Statistical Physics

G. Falkovich
http://www.weizmann.ac.il/home/fnfal/

Contents

1 Basic principles.

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

1.1 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: $w = \lim_{T\to\infty} \Delta t/T$. 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:

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

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. Statistical independence of macroscopic subsystems at the absence of long-range forces means that the distribution for a composite system ρ_{12} is factorized: $\rho_{12} = \rho_1\rho_2$.

Now, we take the ensemble of identical systems starting from different points in phase space. If the motion is considered for not very large time it is conservative and can be described by the Hamiltonian dynamics (that is $\dot{q}_i = \partial\mathcal{H}/\partial p_i$ and $\dot{p}_i = -\partial\mathcal{H}/\partial q_i$ then the flow in the phase space is incompressible: $div\,\mathbf{v} = \partial\dot{q}_i/\partial q_i + \partial\dot{p}_i/\partial p_i = 0$. That gives the Liouville theorem: $d\rho/dt = \partial\rho/\partial t + (\mathbf{v}\cdot\nabla)\rho = 0$ that is the statistical distribution is conserved along the phase trajectories of any subsystem. As a result, equilibrium ρ 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 $\mathbf{M}(p,q)$:

$$\ln\rho_a = \alpha_a + \beta E_a(p,q) + \mathbf{c}\cdot\mathbf{P}_a(p,q) + \mathbf{d}\cdot\mathbf{M}(p,q) \ . \tag{2}$$

Here α_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*

$$\rho(p,q) = A\exp[-\beta E(p,q)]\ . \tag{3}$$

For 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*:

$$\rho(p,q) = A\delta[E(p,q)-E_0]\ . \tag{4}$$

Usually one considers the energy fixed with the accuracy Δ so that the microcanonical distribution is

$$\rho = \begin{cases} 1/\Gamma & \text{for } E \in (E_0, E_0+\Delta) \\ 0 & \text{for } E \notin (E_0, E_0+\Delta)\,, \end{cases} \tag{5}$$

where Γ is the volume of the phase space occupied by the system

$$\Gamma(E,V,N,\Delta) = \int_{E<\mathcal{H}<E+\Delta} d^{3N}p\,d^{3N}q\ . \tag{6}$$

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

$$\Gamma(E,V,N,\Delta) \propto E^{3N/2-1}V^N\Delta/(3N/2-1)! \approx (E/N)^{3N/2}V^N\Delta\ . \tag{7}$$

See Landau & Lifshitz, Sects 1-4.

1.2 Microcanonical distribution

One can link statistical physics with thermodynamics using either canonical or microcanonical distribution. We start from the latter and introduce the entropy as

$$S(E,V,N) = \ln\Gamma(E,V,N)\ . \tag{8}$$

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

Noninteracting subsystems are statistically independent so that the statistical weight of the composite system is a product and 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, Δ, is much less than the total energy E. Then

$$\Gamma(E) = \sum_{i=1}^{E/\Delta} \Gamma_1(E_i)\Gamma_2(E - E_i)\ . \tag{9}$$

We denote $\bar{E}_1, \bar{E}_2 = E - \bar{E}_1$ the values that correspond to the maximal term in the sum (9), the extremum condition is evidently $(\partial S_1/\partial E_1)_{\bar{E}_1} = (\partial S_2/\partial E_2)_{\bar{E}_2}$. 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.

Identification with the thermodynamic entropy can be done considering any system, for instance, an ideal gas (7): $S(E, V, N) = (3N/2)\ln E + f(N, V)$. 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 (8) depends on the units. Proper quantitative definition comes from quantum physics with Γ 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 $\Delta p \Delta q = 2\pi\hbar$. Note in passing that quantum particles (atoms and molecules) are indistinguishable so one needs to divide Γ (7) by the number of transmutations $N!$ which makes the resulting entropy of the ideal gas extensive: $S(E, V, N) = (3N/2)\ln E/N + N\ln V/N + \text{const}$[1].

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 ϵ. 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:

---

[1] 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

$\Gamma(N,L) = C_N^L = N!/L!(N-L)!$. 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) $T^{-1} = \partial S/\partial E = \epsilon^{-1}(\partial/\partial L)\ln[N!/L!(N-L)!] = \epsilon^{-1}\ln(N-L)/L$ and specific heat $C = dE/dT = N(\epsilon/T)^2 2\cosh^{-1}(\epsilon/T)$. Note that the ratio of the number of particles on the upper level to those on the lower level is $\exp(-\epsilon/T)$ (Boltzmann relation). Specific heat turns into zero both at low temperatures (too small portions of energy are "in circulation") and in high temperatures (occupation numbers of two levels already close to equal).

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

1.3 Canonical distribution

We now consider small subsystem or 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. Assuming that all the states of the thermostat are equally likely to occur we see that the probability should be directly proportional to the statistical weight of the thermostat $\Gamma_0(E_0-E)$ where we evidently assume that $E \ll E_0$, expand $\Gamma_0(E_0-E) = \exp[S_0(E_0-E)] \approx \exp[S_0(E_0) - E/T)]$ and obtain

$$w_a(E) = Z^{-1}\exp(-E/T) \ , \tag{10}$$

$$Z = \sum_a \exp(-E_a/T) \ . \tag{11}$$

Note that there is no trace of 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. This is the derivation of the canonical distribution from the microcanonical one which allows us to specify $\beta = 1/T$ in (2,3).

The probability to have a given energy is the probability of the state (10) times the number of states:

$$W(E) = \Gamma(E)w_n = \Gamma(E)Z^{-1}\exp(-E/T) \ . \tag{12}$$

Here $\Gamma(E)$ grows fast while $\exp(-E/T)$ decays fast when the energy E grows. As a result, $W(E)$ is concentrated in a very narrow peak and the energy fluctuations around $\bar{E}$ are very small (see Sect. 1.6 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 (10) tells that the probability of a given microstate exponentially decays with the energy of the state while (12) 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 system in the state a is then given by the average number of systems $\bar{n}_a$ in this state divided by the total number of systems N. The set of occupation numbers $\{n_a\} = (n_0, n_1, n_2 \ldots)$ satisfies obvious conditions

$$\sum_a n_a = N \ , \qquad \sum_a E_a n_a = E = \epsilon N \ . \tag{13}$$

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:

$$\bar{n}_a = \frac{\sum n_a W\{n_a\}}{\sum W\{n_a\}} \ , \tag{14}$$

where summation goes over all the sets that satisfy (13). We assume that in the limit when $N, n_a \to \infty$ the main contribution into (14) is given by the most probable distribution which is found by looking at 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 $\ln W = N \ln N - \sum_a n_a \ln n_a$ and the extremum n_a^* corresponds to $\ln n_a^* = -\alpha - 1 - \beta E_a$ which gives

$$\frac{n_a^*}{N} = \frac{\exp(-\beta E_a)}{\sum_a \exp(-\beta E_a)} \ . \tag{15}$$

The parameter β is given implicitly by the relation

$$\frac{E}{N} = \epsilon = \frac{\sum_a E_a \exp(-\beta E_a)}{\sum_a \exp(-\beta E_a)} \ . \tag{16}$$

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

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):

$$F(T,V,N) = -T\ln Z(T,V,N)\ . \tag{17}$$

To prove that, differentiate the identity $\sum_a \exp[(F-E_a)/T] = 1$ 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 is $w_a = 1/\Gamma$ that is

$$\begin{aligned} S &= -\langle w_a\rangle = -\sum w_a \ln w_a \qquad (18)\\ &= \sum w_a(E_a/T + \ln Z) = E/T + \ln Z\ . \end{aligned}$$

See Landau & Lifshitz (Sects 31,36).

1.4 Two simple examples

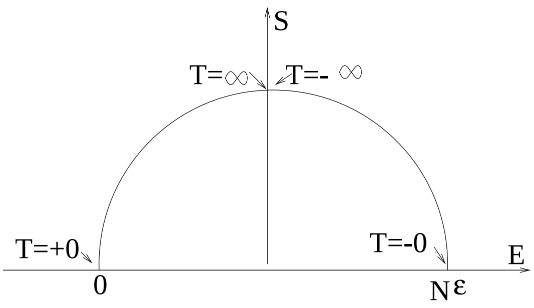
Here we consider two examples with the simplest structures of energy levels to illustrate the use of microcanonical and canonical distributions.

1.4.1 Two-level system

Assume levels 0 and ϵ. Remind that in Sect. 1.2 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:

$$T^{-1} = \frac{\partial S}{\partial E} = \epsilon^{-1}(\partial/\partial L)\ln[N!/L!(N-L)!] = \epsilon^{-1}\ln(N-L)/L\ . \tag{19}$$

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. Available (non-equilibrium) states lie below the $S(E)$ plot, notice that for the right (negative-temperature) part the entropy maximum corresponds to the energy maximum as well. 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.

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. Let us derive the generalized force M that corresponds to the magnetic field and determines the work done under the change of magnetic field: $dE = TdS - MdH$. 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 is calculated according to (34) and is called magnetization (or magnetic moment of the system):

$$M=-\left(\frac{\partial E}{\partial H}\right)_S=\mu(N_+-N_-)=N\mu\frac{\exp(\mu H/T)-\exp(-\mu H/T)}{\exp(\mu H/T)+\exp(-\mu H/T)}\ . \qquad (20)$$

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: $\chi(T)=(\partial M/\partial H)_{H=0}=N\mu^2/T$.

