thus $M$) is equivalent to the change of temperature even though spins do not exchange energy. One can say that under the change of the value of the homogeneous magnetic field the relaxation is instantaneous in the spin system. This property is used in cooling the substances that contain paramagnetic impurities. For the entropy of the spin system to be preserved, one needs to change the field slowly comparatively to the spin-spin relaxation and fast comparatively to the spin-lattice relaxation. The first condition means that one cannot reach negative temperatures by adiabatically reversing magnetic field since the relaxation times of spins grow when field decreases; indeed, negative temperatures must be reached through $T \to \infty$, not zero. In practice, negative temperatures were reached (by Purcell, Pound and Ramsey in 1951) by *fast* reversal of the magnetic field.

To conclude let us treat the two-level system by the canonical approach where we calculate the partition function and the free energy:

$$Z(T,N) = \sum_{L=0}^{N} C_N^L \exp[-L\epsilon/T] = [1+\exp(-\epsilon/T)]^N \ , \tag{38}$$

$$F(T,N) = -T\ln Z = -NT\ln[1+\exp(-\epsilon/T)] \ . \tag{39}$$

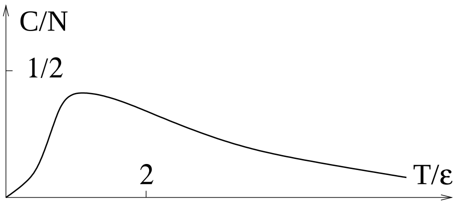
We can now re-derive the entropy as $S = -\partial F/\partial T$ and derive the (mean) energy and specific heat:

$$\bar{E} \ = \ Z^{-1}\sum_a E_a \exp(-\beta E_a) = -\frac{\partial \ln Z}{\partial \beta} = T^2 \frac{\partial \ln Z}{\partial T} \tag{40}$$

$$
\begin{aligned}
&= \frac{N\epsilon}{1+\exp(\epsilon/T)} \,, &(41)\\
C &= \frac{dE}{dT} = \frac{N\exp(\epsilon/T)}{[1+\exp(\epsilon/T)]^2}\frac{\epsilon^2}{T^2} \,. &(42)
\end{aligned}
$$

Here (40) is a general formula which we shall use in the future. Remark that even though canonical approach corresponds to a system in a thermostat, which necessary has positive temperature, all the formulas make sense at negative $T$ too.

Specific heat is one of the second derivatives of the thermodynamic potentials; such quantities characterize the response of the system to change of parameters and will feature prominently in the course. Specific heat tells us how much one raises the energy of the system when increasing the temperature by one degree (or, alternatively, how much energy one needs to increase the temperature by one degree). Specific heat of a two-level system turns into zero both at low temperatures (too small portions of energy are "in circulation") and at high temperatures (occupation numbers of two levels already close to equal so changing temperature does not change energy).



A specific heat of this form characterized by a peak is observed in all systems with an excitation gap.

More details can be found in Kittel, Section 24 and Pathria, Section 3.9.

### 2.5.2 Harmonic oscillators

Small oscillations around the equilibrium positions (say, of atoms in the lattice or in the molecule) can be treated as harmonic and independent. The harmonic oscillator is a particle in the quadratic potential $U(q) = m\omega^2 q^2/2$, it is described by the Hamiltonian

$$
\mathcal{H}(q,p) = \frac{1}{2m}\Big(p^2 + \omega^2 q^2 m^2\Big) \,. \tag{43}
$$

We start from the quasi-classical limit, $\hbar\omega \ll T$, when the single-oscillator partition function is obtained by Gaussian integration:

$$Z_1(T) = (2\pi\hbar)^{-1} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dq \exp(-\mathcal{H}/T) = \frac{T}{\hbar\omega} \ . \tag{44}$$

We can now get the partition function of $N$ independent oscillators as $Z(T,N) = Z_1^N(T) = (T/\hbar\omega)^N$, the free energy $F = NT\ln(\hbar\omega/T)$ and the mean energy from (40): $E = NT$ — this is an example of the equipartition (every oscillator has two degrees of freedom with $T/2$ energy for each)[8]. The thermodynamic equations of state are $\mu(T) = \partial F/\partial N = T\ln(\hbar\omega/T)$ and $S = N[\ln(T/\hbar\omega) + 1]$ while the pressure is zero because there is no volume dependence. The specific heat $C_P = C_V = N$.

Apart from thermodynamic quantities one can write the probability distribution of coordinate of the particle with a finite temperature (i.e. in contact with the thermostat). The distribution is given by the Gibbs distribution using the potential energy:

$$dw_q = \sqrt{m\omega^2/2\pi T}\exp(-m\omega^2 q^2/2T)dq \ . \tag{45}$$

Using kinetic energy and simply replacing $q \to p/m\omega$ one obtains a similar formula $dw_p = (2\pi mT)^{-1/2}\exp(-p^2/2mT)dp$ which is the Maxwell distribution.

For a quantum case, the energy levels are given by $E_n = \hbar\omega(n + 1/2)$. The single-oscillator partition function

$$Z_1(T) = \sum_{n=0}^{\infty} \exp[-\hbar\omega(n + 1/2)/T] = \frac{1}{2\sinh(\hbar\omega/2T)} \tag{46}$$

gives again $Z(T,N) = Z_1^N(T)$ and $F(T,N) = NT\ln[2\sinh(\hbar\omega/2T)] = N\hbar\omega/2 + NT\ln[1 - \exp(-\hbar\omega/T)$. The chemical potential,

$$\mu(T) = T\ln[\sinh(\hbar\omega/2T)/2] \,,$$

is negative in the classical region $T \gg \hbar\omega$ and is positive in the quantum region $T \ll \hbar\omega$.

The energy of the quantum oscillator is

$$E = N\hbar\omega/2 + N\hbar\omega[\exp(\hbar\omega/T) - 1]^{-1}$$

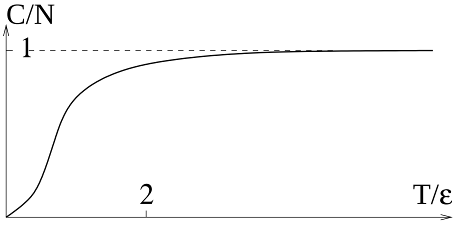
---

[8] If some variable $x$ enters energy as $x^{2n}$ then the mean energy associated with that degree of freedom is $\int x^{2n}\exp(-x^{2n}/T)dx/\int\exp(-x^{2n}/T)dx = T2^{-n}(2n-1)!!$.

where one sees the contribution of zero quantum oscillations and the breakdown of classical equipartition. The specific heat is as follows:

$$C_P = C_V = N(\hbar\omega/T)^2 \exp(\hbar\omega/T)[\exp(\hbar\omega/T) - 1]^{-2} \ . \tag{47}$$

Note that zero oscillations do not contribute the specific heat. Comparing (47) with (42) we see the same behavior at $T \ll \hbar\omega$: $C_V \propto \exp(-\hbar\omega/T)$ because "too small energy portions are in circulation" and they cannot move system to the next level. At large $T$ the specific heat of two-level system turns into zero because the occupation numbers of both levels are almost equal while for oscillator we have classical equipartition (every oscillator has two degrees of freedom so it has $T$ in energy and 1 in $C_V$).



Quantum analog of (45) must be obtained by summing the wave functions of quantum oscillator with the respective probabilities:

$$dw_q = a dq \sum_{n=0}^{\infty} |\psi_n(q)|^2 \exp[-\hbar\omega(n + 1/2)/T] \ . \tag{48}$$

Here $a$ is the normalization factor. Straightforward (and beautiful) calculation of (48) can be found in Landau & Lifshitz Sect. 30. Here we note that the distribution must be Gaussian $dw_q \propto \exp(-q^2/2\overline{q^2})$ where the mean-square displacement $\overline{q^2}$ can be read from the expression for energy so that one gets:

$$dw_q = \left(\frac{\omega}{\pi\hbar} \tanh \frac{\hbar\omega}{2T}\right)^{1/2} \exp\left(-q^2 \frac{\omega}{\hbar} \tanh \frac{\hbar\omega}{2T}\right) dq \ . \tag{49}$$

At $\hbar\omega \ll T$ it coincides with (45) while at the opposite (quantum) limit gives $dw_q = (\omega/\pi\hbar)^{1/2} \exp(-q^2\omega/\hbar) dq$ which is a purely quantum formula $|\psi_0|^2$ for the ground state of the oscillator.

See also Pathria Sect. 3.8 for more details.

# 3 Ideal Gases

We now go on to apply a general theory given in the Chapter 2. Here we consider systems with the kinetic energy exceeding the potential energy of inter-particle interactions: $\langle U(\mathbf{r}_1 - \mathbf{r}_2)\rangle \ll \langle mv^2/2\rangle$.

We start from neglecting the potential energy of interaction completely. Note though that molecules in the same state do have quantum interaction so generally one cannot consider particles completely independent. If however we consider all molecules in the same state as a subsystem then such subsystems do not interact. Since the number of particles in a given state $n_a$ is not fixed, we need to describe such set by a grand canonical ensemble (with $V, \mu$ and $T = 1/\beta$ as variables). Using the distribution (30) with $N = n_a$ and $E = n_a\epsilon_a$ one expresses the probability of occupation numbers:

$$w(n_a) = \exp\{\,\beta[\Omega_a + n_a(\mu - \epsilon_a)]\} \ . \tag{50}$$

Consider now a dilute gas, when all $n_a \ll 1$. Then the probability of no particles in the given state is close to unity, $w_0 = \exp(\beta\Omega_a) \approx 1$, and the probability of having one particle and the average number of particles is given by

$$\bar{n}_a = \sum_{n_a} w(n_a) n_a \approx w_1 \approx \exp\left(\frac{\mu - \epsilon_a}{T}\right) \ , \tag{51}$$

which is called Boltzmann distribution. It is the same as (27) for independent systems, only the normalization factor is expressed here via the chemical potential.

## 3.1 Boltzmann (classical) gas

is such that one can also neglect quantum exchange interaction of particles (atoms or molecules) in the same state which requires the occupation numbers of any quantum state to be small, which in turn requires the number of states $Vp^3/\hbar^3$ to be much larger than the number of molecules $N$. Since the typical momentum is $p \simeq \sqrt{mT}$ we get the condition

$$(mT)^{3/2} \gg \hbar^3 n \ . \tag{52}$$

To get the feeling of the order of magnitudes, one can make an estimate with $m = 1.6 \cdot 10^{-24} g$ (proton) and $n = 10^{21} cm^{-3}$ which gives $T \gg \hbar^2 n^{2/3}/km \simeq 0.5K$. Another way to interpret (52) is to say that the mean distance between molecules $n^{-1/3}$ must be much larger than the wavelength $\lambda = (2\pi\hbar^2/mT)^{1/2}$. In this case, one can pass from the distribution over the quantum states to the distribution in the phase space:

$$\bar{n}(p,q) = \exp\left[\frac{\mu - \epsilon(p,q)}{T}\right] . \tag{53}$$

In particular, the distribution over momenta is always quasi-classical for the Boltzmann gas. Indeed, the distance between energy levels is determined by the size of the box, $\Delta E \simeq h^2 m^{-1} V^{-2/3} \ll h^2 m^{-1} (N/V)^{2/3}$ which is much less than temperature according to (52). To put it simply, if the thermal quantum wavelength $h/p \simeq h(mT)^{-1/2}$ is less than the distance between particles it is also less than the size of the box. We conclude that the Boltzmann gas has the Maxwell distribution over momenta. If such is the case even in the external field then $n(q,p) = \exp\{[\mu - \epsilon(p,q)]/T\} = \exp\{[\mu - U(q) - p^2/2m]/T\}$. That gives, in particular, the particle density in space $n(\mathbf{r}) = n_0 \exp[-U(\mathbf{r})/T]$ where $n_0$ is the concentration without field. In the uniform gravity field we get the barometric formula $n(z) = n(0)\exp(-mgz/T)$.

Since now molecules do not interact then we can treat them as members of the Gibbs canonical ensemble (you probably noticed by now that we are consistently looking for ways to divide every new system we consider into independent parts). The partition function of the Boltzmann gas can be obtained from the partition function of a single particle (like we did for two-level system and oscillator) with the only difference that particles are now real and indistinguishable so that we must divide the sum by the number of transmutations:

$$Z = \frac{1}{N!}\left[\sum_a \exp(-\epsilon_a/T)\right]^N .$$

Using the Stirling formula $\ln N! \approx N \ln(N/e)$ we write the free energy

$$F = -NT \ln\left[\frac{e}{N}\sum_a \exp(-\epsilon_a/T)\right] . \tag{54}$$

Since the motion of the particle as a whole is always quasi-classical for the Boltzmann gas, one can single out the kinetic energy: $\epsilon_a = p^2/2m + \epsilon'_a$. If in addition there is no external field (so that $\epsilon'_a$ describes rotation and the internal degrees of freedom of the particle) then one can integrate over $d^3p d^3q/h^3$ and get for the ideal gas:

$$F = -NT \ln\left[\frac{eV}{N}\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\sum_a \exp(-\epsilon'_a/T)\right] . \tag{55}$$

To complete the computation we need to specify the internal structure of the particle. Note though that $\sum_a \exp(-\epsilon'_a/T)$ depends only on temperature so that we can already get the equation of state $P = -\partial F/\partial V = NT/V$.

**Mono-atomic gas**. At the temperatures much less than the distance to the first excited state all the atoms will be in the ground state (we put $\epsilon_0 = 0$). That means that the energies are much less than Rydberg $\varepsilon_0 = e^2/a_B = me^4/\hbar^2 \simeq 4 \cdot 10^{-11} erg$ and the temperatures are less than $\varepsilon_0/k \simeq 3 \cdot 10^5 K$ (otherwise atoms are ionized).

If there is neither orbital angular momentum nor spin ($L = S = 0$ — such are the atoms of noble gases) we get $\sum_a \exp(-\epsilon'_a/T) = 1$ as the ground state is non-degenerate and

$$F = -NT \ln\left[\frac{eV}{N}\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\right] = -NT \ln\frac{eV}{N} - Nc_v T \ln T - N\zeta T\,, \tag{56}$$

$$c_v = 3/2\,, \quad \zeta = \frac{3}{2}\ln\frac{m}{2\pi\hbar^2}\ . \tag{57}$$

Here $\zeta$ is called the chemical constant. Note that for $F = AT + BT\ln T$ the energy is linear $E = F - T\partial F/\partial T = BT$ that is the specific heat, $C_v = B$, is independent of temperature. The formulas thus derived allow one to derive the conditions for the Boltzmann statistics to be applicable which requires $\bar{n}_a \ll 1$. Evidently, it is enough to require $\exp(\mu/T) \ll 1$ where

$$\mu = \frac{E-TS+PV}{N} = \frac{F+PV}{N} = \frac{F+NT}{N} = T\ln\left[\frac{N}{V}\left(\frac{2\pi\hbar^2}{mT}\right)^{3/2}\right] = T\ln(n\lambda^3)\,.$$

Using such $\mu$ we get $(mT)^{3/2} \gg h^3 n$. Note that $\mu < 0$. As $T$ grows, $\mu/T$ is getting very large negative.

If there is a nonzero spin, it can be in $2S+1$ states. Even though all these states have the same energy, they must be counted in the partition function, which adds $\zeta_S = \ln(2S+1)$ to the chemical constant (57). If both $L$ and $S$ are nonzero then the total angular momentum $J$ determines the fine structure of levels $\epsilon_J$. This is the energy of spin-orbital and spin-spin interactions, both relativistic effects, so that the energy can be estimated as $\epsilon_J \simeq \varepsilon_0(v/c)^2 \simeq \varepsilon_0(Z_n e^2/\hbar c)^2$. For not very high nuclei charge $Z_n$, it is generally comparable with the room temperature $\epsilon_J/k \simeq 200\div 300K$. Every such level has a degeneracy $2J+1$ so that the respective partition function

$$z = \sum_J (2J+1)\exp(-\epsilon_J/T)\ .$$

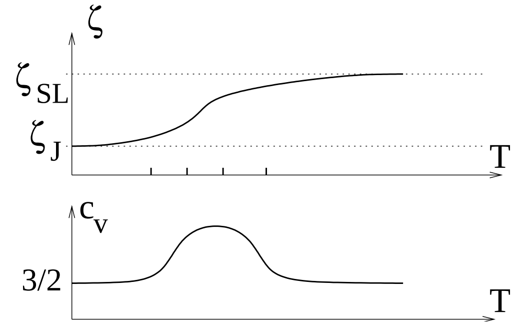
Without actually specifying $\epsilon_J$ we can determine this sum in two limits of large and small temperature. If $\forall J$ one has $T \gg \epsilon_J$, then $\exp(-\epsilon_J/T) \approx 1$ and $z = (2S+1)(2L+1)$ which is the total number of components of the fine level structure. In this case

$$\zeta_{SL} = \ln(2S+1)(2L+1) \ .$$

In the opposite limit of temperature smaller than all the fine structure level differences, only the ground state with $\epsilon_{J_0} = 0$ contributes and one gets

$$\zeta_J = \ln(2J_0 + 1) \ ,$$

where $J_0$ is the total angular momentum in the ground state.



Note that $c_v = 3/2$ in both limits that is the specific heat is constant at low and high temperatures (no contribution of electron degrees of freedom) having some maximum in between (due to contributions of the electrons). We have already seen this in considering two-level system and the lesson is general: if one has a finite number of levels then they do not contribute to the specific heat both at low and high temperatures. Indeed, the partition function $\zeta(T) = \sum_{a=0}^{k} \exp(-\epsilon_a/T)$ is constant both at low at high temperatures: it is equal to the degeneracy of the lowest level at $T \ll \epsilon_0$ and to the total number of states (with their degeneracies) at $T \gg \epsilon_k$.

**Specific heat of diatomic molecules**. We need to calculate the sum over the internal degrees of freedom in (55). We assume the temperature to be smaller than the energy of dissociation (which is typically of the order of electronic excited states). Since most molecules have $S = L = 0$ in the ground state we disregard electronic states in what follows. The internal excitations of the molecule are thus vibrations and rotations with the energy $\epsilon'_a$ characterized by two quantum numbers, $j$ and $K$:

$$\epsilon_{jK} = \hbar\omega(j + 1/2) + \left(\hbar^2/2I\right) K(K+1) \ . \tag{58}$$

Here $\omega$ is the frequency of vibrations and $I$ is the moment of inertia for rotations. We estimate the parameters here assuming the typical scale to be Bohr radius

$a_B = \hbar^2/me^2 \simeq 0.5 \cdot 10^{-8} cm$ and the typical energy to be Rydberg $\varepsilon_0 = e^2/a_B = me^4/\hbar^2 \simeq 4 \cdot 10^{-11} erg$. Note that $m = 9 \cdot 10^{-28} g$ is the electron mass here. Now the frequency of the *atomic* oscillations is given by the ratio of the Coulomb restoring force and the mass of the ion:

$$\omega \simeq \sqrt{\frac{\varepsilon_0}{a_B^2 M}} = \sqrt{\frac{e^2}{a_B^3 M}} \; .$$

Rotational energy is determined by the moment of inertia $I \simeq M a_B^2$. We may thus estimate the typical energies of vibrations and rotations as follows:

$$\hbar\omega \simeq \varepsilon_0 \sqrt{\frac{m}{M}} \, , \qquad \frac{\hbar^2}{I} \simeq \varepsilon_0 \frac{m}{M} \; . \tag{59}$$

Since $m/M \simeq 10^{-4}$ then both energies are much smaller than the energy of dissociation $\simeq \epsilon_0$ and the rotational energy is smaller than the vibrational one so that rotations start to contribute at lower temperatures: $\varepsilon_0/k \simeq 3 \cdot 10^5 K$, $\hbar\omega/k \simeq 3 \cdot 10^3 K$ and $\hbar^2/Ik \simeq 30\,K$.

To calculate the contribution of rotations one ought to calculate the partition function

$$z_{rot} = \sum_K (2K+1) \exp\left(-\frac{\hbar^2 K(K+1)}{2IT}\right) \; . \tag{60}$$

Again, when temperature is much smaller than the distance to the first level, $T \ll \hbar^2/2I$, the specific heat must be exponentially small. Indeed, retaining only two first terms in the sum (60), we get $z_{rot} = 1 + 3\exp(-\hbar^2/IT)$ which gives in the same approximation $F_{rot} = -3NT\exp(-\hbar^2/IT)$ and $c_{rot} = 3(\hbar^2/IT)^2 \exp(-\hbar^2/IT)$. We thus see that at low temperatures diatomic gas behaves an mono-atomic.

At large temperatures, $T \gg \hbar^2/2I$, the terms with large $K$ give the main contribution to the sum (60). They can be treated quasi-classically replacing the sum by the integral:

$$z_{rot} = \int_0^\infty dK (2K+1) \exp\left(-\frac{\hbar^2 K(K+1)}{2IT}\right) = \frac{2IT}{\hbar^2} \; . \tag{61}$$

That gives the constant specific heat $c_{rot} = 1$. The harmonic oscillator was considered in Sect. 2.5.2. In the quasi-classical limit, $\hbar\omega \ll T$, the partition function of $N$ independent oscillators is $Z(T,N) = Z_1^N(T) = (T/\hbar\omega)^N$, the free energy $F = NT\ln(\hbar\omega/T)$ and the mean energy from (40): $E = NT$. The specific heat $C_V = N$.

For a quantum case, the energy levels are given by $E_n = \hbar\omega(n + 1/2)$. The single-oscillator partition function

$$Z_1(T) = \sum_{n=0}^{\infty} \exp[-\hbar\omega(n+1/2)/T] = \frac{1}{2\sinh(\hbar\omega/2T)} \tag{62}$$

gives again $Z(T,N) = Z_1^N(T)$ and $F(T,N) = NT\ln[2\sinh(\hbar\omega/2T)] = N\hbar\omega/2 + NT\ln[1-\exp(-\hbar\omega/T)$. The energy now is

$$E = N\hbar\omega/2 + N\hbar\omega[\exp(\hbar\omega/T)-1]^{-1}$$

where one sees the contribution of zero quantum oscillations and the breakdown of classical equipartition. The specific heat (per molecule) of vibrations is thus as follows: $c_{vib} = (\hbar\omega/T)^2\exp(\hbar\omega/T)[\exp(\hbar\omega/T)-1]^{-2}$. At $T \ll \hbar\omega$: we have $C_V \propto \exp(-\hbar\omega/T)$. At large $T$ we have classical equipartition (every oscillator has two degrees of freedom so it has $T$ in energy and 1 in $C_V$).

The resulting specific heat of the diatomic molecule, $c_v = 3/2 + c_{rot} + c_{vibr}$, is shown on the figure: Note that for $\hbar^2/I < T \ll \hbar\omega$ the specific heat (weakly) decreases because the distance between rotational levels increases with energy so that the level density (which is actually $c_v$) decreases with temperature.

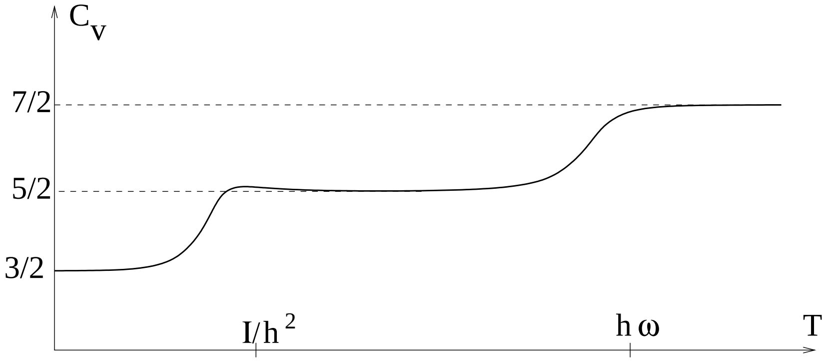


Figure 1: Specific heat of a diatomic molecule as a function of temperature.

For (non-linear) molecules with $\mathcal{N} > 2$ atoms we have 3 translations, 3 rotations and $6\mathcal{N}-6$ vibrational degrees of freedom ($3\mathcal{N}$ momenta and out of total $3\mathcal{N}$ coordinates one subtracts 3 for the motion as a whole and 3 for rotations). That makes for the high-temperature specific heat $c_v = c_{tr} + c_{rot} + c_{vib} = 3/2+3/2+3\mathcal{N}-3 = 3\mathcal{N}$. Indeed, every variable (i.e. every degree of freedom) that enters $\epsilon(p,q)$, which is quadratic in $p,q$, contributes $1/2$ to $c_v$. Translation and rotation each contributes only momentum and thus gives $1/2$ while each vibration contributes both momentum and coordinate (i.e. kinetic and potential energy) and gives 1.

Landau & Lifshitz, Sects. 47, 49, 51.

## 3.2 Fermi and Bose gases

Like we did at the beginning of the Section 3 we consider all particles at the same quantum state as Gibbs subsystem and apply the grand canonical distribution with the potential

$$\Omega_a = -T \ln \sum_{n_a} \exp[n_a(\mu - \epsilon_a)/T] \ . \tag{63}$$

Here the sum is over all possible occupation numbers $n_a$. For fermions, there are only two terms in the sum with $n_a = 0, 1$ so that

$$\Omega_a = -T \ln \{1 + \exp[\beta(\mu - \epsilon_a)]\} \ .$$

For bosons, one must sum the infinite geometric progression (which converges when $\mu < 0$) to get $\Omega_a = T \ln \{1 - \exp[\beta(\mu - \epsilon_a)]\}$. Remind that $\Omega$ depends on $T, V, \mu$. The average number of particles in the state with the energy $\epsilon$ is thus

$$\bar{n}(\epsilon) = -\frac{\partial \Omega_a}{\partial \mu} = \frac{1}{\exp[\beta(\epsilon - \mu)] \pm 1} \ . \tag{64}$$

Upper sign here and in the subsequent formulas corresponds to the Fermi statistics, lower to Bose. Note that at $\exp[\beta(\epsilon - \mu)] \gg 1$ both distributions turn into Boltzmann distribution (51). The thermodynamic potential of the whole system is obtained by summing over the states

$$\Omega = \mp T \sum_a \ln \left[1 \pm e^{\beta(\mu - \epsilon_a)}\right] \ . \tag{65}$$

Fermi and Bose distributions are generally applied to elementary particles (electrons, nucleons or photons) or quasiparticles (phonons) since atomic and molecular gases are described by the Boltzmann distribution (with the exception of ultra-cold atoms in optical traps). For elementary particle, the energy is kinetic energy, $\epsilon = p^2/2m$, which is always quasi-classical (that is the thermal wavelength is always smaller than the size of the box but can now be comparable to the distance between particles). In this case we may pass from summation to the integration over the phase space with the only addition that particles are also distinguished by the direction of the spin $\mathbf{s}$ so there are $g = 2s + 1$ particles in the elementary sell of the phase space. We thus replace (64) by

$$dN(p, q) = \frac{g dp_x dp_y dp_z dx dy dx h^{-3}}{\exp[\beta(\epsilon - \mu)] \pm 1} \ . \tag{66}$$

Integrating over volume we get the quantum analog of the Maxwell distribution:

$$dN(\epsilon) = \frac{gVm^{3/2}}{\sqrt{2}\pi^2\hbar^3} \frac{\sqrt{\epsilon}\, d\epsilon}{\exp[\beta(\epsilon-\mu)] \pm 1} \ . \tag{67}$$

In the same way we rewrite (65):

$$\begin{aligned} \Omega &= \mp \frac{gVTm^{3/2}}{\sqrt{2}\pi^2\hbar^3} \int_0^\infty \sqrt{\epsilon} \ln\left[1 \pm e^{\beta(\mu-\epsilon)}\right] d\epsilon \\ &= -\frac{2}{3} \frac{gVm^{3/2}}{\sqrt{2}\pi^2\hbar^3} \int_0^\infty \frac{\epsilon^{3/2} d\epsilon}{\exp[\beta(\epsilon-\mu)] \pm 1} = -\frac{2}{3} \int_0^\infty \epsilon\, dN(\epsilon) = -\frac{2}{3} E. \end{aligned} \tag{68}$$

Since also $\Omega = -PV$ we get the equation of state

$$PV = \frac{2}{3} E \ . \tag{69}$$

We see that this relation is the same as for a classical gas, it actually is true for any non-interacting particles with $\epsilon = p^2/2m$ in 3-dimensional space. Indeed, consider a cube with the side $l$. Every particle hits a wall $|p_x|/2ml$ times per unit time transferring the momentum $2|p_x|$ in every hit. The pressure is the total momentum transferred per unit time $p_x^2/ml$ divided by the wall area $l^2$ (see Kubo, p. 32):

$$P = \sum_{i=1}^{N} \frac{p_{ix}^2}{ml^3} = \sum_{i=1}^{N} \frac{p_i^2}{3ml^3} = \frac{2E}{3V} \ . \tag{70}$$

So at the same energy all ideal gases have the same pressure. How different are their pressures at the same *temperature*? In the limit of Boltzmann statistics we have $E = 3NT/2$ so that (69) reproduces $PV = NT$. Let us obtain the (small) quantum corrections to the pressure assuming $\exp(\mu/T) \ll 1$. Expanding integral in (68)

$$\int\limits_0^\infty \frac{\epsilon^{3/2} d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1} \approx \int\limits_0^\infty \epsilon^{3/2} e^{\beta(\mu-\epsilon)} \left[1 \mp e^{\beta(\mu-\epsilon)}\right] d\epsilon = \frac{3\sqrt{\pi}}{4\beta^{5/2}} e^{\beta\mu} \left(1 \mp 2^{-5/2} e^{\beta\mu}\right) \ ,$$

and substituting Boltzmann expression for $\mu$ we get

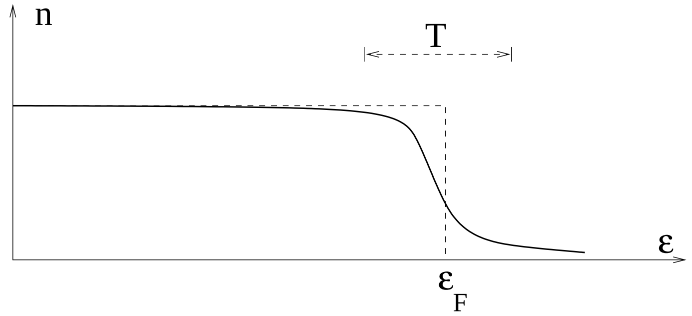
$$PV = NT \left[1 \pm \frac{\pi^{3/2}}{2g} \frac{N}{V} \frac{h^3}{(mT)^{3/2}}\right] \ . \tag{71}$$

Non-surprisingly, the small factor here the ratio of the thermal wavelength cube to the specific volume. We see that quantum effects give some effective attraction between bosons and repulsion between fermions. One should resist temptation to interpret it dynamically assuming that fermions repel each other while bosons attract. Two fermions at different levels has not interaction whatsoever. Of course, this effect is of purely entropic nature, it can be interpreted as follows. As we saw from (69), for a given energy the pressure is the same for bosons and fermions. However, relative to the classical gas, the entropy grows faster with energy for fermions and slower for bosons so that the temperature, $T = \partial E/\partial S$, is lower for fermions and higher for bosons at the same energy. Since $\partial T/\partial E = 1/C > 0$ for stability, then *at a given temperature* the energy is higher for fermions.

Landau & Lifshitz, Sects. 53-56.

### 3.2.1 Degenerate Fermi Gas

The main goal of the theory here is to describe the electrons in the metals (it is also applied to the Thomas-Fermi model of electrons in large atoms, to protons and neutrons in large nucleus, to electrons in white dwarf stars, to neutron stars and early Universe). Drude and Lorents at the beginning of 20th century applied Boltzmann distribution and obtained decent results for conductivity but disastrous discrepancy for the specific heat (which they expected to be 3/2 per electron). That was cleared out by Sommerfeld in 1928 with the help of Fermi-Dirac distribution. The energy of an electron in a metal is comparable to Rydberg and so is the chemical potential (which is positive for degenerate Fermi gas in distinction from Boltzmann and Bose gases, since one increases energy by putting extra particle into the system, see below). Therefore, for most temperatures we may assume $T \ll \mu$ so that the Fermi distribution is close to the step function:



At $T=0$ electrons fill all the momenta up to $p_F$ that can be expressed via the concentration ($g=2$ for $s=1/2$):

$$\frac{N}{V}=2\frac{4\pi}{h^3}\int_0^{p_F}p^2dp=\frac{p_F^3}{3\pi^2\hbar^3}\ , \tag{72}$$

which gives the Fermi energy

$$\epsilon_F=(3\pi^2)^{2/3}\frac{\hbar^2}{2m}\left(\frac{N}{V}\right)^{2/3}\ . \tag{73}$$

Indeed, the Fermi distribution $\bar n(\epsilon)=(\exp[\beta(\epsilon-\mu)]+1)^{-1}$ at $T=0$ turns into a step function $\bar n(\epsilon)=1$ for $\epsilon<\mu$ and $\bar n(\epsilon)=0$ for $\epsilon>\mu$. We thus see that the chemical potential at $T=0$ coincides with the Fermi energy. Putting already one electron per unit cell and taking $N/V\simeq 10^{21}$ one estimates

$$\frac{\epsilon_F}{k}\simeq 10\frac{10^{-54}}{2\cdot 10^{-27}}\frac{10^{14}}{1.3\cdot 10^{-16}}\simeq 10^4K.$$

Condition $T\ll\epsilon_F$ is evidently opposite to (52). Note that the condition of ideality requires that the electrostatic energy $Ze^2/a$ is much less than $\epsilon_F$ where $Ze$ is the charge of ion and $a\simeq(ZV/N)^{1/3}$ is the mean distance between electrons and ions. We see that the condition of ideality, $N/V\gg(e^2m/\hbar^2)^3Z^2$, surprisingly improves with increasing concentration because the kinetic energy $p_F^2/2m\propto n^{2/3}$ grows faster than the potential energy $Ze^2/a\propto n^{1/3}$. Note nevertheless that in most metals the interaction is substantial, why one can still use Fermi distribution (only introducing an effective electron mass) is the subject of Landau theory of Fermi liquids to be described in the course of condensed matter physics (in a nutshell, it is because the main effect of interaction is reduced to some mean effective periodic field).

To obtain the specific heat, $C_v=(\partial E/\partial T)_{V,N}$ one must find $E(T,V,N)$. Since it corresponds to the canonical description, a straightforward way is to find the free energy and then express everything through it. We shall use it in the next subsection. Here we demonstrate another way: exclude $\mu$ from two relations, (67) and (68):

$$N=\frac{2Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\int_0^\infty\frac{\sqrt{\epsilon}d\epsilon}{\exp[\beta(\epsilon-\mu)]+1}\,,$$
$$E=\frac{2Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\int_0^\infty\frac{\epsilon^{3/2}d\epsilon}{\exp[\beta(\epsilon-\mu)]+1}\ .$$

At $T \ll \mu \approx \epsilon_F$ this can be done perturbatively using the formula

$$\int_0^\infty \frac{f(\epsilon)\,d\epsilon}{e^{\beta(\epsilon-\mu)}+1} = \int_0^\mu \frac{f(\mu-z)dz}{e^{-\beta z}+1} + \int_0^\infty \frac{f(\mu+z)dz}{e^{\beta z}+1} = \int_0^\mu f(\epsilon)\,d\epsilon$$
$$+\int_0^\mu \frac{f(\mu-z)dz}{e^{\beta z}+1} + \int_0^\infty \frac{f(\mu+z)dz}{e^{\beta z}+1} \approx \int_0^\mu f(\epsilon)\,d\epsilon + \frac{\pi^2 T^2 f'(\mu}{6}) \ , \qquad (74)$$

which gives

$$N = \frac{2Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\frac{2}{3}\mu^{3/2}\left(1+\pi^2T^2/8\mu^2\right) \ ,$$
$$E = \frac{2Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\frac{2}{5}\mu^{5/2}\left(1+5\pi^2T^2/8\mu^2\right) \ .$$

From the first equation we find $\mu(N,T)$ perturbatively

$$\mu = \epsilon_F\left(1-\pi^2T^2/8\epsilon_F^2\right)^{2/3} \approx \epsilon_F\left(1-\pi^2T^2/12\epsilon_F^2\right) \ .$$

We see that the chemical potential of the Fermi gas decreases with temperature. Since it must be negative at large $T$ when Boltzmann statistics apply, it changes sign at $T \simeq \epsilon_F$. We now substitute $\mu(T)$ into the second equation:

$$E = \frac{3}{5}N\epsilon_F\left(1+5\pi^2T^2/12\epsilon_F^2\right) \ , \qquad (75)$$
$$C_V = \frac{\pi^2}{2}N\frac{T}{\epsilon_F} \ . \qquad (76)$$

We see that $C_V \ll N$ and it goes to zero when $T \to 0$ (as well as entropy) since the number of electrons that can participate in thermal motion decreases $\propto T$. One sees that only fraction $\simeq T/\epsilon_F$ of all electrons participates in thermal exchange. Another important point to stress is that the energy (and $PV$) are much larger than $NT$, the consequence is that the fermionic nature of electrons is what actually determines the resistance of metals (and neutron stars) to compression. For an electron density in metals, $n \simeq 10^{22} cm^{-3}$, we get

$$P \approx \frac{2n\epsilon_F}{5} = (3\pi^2)^{2/3}\frac{\hbar^2}{5m}n^{5/3} \simeq 10^4 atm \ .$$

Landau & Lifshitz, Sects. 57, 58 and Pathria 8.3.

### 3.2.2 Photons

Consider electromagnetic radiation in an empty cavity kept at the temperature $T$. Since electromagnetic waves are linear they do not interact. Without interaction, any distribution is possible. To establish equilibrium distribution, waves must interact with something that has a finite temperature. Thermalization of radiation comes from interaction with walls (absorption and re-emission)[9]. One can derive the equation of state without all the formalism of the partition function. Indeed, consider the plane electromagnetic wave with the fields having amplitudes $\mathbf{E}$ and $\mathbf{B}$. The average energy density is $(E^2+B^2)/2 = E^2$ while the momentum flux modulus is $|\mathbf{E}\times\mathbf{B}| = E^2$. The radiation field in the box can be considered as incoherent superposition of plane wave propagating in all directions. Since all waves contribute the energy density and only one-third of the waves contribute the radiation pressure on any wall then

$$PV = E/3\ . \tag{77}$$

In a quantum consideration we treat electromagnetic waves as photons which are massless particles with the spin 1 that can have only two independent orientations (correspond to two independent polarizations of a classical electromagnetic wave). The energy is related to the momentum by $\epsilon = cp$. Now, exactly as we did for particles [where the law $\epsilon = p^2/2m$ gave $PV = 2E/3$ — see (70)] we can derive (77) considering[10] that every incident photon brings momentum $2p\cos\theta$ to the wall, that the normal velocity is $c\cos\theta$ and integrating $\int \cos^2\theta \sin\theta\, d\theta$. Photon pressure is relevant inside the stars, particularly inside the Sun.

Let us now apply the Bose distribution to the system of photons in a cavity. Since the number of photons is not fixed then a minimum of the free energy, $F(T,V,N)$, requires zero chemical potential: $(\partial F/\partial N)_{T,V} = \mu = 0$. The Bose distribution over the quantum states with fixed polarization,

---

[9] It is meaningless to take perfect mirror walls which do not change the frequency of light under reflection and formally correspond to zero $T$.

[10] This consideration is not restricted to bosons. Indeed, ultra-relativistic fermions have $\epsilon = cp$ and $P = E/3V$, e.g. electrons in graphene. In the relativistic theory, energy and momentum are parts of the energy-momentum tensor whose trace must be positive, which requires $cp \le \epsilon$ and $P \le E/3V$, where $E$ is the total energy including the rest mass $Nmc^2$, L&L 61.

momentum $\hbar\mathbf{k}$ and energy $\epsilon = \hbar\omega = \hbar ck$ is called Planck distribution

$$\bar{n}_{\mathbf{k}} = \frac{1}{e^{\hbar\omega/T} - 1} \ . \tag{78}$$

At $T \gg \hbar\omega$ it gives the Rayleigh-Jeans distribution $\hbar\omega\bar{n}_k = T$ which is classical equipartition. Assuming cavity large we consider the distribution over wave vectors continuous. Multiplying by 2 (the number of polarizations) we get the spectral distribution and the total energy

$$dE_\omega = \hbar ck \frac{2V}{(2\pi)^3} \frac{4\pi k^2 dk}{e^{\hbar ck/T} - 1} = \frac{V\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{e^{\hbar\omega/T} - 1} \ . \tag{79}$$

$$E = \frac{4\sigma}{c} VT^4 \ , \tag{80}$$

where $\sigma = \pi^2/60\hbar^3c^2$ is the Stephan-Boltzmann constant. The specific heat $c_v \propto T^3$ - indeed, the phase volume is $k^3$ and the typical wavenumber $k \propto T$. Since $P = 4\sigma T^4/3c$ depends only on temperature, $c_P$ does not exist (may be considered infinite). We consider fixed temperature so that the relevant potential is $F$ $(= \Omega$ for $\mu = 0)$. It is derived from energy using $S = -\partial F/\partial T$ and $F + TS = F - T\partial F/\partial T = -T^2\partial(F/T)/\partial T = E$, which gives $F(T) = -T\int^T E(T_1)T_1^{-2}dT_1 = -E/3 \propto VT^4$ and entropy $S = -\partial F/\partial T \propto VT^3$ that is the Nernst law is satisfied: $S \to 0$ when $T \to 0$. Under adiabatic compression or expansion of radiation, entropy constancy requires $VT^3 =$ const and $PV^{4/3} =$ const.

A small orifice in the cavity absorbs all the incident light like a black body. Of course, such orifice is only black when illuminated. When we do not shine a light on it, but observe what comes out of such a hole, it is called black-body radiation, yet it is not black but has a color depending on the temperature. Indeed, the distribution (79) has a maximum at $\hbar\omega_m = 2.8T$ (one can estimate the Sun surface temperature by looking at its color). The energy flux from a unit surface of the hole is the energy density times $c$ and times the geometric factor

$$I = \frac{cE}{V}\int_0^{\pi/2} \cos\theta \sin\theta \, d\theta = \frac{c}{4}\frac{E}{V} = \sigma T^4 \ . \tag{81}$$

Landau & Lifshitz, Sect. 63 and Huang, Sect. 12.1.

### 3.2.3 Phonons

The specific heat of a crystal lattice can be calculated using the powerful idea of quasi-particles: turning the set of strongly interacting atoms into a set of weakly interacting waves. In this way one considers the oscillations of the atoms as acoustic waves with three branches (two transversal and one longitudinal) $\omega_i = u_i k$ where $u_i$ is the respective sound velocity. Debye took this expression for the spectrum and imposed a maximal frequency $\omega_{max}$ so that the total number of degrees of freedom is equal to 3 times the number of atoms:

$$\frac{4\pi V}{(2\pi)^3}\sum_{i=1}^{3}\int\limits_0^{\omega_{max}}\frac{\omega^2 d\omega}{u_i^3} = \frac{V\omega_{max}^3}{2\pi^2 u^3} = 3N\ . \tag{82}$$

Here we introduced some effective sound velocity $u$ defined by $3u^{-3} = 2u_t^{-3} + u_l^{-3}$. One usually introduces the Debye temperature

$$\Theta = \hbar\omega_{max} = \hbar u(6\pi^2 N/V)^{1/3} \simeq \hbar u/a\ , \tag{83}$$

where $a$ is the lattice constant.

We can now write the energy of lattice vibrations using the Planck distribution (since the number of phonons is indefinite, $\mu = 0$)

$$E = \frac{3V}{2\pi^2 u^3}\int\limits_0^{\omega_{max}}\hbar\omega\left(\frac{1}{2} + \frac{1}{\exp(\hbar\omega/T)-1}\right)\omega^2 d\omega = \frac{9N\Theta}{8} + 3NTD\left(\frac{\Theta}{T}\right), \tag{84}$$

$$D(x) = \frac{3}{x^3}\int_0^x \frac{z^3 dz}{e^z - 1} = \begin{cases} 1 & \text{for} \quad x \ll 1\,, \\ \pi^4/5x^3 & \text{for} \quad x \gg 1\,. \end{cases}$$

At $T \ll \Theta$ for the specific heat we have the same cubic law as for photons:

$$C = N\frac{12\pi^4}{5}\frac{T^3}{\Theta^3}\ . \tag{85}$$

For liquids, there is only one (longitudinal) branch of phonons so $C = N(4\pi^4/5)(T/\Theta)^3$ which works well for He IV at low temperatures.

At $T \gg \Theta$ we have classical specific heat (Dulong-Petit law) $C = 3N$. Debye temperatures of different solids are between 100 and 1000 degrees Kelvin. We can also write the free energy of the phonons as a sum/integral over frequencies of the single oscillator expression:

$$F = 9NT\left(\frac{T}{\Theta}\right)^3\int\limits_0^{\Theta/T} z^2\ln\left(1 - e^{-z}\right)dz = NT\left[3\ln\left(1 - e^{-\Theta/T}\right) - D(\Theta/T)\right], \tag{86}$$

and find that, again, at low temperatures $S = -\partial F/\partial T \propto T^3$ i.e. Nernst theorem. An interesting quantity is the coefficient of thermal expansion $\alpha = (\partial \ln V/\partial T)_P$. To get it one must pass to the variables $P, T, \mu$ introducing the Gibbs potential $G(P,T) = E - TS + PV$ and replacing $V = \partial G/\partial P$. At high temperatures, $F \approx 3NT \ln(\Theta/T)$. It is the Debye temperature here which depends on $P$, so that the part depending on $T$ and $P$ in both potentials is linearly proportional to $T$: $\delta F(P,T) = \delta G(P,T) = 3NT \ln \Theta$. That makes the mixed derivative

$$\alpha = V^{-1}\frac{\partial^2 G}{\partial P \partial T} = 3\frac{N}{V}\frac{\partial \ln \Theta}{\partial P}$$

independent of temperature. One can also express it via so-called mean geometric frequency defined as follows: $\ln \bar{\omega} = (3N)^{-1}\sum \ln \omega_a$. Then $\delta F = \delta G = T\sum_a \ln(\hbar\omega_a/T) = NT \ln \hbar\bar{\omega}(P)$ , and $\alpha = (N/V\bar{\omega})d\bar{\omega}/dP$. When the pressure increases, the atoms are getting closer, restoring force increases and so does the frequency of oscillations so that $\alpha \geq 0$.

Note that we've got a constant contribution $9N\Theta/8$ in (84) which is due to quantum zero oscillations. While it does not contribute the specific heat, it manifests itself in X-ray scattering, Mössbauer effect etc. Incidentally, this is not the whole energy of a body at zero temperature, this is only the energy of excitations due to atoms shifting from their equilibrium positions. There is also a negative energy of attraction when the atoms are precisely in their equilibrium position. The total (so-called binding) energy is negative for crystal to exists at $T = 0$. One may ask why we didn't account for zero oscillations when considered photons in (79,80). Since the frequency of photons is not restricted from above, the respective contribution seems to be infinite. How to make sense out of such infinities is considered in quantum electrodynamics; note that the zero oscillations of the electromagnetic field are real and manifest themselves, for example, in the Lamb shift of the levels of a hydrogen atom. In thermodynamics, zero oscillations of photons are of no importance.

To conclude, let us reiterate the lesson we learnt about the specific heat which is essentially the number of active degrees of freedom. For solids at room temperatures, that number is much smaller than the either the number of atoms or electrons because only $(T/\Theta)^3$ fraction of atoms and $T/\epsilon_F$ fraction of electrons effectively participate.

Landau & Lifshitz, Sects. 64–66; Huang, Sect. 12.2

### 3.2.4 Bose gas of particles and Bose-Einstein condensation

Consider an ideal Bose gas of massive particles, like atoms [11]. Now, the number of particles is fixed in a closed system. The chemical potential then must not be identically zero, as it was for massless particles, whose number was not conserved. Moment reflection upon $n(\epsilon) = [\exp(\beta(\epsilon-\mu)) - 1]^{-1}$ tells that the chemical potential of massive bosons must be non-positive, otherwise one would have negative or infinite occupation numbers. The chemical potential absolute value is large at high temperatures where it coincides with that of the Boltzmann gas, $\mu = T\ln(n\lambda^3)$, where the thermal wavelength is $\lambda = (2\pi\hbar^2/mT)^{1/2}$. The magnitude apparently decreases when the temperature decreases (at fixed concentration $n = N/V$). The question is whether the chemical potential goes to zero only at zero temperature or it can turn into zero at a finite $T$. At $\mu = 0$, the Bose distribution $n(\epsilon) = [\exp(\beta\epsilon) - 1]^{-1}$ has singularity at $\epsilon = 0$. We didn't care about that considering photons or phonons, since zero-energy-zero-frequency waves are no waves at all. Now, however, massive particles are real even at zero energy. Of course, in a finite box, the lowest kinetic energy is finite, $\hbar^2/2mL^2$, yet it goes to zero at the thermodynamic limit $L \to \infty$. Therefore, having zero chemical potential would mean macroscopic population of the lowest level at the thermodynamic limit. Let us see if it is possible.

As usual, one determines the chemical potential as a function of temperature and the specific volume by equaling the total number of particles to the sum of Bose distribution over all states (the same as finding a thermodynamic potential and differentiating it with respect to $N$). It is more convenient here to work with the function $z = \exp(\mu/T)$ which is called fugacity:

$$N=\sum_{\mathbf{p}}\frac{1}{e^{\beta(\epsilon_p-\mu)}-1}=\frac{4\pi V}{h^3}\int_0^\infty\frac{p^2dp}{z^{-1}e^{p^2/2mT}-1}+\frac{z}{1-z}=\frac{Vg_{3/2}(z)}{\lambda^3}+\frac{z}{1-z}\,. \quad (87)$$

We introduced the function

$$g_a(z) = \frac{1}{\Gamma(a)}\int_0^\infty \frac{x^{a-1}dx}{z^{-1}e^x - 1} = \sum_{i=1}^\infty \frac{z^i}{i^a}\ . \quad (88)$$

We singled out the contribution of zero-energy level even though it is not supposed to contribute at the thermodynamic limit $N \to \infty$. But, as we

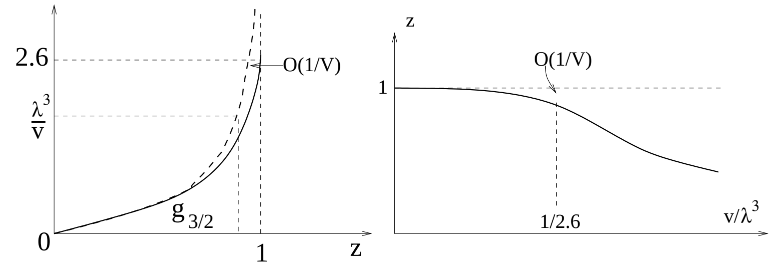
---

[11] The manuscript of Bose, devoted to re-deriving the Planck distribution by the approach based on (50,63), was rejected by the Philosophical Magazine. Bose then sent it to Einstein, who translated it into German, published and used the method to treat atoms.

discussed, thermodynamic limit also means that $V \to \infty$ which allowed us to consider the lowest level to have zero energy. To analyze the behavior of the term $z/(1-z)$ at $\mu \to 0$ and $z \to 1$, let us rewrite (87) denoting $n_0 = z/(1-z)$ the number of particles at $p = 0$

$$\frac{n_0}{V} = \frac{1}{v} - \frac{g_{3/2}(z)}{\lambda^3} \; . \tag{89}$$

The graphic solution of (89) for a finite $V$ can be seen in the Figure below by plotting $g_{3/2}(z)$ (solid line). The function $g_{3/2}(z)$ monotonically grows while $z$ changes from zero ($\mu = -\infty$) to unity ($\mu = 0$). At $z = 1$, the value is $g_{3/2}(1) = \zeta(3/2) \approx 2.6$ and the derivative is infinite (the integral diverges at zero as $\int_0 x^{1/2}dx/x^2$). For a given $v, T$ one knows $\lambda^3/v$ and finds $z$ as a point when the horizontal line $\lambda^3/v$ cross the solid line (at $V \to \infty$) or the broken line (at finite $V$). The distance between the solid and broken lines determines $n_0$. For high temperatures, $\lambda^3/v$ is small and so is $z$. Note, however, that by lowering temperature we can make $\lambda$ arbitrarily large. When the temperature and the specific volume $v = V/N$ are such that $\lambda^3/v > g_{3/2}(1)$ (the thermal wavelength is now larger than the inter-particle distance) then there is a finite fraction of particles that occupies the zero-energy level.



When $V \to \infty$ we have a sharp transition at $\lambda^3/v = g_{3/2}(1)$ i.e. at

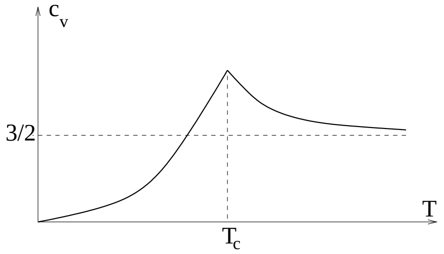
$$T = T_c = 2\pi\hbar^2/m[vg_{3/2}(1)]^{2/3} \; .$$

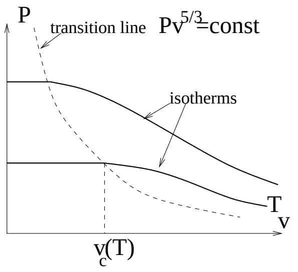
At $T \le T_c$ we have $z \equiv 1$ that is $\mu \equiv 0$. At $T > T_c$ we obtain $z$ solving $\lambda^3/v = g_{3/2}(z)$. Therefore, at the thermodynamic limit we put $n_0 = 0$ at $T > T_c$ and $n_0/N = 1 - vg_{3/2}(1)/\lambda^3 = 1 - (T/T_c)^{3/2}$ as it follows from (89). All thermodynamic relations have now different expressions above and below $T_c$ (upper and lower cases respectively):

$$E = \frac{3}{2}PV = \frac{2\pi V}{mh^3}\int_0^\infty \frac{p^4 dp}{z^{-1}\exp(p^2/2mT) - 1} = \begin{cases} (3VT/2\lambda^3)g_{5/2}(z) \\ (3VT/2\lambda^3)g_{5/2}(1) \end{cases} , \tag{90}$$

$$c_v = \begin{cases} (15v/4\lambda^3)g_{5/2}(z) - 9g_{3/2}(z)/4g_{1/2}(z) \\ (15v/4\lambda^3)g_{5/2}(1) \end{cases} \tag{91}$$

At low $T$, $c_v \propto \lambda^{-3} \propto T^{3/2}$, it decreases faster than $c_v \propto T$ for electrons yet slower than $c_v \propto T^3$ (that we had for $\epsilon_p = cp$). Note that the number of over-condensate particles now changes with $T$ as for phonons and photons, and $\mu = 0$ too. So the specific heat (as the effective number of degrees of freedom) of the ideal Bose gas behaves as the phase volume $p_T^3$; for massless bosons, $p_T \propto T$ and $c_v \propto T^3$ while for massive $p_T \propto T^{1/2}$ and $c_v \propto T^{3/2}$. One can also say that the particle levels, $\epsilon_p = p^2/2m$, are denser at lower energies, that is why the specific heat is larger for massive particles. The other side of the coin is that the same increase of the distance between levels $\epsilon_p = p^2/2m$ with $p$ makes $c_v$ decreasing with $T$ at high temperatures, as for rotators in Sect. 3.1:





The pressure given by the upper line of (90) depends on $V$ via $z$. However, the lower line shows that the pressure is independent of the volume at $T < T_c$. That prompts the analogy with a phase transition of the first order. Indeed, this reminds the properties of the saturated vapor (particles with nonzero energy) in contact with the liquid (particles with zero energy): changing volume at fixed temperature we change the fraction of the particles in the liquid but not the pressure. This is why the phenomenon is called the Bose-Einstein condensation. Increasing temperature we cause evaporation (particle leaving condensate in our case) which increases $c_v$; after all liquid evaporates (at $T = T_c$) $c_v$ starts to decrease. It is sometimes said that it is a "condensation in the momentum space" but if we put the system in a gravity field then there will be a spatial separation of two phases just like in a gas-liquid condensation (liquid at the bottom). On P-V diagram, the transition isotherm $T = T_c$ corresponds to $P = g_{5/2}(1)T_c/\lambda^3(T_c) \propto T_c^{5/2} \propto V^{-5/3}$.

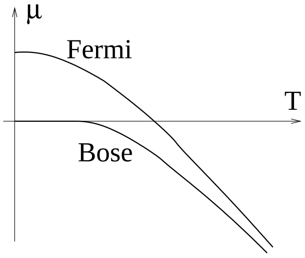
We can also obtain the entropy, above $T_c$ by usual formulas that follow from (68) and below $T_c$ just integrating specific heat $S = \int dE/T = N\int c_v(T)dT/T = 5E/3T = 2Nc_v/3$:

$$\frac{S}{N} = \begin{cases} (5v/2\lambda^3)g_{5/2}(z) - \log(z) \\ (5v/2\lambda^3)g_{5/2}(1) \end{cases} \tag{92}$$

The entropy is zero at $T = 0$ which means that the condensed phase has no entropy - it is a coherent quantum state of a macroscopic number of atoms. At finite $T$ all the entropy is due to gas phase. Below $T_c$ we can write $S/N = (T/T_c)^{3/2}s = (v/v_c)s$ where $s$ is the entropy per gas particle: $s = 5g_{5/2}(1)/2g_{3/2}(1)$. The latent heat of condensation per particle is $Ts$ that it is indeed phase transition of the first order. Usual gas-liquid condensation is caused by the interaction - molecules attract each other at large distances. We see that the Bose-Einstein condensation takes place already in an ideal gas and is due to the discreteness of the energy levels, which indeed leads to some effective attraction as was seen from (71).

The Bose-Einstein condensation has been observed for atoms in electromagnetic traps at low temperatures. Transition to superfluidity in liquid $\mathrm{He}^4$ is of the second order, but then a liquid is not an ideal gas.

To conclude, we have seen in this Section how quantum effects lead to switching off degrees of freedom at low temperatures. Fermi and Bose systems reach the zero-entropy state at $T = 0$ in different ways. It is also instructive to compare their chemical potentials:



We also saw that an effective interaction brought by quantum effect lead to phase transition only for the case of attraction i.e. for bosons.

Landau & Lifshitz, Sect. 62; Huang, Sect. 12.3.

# 4 Interacting systems and phase transitions

Here we take into account a weak interaction between particles. There are two limiting cases when the consideration is simplified:
i) when the interaction is long-range so that every particle effectively interact with many other particles and one can apply some mean-field description, this will be considered in Section 4.1 below;
ii) when the typical range of interaction is much smaller than the mean distance between particles so that it is enough to consider only two-particle interactions, this is considered afterwards.

## 4.1 Coulomb interaction and screening

Here we consider a three-dimensional system of charged particles (plasma or electrolyte), and describe its thermodynamic properties and the correlation functions of the charge-density fluctuations. The Coulomb energy of interaction is $u_{ab} = e^2 z_a z_b / r_{ab}$ where $z_a, z_b$ are the charges and $r_{ab}$ is the distance between two particles.

Do we expect the interaction to change the total energy of the system and thus change thermodynamics in a neutral system with the zero total charge? Indeed, if the ions and electrons were distributed uniformly then the total Coulomb energy of interaction is zero. However, interaction leads to correlations in particle positions (particle prefer to be surrounded by the particles of the opposite charge) which makes for a nonzero (negative) contribution to the energy and other thermodynamic quantities.

Interaction of charged particles is long-range and one may wonder how at all one may use a thermodynamic approach (divide a system into independent subsystems, for instance). The answer is that the preference to be surrounded by opposite charges leads to charge screening, making interaction effectively finite-range. The semi-phenomenological description of such systems has been developed by Debye and Hückel (1923) and it works for plasma and electrolytes. Consider the simplest situation when we have electrons of the charge $-e$ and ions of the charge $+e$.

We start from a rough estimate for the screening radius $r_D$ which we define as that of a sphere around an ion where the total charge of all particles is of order $-e$ i.e. compensates the charge of the ion. Particles are distributed in the field $U(r)$ according to the Boltzmann formula $n(r) = n_0 \exp[-U(r)/T]$

and the estimate is as follows:

$$r_D^3 n_0[\exp(e^2/r_D T) - \exp(-e^2/r_D T)] \simeq 1 \ . \tag{93}$$

We obtain what is called the Debye radius

$$r_D \sim \sqrt{\frac{T}{n_0 e^2}} \tag{94}$$

under the condition of interaction weakness, $e^2/r_D T = (e^2 n_0^{1/3}/T)^{3/2} \ll 1$. Note that under that condition there are many particles inside the Debye sphere: $n_0 r_d^3 \gg 1$ (in electrolytes $r_D$ is of order $10^{-3} \div 10^{-4}\,\mathrm{cm}$ while in ionosphere plasma it can be kilometers). Everywhere $n_0$ is the mean density of either ions or electrons.

We can now estimate the electrostatic contribution to the energy of the system of $N$ particles (what is called correlation energy):

$$\bar{U} \simeq -N\frac{e^2}{r_D} \simeq -\frac{N^{3/2}e^3}{\sqrt{VT}} = -\frac{A}{\sqrt{VT}} \ . \tag{95}$$

The (positive) addition to the specific heat

$$\Delta C_V = \frac{A}{2V^{1/2}T^{3/2}} \simeq N\frac{e^2}{r_D T} \ll N \ . \tag{96}$$

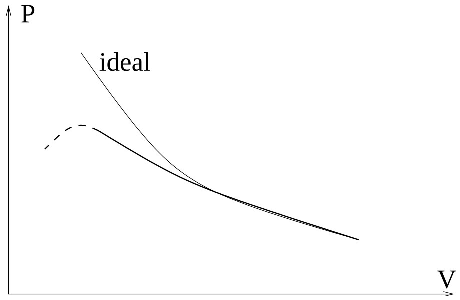
One can get the correction to the entropy by integrating the specific heat:

$$\Delta S = -\int_T^\infty \frac{C_V(T)dT}{T} = -\frac{A}{3V^{1/2}T^{3/2}} \ . \tag{97}$$

We set the limits of integration here as to assure that the effect of screening disappears at large temperatures. We can now get the correction to the free energy and pressure

$$\Delta F = \bar{U} - T\Delta S = -\frac{2A}{3V^{1/2}T^{1/2}} \,, \quad \Delta P = -\frac{A}{3V^{3/2}T^{1/2}} \ . \tag{98}$$

Total pressure is $P = NT/V - A/3V^{3/2}T^{1/2}$ — a decrease at small $V$ (see figure) hints about the possibility of phase transition which indeed happens (droplet creation) for electron-hole plasma in semiconductors even though our calculation does not work at those concentrations.



The correlation between particle positions (around every particle there are more particles of opposite charge) means that attraction prevails over repulsion so that it is necessary that corrections to energy, entropy, free energy and pressure are negative. Positive addition to the specific heat could be interpreted as follows: increasing temperature one decreases screening and thus increases energy.

Now, we can do all the consideration in a more consistent way calculating exactly the value of the constant $A$. To calculate the correlation energy of electrostatic interaction one needs to multiply every charge by the potential created by other charges at its location. In estimates, we took the Coulomb law for the potential around every charge, while it must differ as the distance increases. Indeed, the electrostatic potential $\phi(r)$ around an ion determines the distribution of ions (+) and electrons (-) by the Boltzmann formula $n_\pm(r) = n_0 \exp[\mp e\phi(r)/T]$ while the charge density $e(n_+ - n_-)$ in its turn determines the potential by the Poisson equation:

$$\Delta\phi = -4\pi e(n_+ - n_-) = -4\pi e n_0 \left(e^{-e\phi/T} - e^{e\phi/T}\right) \approx \frac{8\pi e^2 n_0}{T}\phi \ , \qquad (99)$$

where we expanded the exponents assuming the weakness of interaction. We also need to impose the condition that when $r \to 0$ the potential is that of the ion $\phi \to e/r$. Under such a condition the equation (99) has a central-symmetric solution $\phi(r) = (e/r)\exp(-\kappa r)$ where $\kappa^2 = 8\pi r_D^{-2}$. We are interesting in this potential near the ion i.e. at small $r$: $\phi(r) \approx e/r - e\kappa$ where the first term is the field of the ion itself while the second term is precisely what we need i.e. contribution of all other charges. We can now write the energy of every ion and electron as $-e^2\kappa$ and get the total electrostatic energy multiplying by the number of particles ($N = 2n_0V$) and dividing by 2 so as not to count every couple of interacting charges twice:

$$\bar{U} = -n_0 V \kappa e^2 = -\sqrt{\pi}\,\frac{N^{3/2}e^3}{\sqrt{VT}} \ . \qquad (100)$$

Comparing with the rough estimate (95), we just added the factor $\sqrt{\pi}$.

The consideration by Debye-Hückel is the right way to account for the first-order corrections in the small parameter $e^2 n_0^{1/3}/T$. The method replaces interactions by an effective mean field $\phi$, it is thus a variant of a mean-field approach. One cannot though get next corrections within the method [further expanding the exponents in (99)]. That would miss multi-point correlations which contribute the next orders. Indeed, the existence of an ion at some point influences not only the probability of having an electron at some other point but they together influence the probability of charge distribution in the rest of the space. To account for multi-point correlations, one needs Bogolyubov's method of correlation functions. Such functions are multi-point joint probabilities to find simultaneously particles at given places. The correlation energy is expressed via the two-point correlation function $w_{ab}$ where the indices mark both the type of particles (electrons or ions) and the positions $\mathbf{r}_a$ and $\mathbf{r}_b$:

$$E = \frac{1}{2}\sum_{a,b} \frac{N_a N_b}{V^2} \int\int u_{ab} w_{ab} dV_a dV_b \ . \tag{101}$$

Here $u_{ab}$ is the energy of the interaction. The pair correlation function is determined by the Gibbs distribution integrated over the positions of all particles except the given pair:

$$w_{ab} = V^{2-N} Z_i^{-1} \int \exp\left[\frac{-U(r_1 \ldots r_N)}{T}\right] dV_1 \ldots dV_{N-2} \ . \tag{102}$$

Here the normalization factor can be expressed via the difference between total free energy and the free energy of an ideal gas: $Z_i^{-1} = \exp[(F - F_{id})/T]$. The energy of interaction is the sum over pairs:

$$U = u_{ab} + \sum_c (u_{ac} + u_{bc}) + \sum_{c,d \neq a,b} u_{cd} \ .$$

Note that $w_{ab} = 1$ for an ideal gas. Assuming the interaction weak and expanding (102) in $U/T$ we get terms like $u_{ab} w_{ab}$ and in addition $(u_{ac} + u_{bc}) w_{abc}$ which involves the third particle c and the triple correlation function that one can express via the integral similar to (102):

$$w_{abc} = V^{3-N} Z_i^{-1} \int \exp\left[\frac{-U(r_1 \ldots r_N)}{T}\right] dV_1 \ldots dV_{N-3} \ . \tag{103}$$

We can also see this (so-called closure problem) by differentiating

$$\frac{\partial w_{ab}}{\partial \mathbf{r}_b} = -\frac{w_{ab}}{T}\frac{\partial u_{ab}}{\partial \mathbf{r}_b} - (VT)^{-1}\sum_c N_c \int \frac{\partial u_{bc}}{\partial \mathbf{r}_b} w_{abc}\, dV_c \ , \tag{104}$$

and observing that the equation on $w_{ab}$ is not closed, it contains $w_{abc}$; the similar equation on $w_{abc}$ will contain $w_{abcd}$ etc. Debye-Hückel approximation corresponds to closing this hierarchical system of equations already at the level of the first equation (104) putting $w_{abc} \approx w_{ab}w_{bc}w_{ac}$ and assuming $\omega_{ab} = w_{ab} - 1 \ll 1$, that is assuming that two particles rarely come close while three particles never come together:

$$\frac{\partial \omega_{ab}}{\partial \mathbf{r}_b} = -\frac{1}{T}\frac{\partial u_{ab}}{\partial \mathbf{r}_b} - (VT)^{-1}\sum_c N_c \int \frac{\partial u_{bc}}{\partial \mathbf{r}_b} \omega_{ac}\, dV_c \ , \tag{105}$$

For other contributions to $w_{abc}$, the integral (which is a vector) turns into zero due to isotropy. This is the general equation valid for any form of interaction. For Coulomb interaction, we can turn the integral equation (105) into the differential equation by using $\Delta r^{-1} = -4\pi\delta(\mathbf{r})$ [12]. For that we differentiate (105) once more:

$$\Delta\omega_{ab}(r) = \frac{4\pi z_a z_b e^2}{T}\delta(\mathbf{r}) + \frac{4\pi z_b e^2}{TV}\sum_c N_c z_c \omega_{ac}(r) \ . \tag{106}$$

The dependence on ion charges and types is trivial, $\omega_{ab}(r) = z_a z_b \omega(r)$ and we get $\Delta\omega = 4\pi e^2\delta(\mathbf{r})/T + \kappa^2\omega$ which is (99) with delta-function enforcing the condition at zero. We see that the pair correlation function satisfies the same equation as the potential. Substituting the solution $\omega(r) = -(e^2/rT)\exp(-\kappa r)$ into $w_{ab}(r) = 1 + z_a z_b\omega(r)$ and that into (101) one gets contribution of 1 vanishing because of electro-neutrality and the term linear in $\omega$ giving (100). To get to the next order, one considers (104) together with the equation for $w_{abc}$, where one expresses $w_{abcd}$ via $w_{abc}$.

The correlation function of electron and ion densities are proportional to $w_{ab}$, for instance, $\langle n_-(0)n_-(r)\rangle = n_0^2 w_{--}$. The correlation functions of the fluctuations $\Delta n(r) = n(r) - n_0$ all have the form:

$$\begin{aligned}\langle \Delta n_-(0)\Delta n_-(r)\rangle &= \langle \Delta n_+(0)\Delta n_+(r)\rangle = -\langle \Delta n_-(0)\Delta n_+(r)\rangle \\ &= n_0^2\omega(r) = -\frac{(n_0 e)^2}{rT}\exp(-\kappa r) \ . \end{aligned} \tag{107}$$

---

[12] Integrate Laplacian over a region that includes zero: $4\pi\int_0 r^2 dr r^{-2}\partial_r r^2 \partial_r r^{-1} = -1$.

This screened Coulomb law is encountered in many situation, it is called Yukawa potential in quantum mechanics and Ornstein-Zernicke correlation function in statistical physics — we show later that this is a universal form of correlation function at large distances, see (161). The correlation function of the charge $q = e(n_i - n_e)$ is $4n_0^2\omega(r)$. We see that the densities of the particles of the same charge and the density of the total charge anti-correlate: if there is more in one place then there is less around. Densities of the different-charge particles have positive correlation function because the particles prefer to be close. For more details, see Landau & Lifshitz, Sects. 78,79.

After seeing how screening works, it is appropriate to ask what happens when there is no screening in a system with a long-range interaction. One example of that is gravity. Indeed, thermodynamics is very peculiar in this case. Arguably, the most dramatic manifestation is the Jeans instability of sufficiently large systems which leads to gravitational collapse, creation of stars and planets and, eventually, appearance of life.

The quantum (actually quasi-classical) variant of such mean-field consideration is called Thomas-Fermi method (1927) and is traditionally studied in the courses of quantum mechanics as it is applied to the electron distribution in large atoms (such placement is probably right despite the method is stat-physical because for physicists objects of study are more important than methods). In this method we consider the effect of electrostatic interaction on a degenerate electron gas at zero temperature. According to the Fermi distribution (72), the maximal kinetic energy (Fermi energy) is related to the local concentration $n(r)$ by $p_0^2/2m = (3\pi^2 n)^{2/3}\hbar^2/2m$, which is also the expression for the chemical potential at the zero temperature. We need to find the electrostatic potential $\phi(r)$ which determines the interaction energy for every electron, $-e\phi$. The sum of the chemical potential and the interaction energy, $p_0^2/2m - e\phi = -e\phi_0$, must be space-independent otherwise the electrons drift to the places with lower $-\phi_0$. The constant $\phi_0$ is zero for neutral atoms and negative for ions. We can now relate the local electron density $n(r)$ to the local potential $\phi(r)$: $p_0^2/2m = e\phi - e\phi_0 = (3\pi^2 n)^{2/3}\hbar^2/2m$ — that relation one must now substitute into the Poisson equation $\Delta\phi = 4\pi e n \propto (\phi - \phi_0)^{3/2}$. The boundary condition is $\phi \to Z/r$ at $r \to 0$ where $Z$ is the charge of the nuclei. The solution of this equation is a monotonically decreasing function, which has another power law at large distances: $\phi \propto r^{-4}$, $n \propto r^{-6}$. This solution cannot be extended to infinity as atoms are supposed to have finite sizes. Indeed, at large distances ($\propto Z^{1/3}$), the quasi-classical description breaks where the quantum wavelength is comparable to the distance $r$. The description also is inapplicable below the Bohr radius. The Thomas-Fermi approximation works well for large-atoms where there is an intermediate interval of distances (Landau&Lifshitz, Quantum Mechanics, Sect. 70).

## 4.2 Cluster and virial expansions

Consider a dilute gas with the short-range inter-particle energy of interaction $u(r)$. We assume that $u(r)$ decays on the scale $r_0$ and

$$\epsilon \equiv (2/3)\pi r_0^3 N/V \equiv bN/V \ll 1 .$$

Integrating over momenta we get the partition function $Z$ and the grand partition function $\mathcal{Z}$ as

$$\begin{aligned} Z(N,V,T) &= \frac{1}{N!\lambda_T^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \exp[-U(\mathbf{r}_1,\dots,\mathbf{r}_N)/T] \equiv \frac{Z_N(V,T)}{N!\lambda_T^{3N}} . \\ \mathcal{Z}(z,V,T) &= \sum_{N=0}^{\infty} \frac{z^N Z_N}{N!\lambda_T^{3N}} . \end{aligned} \tag{108}$$

Here we use fugacity $z = \exp(\mu/T)$ instead of the chemical potential. The terms with $N = 0, 1$ give unit integrals, with $N = 2$ we shall have $U_{12} = u(r_{12})$, then $U_{123} = u(r_{12}) + u(r_{13}) + u(r_{23})$, etc. In every term we may integrate over the coordinate of the center of mass of $N$ particles and obtain

$$\begin{aligned} \mathcal{Z}(\mu,V,T) = 1 + V\frac{z}{\lambda_T^3} + \frac{V}{2!}\left(\frac{z}{\lambda_T^3}\right)^2 \int d\mathbf{r} \exp[-u(r)/T] \\ + \frac{V}{3!}\left(\frac{z}{\lambda_T^3}\right)^3 \int \exp\{-[u(r_{12}) + u(r_{13}) + u(r_{23})]/T\}\, d\mathbf{r}_2 d\mathbf{r}_3 + \dots \end{aligned} \tag{109}$$

The first terms does not account for interaction. The second one accounts for the interaction of only one pair (under the assumption that when one pair of particles happens to be close and interact, this is such a rare event that the rest can be considered non-interacting). The third term accounts for simultaneous interaction of three particles etc. We can now write the Gibbs potential $\Omega = -PV = -T \ln \mathcal{Z}$ and expand the logarithm in powers of $z/\lambda_T^3$:

$$P = \lambda_T^{-3} \sum_{l=1}^{\infty} b_l z^l . \tag{110}$$

When expressing the expansion of the logarithm via the expansion of the function itself, one encounters so-called cumulants. We shall consider the general form of the cumulant expansion below in (231). Here we just notice that if we introduce $\varepsilon = \sum_{n=1} \alpha_n z^n$, then

$$\begin{aligned}
\ln(1+\varepsilon) &= 1+\varepsilon-\frac{\varepsilon^2}{2}+\frac{\varepsilon^3}{3}+\ldots \\
&= 1+\alpha_1 z+(\alpha_2-\frac{\alpha_1^2}{2})z^2+(\alpha_3-\alpha_1\alpha_2+\frac{\alpha_1^3}{3})z^3\ldots
\end{aligned}$$

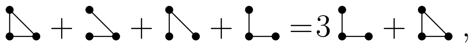
It is then convenient to introduce the two-particle function, called interaction factor, $f_{ij} = \exp[-u(r_{ij})/T]-1$, which is zero outside the range of interaction. Terms containing integrals of $k$ functions $f_{ij}$ are proportional to $\epsilon^k$. The coefficients $b_l$ can be expressed via $f_{ij}$:

$$\begin{aligned}
b_1 &= 1\,, \quad b_2 = (1/2)\lambda_T^{-3}\int f_{12}\, d\mathbf{r}_{12}\,, \\
b_3 &= (1/6)\lambda_T^{-6}\int\Big(e^{-U_{123}/T}-e^{-U_{12}/T}-e^{-U_{23}/T}-e^{-U_{13}/T}+2\Big)\, d\mathbf{r}_{12}d\mathbf{r}_{13} \\
&= (1/6)\lambda_T^{-6}\int(3f_{12}f_{13}+f_{12}f_{13}f_{23})\, d\mathbf{r}_{12}d\mathbf{r}_{13}\ . \qquad (111)
\end{aligned}$$

We see that in this case the integrals over different $d\mathbf{r}_i$ are not independent so that it is pretty cumbersome to analyze higher orders in analytical expressions. Instead, every term in

$$\begin{aligned}
Z_N(V,T) &= \int \textstyle\prod_{i<j}(1+f_{ij})\, d\mathbf{r}_1\ldots d\mathbf{r}_N \\
&= \int\Big(1+\sum f_{ij}+\sum f_{ij}f_{kl}+\ldots\Big)\, d\mathbf{r}_1\ldots d\mathbf{r}_N\ .
\end{aligned}$$

can be represented as a graph with $N$ points and lines connecting particles which interaction we account for. In this way, $Z_N$ is a sum of all distinct $N$-particle graphs. Since most people are better in manipulating visual (rather than abstract) objects then it is natural to use graphs to represent analytic expressions which is called *diagram technique*. For example, the three-particle clusters are as follows:



(112)

which corresponds to (111). Factorization of terms into independent integrals corresponds to decomposition of graphs into $l$-clusters i.e. $l$-point graphs where all points are directly or indirectly connected. Associated with the $l$-th cluster we may define dimensionless factors $b_l$ (called cluster integrals):

$$b_l = \frac{1}{l! V \lambda_T^{3(l-1)}} \times [\text{sum of all } l - \text{clusters}] \; . \tag{113}$$

In the square brackets here stand integrals like

$$\int d\mathbf{r} = V \text{ for } l = 1, \quad \int f(r_{12})\, d\mathbf{r}_1 d\mathbf{r}_2 = V \int f(r)\, d\mathbf{r} \text{ for } l = 2 \text{ ,etc .}$$

Using the cluster expansion one can now show that the cluster integrals $b_l$ indeed appear in the expansion (110) (see e.g. Pathria, Sects. 9.1-2, second edition). For $l = 1, 2, 3$ we saw that this is indeed so.

We can now use (110) and write the total number of particles:

$$PV = -\Omega = T \ln \mathcal{Z}(z, V) = (V/\lambda_T^3) \sum_{l=1}^{\infty} b_l z^l \tag{114}$$

$$n = \frac{1}{v} = \frac{Tz}{V} \frac{\partial \ln \mathcal{Z}}{\partial z} = \lambda_T^{-3} \sum_{l=1}^{\infty} l b_l z^l \; . \tag{115}$$

To get the equation of state we must express $z$ via $n\lambda_T^3$ from (115):

$$z = n\lambda_T^3/b_1 - 2b_2 b_1^{-3} (n\lambda_T^3)^2 + (8b_2^2 b_1^{-5} - 3b_3 b_1^{-4})(n\lambda_T^3)^3 + \ldots$$

We now substitute it into (114), which generates the series called the virial expansion

$$\frac{Pv}{T} = \sum_{l=1}^{\infty} a_l(T) \left( \frac{\lambda_T^3}{v} \right)^{l-1} \; . \tag{116}$$

Dimensionless virial coefficients are expressed via cluster coefficients i.e. they depend on the interaction potential and temperature:

$$a_1 = b_1 = 1 \; , \; a_2 = -b_2 \; , \; a_3 = 4b_2^2 - 2b_3 = -\lambda_T^{-6} \int f_{12} f_{13} f_{23} \, d\mathbf{r}_{12} d\mathbf{r}_{13}/3 \ldots \; .$$

In distinction from the cluster coefficients $b_l$, which contain terms of different order in $f_{ij}$, we now have $a_l \propto \epsilon^l$ i.e. $a_l$ comes from simultaneous interaction of $l$ particles. Using graph language, virial coefficients $a_l$ are determined by irreducible clusters i.e. such that there are at least two entirely independent non-intersecting paths that connect any two points. Tentative physical interpretation of the cluster expansion (116) is that we consider an ideal gas of clusters whose pressures are added.