At weak fields and positive temperature, $\mu H \ll T$, (20) gives the formula for the so-called Pauli paramagnetism

$$\frac{M}{N\mu}=\frac{\mu H}{T}\ . \qquad (21)$$

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 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. Note that the entropy of the spin system does not change when the field changes slowly comparatively to the spin-spin relaxation and fast comparatively to the spin-lattice relaxation.

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

$$Z(T,N)=\sum_{L=0}^{N}C_N^L\exp[-L\epsilon/T]=[1+\exp(-\epsilon/T)]^N\ , \qquad (22)$$

$$F(T,N)=-T\ln Z=-NT\ln[1+\exp(-\epsilon/T)]\ . \qquad (23)$$

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

$$\bar{E} \ = \ Z^{-1}\sum_a E_a\exp(-\beta E_a)=-\frac{\partial\ln Z}{\partial\beta}=T^2\frac{\partial\ln Z}{\partial T} \qquad (24)$$

$$
\begin{aligned}
&= \frac{N\epsilon}{1+\exp(\epsilon/T)}\ , && (25)\\
C &= \frac{dE}{dT} = \frac{N\exp(\epsilon/T)}{[1+\exp(\epsilon/T)]^2}\frac{\epsilon^2}{T^2}\ . && (26)
\end{aligned}
$$

Note that (24) is a general formula which we shall use in the future. Specific heat turns into zero both at low temperatures (too small portions of energy are "in circulation") and in high temperatures (occupation numbers of two levels already close to equal). More details can be found in Kittel, Section 24 and Pathria, Section 3.9.

1.4.2 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 described by the Hamiltonian

$$
\mathcal{H}(q,p) = \frac{1}{2m}\Big(p^2+\omega^2 q^2 m^2\Big)\ . \qquad (27)
$$

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

$$
Z_1(T) = (2\pi\hbar)^{-1}\int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dq \exp(-\mathcal{H}/T) = \frac{T}{\hbar\omega}\ . \qquad (28)
$$

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 (24): $E = NT$ — this is an example of the equipartition (every oscillator has two degrees of freedom with $T/2$ energy for each)[2]. The thermodynamic equations of state are $\mu(T) = T\ln(\hbar\omega/T)$ 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 which is given by the Gibbs distribution using the potential energy:

$$
dw_q = \omega(2\pi T)^{-1/2}\exp(-\omega^2 q^2/2T)dq\ . \qquad (29)
$$

---

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

Using kinetic energy and simply replacing $q \to p/\omega$ one obtains a similar formula $dw_p = (2\pi T)^{-1/2} \exp(-p^2/2T)dp$ 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

$$Z_1(T) = \sum_{n=0}^{\infty} \exp[-\hbar\omega(n + 1/2)/T] = 2\sinh^{-1}(\hbar\omega/2T) \tag{30}$$

gives again $Z(T, N) = Z_1^N(T)$ and $F(T, N) = NT \ln[\sinh(\hbar\omega/2T)/2] = N\hbar\omega/2 + NT \ln[1 - \exp(-\hbar\omega/T)$. 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 is as follows: $C_P = C_V = N(\hbar\omega/T)^2 \exp(\hbar\omega/T)[\exp(\hbar\omega/T) - 1]^{-2}$. Comparing with (26) 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 (29) must be obtained by summing the wave functions of quantum oscillator with the respective probabilities:

$$dw_q = a\,dq \sum_{n=0}^{\infty} |\psi_n(q)|^2 \exp[-\hbar\omega(n + 1/2)/T] \; . \tag{31}$$

Here a is the normalization factor. Straightforward (and beautiful) calculation of (31) 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:

$$dw_q = \left(\frac{\omega}{\pi\hbar} \tanh \frac{\hbar\omega}{2T}\right)^{1/2} \exp\left(-q^2 \frac{\omega}{\hbar} \tanh \frac{\hbar\omega}{2T}\right) dq \; . \tag{32}$$

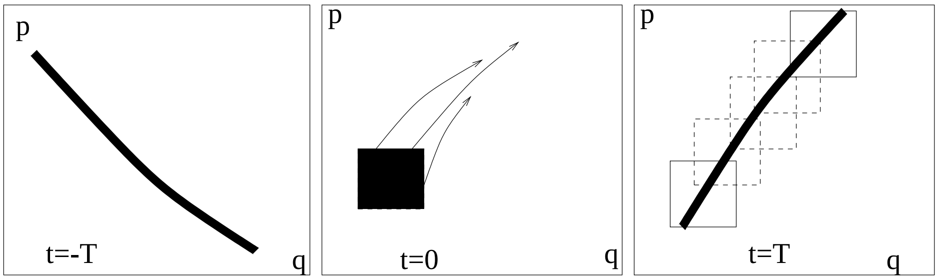
At $\hbar\omega \ll T$ it coincides with (29) while at the opposite (quantum) limit gives $dw_q = (\omega/\pi\hbar)^{1/2} \exp(-q^2\omega/\hbar)dq$ which is a purely quantum formula $|\psi_0|^2$ for the ground state of the oscillator.

See also Pathria Sect. 3.7 for more details.

1.5 Entropy

By definition, entropy 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 (16)] 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. Note that 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.

Dynamics (classical and quantum) is time reversible. 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). 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. The dynamical background of entropy growth is the separation of trajectories in phase space so that trajectories started from a small finite region fill larger and larger regions of phase space as time proceeds. On the figure, one can see how the 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). Rectangles in the right box show finite resolution (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 but the fact of space filling persists.



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. If temporal changes are slow enough then the entropy does not change i.e. the process is adiabatic. Indeed, if we have some parameter $\lambda(t)$ slowly changing with time then positivity of $\dot S = dS/dt$ requires that the expansion of $\dot S(\dot\lambda)$ starts from the second term,

$$\frac{dS}{dt}=\frac{dS}{d\lambda}\cdot\frac{d\lambda}{dt}=A\left(\frac{d\lambda}{dt}\right)^2 \qquad \Rightarrow \qquad \frac{dS}{d\lambda}=A\frac{d\lambda}{dt}\,. \tag{33}$$

We see that when $d\lambda/dt$ goes to zero, entropy is getting independent of λ. That means that we can change λ (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 λ (called coordinate) one changes the energy levels E_a and the total energy. Respective force (pressure when λ is volume, magnetic or electric moments when λ is the respective field) is obtained as the average (over the equilibrium statistical distribution) of the energy derivative with respect to λ and is equal to the derivative of the thermodynamic energy at constant entropy because the probabilities $w_a = 1/\Gamma$ does not change:

$$\overline{\frac{\partial\mathcal{H}(p,q,\lambda)}{\partial\lambda}}=\sum_a w_a\frac{\partial E_a}{\partial\lambda}=\frac{\partial}{\partial\lambda}\sum_a w_aE_a=\left(\frac{\partial E(S,\lambda,\ldots)}{\partial\lambda}\right)_S\,. \tag{34}$$

Here $\mathcal{H}(p,q,\lambda)$ is the microscopic Hamiltonian while $E(S,\lambda,\ldots)$ is the thermodynamic energy. Note that in an adiabatic process all w_a are assumed to be constant 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.

The last statement we make here about entropy is the third law of thermodynamics (Nernst theorem) which 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 $S = 0$. Since generally the degeneracy of the ground state 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. We shall discuss it later considering Debye theory of solids. See Huang (Section 9.4) for more details.

1.6 Grand canonical ensemble

Let us now repeat the derivation we done in Sect. 1.3 but in more detail and considering also the fluctuations in the particle number N. 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 whole system. Let us first do the expansion up to the first-order terms as in (10,11)

$$\begin{aligned} w_{aN} &= A\exp[S(E_0 - E_{aN}, N_0 - N)] = A\exp[S(E_0, N_0) + (\mu N - E_{aN})/T] \\ &= \exp[(\Omega + \mu N - E_{aN})/T] \ . \end{aligned} \tag{35}$$

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

$$\Omega(T, V, \mu) = -T \ln \sum_N \exp(\mu N/T) \sum_a \exp(-E_{aN})/T) \ . \tag{36}$$

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 (35) coincides with the canonical distribution $w_a = \exp[(F - E_a)/T]$.

To describe fluctuations one needs to expand further 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 β

the identity $\overline{E-\bar{E}}=0$. For this purpose one can use canonical distribution and get

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

$$\overline{(E-\bar{E})^2}=-\frac{\partial\bar{E}}{\partial\beta}=T^2C_V\ . \tag{37}$$

Since both $\bar{E}$ and C_V are proportional to N then the relative fluctuations are small indeed: $\overline{(E-\bar{E})^2}/\bar{E}^2\propto N^{-1}$. In what follows (as in the most of what preceded) we do not distinguish between E and $\bar{E}$.

Let us now discuss the fluctuations of particle number. One gets the probability to have N particles by summing (35) over a: $W(N)\propto\exp[\mu(T,V)N-F(T,V,N)]$ 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:

$$\overline{(N-\bar{N})^2}=2T\left(\frac{\partial^2F}{\partial N^2}\right)^{-1}=2TNv^{-2}\left(\frac{\partial P}{\partial v}\right)^{-1}\propto N\ . \tag{38}$$

Here we used the fact that $F(T,V,N)=Nf(T,v)$ with $v=V/N$ and substituted the derivatives calculated at fixed V: $(\partial F/\partial N)_V=f(v)-v\partial f/\partial v$ and $(\partial^2F/\partial N^2)_V=N^{-1}v^2\partial^2f/\partial v^2=-N^{-1}v^2\partial P(v)/\partial v$. 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. Note that any extensive quantity $f=\sum_{i=1}^N f_i$ which is a sum over independent subsystems (i.e. $\overline{f_if_k}=\bar{f}_i\bar{f}_k$) have a small relative fluctuation: $(\overline{f^2}-\bar{f}^2)/\bar{f}^2\propto 1/N$.

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

2 Gases

Here we consider systems with the kinetic energy exceeding the potential energy of inter-particle interactions: $\langle U(\mathbf{r}_1 - \mathbf{r}_2)\rangle \ll \langle mv^2/2\rangle$.

2.1 Ideal Gases

We start from neglecting the potential energy of interaction completely. Note though that quantum effect does not allow one to consider particles completely independent. The absence of any interaction allows one to treat any molecule as subsystem and apply to it the Gibbs canonical distribution: the average number of molecules in a given state is $\bar{n}_a = NZ^{-1}\exp(-\epsilon_a/T)$ which is called Boltzmann distribution. One can also use grand canonical ensemble considering all molecules in the same state as a subsystem with a non-fixed number of particles. Using the distribution (35) with $N = n_a$ and $E = n_a\epsilon_a$ one expresses the probability of occupation numbers via the chemical potential: $w(n_a) = \exp\{\,\beta[\Omega_a + n_a(\mu - \epsilon_a)]\}$.

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 are given by the Boltzmann distribution in the form

$$\bar{n}_a = \exp\left(\frac{\mu - \epsilon_a}{T}\right) . \tag{39}$$

2.1.1 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/h^3 to be much larger than the number of molecules N. Since the typical momentum is $p \simeq \sqrt{mT}$ we get the condition

$$(mT)^{3/2} \gg h^3 n . \tag{40}$$

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 $T \gg 0.5K$. Another way to interpret (40) is to say that the mean distance between molecules $n^{-1/3}$ must be much larger than the wavelength h/p. In this case, one can pass from the distribution over the quantum states to the distribution in the phase space:

$$\bar{n}(p,q) = \exp\left[\frac{\mu - \epsilon(p,q)}{T}\right] . \tag{41}$$

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 (40). 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 $n(z) = n(0)\exp(-mgz/T)$.

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

$$F = -NT\ln\left[\frac{e}{N}\sum_a \exp(-\epsilon_a/T)\right] . \tag{42}$$

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 ϵ'_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:

$$F = -NT\ln\left[\frac{eV}{N}\left(\frac{mT}{2\pi\hbar^2}\right)^{3/2}\sum_a \exp(-\epsilon'_a/T)\right] . \tag{43}$$

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 $\varepsilon_0 = e^2/a_B = me^4/\hbar^2$ 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

$$F = -NT \ln \left[\frac{eV}{N} \left(\frac{mT}{2\pi\hbar^2} \right)^{3/2} \right] = -NT \ln \frac{eV}{N} - Nc_v T \ln T - N\zeta T \, , \quad (44)$$

$$c_v = 3/2 \, , \quad \zeta = \frac{3}{2} \ln \frac{m}{2\pi\hbar^2} \; . \quad (45)$$

Here ζ 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

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

Using such μ we get $(mT)^{3/2} \gg h^3 n$. Note that $\mu < 0$.

If there is a nonzero spin the level has a degeneracy $2S + 1$ which adds $\zeta_S = \ln(2S + 1)$ to the chemical constant (45). If both L and S are nonzero then the total angular momentum J determines the fine structure of levels ϵ_J (generally comparable with the room temperature — typically, $\epsilon_J/k \simeq 200 \div 300K$). Every such level has a degeneracy $2J + 1$ so that the respective partition function

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

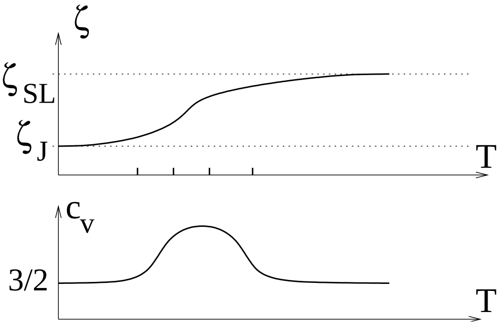
Without actually specifying ϵ_J we can determine this sum in two limits of large and small temperature. If $\forall J$ one has $T \gg \epsilon_J$, then $\exp(-\epsilon_J/T) \approx 1$ and $z = (2S + 1)(2L + 1)$ which is the total number of components of the fine level structure. In this case

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

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

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

where J 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.

Specific heat of diatomic molecules. We need to calculate the sum over the internal degrees of freedom in (43). 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 ϵ'_a characterized by two quantum numbers, j and K:

$$\epsilon_{jk} = \hbar\omega(j + 1/2) + \Big(\hbar^2/2I\Big) K(K+1) \ . \tag{46}$$

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 $I \simeq Ma_B^2$. We may thus estimate the typical energies of vibrations and rotations as follows:

$$\hbar\omega \simeq \varepsilon_0 \sqrt{\frac{m}{M}}\,, \qquad \frac{\hbar^2}{I} \simeq \varepsilon_0 \frac{m}{M}\ . \tag{47}$$

Since $m/M \simeq 10^{-4}$ then that 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 $\hbar^2/Ik \simeq 30\,K$.

The harmonic oscillator was considered in in Sect. 1.4.2. 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 (24): $E = NT$. 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

$$Z_1(T) = \sum_{n=0}^{\infty} \exp[-\hbar\omega(n+1/2)/T] = 2\sinh^{-1}(\hbar\omega/2T) \tag{48}$$

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

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

$$z_{rot} = \sum_K (2K+1)\exp\left(-\frac{\hbar^2 K(K+1)}{2IT}\right)\ . \tag{49}$$

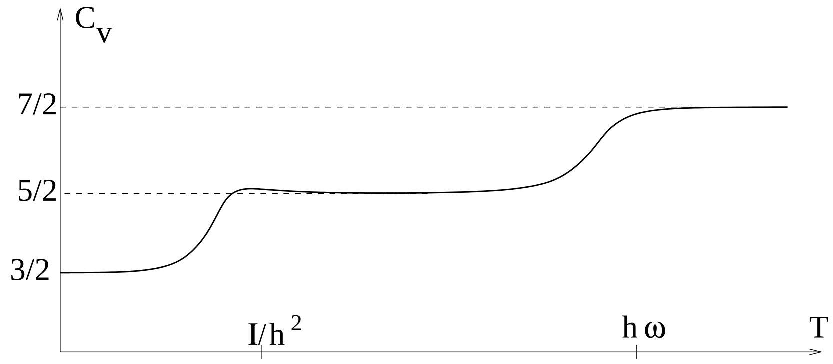
Again, when temperature is much smaller than the distance to the first level, $T \ll \hbar^2/2I$, the specific heat must be exponentially small. Indeed, retaining only two first terms in the sum (49), we get $z_{rot} = 1 + 3\exp(-\hbar^2/IT)$ which gives in the same approximation $F_{rot} = -3NT\exp(-\hbar^2/IT)$ 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 (49). They can be treated quasi-classically replacing the sum by the integral:

$$z_{rot} = \int_0^\infty dK(2K+1)\exp\left(-\frac{\hbar^2 K(K+1)}{2IT}\right) = \frac{2IT}{\hbar^2}\ . \tag{50}$$

That gives the constant specific heat $c_{rot} = 1$. 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 so that the level density (which is actually c_v) decreases.

For (non-linear) molecules with $\mathcal{N} > 2$ atoms we have 3 translations, 3 rotations and $6\mathcal{N} - 6$ vibrational degrees of freedom ($3n$ momenta and out of total $3n$ 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.

2.2 Fermi and Bose gases

Like we did at the beginning of the Section 2.1 we consider all particles at the same quantum state as Gibbs subsystem and apply the grand canonical distribution with the potential

$$\Omega_a = -T \ln \sum_{n_a} \exp[n_a(\mu - \epsilon_a)/T] \ . \tag{51}$$

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 Ω depends on T, V, μ. The average number of particles in the state with the energy ϵ is thus

$$\bar{n}(\epsilon) = -\frac{\partial \Omega_a}{\partial \mu} = \frac{1}{\exp[\beta(\epsilon - \mu)] \pm 1} \ . \tag{52}$$

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 (39). The thermodynamic potential of the whole system is obtained by summing over the states

$$\Omega = \mp T \sum_a \ln \left[1 \pm e^{\beta(\mu - \epsilon_a)}\right] \ . \tag{53}$$

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 recent 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 (52) by

$$dN(p, q) = \frac{g dp_x dp_y dp_z dx dy dx h^{-3}}{\exp[\beta(\epsilon - \mu)] \pm 1} \ . \tag{54}$$

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

$$dN(\epsilon) = \frac{gVm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\frac{\sqrt{\epsilon}\,d\epsilon}{\exp[\beta(\epsilon-\mu)]\pm 1}\ . \tag{55}$$

In the same way we rewrite (53):

$$\begin{aligned}\Omega &= \mp\frac{gVTm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\int_0^\infty \sqrt{\epsilon}\ln\left[1\pm e^{\beta(\mu-\epsilon)}\right]\,d\epsilon \\ &= -\frac{2}{3}\frac{gVm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\int_0^\infty \frac{\epsilon^{3/2}d\epsilon}{\exp[\beta(\epsilon-\mu)]\pm 1} = -\frac{2}{3}E.\end{aligned} \tag{56}$$

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

$$PV = \frac{2}{3}E\ . \tag{57}$$

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):

$$P = \sum_{i=1}^N \frac{p_{ix}^2}{ml^3} = \sum_{i=1}^N \frac{p_i^2}{3ml^3} = \frac{2E}{3V}\ . \tag{58}$$

In the limit of Boltzmann statistics we have $E = 3NT/2$ so that (57) reproduces $PV = NT$. Let us obtain the (small) quantum corrections to the pressure assuming $\exp(\mu/T) \ll 1$. Expanding integral in (56)

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

and substituting Boltzmann expression for μ we get

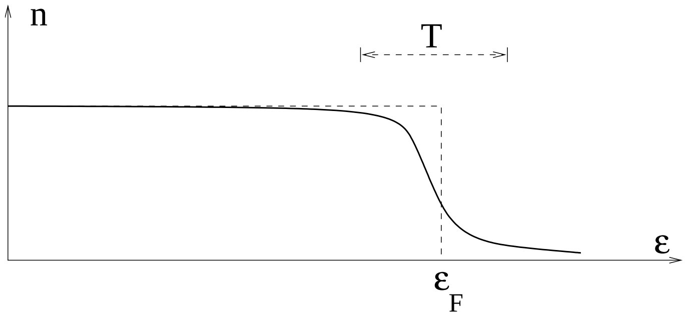
$$PV = NT\left[1\pm\frac{\pi^{3/2}}{2g}\frac{N}{V}\frac{h^3}{(mT)^{3/2}}\right]\ . \tag{59}$$

Non-surprisingly, the small factor here is the ratio of the thermal wavelength to the distance between particles. We see that quantum effects give some effective attraction between bosons and repulsion between fermions.