## 4.3 Van der Waals equation of state

We thus see that the cluster expansion in powers of $f$ generates the virial expansion of the equation of state in powers of $n = N/V$. Here we account only for pairs of the interacting particles. The second virial coefficient

$$B(T) = a_2\lambda_T^3 = 2\pi \int \Big\{1 - \exp[-u(r)/T]\Big\} r^2\, dr \tag{117}$$

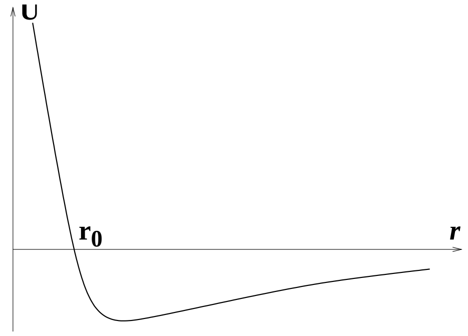
can be estimated by splitting the integral into two parts, from 0 to $r_0$ (where we can neglect the exponent assuming $u$ large positive) and from $r_0$ to $\infty$ (where we can assume small negative energy, $u \ll T$, and expand the exponent). That gives

$$B(T) = b - \frac{a}{T}\,, \quad a \equiv 2\pi \int_{r_0}^{\infty} u(r) r^2 dr\ . \tag{118}$$

with $b = (2/3)\pi r_0^3$ introduced above. Of course, for any particular $u(r)$ it is pretty straightforward to calculate $a_2(T)$ but (118) gives a good approximation for most cases. We can now get the first correction to the equation of state:

$$P = \frac{NT}{V}\left[1 + \frac{NB(T)}{V}\right] = nT(1 + bn) - an^2\ . \tag{119}$$

Generally, $B(T)$ is negative at low and positive at high $T$. For Coulomb interaction the correction to pressure (99) is always negative, while in this case it is positive at high temperature where molecules hit each other often and negative at low temperatures when long-range attraction between molecules decreases the pressure. Since $NB/V < Nb/V \ll 1$ the correction is small. Note that $a/T \ll r_0^3$ since we assume weak interaction.



While by its very derivation the formula (119) is derived for a dilute gas one may desire to change it a bit so that it can (at least qualitatively) describe the limit of incompressible liquid. That would require the pressure to go to infinity when density reaches some value. This is usually done by replacing in (119) $1 + bn$ by $(1 - bn)^{-1}$ which is equivalent for $bn \ll 1$ but for $bn \to 1$ gives $P \to \infty$. The resulting equation of state is called van der Waals equation:

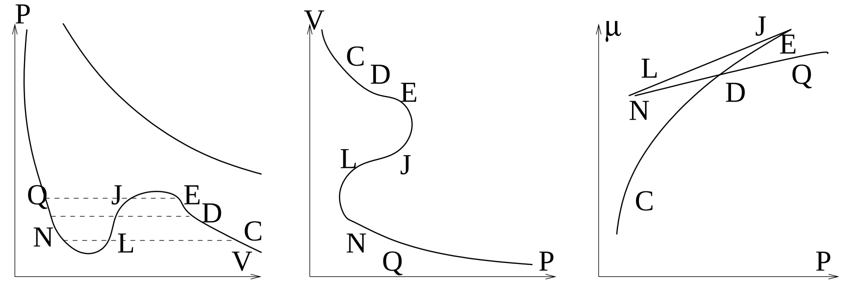
$$\Big(P + an^2\Big)(1 - nb) = nT \ . \tag{120}$$

There is though an alternative way to obtain (120) without assuming the gas dilute. This is some variant of the mean field even though it is not a first step of any consistent procedure. The essence of a mean-field approach is to replace calculating the partition function of $N$ interacting particles by a partition function of a single particle in some mean field created by all other particles. One version of it is to assume that every molecule moves in some effective field $U_e(r)$ which is a strong repulsion ($U_e \to +\infty$) in some region of volume $bN$ and is an attraction of order $-aN$ outside:

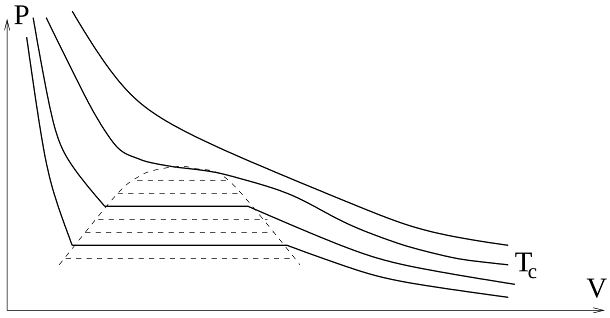
$$F - F_{id} \approx -TN \ln\left\{\int e^{-U_e(r)/T}\, d\mathbf{r}/V\right\} = -TN\left[\ln(1 - bn) + \frac{aN}{VT}\right] \ . \tag{121}$$

Differentiating (121) with respect to $V$ gives (120). That "derivation" also helps understand better the role of the parameters $b$ (excluded volume) and $a$ (mean interaction energy per molecule). From (121) one can also find the entropy of the van der Waals gas $S = -(\partial F/\partial T)_V = S_{id} + N\ln(1 - nb)$ and the energy $E = E_{id} - N^2a/V$, which are both lower than those for an ideal gas, while the sign of the correction to the free energy depends on the temperature. Since the correction to the energy is $T$-independent then $C_V$ is the same as for the ideal gas.

Let us now look closer at the equation of state (120). The set of isotherms is shown on the figure:



Since it is expected to describe both gas and liquid then it must show phase transition. Indeed, we see the region with $(\partial P/\partial V)_T > 0$ at the lower isotherm in the first figure. When the pressure corresponds to the level NLC, it is clear that L is an unstable point and cannot be realized. But which stable point is realized, N or C? To get the answer, one must minimize the Gibbs potential $G(T, P, N) = N\mu(T, P)$ since we have $T$ and $P$ fixed. For one mole, integrating the relation $d\mu(T, P) = -sdT + vdP$ at constant temperature we find: $G = \mu = \int v(P)dP$. It is clear that the pressure that corresponds to D (having equal areas before and above the horizontal line) separates the absolute minimum at the left branch Q (liquid-like) from that on the right one C (gas-like). The states E (over-cooled or over-compressed gas) and N (overheated or overstretched liquid) are metastable, that is they are stable with respect to small perturbations but they do not correspond to the global minimum of chemical potential. We thus conclude that the true equation of state must have isotherms that look as follows:



The dependence of volume on pressure is discontinuous along the isotherm in the shaded region (which is the region of phase transition). True partition function and true free energy must give such an equation of state. We were unable to derive it because we restricted ourselves by the consideration of the uniform systems while in the shaded region the system is nonuniform being the mixture of two phases. For every such isotherm $T$ we have a value of pressure $P(T)$, that corresponds to the point D, where the two phases coexist. On the other hand, we see that if temperature is higher than some $T_c$ (critical point), the isotherm is monotonic and there is no phase transition. Critical point was discovered by Mendeleev (1860) who also built the periodic table of elements. At critical temperature the dependence $P(V)$ has an inflection point: $(\partial P/\partial V)_T = (\partial^2 F/\partial V^2)_T = 0$. According to (34) the fluctuations must be large at the critical point (more detail in the next Chapter).

## 4.4 Thermodynamic description of phase transitions

The main theme of this Section is the competition (say, in minimizing the free energy) between the interaction energy, which tends to order systems, and the entropy, which brings a disorder. Upon the change of some parameters, systems can undergo a phase transition from more to less ordered states. We start from the phenomenological approach to the transitions of both first and second orders. We then proceed to develop a microscopic statistical theory based on Ising model.

### 4.4.1 Necessity of the thermodynamic limit

So far we got the possibility of a phase transition almost for free by cooking the equation of state for the van der Waals gas. But can one really derive the equations of state that have singularities or discontinuities? It clear that this is impossible in a finite system since the partition function is a finite sum of exponents, so it must be an analytic function of temperature. Indeed, the classical grand partition function (expressed via fugacity $z = \exp(\mu/T)$) is as follows:

$$\mathcal{Z}(z,V,T) = \sum_{N=0}^{\infty} z^N Z(N,V,T) \ . \tag{122}$$

Here the classical partition function of the $N$-particle system is

$$Z(N,V,T) = \frac{1}{N!\lambda^{3N}} \int \exp[-U(\mathbf{r}_1,\ldots,\mathbf{r}_N)/T] d\mathbf{r}_1,\ldots,\mathbf{r}_N \tag{123}$$

and the thermal wavelength is $\lambda^2 = 2\pi\hbar^2/mT$. Interaction means that for a finite volume $V$ there is a maximal number of molecules $N_m(V)$ that can be accommodated in $V$. That means that the sum in (122) actually goes until $N_m$ so that the grand partition function is a polynomial in fugacity with all coefficients positive[13]. The equation of state can be obtained by eliminating $z$ from the equations that give $P(v)$ in a parametric form — see (114,115):

$$\frac{P}{T} = \frac{1}{V}\ln\mathcal{Z}(z,V) = \frac{1}{V}\ln\left[1 + zZ(1,V,T) + \ldots + z^{N_m}Z(N_m,V,T)\right] \ ,$$

$$\frac{1}{v} = \frac{z}{V}\frac{\partial\ln\mathcal{Z}(z,V)}{\partial z} = \frac{zZ(1,V,T) + 2z^2Z(2,V,T)\ldots + N_m z^{N_m}Z(N_m,V,T)}{V[1 + zZ(1,V,T) +\ldots+ z^{N_m}Z(N_m,V,T)]} \ .$$

---

[13] Even when one does not consider hard-core models, the energy of repulsion grows so fast when the distance between molecules are getting less than some scale that Boltzmann factor effectively kills the contribution of such configurations.

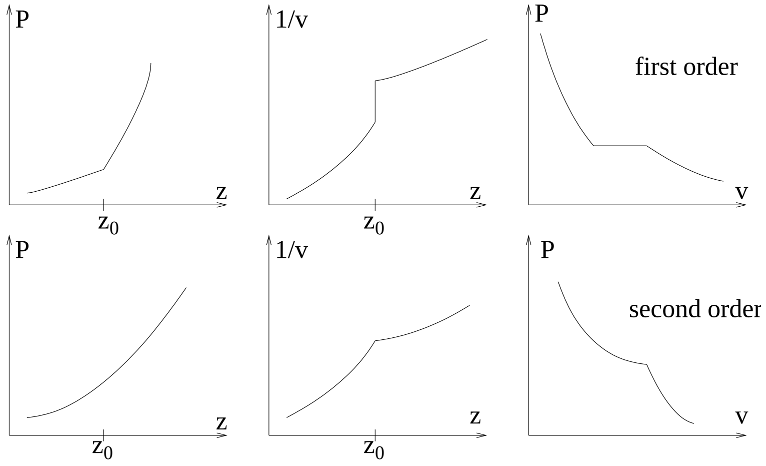
Polynomial with all coefficients positive does not turn into zero for real $z$. Therefore, for $\mathcal{Z}(z)$ being a polynomial, both $P(z)$ and $v(z)$ are analytic functions of $z$ in a region of the complex plane that includes the real positive axis. Therefore, $P(v)$ is an analytic function in a region of the complex plane that includes the real positive axis. Note that $V/N_m \leq v < \infty$. One can also prove that $\partial v^{-1}/\partial z > 0$ by using $\sum A_i i^2 \sum A_i \geq (\sum A_i i)^2$, where $A_i = z^i Z(i, V, T) > 0$ or by noticing that $z\partial_z n = \langle n^2 \rangle - \langle n \rangle^2 > 0$ for $n = 1/v$. That gives $\partial P/\partial v = (\partial P/\partial z)/(\partial v/\partial z) < 0$ for all $v$.

For a first-order transition, the pressure must be independent of $v$ in the transition region. We see that strictly speaking in a finite volume we cannot have that since $P(v)$ is analytic, nor we can have $\partial P/\partial v > 0$. That means that singularities, jumps etc can appear only in the thermodynamic limit $N \to \infty, V \to \infty$ (where, formally speaking, the singularities that existed in the complex plane of $z$ can come to the real axis). Such singularities are related to zeroes of $\mathcal{Z}(z)$. When such a zero $z_0$ tends to a real axis at the limit $N \to \infty$ (like the root $e^{-i\pi/2N}$ of the equation $z^N + 1 = 0$) then $1/v(z)$ and $P(z)$ are determined by two different analytic functions in two regions: one, including the part of the real axis with $z < z_0$ and another with $z > z_0$. These two different regions correspond to different phases (like liquid at $z > z_0$ and gas at $0 \leq z < z_0$); changing chemical potential we pass from one phase to another. Depending on the order of zero of $\mathcal{Z}(z)$, $1/v$ itself may have a jump or its $m$-th derivative may have a jump, which corresponds to the $m + 1$ order of phase transition. For a simple zero, $m = 0$, the singular part of the pressure $P \propto \lim_{V\to\infty} V^{-1} \ln |z - z_0 - O(V^{-1})|$ turns to zero at $z = z_0$ and $V \to \infty$; therefore the pressure is continuous at $z \to z_0$ but $\partial P/\partial z$ and $1/v$ are discontinuous - this is the transition of the first order[14]. For a second-order transition, volume is continuous but its derivative jumps. We see now what happens as we increase $T$ towards $T_c$: another zero comes from the complex plane into real axis and joins the zero that existed there before, turning 1st order phase transition into the 2nd order transition; at $T > T_c$ the zeroes leave the real axis. To conclude, we see that the system (114,115) in principle can detect the phase transition, is we are able to solve it and see that in the limit $N \to \infty$ the zero of $\mathcal{Z}(z)$ comes to the real axis.

---

[14] For a finite $N$, zeroes exist as complex-conjugated pairs, so one may wonder how a simple zero can appear. It can happen, for instance, when at $N \to \infty$ discrete zeroes merge into a cut and then two cuts $\sqrt{z - z_0 + i/N}$ and $\sqrt{z - z_0 - i/N}$ merge into a simple zero: $\sqrt{(z - z_0 + i/N)(z - z_0 - i/N)} \to z - z_0$

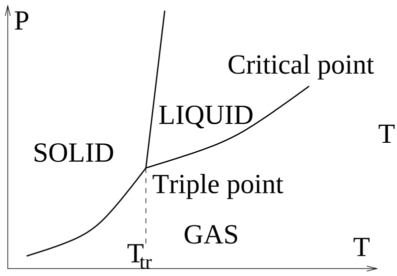
That one needs $N \to \infty$ for a phase transition is one of the manifestation of the "more is different" principle. See more details in Huang, Sect. 15.1-2.

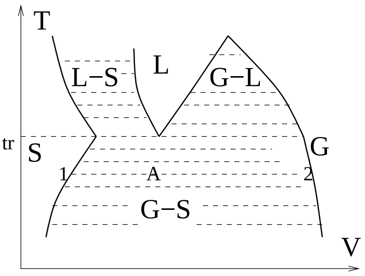


### 4.4.2 First-order phase transitions

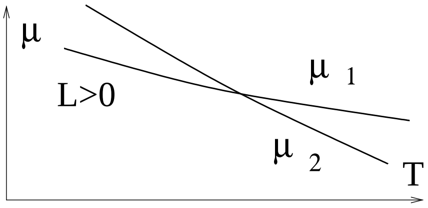
Let us now consider equilibrium between phases from a general viewpoint. We must have $T_1 = T_2$ and $P_1 = P_2$. Requiring $dG/dN_1 = \partial G_1/\partial N_1 + (\partial G_2/\partial N_2)(dN_2/dN_1) = \mu_1(P,T) - \mu_2(P,T) = 0$ we obtain the curve of the phase equilibrium $P(T)$. We thus see on the $P-T$ plane the states outside the curve are homogeneous while on the curve we have the coexistence of two different phases. If one changes pressure or temperature crossing the curve then the phase transition happens. Three phases can coexist only at a point.

On the $T-V$ plane the states with phase coexistence fill whole domains (shaded on the figure) since different phases have different specific volumes. Different points on the $V-T$ diagram inside the coexistence domains correspond to different fractions of phases. Consider, for instance, the point A inside the gas-solid coexistence domain. Since the specific volumes of the solid and the gas are given by the abscissas of the points 1 and 2 respectively then the fractions of the phases in the state A are inversely proportional to the lengthes A1 and A2 respectively (the lever rule).





Changing $V$ at constant $T$ in the coexistence domain (say, from the state 1 to the state 2) we realize the phase transition of the first order. Phase transitions of the first order are accompanied by an absorption or release of some (latent) heat $L$. Since the transition happens in equilibrium (between phases) and at fixed temperature then the heat equals simply to $L = \int T ds = T(s_2 - s_1)$ (per mole). If 2 is preferable to 1 at higher $T$, then $s_2 = -\partial\mu_2/\partial T > s_1 = -\partial\mu_1/\partial T$ and $L > 0$ (heat absorbed) upon the $1 \to 2$ transition:



It must be so according to the Le Chatelier principle: adding heat we increase temperature which causes the phase transition, which absorbs heat. On the other hand, differentiating $\mu_1(P,T) = \mu_2(P,T)$ with respect to $T$ and using $s = -(\partial\mu/\partial T)_P$, $v = (\partial\mu/\partial P)_T$, one gets the Clausius-Clapeyron equation

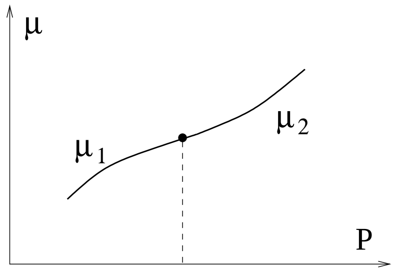
$$\frac{dP}{dT} = \frac{s_1 - s_2}{v_1 - v_2} = \frac{L}{T(v_2 - v_1)} \ . \tag{124}$$

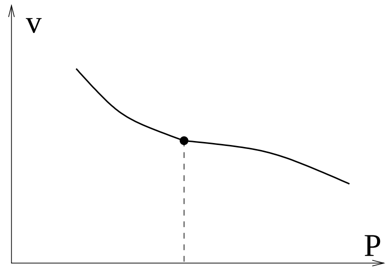
Since the entropy of a liquid is usually larger than that of a solid then $L > 0$ that is the heat is absorbed upon melting and released upon freezing. Most of the substances also expand upon melting then the solid-liquid equilibrium line has $dP/dT > 0$, as on the P-T diagram above. Water, on the contrary, contracts upon melting so the slope of the melting curve is negative (fortunate for fish and unfortunate for Titanic, ice floats on water). Note that symmetries of solid and liquid states are different so that one cannot continuously transform solid into liquid. That means that the melting line starts on another line and goes to infinity since it cannot end in a critical point (like the liquid-gas line).

Clausius-Clapeyron equation allows one, in particular, to obtain the pressure of vapor in equilibrium with liquid or solid. In this case, $v_1 \ll v_2$. We may treat the vapor as an ideal so that $v_2 = T/P$ and (124) gives $d\ln P/dT = L/T^2$. We may further assume that $L$ is approximately independent of $T$ and obtain $P \propto \exp(-L/T)$ which is a fast-increasing function of temperature. Landau & Lifshitz, Sects. 81–83.

### 4.4.3 Second-order phase transitions

As we have seen, in the critical point, the differences of specific entropies and volumes turn into zero. Considering $\mu(P)$ at $T = T_c$ one can say that the chemical potential of one phase ends where another starts and the derivative $v(P) = (\partial\mu/\partial P)_{T_c}$ is continuous.





Another examples of continuous phase transitions (i.e. such that correspond to a continuous change in the system) are all related to the change in symmetry upon the change of $P$ or $T$. Since symmetry is a qualitative characteristics, it can change even upon an infinitesimal change (for example, however small ferromagnetic magnetization breaks isotropy). Here too every phase can exist only on one side of the transition point. The transition with first derivatives of the thermodynamic potentials continuous is called second order phase transition. Because the phases end in the point of transition, such point must be singular for the thermodynamic potential, and indeed second derivatives, like specific heat, are generally discontinuous. One set of such transitions is related to the shifts of atoms in a crystal lattice; while close to the transition such shift is small (i.e. the state of matter is almost the same) but the symmetry of the lattice changes abruptly at the transition point. Another set is a spontaneous appearance of macroscopic magnetization (i.e. ferromagnetism) below Curie temperature. Transition to superconductivity is of the second order. Variety of second-order phase transitions happen in liquid crystals etc. Let us stress that some transitions with a symmetry change are first-order (like melting) but all second-order phase transitions correspond to a symmetry change.

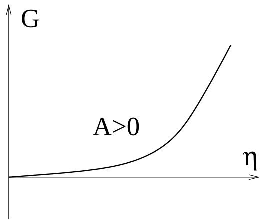
It is important to stress that we are talking about symmetries (i.e. invariance under transformations) of the thermodynamic and statistical properties. In this respect, solid phase is *less symmetric* than liquid or gas. Indeed, the properties of a disordered medium are invariant under arbitrary translations and rotations in space, while the properties of a crystal are invariant only under translations by lattice spacings and rotation by the angles of the crystal lattice. In other words, disordered media are invariant with respect to continuous symmetries, while crystals have discrete symmetries. Of course, would we be interested in the symmetries of a medium as a geometric set of objects then crystal coincides with itself after discrete transformations while disordered medium does not have any exact symmetry at all, this is why we think about crystals as more symmetric. In statistical physics, however, appearance of any order makes the state less symmetric.

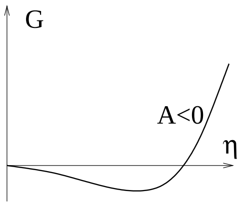
### 4.4.4 Landau theory

To describe general properties of the second-order phase transitions Landau suggested to characterize symmetry breaking by some order parameter $\eta$ which is zero in the symmetrical phase and is nonzero in nonsymmetric phase. Example of an order parameter is magnetization. The choice of order parameter is non-unique; to reduce arbitrariness, it is usually required to transform linearly under the symmetry transformation. The thermodynamic potential can be formally considered as $G(P,T,\eta)$ even though $\eta$ is not an independent parameter and must be found as a function of $P,T$ from requiring the minimum of $G$. We can now consider the thermodynamic potential near the transition as a series in small $\eta$:

$$G(P,T,\eta) = G_0 + A(P,T)\eta^2 + B(P,T)\eta^4 \ . \tag{125}$$

The linear term is absent to keep the first derivative continuous at $\eta = 0$. The coefficient $B$ must be positive since arbitrarily large values of $\eta$ must cost a lot of energy. The coefficient $A$ must be positive in the symmetric phase when minimum in $G$ corresponds to $\eta = 0$ (left figure below) and negative in the non-symmetric phase where $\eta \neq 0$. Therefore, at the transition $A_c(P,T) = 0$ and $B_c(P,T) > 0$:





We assume that the symmetry of the system requires the absence of $\eta^3$-term, then the only requirement on the transition is $A_c(P,T) = 0$ so that the transition points fill the line in $P - T$ plane. If the transition happens at some $T_c$ then generally near transition[15] $A(P,T) = a(P)(T - T_c)$. Writing then the potential

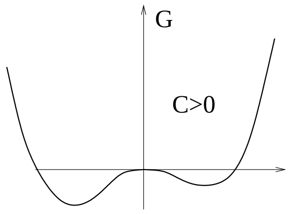
$$G(P,T,\eta) = G_0 + a(P)(T - T_c)\eta^2 + B(P,T)\eta^4 \ , \tag{126}$$

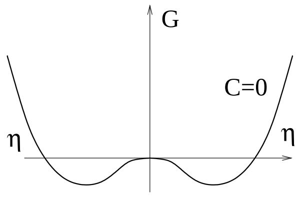
and requiring $\partial G/\partial \eta = 0$ we get

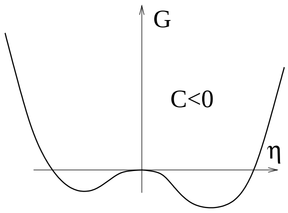
$$\bar{\eta}^2 = \frac{a}{2B}(T_c - T) \ . \tag{127}$$

In the lowest order in $\eta$ the entropy is $S = -\partial G/\partial T = S_0 + a^2(T - T_c)/2B$ at $T < T_c$ and $S = S_0$ at $T > T_c$. Entropy is lower at lower-temperature phase (which is generally less symmetric). Specific heat $C_p = T\partial S/\partial T$ has a jump at the transitions: $\Delta C_p = a^2 T_c/2B$. Specific heat increases when symmetry is broken since more types of excitations are possible.

If symmetries allow the cubic term $C(P,T)\eta^3$ (like in a gas or liquid near the critical point discussed in Sect. 5.2 below) then one generally has a first-order transition, say, when $A < 0$ and $C$ changes sign:







It turns into a second-order transition, for instance, when $A = C = 0$ i.e. only in isolated points in $P - T$ plane.

Consider now what happens when there is an external field (like magnetic field) which contributes the energy (and thus the thermodynamic potential)

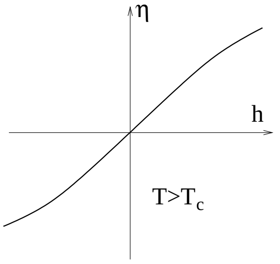
---

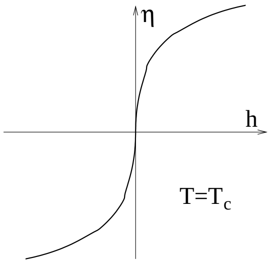
[15] We assume $a > 0$ since in almost all cases the more symmetric state corresponds to higher temperatures; rare exceptions exist so this is not a law.

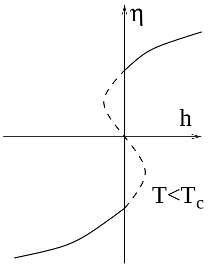
by the term $-h\eta V$. Equilibrium condition,

$$2a(T-T_c)\eta + 4B\eta^3 = hV\ , \tag{128}$$

has one solution $\eta(h)$ above the transition and may have three solutions (one stable, two unstable) below the transition:







The similarity to the van der Waals isotherm is not occasional: changing the field at $T < T_c$ one encounters a first-order phase transition at $h = 0$ where the two phases with $\eta = \pm\sqrt{a(T_c - T)/2B}$ coexist. We see that $p - p_c$ is analogous to $h$ and $1/v - 1/v_c$ to the order parameter (magnetization) $\eta$.

Susceptibility diverges at $T \to T_c$:

$$\chi = \left(\frac{\partial\eta}{\partial h}\right)_{h=0} = \frac{V}{2a(T-T_c)+12B\eta^2} = \begin{cases} [2\alpha(T-T_c)]^{-1} \text{ at } T > T_c \\ [4\alpha(T_c-T)]^{-1} \text{ at } T < T_c \end{cases} \tag{129}$$

Comparing to $\chi \propto 1/T$ obtained in (37) for noninteracting spins we see that the paramagnetic phase corresponds to $T \gg T_c$. Experiments support the Curie law (129). Since $a \propto V$ we have introduced $\alpha = a/V$ in (129).

We see that Landau theory (based on the only assumption that the thermodynamic potential must be an analytic function of the order parameter) gives universal predictions independent on space dimensionality and of all the details of the system except symmetries. Is it true? Considering specific systems we shall see that Landau theory actually corresponds to the behavior near $T_c$ of a mean-field approximation i.e. to neglecting the fluctuations. The potential is getting flat near the transition then the fluctuations grow. In particular, the probability of the order parameter fluctuations around the equilibrium value $\bar\eta$ behaves as follows

$$\exp\left[-\frac{(\eta-\bar\eta)^2}{T_c}\left(\frac{\partial^2 G}{\partial\eta^2}\right)_{T,P}\right]\ ,$$

so that the mean square fluctuation of the order parameter, $\langle(\eta-\bar\eta)^2\rangle = T_c/2A = T_c/2a(T-T_c)$. Remind that $a$ is proportional to the volume under consideration. Fluctuations are generally inhomogeneous and are correlated on some scale. To establish how the correlation radius depends on $T - T_c$ one can generalize the Landau theory for inhomogeneous $\eta(\mathbf{r})$, which is done in Sect. 5.2 below, where we also establish the validity conditions of the mean field approximation and of the Landau theory. Landau & Lifshitz, Sects. 142, 143, 144, 146.

## 4.5 Ising model

We now descend from phenomenology to real microscopic statistical theory. Our goal is to describe how disordered systems turn into ordered one when interaction prevails over thermal motion. Different systems seem to be having interaction of different nature with their respective difficulties in the description. For example, for the statistical theory of condensation one needs to account for many-particle collisions. Magnetic systems have interaction of different nature and the technical difficulties related with the commutation relations of spin operators. It is remarkable that there exists one highly simplified approach that allows one to study systems so diverse as ferromagnetism, condensation and melting, order-disorder transitions in alloys, phase separation in binary solutions, and also model phenomena in economics, sociology, genetics, to analyze the spread of forest fires etc. This approach is based on the consideration of lattice sites with the nearest-neighbor interaction that depends upon the manner of occupation of the neighboring sites. We shall formulate it initially on the language of ferromagnetism and then establish the correspondence to some other phenomena.

### 4.5.1 Ferromagnetism

Experiments show that ferromagnetism is associated with the spins of electrons (not with their orbital motion). Spin 1/2 may have two possible projections, so we consider lattice sites with elementary magnetic moments $\pm\mu$. In Sect. 2.5.1, we considered this system in an external magnetic field $H$ without any interaction between moments and got the magnetization (36):

$$M = N\mu \frac{\exp(\mu H/T) - \exp(-\mu H/T)}{\exp(\mu H/T) + \exp(-\mu H/T)} \ . \tag{130}$$

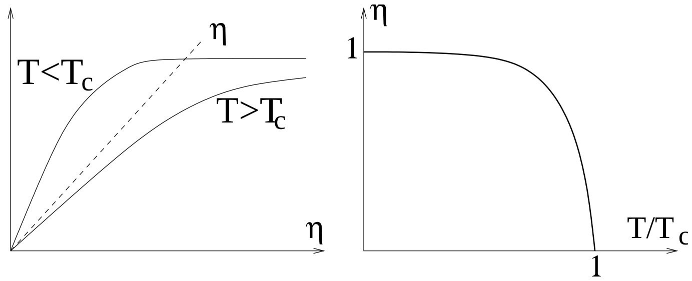
Of course, this formula gives no first-order phase transition upon the change of the sign of $H$, in distinction from the Landau theory (128). The reason is that (130) does not account for interaction (i.e. $B$-term in the Landau theory). First phenomenological treatment of the interacting system was done by Weiss who *assumed* that there appears some extra magnetic field, proportional to magnetization per particle. One adds that field to $H$ and thus describes the influence that $M$ causes upon itself:

$$M = N\mu \tanh \frac{\mu(H+\xi M/N)}{T}) \ . \tag{131}$$

And now put the external field to zero $H = 0$. The resulting equation can be written as

$$\eta = \tanh \frac{T_c \eta}{T} \ , \tag{132}$$

where we denoted $\eta = M/\mu N$ and $T_c = \xi\mu^2$ - that this is indeed the critical temperature must follow from the fact that at higher temperatures only zero magnetization is possible. Indeed, there is a single solution $\eta = 0$ at $T > T_c$, while at $T < T_c$ there are two more nonzero solutions which exactly means the appearance of the spontaneous magnetization. In the figure below, the left panel shows graphic solution of (132): the broken line shows the left side and the two solid lines show the right side respectively at $T > T_c$ and $T < T_c$. At $T_c - T \ll T_c$ one has $\eta^2 = 3(T_c - T)$ exactly as in Landau theory (127).



One can compare $T_c$ with experiments and find surprisingly high $\xi \sim 10^3 \div 10^4$. That means that the real interaction between moments is much higher than the interaction between neighboring dipoles $\mu^2 n = \mu^2/a^3$. Frenkel and Heisenberg solved this puzzle (in 1928): it is not the magnetic energy but the difference of electrostatic energies of electrons with parallel and antiparallel spins, so-called exchange energy, which is responsible for the interaction (parallel spins have antisymmetric coordinate wave function and much lower energy of interaction than antiparallel spins).

To develop a regular procedure starting from the mean-field approximation, one must formulate the microscopic model. The so-called Ising model was formulated by Lenz in 1920 and solved in one dimension by his student Ising in 1925. It deals with the discrete spin variable $\sigma_i = \pm 1$ at every lattice site. The energy includes interaction with the external field and between nearest neighbors (n.n.):

$$\mathcal{H} = -\mu H \sum_i^N \sigma_i + \frac{J}{2} \sum_{ij-n.n.} (1 - \sigma_i \sigma_j) \ . \tag{133}$$

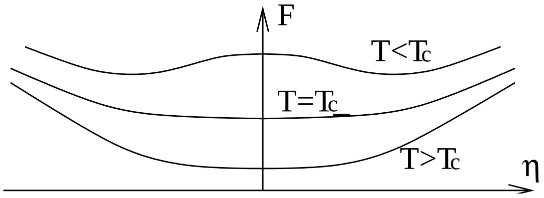
We assume that every spin has $\gamma$ neighbors ($\gamma = 2$ in one-dimensional chain, 4 in two-dimensional square lattice, 6 in three dimensional simple cubic lattice etc). We see that parallel spins have zero interaction energy while antiparallel have $J$ (which is comparable to Rydberg).

Let us start from $H = 0$. Magnetization is completely determined by the numbers of spins up: $M = \mu(N_+ - N_-) = \mu(2N_+ - N)$. We need to write the free energy $F = E - TS$ and minimizing it find $N_+$. The competition between energy and entropy determines the phase transition. Entropy is easy to get: $S = \ln C_N^{N_+} = \ln[N!/N_+!(N - N_+)!]$. The energy of interaction depends on the number of neighbors with opposite spins $N_{+-}$. The crudest approximation (Bragg and Williams, 1934) is, of course, mean-field, i.e. replacing N-particle problem with a single-particle problem. It consists of saying that every up spin has the number of down neighbors equal to the mean value $\gamma N_-/N$ so that the energy $\langle \mathcal{H} \rangle = E = JN_{+-} \approx \gamma N_+(N - N_+)J/N$. Requiring the minimum of the free energy, $\partial F/\partial N_+ = 0$, we get:

$$\gamma J \frac{N - 2N_+}{N} - T \ln \frac{N - N_+}{N_+} = 0 \ . \tag{134}$$

Here we can again introduce the variables $\eta = M/\mu N$ and $T_c = \gamma J/2$ and reduce (134) to (132). We thus see that indeed Weiss guess gave the same equation of state as the mean field approximation. The addition is that now we have the expression for the free energy, $F/2N = T_c(1 - \eta^2) - T(1 + \eta)\ln(1 + \eta) - T(1 - \eta)\ln(1 - \eta)$, so that we can indeed make sure that the nonzero $\eta$ at $T < T_c$ correspond to minima. Here is the free energy plotted as a function of magnetization, we see that it has exactly the form we assumed in the Landau theory (which corresponds to the mean field approximation at $T$ close to $T_c$). The energy is symmetrical with respect to flipping all the spins simultaneously. The free energy is symmetric with respect to $\eta \leftrightarrow -\eta$. But the system at $T < T_c$ lives in one of the minima (positive or negative $\eta$).

When the symmetry of the state is less than the symmetry of the potential (or Hamiltonian) it is called spontaneous symmetry breaking.



We can also calculate the specific heat using $E = \gamma N_+(N - N_+)J/N = (T_c N/2)(1 - \eta^2)$ and obtain the jump exactly like in Landau theory:

$$\eta = \sqrt{3(T_c - T)}\,, \quad \Delta C = C(T_c - 0) = -T_c N \eta \frac{d\eta}{dT} = 3N/2 \ .$$

At $T \to 0$, we have $\eta \approx 1 - 2\exp(-2T_c/T)$ and the specific heat vanishes exponentially: $C(T) \approx 4N(T_c/T)^2 \exp(-2T_c/T)$.

Note that in our approximation, when the long-range order (i.e. $N_+$) is assumed to completely determine the short-range order (i.e. $N_{+-}$), the energy is independent of temperature at $T > T_c$ since $N_+ \equiv N/2$. We do not expect this in reality. Moreover, let us not delude ourselves that we proved the existence of the phase transition. How wrong is the mean-field approximation one can see comparing it with the exact solution for the one-dimensional chain. Indeed, consider again $H = 0$. It is better to think not about spins but about the links between spins. Starting from the first spin, the state of the chain can be defined by saying whether the next one is parallel to the previous one or not. If the next spin is opposite it gives the energy $J$ and if it is parallel the energy is zero. There are $N - 1$ links. The partition function is that of the $N - 1$ two-level systems (38):

$$Z = 2[1 + \exp(-\beta J)]^{N-1} \ . \tag{135}$$

Here 2 because there are two possible orientations of the first spin.

One can do it for a more general Hamiltonian

$$\mathcal{H} = -\sum_{i=1}^{N}\left[H\sigma_i + \frac{J}{2}(1 - \sigma_i\sigma_{i+1})\right] = -\sum_{i=1}^{N}\left[\frac{H}{2}(\sigma_i + \sigma_{i+1}) + \frac{J}{2}(1 - \sigma_i\sigma_{i+1})\right] \tag{136}$$

To avoid considering the open ends of the chain (which must be irrelevant in the thermodynamic limit), we consider it on a ring so that $\sigma_{N+1} = \sigma_1$ and write the partition function as a simple sum over spin value at every cite:

$$
\begin{aligned}
Z &= \sum_{\{\sigma_i\}} \exp\left[\beta \sum_{i=1}^{N} \left\{ \frac{H}{2}(\sigma_i+\sigma_{i+1}) - \frac{J}{2}(1-\sigma_i\sigma_{i+1}) \right\}\right] && (137)\\
&= \sum_{\{\sigma_i\}} \prod_{i=1}^{N} \exp\left[\beta \left\{ \frac{H}{2}(\sigma_i+\sigma_{i+1}) - \frac{J}{2}(1-\sigma_i\sigma_{i+1}) \right\}\right] && (138)
\end{aligned}
$$

Every factor in the product can have four values, which correspond to four different choices of $\sigma_i = \pm 1, \sigma_{i+1} = \pm 1$. Therefore, every factor can be written as a matrix element of $2 \times 2$ matrix: $\langle \sigma_j | \hat{T} | \sigma_{j+1} \rangle = T_{\sigma_j \sigma_{j+1}} = \exp[\beta\{H(\sigma_i + \sigma_{i+1})/2 - J(1 - \sigma_i\sigma_{i+1})/2\}$. It is called the transfer matrix because it *transfers* us from one cite to the next.

$$
T = \begin{pmatrix} T_{1,1} & T_{1,-1} \\ T_{-1,1} & T_{-1,-1} \end{pmatrix} \tag{139}
$$

where $T_{11} = e^{\beta H}$, $T_{-1,-1} = e^{-\beta H}$, $T_{-1,1} = T_{1,-1} = e^{-\beta J}$. For any matrices $\hat{A}, \hat{B}$ the matrix elements of the product are $[AB]_{ik} = A_{ij}B_{jk}$. Therefore, when we sum over the values of the intermediate spin, we obtain the matrix elements of the matrix squared: $\sum_{\sigma_i} T_{\sigma_{i-1}\sigma_i} T_{\sigma_i \sigma_{i+1}} = [T^2]_{\sigma_{i-1}\sigma_{i+1}}$. The sum over $N-1$ spins gives $T^{N-1}$. Because of periodicity we end up with summing over a single spin which corresponds to taking trace of the matrix:

$$
Z = \sum_{\{\sigma_i\}} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \dots T_{\sigma_N\sigma_1} = \sum_{\sigma_1 = \pm 1} \langle \sigma_1 | \hat{T}^{N-1} | \sigma_1 \rangle = \mathrm{trace}\, T^{N-1} \ . \tag{140}
$$

The eigenvalues $\lambda_1, \lambda_2$ of $T$ are given by

$$
\lambda_{1,2} = \cosh(\beta H) \pm \sqrt{\sinh^2(\beta H) + e^{-2\beta J}} \ . \tag{141}
$$

The trace is the sum of the eigenvalues

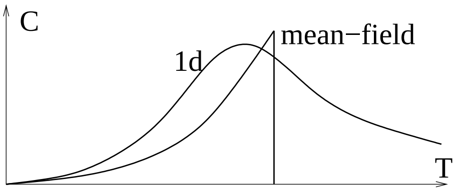
$$
Z = \lambda_1^{N-1} + \lambda_2^{N-1} \, . \tag{142}
$$

Therefore

$$
\begin{aligned}
F = -T\log(\lambda_1^{N-1} + \lambda_2^{N-1}) = -T\Bigg[&(N-1)\log(\lambda_1) \\
+\log\left(1 + \left(\frac{\lambda_2}{\lambda_1}\right)^{N-1}\right)\Bigg] &\to -NT\log\lambda_1 \quad \text{as} \quad N \to \infty
\end{aligned} \tag{143}
$$

Note that the partition functions (142) at $H = 0$ and (135) give the same free energies only at the thermodynamics limit when a ring is indistinguishable from a chain with open ends.