Landau & Lifshitz, Sects. 53, 54, 56.

2.2.1 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. Since the energy of an electron in a metal is comparable to Rydberg and so is the chemical potential (see below) then 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 $s = 1/2$):

$$\frac{N}{V} = 2\frac{4\pi}{h^3}\int_0^{p_F} p^2 dp = \frac{p_F^3}{3\pi^2\hbar^3} , \tag{60}$$

which gives the Fermi energy

$$\epsilon_F = (3\pi^2)^{2/3}\frac{\hbar^2}{2m}\left(\frac{N}{V}\right)^{2/3} . \tag{61}$$

The chemical potential at $T = 0$ coincides with the Fermi energy (putting already one electron per unit cell one obtains $\epsilon_F/k \simeq 10^4 K$). Condition $T \ll \epsilon_F$ is evidently opposite to (40). Note that the condition of ideality requires that the electrostatic energy Ze^2/a is much less than ϵ_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, $N/V \gg (e^2 m/\hbar^2)^3 Z^2$, surprisingly improves with increasing concentration. 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 $E(T, V, N)$ i.e. exclude μ from two relations, (55) and (56):

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

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

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

$$\int_0^\infty \frac{f(\epsilon)\,d\epsilon}{\exp[\beta(\epsilon-\mu)]+1} \approx \int_0^\mu f(\epsilon)\,d\epsilon + \frac{\pi^2}{6}T^2 f'(\mu) , \qquad (62)$$

which gives

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

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

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

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

and substitute it into the second equation:

$$E = \frac{3}{5}N\epsilon_F\left(1+\frac{5\pi^2}{12}\frac{T^2}{\epsilon_F^2}\right) , \qquad (63)$$

$$C_V = \frac{\pi^2}{2}N\frac{T}{\epsilon_F} . \qquad (64)$$

We see that $C_V \ll N$.

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

2.2.2 Photons

Consider electromagnetic radiation in an empty cavity kept at the temperature T. Since electromagnetic waves are linear (i.e. they do not interact) thermalization of radiation comes from interaction with walls (absorption and re-emission)[3]. 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

$$PV = E/3 \ . \tag{65}$$

In a quantum consideration we treat electromagnetic waves as photons which are massless particles with the spin 1 which 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 (58)] we can derive (65) considering[4] 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 minimumality 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, momentum $\hbar\mathbf{k}$ and energy $\epsilon = \hbar\omega = \hbar ck$ is called Planck distribution

$$\bar n_{\mathbf{k}} = \frac{1}{e^{\hbar\omega/T}-1} \ . \tag{66}$$

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

---

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

[4] This consideration is not restricted to bosons. Indeed, ultra-relativistic fermions have $\epsilon = cp$ and $P = E/3V$. Note that in the field 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.

over wave vectors continuous. Multiplying by 2 (the number of polarizations) we get the spectral distribution of energy

$$dE_\omega = \hbar ck \frac{2V}{(2\pi)^3} \frac{4\pi k^2 dk}{e^{\hbar ck/T} - 1} = \frac{V\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{e^{\hbar\omega/T} - 1} \ . \tag{67}$$

It has a maximum at $\hbar\omega_m = 2.8T$. The total energy

$$E = \frac{4\sigma}{c} VT^4 \ , \tag{68}$$

where the Stephan-Boltzmann constant is as follows: $\sigma = \pi^2/60\hbar^3 c^2$. The specific heat $c_v \propto T^3$. Since $P = 4\sigma T^4/3c$ depends only on temperature, c_P does not exist (may be considered infinite). One can also derive the free energy (which coincides with Ω for $\mu = 0$), $F = -E/3 \propto VT^4$ 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.

If one makes a small orifice in the cavity then it absorbs all the incident light like a black body. Therefore, what comes out of such a hole is called black-body radiation. The energy flux from a unit surface is the energy density times c and times the geometric factor

$$I = \frac{cE}{V} \int_0^{\pi/2} \cos\theta \sin\theta \, d\theta = \frac{c}{4} \frac{E}{V} = \sigma T^4 \ . \tag{69}$$

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

2.2.3 Phonons

The specific heat of a crystal lattice can be calculated considering 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 ω_{max} so that the total number of degrees of freedom is equal to 3 times the number of atoms:

$$\frac{4\pi V}{(2\pi)^3} \sum_{i=1}^{3} \int\limits_0^{\omega_{max}} \frac{\omega^2 d\omega}{u_i^3} = \frac{V\omega_{max}^3}{2\pi^2 u^3} = 3N \ . \tag{70}$$

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

$$\Theta = \hbar\omega_{max} = \hbar u(6\pi^2 N/V)^{1/3} \simeq \hbar u/a \ , \tag{71}$$

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, $\mu = 0$)

$$E = \frac{3V}{2\pi^2 u^3} \int\limits_0^{\omega_{max}} \hbar\omega \left(\frac{1}{2} + \frac{1}{\exp(\hbar\omega/T) - 1}\right) \omega^2 d\omega = \frac{9N\Theta}{8} + 3NTD\left(\frac{\Theta}{T}\right), \tag{72}$$

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

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

$$C = N\frac{12\pi^4}{5}\frac{T^3}{\Theta^3}\ . \tag{73}$$

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

$$F = 9NT\left(\frac{T}{\Theta}\right)^3 \int\limits_0^{T/\Theta} z^2 \ln\left(1 - e^{-z}\right) dz = NT\left[3\ln\left(1 - e^{-\Theta/T}\right) - D(\Theta/T)\right], \tag{74}$$

and find that, again, al 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, μ 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 (72) 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 (67,68). 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.

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

2.2.4 Bose gas of particles and Bose-Einstein condensation

We consider now an ideal Bose gas of massive particles with the fixed number of particles. This is applied to atoms at very low temperatures. As usual, equaling the total number of particles to the sum of Bose distribution over all states gives the equation that determines the chemical potential as a function of temperature and the specific volume. It is more convenient here to work with the function $z = \exp(\mu/T)$ which is called fugacity:

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

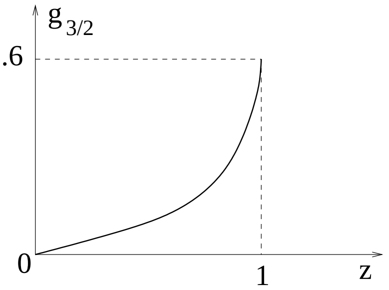
We introduced the thermal wavelength $\lambda = (2\pi\hbar^2/mT)^{1/2}$ and the function

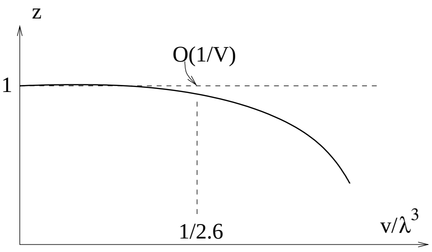
$$g_a(z) = \frac{1}{\Gamma(a)}\int_0^\infty \frac{x^{a-1}dx}{z^{-1}e^x - 1} = \sum_{i=1}^\infty \frac{z^i}{i^a}\,. \tag{75}$$

One may wonder why we split the contribution of zero-energy level as it is not supposed to contribute at the thermodynamic limit $V \to \infty$. Yet this is not true at sufficiently low temperatures. Indeed, let us rewrite it denoting $n_0 = z/(1-z)$ the number of particles at $p = 0$

$$\frac{n_0}{V} = \frac{1}{v} - \frac{g_{3/2}(z)}{\lambda^3} \ . \tag{76}$$

The function $g_{3/2}(z)$ behaves as shown at the figure, it monotonically grows while z changes from zero ($\mu = -\infty$) to unity ($\mu = 0$). Remind that the chemical potential of bosons is non-positive (otherwise one would have infinite occupation numbers). At $z = 1$, the value is $g_{3/2}(1) = \zeta(3/2) \approx 2.6$ and the derivative is infinite. When the temperature and the specific volume $v = V/N$ are such that $\lambda^3/v > g_{3/2}(1)$ (notice that 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. The solution of (76) looks as shown in the figure. 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 $\lambda^3/v = g_{3/2}(z)$.