Now, as we know, there is no phase transitions for a two-level system. In particular one can compare the mean-field energy $E = T_c(1-\eta^2)$ with the exact 1d expression (41) which can be written as $E(T) = NJ/(1+e^{J/T})$ and compare the mean field specific heat with the exact 1d expression:



We can improve the mean-field approximation by accounting exactly for the interaction of a given spin $\sigma_0$ with its $\gamma$ nearest neighbors and replacing the interaction with the rest of the lattice by a new mean field $H'$ (this is called Bethe-Peierls or BP approximation):

$$\mathcal{H}_{\gamma+1} = -\mu H' \sum_{j=1}^{\gamma} \sigma_j - (J/2) \sum_{j=1}^{\gamma} \sigma_0 \sigma_j \ . \tag{144}$$

The external field $H'$ is determined by the condition of self-consistency, which requires that the mean values of all spins are the same: $\bar\sigma_0 = \bar\sigma_i$. To do that, let us calculate the partition function of this group of $\gamma + 1$ spins:

$$Z = \sum_{\sigma_0, \sigma_j = \pm 1} \exp\Big(\eta \sum_{j=1}^{\gamma} \sigma_j + \nu \sum_{j=1}^{\gamma} \sigma_0 \sigma_j\Big) = Z_+ + Z_- \ ,$$

$$Z_\pm = \sum_{\sigma_j = \pm 1} \exp\Big[(\eta \pm \nu) \sum_{j=1}^{\gamma} \sigma_j\Big] = [2\cosh(\eta \pm \nu)]^\gamma \ , \quad \eta = \mu H'/T \ , \ \nu = J/2T \ .$$

$Z_\pm$ correspond to $\sigma_0 = \pm 1$. Requiring $\bar\sigma_0 = (Z_+ - Z_-)/Z$ to be equal to

$$\begin{aligned} \bar\sigma_j &= \frac{1}{\gamma}\Big\langle \sum_{j=1}^{\gamma} \sigma_j \Big\rangle = \frac{1}{\gamma Z}\frac{\partial Z}{\partial \eta} \\ &= Z^{-1}\Big\{[2\cosh(\eta+\nu)]^{\gamma-1}\sinh(\eta+\nu) + [2\cosh(\eta-\nu)]^{\gamma-1}\sinh(\eta-\nu)\Big\} \ , \end{aligned}$$

we obtain

$$\eta = \frac{\gamma - 1}{2} \ln\left[\frac{\cosh(\eta+\nu)}{\cosh(\eta-\nu)}\right] \tag{145}$$

instead of (132) or (134). Condition that the derivatives with respect to $\eta$ at zero are the same, $(\gamma-1)\tanh\nu = 1$, gives the critical temperature:

$$T_c = J\ln^{-1}\Big(\frac{\gamma}{\gamma-2}\Big)\,, \quad \gamma \geq 2\ . \tag{146}$$

It is lower than the mean field value $\gamma J/2$ and tends to it when $\gamma \to \infty$ — mean field is exact in an infinite-dimensional space. More important, it shows that there is no phase transition in 1d when $\gamma = 2$ and $T_c = 0$ (in fact, BP is exact in 1d). Note that now the magnetization is not $\eta$ but the mean spin $\bar\sigma_0 = \sinh(2\eta)/[\cosh(2\eta) + \exp(-2\nu)]$. BP gives nonzero energy and specific heat at $T > T_c$ (despite $\eta = \bar\sigma_0 = 0$): $Z = 2^{\gamma+1}\cosh^{\gamma}(\beta J/2)$ and $F = -T\ln Z = -(\gamma/\beta)\ln\cosh(\beta J/2)$. The energy is $E = \partial(F\beta)/\partial\beta = (\gamma J/2)\tanh(\beta J/2)$ and the specific heat, $C = (\gamma J^2/8T^2)\cosh^2(J/2T)$, such that $C \propto T^{-2}$ at $T \to \infty$ (see Pathria 11.6 for more details):

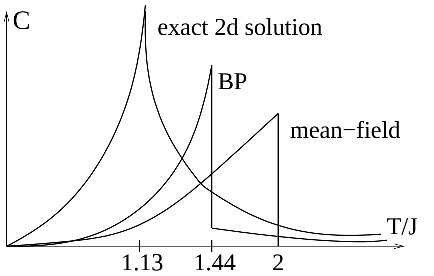


Figure 2: Specific heat of the Ising model in two dimensions.

Bragg-Williams and Bethe-Peierls approximations are the first and the second steps of some consistent procedure. When the space dimensionality is large, then $1/d$ is a small parameter whose powers determine the contributions of the subsequent approximations. Mean-field corresponds to the total neglect of fluctuations, while BP accounts for them in the first approximation. One can also say that it corresponds to the account of correlations: indeed, correlations make fluctuations (like having many spins with the same direction in a local neighborhood) more probable and require us to account for them. We see that subsequent approximations shift the critical temperature down and make the specific heat at the transition higher. This is natural as we increase the number of the degrees of freedom we account for. The two-dimensional Ising model was solved exactly by Onsager (1944). The exact solution shows the phase transition in two dimensions. The main qualitative difference from the mean field and BP is the divergence of the specific heat at the transition: $C \propto -\ln|1-T/T_c|$. This is the result of fluctuations: the closer one is to $T_c$ the stronger are the fluctuations. The singularity of the specific heat is integrable that is, for instance, the entropy change $S(T_1)-S(T_2)=\int_{T_1}^{T_2} C(T)dT/T$ is finite across the transition (and goes to zero when $T_1 \to T_2$) and so is the energy change. Note also that the true $T_c = J/2\ln[(\sqrt{2}-1)^{-1}]$ is less than both the mean-field value $T_c = \gamma J/2 = 2J$ and BP value $T_c = J/\ln 2$ also because of fluctuations (one needs lower temperature to "freeze" the fluctuations and establish the long-range order).

### 4.5.2 Equivalent models

The anti-ferromagnetic case has $J < 0$ and the ground state at $T = 0$ corresponds to the alternating spins. We may break our lattice into two sublattices, one having all spins up, another down. Without an external magnetic field, the magnetization of every sublattice is the same as for Ising model with $J > 0$. That follows from the fact that the energy is invariant with respect to the transformation $J \to -J$ and flipping all the spins of one of the sublattices. Therefore we have the second-order phase transition at zero field and at the temperature which is called Neel temperature. The difference from ferromagnetic is that there is a phase transition also at a nonzero external field (there is a line of transition in $H-T$ plane.

One can try to describe the condensation transition by considering a regular lattice with $N$ cites that can be occupied or not. We assume our lattice to be in a contact with a reservoir of atoms so that the total number of atoms, $N_a$, is not fixed. We thus use a grand canonical description with $\mathcal{Z}(z,N,T)$ given by (108). We model the hard-core repulsion by requiring that a given cite cannot be occupied by more than one atom. The number of cites plays the role of volume (choosing the volume of the unit cell unity). If the neighboring cites are occupied by atoms it corresponds to the (attraction) energy $-2J$ so we have the energy $E = -2JN_{aa}$ where $N_{aa}$ is the total number of nearest-neighbor pairs of atoms. The partition function is

$$Z(N_a,T) = \sum^{a} \exp(2JN_{aa}/T)\,, \tag{147}$$

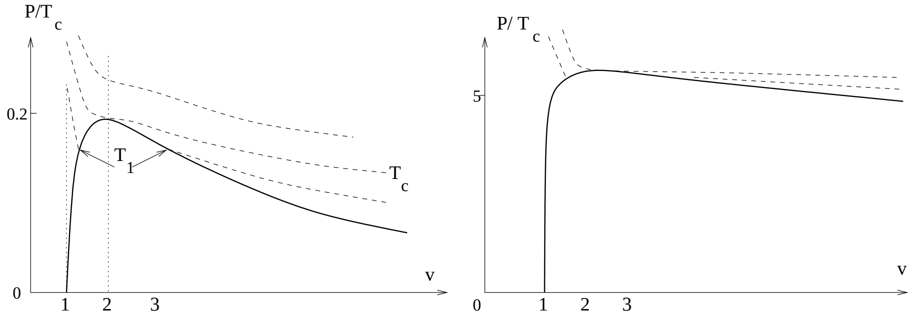
where the sum is over all ways of distributing $N_a$ indistinguishable atoms over $N$ cites. Of course, the main problem is in calculating how many times one finds the given $N_{aa}$. The grand partition function,

$$\mathcal{Z}(z, V, T) = \sum_{N_a}^{\infty} z^{N_a} Z(N_a, T)) \,, \tag{148}$$

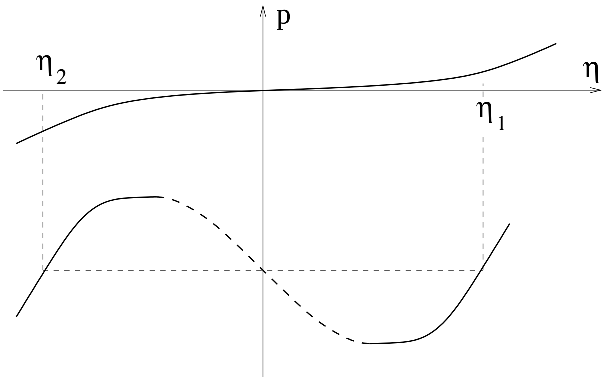
gives the equation of state in the implicit form (like in Sect. 4.4.1): $P = T \ln \mathcal{Z}/N$ and $1/v = (z/V)\partial \ln \mathcal{Z}/\partial z$. The correspondence with the Ising model can be established by saying that an occupied site has $\sigma = 1$ and unoccupied one has $\sigma = -1$. Then $N_a = N_+$ and $N_{aa} = N_{++}$. Recall that for Ising model, we had $E = -\mu H(N_+ - N_-) + JN_{+-} = \mu HN + (J\gamma - 2\mu H)N_+ - 2JN_{++}$. Here we used the identity $\gamma N_+ = 2N_{++} + N_{+-}$ which one derives counting the number of lines drawn from every up spin to its nearest neighbors. The partition function of the Ising model can be written similarly to (148) with $z = \exp[(\gamma J - 2\mu H)/T]$. Further correspondence can be established: the pressure $P$ of the lattice gas can be expressed via the free energy per cite of the Ising model: $P \leftrightarrow -F/N + \mu H$ and the inverse specific volume $1/v = N_a/N$ of the lattice gas is equivalent to $N_+/N = (1 + M/\mu N)/2 = (1+\eta)/2$. We see that generally (for given $N$ and $T$) the lattice gas corresponds to the Ising model with a nonzero field $H$ so that the transition is generally of the first-order in this model. Indeed, when $H = 0$ we know that $\eta = 0$ for $T > T_c$ which gives a single point $v = 2$, to get the whole isotherm one needs to consider the nonzero $H$ i.e. the fugacity different from $\exp(\gamma J)$. In the same way, the solutions of the zero-field Ising model at $T < T_c$ gives us two values of $\eta$ that is two values of the specific volume for a given pressure $P$. Those two values, $v_1$ and $v_2$, precisely correspond to two phases in coexistence at the given pressure. Since $v = 2/(1+\eta)$ then as $T \to 0$ we have two roots $\eta_1 \to 1$ which correspond to $v_1 \to 1$ and $\eta_1 \to -1$ which corresponds to $v_1 \to \infty$. For example, in the mean field approximation (134) we get (denoting $B = \mu H$)

$$\begin{aligned} P &= B - \frac{\gamma J}{4}(1+\eta^2) - \frac{T}{2}\ln\frac{1-\eta^2}{4} \,, \quad B = \frac{\gamma J}{2} - \frac{T}{2}\ln z \,, \\ v &= \frac{2}{1+\eta} \,, \quad \eta = \tanh\left(\frac{B}{T} + \frac{\gamma J \eta}{2T}\right) \,. \end{aligned} \tag{149}$$

As usual in a grand canonical description, to get the equation of state one expresses $v(z)$ [in our case $B(\eta)$] and substitutes it into the equation for the pressure. On the figure, the solid line corresponds to $B = 0$ at $T < T_c$ where we have a first-order phase transition with the jump of the specific volume, the isotherms are shown by broken lines. The right figure gives the exact two-dimensional solution.



The mean-field approximation (149) is equivalent to the Landau theory near the critical point. In the variables $t = T - T_c, \eta = n - n_c$ the equation of state takes the form $p = P - P_c = bt + 2at\eta + 4C\eta^3$ with $C > 0$ for stability and $a > 0$ to have a homogeneous state at $t > 0$. In coordinates $p, \eta$ the isotherms at $t = 0$ (upper curve) and $t < 0$ (lower curve) look as follows:



The densities of the two phases in equilibrium, $\eta_1, \eta_2$ are given by the condition

$$\int_1^2 v\, dp = 0 \;\Rightarrow \int_1^2 \eta\, dp = \int_{\eta_1 1}^{\eta_2} \eta \left(\frac{\partial p}{\partial \eta}\right)_t d\eta = \int_{\eta_1}^{\eta_2} \eta \Big(2at + 12C\eta^2\Big)\, d\eta = 0\,, \quad (150)$$

where we have used $v = n^{-1} \sim n_c^{-1} - \eta n_c^{-2}$. We find from (150) $\eta_1 = -\eta_2 = (-at/2C)^{1/2}$. According to Clausius-Clapeyron equation (124) we get the latent heat of the transition $q \approx bT_c(\eta_1 - \eta_2)/n_c^2 \propto \sqrt{-t}$. We thus have the phase transition of the first order at $t < 0$. As $t \to -0$ this transition is getting close to the phase transitions of the second order. See Landau & Lifshitz, Sect. 152.

As $T \to T_c$ the mean-field theory predicts $1/v_1 - 1/v_2 \propto (T_c - T)^{1/2}$ while the exact Onsager solution gives $(T_c - T)^{1/8}$. Real condensation transition gives the power close 1/3. Also lattice theories give always (for any $T$) $1/v_1 + 1/v_2 = 1$ which is also a good approximation of the real behavior (the sum of vapor and liquid densities decreases linearly with the temperature increase but very slowly). One can improve the lattice gas model considering the continuous limit with the lattice constant going to zero and adding the pressure of the ideal gas.

Another equivalent model is that of the binary alloy that is consisting of two types of atoms. X-ray scattering shows that below some transition temperature there are two crystal sublattices while there is only one lattice at higher temperatures. Here we need the three different energies of inter-atomic interaction: $E = \epsilon_1 N_{11} + \epsilon_2 N_{22} + \epsilon_{12} N_{12} = (\epsilon_1 + \epsilon_2 - 2\epsilon_{12}) N_{11} + \gamma(\epsilon_{12} - \epsilon_2) N_1 + \gamma \epsilon_2 N/2$. This model described canonically is equivalent to the Ising model with the free energy shifted by $\gamma(\epsilon_{12} - \epsilon_2) N_1 + \gamma \epsilon_2 N/2$. We are interested in the case when $\epsilon_1 + \epsilon_2 > 2\epsilon_{12}$ so that it is indeed preferable to have alternating atoms and two sublattices may exist at least at low temperatures. The phase transition is of the second order with the specific heat observed to increase as the temperature approaches the critical value. Huang, Chapter 16 and Pathria, Chapter 12.

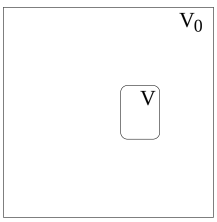
As we have seen, to describe the phase transitions of the second order near $T_c$ we need to describe strongly fluctuating systems. We shall study fluctuations more systematically in the next section and return to critical phenomena in Sects. 5.2 and 5.4.

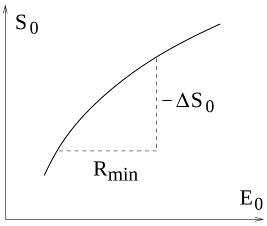
# 5 Fluctuations

In this Chapter, we first consider macroscopic fluctuations of thermodynamic parameters, then we treat spatially non-uniform large-scale fluctuations; we consider microscopic fluctuations in the next Chapter.

## 5.1 Thermodynamic fluctuations

Here we consider fluctuations from a thermodynamic perspective with an added knowledge that the probability is proportional to the phase volume. We briefly treated the subject in Section 2.4, here we develop a general treatment. We shall ask in particular whether the fluctuations of different quantities correlate with each other. Consider fluctuations of energy and volume of a given (small) subsystem. The probability of a fluctuation is determined by the (small) entropy change of the whole system $w \propto \exp(\Delta S_0)$ which is determined by the minimal work needed for a reversible creation of such a fluctuation: $T\Delta S_0 = -R_{min}$ - see the figure. Just the same, $R_{min}$ is the maximal work one can extract from the system by bringing it into equilibrium at a constant total entropy. For example, if the fluctuation is that the subsystem starts moving as a whole with the velocity $v$ then the minimal work is the kinetic energy $Mv^2/2$, so that the probability of such a fluctuation is $w(v) \propto \exp(-Mv^2/2T)$. Generally, we can express $R_{min}$ via the fluctuations $\Delta S, \Delta E, \Delta V$ related to the subsystem. To do that, note that the energy change of the subsystem, $\Delta E = R + P_0\Delta V_0 - T_0\Delta S_0$, is equal to the work $R$ done on it (by something from outside the system) plus the work done by the rest of the system $P_0\Delta V_0 = -P_0\Delta V$ plus the heat received from the rest of the system $-T_0\Delta S_0$. The signs are correct: positive $\Delta V_0$ means contraction of the subsystem i.e. positive work done on it, and positive $\Delta S_0$ means that the subsystem gives away heat. Minimal work corresponds to $\Delta S_0 = -\Delta S$ so that $R_{min} = \Delta E + P_0\Delta V - T_0\Delta S$. In calculating variations we also assume $P, T$ equal to their mean values which are $P_0, T_0$. Stress that we only assumed the subsystem to be small i.e. $\Delta S_0 \ll S_0$, $E \ll E_0$, $V \ll V_0$ while fluctuations can be substantial, i.e. $\Delta E$ can be comparable with $E$.





If, in addition, we assume the fluctuations to be small ($\Delta E \ll E$) we can expand $\Delta E(S,V)$ up to the first non-vanishing terms. The linear terms of this expansion cancel $P\Delta V - T\Delta S$, while the quadratic terms give the Gaussian probability distribution:

$$R_{min} = \Delta E + P\Delta V - T\Delta S = [E_{SS}(\Delta S)^2 + 2E_{SV}\Delta S\Delta V + E_{VV}(\Delta V)^2]/2$$
$$= (1/2)(\Delta S\Delta E_S + \Delta V\Delta E_V) = (1/2)(\Delta S\Delta T - \Delta P\Delta V)\,. \tag{151}$$

Written in such a way, it shows a sum of contributions of hidden and mechanical degrees of freedom. To avoid misunderstanding, note that $\Delta P$ is not the difference $P - P_0$ (assumed zero) but $\Delta P = P - \bar{P} = P_0$ i.e. deviation from the mean (equilibrium) value. Of course, only two variables are independent. From that general formula one obtains different cases by choosing different pairs of independent variables. In particular, choosing an extensive variable from one pair and an intensive variable from another pair (i.e. either $V, T$ or $P, S$), we get cross-terms canceled because of the Maxwell identities like $(\partial P/\partial T)_V = (\partial S/\partial V)_T = \partial^2 F/\partial T\partial V$. That means the absence of cross-correlation i.e. respective quantities fluctuate independently[16]: $\langle \Delta T\Delta V\rangle = \langle \Delta P\Delta S\rangle = 0$. Indeed, choosing $T$ and $V$ as independent variables we must express

$$\Delta S = \left(\frac{\partial S}{\partial T}\right)_V \Delta T + \left(\frac{\partial S}{\partial V}\right)_T \Delta V = \frac{C_v}{T}\Delta T + \left(\frac{\partial P}{\partial T}\right)_V \Delta V\,,$$
$$\Delta P = \left(\frac{\partial P}{\partial T}\right)_V \Delta T + \left(\frac{\partial P}{\partial V}\right)_T \Delta V\,,$$

and obtain

$$w \propto \exp\left[-\frac{C_v}{2T^2}(\Delta T)^2 + \frac{1}{2T}\left(\frac{\partial P}{\partial V}\right)_T (\Delta V)^2\right]\,. \tag{152}$$

---

[16] Remind that the Gaussian probability distribution $w(x,y) \sim \exp(-ax^2 - 2bxy - cy^2)$ corresponds to the second moments $\langle x^2\rangle = 2c/(ac-b^2)$, $\langle y^2\rangle = 2a/(ac-b^2)$ and to the cross-correlation $\langle xy\rangle = 2b/(b^2 - ac)$.

Mean squared fluctuation of the temperature

$$\langle (\Delta T)^2 \rangle = T^2/C_v\,.$$

In particular, with that accuracy one can define temperature for an isolated system. Remind that at low temperatures, the specific heat decreases when temperature decreases due to "freezing-out" of the degrees of freedom. For instance, $C_v \propto T$ for solids (due to electrons) and $C_v \propto T^{3/2}$ for the gas of cold atoms. In all cases the relative fluctuations of the temperature, $\langle (\Delta T)^2 \rangle / T^2 = 1/C_v$, increases when $T \to 0$, which makes difficult the life of experimentalists dealing with low temperatures.

Mean squared fluctuation of the volume (for a given number of particles),

$$\langle (\Delta V)^2 \rangle = -T(\partial V/\partial P)_T\,,$$

gives the fluctuation of the specific volume

$$\langle (\Delta v)^2 \rangle = N^{-2} \langle (\Delta V)^2 \rangle$$

which can be converted into the mean squared fluctuation of the number of particles in a *fixed* volume:

$$\begin{aligned}
&\Delta v = \Delta \frac{V}{N} = V \Delta \frac{1}{N} = -\frac{V \Delta N}{N^2}\,,\\
&\langle (\Delta N)^2 \rangle = -T \frac{N^2}{V^2} \left(\frac{\partial V}{\partial P}\right)_T = NT \left(\frac{\partial n}{\partial P}\right)_{T,N}\\
&= nT \left(\frac{\partial N}{\partial P}\right)_{T,V} = T \left(\frac{\partial P}{\partial \mu}\right)_{T,V} \left(\frac{\partial N}{\partial P}\right)_{T,V} = T \left(\frac{\partial N}{\partial \mu}\right)_{T,V}\,. \qquad (153)
\end{aligned}$$

We expressed $n = (\partial P/\partial \mu)_{T,V}$ using $-d\Omega(T,\mu) = d(PV) = SdT + Nd\mu$. The mean squared fluctuations of $N, V, P$ (and similar thermodynamic quantities having mechanical nature) turn to zero linearly as $T \to 0$. The relative fluctuation, $\langle (\Delta N)^2 \rangle / N^2 = N^{-1} T(\partial n/\partial P)$, generally decreases as $1/N$ in the thermodynamic limit. This is not true near the critical point where $\partial n/\partial P$ is large. As we discussed in Sect. 4.4.1, $\partial P/\partial n$ cannot be exactly zero in a finite system. We then expect at criticality $\partial P/\partial n \propto 1/N$ and the relative fluctuation to be independent of $N$.

For a classical **ideal gas** with $V = NT/P$ (153) gives $\langle (\Delta N)^2 \rangle = N$. In this case, we can do more than considering small fluctuations (or large volumes). We can find the probability of large fluctuations, even comparable to the mean value $\bar{N} = N_0V/V_0$. The probability for $N$ (noninteracting) particles to be inside some volume $V$ out of the total volume $V_0$ is

$$\begin{aligned} w_N &= \frac{N_0!}{N!(N_0-N)!}\left(\frac{V}{V_0}\right)^N\left(\frac{V_0-V}{V_0}\right)^{N_0-N} \\ &\approx \frac{\bar{N}^N}{N!}\left(1-\frac{\bar{N}}{N_0}\right)^{N_0} \approx \frac{\bar{N}^N\exp(-\bar{N})}{N!} \ . \end{aligned} \tag{154}$$

Here we assumed that $N_0 \gg N$ and $N_0! \approx (N_0-N)!N_0^N$. Note that $N_0$ disappeared from (154). The distribution (154) is called Poisson law which takes place for independent events. It satisfies normalization $\sum_N w_N = 1$. Mean squared fluctuation is the same as for small fluctuations:

$$\begin{aligned} \langle(\Delta N)^2\rangle &= \langle N^2\rangle - \bar{N}^2 = \exp(-\bar{N})\sum_{N=1}\frac{\bar{N}^N N}{(N-1)!} - \bar{N}^2 \\ &= \exp(-\bar{N})\left[\sum_{N=2}\frac{\bar{N}^N}{(N-2)!} + \sum_{N=1}\frac{\bar{N}^N}{(N-1)!}\right] - \bar{N}^2 \\ &= \exp(-\bar{N})\left[\bar{N}^2\sum_{N=0}\frac{\bar{N}^N}{N!} + \bar{N}\sum_{N=1}\frac{\bar{N}^N}{N!}\right] - \bar{N}^2 = \bar{N} \ . \end{aligned} \tag{155}$$

This formula also can be used for reading the dependence on the volume. Recall that the measurement volume is proportional to $\bar{N}$. In particular, the probability that a given volume is empty ($N = 0$) decays exponentially with the volume. This is the same as the fact that the probability of a radioactive decay not happening during time $t$ decays exponentially with $t$. On the other hand, the probability to cram more than average number of particles into the volume decays with $N$ in a factorial way, i.e. faster than exponential: $w_N \propto \exp[-N\ln(N/\bar{N})]$. One can check that near the maximum, at $|N-\bar{N}| \ll \bar{N}$, the Poisson distribution coincide with the Gaussian distribution: $w_N = (2\pi\bar{N})^{-1/2}\exp[-(N-\bar{N})^2/2\bar{N}]$.

Of course, real molecules do interact, so that the statistics of their density fluctuations deviate from the Poisson law, particularly near the critical point where the interaction energy is getting comparable to the entropy contribution into the free energy. Spatial correlation of fluctuations in the interacting systems is considered in the next Section 5.2.

Landau & Lifshitz, Sects. 20, 110–112, 114.

## 5.2 Spatial correlation of fluctuations

We now expand the phenomenological consideration to systems with interaction and discuss a spatial correlation of fluctuations. Our particular interest is when we can have a long-range order in a system, and when, on the contrary, fluctuations destroy such an order. As always, we consider systems with a short radius of interaction and treat correlations on the scales much larger than this radius. It is clear that whether fluctuations destroy a long-range order must depend on the dimensionality. For example, if correlations are along the line, it is easier to break the order, than when it is supported from different directions. We treat 3d here and then consider lower dimensionalities.

Let us now regularly consider the correlation function of fluctuations at the distances much larger than the radius of interaction. Consider, for instance, the concentration $n = N/V$. Since the fluctuations of $n$ and $T$ are independent, we assume $T =$const so that the minimal work is the change in the free energy, which we again expand to the quadratic terms

$$w \propto \exp(-\Delta F/T)\,, \quad \Delta F = \frac{1}{2}\int \phi(r_{12})\Delta n(\mathbf{r}_1)\Delta n(\mathbf{r}_2)\, dV_1 dV_2\,. \tag{156}$$

Here $\phi$ is the second (variational) derivative of $F$ with respect to $n(r)$. Since we restrict ourselves by the quadratic terms in $\Delta n$ then the probability distribution is Gaussian (in field theory, it is called the case of the free field). We assumed translational invariance so that $\phi(r_{12})$ depends only on a relative distance. In this case, the free energy can be diagonalized, i.e. presented as a single (not double) sum, by the Fourier transform:

$$\Delta n(r) = \sum_k \Delta n_k e^{i\mathbf{kr}}\,,\ \Delta n_k = \frac{1}{V}\int \Delta n(r) e^{-i\mathbf{kr}}\, d\mathbf{r}\,,\ \phi(k) = \int \phi(r) e^{-i\mathbf{kr}}\, d\mathbf{r}\,.$$

$$\Delta F = \frac{V}{2}\sum_k \phi(k)|\Delta n_k|^2\,,$$

which corresponds to a Gaussian probability distribution of independent variables - amplitudes of the harmonics. The mean squared fluctuation is as follows

$$\langle|\Delta n_k|^2\rangle = \frac{T}{V\phi(k)}\,. \tag{157}$$

The free energy must increase when the concentration is modulated in space. That means that $\phi(k)$ must be growing function of $k$. Therefore, the largest fluctuations correspond to small $k$, where we can use the expansion called the Ornshtein-Zernicke approximation

$$\phi(k) \approx \phi_0 + 2gk^2 \,. \tag{158}$$

The coefficient $g$ is assumed positive. Such expansion presumes short-range interaction which makes large-scale limit regular. From the previous section, $\phi_0(T) = n^{-1}(\partial P/\partial n)_T$.

Not only for the concentration but also for other quantities (like magnetization, for instance), (158) is a general form at small $k$. Let us show how it appears in a continuous limit of a lattice model. Consider some variable $\eta_{\mathbf{i}}$ defined on the sites $\mathbf{i}$ of, say, cubic lattice with the spacing $a$. The free energy contains on-cite term and the interaction between neighboring cites $\mathbf{i}$ and $\mathbf{i} \pm \mathbf{e}_j$ $(j = 1, \dots d)$:

$$\Delta F = \frac{1}{2}\sum(\lambda^{-1}\eta_{\mathbf{i}}^2 - \eta_{\mathbf{i}}\eta_{\mathbf{i}\pm\mathbf{e}_j}/2d) \,. \tag{159}$$

To obtain the continuous limit, we introduce the coordinates $\mathbf{x} = \mathbf{i}a$ and re-scale the variable: $\eta_{\mathbf{i}} = \sqrt{2d}a^{d/2-1}\eta(\mathbf{x})$ to eliminate $a$ from the gradient term below. We then denote $\lambda^{-1} - 1 = m^2a^2/2d$ and obtain

$$\Delta F = \frac{1}{2}\int \frac{d\mathbf{x}2da^d}{a^d}\left\{\lambda^{-1}\eta^2(\mathbf{x}) - \eta(\mathbf{x})\frac{1}{d}\sum_{j=1}^{d}\left[\eta(\mathbf{x}) + \frac{1}{2}\partial_j^2\eta(\mathbf{x}) + \dots\right]\right\}$$
$$= \int\frac{d\mathbf{x}}{2}\,\eta(\mathbf{x})(m^2 - \Delta)\eta(\mathbf{x}) + \mathrm{O}(a^4) = \int\frac{d\mathbf{x}}{2}(m^2\eta^2 + |\nabla\eta|^2) + \mathrm{O}(a^4) \,. \tag{160}$$

That is we obtained (158) with $\phi_0(T) = m^2$ and $2g = 1$. Generally, the gradient term can be neglected when $\eta$ changes on the scale far exceeding $a$, but this is not the case near phase transition when $\lambda \to 1$ and $\phi_0 = m^2 \to 0$.

Making the inverse Fourier transform of (157) with (158), we find (the large-scale part of) the pair correlation function of the concentration in 3d:

$$\langle \Delta n(0)\Delta n(r)\rangle = \frac{1}{V}\int d\mathbf{r}_1 \Delta n(\mathbf{r}_1)\Delta n(\mathbf{r}_1 + \mathbf{r}) = \sum_k |\Delta n_k|^2 e^{i\mathbf{kr}}$$
$$= \int |\Delta n_k|^2 e^{i\mathbf{kr}}\frac{Vd^3k}{(2\pi)^3} = \int_0^\infty \frac{Vk^2dk}{(2\pi)^2}|\Delta n_k|^2 \int_0^\pi \sin\theta d\theta e^{ikr\cos\theta}$$
$$= \int_0^\infty \frac{T}{\phi_0 + 2gk^2}\frac{e^{ikr} - e^{-ikr}}{ikr}\frac{k^2dk}{(2\pi)^2} = \frac{T\exp(-r/r_c)}{8\pi gr} \,. \tag{161}$$

One can derive that by using $(\kappa^2 - \Delta)\exp(-\kappa r)/r = 4\pi\delta(\mathbf{r})$ or directly: expand the integral to $-\infty$ and then close the contour in the complex upper half plane for the first exponent and in the lower half plane for the second exponent so that the integral is determined by the respective poles $k = \pm i\kappa = \pm ir_c^{-1}$. We defined the correlation radius of fluctuations $r_c = [2g(T)/\phi_0(T)]^{1/2} = 1/m$. Far from any phase transition, the correlation radius is typically the mean distance between molecules i.e. $a$.

Note that (160,161) require that $\phi_0(T) = m^2(T)$ is non-negative. Near the critical point, $\phi_0(T)$ decreases, as $\phi_0(T) \propto A(T)/V = \alpha(T - T_c)$ in the Landau theory, so that the correlation radius increases. To generalize the Landau theory for inhomogeneous $\eta(\mathbf{r})$ one writes the thermodynamic potential uniting (126) and (158) as

$$F\{\eta(\mathbf{r})\} = \int d\mathbf{r}[g|\nabla\eta|^2 + \alpha(T - T_c)\eta^2 + b\eta^4] \ . \tag{162}$$

The thermodynamic potential is now a functional since it depends on the whole function $\eta(\mathbf{r})$. Addition of the extra term $g|\nabla\eta|^2$ means that having an inhomogeneous state costs us extra value of the thermodynamic potential. We accounted for that extra cost using only the first spatial derivative, which means that this expression is obtained by effectively averaging the potential over the scales large comparing to the inter-atom distances (but small comparing to the correlation radius, so that this expression makes sense only not far from the critical temperature). We again assume that only the coefficient at $\eta^2$ turns into zero at the transition. The correlation radius diverges at the transition[17]: $r_c(T) = \sqrt{g/\alpha(T - T_c)} = a\sqrt{T_c/(T - T_c)}$. Here we expressed $g \simeq \alpha T_c a^2$ via the correlation radius $a$ far from the transition. We must now estimate the typical size of fluctuations for the volume $r_c^3$: $\langle(\Delta\eta)^2\rangle \simeq T/A \simeq T/\phi_0 r_c^3$. As any thermodynamic approach, the Landau theory is valid only if the mean square fluctuation on the scale of $r_c$ is much less than $\bar{\eta}$, so-called Ginzburg-Levanyuk criterium:

$$\frac{T_c}{2\alpha(T - T_c)r_c^3} \ll \frac{\alpha(T - T_c)}{b} \ \Rightarrow \ \frac{T - T_c}{T_c} \gg \frac{b^2}{\alpha^4 T_c^2 a^6} \equiv \left(\frac{r_i}{a}\right)^6 . \tag{163}$$

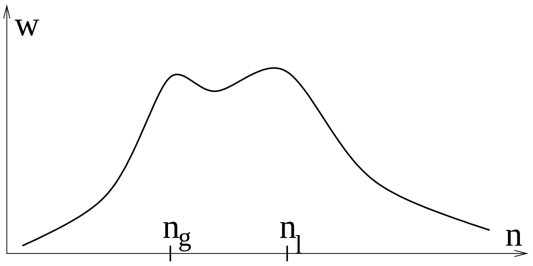
We introduced $r_i^3 = b/\alpha^2 T_c$ which can be interpreted as the effective volume of interaction per degree of freedom: if we divide the energy density of in-

---

[17] The correlation radius generally stays finite at a first-order phase transition. The divergence of $r_c$ at $T \to T_c$ means that fluctuations are correlated over all distances so that the whole system is in a unique critical phase at a second-order phase transition.

teraction $b\eta^4 \simeq b(\alpha T_c/b)^2$ by the energy of a single degree of freedom $T_c$ we get the number of degrees of freedom per unit volume i.e. $r_i^{-3}$. Since the Landau theory is built at $T - T_c \ll T_c$ then it has validity domain only when $r_i/r_{c0} \ll 1$ which often takes place (in superconductors, this ratio can be less than $10^{-2}$, since the interaction scale is a lattice size and the correlation radius is the size of the electron-hole Cooper pair). In a narrow region near $T_c$ fluctuations dominate. We thus must use the results of the Landau theory only outside the fluctuation region, in particular, the jump in the specific heat $\Delta C_p$ is related to the values on the boundaries of the fluctuation region; inside the region $C_p$ can change much, even diverge as we have seen.

Landau theory predicts $r_c \propto (T - T_c)^{-1/2}$ and the correlation function approaching the power law $1/r$ as $T \to T_c$ in 3d. Those scalings are valid under the condition (163), that is not very close to $T_c$. As one can see from the exact solution of 2d Ising model, the true asymptotic at $T \to T_c$ (i.e. inside the fluctuation region) are different: $r_c \propto (T - T_c)^{-1}$ and $\varphi(r) = \langle \sigma(0)\sigma(r) \rangle \propto r^{-1/4}$ at $T = T_c$ in that case. Yet the fact of the radius divergence remains. It means the breakdown of the Gaussian approximation for the probability of fluctuations since we cannot divide the system into independent subsystems. Indeed, far from the critical point, the probability distribution of the density has two approximately Gaussian peaks, one at the density of liquid $n_l$, another at the density of gas $n_g$. As we approach the critical point and the distance between peaks is getting comparable to their widths, the distribution is non-Gaussian. In other words, one needs to describe a strongly interaction system near the critical point which makes it similar to other great problems of physics (quantum field theory, turbulence).



## 5.3 Different order parameters and space dimensionalities

Phenomenological treatment in the previous section have shown that above the critical point (in the symmetric phase) $\phi_0(T)$ is positive and the critical radius $r_c \propto \phi_0^{-1/2}$ is finite. At criticality $\phi_0 = 0$ and the energy of the perturbation goes to zero with the wave number. What will be below the critical temperature in the symmetry-broken phase when $\phi_0 < 0$? Of course, that would be silly to assume that we have imaginary $r_c$ and oscillatory correlation function according to (161) - to describe fluctuations we now need to expand the free energy near the new minimum of the thermodynamic potential, where the mean value of the order parameter $\bar\eta$ is nonzero. At $T < T_c$ the second derivative of the thermodynamic potential is finite positive at $\bar\eta$ so the correlation radius must be finite, right? Actually, this is true only when the order parameter is a real scalar like in the Ising model or in the simple version of the Landau theory we considered so far. This is because in the Ising model the symmetry was discrete (up-down, present-absent) and so was described by a scalar real order parameter. Whether the correlation radius is finite or infinite depends on the dimensionality of the order parameter, which is determined by the type of the symmetry. What if the symmetry broken by the phase transition is continuous?

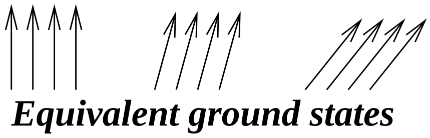
### 5.3.1 Goldstone mode and Mermin-Wagner theorem

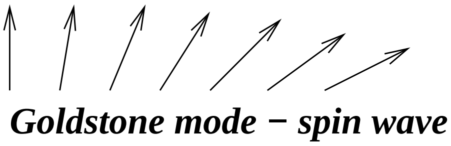
Here we consider the case when the Hamiltonian is invariant under $O(n)$ rotations. Now to break a symmetry we need to choose a direction in $n$-dimensional space, and the order parameter is a vector $(\eta_1, \ldots, \eta_n)$. The rotational symmetry is continuous since one can rotate by an infinitesimally small angle. Then the analog of the Landau thermodynamic potential must have a form $g\sum_i |\nabla\eta_i|^2 + \alpha(T-T_c)\sum_i \eta_i^2 + b\Big(\sum_i \eta_i^2\Big)^2$. Here we assumed that the interaction is short-range and used the Ornshtein-Zernicke approximation for the spatial dependence. When $T < T_c$ the minimum corresponds to breaking the $O(n)$ symmetry, for example, by taking the first component $\bar\eta = [\alpha(T_c - T)/2b]^{1/2}$ and the other components zero. This is an example of what is called *spontaneous symmetry breaking*: the symmetry of the thermodynamic potential (and of the most probable configuration) is less than the symmetry of the Hamiltonian. Thermodynamic potential is the log of the partition function which is the sum (integral in a continuous limit) over configurations of the exponent of the Hamiltonian. The symmetry of an integral can be less than the symmetry of the integrand if the integration measure (choice of configurations over which we sum) breaks the symmetry.

Considering fluctuations we put $([\alpha(T_c - T)/2b]^{1/2} + \eta_1, \eta_2, \ldots, \eta_n)$ and obtain the quadratic terms in the thermodynamic potential as follows:

$$g\sum_i |\nabla \eta_i|^2 + 4b\bar{\eta}^2\eta_1^2 + [\alpha(T - T_c) + 2b\bar{\eta}^2]\sum_i \eta_i^2 = g\sum_i |\nabla \eta_i|^2 + 2\alpha(T_c - T)\eta_1^2 .$$

That form means that only the longitudinal mode $\eta_1$ has a finite correlation length $r_c = [2\alpha(T_c - T)]^{-1/2}$. Almost uniform fluctuations of the transverse modes do not cost any free energy and thus have non-vanishing probability. This is an example of the Goldstone theorem which claims that whenever continuous symmetry is spontaneously broken then the mode must exist with the energy going to zero with the wavenumber. This statement is true beyond the mean-field approximation or Landau theory as long as the force responsible for symmetry breaking is short-range. For a spin system, the broken symmetry is rotational and the Goldstone mode is the excited state where the spin turns as the location changes, as shown in the Figure. That excitation propagates as a spin wave. For a solid, the broken symmetry is translational and the Goldstone mode is a phonon.





As we have seen, if the symmetry broken is continuous then even below the critical temperature we have an infinite correlation radius and fluctuations on all scales like at criticality. Indeed, Goldstone modes are necessarily excited by thermal fluctuations. Does it mean that there is effectively no long-range order even below $T_c$ since we can always meet arbitrary long fluctuation? The answer depends on the space dimensionality. Let us show that Goldstone modes destroy long-range order for $d \leq 2$. Indeed the integral (161) at $\phi_0 = 0$

describes fluctuations whose variance *grows* with the system size $L$:

$$\langle(\Delta\eta)^2\rangle \propto \int_{1/L}^{1/a} k^{-2} d^d k = \frac{a^{2-d} - L^{2-d}}{d-2} \ . \tag{164}$$

For $d > 2$ that growth saturates so the variance is independent of $L$ at the thermodynamic limit. On the contrary, the mean variance of the order parameter diverges when $L \to \infty$ at $d \le 2$. For example, $\langle(\Delta\eta)^2\rangle \propto \ln(L/a)$ in 2d. That means that in the thermodynamic limit the system is actually disordered: soft (Goldstone) modes with no energy price for long fluctuations ($\phi_0 = 0$) destroy long order (this statement is called Mermin-Wagner theorem). In exactly the same way phonons with $\omega_k \propto k$ make 2d crystals impossible: the energy of the lattice vibrations is proportional to the squared atom velocity (which is the frequency $\omega_k$ times displacement $u_k$), $T \simeq \omega_k^2 u_k^2$; that makes mean squared displacement proportional to $\langle u^2\rangle \propto \int d^d k u_k^2 = \int d^d k T/\omega_k^2 \propto L^{2-d}$ — in large enough samples the amplitude of displacement is getting comparable to the distance between atoms in the lattice, which means the absence of a long-range order.

Another example of the Goldstone mode destroying the long-range coherence is the case of the complex scalar order parameter $\Psi = \phi_1 + \imath\phi_2$ (say, the amplitude of the quantum condensate). In this case, the density of the Landau thermodynamic potential invariant with respect to the phase change $\Psi \to \Psi\exp(i\alpha)$ (called global gauge invariance) has the (so-called Landau-Ginzburg) form

$$F = g|\nabla\Psi|^2 + \alpha(T - T_c)|\Psi|^2 + b|\Psi|^4 \ . \tag{165}$$

At $T < T_c$, the (space-independent) minima of the potential form a circle in $\phi_1 - \phi_2$ plane: $|\Psi|^2 = \phi_1^2 + \phi_2^2 = \alpha(T_c - T)/2b = \phi_0^2$. Any ordered (coherent) state would correspond to some choice of the phase, say $\phi_1 = \phi_0, \phi_2 = 0$. For small perturbations around this state, $\Psi = \phi_0 + \varphi + \imath\xi$, the quadratic part of the potential takes the form $g|\nabla\xi|^2 + g|\nabla\varphi|^2 + 4b\phi_0^2|\varphi|^2$. Since there is no term proportional to $\xi^2$, then fluctuations of $\xi$ have an infinite correlation radius even at $T < T_c$; their variance (164) diverges with $L \to \infty$ at $d \le 2$, which means that phase fluctuations destroy coherence of 2d quantum condensate. That was a macroscopic way to see that the Bose-Einstein condensation is impossible in 2d. One can also see that from microscopic description, observing that the integral (87) at $z = 1$ diverges as $\int_0 pdp/p^2$, that is there is enough states at small momenta to accommodate all particles, no need to have macroscopic population at zero momentum.

Let us summarize. If the system Hamiltonian has a continuous symmetry, like rotation for spins, translation for atoms or gauge transformation for quantum wave function, this symmetry must be spontaneously broken to provide for an ordered state at low temperatures. In an ordered state, spins look predominantly in one directions in a ferromagnetic, atoms sit in preferred positions in the crystal lattice and the quantum phase is fixed. Then the Goldstone theorem claims that since those ordered states appeared by spontaneously breaking a continuous symmetry then the mode must exist whose energy goes to zero with the wavenumber (rotation or translation of the system as a whole costs no energy). This statement is true as long as the force responsible for symmetry breaking is short-range. For a spin system, the broken symmetry is rotational and the Goldstone mode is the excited state where the spin turns as the location changes. That excitation propagates as a spin wave. For a solid, the broken symmetry is translational and the Goldstone mode is a phonon. Long -wave variations of quantum phase also propagate as phonons.

So how actually the correlations look like at low temperatures at $d = 2$, when the correlation radius of fluctuations is infinite, on the one hand, but no long-range order exists, on the other hand? Consider the so-called XY model which describes a system of two-dimensional spins $\mathbf{s} = (s\cos\varphi, s\sin\varphi)$ with ferromagnetic interaction i.e. with the Hamiltonian

$$\mathcal{H} = -J\sum_{i,j}\mathbf{s}(i)\cdot\mathbf{s}(j) = -Js^2\sum_{i,j}\cos(\varphi_i - \varphi_j)\ . \tag{166}$$

At low enough temperature, we expect $|\varphi_i - \varphi_j| \ll 2\pi$, which allows one to approximate $\cos(\varphi_i - \varphi_j) \approx 1 - (\varphi_i - \varphi_j)^2/2$ and go into a continuous limit (spin-wave approximation):

$$\begin{aligned}\mathcal{H} &\approx -\frac{\gamma NJs^2}{2} + \frac{Js^2}{2}\sum_{i,j}|\varphi_i - \varphi_j|^2 \\ &\approx -\frac{\gamma NJs^2}{2} + \frac{Js^2}{2}\int|\nabla\varphi(\mathbf{r})|^2 d^2r\ . \end{aligned}\tag{167}$$

That Hamiltonian can be also written as a sum over Fourier harmonics $\mathcal{H} + \gamma NJs^2/2 = \sum_{\mathbf{k}} H_{\mathbf{k}} = Na^2Js^2\sum_k k^2|\varphi_k|^2/2 = \sum_k |\varphi_k|^2/2\langle|\varphi_k|^2\rangle$ with each term having an Ornstein-Zernike form. Here $a$ is the lattice spacing. There is no $\varphi_0^2$ term because of the O(2) rotational symmetry of the spins which corresponds to translational symmetry of the phases $\varphi$. In this (low-$T$)

approximation, the phases have Gaussian statistics with the pair correlation function which is logarithmic for distance $r \ll L$:

$$\begin{aligned}\langle\varphi(\mathbf{r})\varphi(\mathbf{0})\rangle &= \sum_k |\varphi_k|^2 e^{i\mathbf{kr}} = Na^2\int\langle|\varphi_k|^2\rangle e^{i\mathbf{kr}}\frac{d\mathbf{k}}{(2\pi)^2}\\ &= \int\frac{T}{Js^2k^2}J_0(kr)\frac{d^2k}{2\pi}\approx\frac{T}{2\pi Js^2}\ln(L/r)\,. \qquad (168)\end{aligned}$$

The variance is $\langle\varphi^2\rangle = (2\pi\beta Js^2)^{-1}\ln(L/a)$ and $\langle[\varphi(\mathbf{r})-\varphi(\mathbf{0})]^2\rangle = \langle[\delta\varphi(\mathbf{r})]^2\rangle = (4\pi\beta Js^2)^{-1}\ln(r/a)$ - this is true, of course, for $r \gg a$. Let us calculate now the correlation function between two spins distance $r$ apart:

$$\begin{aligned}&\langle\mathbf{s}(\mathbf{0})\cdot\mathbf{s}(\mathbf{r})\rangle/s^2 = \langle\cos[\varphi(\mathbf{r})-\varphi(\mathbf{0})]\rangle = Re\langle e^{i\varphi(\mathbf{r})-i\varphi(\mathbf{0})}\rangle\\ &= Re\int D\varphi e^{i\delta\varphi(\mathbf{r})-[\delta\varphi(\mathbf{r})]^2/2\langle[\delta\varphi(\mathbf{r})]^2\rangle} = e^{-\langle[\delta\varphi(\mathbf{r})]^2\rangle/2} = \left(\frac{r}{a}\right)^{-1/2\pi\beta Js^2}\,.\end{aligned}$$

One can do it using Fourier representation from the very beginning:

$$\begin{aligned}&Re\langle\exp[i\varphi(\mathbf{r})-i\varphi(\mathbf{0})]\rangle = \int d\varphi_k d\varphi_k^*\, Re\exp\left\{\sum_{\mathbf{k}}\left[i\varphi_k\left(e^{i(\mathbf{kr})}-1\right)-\beta H_{\mathbf{k}}\right]\right\}\\ &= \exp\left[-\frac{T}{Na^2Js^2}\sum_{\mathbf{k}}\frac{1-\cos(\mathbf{kr})}{k^2}\right] = \exp\left[-\frac{T}{Js^2}\int\frac{d\mathbf{k}}{(2\pi)^2}\frac{1-\cos(\mathbf{kr})}{k^2}\right]\\ &\approx \exp\left[-(2\pi\beta Js^2)^{-1}\ln(\pi r/a)\right] = \left(\frac{\pi r}{a}\right)^{-1/2\pi\beta Js^2}\,. \qquad (169)\end{aligned}$$

Here we used the formula of Gaussian integration

$$\int_{-\infty}^{\infty} d\varphi e^{-A\varphi^2/2+iJ\varphi} = \sqrt{2\pi/A}e^{-J^2A^{-1}/2}\,. \qquad (170)$$

We see that the correlation function falls of by a power law at however low temperature, and it does not approach constant as in a state with a long-range order. We thus conclude that there is no long-range order at all temperatures. Description looks practically the same for two dimensional crystals where the energy is proportional to the squared difference between the displacements of the neighboring atoms.

### 5.3.2 Berezinskii-Kosterlitz-Thouless phase transition

Still, a power-law decay of correlations (169), found at low $T$, is very much different from the exponential decay in a state with a finite correlation radius.

That is the state with a power-law decay formally corresponds to an infinite correlation radius. A long-range order is absent in that state yet a local order exists, which means that at sufficiently low temperatures superfluidity and superconductivity can exist in 2d films, and 2d crystals can support transverse sound (recall that longitudinal sound exists in fluids as well, so it is transverse sound which is a defining property of a solid). Remind that our consideration (166-169) was for sufficiently low temperatures. One then asks if the power-law behavior disappears and a finite correlation radius appears above some temperature.

Consider XY model at high temperatures. In this case, the spin correlation function can be written as follows:

$$\langle\cos[\varphi_{\mathbf{r}}-\varphi_{\mathbf{0}}]\rangle = Z^{-1}\prod_{i=1}^{N}\int\frac{d\varphi_i}{2\pi}\cos(\varphi_{\mathbf{r}}-\varphi_{\mathbf{0}})e^{\beta Js^2\sum_{i,j}\cos(\varphi_i-\varphi_j)}$$

$$\approx Z^{-1}\prod_{i=1}^{N}\int\frac{d\varphi_i}{2\pi}\cos[\varphi_{\mathbf{r}}-\varphi_{\mathbf{0}}]\prod_{i,j}[1+\beta Js^2\cos(\varphi_i-\varphi_j)]+o(\beta Js^2)\ .$$

In the lowest order in $\beta Js^2$, every bond between neighboring spins contribute either 1 or $\cos(\varphi_i-\varphi_j)$. Since the integral of cosine is zero, any set of bonds (i.e. graph) with a single bond sticking out of the site, which is neither **0** nor **r**, is zero. Only those graphs contribute that start at **0** and end at **r**. The leading contribution is from the shortest path of the length $r$. Internal and end-point integrations are respectively as follows:

$$\int\frac{d\varphi_i}{2\pi}\cos(\varphi_i-\varphi_{i-1})\cos(\varphi_{i+1}-\varphi_i)=\frac{1}{2}\cos(\varphi_{i+1}-\varphi_{i-1})\,,$$

$$\int\frac{d\varphi_{\mathbf{r}}d\varphi_{\mathbf{0}}}{(2\pi)^2}\cos^2(\varphi_{\mathbf{r}}-\varphi_{\mathbf{0}})=\frac{1}{2}\ .$$

Each cosine comes with $\beta Js^2$ factor so that each bond along the path contributes $\beta Js^2/2$:

$$\frac{\langle\mathbf{s}(\mathbf{0})\cdot\mathbf{s}(\mathbf{r})\rangle}{s^2}=\langle\cos(\varphi_{\mathbf{r}}-\varphi_{\mathbf{0}})\rangle=\left(\frac{\beta Js^2}{2}\right)^r=e^{-r/r_c}\text{ with }r_c=\ln^{-1}(2/\beta Js^2)\,.$$

We see that there is a finite correlation radius at high temperatures. To get a power-law decay of correlations at low temperatures we need to turn the correlation radius into infinity at some finite temperature. That must be a phase transition. What causes it? The answer is that, apart from

Goldstone modes, there is another set of excitations, which destroy an order even more effectively and lead to a finite correlation radius. For a 2d crystal, an order is destroyed by randomly placed dislocations while for a spin system or condensate by vortices. In considering spin waves in the XY model we neglected the periodic nature of the angle $\varphi$. Since fluctuations of the angle increase (logarithmically) at large distances, one needs to account for large variations of the angle and for topological effects due to angle periodicity. Vortex is a configurations of spins for which the angle rotates by $2\pi m$ when one goes around a closed path, see Figure 3. We can consider a large circle around the vortex center, then the variations of the angle $\varphi$ along the circle are smooth and can be described in the continuous approximation (167). The energy can be minimized not only by a uniform $\varphi$ =const but also by an inhomogeneous configuration, which is a vortex. Indeed, let us look at the variation:

$$\frac{\delta}{\delta\varphi(\mathbf{r})}\int|\nabla\varphi(\mathbf{r}')|^2\,d\mathbf{r}'=\Delta\varphi=0\ .$$

The vortex is the solution of this equation, which in the polar coordinates $r,\theta$ is simply $\varphi(r,\theta)=m\theta$ so that $(\nabla\varphi)_\theta=r^{-1}\partial\varphi/\partial\theta=m/r$ and the energy $E=(Js^2/2)\int|\nabla\varphi(\mathbf{r})|^2d^2r=\pi Js^2m^2\ln(L/a)$. We see that the energy of a single vortex is proportional to the logarithm of a sample size $L$. But the entropy $S$ associated with a single vortex is a log of the number of places one can put the center of the vortex, so it is also logarithmic in area and in size: $S=\ln(L/a)^2$. That means that there exists some sufficiently high temperature $T=\pi Js^2m^2/2$ when the energy contribution into the free energy $E-TS$ is getting less than the entropy contribution $TS$. The lowest temperature is for $m=1$: $T_{BKT}=\pi Js^2/2$ - it corresponds to the so-called Berezinskii-Kosterlitz-Thouless (BKT) phase transition from a low-$T$ where there are no isolated vortices to a high-$T$ state of free dislocations/vortices. Free vortices provide for a screening (a direct analog of the Debye screening of charges) so that the high-$T$ state has a finite correlation length equal to the Debye radius. Vortices are present at all temperatures. However, at low temperatures, one does not meet isolated vortices but only dipoles, i.e. pairs with $m=\pm1$, whose energy is determined by the inter-vortex distance and is independent of the system size $L$. Indeed, far from the dipole, the field $\nabla\varphi=\nabla\varphi_++\nabla\varphi_-$ decays as $r^{-2}$ so in the energy integral only shorter distances contribute, $E\propto\ln(r/a)$, where $r$ is the distance between the vortices. As opposite to initial spins or atoms that create crystal or condensate and interact locally, vortices and dislocations have a long-range

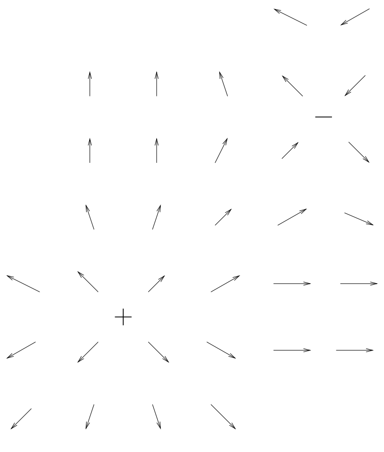


Figure 3: Two vortices with $m = 1$ and $m = -1$ in the XY model.

(logarithmic) interaction. Dipoles cannot screen which leads to power-law correlations at low $T$.

Let us now consider a different system: complex scalar field, which describes, for instance, quantum wave function of the set of atoms. Close to Bose-Einstein condensation in 2d it corresponds to the thermodynamic Landau-Ginzburg potential (165), which contains nonlinearity. That system can have vortices as well. A vortex is an inhomogeneous state that corresponds to an extremum of the potential. Varying the potential (165) with respect to $\Psi^*$ we get the equation

$$\delta F/\delta\Psi^* = -g\Delta\Psi + 2b\left(|\Psi|^2 - \phi_0^2\right)\Psi = 0 \ . \tag{171}$$

In polar coordinates, the Laplacian takes the form $\Delta = r^{-1}\partial_r r\partial_r + r^{-2}\partial_\theta^2$. Vortex corresponds to $\Psi(r,\theta) = A(r)\exp(im\theta)$ where $A$ can be chosen real and integer $m$ determines the circulation. Going around the vortex point in 2d (vortex line in 3d), the phase acquires $2m\pi$. The dependence on the angle $\theta$ makes the condensate amplitude turning into zero on the axis: $A \propto r^m$ as $r \to 0$. At infinity, we have a homogeneous condensate: $\lim_{r\to\infty}(A - \phi_0) \propto r^{-2}$. That this is a vortex is also clear from the fact that there is a current $J = i(\psi\nabla\psi^* - \psi^*\nabla\psi)$ and the velocity around it:

$$J_\theta \propto i\left(\psi\frac{\partial\psi^*}{r\partial\theta} - \psi^*\frac{\partial\psi}{r\partial\theta}\right) = \frac{A^2}{r} \ ,$$

so that the circulation is independent of the distance from the vortex. Second, notice that the energy of a vortex indeed diverges logarithmically with the sample size (as well as the energy of dislocations and 2d Coulomb charges): velocity around the vortex decays as $v \propto 1/r$ so that the kinetic energy diverges as $A^2 \int v^2 \, d^2r \propto \int d^2r/r^2$. A single vortex has its kinetic energy proportional to the logarithm of the sample size $L$. As we argued, since any charge can be placed in roughly $(L/a)^2$ different positions, where $a$ is the charge size (vortex core), then the entropy $S = 2\ln(L/a)$ is also logarithmic. The free energy is as follows:

$$F = E - TS = (A^2 - 2T)\ln(L/a) = \left[\frac{\alpha(T_c - T)}{2b} - 2T\right]\ln(L/a) \ .$$

We find that the second (BKT) phase transition happens at

$$T_{BKT} = T_c \frac{\alpha}{\alpha + 4b} \ ,$$

i.e. at the temperature less than $T_c$, when the condensate appears. At $T_c > T > T_{BKT}$, the entropy term $TS$ wins so that the minimum of $F$ corresponds to proliferation of isolated charges which provide for screening and an exponential decay of correlations.

The statistics of vortices and similar systems can be studied using a simple model of 2d Coulomb gas, which is described by the Hamiltonian

$$\mathcal{H} = -\sum_{i \neq j} m_i m_j \ln|\mathbf{r}_i - \mathbf{r}_j| \tag{172}$$

with $m_i = \pm 1$. Since we are interested only in the mutual positions of the vortices, we accounted only for the energy of their interaction. One can understand it better if we recall that the Hamiltonian describes also dynamics. The coordinates $\mathbf{r}_i = (x_i, y_i)$ of the point vortices (charges) satisfy the Hamiltonian equations

$$m_i \dot{x}_i = \frac{\partial \mathcal{H}}{\partial y_i} = m_i \sum_j m_j \frac{y_i - y_j}{|\mathbf{r}_i - \mathbf{r}_j|^2} \ , \quad m_i \dot{y}_i = -\frac{\partial \mathcal{H}}{\partial x_i} = m_i \sum_j m_j \frac{x_j - x_i}{|\mathbf{r}_i - \mathbf{r}_j|^2} \ ,$$

which simply tell that every vortex moves in the velocity field induced by all other vortices.

For example, the statistics of the XY-model at low temperatures can be considered with the Hamiltonian which is a sum of (167) and (172), and the partition function respectively a product of those due to spin waves and vortices. In the same way, condensate perturbations are sound waves and vortices. Goldstone modes(spin and sound waves) destroy long-range-order in 2d systems with a continuous symmetry. Since the Gaussian partition function of the waves is analytic, the phase transition can only originate from the vortices; near such phase transition the vortex degrees of freedom can be treated separately. Note that this is the phase transition without symmetry breaking.

In the canonical approach, one writes the Gibbs distribution $\mathcal{P} = Z^{-1}e^{-\beta\mathcal{H}}$. The partition function,

$$Z(\beta) = \int \exp(-\beta\mathcal{H})\,\Pi_i d^2 r_i = V \prod_{i\neq j} \int r_{ij}^{\beta m_i m_j} d^2 r_{ij}\ ,$$

behaves at $r_{ij} \to \epsilon \to 0$ as $\epsilon^{2(N-1)-\beta(N^2-N(N-2))/4}$. Here we took into account that a neutral system with $N$ vortices must have the following number of pairs: $N_{+-} = (N/2)^2$ and $N_{++} = N_{--} = N(N-2)/8$. That is the integral converges only for $\beta < 4(N-1)/N$. In the thermodynamic limit $N \to \infty$, the convergence condition is $\beta < 4$. When $\beta \to 4$ with overwhelming probability the system of vortices collapses, which corresponds to the BKT phase transition.

In the microcanonical approach one considers vortices in a domain with the area $A$. The density of states,

$$g(E, A) = \int \delta(E - H)\Pi_i dz_i\ , \tag{173}$$

must have a maximum since the phase volume $\Gamma(E) = \int^E g(E)dE$ increases monotonically with $E$ from zero to $A^N$. The entropy, $S(E) = \ln g(E)$, therefore, gives a negative temperature $T = g/g'$ at sufficiently high energy. Since $g(E, A) = A^N g(E', 1)$ where $E' = E + 2\ln A \sum_{i<j}^N m_i m_j$ we can get the equation of state

$$P = T\left(\frac{\partial S}{\partial A}\right)_E = \frac{NT}{A}\left(1 + \frac{1}{2NT}\sum_{i<j}^N m_i m_j\right)\ . \tag{174}$$

For a neutral system with $m_i = \pm 1$ we find $\sum_{i<j}^N m_i m_j = N_{++} + N_{--} - N_{+-} = -N/2$ and

$$P = \frac{NT}{A}\left(1 - \frac{\beta}{4}\right)\ , \tag{175}$$

which shows that pressure turns to zero at the BKT transition temperature.

Note that the initial system had a short-range interaction, which allowed to use the potential (165) in the long-wave limit. In other words, particles that create condensate interact locally. However, condensate vortices have a long-range interaction (172). While BKT transition may seem exotic and particular for 2d, the lesson is quite general: when there are entities (like vortices, monopoles, quarks) whose interaction energy depends on the distance similarly to entropy then the phase transition from free entities to pairs is possible at some temperature. In particular, it is argued that a similar transition took place upon cooling of the early universe in which quarks (whose interaction *grows* with the distance, just like entropy) bind together to form the hadrons and mesons we see today.

See Kardar, Fields, sect. 8.2.

### 5.3.3 Higgs mechanism

Let us stress that the previous subsection treated the condensate made of particles with local interaction. The physics is different, however, if the condensate is made of charged particles which can interact with the electromagnetic field, defined by the vector potential $\mathbf{A}$ and the field tensor $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$. That corresponds to the thermodynamic potential $|(\nabla - \imath e\mathbf{A})\Psi|^2 + \alpha(T-T_c)|\Psi|^2 + b|\Psi|^4 - F_{\mu\nu}F^{\mu\nu}/4$, which is invariant with respect to the local (inhomogeneous) gauge transformations $\Psi \to \Psi\exp[\imath\varphi(\mathbf{r})]$, $\mathbf{A} \to \mathbf{A} + \nabla\varphi/e$ for any differentiable function $\varphi(\mathbf{r})$. Note that the term $F_{\mu\nu}F^{\mu\nu} = (\nabla\times\mathbf{A})^2$ does not change when one adds gradient to $\mathbf{A}$. The equilibrium state is $\Psi_0 = \phi_0 + \exp[\imath\varphi/\phi_0]$ and $\mathbf{A}_0 = \nabla\varphi/e\phi_0$, absorbing also the phase shift into the vector potential. Perturbations can now be defined as $\Psi = (\phi_0 + h)\exp[\imath\varphi/\phi_0]$ and $\mathbf{A}_0 + \mathbf{A}$. The quadratic part of the thermodynamic potential is now

$$|\nabla h|^2 + \alpha(T-T_c)h^2 + e^2\phi_0^2A^2 - F_{\mu\nu}F^{\mu\nu}/4\,, \qquad (176)$$

i.e. the correlation radii of both modes $h, \mathbf{A}$ are finite. The correlation radius of the gauge-field mode is $r_c = 1/2e\phi_0 = \sqrt{b/\alpha(T_c - T)}$. This case does not belong to the validity domain of the Mermin-Wagner theorem since the Coulomb interaction is long-range.

Thinking dynamically, if the energy of a perturbation contains a nonzero homogeneous term quadratic in the perturbation amplitude, then the spectrum of excitations has a gap at $k \to 0$. For example, the (quadratic part of the) Hamiltonian $\mathcal{H} = g|\nabla\psi|^2 + \phi_0^2|\psi|^2$ gives the dynamical equation $\imath\dot{\psi} = \delta\mathcal{H}/\delta\psi* = -g\Delta\psi + \phi_0^2\psi$, which gives $\psi \propto \exp(\imath kx - \imath\omega_k t)$ with $\omega_k = \phi_0^2 + gk^2$. Therefore, what we have just seen is that the Coulomb interaction leads to a gap in the plasmon spectrum in superconducting transition. In other words, photon is massless in a free space but has a mass if it has to move through the condensate of Cooper pairs in a superconductor (Anderson, 1963). Another manifestation of this simple effect comes via the analogy between quantum mechanics and statistical physics (which we discuss in more detail in the subsection 6.1). On the language of relativistic quantum field theory, the system Lagrangian plays a role of the thermodynamic potential and the vacuum plays a role of the equilibrium. Here, the presence of a constant term means a finite energy ($mc^2$) at zero momentum, which means a finite mass. Interaction with the gauge field described by (176) is called Higgs mechanism by which particles acquire mass in quantum field theory, the excitation analogous to that described by $\mathbf{A}$ is called vector Higgs boson (Nambu, 1960; Englert, Brout, Higgs, 1964). What is dynamically a finite inertia is translated statistically into a finite correlation radius. When you hear about the infinite correlation radius, gapless excitations or massless particles, remember that people talk about the same phenomenon. See also Kardar, Fields, p. 285-6.

### 5.3.4 Impossibility of long-range order in 1d

The Goldstone mode destroys a long-range order in two dimensions and even more so in one dimension. The mode, however, does not exist if the order parameter cannot change continuously but takes only discrete values as in the Ising model. In this case, ferromagnetic order appears at low temperatures in two and higher dimensions but in 1d no temperature is low enough to have a nonzero magnetization $N\langle\sigma\rangle$. Is that a general property of one-dimensional systems?

The state of lowest energy has all spins parallel. The first excited state correspond to one spin flip and has an energy higher by $\Delta E = \gamma J$, where $\gamma$ is the number of nearest neighbors ($\gamma = 2d$ for a cubic lattice). The concentration of such opposite spins is proportional to $\exp(-\gamma J/T)$ and is low at low temperatures so that $\langle\sigma\rangle \approx 1$ and the magnetization is close to $N$. In one dimension, however, the lowest excitation is not the flip of one spin (energy $2J$) but flipping all the spins to the right or left from some site (energy $J$). Again the mean number of such flips is $N\exp(-J/T)$ but now long-range is determined by comparing this number to unity, not to the total number of spins. In sufficiently long chain the number $N\exp(-J/T)$ is larger than unity. That means that long chains consist of pieces with different spin signs and the mean magnetization is zero. Note that short pieces with $N < \exp(J/T)$ are magnetized. Yet the ferromagnetism of the whole lattice in the thermodynamic limit is possible only starting from two dimensions and only without the Goldstone mode.

It is physically natural that fluctuations has much influence in one dimension: it is enough to have a fluctuation in a domain exceeding the radius of interaction to loose completely the information of the order. It is thus not surprising that long-range order is impossible in one-dimensional systems with short-range interaction.

That argument can be generalized for arbitrary systems with the short-range interaction in the following way (Landau, 1950; Landau & Lifshitz, Sect. 163): assume we have $n$ contact points of two different phases and that every point costs the energy $\epsilon$. Those points add $n\epsilon - TS$ to the thermodynamic potential. The entropy is $\ln C_L^n$ where $L$ is the length of the chain. Evaluating entropy at $1 \ll n \ll L$ we get the addition to the potential $n\epsilon - Tn\ln(eL/n)$. The derivative of the thermodynamic potential with respect to $n$ is thus $\epsilon - T\ln(L/n)$ and it is negative for sufficiently small $n/L$. That means that one decreases the thermodynamic potential creating the mixture of two phases all the way until the derivative comes to zero which happens at $L/n = \exp(\epsilon/T)$ — this length can be called the correlation scale of fluctuations and it is always finite in 1d at a finite temperature as in a disordered state. We then expect the spin-spin correlation function in 1d to behave as $\langle\sigma(0)\sigma(r)\rangle = exp[-r\exp(\epsilon/T)]$. Let us stress that for the above arguments it is important that the ground state is non-degenerate so that the first excited state has a higher energy (degeneracy leads to criticality).

Let us conclude this section by restating the relation between the space dimensionality and the possibility of phase transition and ordered state for a short-range interaction. At $d \geq 3$ an ordered phase is always possible for a continuous as well as discrete broken symmetry. For $d = 2$, one type of a possible phase transition is that which breaks a discrete symmetry and corresponds to a real scalar order parameter like in Ising model; another type is the BKT phase transition between power-law and exponential decrease of correlations. At $d = 1$, neither symmetry breaking nor ordered state is possible for system with a short-range interaction.

Landau & Lifshitz, Sects. 116, 152.

## 5.4 Universality classes and renormalization group

Since the correlation radius diverges near the critical point, then fluctuations of all scales (from the lattice size to $r_c$) contribute the free energy. One therefore may *hope* that the particular details of a given system (type of atoms, their interaction, etc) are unimportant in determining the most salient features of the phase transitions, what is important is the space dimensionality and which type of symmetry is broken — for instance, whether it is described by scalar, complex or vector order parameter. Those salient features must be related to the nature of singularities that is to the critical exponents which govern the power-law behavior of different physical quantities as functions of $t = (T - T_c)/T_c$ and the external field $h$. Every physical quantity may have its own exponent, for instance, specific heat $C \propto t^{-\alpha}$, order parameter $\eta \propto (-t)^{\beta}$ and $\eta \propto h^{1/\delta}$, susceptibility $\chi \propto t^{-\gamma}$, correlation radius $r_c \propto t^{-\nu}$, the pair correlation function $\langle \sigma_i \sigma_j \rangle \propto |i - j|^{2-d-\eta}$, etc. Only two exponents are independent since all quantities must follow from the free energy which is a function of two thermodynamic parameters, for instance $F(t, h)$. Moreover, a picture of fluctuations at all scales suggests the *scaling hypothesis*: if we re-scale the lengthes by the factor $k$ then one can find such numbers $a, b$ that the free energy is transformed under re-scaling of arguments as follows: $F(k^a t, k^b h) = k^d F(t, h)$. In other words, a change in $t$ can be compensated up to an overall factor by a change in $h$. But that means that the function $F(t, h)$ actually depends on one variable, rather than two:

$$F(t, h) = t^{d/a} g(h/t^{b/a}) \ . \tag{177}$$

One can now express the quantities of interest via the derivatives at $h = 0$: $C = \partial^2 F/\partial t^2$, $\eta = \partial F/\partial h$, $\chi = \partial^2 F/\partial h^2$ and relate $\beta = (d - b)/a$ etc.

A general formalism which describes how to make a coarse-graining to keep only most salient features in the description is called the renormalization group (RG). It consists in subsequently eliminating small-scale degrees of freedom and looking for fixed points of such a procedure. Next two subsections present two variants of RG, discrete and continuous.

### 5.4.1 Block spin transformation

For Ising model, to eliminate small-scale degrees of freedom, we divide all the spins into groups (blocks) with the side $k$ so that there are $k^d$ spins in every block ($d$ is space dimensionality). We then assign to any block a new variable $\sigma'$ which is $\pm 1$ when respectively the spins in the block are predominantly up or down. We *assume* that the phenomena very near critical point can be described equally well in terms of block spins with the energy of the same form as original, $E' = -h'\sum_i \sigma'_i + J'/4\sum_{ij}(1-\sigma'_i\sigma'_j)$, but with different parameters $J'$ and $h'$. Let us demonstrate how it works using 1d Ising model with $h=0$ and $J/2T \equiv K$. Let us transform the partition function $\sum_{\{\sigma\}} \exp\Big[K\sum_i \sigma_i\sigma_{i+1}\Big]$ by the procedure (called decimation[18]) of eliminating degrees of freedom by ascribing (undemocratically) to every block of $k=3$ spins the value of the central spin. Consider two neighboring blocks $\sigma_1, \sigma_2, \sigma_3$ and $\sigma_4, \sigma_5, \sigma_6$ and sum over all values of $\sigma_3, \sigma_4$ keeping $\sigma'_1 = \sigma_2$ and $\sigma'_2 = \sigma_5$ fixed. The respective factors in the partition function can be written as follows: $\exp[K\sigma_3\sigma_4] = \cosh K + \sigma_3\sigma_4 \sinh K$, which is true for $\sigma_3\sigma_4 = \pm 1$. Denote $x = \tanh K$. Then only the terms with even powers of $\sigma_3$ and $\sigma_4$ contribute the factors in the partition function that involve these degrees of freedom:

$$\begin{aligned}
&\sum_{\sigma_3,\sigma_4=\pm 1} \exp[K(\sigma'_1\sigma_3 + \sigma_3\sigma_4 + \sigma_4\sigma'_2)] \\
&= \cosh^3 K \sum_{\sigma_3,\sigma_4=\pm 1} (1+x\sigma'_1\sigma_3)(1+x\sigma_4\sigma_3)(1+x\sigma'_2\sigma_4) \\
&= 4\cosh^3 K(1+x^3\sigma'_1\sigma'_2) = e^{-g(K)}\cosh K'(1+x'\sigma'_1\sigma'_2)\,, \qquad (178) \\
&g(K) = \ln\left(\frac{\cosh K'}{4\cosh^3 K}\right)\ . \qquad (179)
\end{aligned}$$

The expression (178) has the form of the Boltzmann factor $\exp(K'\sigma'_1\sigma'_2)$ with the re-normalized constant $K' = \tanh^{-1}(\tanh^3 K)$ or $x' = x^3$ — this formula and (179) are called recursion relations. The partition function of the whole system in the new variables can be written as

$$\sum_{\{\sigma'\}} \exp\Big[-g(K)N/3 + K'\sum_i \sigma'_i\sigma'_{i+1}\Big]\ .$$

The term proportional to $g(K)$ represents the contribution into the free energy of the short-scale degrees of freedom which have been averaged out. This term does not affect the calculation of any spin correlation function. Yet the renormalization of the constant, $K \to K'$, influences the correlation

---

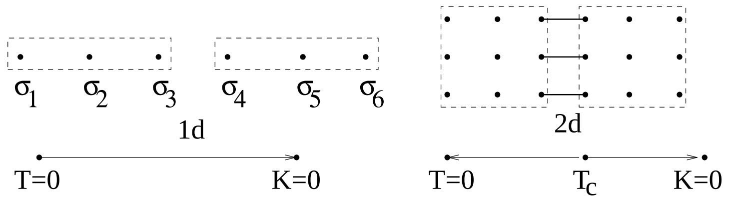
[18] the term initially meant putting to death every tenth soldier of a Roman army regiment that run from a battlefield.

functions. Let us discuss this renormalization. Since $K \propto 1/T$ then $T \to \infty$ correspond to $x \to 0+$ and $T \to 0$ to $x \to 1-$. One is interested in the set of the parameters which does not change under the RG, i.e. represents a fixed point of this transformation. Both $x = 0$ and $x = 1$ are fixed points of the transformation $x \to x^3$, the first one stable and the second one unstable. Indeed, iterating the process for $0 < x < 1$, we see that $x$ approaches zero and effective temperature infinity. That means that large-scale degrees of freedom are described by the partition function where the effective temperature is high so the system is in a paramagnetic state in agreement with the general argument of Sec. 5.3.4. At this limit we have $K, K' \to 0$ so that the contribution of the small-scale degrees of freedom is getting independent of the temperature: $g(K) \to -\ln 4$.

We see that there is no phase transition since there is no long-range order for any $T$ (except exactly for $T = 0$). RG can be useful even without critical behavior, for example, the correlation length measured in lattice units must satisfy $r_c(x') = r_c(x^3) = r_c(x)/3$ which has a solution $r_c(x) \propto -\ln^{-1} x$, an exact result for 1d Ising. It diverges at $x \to 1$ ($T \to 0$) as

$$r_c(T) \propto -\ln^{-1}(\tanh K) \approx -\ln^{-1}(1 - 2e^{-2K}) \propto \exp(2K) = \exp(J/T)$$

again in agreement with Sec. 5.3.4. The same formulas also describe the antiferromagnetic case of $K < 0$.



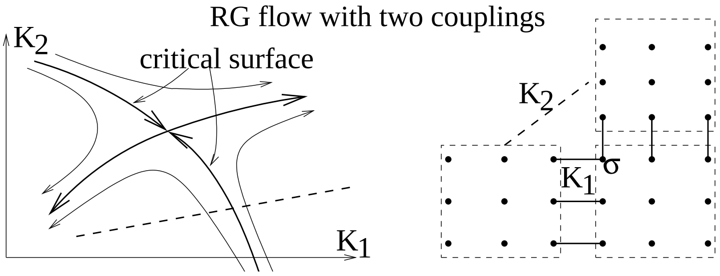
We see that spatial re-scaling leads to the renormalization of temperature, which explains the scaling hypothesis leading to (177); here we used the renormalization factor $k = 3$.

The picture of RG flow is different in higher dimensions. Indeed, in 1d the zero-temperature fixed point is unstable i.e. $K$ decreases under RG transformation. Yet in the low-temperature region ($x \approx 1, K \to \infty$) it decreases very slow so that it does not change in the main order: $K' = K - const \approx K'$. This can be interpreted as due to the interaction between $k$-blocks being mediated by their boundary spins that all look at the same direction: $K' \approx K\langle\sigma_3\rangle_{\sigma_2=1}\langle\sigma_4\rangle_{\sigma_5=1} \approx K$ (by the same token, at high temperatures $\langle\sigma\rangle \propto K$

so that $K' \propto K^3$). However, in $d$ dimensions, there are $k^{d-1}$ spins at the block side so that $K' \propto k^{d-1}K$ as $K \to \infty$ (in the case $k = 3$ and $d = 2$ we have $K' \approx 3K$). That means that $K' > K$ that is the low-temperature fixed point is stable at $d > 1$. On the other hand, the paramagnetic fixed point $K = 0$ is stable too, so that there must be an unstable fixed point in between at some $K_c$ which precisely corresponds to $T_c$. Indeed, consider $r_c(K_0) \sim 1$ at some $K_0$ that corresponds to sufficiently high temperature, $K_0 < K_c$. Since $r_c(K) \sim k^{n(K)}$, where $n(K)$ is the number of RG iterations one needs to come from $K$ to $K_0$, and $n(K) \to \infty$ as $K \to K_c$ then $r_c \to \infty$ as $T \to T_c$. Critical exponent $\nu = -d \ln r_c / d \ln t$ is expressed via the derivative of RG at $T_c$. Indeed, denote $dK'/dK = k^y$ at $K = K_c$ so that $K' - K_c \approx k^y(K - K_c)$. Since $kr_c(K') = r_c(K)$ then we may present:

$$r_c(K) \propto (K - K_c)^{-\nu} = k(K' - K_c)^{-\nu} = k\left[k^y(K - K_c)\right]^{-\nu} ,$$

which requires $\nu = 1/y$. We see that in general, the RG transformation of the set of parameters $\mathbf{K}$ is nonlinear. Linearizing it near the fixed point one can find the critical exponents from the eigenvalues of the linearized RG and, more generally, classify different types of behavior. That requires generally the consideration of RG flows in multi-dimensional spaces.