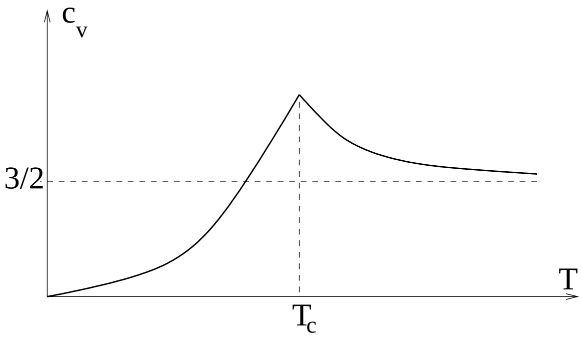


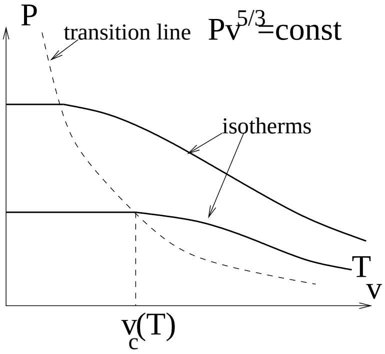
Therefore, at the thermodynamic limit we put $n_0 = 0$ at $T > T_c$ and $n_0/N = 1 - (T/T_c)^{3/2}$ as it follows from (76). All thermodynamic relations have now different expressions above and below T_c (upper and lower cases respectively):

$$E = \frac{3}{2}PV = \frac{2\pi V}{mh^3}\int_0^\infty \frac{p^4 dp}{z^{-1}\exp(p^2/2mT) - 1} = \begin{cases} (3VT/2\lambda^3)g_{5/2}(z) \\ (3VT/2\lambda^3)g_{5/2}(1) \end{cases} \tag{77}$$

$$c_v = \begin{cases} (15v/4\lambda^3)g_{5/2}(z) - 9g_{3/2}(z)/4g_{1/2}(z) \\ (15v/4\lambda^3)g_{5/2}(1) \end{cases} \tag{78}$$

At low T, $c_v \propto \lambda^{-3} \propto T^{3/2}$, it decreases faster than $c_v \propto T$ for electrons (since the number of over-condensate particles now changes with T as for phonons and photons and $\mu = 0$ too) yet slower than $c_v \propto T^3$ (that we had for $\epsilon_p = cp$) because the particle levels, $\epsilon_p = p^2/2m$, are denser at lower energies. On the other hand, since the distance between levels increases with energy so that at high temperatures c_v decreases with T as for rotators in Sect. 2.1.1:





At $T < T_c$ the pressure is independent of the volume which promts 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).

We can also obtain the entropy [above T_c by usual formulas that follow from (56) and below T_c just integrating specific heat $S = \int dE/T = N \int c_v(T)dT/T = 5E/3T = 2Nc_v/3$):

$$\frac{S}{N} = \begin{cases} (5v/2\lambda^3)g_{5/2}(z) - \log(z) \\ (5v/2\lambda^3)g_{5/2}(1) \end{cases} \tag{79}$$

The entropy is zero at $T = 0$ which means that the condensed phase has no entropy. 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.

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

2.3 Chemical reactions

Time to learn why μ is called *chemical* potential. Reactions in the mixture of ideal gases. Law of mass action. Heat of reaction. Ionization equilibrium.

Landau & Lifshitz, Sects. 101–104.

3 Non-ideal gases

Here we take into account a weak interaction between particles. There are two limiting cases when the consideration is simplified:
i) 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,
ii) when the interaction is long-range so that every particle effectively interact with many other particles and one can apply some mean-field description.

We start from ii) (even though it is conceptually more complicated) so that after consideration of i) we can naturally turn to phase transitions.

3.1 Coulomb interaction and screening

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 in screening. Indeed, if the system is neutral and the ions and electrons are distributed uniformly then the total Coulomb energy of interaction is zero. Of course, interaction leads to correlations in particle positions (particle prefer to be surrounded by the particles of the opposite charge) which makes for a nonzero contribution to the energy and other thermodynamic quantities. 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:

$$r_D^3 n_0 [\exp(e^2/r_D T) - \exp(-e^2/r_D T)] \simeq 1 \; . \tag{80}$$

We obtain what is called the Debye radius

$$r_D \sim \sqrt{\frac{T}{n_0 e^2}} \tag{81}$$

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 $10^{-3} \div 10^{-4}\,$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):

$$\bar{U} \simeq -N\frac{e^2}{r_D} \simeq -\frac{N^{3/2}e^3}{\sqrt{VT}} = -\frac{A}{\sqrt{VT}} \ . \tag{82}$$

The (positive) addition to the specific heat

$$\Delta C_V = \frac{A}{2V^{1/2}T^{3/2}} \simeq N\frac{e^2}{r_D T} \ll N \ . \tag{83}$$

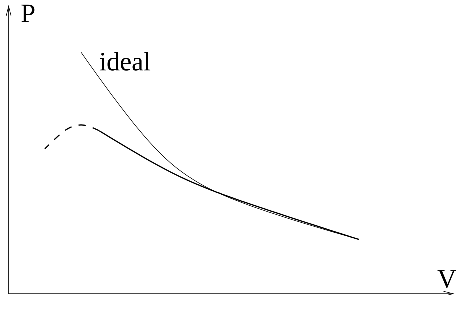
One can get the correction to the entropy by integrating the specific heat:

$$\Delta S = -\int_T^\infty \frac{C_V(T)dT}{T} = -\frac{A}{3V^{1/2}T^{3/2}} \ . \tag{84}$$

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

$$\Delta F = \bar{U} - T\Delta S = -\frac{2A}{3V^{1/2}T^{1/2}} \ , \quad \Delta P = -\frac{A}{3V^{3/2}T^{1/2}} \ . \tag{85}$$

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 that 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. 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

$$\Delta\phi = -4\pi e(n_+ - n_-) = -4\pi e n_0 \left(e^{-e\phi/T} - e^{e\phi/T}\right) \approx \frac{8\pi e^2 n_0}{T}\phi \ , \qquad (86)$$

where we expanded the exponents assuming the weakness of interaction. This equation 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 by multiplying by the number of particles ($N = 2n_0V$) and dividing by 2 so as not to count every couple of interacting charges twice:

$$\bar{U} = -n_0 V \kappa e^2 = -\sqrt{\pi}\,\frac{N^{3/2}e^3}{\sqrt{VT}} \ . \qquad (87)$$

Comparing with the rough estimate (82), 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 $e^2 n_0^{1/3}/T$. One cannot though get next corrections within the method [further expanding the exponents in (86)]. That would miss multi-point correlations which contribute the next orders. To do this 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 $\mathbf{r}_b$:

$$E = \frac{1}{2}\sum_{a,b}\frac{N_a N_b}{V^2}\int\int u_{ab} w_{ab}\, dV_a dV_b \ . \qquad (88)$$

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:

$$w_{ab} = V^{2-N} \int \exp\left[\frac{F - F_{id} - U(r_1 \ldots r_N)}{T}\right] dV_1 \ldots dV_{N-2} \ . \tag{89}$$

Expanding this equation in U/T we get terms like $u_{ab}w_{ab}$ and in addition $u_{bc}w_{abc}$ which involves the triple correlation function that one can express via the integral similar to (89). Debye-Hückel approximation corresponds to putting $w_{abc} = w_{ab}w_{bc}w_{ac}$ and assuming $\omega_{ab} = w_{ab} - 1 \ll 1$. To get r-dependence we take Laplacian of (89) which gives the equation

$$\Delta\omega_{ab}(r) = \frac{4\pi z_a z_b e^2}{T}\delta(\mathbf{r}) + \frac{4\pi z_b e^2}{TV}\sum_c N_c z_c \omega_{ac}(r) \ . \tag{90}$$

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 (86) 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 88) one gets contribution of 1 vanishing because of electro-neutrality and the term linear in ω giving (87). To get to the next order, one must expand the expression for w_{abc}, express it via w_{abcd} which can be reduced similarly to w_{abc}.

The quantum 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 objects of study are more important than methods). In this method we consider the effect of electrostatic interaction on a degenerate electron gas. According to the Fermi distribution (60) the maximal kinetic energy is related to the local concentration n by $p_0^2/2m = (3\pi^2 n)^{2/3}\hbar^2/2m$. Denote $-e\phi_0$ the maximal value of the total energy of the electron (it 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 $\Delta\phi = 4\pi en \propto (\phi - \phi_0)^{3/2}$.

See Landau & Lifshitz, Sects. 78,79 for more details.

3.2 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

$$\begin{aligned} Z(N,V,T) &= \frac{1}{N!\lambda_T^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N \exp[-U(\mathbf{r}_1,\dots,\mathbf{r}_N)] \equiv \frac{Z_N(V,T)}{N!\lambda_T^{3N}} \ . \\ \mathcal{Z}(z,V,T) &= \sum_{N=0}^{\infty} \frac{z^N Z_N}{N!\lambda_T^{3N}} \ . \end{aligned} \tag{91}$$

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

$$\begin{aligned} \mathcal{Z}(\mu,V,T) = 1 + V\frac{z}{\lambda_T^3} + \frac{V}{2!}\left(\frac{z}{\lambda_T^3}\right)^2 \int d\mathbf{r} \exp[-u(r)/T] \\ + \frac{V}{3!}\left(\frac{z}{\lambda_T^3}\right)^3 \int \exp\{-[u(r_{12}) + u(r_{13}) + u(r_{23})]/T\}\, d\mathbf{r}_2 d\mathbf{r}_3 + \dots \ . \end{aligned} \tag{92}$$

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/λ_T^3:

$$P = \lambda_T^{-3} \sum_{l=1}^{\infty} b_l z^l \ . \tag{93}$$

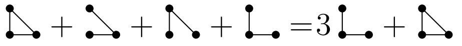
It is convenient to introduce the two-particle function, called interaction factor, $f_{ij} = \exp[-u(r_{ij})/T] - 1$, which is zero outside the range of interaction. Terms containing integrals of k functions f_{ij} are proportional to ϵ^k. The coefficients b_l can be expressed via f_{ij}:

$$
\begin{aligned}
b_1 &= 1\,, \quad b_2 = (1/2)\lambda_T^{-3}\int f_{12}\,d\mathbf{r}_{12}\,,\\
b_3 &= (1/6)\lambda_T^{-6}\int\left(e^{-U_{123}/T}-e^{-U_{12}/T}-e^{-U_{23}/T}-e^{-U_{13}/T}+2\right)d\mathbf{r}_{12}d\mathbf{r}_{13}\\
&= (1/6)\lambda_T^{-6}\int(3f_{12}f_{13}+f_{12}f_{13}f_{23})\,d\mathbf{r}_{12}d\mathbf{r}_{13}\;. \qquad (94)
\end{aligned}
$$

It is pretty cumbersome to analyze higher orders in analytical expressions. Instead, every term in

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

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 visible (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:



(95)

which corresponds to (94). 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):

$$
b_l = \frac{1}{l!V\lambda_T^{3(l-1)}}\times[\text{sum of all } l-\text{clusters}]\;. \qquad (96)
$$

In the square brackets here stand integrals like

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

Using the cluster expansion we can now show that the cluster integrals b_l indeed appear in the expansion (93). For $l = 1, 2, 3$ we saw that this is indeed so.

Denote m_l the number of l-clusters and by $\{m_l\}$ the whole set of $m_1,\ldots$. In calculating Z_N we need to include the number of ways to distribute N particles over clusters which is $N!/\prod_l (l!)^{m_l}$. We then must multiply it by the sum of all possible clusters in the power m_l divided by $m_l!$ (since an exchange of all particles from one cluster with another cluster of the same size does not matter). Since the sum of all l-clusters is $b_l l! \lambda_T^{3(l-1)} V$ then

$$Z_N = N!\lambda^{3N} \sum_{\{m_l\}} \prod_l (b_l \lambda_T^{-3} V)^{m_l}/m_l! \ .$$

Here we used $N = \sum l m_l$. The problem here is that the sum over different partitions $\{m_l\}$ needs to be taken under this restriction too and this is technically very cumbersome. Yet when we pass to calculating the grand canonical partition function[5] and sum over all possible N we obtain an unrestricted summation over all possible $\{m_l\}$. Writing $z^N = z^{\sum l m_l} = \prod_l (z^l)^{m_l}$ we get

$$\begin{aligned}
\mathcal{Z} &= \sum_{m_1,m_2,\ldots=0}^{\infty} \prod_{l=1}^{\infty} \left(\frac{V b_l z^l}{\lambda_T^3}\right)^{m_l} \frac{1}{m_l!} \\
&= \sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \cdots \left[\frac{1}{m_1!}\left(\frac{V b_1}{\lambda_T^3} z\right)^{m_1} \frac{1}{m_2!}\left(\frac{V b_2}{\lambda_T^3} z\right)^{m_2} \cdots\right] \\
&= \exp\left(\frac{V}{\lambda_T^3} \sum_{l=1}^{\infty} b_l z^l\right) \ . \qquad (97)
\end{aligned}$$

We can now reproduce (93) and write the total number of particles:

$$PV = -\Omega = T \ln \mathcal{Z}(z,V) = (V/\lambda_T^3) \sum_{l=1}^{\infty} b_l z^l \qquad (98)$$

$$\frac{1}{v} = \frac{z}{V} \frac{\partial \ln \mathcal{Z}}{\partial z} = \lambda_T^{-3} \sum_{l=1}^{\infty} l b_l z^l \ . \qquad (99)$$

To get the equation of state of now must express z via v/λ^3 from (99) and substitute into (98). That will generate the series called the virial expansion

$$\frac{Pv}{T} = \sum_{l=1}^{\infty} a_l(T) \left(\frac{\lambda_T^3}{v}\right)^{l-1} \ . \qquad (100)$$

---

[5] Sometimes the choice of the ensemble is dictated by the physical situation, sometimes by a technical convenience like now. The equation of state must be the same in the canonical and microcanonical as we expect the pressure on the wall restricting the system to be equal to the pressure measured inside.

Dimensionless virial coefficients can be 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. Further details can be found in Pathria, Sects. 9.1-2 (second edition).

3.3 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

$$B(T) = a_2\lambda_T^3 = 2\pi\int\Big\{1 - \exp[-u(r)/T]\Big\}\,dr \tag{101}$$

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 ∞ (where we can assume small negative energy, $u \ll T$, and expand the exponent). That gives

$$B(T) = b - \frac{a}{T} \tag{102}$$

with b introduced above and

$$a \equiv 2\pi\int_{r_0}^{\infty} u(r)r^2dr\ .$$

Of course, for any particular $u(r)$ it is pretty straightforward to calculate $a_2(T)$ but (102) gives a good approximation for most cases. Generally, $B(T)$ is negative at low and positive at high temperatures. We can now get the first correction to the equation of state:

$$P = \frac{NT}{V}\left[1 + \frac{NB(T)}{V}\right] = nT(1 + bn) - an^2\ . \tag{103}$$

Since $NB/V < Nb/V \ll 1$ the correction is small. Note that $a/T \ll 1$ since we assume weak interaction.

While by its very derivation the formula (103) 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 (103) $1 + bn$ 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:

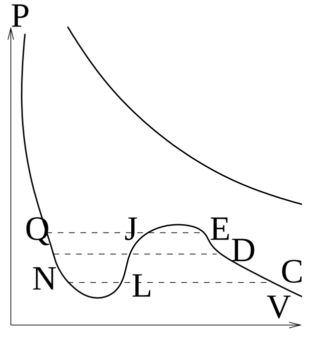
$$\Big(P + an^2\Big)(1 - nb) = nT \; . \tag{104}$$

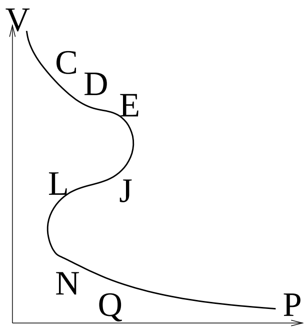
There is though an alternative way to obtain (104) without assuming the gas dilute. This is some variant of the mean field even though it is not a first step of any consistent procedure. Namely, we assume that every molecule moves in some effective field $U_e(r)$ which is a strong repulsion ($U_e \to \infty$) in some region of volume bN and is of order aN outside:

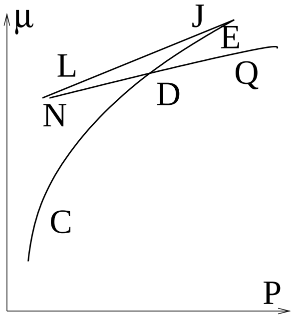
$$F - F_{id} \approx -TN \ln\left\{\int e^{-U_e(r)/T}\, d\mathbf{r}/V\right\} = -TN\left[\ln(1 - bn) - \frac{aN}{VT}\right] \; . \tag{105}$$

Differentiating (105) with respect to V gives (104). That “derivation” also helps understand better the role of the parameters b (excluded volume) and a (mean interaction energy per molecule). From (105) 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 $E = E_{id} - N^2a/V$. Since the correction to the energy is T-independent then C_V is the same as for the ideal gas.

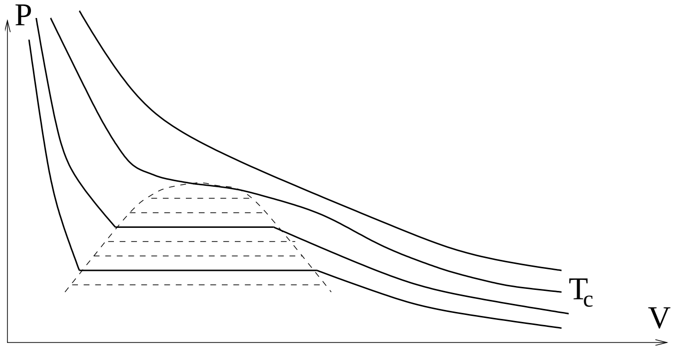
Let us now look closer at the equation of state (104). The set of isotherms is shown on the figure:







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 correspond 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 $d\mu(T,P) = -sdT + vdP$ under the 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 enable 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: $(\partial P/\partial V)_T = (\partial^2 P/\partial V^2)_T = 0$. According to (38) the fluctuations must be large at the critical point (more detail in the next chapter).

4 Phase transitions

4.1 Thermodynamic approach

We start this chapter 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.