Already in 2d, summing over corner spin $\sigma$ produces diagonal coupling between blocks. In addition to $K_1$, that describes an interaction between neighbors, we need to introduce another parameter, $K_2$, to account for a next-nearest neighbor interaction. In fact, RG generates all possible further couplings so that it acts in an infinite-dimensional $\mathbf{K}$-space. An unstable fixed point in this space determines critical behavior. We know, however, that we need to control a finite number of parameters to reach a phase transition; for Ising at $h = 0$ and many other systems it is a single parameter, temperature. For all such systems (including most magnetic ones), RG flow has only one unstable direction (with positive $y$), all the rest (with negative $y$) must be contracting stable directions, like the projection on $K_1, K_2$ plane shown in the Figure. The line of points attracted to the fixed point is the projection of the critical surface, so called because the long-distance properties of each system corresponding to a point on this surface are controlled by the fixed point. The critical surface is a separatrix, dividing points that flow to high-$T$ (paramagnetic) behavior from those that flow to low-$T$ (ferromagnetic) behavior at large scales. We can now understand universality of long-distance critical behavior in a sense that systems in different regions of the parameter **K**-space flow to the same fixed point and have thus the same exponents. Indeed, changing the temperature in a system with only nearest-neighbor coupling, we move along the line $K_2 = 0$. The point where this line meets critical surface defines $K_{1c}$ and respective $T_{c1}$. At that temperature, the large-scale behavior of the system is determined by the RG flow i.e. by the fixed point. In another system with nonzero $K_2$, changing $T$ we move along some other path in the parameter space, indicated by the broken line at the figure. Intersection of this line with the critical surface defines some other critical temperature $T_{c2}$. But the long-distance properties of this system are again determined by the same fixed point i.e. all the critical exponents are the same. For example, the critical exponents of a simple fluid are the same as of a uniaxial ferromagnetic. In this case (of the Ising model in an external field), RG changes both $t$ and $h$, the free energy (per block) is then transformed as $f(t,h) = g(t,h) + k^{-d} f(t', h')$. The part $g$ originates from the degrees of freedom within each block so it is an analytic function even at criticality. If we are interested in extracting the singular behavior (like critical exponents) near the critical point, we consider only singular part, which has the form (177) with $a, b$ determined by the derivatives of the RG.

See Cardy, Sect 3.

### 5.4.2 Renormalization group in 4d and $\epsilon$-expansion

Here we present a more general consideration of the universality class described by a single scalar parameter $\eta$ and having a symmetry $\eta \to -\eta$. In particular, Ising model and all its relatives with $K_1, K_2, ...$ belong to this class, when the external magnetic field is absent. Note that in RG treatment the model with discrete variables $\sigma = \pm 1$ turns to be equivalent to the model with continuous $\eta$, what matters is the type of symmetry (discrete symmetry in this case). One can understand it, thinking about block spin renormalization which takes the mean spin over block (instead of choosing the middle spin as we did in the previous section), the spins of blocks will then be non-integer; after many steps of such procedure we obtain a model with a continuous macroscopic magnetization.

Let us describe the formal scheme. We start from the Gibbs distribution that contains microscopic Hamiltonian: $\exp[-\beta(F-\mathcal{H})]$. After averaging it over microscopic degrees of freedom *at a given order parameter* $\eta(\mathbf{r})$, we obtain macroscopic probability distribution $\exp[-\beta(F-\mathcal{F})]$, where the Landau functional $\mathcal{F}$ must depend smoothly on the order parameter as long as $\eta(\mathbf{r})$ is smooth on atomic scales (that smoothness also allowed us to use locally equilibrium Gibbs distribution). Near critical point, when $\eta$ is small, we can expand

$$\mathcal{F}=\frac{1}{2}\int d\mathbf{r}(a\eta^2+g|\nabla\eta|^2+b\eta^4)\ . \tag{180}$$

The partition function and the free energy are given now by the normalization condition:

$$\exp(-\beta F)=\int D\eta\exp[-\beta\mathcal{F}(\eta)]\ . \tag{181}$$

In the right-hand side we have a functional (or path) integral since we need to integrate over all possible (smooth) functions $\eta(\mathbf{r})$. Such functions can be expanded into Fourier series, $\eta(\mathbf{r})=\sum_{\mathbf{q}}\eta_{\mathbf{q}}\exp(i\mathbf{qr})$, containing large but finite number of wavevectors, from the inverse box size $L^{-1}$ to the inverse atomic distance $\Lambda=1/a$. We can thus treat functional integral as a usual multiple integral $\int d\eta_{\mathbf{q}_1}\dots d\eta_{\mathbf{q}_N}$ with $N=(L\Lambda)^d$.

If fluctuations are small and in the functional integral we can neglect all the realizations except those that correspond to the mean value, then we obtain $F=\mathcal{F}(\bar\eta)$ as was written in (162). That corresponds to Landau theory that is to the mean field approximation. However, when fluctuations are substantial, integration over them can make the true free energy very much different from Landau functional, in particular, not an analytic function of the order parameter (as indeed we saw in 2d Ising model). Even more important, while the initial microscopic $a,b,g$ are expected to depend on temperature in an analytic way, the integration over fluctuations may change this dependence. The renormalization group is the procedure to integrate over fluctuations step-by-step, starting from highest wave numbers, and looking for the renormalization of the coefficients $a,g,b$. It is clear that the main difficulty is in the presence of the nonlinear $b$-term, which couples harmonics with different $k$. We assume that $b$ is small (precise criterium will be formulated below) but it may increase as we go towards larger and larger scales doing renormalization group procedure. As we shall now see, the RG evolution of $b$ and the possibility of perturbative integration over small-scale fluctuations depend on the space dimensionality. Indeed, let us look at the criterium (163) for the validity of the Landau theory at arbitrary dimensionality $d$:

$$\langle(\Delta\eta)^2\rangle = \frac{T_c}{2\alpha(T-T_c)r_c^d} \ll \langle\eta\rangle^2 = \frac{\alpha(T-T_c)}{b}. \tag{182}$$

Since in the Landau theory $r_c \propto (T-T_c)^{-1/2}$, then the lhs is proportional to $(T-T_c)^{d/2-1}$. We see that $d=4$ plays a very special role. For $d>4$, the relative fluctuations are getting smaller as one approaches the critical point so that the Landau theory (i.e. the mean field approach) is exact near the critical point. On the contrary, when $d<4$, the fluctuations grow relative to the mean value and the mean-field approximation is invalid at criticality. The case $d=4$ is marginal and is amenable to an explicit renormalization-group analysis as long as the nonlinear term in the Landau functional is small so that the interaction between fluctuations of different scales can be considered by a perturbation theory.

Consider now $d=4$ (Larkin and Khmelnitskii 1969) and divide $\eta=\eta'+\phi$, where $\phi$ contains the largest wavenumbers, $\Lambda'<q<\Lambda$, and $\eta'$ contains the rest. Our goal is to obtain the new functional $\mathcal{F}'(\eta')$ determined by

$$e^{-\beta\mathcal{F}'(\eta')} = \int D\phi e^{-\beta\mathcal{F}(\eta'+\phi)} = e^{-\beta\mathcal{F}(\eta')}\int D\phi e^{-\beta[\mathcal{F}(\phi)+F_{int}]}\ , \tag{183}$$

where $F_{int}$ contains cross terms i.e. describes interaction between fluctuations of different scales. If such term was absent, we would have $\mathcal{F}'=\mathcal{F}$ that is $\int D\phi e^{-\beta\mathcal{F}(\phi)}=1$. We shall make small steps in the elimination procedure, so that the change of the functional is expected to be small and we can expand it in powers of $F_{int}$:

$$\mathcal{F}'(\eta')-\mathcal{F}(\eta') = \frac{1}{\beta}\log\left(\int D\phi e^{-\beta[\mathcal{F}(\phi)+F_{int}]}\right) \approx \langle F_{int}\rangle - \frac{\beta}{2}\langle F_{int}^2\rangle_c\ , \tag{184}$$

As usual, expanding the log of the exponent we obtained the cumulant $\langle F_{int}^2\rangle_c = \langle F_{int}^2\rangle - \langle F_{int}\rangle^2$. The angular brackets here mean the averaging over the statistics of the fast field: $\langle F_{int}\rangle = \int D\phi F_{int} e^{-\beta\mathcal{F}(\phi)}$. Since we assumed that $b$ is small then this statistics is close to Gaussian i.e. determined by the quadratic part of the Landau functional. In particular, the spectral density of the second moment is given by the Ornstein-Zernike expression $|\bar{\phi}_q|^2 = T(a+gk^2)^{-1}$. From now on let us consider the temperature to be at the critical point where $a=0$ and use $|\bar{\phi}_q|^2 = T/gk^2$. In the interaction term $F_{int}$ we omit the linear term $\int d^4r\phi\eta^3 = \sum_{k,p,l}\eta_k\eta_p\eta_l\phi_{\mathbf{k}+\mathbf{p}+\mathbf{l}}$ whose contributions into (184) are nonzero only for $\phi_q$ with $q \simeq \Lambda'$ and are small comparing to those of the term $F_{int} = 3\int d^4rb\eta^2\phi^2$, which we take into account. Indeed, the first term in (184) is as follows:

$$\langle F_{int}\rangle = 3b\langle\phi^2\rangle\int d^4r\eta^2\,,$$
$$\langle\phi^2\rangle = \int_{\Lambda'}^{\Lambda} d^4q|\bar{\phi}_q|^2 = \int_{\Lambda'}^{\Lambda}\frac{d^4q}{(2\pi)^4}\frac{T}{gq^2} = \frac{T\Lambda^2}{4\pi^2 g}\,. \tag{185}$$

This correction to the $a$-term is determined by atomic scales, independent of $\Lambda'$, and corresponds to the renormalization of $T_c$ due to fluctuations. Next term in (184) is much more interesting:

$$\begin{aligned}\langle F_{int}^2\rangle_c &= 24b^2\int d^4r_1d^4r_2\left[\langle\phi^2(\mathbf{r}_1)\phi^2(\mathbf{r}_2)\rangle - \langle\phi^2(\mathbf{r}_1)\rangle\langle\phi^2(\mathbf{r}_2)\rangle\right]\eta'^2(\mathbf{r}_1)\eta'^2(\mathbf{r}_2)\\ &= 24b^2\int d^4r_1d^4r_2\langle\phi(\mathbf{r}_1)\phi(\mathbf{r}_2)\rangle^2\eta'^2(\mathbf{r}_1)\eta'^2(\mathbf{r}_2)\,.\end{aligned} \tag{186}$$

Here we expressed the fourth moment of the Gaussian fast fields $\phi$ through the pair correlation function, which is as follows:

$$G(\mathbf{r}_1-\mathbf{r}_2) = \langle\phi(\mathbf{r}_1)\phi(\mathbf{r}_2)\rangle = \int_{\Lambda'}^{\Lambda} d^4q\frac{Te^{i\mathbf{qr}}}{gq^2} \approx \frac{T}{4\pi^2 gr^2}\ \text{for}\ \Lambda' \ll r^{-1} \ll \Lambda\,. \tag{187}$$

In the integral (186), the function $G(\mathbf{r}_1-\mathbf{r}_2)$ decays starting from the scale $1/\Lambda'$ while the field $\eta$ changes over larger scales. Therefore, we can put $\eta'^2(\mathbf{r}_2)\to\eta'^2(\mathbf{r}_1)$ which gives

$$\begin{aligned}\mathcal{F}'(\eta') - \mathcal{F}(\eta') &= -12\beta b^2\int d^4r\eta'^4(\mathbf{r})\int d^4RG^2(\mathbf{R})\\ &= -\frac{96Tb^2}{\pi^2g^2}\ln(\Lambda/\Lambda')\int d^4r\eta'^4(\mathbf{r})\,,\end{aligned} \tag{188}$$

This is apparently the renormalization of the nonlinear term, that is the change in the parameter $b$ due to step in $\Delta\xi = \ln(\Lambda/\Lambda')$:

$$\Delta b = -\Delta\xi\frac{96Tb^2}{\pi^2g^2}\,.$$

Since the changes at every step are small, we can use differential approximation. Introducing dimensionless interaction parameter (also called coupling constant or charge), $\lambda = 96bT/\pi^2 g^2$ we obtain the renormalization group equation

$$\frac{d\lambda}{d\xi} = -\lambda^2 \,, \quad \lambda(\xi) = \frac{\lambda_0}{1+\lambda_0\xi} \,. \tag{189}$$

We see that the coupling constant decreases when the scale of averaging grows, such situation is called "zero charge", since it was observed first in quantum electrodynamics[19] (Abrikosov, Khalatnikov and Landau 1956). That means that for $d = 4$ large-scale degrees of freedom interact weakly and thus have Gaussian statistics, the situation referred as the Gaussian fixed point. In a similar way, by collecting terms that contain logarithmic integration, one can show that the functional $\mathcal{F}'$ preserves the form of the Landau functional and find the renormalization of the parameters $a, g$. We thus conclude that our approach works uniformly well when the coupling constant at the atomic scale, $\lambda_0$, is small, then it only gets smaller at larger scales. That allows us to choose $\Lambda, \Lambda'$ in such a way that $\Delta\xi$ is large (so we can pick up only terms with logarithmic integration) and yet $\lambda\Delta\xi$ is small so we can apply the renormalization group approach.

The case $d = 3$ presents that unfortunate but usual combination of being most interesting and yet neither amenable to exact solution (as $d = 1, 2$) nor to a perturbative approach (as $d = 4$). Following the ingenious idea of Wilson (1971), we can consider $d = 4 - \epsilon$, develop perturbation theory in $\epsilon$ starting from the known solution for $d = 4$ and then put $\epsilon = 1$ praying for the better. In the space with $d = 4 - \epsilon$ the integrals will diverge by a power law instead of logarithmic, but as long as $\epsilon$ is small, the same terms dominate. The main difference is that now we also need to rescale the field $\eta$ to preserve the form of the free energy after one step of the transformation (integrating over $\phi$ and rescaling $r \to r\Lambda/\Lambda'$). Indeed, the main term in the functional is the quadratic term $\int d^d k g k^2 |\eta_{\mathbf{k}}|^2$, which preserves its form under transformation $k \to k\Lambda'/\Lambda$, $\eta \to \eta(\Lambda/\Lambda')^{1+d/2}$. Under that transformation, the nonlinear term $b \int d^d k d^d p d^d q \eta_{\mathbf{k}} \eta_{\mathbf{p}} \eta_{\mathbf{q}} \eta_{-\mathbf{k}-\mathbf{p}-\mathbf{q}}$ acquires the factor $(\Lambda'/\Lambda)^{4-d}$, so to preserve its form we need to rescale also

---

[19] Effective charge decreases with distance due to screening by vacuum fluctuations. Opposite situation, when the interaction increases with the scale, is called "asymptotic freedom", it takes place in quantum chromodynamics and is responsible for confinement of quarks.

the coupling constant: $b \to b(\Lambda/\Lambda')^{4-d} = be^{\epsilon\xi}$. Such re-scaling was absent in 4d. Therefore, the dimensionless coupling constant must now be defined as $\lambda(\Lambda') = (\Lambda'/\Lambda)^{4-d}(96Tb/\pi^2 g^2) = 96Tbe^{\epsilon\xi}/\pi^2 g^2$. Differentiating this with respect to $\xi$ adds an extra term to the renormalization group equation:

$$\frac{d\lambda}{d\xi} = \epsilon\lambda - \lambda^2 \ . \tag{190}$$

For $d > 4$ that is for negative $\epsilon$, one has $\lambda \to 0$ as $\xi \to \infty$, which corresponds to the Gaussian fixed point, as for $d = 4$. On the contrary, when the space dimensionality is less than four and $\epsilon > 0$, the equation (190) has a nontrivial (i.e. non-Gaussian) fixed point $\lambda = \epsilon$, which is stable so that $\lambda \to \epsilon$ as $\xi \to \infty$. Remind that we are already at the critical temperature i.e. study RG flow on the critical surface. Using $\epsilon$-expansion, one can similarly find the scaling exponents of $a, g$, which are $\epsilon$-dependent and thus different from those suggested by the Landau theory. But the main result of the above consideration is not the specific value of this or that exponent at $d = 3$ and $\epsilon = 1$, but a compelling argument for the existence of a non-trivial fixed point of the renormalization group at $d = 3$. That fixed point, that is the form of the free energy, is, of course, very much different from (180) which reproduces itself under RF transformation only at $\epsilon \ll 1$. See Wilson and Kogut for more details.

At the end of this section, let us reiterate the dramatic shift of paradigm brought by the renormalization group approach. Instead of being interested in this or that Hamiltonian, we are interested in different RG-flows in the space of Hamiltonians. Whole families (universality classes) of different systems described by different Hamiltonians flow under RG transformation to the same fixed point i.e. have the same large-scale behaviour.

# 6 Fluctuating fields and random walks

Many of the properties of the statistical systems, in particularly, the statistics of fluctuations can be related to the fundamental problem of a random walk. It is interesting both for fundamental physics and for numerous modern applications related to nano-particles, macro-molecules etc. Our emphasis now will be on fluctuations.

## 6.1 Random walk and diffusion

Consider a particle that can hop randomly to a neighboring cite of $d$-dimensional cubic lattice, starting from the origin at $t = 0$. We denote $a$ the lattice spacing, $\tau$ the time between hops and $\mathbf{e}_i$ the orthogonal lattice vectors that satisfy $\mathbf{e}_i \cdot \mathbf{e}_j = a^2\delta_{ij}$. The probability to be in a given cite $\mathbf{x}$ satisfies the equation

$$P(\mathbf{x}, t+\tau) = \frac{1}{2d}\sum_{i=1}^{d} P(\mathbf{x} \pm \mathbf{e}_i, t) \ . \tag{191}$$

The first (painful) way to solve this equation is to turn it into averaging exponents as we always do in statistical physics. This is done using the Fourier transform, $P(\mathbf{x}) = (a/2\pi)^d \int e^{i\mathbf{kx}} P(\mathbf{k})\, d^dk$, which gives

$$P(\mathbf{k}, t+\tau) = \frac{1}{d}\sum_{i=1}^{d} \cos ak_i \, P(\mathbf{k}, t) \ . \tag{192}$$

The initial condition for (191) is $P(\mathbf{x}, 0) = \delta(\mathbf{x})$, which gives $P(\mathbf{k}, 0) = 1$ and $P(\mathbf{k}, t) = \left(d^{-1}\sum_{i=1}^{d} \cos ak_i\right)^{t/\tau}$. That gives the solution in space as an integral

$$P(\mathbf{x}, t) = (a/2\pi)^d \int e^{i\mathbf{kx}} \left(\frac{1}{d}\sum_{i=1}^{d} \cos ak_i\right)^{t/\tau} d^dk \ . \tag{193}$$

We are naturally interested in the continuous limit $a \to 0, \tau \to 0$. If we take $\tau \to 0$ first, the integral tends to zero and if we take $a \to 0$ first, the answer remains delta-function. A non-trivial evolution appears when the lattice constant and the jump time go to zero simultaneously. Consider the cosine expansion,

$$\left(\frac{1}{d}\sum_{i=1}^{d} \cos ak_i\right)^{t/\tau} = \left(1 - a^2k^2/2d + \ldots\right)^{t/\tau} ,$$

where $k^2 = \sum_{i=1}^{d} k_i^2$. The finite answer $\exp(-\kappa t k^2)$ appears only if one takes the limit keeping constant the ratio $\kappa = a^2/2d\tau$. In this limit, the space density of the probability stays finite and is given by the saddle-point calculation of the integral:

$$\rho(\mathbf{x},t) = P(\mathbf{x},t)a^{-d} \approx (2\pi)^{-d}\int e^{i\mathbf{kx}-t\kappa k^2}\, d^dk = (4\pi\kappa t)^{-d/2}\exp\left(-\frac{x^2}{4\kappa t}\right) \tag{194}$$

The second (painless) way to get this answer is to pass to the continuum limit already in the equation (191):

$$P(\mathbf{x},t+\tau)-P(\mathbf{x},t) = \frac{1}{2d}\sum_{i=1}^{d}\left[P(\mathbf{x}+\mathbf{e}_i,t)+P(\mathbf{x}-\mathbf{e}_i,t)-2P(\mathbf{x},t)\right] \ . \tag{195}$$

This is a finite difference approximation to the diffusion equation

$$(\partial_t - \kappa\Delta)P(\mathbf{x},t) = 0 \ . \tag{196}$$

Of course, $\rho$ satisfies the same equation, and (194) is its solution. Note that (194,196) are isotropic and translation invariant while the discrete version respected only cubic symmetries. Also, the diffusion equation conserves the total probability, $\int \rho(\mathbf{x},t)\, d\mathbf{x}$, because it has the form of a continuity equation, $\partial_t\rho(\mathbf{x},t) = -\mathrm{div}\,\mathbf{j}$ with the current $\mathbf{j} = -\kappa\nabla\rho$.

Another way to describe it is to treat $\mathbf{e}_i$ as a random variable with $\langle\mathbf{e}_i\rangle = 0$ and $\langle\mathbf{e}_i\mathbf{e}_j\rangle = a^2\delta_{ij}$, so that $\mathbf{x} = \sum_{i=1}^{t/\tau}\mathbf{e}_i$.

Random walk is a basis of the statistical theory of fields. One way to see the relation is to consider the Green function which is the mean time spent on a given site $\mathbf{x}$:

$$G(\mathbf{x}) = \sum_{t=0}^{\infty} P(\mathbf{x},t) \ . \tag{197}$$

The most natural question is whether this time if finite or infinite. From, (194) it is immediately clear that the answer depends on the space dimensionality: $\int^\infty t^{-d/2}$ diverges for $d \le 2$. It is instructive to see see it from the discrete version too. Indeed, summing (193) over $t$ as a geometric series one gets

$$G(\mathbf{x}) = \int \frac{e^{i\mathbf{kx}}\, d^dk}{1 - d^{-1}\sum\cos(ak_i)} \ . \tag{198}$$

It diverges at $k \to 0$ when $d \le 2$. In the limit $k \to 0$ one can also use the continuous limit, where the Green function has a form

$$g(\mathbf{x}) = \lim_{a\to 0}(a^{2-d}/2d)G(\mathbf{x}/a) = \int_0^\infty dt \int e^{i\mathbf{kx}-tk^2}\, d^dk = \int \frac{e^{i\mathbf{kx}}\, d^dk}{(2\pi)^d k^2}\ . \qquad (199)$$

We have seen this integral calculating the correlation function of fluctuation (164). The divergence of this integral in one and two dimensions meant before that the fluctuation are so strong that they destroy any order. Now, (199) suggests another interpretation: integral divergence means that the mean time spent by a random walker on any given site is infinite. In other words, it means that the walker in 1d and 2d returns to any point infinite number of times. Analogy between the Green function of a random walker and the correlation function of fluctuating fields appear in the large-scale limit exploiting a short-correlated nature of a random walk and short-range interaction of fields.

A path of a random walker behaves rather like a surface than a line. Two-dimensionality of the random walk is a reflection of the square-root diffusion law: $\langle x\rangle \propto \sqrt{t}$. Indeed, one can define the dimensionality of a geometric object as a relation between its size $R$ and the number $N$ of standard-size elements(with fixed volume or area) needed to cover it . For a line, $N \propto R$, generally $N \propto R^d$. For a random walk, the number of elements is of order of the number of steps, $N \propto t$, while $R \propto x$ so that $d = 2$. Surfaces generally intersect along curves in 3d, they meet at isolated points in 4d and do not meet at $d > 4$. That is reflected in special properties of critical phenomena in 2d (where random walker fills the surface) and 4d (where random walkers do not meet and hence do not interact)..

To describe the whole statistics of the (fluctuating) time on a site, one can use a slight generalization which gives the generating function for the time moments:

$$G(\mathbf{x},\lambda) = \lambda\sum_{t=0}^{\infty}\lambda^{t/\tau}P(\mathbf{x},t) = \int \frac{e^{i\mathbf{kx}}\, d^dk}{\lambda^{-1} - d^{-1}\sum\cos(ak_i)}\ . \qquad (200)$$

At $\lambda = 1$ it coincides with the Green functions while its derivatives give moments of time:

$$\langle(1+t/\tau)^n\rangle = \left(\frac{\partial^n G}{\partial\lambda^n}\right)_{\lambda=1}\ .$$

The continuous limit of the generating function,

$$g(\mathbf{x},m) = \lim_{a\to 0}(a^{2-d}/2d)G(\mathbf{x}/a,\lambda) = \int \frac{e^{i\mathbf{kx}}\, d^dk}{(2\pi)^d(k^2+m^2)}\ , \qquad (201)$$

exactly corresponds to the Ornstein-Zernike approximation of the correlation function of fluctuations of order parameter away from criticality (with a finite correlation radius). Here we denoted $1/\lambda = 1 + m^2a^2/2d$ so that $m$ plays the role of the inverse radius of correlation or mass of the quantum particle. Note that this Green function can be presented as an integral of the probability density (194) taken with $\kappa = 1$:

$$g(\mathbf{x}, m) = \int_0^\infty e^{-m^2 t}\rho(\mathbf{x}, t)\, dt \ . \tag{202}$$

The properties of random walks can be expressed alternatively in terms of sums over different paths. Let us write the transition probability indicating explicitly the origin: $\rho(\mathbf{x}, t; 0, 0)$. Then we can write the convolution identity which simply states that the walker was certainly somewhere at an intermediate time $t_1$:

$$\rho(\mathbf{x}, t; 0, 0) = \int \rho(\mathbf{x}, t; \mathbf{x}_1, t_1)\rho(\mathbf{x}_1, t_1; 0, 0)\, d\mathbf{x}_1 \ . \tag{203}$$

We now divide the time interval $t$ into an arbitrary large number of intervals and using (194) we write

$$\begin{aligned}\rho(\mathbf{x}, t; 0, 0) &= \int \Pi_{i=0}^{n} \frac{d\mathbf{x}_{i+1}}{[4\pi\kappa(t_{i+1} - t_i)]^{d/2}} \exp\left[-\frac{(\mathbf{x}_{i+1} - \mathbf{x}_i)^2}{4\kappa(t_{i+1} - t_i)}\right] \\ &\to \int \mathcal{D}\mathbf{x}(t') \exp\left[-\frac{1}{4\kappa}\int_0^t dt' \dot{x}^2(t')\right] \ . \end{aligned} \tag{204}$$

The last expression is an integral over paths that start at zero and end up at $\mathbf{x}$ at $t$. We concentrate on the exponential here: it gives the weight of every trajectory. By virtue of (202) it leads to a path-integral expression for the Green function:

$$g(\mathbf{x}, m) = \int_0^\infty dt \int \mathcal{D}\mathbf{x}(t') \exp\left\{-\int_0^t dt' \left[m^2 + \frac{1}{4}\dot{x}^2(t')\right]\right\} \ . \tag{205}$$

Comparison of the Green functions (164,201) shows the relation between the random walk and a free field. This analogy goes beyond the correlation function to all the statistics. Indeed, much in the same way, the partition function of a fluctuating field $\eta(\mathbf{x})$ that takes continuous values in a continuum limit can be written as a path integral over all realizations of the field:

$$Z = \int d\eta_1 \ldots d\eta_N \exp\left[-\beta\mathcal{H}(\eta_1 \ldots \eta_N)\right] \to \int \mathcal{D}\eta \exp\left[-\beta\mathcal{H}\{\eta(\mathbf{x})\}\right] \ . \tag{206}$$

For a Gaussian free field in a discrete case one takes

$$\beta\mathcal{H} = \frac{1}{2}\sum_{\mathbf{x},\mathbf{x}'} \eta(\mathbf{x})[\lambda^{-1}\delta_{\mathbf{x},\mathbf{x}'} - J_{\mathbf{x},\mathbf{x}'}]\eta(\mathbf{x}') \tag{207}$$

$$= \frac{1}{2}\int d^d k \eta(\mathbf{k})[\lambda^{-1} - d^{-1}\sum \cos(ak_i)]\eta(-\mathbf{k}) \ , \tag{208}$$

where $J_{\mathbf{x},\mathbf{x}'} = 1/2d$ when $|\mathbf{x} - \mathbf{x}'| = a$ and $J_{\mathbf{x},\mathbf{x}'} = 0$ otherwise. In the continuous limit one re-normalizes $\eta(\mathbf{x}/a) \to \sqrt{2d}a^{d/2-1}\eta(\mathbf{x})$, and obtains (206) with

$$\beta\mathcal{H} \;=\; \frac{1}{2}\int d\mathbf{x}\left(m^2\eta^2 + |\nabla\eta|^2\right) \, . \tag{209}$$

We did it before in (160). In particular, it gives the Ornstein-Zernike correlation function (164,201).

Stochastic dynamics of a random walk can thus be seen as statistical physics in space-time with trajectories playing the role of configurations.

Looking at the transition probability (204), one can also see the analogy between the statistical physics of a random walk and quantum mechanics. According to Feynman, for a quantum non-relativistic particle with a mass $M$ in the external potential $U(\mathbf{x})$, the transition amplitude $T(\mathbf{x}, t; 0, 0)$ from zero to $\mathbf{x}$ during $t$ is given by the sum over all possible paths connecting the points. Every path is weighted by the factor $\exp(iS/\hbar)$ where $S$ is the classical action:

$$T(\mathbf{x}, t; 0, 0) \;=\; \int \mathcal{D}\mathbf{x}(t') \exp\left[\frac{i}{\hbar}\int_0^t dt'\left(\frac{M\dot{x}^2}{2} + U(x)\right)\right] \, . \tag{210}$$

Comparing with (204), we see that the transition probability of a random walk is given by the transition amplitude of a free quantum particle during an imaginary time. In quantum theory, one averages over quantum rather than thermal fluctuations, yet the formal structure of the theory is similar.

This similarity can be also revealed by using the formalism of the transfer matrix for the systems with nearest neighbor interaction. Indeed, in a nutshell, quantum mechanics is done by specifying two sets of states $|q\rangle$ and $\langle p|$, which has ortho-normality and completeness: $\langle p|q\rangle = \delta_{qp}$ and $\sum_q |q\rangle\langle q| = 1$. Physical quantities are represented by operators, and measurement corresponds to taking a trace of the operator over the set of states: $\mathrm{trace}\, P = \sum_q \langle q|P|q\rangle$. One special operator, called Hamiltonian $\mathcal{H}$, determines the temporal evolution of all other operators according to $P(t) = \exp(i\mathcal{H}t)P(0)\exp(-i\mathcal{H}t)$. The operator $T(t) = \exp(i\mathcal{H}t)$ is called time translation operator also called evolution operator. The quantum-mechanical average of any operator $Q$ is calculated as a trace with the evolution operator normalized by the trace of the evolution operator:

$$\langle Q\rangle = \frac{\mathrm{trace}\, T(t)Q}{Z(t)} \, , \qquad Z(t) = \mathrm{trace}\, T(t) = \sum_a e^{-itE_a} \, . \tag{211}$$

The normalization factor is naturally to call the partition function, all the more if we formally consider it for an imaginary time $t = i\beta$

$$Z(\beta) = \mathrm{trace}\, T(i\beta) = \sum_a e^{-\beta E_a} \, . \tag{212}$$

If the inverse "temperature" $\beta$ goes to infinity then all the sums are dominated by the ground state, $Z(\beta) \approx \exp(-\beta E_0)$ and the average in (212) are just expectation values in the ground state.

That quantum mechanical description can now be compared with the transfer-matrix description of the Ising model in Section 4.5.1. For the Ising model, the sum over two values of $\sigma$ at every cite is the analog of taking trace in quantum-mechanical average. If there are $m$ values on the cite, then $T$ is $m \times m$ matrix. For a spin in $n$-dimensional space (described by so-called $O(n)$ model), trace means integrating over orientations. We see that the translations along the chain are analogous to quantum-mechanical translations in (imaginary) time. This analogy is not restricted to 1d systems, one can consider 2d strips that way too.

## 6.2 Brownian motion

Let us see how the properties of the random walk appear for motion of a small particle in a fluid. The momentum of the particle, $\mathbf{p} = M\mathbf{v}$, changes because of collisions with the molecules. When the particle $M$ is much larger than the molecular mass $m$ then the particle velocity $v$ is small comparing to the typical velocities of the molecules $v_T \simeq \sqrt{T/m}$. Then one can write the force acting on it as Taylor expansion with the parts independent $\mathbf{p}$ and linear in $\mathbf{p}$:

$$\dot{\mathbf{p}} = \mathbf{f} - \lambda \mathbf{p} \ . \tag{213}$$

If the particle radius $R$ is larger than the mean free path $\ell$, in calculating resistance, we can consider fluid as a continuous medium and characterize it by the viscosity $\eta \simeq mnv_T\ell$, where $n$ is the concentration of the molecules. For a slow moving particle, $v \ll v_T\ell/R$, the resistance is given by the Stokes formula

$$\lambda = 6\pi\eta R/M \ . \tag{214}$$

The solution of the linear equation (213) is

$$\mathbf{p}(t) = \int_{-\infty}^{t} \mathbf{f}(t')e^{\lambda(t'-t)}dt' \ . \tag{215}$$

We must treat the force $\mathbf{f}(t)$ as a random function since we do not track molecules hitting the particle, which makes (213) Langevin equation. We assume that $\langle \mathbf{f} \rangle = 0$ and that $\langle \mathbf{f}(t') \cdot \mathbf{f}(t'+t) \rangle = 3C(t)$ decays with $t$ during the correlation time $\tau$ which is much smaller than $\lambda^{-1}$. Since the integration time in (215) is of order $\lambda^{-1}$ then the condition $\lambda\tau \ll 1$ means that the momentum of a Brownian particle can be considered as a sum of many independent random numbers (integrals over intervals of order $\tau$) and so it must have a Gaussian statistics $\rho(\mathbf{p}) = (2\pi\sigma^2)^{-3/2}\exp(-p^2/2\sigma^2)$ where

$$\begin{aligned}\sigma^2 &= \langle p_x^2\rangle = \langle p_y^2\rangle = \langle p_z^2\rangle = \int_0^\infty C(t_1-t_2)e^{-\lambda(t_1+t_2)}dt_1dt_2\\ &\approx \int_0^\infty e^{-2\lambda t}\,dt\int_{-2t}^{2t}C(t')\,dt' \approx \frac{1}{2\lambda}\int_{-\infty}^{\infty}C(t')\,dt' \ . \qquad (216)\end{aligned}$$

On the other hand, equipartition guarantees that $\langle p_x^2\rangle = MT$ so that we can express the friction coefficient via the correlation function of the force fluctuations (a particular case of the fluctuation-dissipation theorem to be discussed below in Sect. 7.2):

$$\lambda = \frac{1}{2TM}\int_{-\infty}^{\infty}C(t')\,dt' \ . \qquad (217)$$

Displacement

$$\Delta\mathbf{r} = \mathbf{r}(t+t') - \mathbf{r}(t) = \int_0^{t'}\mathbf{v}(t'')\,dt''$$

is also Gaussian with a zero mean. To get its second moment we need the different-time correlation function of the velocities

$$\langle\mathbf{v}(t)\cdot\mathbf{v}(0)\rangle = (3T/M)\exp(-\lambda|t|) \qquad (218)$$

which can be obtained from (215). Note that the friction makes velocity correlated on a longer timescale than the force. That gives

$$\langle(\Delta\mathbf{r})^2\rangle = \int_0^{t'}dt_1\int_0^{t'}dt_2\langle\mathbf{v}(t_1)\mathbf{v}(t_2)\rangle = \frac{6T}{M\lambda^2}(\lambda t' + e^{-\lambda t'} - 1) \ .$$

The mean squared distance initially grows quadratically (so-called ballistic regime at $\lambda t' \ll 1$). In the limit of a long time (comparing to the relaxation time $\lambda^{-1}$ rather than to the force correlation time $\tau$) we have the diffusive growth $\langle(\Delta\mathbf{r})^2\rangle \approx 6Tt'/M\lambda$. Generally $\langle(\Delta\mathbf{r})^2\rangle = 2d\kappa t$ where $d$ is space dimensionality and $\kappa$ - diffusivity. In our case $d = 3$ and then the diffusivity is as follows: $\kappa = T/M\lambda$ — the Einstein relation. Using (214) one can rewrite it as follows

$$\kappa = \frac{T}{M\lambda} = \frac{T}{6\pi\eta R} \ . \qquad (219)$$

Note that the diffusivity depends on particle radius, but not mass. Heavier particles are slower both to start and to stop moving. Measuring diffusion of particles with a known size one can determine the temperature[20].

The probability distribution of displacement at $\lambda t' \gg 1$,

$$\rho(\Delta \mathbf{r}, t') = (4\pi\kappa t')^{-3/2} \exp[-(\Delta \mathbf{r})^2/4\kappa t'] \,,$$

satisfies the diffusion equation $\partial\rho/\partial t' = \kappa\nabla^2\rho$. If we have many particles initially distributed according to $n(\mathbf{r}, 0)$ then their distribution $n(\mathbf{r}, t) = \int \rho(\mathbf{r} - \mathbf{r}', t) n(\mathbf{r}', 0)\, d\mathbf{r}'$, also satisfies the diffusion equation: $\partial n/\partial t' = \kappa\nabla^2 n$.

In the external field $V(\mathbf{q})$, the particle satisfies the equations

$$\dot{\mathbf{p}} = -\lambda\mathbf{p} + \mathbf{f} - \partial_q V \,, \quad \dot{\mathbf{q}} = p/M \,. \tag{220}$$

Note that these equations characterize the system with the Hamiltonian $\mathcal{H} = p^2/2M + V(\mathbf{q})$, that interact with the thermostat, which provides friction $-\lambda\mathbf{p}$ and agitation $\mathbf{f}$ - the balance between these two terms expressed by (217) means that the thermostat is in equilibrium.

We now pass from considering individual trajectories to the description of the "cloud" of trajectories. Consider the over-damped limit, $\lambda\mathbf{p} \gg \dot{\mathbf{p}}$, which gives the first-order equation:

$$\lambda\mathbf{p} = \lambda M\dot{\mathbf{q}} = \mathbf{f} - \partial_q V \,. \tag{221}$$

We can now derive the equation on the probability distribution $\rho(q, t)$, which changes with time due to random noise and evolution in the potential, the two mechanisms can be considered additively. We know that without $V$,

$$\mathbf{q}(t) - \mathbf{q}(0) = (\lambda M)^{-1} \int_0^t \mathbf{f}(t')dt' \,, \quad \langle |\mathbf{q}(t) - \mathbf{q}(0)|^2 \rangle = 2Dt \,,$$

and the density $\rho(q, t)$ satisfies the diffusion equation. The dynamical equation without any randomness, $\lambda M\dot{\mathbf{q}} = -\partial_q V$, corresponds to a flow in $\mathbf{q}$-space with the velocity $\mathbf{w} = -\partial_q V/\lambda M$. In a flow, density satisfies the continuity equation $\partial_t \rho = -div\, \rho\mathbf{w}$. Together, diffusion and advection give the so-called Fokker-Planck equation

$$\frac{\partial\rho}{\partial t} = \kappa\nabla^2\rho + \frac{1}{\lambda M}\frac{\partial}{\partial q_i}\rho\frac{\partial V}{\partial q_i} = -\mathrm{div}\,\mathbf{J} \,. \tag{222}$$

---

[20] With temperature in degrees, (219) contains the Boltzmann constant, $k = \kappa M\lambda/T$, which was actually determined by this relation and found constant indeed, i.e. independent of the medium and the type of particle. That proved the reality of atoms - after all, $kT$ is the kinetic energy of a single atom.