4.1.1 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? Let us show that this is impossible in a finite system. Indeed, the classical grand partition function (expressed via fugacity $z = \exp(\mu/T)$) is as follows:

$$\mathcal{Z}(z, V, T) = \sum_{N=0}^{\infty} z^N Z(N, V, T) \ . \tag{106}$$

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

$$Z(N, V, T) = \frac{1}{N!\lambda^{3N}} \int \exp[-U(\mathbf{r}_1, \dots, \mathbf{r}_N)/T] d\mathbf{r}_1, \dots, \mathbf{r}_N \tag{107}$$

and the thermal wavelength is given by $\lambda^2 = 2\pi\hbar^2/mT$. Now, 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 (106) actually goes until N_m so that the grand partition function is a polynomial in fugacity with all coefficients positive[6]. The equation of state can be obtained by eliminating z from the parametric equations that give $P(v)$ in a parametric form — see (98,99):

$$\frac{P}{T} = \frac{1}{V} \ln \mathcal{Z}(z, V) \, , \quad \frac{1}{v} = \frac{z}{V} \frac{\partial \ln \mathcal{Z}(z, V)}{\partial z} \ .$$

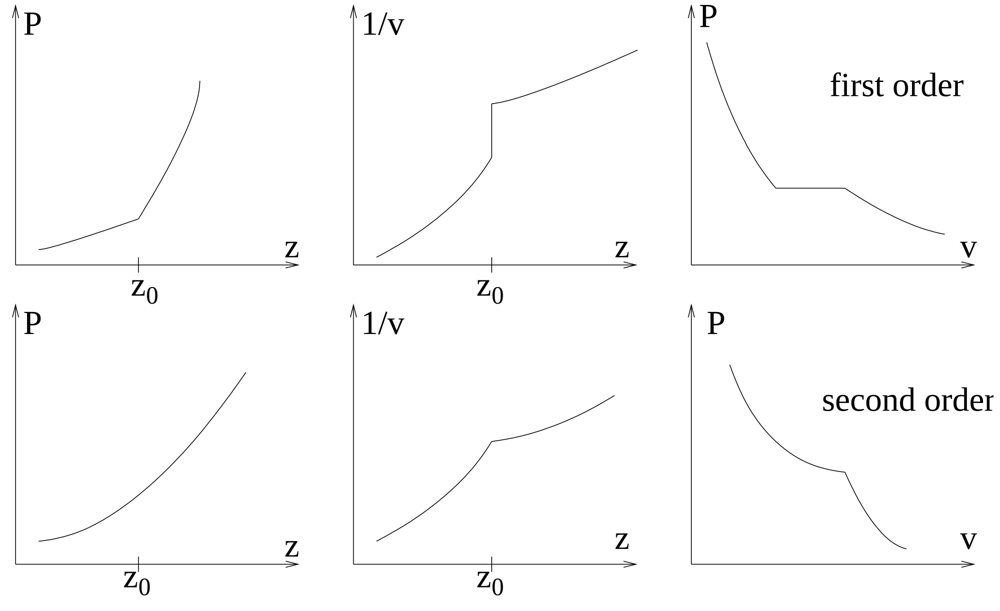
For $\mathcal{Z}(z)$ being a polynomial, both P and v are analytic functions of z in a region of the complex plane that includes the real positive axis. Therefore,

---

[6] 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.

$P(v)$ 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 $\partial v^{-1}/\partial z > 0$ so that $\partial P/\partial v = (\partial P/\partial z)/(\partial v/\partial z) < 0$.

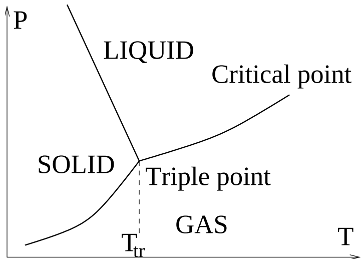
For a phase transition of the first order the pressure must be independent of v in the transition region. We see that strictly speaking in a finite volume we cannot have that since $P(v)$ is analytic, nor we can have $\partial P/\partial v > 0$. That means that singularities, jumps etc can appear only 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 at the limit $N \to \infty$ such a zero z_0 tend to a real axis of z (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 $z > z_0$. Depending on the order of zero of $\mathcal{Z}(z)$, $1/v$ itself may have a jump or its n-th derivative may have a jump, which corresponds to the $n+1$ order of phase transition. For $n = 0$ $1/v$ is discontinuous in the transition of the first order. For the second-order phase transition, volume is continuous but its derivative jumps. Huang, Sect. 15.1-2.

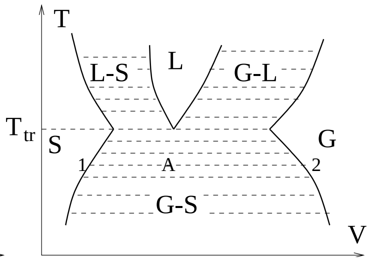


4.1.2 First-order phase transitions

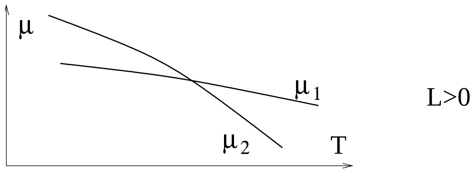
Let us now consider equilibrium between phases from a general viewpoint. We must have $T_1 = T_2$ and $P_1 = P_2$. Requiring $dG/dN_1 = \partial G_1/\partial N_1 + (\partial G_2/\partial N_2)(dN_2/dN_1) = \mu_1(P,T) - \mu_2(P,T) = 0$ we obtain the curve of the phase equilibrium $P(T)$. We thus see on 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 point 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 at fixed temperature and pressure then the heat equals to the enthalpy change or simply to $L = T\Delta s$ (per mole). If 2 is preferable to 1 at higher T (see the figure below) then $L > 0$ (heat absorbed) as it must be according to the Le Chatelier principle:



On the other hand, differentiating $\mu_1(P,T) = \mu_2(P,T)$ and using $s =$

$-(\partial\mu/\partial T)_P$, $v = (\partial\mu/\partial P)_T$, one gets the Clausius-Clapeyron equation

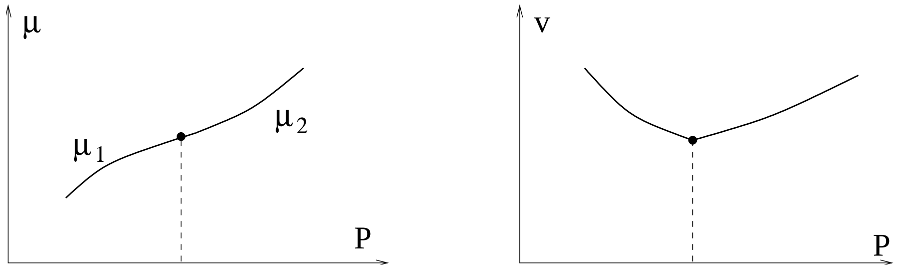
$$\frac{dP}{dT} = \frac{s_1 - s_2}{v_1 - v_2} = \frac{L}{T(v_2 - v_1)} . \tag{108}$$

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 $dP/dT > 0$. Water, on the contrary, contracts upon melting so the slope of the melting curve is negative as on the P-T diagram above. 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 (108) 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.

4.1.3 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.



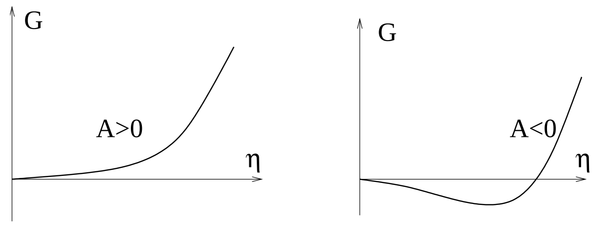
Another examples of continuous phase transitions (i.e. such that correspond to a continuous change in the system) are all related to the change in symmetry upon the change of P or T. Since symmetry is a qualitative characteristics, it can change even upon an infinitesimal change (for example, however 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 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.

4.1.4 Landau theory

To describe general properties of the second-order phase transitions Landau suggested to characterize symmetry breaking by some order parameter η 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 η 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 η:

$$G(P,T,\eta) = G_0 + A(P,T)\eta^2 + B(P,T)\eta^4 \ . \tag{109}$$

The linear term is absent to keep the first derivative continuous. 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 η^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 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.

If the transition happens at some T_c then generally near transition[7] $A(P,T)=a(P)(T-T_c)$. Writing then the potential

$$G(P,T,\eta)=G_0+a(P)(T-T_c)\eta^2+B(P,T)\eta^4\ , \tag{110}$$

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

$$\bar{\eta}^2=\frac{a}{2B}(T_c-T)\ . \tag{111}$$

In the lowest order in η 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: $\Delta C_p=a^2T_c/2B$. Specific heat increases when symmetry is broken since more types of excitations are possible.

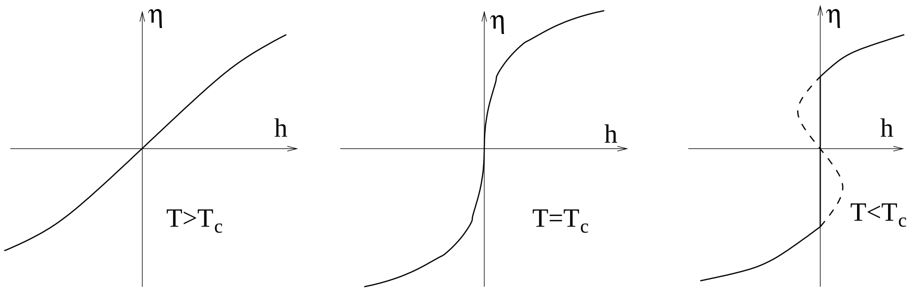
Consider now what happens when there is an external field (like magnetic field) which contributes the energy (and thus the thermodynamic potential) by the term $-h\eta V$. Equilibrium condition,

$$2a(T-T_c)\eta+4B\eta^3=hV\ , \tag{112}$$

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

---

[7] 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.



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) η.