More formally, one can derive this equation by writing the Langevin equation (221) as $\dot{q}_i - w_i = \eta_i$ and taking the random force Gaussian delta-correlated: $\langle \eta_i(0)\eta_j(t)\rangle = 2\kappa\delta_{ij}\delta(t)$. Since it is the quantity $\dot{\mathbf{q}} - \mathbf{w}$ which is Gaussian now, then the path integral representation (204) changes into

$$\rho(\mathbf{q}, t; 0, 0) \;=\; \int \mathcal{D}\mathbf{q}(t') \exp\left[-\frac{1}{4\kappa}\int_0^t dt' |\dot{\mathbf{q}} - \mathbf{w}|^2\right] , \tag{223}$$

To describe the time change, consider the convolution identity (203) for an infinitesimal time shift $\epsilon$, then instead of the path integral we get simply the integral over the initial position $\mathbf{q}'$. We substitute $\dot{\mathbf{q}} = (\mathbf{q} - \mathbf{q}')/\epsilon$ into (223) and obtain

$$\rho(\mathbf{q}, t) = \int d\mathbf{q}'(4\pi\kappa\epsilon)^{-d/2} \exp\left[-\frac{[\mathbf{q} - \mathbf{q}' - \epsilon\mathbf{w}(\mathbf{q}')]^2}{4\kappa\epsilon}\right] \rho(\mathbf{q}', t - \epsilon) \ . \tag{224}$$

What is written here is simply that the transition probability is the Gaussian probability of finding the noise $\eta$ with the right magnitude to provide for the transition from $\mathbf{q}'$ to $\mathbf{q}$. We now change integration variable, $\mathbf{y} = \mathbf{q} - \mathbf{q}' - \epsilon\mathbf{w}(\mathbf{q}')$, and keep only the first term in $\epsilon$: $d\mathbf{q}' = d\mathbf{y}[1 - \epsilon\partial_{\mathbf{q}} \cdot \mathbf{w}(\mathbf{q})]$. Here $\partial_{\mathbf{q}} \cdot \mathbf{w} = \partial_i w_i = div\, \mathbf{w}$. In the resulting expression, we expand the last factor,

$$\begin{aligned}
&\rho(\mathbf{q}, t) \approx (1 - \epsilon\partial_{\mathbf{q}} \cdot \mathbf{w}) \int d\mathbf{y}(4\pi\kappa\epsilon)^{-d/2} e^{-y^2/4\kappa\epsilon} \rho(\mathbf{q} + \mathbf{y} - \epsilon\mathbf{w}, t - \epsilon) \\
&\approx (1 - \epsilon\partial_{\mathbf{q}} \cdot \mathbf{w}) \int d\mathbf{y}(4\pi\kappa\epsilon)^{-d/2} e^{-y^2/4\kappa\epsilon} \Big[\rho(\mathbf{q}, t) + (\mathbf{y} - \epsilon\mathbf{w}) \cdot \partial_{\mathbf{q}}\rho(\mathbf{q}, t) \\
&+ (y_i y_j - 2\epsilon y_i w_j + \epsilon^2 w_i w_j)\partial_i\partial_j \rho(\mathbf{q}, t)/2 - \epsilon\partial_t\rho(\mathbf{q}, t)\Big] \\
&= (1 - \epsilon\partial_{\mathbf{q}} \cdot \mathbf{w})[\rho - \epsilon\mathbf{w} \cdot \partial_{\mathbf{q}}\rho + \epsilon\kappa\Delta\rho - \epsilon\partial_t\rho + O(\epsilon^2)] \ ,
\end{aligned} \tag{225}$$

and obtain (222) collecting terms linear in $\epsilon$. Note that it was necessary to expand until the quadratic terms in $y$, which gave the contribution linear in $\epsilon$, namely the Laplacian, i.e. the diffusion operator.

The Fokker-Planck equation has a stationary solution which corresponds to the particle in an external field and in thermal equilibrium with the surrounding molecules:

$$\rho(q) \propto \exp[-V(q)/\lambda M\kappa] = \exp[-V(q)/T] \ . \tag{226}$$

Apparently it has a Boltzmann-Gibbs form, and it turns into zero the probability current: $\mathbf{J} = -\rho\partial V/\partial\mathbf{q} - \kappa\partial\rho/\partial\mathbf{q} = 0$. We shall return to the Fokker-Planck equation in the next Chapter for the consideration of the detailed balance and fluctuation-dissipation relations.

Ma, Sect. 12.7; Kardar Fields, Sect 9.1.

# 7 Response and fluctuations

As we learnt above (in particular, in Sections 5.1,5.2), the probability of the fluctuation is determined by the entropy change which is the minimal work (divided by temperature) which is the change in the respective thermodynamics potential. For small fluctuations, we can expand the thermodynamic potential up to quadratic terms thus getting a Gaussian probability distribution. Therefore, the mean squared thermodynamic fluctuation of any quantity is determined by the second derivative of the thermodynamic potential with respect to this quantity. Those second derivatives are related to susceptibilities with respect to the properly defined external forces. Here we formulate a general relation between susceptibilities and the second moments of fluctuations.

## 7.1 Static response

Consider a system with the Hamiltonian $\mathcal{H}$ and add some small static external force $f$ so that the Hamiltonian becomes $\mathcal{H} - xf$ where $x$ is called the coordinate. The examples of force-coordinate pairs are magnetic field and magnetization, pressure and volume etc. The mean value of any other variable $B$ can be calculated by the canonical distribution with the new Hamiltonian

$$\bar{B} = \frac{\sum B \exp[(xf - \mathcal{H})/T]}{\sum \exp[(xf - \mathcal{H})/T]} \ .$$

Note that we assume that the perturbed state is also in equilibrium. The susceptibility of $B$ with respect to $f$ is as follows

$$\chi \equiv \frac{\partial \bar{B}}{\partial f} = \frac{\langle Bx \rangle - \bar{B}\bar{x}}{T} \equiv \frac{\langle Bx \rangle_c}{T} \ . \tag{227}$$

Here the cumulant (also called the irreducible correlation function) is defined for quantities with the subtracted mean values $\langle xy \rangle_c \equiv \langle (x - \bar{x})(y - \bar{y}) \rangle$ and it is thus the measure of statistical correlation between $x$ and $y$. We thus learn that the susceptibility is the measure of the statistical coherence of the system, increasing with the statistical dependence of variables. Consider few examples of this relation.

Example 1. If $x = \mathcal{H}$ is energy itself then $f$ represents the fractional increase in the temperature: $\mathcal{H}(1-f)/T \approx \mathcal{H}/(1+f)T$. Formula (227) then gives the relation between the specific heat (which is a kind of susceptibility) and the squared energy fluctuation which can be written via the irreducible correlation function of the energy density $\epsilon(\mathbf{r})$:

$$\begin{aligned}\frac{\partial\bar{\mathcal{H}}}{\partial f} &= T\frac{\partial E}{\partial T} = TC_v = \langle(\Delta E)^2\rangle/T \\ &= \frac{1}{T}\int\langle\epsilon(\mathbf{r})\epsilon(\mathbf{r}')\rangle_c\, d\mathbf{r}d\mathbf{r}' = \frac{V}{T}\int\langle\epsilon(\mathbf{r})\epsilon(0)\rangle_c\, d\mathbf{r} \qquad (228)\end{aligned}$$

Growth of the specific heat when the temperature approached criticality is related to the growth of the correlation function of fluctuations. As we discussed before, the specific heat is extensive i.e. proportional to the volume (or number of particles), but the coefficient of proportionality actually tells us how many degrees of freedom are effective in absorbing energy at a given temperature (recall two-level system where specific heat was small for high and low temperatures). We see from (228) that the higher the correlation the larger is the specific heat that is the more energy one needs to spend to raise the temperature by one degree. In other words, system with more correlation absorbs more energy under a given temperature difference.

Example 2. If $f = h$ is a magnetic field then the coordinate $x = M$ is the magnetization and (227) gives the magnetic susceptibility

$$\chi = \frac{\partial M}{\partial h} = \frac{\langle M^2\rangle_c}{T} = \frac{V}{T}\int\langle m(\mathbf{r})m(0)\rangle_c\, d\mathbf{r}\ .$$

Divergence of $\chi$ near the Curie point means the growth of correlations between distant spins i.e. the growth of correlation length. For example, the Ornshtein-Zernicke correlation function (161) gives $\langle m(\mathbf{r})m(0)\rangle_c \propto r^{2-d}$ so that in the mean-field approximation $\chi \propto \int_a^{r_c} d^d r r^{2-d} \propto r_c^2 \propto |T - T_c|^{-1}$.

General remark. These fluctuation-response relations can be related to the change of the thermodynamic potential (free energy) under the action of the force:

$$\begin{aligned}F &= -T\ln Z = -T\ln\sum\exp[(xf - \mathcal{H})/T] \\ &= -T\ln Z_0 - T\ln\langle\exp(xf/T)\rangle_0 = F_0 - f\langle x\rangle_0 - \frac{f^2}{2T}\langle x^2\rangle_{0c} + \ldots \qquad (229)\\ \langle x\rangle &= -\partial F/\partial f, \quad \langle x^2\rangle_c/T = \partial\langle x\rangle/\partial f = -\partial F^2/\partial f^2\,. \qquad (230)\end{aligned}$$

Subscript 0 means an average over the state with $f = 0$, like $\langle\exp(xf/T)\rangle_0 = \sum\exp(xf/T)\exp(-\mathcal{H}/T)/\sum\exp(-\mathcal{H}/T)$; we don't write the subscript in the expansion (230), which can take place around any value of $f$. Formula (229) is based on the cumulant expansion theorem:

$$\begin{aligned}
&\langle \exp(ax)\rangle = 1 + \sum_{n=1}^{\infty} \frac{a^n}{n!}\langle x^n\rangle \,, \\
&\ln\langle \exp(ax)\rangle = -\sum_{n=1} \frac{1}{n}\left(1-\langle \exp(ax)\rangle\right)^n = -\sum_{n=1} \frac{1}{n}\left(-\sum_{m=1}^{\infty}\frac{a^m}{m!}\langle x^m\rangle\right)^n \\
&= a\langle x\rangle + \left(\langle x^2\rangle - \langle x\rangle^2\right)\frac{a^2}{2!} + \ldots = \sum_{n=1}^{\infty}\frac{a^n}{n!}\langle x^n\rangle_c = \langle e^{ax}-1\rangle_c \,. \qquad (231)
\end{aligned}$$

In other words, $\langle \exp(ax)\rangle$ is the generating function of the moments $\langle x^n\rangle$ while $\ln\langle \exp(ax)\rangle$ is the generating function of the cumulants. We already used that in making virial expansion in Sect. 4.2, see (111), in RG expansion (184) and will use below in (288).

Example 3. Consider now the inhomogeneous force $f(\mathbf{r})$ and denote $a(\mathbf{r}) \equiv x(\mathbf{r}) - x_0$. The Hamiltonian change is now the integral

$$\int f(\mathbf{r})a(\mathbf{r})\,d\mathbf{r} = \sum_{kk'} f_k a_{k'} \int e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}}\,d\mathbf{r} = V\sum_k f_k a_{-k} \ .$$

The mean linear response $\bar{a}(\mathbf{r})$ can be written as an integral of the force with the response (Green) function; again, translation invariance makes the relation between the response and force diagonal in the Fourier representation:

$$\bar{a}(\mathbf{r}) = \int G(\mathbf{r}-\mathbf{r}')f(\mathbf{r}')\,d\mathbf{r}' \,, \quad \bar{a}_k = G_k f_k \ . \qquad (232)$$

One relates the Fourier components of the Green function and the correlation function of the coordinate fluctuations choosing $B = a_k$, $x = a_{-k}$ in (227):

$$\begin{aligned}
&VG_k = \frac{\langle a_k a_{-k}\rangle_c}{T} = \frac{1}{T}\int \langle a(\mathbf{r})a(\mathbf{r}')\rangle_c e^{i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})}\,d\mathbf{r}d\mathbf{r}' = \frac{V}{T}\int \langle a(\mathbf{r})a(0)\rangle_c e^{-i\mathbf{k}\cdot\mathbf{r}}\,d\mathbf{r} \,, \\
&\quad TG_k = (a^2)_k \ .
\end{aligned}$$

Example 4. If $B = x = N$ then $f$ is the chemical potential $\mu$:

$$\left(\frac{\partial N}{\partial \mu}\right)_{T,V} = \frac{\langle N^2\rangle_c}{T} = \frac{\langle(\Delta N)^2\rangle}{T} = \frac{V}{T}\int \langle n(\mathbf{r})n(0)\rangle_c d\mathbf{r} \ . \qquad (233)$$

This formula coincides with (153) if one accounts for

$$\begin{aligned}-n^2\left(\frac{\partial V}{\partial P}\right)_{T,N} &= N\left(\frac{\partial n}{\partial P}\right)_{T,N} = n\left(\frac{\partial N}{\partial P}\right)_{T,V} \\ &= \left(\frac{\partial P}{\partial \mu}\right)_{T,V}\left(\frac{\partial N}{\partial P}\right)_{T,V} = \left(\frac{\partial N}{\partial \mu}\right)_{T,V} .\end{aligned} \tag{234}$$

Here we used the fact $\Omega(T,\mu) = PV$ and $N = \partial\Omega/\partial\mu$. We conclude that the response of the density to the pressure is expressed via the density fluctuations.

The correlation function $\langle n(\mathbf{r})n(0)\rangle$ is $n^2$ times the probability to find particle at $\mathbf{r}$ under the condition that there is a particle at zero. In the simplest case of an ideal gas with $\langle n(\mathbf{r})n(0)\rangle_c = n\delta(\mathbf{r})$, (233,234) give $dn/dP = 1/T$. To account for the pair interaction energy $U(r)$ in the first approximation (neglecting multiple correlations) one just uses the Gibbs probability $e^{-U(r)/T}$ so that the cumulant is $\langle n(\mathbf{r})n(0)\rangle_c = n\{\delta(\mathbf{r}) + n[e^{-U(r)/T} - 1]\}$ and the corrected equation of state is

$$p = nT + n^2T\int[1 - e^{-U(r)/T}]\,d\mathbf{r} ,$$

which we already derived in Sect. 4.3 in the virial expansion (116-117). We now see that the second virial coefficient is actually the second density cumulant.

More details in Shang-Keng Ma, Statistical Mechanics, Sect. 13.1

## 7.2 Temporal correlation of fluctuations

We now consider the time-dependent force $f(t)$ so that the Hamiltonian is $\mathcal{H} = \mathcal{H}_0 - xf(t)$. Time dependence requires more elaboration than space inhomogeneity[21] because one must find the non-equilibrium time-dependent probability density in the phase space solving the Liouville equation

$$\frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial x}\frac{dx}{dt} + \frac{\partial\rho}{\partial p}\frac{dp}{dt} = \frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial x}\frac{\partial\mathcal{H}}{\partial p} - \frac{\partial\rho}{\partial p}\frac{\partial\mathcal{H}}{\partial x} = 0 . \tag{235}$$

Here $p$ is the canonical momentum conjugated to the coordinate $x$. One can solve the equation (235) perturbatively in $f$, denoting $\rho(x,p,t) = \rho_0(x,p) +$

---

[21] As the poet (Brodsky) said, "Time is bigger than space: space is an entity, time is in essence a thought of an entity."

$\rho_1(x,p,t)$ and assuming $\rho_1(x,p,t) \ll \rho_0 = Z^{-1}\exp(-\beta\mathcal{H}_0)$ by virtue of $xf \ll \mathcal{H}_0$:

$$\frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial x}\frac{\partial\mathcal{H}_0}{\partial p} - \frac{\partial\rho}{\partial p}\frac{\partial\mathcal{H}_0}{\partial x} = \frac{d\rho_1}{dt} = -\frac{\partial\rho_0}{\partial p}f = f\beta\frac{\partial\mathcal{H}_0}{\partial p}\rho_0 \ . \tag{236}$$

Here, like in Sect. 2.1, $d/dt$ denotes the derivative in the moving reference frame along an unperturbed trajectory determined by $\mathcal{H}_0$. Recall now that $\partial\mathcal{H}_0/\partial p = dx/dt$ (calculated at $f = 0$ i.e. also along an unperturbed trajectory). The formal solution of (236) is written as an integral over the past:

$$\rho_1 = \beta\rho_0\int_{-\infty}^{t} f(\tau)\frac{dx(\tau)}{d\tau}\,d\tau \ . \tag{237}$$

We now use (237) to derive the relation between the fluctuations and response in the time-dependent case. Indeed, the linear response of the coordinate to the force is as follows

$$\langle x(t)\rangle \equiv \int_{-\infty}^{t}\alpha(t,t')f(t')\,dt' = \int x dx\rho_1(x,t)\,, \tag{238}$$

which defines generalized susceptibility (also called response or Green function) $\alpha(t,t') = \alpha(t-t') \equiv \delta\langle x(t)\rangle/\delta f(t')$. Note that causality requires $\alpha(t-t') = 0$ for $t < t'$. Substituting (237) into (238) and taking a variational derivative $\delta/\delta f(t')$ we obtain the fluctuation-dissipation theorem (FDT)

$$T\alpha(t,t') = \int\rho_0 x(t)\frac{dx(t')}{dt'}dx = \left\langle x(t)\frac{dx(t')}{dt'}\right\rangle = \frac{d}{dt'}\langle x(t)x(t')\rangle\,,\ t \ge t' \ . \tag{239}$$

It relates quantity in equilibrium (the decay rate of correlations) to the weakly non-equilibrium quantity (response to a small perturbation). Differentiation and averaging are both linear operations so they commute, which allowed us to take the time derivative outside. Pay attention to the fact that the derivative is with respect to the earlier time, which is related to causality and is also clear looking at (237). While (239) is similar to the fluctuation-response relations obtained in the previous section, it is called the fluctuation-dissipation theorem. To understand (239) better and to see where the word "dissipation" comes from, we introduce the spectral decomposition of the fluctuations:

$$x_\omega = \int_{-\infty}^{\infty} x(t)e^{i\omega t}dt\,, \quad x(t) = \int_{-\infty}^{\infty} x_\omega e^{-i\omega t}\frac{d\omega}{2\pi} \ . \tag{240}$$

The pair correlation function, $\langle x(t')x(t)\rangle$ must be a function of the time difference which requires $\langle x_\omega x_{\omega'}\rangle = 2\pi\delta(\omega+\omega')(x^2)_\omega$ — this relation is the definition of the spectral density of fluctuations $(x^2)_\omega$. Linear response in the spectral form is $\bar{x}_\omega = \alpha_\omega f_\omega$ where

$$\alpha(\omega) = \int_0^\infty \alpha(t)e^{i\omega t}\,dt = \alpha' + \imath\alpha''$$

is analytic in the upper half-plane of complex $\omega$ under the assumption that $\alpha(t)$ is finite everywhere. Since $\alpha(t)$ is real then $\alpha(-\omega^*) = \alpha^*(\omega)$.

We can now make a Fourier transform of (239) and obtain the spectral form of the fluctuation-dissipation theorem (Callen and Welton, 1951):

$$\begin{aligned}
&\frac{T\alpha(\omega)}{i\omega} = \int_0^\infty \langle x(0)x(t)\rangle \exp(i\omega t)dt\,, \\
&(x^2)_\omega = \int_0^\infty \langle x(0)x(t)\rangle \exp(i\omega t)dt + \int_{-\infty}^0 \langle x(0)x(t)\rangle \exp(i\omega t)dt \\
&= \frac{T[\alpha(\omega)-\alpha(-\omega)]}{i\omega} = \frac{2T\alpha''(\omega)}{\omega}\ .
\end{aligned}$$

Let us show that the imaginary part $\alpha''$ determines the energy dissipation,

$$\frac{dE}{dt} = \overline{\frac{d\mathcal{H}}{dt}} = \overline{\frac{\partial\mathcal{H}}{\partial t}} = \overline{\frac{\partial\mathcal{H}}{\partial f}}\frac{df}{dt} = -\bar{x}\frac{df}{dt} \tag{241}$$

For purely monochromatic perturbation, $f(t) = f_\omega\exp(-i\omega t) + f^*_\omega\exp(i\omega t)$, $\bar{x} = \alpha(\omega)f_\omega\exp(-i\omega t) + \alpha(-\omega)f^*_\omega\exp(i\omega t)$, the dissipation averaged over a period is as follows:

$$\overline{\frac{dE}{dt}} = \int_0^{2\pi/\omega}\frac{\omega dt}{2\pi}[\alpha(-\omega)-\alpha(\omega)]\imath\omega|f_\omega|^2 = 2\omega\alpha''_\omega|f_\omega|^2\ . \tag{242}$$

We can now calculate the average dissipation using (237)

$$\begin{aligned}
\overline{\frac{dE}{dt}} &= -\int x\dot{f}\rho_1\,dpdx = -\beta\int x(t)\dot{f}(t)\rho_0\,dpdx\int_{-\infty}^t \dot{x}(\tau-t)f(\tau)\,d\tau \\
&= -i\omega|f_\omega|^2\beta\int_{-\infty}^\infty \langle x(t)\dot{x}(t')\rangle e^{i\omega(t-t')}\,dt' = \beta\omega^2|f_\omega|^2(x^2)_\omega\ , \qquad (243)
\end{aligned}$$

where the spectral density of the fluctuations is calculated with $\rho_0$ (i.e. at unperturbed equilibrium). Comparing (242) and (243) we obtain again:

$$2T\alpha''(\omega) = \omega(x^2)_\omega\ . \tag{244}$$

This truly amazing formula relates the dissipation coefficient that governs non-equilibrium kinetics under the external force with the equilibrium fluctuations. The physical idea is that to know how a system reacts to a force one might as well wait until the fluctuation appears which is equivalent to the result of that force. Note that the force $f$ disappeared from the final result which means that the relation is true even when the (equilibrium) fluctuations of $x$ are not small. Integrating (244) over frequencies we get

$$\langle x^2\rangle = \int_{-\infty}^{\infty} (x^2)_\omega \frac{d\omega}{2\pi} = \frac{T}{\pi}\int_{-\infty}^{\infty} \frac{\alpha''(\omega)d\omega}{\omega} = \frac{T}{\imath\pi}\int_{-\infty}^{\infty} \frac{\alpha(\omega)d\omega}{\omega} = T\alpha(0)\ . \quad (245)$$

Here we added the real part, which is symmetric under the frequency change so does not contribute. The last step used the Cauchy integration formula (for the contour that runs along the real axis, avoids zero by a small semi-circle and closes at the upper half plane). We thus see that the mean squared fluctuation is the zero-frequency response, which is the integral of the response over time:

$$\alpha(\omega = 0) = \int_0^\infty \alpha(t)\, dt\ .$$

We have seen before the simplest case of the relation between fluctuations and relaxation - the relation (217) expressing the friction coefficient of the Brownian particle via the variance of the fluctuating molecular forces and the temperature of the medium. The same Langevin equation (213) is satisfied by many systems, in particular by the current $I$ flowing through an electric L-R circuit at a finite temperature:

$$L\frac{dI}{dt} = -RI + V(T)\ ,$$

where the fluctuating voltage $V(t)$ is due to thermal fluctuations. Similar to (217) we can then relate the resistance to the equilibrium voltage fluctuations on a resistor $R$:

$$\int_{-\infty}^{\infty} \langle V(0)V(t)\rangle dt = 2RT\ .$$

This relation is called Nyquist theorem, note that $L$ does not enter. Equipartition requires $L < I^2 > /2 = T/2$, so that similar to (218) we can write the current auto-correlation function $\langle I(0)I(t)\rangle = (T/L)\exp(-Rt/L)$. The direct analog of the FDT (239) is now

$$\frac{d}{dt}\langle I(0)I(t)\rangle = \frac{TR}{L^2}\exp(-Rt/L)\ ,$$

so that the current response function is $\alpha(t) = (R/L^2)\exp(-Rt/L)$. The exponential correlation function corresponds to the Lorentzian spectral density: $(I^2)_\omega = 2RT/(L^2\omega^2 + R^2)$. At low frequencies it corresponds to a constant spectral density (white noise) - current noise is uncorrelated at large times. Generally, the spectral density has a universal Lorentzian form in the low-frequency limit when the period of the force is much longer than the relaxation time for establishing the partial equilibrium characterized by the given value $\bar{x} = \alpha(0)f$. In this case, the evolution of $x$ is the relaxation towards $\bar{x}$:

$$\dot{x} = -\lambda(x - \bar{x}) \ . \tag{246}$$

For harmonics,

$$\begin{aligned}
&(\lambda - i\omega)x_\omega = \lambda\bar{x} = \lambda\alpha(0)f \ , \\
&\alpha(\omega) = \alpha(0)\frac{\lambda}{\lambda - i\omega} \ , \quad \alpha''(\omega) = \alpha(0)\frac{\omega}{\lambda^2 + \omega^2} \ .
\end{aligned} \tag{247}$$

The spectral density of such (so-called quasi-stationary) fluctuations is as follows:

$$(x^2)_\omega = \langle x^2 \rangle \frac{2\lambda}{\lambda^2 + \omega^2} \ . \tag{248}$$

It corresponds to the long-time exponential decay of the temporal correlation function: $\langle x(t)x(0)\rangle = \langle x^2\rangle \exp(-\lambda|t|)$. That exponent is a temporal analog of the large-scale formula (161). The exponential decay occurs on the scales or times much exceeding typical. Similarly, the probability in the Poisson distribution of not meeting any particle decays exponentially with the volume, the probability of radioactive decay not happening decays exponentially with time etc. Non-smooth behavior at zero is an artefact of the long-time approximation, consistent consideration would give zero derivative at $t = 0$. The susceptibility is $\alpha(t) = \exp(-\lambda t)$.

When several degrees of freedom are weakly deviated from equilibrium, the relaxation must be described by the system of linear equations (consider all $x_i = 0$ at the equilibrium)

$$\dot{x}_i = -\lambda_{ij}x_j \ . \tag{249}$$

The dissipation coefficients are generally non-symmetric: $\lambda_{ij} \neq \lambda_{ji}$. One can however find a proper coordinates in which the coefficients are symmetric. Single-time probability distribution of small fluctuations is Gaussian

$w(\mathbf{x}) \sim \exp(\Delta S) \approx \exp(-\beta_{jk} x_j x_k)$. The matrix $\hat\beta$ is symmetric since it is the matrix of second derivatives. Introduce generalized forces[22] $X_j = -\partial S/\partial x_j = \beta_{ij} x_i$ so that $\dot x_i = \gamma_{ij} X_j$, $\gamma_{ij} = \lambda_{ik}(\hat\beta^{-1})_{kj}$ with $\langle x_i X_j\rangle = \int d\mathbf{x} x_i X_j w = -\int d\mathbf{x} x_i \partial w/\partial x_j = \delta_{ij}$. Indeed, we have seen that the coordinates and the generalized forces do not cross-correlate between different pairs already in the simplest case of uniform fluctuations described by (151), which gave $\langle \Delta T \Delta V\rangle = 0$, for instance. Returning to the general case, note also that $\langle X_j X_j\rangle = \beta_{ij}$ and $\langle x_j x_k\rangle = (\hat\beta^{-1})_{jk}$. If $x_i$ all have the same properties with respect to the time reversal then their correlation function is symmetric too: $\langle x_i(0) x_k(t)\rangle = \langle x_i(t) x_k(0)\rangle$. Differentiating it with respect to $t$ at $t = 0$ we get the Onsager symmetry principle, $\gamma_{ik} = \gamma_{ki}$. For example, the conductivity tensor is symmetric in anisotropic crystals without magnetic field. Also, a temperature difference produces the same electric current as the heat current produced by a voltage. Such symmetry relations due to time reversibility are valid only near equilibrium steady state and are manifestations of the detailed balance (i.e. absence of any persistent currents in the phase space). Let us stress that this is different from the susceptibilities in equilibrium which were symmetric for the simple reason that they were second derivatives of the thermodynamic potential; for instance, dielectric susceptibility $\chi_{ij} = \partial P_i/\partial E_j = \chi_{ji}$ where $P$ is the polarization of the medium - this symmetry is analogous to the (trivial) symmetry of $\hat\beta$, not the (non-trivial) symmetry of $\hat\gamma$.

See Landay & Lifshitz, Sect. 119-120 for the details and Sect. 124 for the quantum case. Also Kittel, Sects. 33-34.

## 7.3 Spatio-temporal correlation function

To have a specific example, let us calculate the correlation function of the density at different points and times for the simplest case of an ideal gas. We have $N$ particles with the time-independent momenta $\mathbf{p}_k$ and the coordinates $\mathbf{R}_k(t) = \mathbf{R}_k(0) + \mathbf{p}_k t/m$. The concentration is just the sum of delta-functions, corresponding to different particles:

$$n(\mathbf{r}, t) = \sum_{k=1}^{N} \delta\Big(\mathbf{r} - \mathbf{R}_k(t)\Big) . \tag{250}$$

---

[22] Here the coordinate-force pair appears in the thermodynamic potential, rather than in Hamiltonian as before, so we call the force *generalized*.

Since the particles do not interact, there is no correlation between different particles. Therefore, the only correlation between densities $n(\mathbf{r},t)$ and $n(\mathbf{r}',t')$ can appear due to a particle that visited both places at respective times:

$$\begin{aligned}
\langle n(\mathbf{r},t)n(\mathbf{r}',t')\rangle_c &= \left\langle \sum_{k=1}^{N} \delta\Big(\mathbf{r}-\mathbf{R}_k(t)\Big)\delta\Big(\mathbf{r}'-\mathbf{R}_k(t')\Big)\right\rangle \\
&= N\left\langle \delta\Big(\mathbf{r}-\mathbf{R}_k(t)\Big)\delta\Big(\mathbf{r}'-\mathbf{R}_k(t)-\mathbf{p}_k(t'-t)/m\Big)\right\rangle \\
&= N\left\langle \delta\Big(\mathbf{r}-\mathbf{R}_k(t)\Big)\delta\Big(\mathbf{r}'-\mathbf{r}-\mathbf{p}_k(t'-t)/m\Big)\right\rangle \ . \qquad (251)
\end{aligned}$$

There are two averages here. The first one is over all possible positions within the space $V$, which gives $\langle\delta\Big(\mathbf{r}-\mathbf{R}_k(t)\Big)\rangle = 1/V$. The second average is over the Maxwell distribution of the momenta:

$$\begin{aligned}
\langle n(\mathbf{r},t)n(\mathbf{r}',t')\rangle_c &= \frac{N}{V}\left\langle \delta\Big(\mathbf{r}'-\mathbf{r}-\mathbf{p}_k(t'-t)/m\Big)\right\rangle \\
&= n(2\pi MT)^{-d/2}\int d\mathbf{p}e^{-p^2/2mT}\delta\Big(\mathbf{r}'-\mathbf{R}_k(t)-\mathbf{p}_k(t'-t)/M\Big) \\
&= n|t-t'|^{-d}\left(\frac{m}{2\pi T}\right)^{d/2}\exp\left(-\frac{m|\mathbf{r}-\mathbf{r}'|^2}{2|t-t'|^2}\right) \ . \qquad (252)
\end{aligned}$$

That function determines the response of the concentration to the change in the chemical potential. In particular, when $t'\to t$ it tends to $n\delta(\mathbf{r}-\mathbf{r}')$, which determines the static response described in Sect. 7.1. For coinciding points it decays by the diffusion law, $\langle n(\mathbf{r},t)n(\mathbf{r},t')\rangle_c \propto |t-t'|^{-d}$, so that the response decays as $|t-t'|^{-d-1}$.

## 7.4 General fluctuation-dissipation relation

Consider again the over-damped Brownian particle with the coordinate $x(t)$ in a time-dependent potential $V(x,t)$:

$$\dot{x} = -\partial_x V + \eta \ . \qquad (253)$$

Here the random function $\eta(t)$ can be thought of as representing interaction with a thermostat with the temperature $T$ so that $\langle\eta(0)\eta(t)\rangle = 2T\delta(t)$. This equation (used very often in different applications) can be applied to any macroscopic observable, where one can distinguish a systematic and random part of the evolution.

The Fokker-Planck equation for the probability $\rho(x,t)$ has the form (222):

$$\partial_t \rho = T\partial_x^2 \rho + \partial_x(\rho \partial_x V) = -\hat{H}_{FP}\rho \ . \tag{254}$$

We have introduced the Fokker-Planck operator,

$$H_{FP} = -\frac{\partial}{\partial x}\left(\frac{\partial V}{\partial x} + T\frac{\partial}{\partial q_i}\right) \ ,$$

which allows one to exploit another instance of the analogy between quantum mechanics and statistical physics. We may say that the probability density is the $\psi$-function is the $x$-representation, $\rho(x,t) = \langle x|\psi(t)\rangle$. In other words, we consider evolution in the Hilbert space of functions so that we may rewrite (254) in a Schrödinger representation as $d|\psi\rangle/dt = -\hat{H}_{FP}|\psi\rangle$, which has a formal solution $|\psi(t)\rangle = \exp(-tH_{FP})|\psi(0)\rangle$. The transition probability is given by the matrix element:

$$\rho(x',t';x,t) = \langle x'|\exp[(t-t')H_{FP})|x\rangle \ . \tag{255}$$

Without the coordinate-dependent field $V(x)$, the transition probability is symmetric, $\rho(x',t;x,0) = \rho(x,t;x',0)$, which is formally manifested by the fact that the respective Fokker-Planck operator $\partial_x^2$ is Hermitian. This property is called the detailed balance. How this is modified in an external field? If the potential $V$ is time independent, then we have a Gibbs steady state which also satisfies the detailed balance: the probability current is the (Gibbs) probability density at the starting point times the transition probability; forward and backward currents must be equal in equilibrium:

$$\rho(x',t;x,0)e^{-V(x)/T} = \rho(x,t;x',0)e^{-V(x')/T} \ . \tag{256}$$

$$\langle x'|e^{-tH_{FP}-V/T}|x\rangle = \langle x|e^{-tH_{FP}-V/T}|x'\rangle = \langle x'|e^{-V/T-tH_{FP}^\dagger}|x\rangle \ .$$

Since this must be true for any $x, x'$ then $e^{-tH_{FP}^\dagger} = e^{V/T}e^{-tH_{FP}}e^{-V/T}$ and

$$H_{FP}^\dagger = e^{V/T}H_{FP}e^{-V/T} \ , \tag{257}$$

i.e. $e^{V/2T}H_{FP}e^{-V/2T}$ is hermitian, which can be checked directly.

Consider now a time-dependent potential and an ensemble of trajectories starting from the initial positions taken with the equilibrium Gibbs distribution corresponding to the initial potential: $\rho(x,0) = Z_0^{-1}\exp[-\beta V(x,0)]$. As time proceeds and the potential continuously changes, the system is not in equilibrium anymore so that $\rho(x,t)$ does not generally have a Gibbs form. Indeed, even though one can define a time-dependent Gibbs state $Z_t^{-1}\exp[-\beta V(x,t)]$ with $Z_t = \int \exp[-\beta V(x,t)]dx$, one can directly check that it is not any longer a solution of the Fokker-Planck equation (254) because of the extra term $\partial_t\rho = -\beta\rho\partial_t V$. Still, can we find some use of the Gibbs factor and also have anything generalizing the detailed balance relation (256) we had in equilibrium? Such relation was found surprisingly recently despite its generality and relative technical simplicity of derivation.

To find the quantity that has a Gibbs form, we need to find an equation which generalizes (254) by having an extra term that will cancel the time derivative. It is achieved by considering, apart from a position $x$, another random quantity defined as the potential energy change (or the external work done) during the time $t$:

$$W_t = \int_0^t dt' \frac{\partial V(x(t'),t')}{\partial t'} \ . \tag{258}$$

The time derivative here is partial i.e. taken only with respect to the second argument. The work is a fluctuating quantity depending on the trajectory $x(t')$. Then the remarkable relation holds (Jarzynski 1997):

$$\langle \exp(-\beta W_t)\rangle = Z_t/Z_0 = \frac{\int \exp[-\beta V(x,t)]\,dx}{\int \exp[-\beta V(x,0)]\,dx} \ . \tag{259}$$

Here the bracket means double averaging, over the initial distribution $\rho(x,0)$ and over the different realizations of the Gaussian noise $\eta(t)$ during the time interval $(0,t)$. In other words, we take many different realizations of the noise $\eta(t)$, choose initial $x(0)$ with the Gibbs probability $\rho_0$ and run (253) many times with every initial data and every noise realization. It will give us many trajectories having different endpoints $x(t)$ and different energy changes $W$ accumulated along the way.

Now consider the joint probability $\rho(x,W,t)$ i.e. the probability to come to $x$ acquiring energy change $W$. This two-dimensional probability distribution satisfies the generalized Fokker-Planck equation, which can be derived as follows: Similar to the argument preceding (222), we note that the flow along $W$ in $x-W$ space proceeds with the velocity $dW/dt = \partial_t V$ so that the respective component of the current is $\rho\partial_t V$ and the equation takes the form

$$\partial_t\rho = \beta^{-1}\partial_x^2\rho + \partial_x(\rho\partial_x V) - \partial_W \rho\partial_t V \ , \tag{260}$$

Since $W_0 = 0$ then the initial condition for (260) is

$$\rho(x, W, 0) = Z_0^{-1} \exp[-V(x, 0)]\delta(W) \ . \tag{261}$$

While we cannot find $\rho(x, W, t)$ for arbitrary $V(t)$ we can multiply (260) by $\exp(-\beta W)$ and integrate over $dW$. Since $V(x, t)$ does not depend on $W$, we get the closed equation for $f(x, t) = \int dW \rho(x, W, t) \exp(-\beta W)$:

$$\partial_t f = \beta^{-1}\partial_x^2 f + \partial_x(f\partial_x V) - \beta f \partial_t V \ , \tag{262}$$

Now, *this* equation does have an exact time-dependent solution $f(x, t) = Z_0^{-1} \exp[-\beta V(x, t)]$ where the factor is chosen to satisfy the initial condition (261). In other words, the distribution weighted by $\exp(-\beta W_t)$ looks like Gibbs state, adjusted to the time-dependent potential at every moment of time. Remark that the entropy is defined only in equilibrium, yet the work divided by temperature is an analog of the entropy change (production), and the exponent of it is an analog of the phase volume change. Let us stress that $f(x, t)$ is not a probability distribution, in particular, its integral over $x$ is not unity. To get (259), what remains is to integrate $f(x, t)$ over $x$. We can also obtain all weighted moments of $x$ like $\langle x^n \exp(-\beta W_t)\rangle$ [23].

Let us reflect. We started from a Gibbs distribution but considered *arbitrary* temporal evolution of the potential. Therefore, our distribution was arbitrarily far from equilibrium during the evolution. Still, to obtain the mean exponent of the work done, it is enough to know the partition functions of the equilibrium Gibbs distributions corresponding to the potential at the beginning and at the end (even though the system is not in equilibrium at the end). Remarkable.

One can obtain different particular results from the general fluctuation-dissipation relation (259). For example, using Jensen inequality $\langle e^A\rangle \geq e^{\langle A\rangle}$ and introducing the free energy $F_t = -T \ln Z_t$, one can obtain the second law of thermodynamics in the following form:

$$\langle W\rangle \geq F_t - F_0 \ .$$

Moreover, Jarzynski relation is a generalization of the fluctuation-dissipation theorem, which can be derived from it for small deviations from equilibrium. Namely, we can consider $V(x, t) = V_0(x) - f(t)x$, consider limit of $f \to 0$, expand (259) up to the second-order terms in $f$ and get (239).

---

[23] I thank R. Chetrite for this derivation

In a multi-dimensional case, there is another way to deviate the system from equilibrium - to apply a non-potential force $\mathbf{f}(\mathbf{q}, t)$ (which is not a gradient of any scalar):

$$\dot{\mathbf{q}} = \mathbf{f} - \partial_{\mathbf{q}} V + \eta \ . \tag{263}$$

The new Fokker-Planck equation has an extra term

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial q_i}\left[\rho\left(\frac{\partial V}{\partial q_i} + f_i\right) + T\frac{\partial \rho}{\partial q_i}\right] = -\hat{H}_{FP}\rho \ . \tag{264}$$

Again, there is no Gibbs steady state and the detailed balance (256,257) is now violated in the following way:

$$H_{FP}^{\dagger} = e^{V/T} H_{FP} e^{-V/T} + (\mathbf{f} \cdot \dot{\mathbf{q}})/T \ , \tag{265}$$

The last term is again the power divided by temperature i.e. the entropy production rate. A close analog of the Jarzynski relation can be formulated for the production rate measured during the time $t$:

$$\sigma_t = -\frac{1}{tT}\int_0^t (\mathbf{f} \cdot \dot{\mathbf{q}})\, dt \ . \tag{266}$$

This quantity fluctuates from realization to realization (of the noise $\eta$). The probabilities $P(\sigma_t)$ satisfy the following relation, which we give without derivation (see Kurchan for details)

$$\frac{P(\sigma_t)}{P(-\sigma_t)} \propto e^{t\sigma_t} \ . \tag{267}$$

The second law of thermodynamics states that to keep the system away from equilibrium, the external force $\mathbf{f}$ must on average do a positive work. Over a long time we thus expect $\sigma_t$ to be overwhelmingly positive, yet fluctuations do happen. The relation (267) shows how low is the probability to observe a negative entropy production rate - this probability decays exponentially with the time of observation. Such fluctuations were unobservable in classical macroscopic thermodynamics, but they are often very important in modern applications to nano and bio objects.

The relation similar to (267) can be derived for any system symmetric with respect to some transformation to which we add anti-symmetric perturbation. Consider a system with the variables $s_1, \ldots, s_N$ and the even energy: $E_0(\mathbf{s}) = E_0(-\mathbf{s})$. Consider the energy perturbed by an odd term,

$E = E_0 - hM/2$, where $M(\mathbf{s}) = \sum s_i = -M(-\mathbf{s})$. The probability of the perturbation $P[M(\mathbf{s})]$ satisfies the direct analog of (267), which is obtained by changing the integration variable $\mathbf{s} \to -\mathbf{s}$:

$$P(a)=\int d\mathbf{s}\delta[M(\mathbf{s})-a]e^{\beta(ha-E_0)}=\int d\mathbf{s}\delta[M(\mathbf{s})+a]e^{-\beta(ha+E_0)}=P(-a)e^{-2\beta ha} \ .$$

Take-home lesson of this Chapter: no fluctuations - no dissipation.

# 8 Entropy and information

By definition, entropy of a closed system determines the number of available states (or, classically, phase volume). Assuming that system spends comparable time in different available states we conclude that since the equilibrium must be the most probable state it corresponds to the entropy maximum. If the system happens to be not in equilibrium at a given moment of time [say, the energy distribution between the subsystems is different from the most probable Gibbs distribution (27)] then it is more probable to go towards equilibrium that is increasing entropy. This is a microscopic (probabilistic) interpretation of the *second law of thermodynamics* formulated by Clausius in 1865. The probability maximum is very sharp in the thermodynamic limit since $\exp(S)$ grows exponentially with the system size. That means that for macroscopic systems the probability to pass into the states with lower entropy is so vanishingly small that such events are never observed.

What often causes confusion here is that the dynamics (classical and quantum) of any given system is time reversible. To avoid the confusion, one must remember that entropy growth is related not to the trajectory of a single point in phase space but to the behavior of finite regions (i.e. sets of such points) or ensembles of systems. The entropy of a single point is zero. The necessity to consider finite regions follows from the insufficiency of information about the true state of the system. Consideration of finite regions is called *coarse graining* and it is the main feature of stat-physical approach responsible for the irreversibility of statistical laws.

## 8.1 Lyapunov exponent

The dynamical background of entropy growth is the separation of trajectories in phase space so that trajectories started from a small finite region are found in larger and larger regions of phase space as time proceeds. The relative motion is determined by the velocity difference between neighboring points in the phase space: $\delta v_i = r_j \partial v_i/\partial x_j = r_j \sigma_{ij}$. Here $\mathbf{x} = (\mathbf{p}, \mathbf{q})$ is the $6N$-dimensional vector of the position and $\mathbf{v} = (\dot{\mathbf{p}}, \dot{\mathbf{q}})$ is the velocity in the phase space. The trace of the tensor $\sigma_{ij}$ is the rate of the volume change which must be zero according to the Liouville theorem (that is a Hamiltonian dynamics imposes an incompressible flow in the phase space). We can decompose the tensor of velocity derivatives into an antisymmetric part (which describes rotation) and a symmetric part (which describes deformation). We are interested here in deformation because it is the mechanism of the entropy growth. The symmetric tensor, $S_{ij} = (\partial v_i/\partial x_j + \partial v_j/\partial x_i)/2$, can be always transformed into a diagonal form by an orthogonal transformation (i.e. by the rotation of the axes), so that $S_{ij} = S_i \delta_{ij}$. The diagonal components are the rates of stretching in different directions. Indeed, the equation for the distance between two points along a principal direction has a form: $\dot{r}_i = \delta v_i = r_i S_i$ . The solution is as follows:

$$r_i(t) = r_i(0) \exp\left[\int_0^t S_i(t')\, dt'\right] . \tag{268}$$

For a time-independent strain, the growth/decay is exponential in time. One recognizes that a purely straining motion converts a spherical element into an ellipsoid with the principal diameters that grow (or decay) in time. Indeed, consider a two-dimensional projection of the initial spherical element i.e. a circle of the radius $R$ at $t = 0$. The point that starts at $x_0, y_0 = \sqrt{R^2 - x_0^2}$ goes into

$$\begin{aligned} & x(t) = e^{S_{11}t} x_0 \,, \\ & y(t) = e^{S_{22}t} y_0 = e^{S_{22}t}\sqrt{R^2 - x_0^2} = e^{S_{22}t}\sqrt{R^2 - x^2(t)e^{-2S_{11}t}} \,, \\ & x^2(t)e^{-2S_{11}t} + y^2(t)e^{-2S_{22}t} = R^2 \,. \end{aligned} \tag{269}$$

The equation (269) describes how the initial circle turns into the ellipse whose eccentricity increases exponentially with the rate $|S_{11} - S_{22}|$. In a multi-dimensional space, any sphere of initial conditions turns into the ellipsoid defined by $\sum_{i=1}^{6N} x_i^2(t) e^{-2S_i t}$ =const.

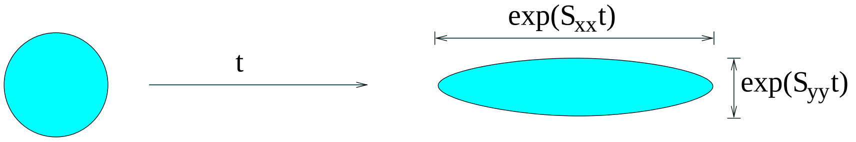


Figure 4: Deformation of a phase-space element by a permanent strain.

Of course, as the system moves in the phase space, both the strain values and the orientation of the principal directions change, so that expanding direction may turn into a contracting one and vice versa. The question is whether averaging over all possibilities gives a zero net result. One can show that in a general case an exponential stretching persists on average and the majority of trajectories separate. Let us first look at that from a temporal perspective: even when the average $\Lambda(t) = \int_0^t S_i(t')dt'$ is zero, the average exponent of it is larger than unity (and generally growing with time):

$$\frac{1}{T}\int_0^T dt \exp\left[\int_0^t S_i(t')dt'\right] \geq 1 \ .$$

This is because the intervals of time with positive $\Lambda(t)$ give more contribution into the exponent than the intervals with negative $\Lambda(t)$, which follows from the concavity of the exponential function. In the simplest case, the average is $(1/2a)\int_a^{-a} e^t dt = (e^a - e^{-a})/2a > 1$. Looking from a spatial perspective, consider the simplest example of a two-dimensional pure strain, which corresponds to an incompressible saddle-point flow: $v_x = \lambda x$, $v_y = -\lambda y$. The vector $\mathbf{r} = (x, y)$ (which is supposed to characterize the distance between two close trajectories) satisfies the equations $\dot{x} = v_x$ and $\dot{y} = v_y$. Whether the vector is stretched or contracted after some time $T$ depends on its orientation and on $T$. Since $x(t) = x_0 \exp(\lambda t)$ and $y(t) = y_0 \exp(-\lambda t) = x_0 y_0 / x(t)$ then every trajectory is a hyperbole. A unit vector initially forming an angle $\varphi$ with the $x$ axis will have its length $[\cos^2\varphi \exp(2\lambda T) + \sin^2\varphi \exp(-2\lambda T)]^{1/2}$ after time $T$. The vector will be stretched if $\cos\varphi \geq [1+\exp(2\lambda T)]^{-1/2} < 1/\sqrt{2}$, i.e. the fraction of stretched directions is larger than half. When along the motion all orientations are equally probable, the net effect is stretching.

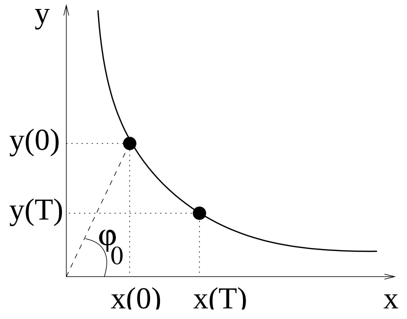


Figure 5: The distance of the point from the origin increases if the angle is less than $\varphi_0 = \arccos[1 + \exp(2\lambda T)]^{-1/2} > \pi/4$. Note that for $\varphi = \varphi_0$ the initial and final points are symmetric relative to the diagonal.

The net stretching and separation of trajectories is formally proved in mathematics by considering random $\hat{\sigma}(t)$ and the transfer matrix $\hat{W}$ defined by $\mathbf{r}(t) = \hat{W}(t, t_1)\mathbf{r}(t_1)$. It satisfies the equation $d\hat{W}/dt = \hat{\sigma}\hat{W}$. The Liouville theorem

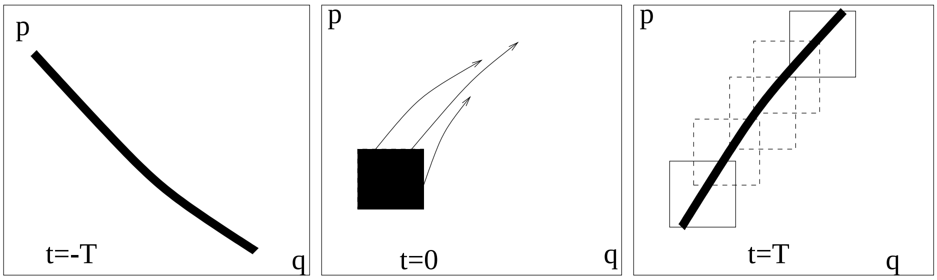
$\mathrm{tr}\,\hat{\sigma} = 0$ means that $\det \hat{W} = 1$. The modulus $r$ of the separation vector may be expressed via the positive symmetric matrix $\hat{W}^T\hat{W}$. The main result (Furstenberg and Kesten 1960; Oseledec, 1968) states that in almost every realization $\hat{\sigma}(t)$, the matrix $\frac{1}{t}\ln \hat{W}^T(t,0)\hat{W}(t,0)$ tends to a finite limit as $t \to \infty$. In particular, its eigenvectors tend to $d$ fixed orthonormal eigenvectors $\mathbf{f}_i$. Geometrically, that precisely means than an initial sphere evolves into an elongated ellipsoid at later times. The limiting eigenvalues

$$\lambda_i = \lim_{t\to\infty} t^{-1} \ln |\hat{W}\mathbf{f}_i| \tag{270}$$

define the so-called Lyapunov exponents. The sum of the exponents is zero due to the Liouville theorem so there exists at least one positive exponent which gives stretching. Therefore, as time increases, the ellipsoid is more and more elongated and it is less and less likely that the hierarchy of the ellipsoid axes will change. Mathematical lesson to learn is that multiplying $N$ random matrices with unit determinant (recall that determinant is the product of eigenvalues), one generally gets some eigenvalues growing (and some decreasing) exponentially with $N$.

The probability to find a ball turning into an exponentially stretching ellipse goes to unity as time increases. The physical reason for it is that substantial deformation appears sooner or later. To reverse it, one needs to contract the long axis of the ellipse, that is the direction of contraction must be inside the narrow angle defined by the ellipse eccentricity, which is less likely than being outside the angle. Randomly oriented deformations on average continue to increase the eccentricity.

Armed with the understanding of the exponential stretching, we now return to the dynamical foundation of the second law of thermodynamics. We assume that our finite resolution does not allow us to distinguish between the states within some square in the phase space. That square is our "grain" in coarse-graining. In the figure below, one can see how such black square of initial conditions (at the central box) is stretched in one (unstable) direction and contracted in another (stable) direction so that it turns into a long narrow strip (left and right boxes). Later in time, our resolution is still restricted - rectangles in the right box show finite resolution (this is coarse-graining). Viewed with such resolution, our set of points occupies larger phase volume (i.e. corresponds to larger entropy) at $t = \pm T$ than at $t = 0$. Time reversibility of any particular trajectory in the phase space does not contradict the time-irreversible filling of the phase space by the set of trajectories considered with a finite resolution. By reversing time we exchange stable and unstable directions (i.e. those of contraction and expansion), but the fact of space filling persists. We see from the figure that the volume and entropy increase both forward and backward in time. To avoid misunderstanding, note that usual arguments that entropy growth provides for time arrow are such: if we already observed an evolution that produces a narrow strip then its time reversal is contraction into a ball; but if we consider a narrow strip as an initial condition, it is unlikely (because of the narrow angle mentioned above) to observe a contraction. Therefore, being shown two movies, one with stretching, another with contraction we conclude that with probability close (but not exactly equal!) to unity the first movie shows the true sequence of events, from the past to the future.



After the strip length reaches the scale of the velocity change (when one already cannot approximate the phase-space flow by a linear profile $\hat{\sigma}r$), strip starts to fold because rotation (which we can neglect for a ball but not for a long strip) is different at different parts of the strip. Still, however long, the strip continues locally the exponential stretching. Eventually, one can find the points from the initial ball everywhere which means that the flow is mixing, also called ergodic. Formal definition is that the flow is called ergodic in the domain if the trajectory of almost every point (except possibly a set of zero volume) passes arbitrarily close to every other point. An equivalent definition is that there are no finite-volume subsets of the domain invariant with respect to the flow except the domain itself. Ergodic flow on an energy surface in the phase space provides for a micro-canonical distribution (i.e. constant), since time averages are equivalent to the average over the surface. While we can prove ergodicity only for relatively simple systems, like the gas of hard spheres, we believe that it holds for most systems of sufficiently general nature (that vague notion can be make more precise by saying that the qualitative systems behavior is insensitive to small variations of its microscopic parameters).

When the density spreads, entropy grows (as the logarithm of the volume occupied). If initially our system was within the phase-space volume $\epsilon^{6N}$, then its density was $\rho_0 = \epsilon^{-6N}$ inside and zero outside. After stretching to some larger volume $A\epsilon^{6N}$ the entropy $S = -\int \rho \ln \rho d\mathbf{x}$ has increased by $\ln A$. The Lyapunov exponent determines the rate of the entropy growth.

At even larger time scales than the time of the velocity change for a trajectory, one can consider the motion as a series of uncorrelated random steps. That produces random walk described in Sect 6.1 above, where we have shown that the spread of the probability density $\rho(\mathbf{r}, t)$ is described by a simple diffusion: $\partial\rho/\partial t = -\kappa\Delta\rho$. The total probability $\int \rho(\mathbf{r}, t)\, d\mathbf{r}$ is conserved but the entropy increases monotonically under diffusion:

$$\frac{dS}{dt} = -\frac{d}{dt}\int \rho(\mathbf{r}, t) \ln \rho(\mathbf{r}, t)\, d\mathbf{r} = -\kappa \int \Delta\rho \ln \rho\, d\mathbf{r} = \kappa \int \frac{(\nabla\rho)^2}{\rho}\, d\mathbf{r} \geq 0\,. \quad (271)$$

Asymptotically in time the density takes the universal form (194); substituting $\rho(\mathbf{r}, t) = (4\pi\kappa t)^{-d/2}\exp\left(-r^2/4\kappa t\right)$ into (271) we obtain a universal entropy production rate, $dS/dt = 1/2t$, independent of $\kappa$ (which is clear from dimensional reasoning).

Two concluding remarks are in order. First, the notion of an exponential separation of trajectories put an end to the old dream of Laplace to be able to predict the future if only all coordinates and momenta are given. Even if we were able to measure all relevant phase-space initial data, we can do it only with a finite precision $\epsilon$. However small is the indeterminacy in the data, it is amplified exponentially with time so that eventually $\epsilon \exp(\lambda T)$ is large and we cannot predict the outcome. Mathematically speaking, limits $\epsilon \to 0$ and $T \to \infty$ do not commute. Second, the above arguments did not use the usual mantra of thermodynamic limit, which means that even the systems with a small number of degrees of freedom need statistics for their description at long times if their dynamics has a positive Lyapunov exponent (which is generic) - this is sometimes called *dynamical chaos*.

## 8.2 Adiabatic processes

The second law of thermodynamics is valid not only for isolated systems but also for systems in the (time-dependent) external fields or under external conditions changing in time as long as there is no heat exchange, that is for systems that can be described by the microscopic Hamiltonian $\mathcal{H}(p, q, \lambda)$ with some parameter $\lambda(t)$ slowly changing with time. That means that the environment is not a macroscopic body with hidden degrees of freedom but is completely determined by the value of the single parameter $\lambda$, that is the entropy of the environment is zero. In particular, $\lambda$ can be the system volume since the walls can be thought of as confining potential. If temporal changes are slow enough then the entropy of the system change only a little i.e. the process is adiabatic. Indeed, the positivity of $\dot{S} = dS/dt$ requires that the expansion of $\dot{S}(\dot{\lambda})$ starts from the second term,

$$\frac{dS}{dt} = \frac{dS}{d\lambda} \cdot \frac{d\lambda}{dt} = A \left( \frac{d\lambda}{dt} \right)^2 \qquad \Rightarrow \qquad \frac{dS}{d\lambda} = A \frac{d\lambda}{dt} \ . \tag{272}$$

We see that when $d\lambda/dt$ goes to zero, entropy is getting independent of $\lambda$. That means that we can change $\lambda$ (say, volume) by finite amount making the entropy change whatever small by doing it slow enough.

During the adiabatic process the system is assumed to be in thermal equilibrium at any instant of time (as in quasi-static processes defined in thermodynamics). Changing $\lambda$ (called coordinate) one changes the energy levels $E_a$ and the total energy. Respective force (pressure when $\lambda$ is volume, magnetic or electric moments when $\lambda$ is the respective field) is obtained as the average (over the equilibrium statistical distribution) of the energy derivative with respect to $\lambda$:

$$\overline{\frac{\partial \mathcal{H}(p,q,\lambda)}{\partial \lambda}} = \sum_a w_a \frac{\partial E_a}{\partial \lambda} = \frac{\partial}{\partial \lambda} \sum_a w_a E_a = \left( \frac{\partial E(S,\lambda,\ldots)}{\partial \lambda} \right)_S \ . \tag{273}$$

We see that the force is equal to the derivative of the thermodynamic energy at constant entropy. It is an important formula since instead of calculating averages over statistical distributions one can just differentiate the thermodynamic potential. Let us also stress that we assumed that in an adiabatic process all probabilities $w_a$ do not change i.e. the entropy of any subsystem us conserved. This is more restrictive than the condition of reversibility which requires only the total entropy to be conserved. In other words, the process can be reversible but not adiabatic. See Landau & Lifshitz (Section 11) for more details.

## 8.3 Information theory approach

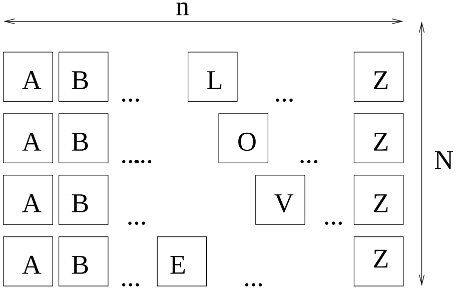
Here I briefly re-tell the story of statistical physics using a different language. It will let us see entropy in a new light. An advantage of using different formulations is that it helps to understand things better and triggers different intuition in different people.

Consider first a simple problem in which we are faced with a choice among $n$ equal possibilities (say, in which of $n$ boxes a candy is hidden). How much we need to know to get the candy? Let us denote the missing information by $I(n)$. Clearly, the information is an increasing function of $n$ and $I(1) = 0$. If we have several independent problems then information must be additive. For example, consider each box to have $m$ compartments: $I(nm) = I(n) + I(m)$. Now, we can write (Shannon, 1948)

$$I(n) = I(e) \ln n = k \ln n \tag{274}$$

That it must be a logarithm is clear also from obtaining the missing information by asking the sequence of questions in which half we find the box with the candy, one then needs $\log_2 n$ of such questions and respective one-bit answers. We can easily generalize the definition (274) for non-integer rational numbers by $I(n/l) = I(n) - I(l)$ and for all positive real numbers by considering limits of the series and using monotonicity. So the single number of the lucky box with the candy brings the information $k \ln n$.

We used to think of information received through words and symbols. If we have an alphabet with $n$ symbols then every symbol we receive is a choice out of $n$ and brings the information $k \ln n$. If symbols come independently then the message of the length $N$ can potentially be one of $n^N$ possibilities so that it brings the information $kN \ln n$. If all the 25 letters of the English alphabet were used with the same frequency then the word "love" would bring the information equal to $4k \ln 25$ or $4 \log_2 25$ bits. Here and below we assume that the receiver has no other prior knowledge on subjects like correlations between letters (for instance, everyone who knows English, can infer that there is only one four-letter word which starts with "lov..." so the last letter brings zero information for such people).



In reality though every letter brings even less information than $k \ln 25$ since we know that letters are used with different frequencies. Indeed, consider the situation when there is a probability $w_i$ assigned to each letter (or box)

$i = 1, \dots, n$. It is then clear that different letters bring different information. When there is randomness, we evaluate the *average* missing information per symbol by repeating our choice, say, $N$ times. As $N \to \infty$ we know that candy in the $i$-th box in $Nw_i$ cases but we do not know the order in which different possibilities appear. Total number of orders is $N!/\,\Pi_i(Nw_i)!$ and the information that we obtained from $N$ symbols is

$$I_N = k \ln\Big(N!/\,\Pi_i(Nw_i)!\Big) \approx -Nk \sum_i w_i \ln w_i + O(lnN) \ . \tag{275}$$

The missing information per problem (or per symbol in the language) coincides with the entropy (29):

$$I(w_1 \dots w_n) = \lim_{N\to\infty} I_N/N = -k \sum_{i=1}^{n} w_i \ln w_i \ . \tag{276}$$

Incidentally for English language the information per symbol is

$$-\sum_{i=a}^{z} w_i \log_2 w_i \approx 4.11\,\text{bits} \ .$$

The information (276) is zero for delta-distribution $w_i = \delta_{ij}$; it is generally less than the information (274) and coincides with it only for equal probabilities, $w_i = 1/n$, when the entropy is maximum. Indeed, equal probabilities we ascribe when there is no extra information, i.e. in a state of maximum ignorance. In this state, we get maximum information per symbol; any prior knowledge can reduce the information. Mathematically, the property

$$I(1/n, \dots, 1/n) \geq I(w_1 \dots w_n) \tag{277}$$

is called convexity. It follows from the fact that the function of a single variable $s(w) = -w \ln w$ is strictly downward convex (concave) since its second derivative, $-1/w$, is everywhere negative for positive $w$. For any concave function, the average over the set of points $w_i$ is less or equal to the function at the average value (so-called Jensen inequality):

$$\frac{1}{n}\sum_{i=1}^{n} s\left(w_i\right) \leq s\left(\frac{1}{n}\sum_{i=1}^{n} w_i\right) \ . \tag{278}$$

From here one gets the entropy inequality:

$$I(w_1 \dots w_n) = \sum_{i=1}^{n} s\left(w_i\right) \leq ns\left(\frac{1}{n}\sum_{i=1}^{n} w_i\right) = ns\left(\frac{1}{n}\right) = I\left(\frac{1}{n}, \dots, \frac{1}{n}\right) \ . \tag{279}$$

The relation (278) can be proven for any concave function. Indeed, the concavity condition states that the linear interpolation between two points $a, b$ lies everywhere below the function graph: $s(\lambda a + b - \lambda b) \geq \lambda s(a) + (1 - \lambda)s(b)$ for any $\lambda \in [0, 1]$. For $\lambda = 1/2$ it corresponds to (278) for $n = 2$. To get from $n = 2$ to arbitrary $n$ we use induction. For that end, we choose $\lambda = (n-1)/n$, $a = (n-1)^{-1} \sum_{i=1}^{n-1} w_i$ and $b = w_n$ to see that

$$\begin{aligned} s\left(\frac{1}{n}\sum_{i=1}^{n} w_i\right) &= s\left(\frac{n-1}{n}(n-1)^{-1}\sum_{i=1}^{n-1} w_i + \frac{w_n}{n}\right) \\ &\geq \frac{n-1}{n} s\left((n-1)^{-1}\sum_{i=1}^{n-1} w_i\right) + \frac{1}{n} s\left(w_n\right) \\ &\geq \frac{1}{n}\sum_{i=1}^{n-1} s\left(w_i\right) + \frac{1}{n} s\left(w_n\right) = \frac{1}{n}\sum_{i=1}^{n} s\left(w_i\right) \ . \end{aligned} \tag{280}$$

In the last line we used the truth of (278) for $n - 1$ to prove it for $n$.

Note that when $n \to \infty$ then (274) diverges while (276) may well be finite. We can generalize (276) for a continuous distribution by dividing into cells (that is considering a limit of discrete points). Here, different choices of variables to define equal cells give different definitions of information. It is in such a choice that physics (or other specific knowledge) enters. Physics (quantum mechanics) requires that for Hamiltonian system the equal volumes in phase space contain equal number of states, so the measure is uniform in canonical coordinates; we then write the missing information in terms of the phase space density, which may also depend on time:

$$I(t) = -\int \rho(p, q, t) \ln[\rho(p, q, t)]\, dp dq \ . \tag{281}$$

If the density of the discrete points in the continuous limit is inhomogeneous, say $m(\mathbf{x})$, then the proper generalization is

$$I(t) = -\int \rho(\mathbf{x}) \ln[\rho(\mathbf{x})/m(\mathbf{x})]\, d\mathbf{x} \ . \tag{282}$$

Note that (282) is invariant with respect to an arbitrary change of variables $\mathbf{x} \to \mathbf{y}(\mathbf{x})$ since $\rho(\mathbf{y})d\mathbf{y} = \rho(\mathbf{x})d\mathbf{x}$ and $m(\mathbf{y})d\mathbf{y} = m(\mathbf{x})d\mathbf{x}$ while (281) was invariant only with respect to canonical transformations (including a time evolution according to a Hamiltonian dynamics) that conserve the element of the phase-space volume.

Mention briefly the application of entropy in communication theory. Inequality (277) means that a communication channel transmitting bits (ones and zeros) on average can transmit no more than one unit of the information (276) per symbol. In other words, $\sum_{i=a}^{z} w_i \log_2 w_i$ gives the minimum number of bits per symbol needed to transmit the ensemble of messages. We can say that the information content of a symbol number $i$ is $\log_2(1/w_i)$, while the entropy is the mean information content per symbol. Note that less probable symbols have larger information content, but the mean information content for a given letter, $-w \ln w$, is maximal for $w = 1/e$. Different probability of letters suggests a way of signal compression by coding common letters by short sequences and infrequent letters by more lengthy combinations - lossless compressions like zip, gz and gif work this way (you may find it interesting to know that jpeg, mpeg, mp3 and telephone use lossy compression which removes information presumed to be unimportant for humans).

Apart from restrictions imposed by the statistics of symbols to be transferred, one also wish to characterize the quality of the channel. Note that in this context one can view measurements as messages about the value of the quantity we measure. Here, the message (measurement) A we receive gives the information about the event (quantity) B. That information must be related to the fact that having observed A increases the probability to have B comparing to the unconditional probability to observe B:

$$I(A, B) = \ln[P(B|A)/P(B)] \ .$$

Here $P(B|A)$ is the so-called conditional probability (of B in the presence of A). The conditional probability is related to the joint probability $P(A, B)$ by the evident formula $P(A, B) = P(B|A)P(A)$, which allows one to write the information in a symmetric form

$$I(A, B) = \ln \left[ \frac{[P(B, A)}{P(A)P(B)} \right] \ . \tag{283}$$

When $A$ and $B$ are independent then the conditional probability is independent of $A$ and information is zero. When they are dependent, $P(B, A) \geq P(A)P(B)$ so that that the information is always positive.

It is interesting to know how much information on average about $B$ one obtains by measuring $A$. Summing over all possible $B_1, \ldots, B_n$ and $A_1, \ldots, A_m$ we obtain Shannon's "mutual information" used to evaluate the quality of communication systems (or measurements)

$$I(A, B) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(A_i, B_j) \ln[P(B_j|A_i)/P(B_j)]$$

$$\to I(Z,Y)=\int dzdyp(z,y)\ln\left[\frac{p(z|y)}{p(y)}\right]=\int dzdy\, p(z,y)\ln\left[\frac{p(z,y)}{p(z)p(y)}\right] . \qquad (284)$$

Here we used $p(z,y) = p(z|y)p(y)$. If one is just interested in the channel as specified by $P(B|A)$ then one maximizes $I(A,B)$ over all choices of the source statistics $P(B)$ and call it channel capacity. Note that (284) is the particular case of multidimensional (282) where one takes $\mathbf{x} = (y,z)$, $m = p(z)p(y)$.

You probably noticed that (274,283) corresponds to the microcanonical description (19) giving information/entropy as a logarithm of the number of states, while (276,284) corresponds to the canonical description (29) giving it as an average.

So far, we defined information via the distribution. Now, we want to use the idea of information to get the distribution. Statistical mechanics is a systematic way of guessing, making use of incomplete information. The main problem is how to get the best guess for the probability distribution $\rho(p,q,t)$ based on any given information presented as $\langle R_j(p,q,t)\rangle = r_j$, i.e. as the expectation (mean) values of some dynamical quantities. Our distribution must contain *the whole truth* (i.e. all the given information) and *nothing but the truth* that is it must maximize the missing information $I$. This is to provide for the widest set of possibilities for future use, compatible with the existing information. Looking for the maximum of

$$I-\sum_j \lambda_j \langle R_j(p,q,t)\rangle = \int \rho(p,q,t)\{\ln[\rho(p,q,t)] - \sum_j \lambda_j R_j(p,q,t)\}\, dpdq\ ,$$

we obtain the distribution

$$\rho(p,q,t) = Z^{-1}\exp\Big[-\sum_j \lambda_j R_j(p,q,t)\Big]\ , \qquad (285)$$

where the normalization factor

$$Z(\lambda_i) = \int \exp\Big[-\sum_j \lambda_j R_j(p,q,t)\Big]\, dpdq\ ,$$

can be expressed via the measured quantities by using

$$\frac{\partial \ln Z}{\partial \lambda_i} = -r_i\ . \qquad (286)$$

For example, consider our initial "candy-in-the-box" problem (think of an impurity atom in a lattice if you prefer physics to candies). Let us denote the number of the box with the candy $j$. Different attempts give different $j$ (for impurity, think of X-ray with wavenumber $k$ scattering on the lattice) but on average after many attempts we find, say, $\langle\cos(kj)\rangle = 0.3$. Then

$$\rho(j) = Z^{-1}(\lambda)\exp[-\lambda\cos(kj)]$$

$$Z(\lambda) = \sum_{j=1}^{n}\exp[\lambda\cos(kj)]\,, \quad \langle\cos(kj)\rangle = d\log Z/d\lambda = 0.3\ .$$

We can explicitly solve this for $k \ll 1 \ll kn$ when one can approximate the sum by the integral so that $Z(\lambda) \approx nI_0(\lambda)$ where $I_0$ is the modified Bessel function. Equation $I_0'(\lambda) = 0.3I_0(\lambda)$ has an approximate solution $\lambda \approx 0.63$.

Note in passing that the set of equations (286) may be self-contradictory or insufficient so that the data do not allow to define the distribution or allow it non-uniquely. If, however, the solution exists then (281,285) define the missing information $I\{r_i\}$ which is analogous to thermodynamic entropy as a function of (measurable) macroscopic parameters. It is clear that $I$ have a tendency to increase whenever a constraint is removed (when we measure less quantities $R_i$).

If we know the given information at some time $t_1$ and want to make guesses about some other time $t_2$ then our information generally gets less relevant as the distance $|t_1 - t_2|$ increases. In the particular case of guessing the distribution in the phase space, the mechanism of loosing information is due to separation of trajectories described in Sect. 8. Indeed, if we know that at $t_1$ the system was in some region of the phase space, the set of trajectories started at $t_1$ from this region generally fills larger and larger regions as $|t_1 - t_2|$ increases. Therefore, missing information (i.e. entropy) increases with $|t_1 - t_2|$. Note that it works both into the future and into the past. Information approach allows one to see clearly that there is really no contradiction between the reversibility of equations of motion and the growth of entropy.

Yet there is one class of quantities where information does not age. They are integrals of motion. A situation in which only integrals of motion are known is called equilibrium. The distribution (285) takes the canonical form (13,14) in equilibrium. On the other hand, taking micro-canonical as constant over the constant-energy surface corresponds to the same approach of not adding any additional information to what is known (energy).

From the information point of view, the statement that systems approach equilibrium is equivalent to saying that all information is forgotten except the integrals of motion. If, however, we possess the information about averages of quantities that are not integrals of motion and those averages do not coincide with their equilibrium values then the distribution (285) deviates from equilibrium. Examples are currents, velocity or temperature gradients like considered in kinetics.

The concept of entropy as missing information[24] allows one to understand that entropy does not really decrease in the system with Maxwell demon or any other information-processing device (indeed, if at the beginning one has an information on position or velocity of any molecule, then the entropy was less by this amount from the start; after using and processing the information the entropy can only increase). Consider, for instance, a particle in the box. If we know that it is in one half then entropy (the logarithm of *available* states) is $\ln(V/2)$. That also teaches us that information has thermodynamic (energetic) value: by placing a piston at the half of the box and allowing particle to hit and move it we can get the work $T\Delta S = T\ln 2$ done; on the other hand, to get such an information one must make a measurement whose minimum energetic cost is $T\Delta S = T\ln 2$ (that was considered by Szilard in 1929 who also introduced "bit" as a unit of information).

More details can be found in Katz, Sects. 2-5, Sethna Sect. 5.3 and Kardar I, Problem 2.6.

## 8.4 Central limit theorem and large deviations

Mathematical statement underlying most of the statistical physics in the thermodynamic limit is the central limit theorem, which states that the sum of many independent random numbers has Gaussian probability distribution. Recently, however, we are more and more interested in the statistics of not very large systems or in the statistics of really large fluctuations. To answer such questions, here we discuss the sum of random numbers in more detail. Consider the variable $X$ which is a sum of many independent identically distributed (iid) random numbers $X = \sum_1^N y_i$. Its mean value $\langle X\rangle = N\langle y\rangle$ grows linearly with $N$. Here we show that its fluctuations $X - \langle X\rangle$ on the scale less and comparable with $\mathcal{O}(N^{1/2})$ are governed by the Central Limit Theorem that states that $(X - \langle X\rangle)/N^{1/2}$ becomes for large $N$ a Gaussian random variable with variance $\langle y^2\rangle - \langle y\rangle^2 \equiv \Delta$. Note that the statistics of the quantities that we sum, $y_i$, can be quite arbitrary, the only requirements

---

[24] that entropy is not a property of the system but of our knowledge about the system

that we must impose is that the first two moments, the mean $\langle y\rangle$ and the variance $\Delta$, are finite. Finally, the fluctuations $X-\langle X\rangle$ on the larger scale $\mathcal{O}(N)$ are governed by the Large Deviation Theorem that states that the PDF of $X$ has asymptotically the form

$$\mathcal{P}(X) \ \propto\ \mathrm{e}^{-NH(X/N-\langle y\rangle)}\,. \tag{287}$$

To show this, let us characterize $y$ by its generating function $\langle \mathrm{e}^{\,zy}\rangle \equiv \mathrm{e}^{S(z)}$ (assuming that the mean value exists for all complex $z$). The derivatives of $S(z)$ at zero determine the cumulants of $y$:

$$\begin{aligned}
&S(z)=\langle\exp(zy)\rangle = 1+\sum_{n=1}^{\infty}\frac{a^n}{n!}\langle y^n\rangle\,,\\
&\ln\langle\exp(zy)\rangle = -\sum_{n=1}\frac{1}{n}\left(1-\langle\exp(zy)\rangle\right)^n = -\sum_{n=1}\frac{1}{n}\left(-\sum_{m=1}^{\infty}\frac{z^m}{m!}\langle y^m\rangle\right)^n\\
&=z\langle y\rangle+\Big(\langle y^2\rangle-\langle y\rangle^2\Big)\frac{z^2}{2!}+\ldots=\sum_{n=1}^{\infty}\frac{z^n}{n!}\langle y^n\rangle_c\ .
\end{aligned} \tag{288}$$

The basic statement is that because all $y$-s in the sum are independent then the generating function $\langle \mathrm{e}^{\,zX}\rangle$ of the moments of $X$ has exponential dependence on $N$: $\langle \mathrm{e}^{\,zX}\rangle = \mathrm{e}^{NS(z)}$. The PDF $\mathcal{P}(X)$ is then given by the inverse Laplace transform $\frac{1}{2\pi i}\int \mathrm{e}^{-z\,X+NS(z)}\,dz$ with the integral over any axis parallel to the imaginary one. For large $N$, the integral is dominated by the saddle point $z_0$ such that $S'(z_0)=X/N$. This is similar to representing the sum (20) by its largest term. If there are several saddle-points, the result is dominated by the one giving the largest probability. We now substitute $X=NS'(z_0)$ into $zX+NS(z)$, and obtain the large deviation relation (287) with

$$H=-S(z_0)+z_0S'(z_0)\ . \tag{289}$$

We see that $-H$ and $S$ are related by the Legendre transform. Note that $NdH/dX=z_0(X)$ and $N^2d^2H/dX^2=Ndz_0/dX=1/S''(z_0)$. The function $H$ of the variable $X/N-\langle y\rangle$ is called Cramér or rate function since it measures the rate of probability decay with the growth of $N$ for every $X/N$. It is also sometimes called entropy function since it is a logarithm of probability.

Several important properties of $H$ can be established independently of the distribution $\mathcal{P}(y)$ or $S(z)$. It is a convex function as long as $S(z)$ is a convex function since their second derivatives have the same sign. It is straightforward to see that the logarithm of the generating function has a positive second derivative (at least for real $z$):

$$\begin{aligned} S''(z) &= \frac{d^2}{dz^2}\ln\int e^{zy}\mathcal{P}(y)\,dy \\ &= \frac{\int y^2e^{zy}\mathcal{P}(y)\,dy\int e^{zy}\mathcal{P}(y)\,dy-[\int ye^{zy}\mathcal{P}(y)\ ,dy]^2}{[\int e^{zy}\mathcal{P}(y)\,dy]^2}\geq 0\ . \qquad (290)\end{aligned}$$

This uses the CauchyBunyakovskySchwarz inequality which is a generalization of $\langle y^2\rangle \geq \langle y\rangle^2$. Also, $H$ takes its minimum at zero, i.e. for $X$ taking its mean value $\langle X\rangle = N\langle y\rangle = NS'(0)$, which corresponds to $z_0 = 0$. Since $S(0) = 0$ then the minimal value of $H$ is zero. The entropy is quadratic around its minimum with $H''(0) = \Delta^{-1}$, where $\Delta = S''(0)$ is the variance of $y$. Quadratic entropy means Gaussian probability near the maximum — this statement is (loosely speaking) the essence of the central limit theorem. In the particular case of Gaussian $y$, $X$ is Gaussian as well. Non-Gaussianity of the $y$'s leads to a non-quadratic behavior of $H$ when deviations of $X/N$ from the mean are large, of the order of $\Delta/S'''(0)$.

A simple example is provided by the statistics of the kinetic energy, $E = \sum_1^N p_i^2/2$, of $N$ classical identical unit-mass particles in 1d. We considered similar problem in the Section 2.2 in the micro-canonical approach and thermodynamic limit $N \to \infty$. Let us now look using canonical Gibbs distribution which is Gaussian for momenta:

$$\rho(p_1,\ldots,p_N) = (2\pi T)^{-N/2}\exp\left(-\sum_1^N p_i^2/2T\right)\ .$$

The energy probability for any $N$ is done by integration, using spherical coordinates in the momentum space:

$$\begin{aligned} \rho(E,N) &= \int \rho(p_1,\ldots,p_N)\delta\left(E-\sum_1^N p_i^2/2\right)dp_1\ldots dp_N \\ &= \left(\frac{E}{T}\right)^{N/2}\frac{\exp(-E/T)}{E\Gamma(N/2)}\ . \qquad (291)\end{aligned}$$

Plotting it for different $N$, one can appreciate how the thermodynamic limit appears. Taking the logarithm and using the Stirling formula one gets the large-deviation form for the energy $R = E/\bar{E}$, normalized by the mean energy

$\bar{E} = NT/2$:

$$\ln \rho(E,N) = \frac{N}{2} \ln \frac{RN}{2} - \ln \frac{N}{2}! - \frac{RN}{2} \approx \frac{N}{2}(1 - R + \ln R) \ . \quad (292)$$

This expression has a maximum at $R = 1$ i.e the most probable value is the mean energy. The probability of $R$ is Gaussian near maximum when $R - 1 \le N^{-1/2}$ and non-Gaussian for larger deviations. Notice that this function is not symmetric with respect to the minimum, it has logarithmic asymptotic at zero and linear asymptotic at infinity.

One can generalize the central limit theorem and the large-deviation approach in two directions: i) for non-identical variables $y_i$, as long as all variances are finite and none dominates the limit $N \to \infty$, it still works with the mean and the variance of $X$ being given by the average of means and variances of $y_i$; ii) if $y_i$ is correlated with a finite number of neighboring variables, one can group such "correlated sums" into new variables which can be considered independent.

Basic books

L. D. Landau and E. M. Lifshitz, Statistical Physics Part 1.
R. K. Pathria, Statistical Mechanics.
R. Kubo, Statistical Mechanics.
K. Huang, Statistical Mechanics.
C. Kittel, Elementary Statistical Physics.

Additional reading

S.-K. Ma, Statistical Mechanics.
A. Katz, Principles of Statistical Mechanics.
J. Cardy, Scaling and renormalization in statistical physics.
K. Wilson and J. Kogut, The renormalization group and the $\epsilon$-expansion, *Physics Reports* **12**, 75-200 (1972).
M. Kardar, Statistical Physics of Particles, Statistical Physics of Fields.
J. Sethna, Entropy, Order Parameters and Complexity.
J. Kurchan, Six out of equilibrium lectures, http://xxx.tau.ac.il/abs/0901.1271