Susceptibility,

$$\chi = \left(\frac{\partial \eta}{\partial h}\right)_{h=0} = \frac{V}{2a(T - T_c) + 12B\eta^2} = \begin{cases} [2\alpha(T - T_c)]^{-1} & \text{at } T > T_c \\ [4\alpha(T_c - T)]^{-1} & \text{at } T < T_c \end{cases} \tag{113}$$

which diverges at $T \to T_c$. That can be compared to $\chi \propto 1/T$ obtained in (21) for the noninteracting spins. Experiments support the Curie law (113). Since $a \propto V$ we have introduced $\alpha = a/V$ in (113).

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 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 $\eta(\mathbf{r})$. This is better done in terms of the space density of G which can be presented similar to (110) as $g|\nabla\eta|^2 + \alpha(T - T_c)\eta^2 + b\eta^4$. That means that having an inhomogeneous state costs us extra value of the thermodynamic potential. We again assumed that only the coefficient at η^2 turns into zero at the transition. We estimate $|\nabla\eta|^2 \simeq (\eta/r)^2$ where r is the size of the fluctuation. When r is large, this term in G does not matter, it starts playing a role at some $r_c \sim \sqrt{g/\alpha(T-T_c)}$ which is called the correlation radius. We thus see that the correlation radius diverges at the transition[8]: $r_c(T) = r_{c0}\sqrt{T_c/(T-T_c)}$. Here we expressed $g \simeq \alpha T_c r_{c0}^2$ via the correlation radius far from the transition (typically r_{c0} is of order of inter-particle distance). 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 $\bar\eta$:

$$\frac{T_c}{2\alpha r_c^3(T-T_c)} \ll \frac{\alpha(T-T_c)}{b} \quad\Rightarrow\quad \frac{T-T_c}{T_c} \gg \frac{b^2}{\alpha^4 r_{c0}^6} \equiv \left(\frac{r_i}{r_{c0}}\right)^6 . \tag{114}$$

We introduced $r_i^3 = b/\alpha^2$ which can be interpreted as the volume of interaction: if we divide the energy density of interaction $b\eta^4 \simeq b(\alpha/b)^2$ 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 is less than 10^{-2}). 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 ΔC_p is related to the values on the boundaries of the fluctuation region.

What if the order parameter is a vector $(\eta_1,\dots,\eta_n)$ and the theory is invariant under $O(n)$ rotations? Then the thermodynamic potential must have a form $g\sum_i|\nabla\eta_i|^2 + \alpha(T-T_c)\sum_i\eta_i^2 + b\Big(\sum_i\eta_i^2\Big)^2$. When $T < T_c$ the minimum corresponds to breaking the $O(n)$ symmetry, for example, by taking the first component $[\alpha(T_c-T)/2b]^{1/2}$ and the other components zero. Considering fluctuations we put $([\alpha(T_c-T)/2b]^{1/2}+\eta_1,\eta_2,\dots,\eta_n)$ and obtain the thermodynamic potential $g\sum_i|\nabla\eta_i|^2 + 2\alpha(T_c-T)\eta_1^2 +$ higher order terms. That form means that only the longitudinal mode η_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 energy. Goldstone theorem claims that whenever continuous symmetry is spontaneously broken (i.e. the symmetry of the state is less than the symmetry of the thermodynamic potential or

---

[8] 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.

Hamiltonian) then the mode must exist with energy going to zero with the wavenumber. This statement is true beyond the mean-field approximation or Landau theory. Goldstone modes are easily excited by thermal fluctuations and one can show that they destroy long-range order for $d \le 2$ (Mermin-Wagner theorem — see e.g. Kardar, Fields 3.3). Still, phase transitions may exist in 2d, for instance, Berezinskii-Kosterlitz-Thouless binding-unbinding transition for vortices or charges: low-temperature state is that of dipole pairs with power-law decay of correlations while high-temperature state has free charges that provide for a Debye screening and exponential decay of correlations.

Landau & Lifshitz, Sects. 142, 143, 144, 146.

4.2 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.

4.2.1 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. We thus consider lattice sites with elementary magnetic moments $\pm\mu$. We already considered (Sect. 1.4.1) this system in an external magnetic field H without any interaction between moments and got the magnetization

(20):

$$M = n\mu\frac{\exp(\mu H/T) - \exp(-\mu H/T)}{\exp(\mu H/T) + \exp(-\mu H/T)}\ . \tag{115}$$

First phenomenological treatment of the interacting system was done by Weiss who *assumed* that there appears some extra magnetic field proportional to magnetization which one adds to H and thus describes the influence that M causes upon itself:

$$M = n\mu \tanh\frac{\mu(H + \beta M)}{T})\ . \tag{116}$$

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

$$\eta = \tanh\frac{T_c\eta}{T}\ , \tag{117}$$

where we denoted $\eta = \beta M\mu/T_c$ and $T_c = \beta\mu^2 n$. At $T > T_c$ there is a single solution $\eta = 0$ while at $T < T_c$ there are two more nonzero solutions which exactly means the appearance of the spontaneous magnetization. At $T_c - T \ll T_c$ one has $\eta^2 = 3(T_c - T)$ exactly as in Landau theory (111). One can compare T_c with experiments and find surprisingly high $\beta \sim 10^3 \div 10^4$. 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).

We can now at last write the Ising model (formulated by Lenz in 1920 and solved in one dimension by his student Ising in 1925): we have the variable $\sigma_i = \pm 1$ at every lattice site. The energy includes interaction with the external field and between neighboring spins:

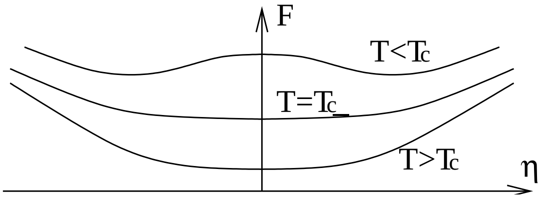
$$\mathcal{H} = -\mu H\sum_i^N \sigma_i + J/4\sum_{ij}(1 - \sigma_i\sigma_j)\ . \tag{118}$$

We assume that every spin has γ 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. 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:

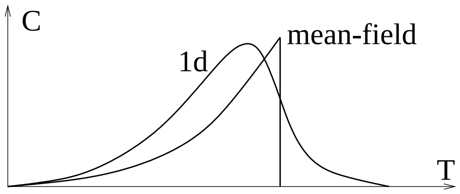
$$\gamma J \frac{N - 2N_+}{N} - T \ln \frac{N - N_+}{N_+} = 0 \ . \tag{119}$$

Here we can again introduce the variables $\eta = M/\mu N$ and $T_c = \gamma J/2$ and reduce (119) to (117). We thus see that indeed Weiss approximation is equivalent to the mean field. The only addition is that now we have the expression for the free energy so that we can indeed make sure that the nonzero η 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 as we see near T_c corresponds to the mean field approximation). 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 η). 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 $E = \gamma N_+(N - N_+)J/N = (\gamma JN/4)(1 - \eta^2 T/T_c)$ and obtain the jump exactly like in Landau theory: $\Delta C = -3\gamma JN/4T_c = -3N/2$.

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. 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. The partition function is that of the two-level system (22): $Z = 2[1 + \exp(-J/T)]^{N-1}$. Here 2 because there are two possible orientations of the first spin. There are $N - 1$ links. Now, as we know, there is no phase transitions for a two-level system. In particular one can compare the specific heat in the mean field with the exact 1d expression (see the figure below)



It is also instructive to compare the exact expression for the energy (25) which can be written as $E(T) = NJ/(1+e^{J/T})$ with the mean-field expression that one gets expressing N_+ from (119) and substituting into $E \approx \gamma N_+(N - N_+)J/N$.

We can improve the mean-field approximation by accounting exactly for the interaction of a given spin σ_0 with its γ 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):

$$\mathcal{H}_{\gamma+1} = -\mu H' \sum_{j=1}^{\gamma} \sigma_j - (J/2) \sum_{j=1}^{\gamma} \sigma_0 \sigma_j \ . \tag{120}$$

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:

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

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

Here $Z_\pm$ correspond to $\sigma_0 = \pm 1$. Requiring $\bar{\sigma}_0 = (Z_+ - Z_-)/Z$ to be equal

$$\bar{\sigma}_j = \frac{1}{\gamma}\Big\langle \sum_{j=1}^{\gamma} \sigma_j \Big\rangle = \frac{1}{\gamma Z}\frac{\partial Z}{\partial \eta} ,$$

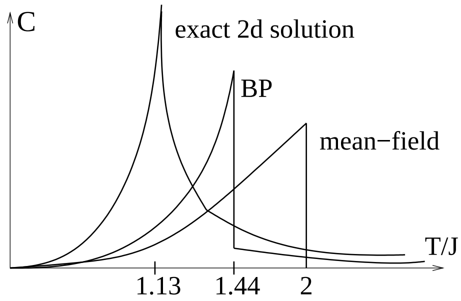
we obtain

$$\eta = \frac{\gamma - 1}{2} \ln \left[\frac{\cosh(\eta + \nu)}{\cosh(\eta - \nu)} \right] \qquad (121)$$

instead of (117) or (119). Condition $(\gamma - 1)\tanh \nu = 1$ gives the critical temperature:

$$T_c = J \ln^{-1}\Big(\frac{\gamma}{\gamma - 2}\Big), \quad \gamma \geq 2 . \qquad (122)$$

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 η is now not a magnetization, which is given by the mean spin $\bar{\sigma}_0 = \sinh(2\eta)/[\cosh(2\eta) + \exp(-2\nu)]$. BP also gives nonzero specific heat at $T > T_c$: $C = \gamma\nu^2/2\cosh^2\nu$ (see Pathria 11.6 for more details):



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 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 wider the scope of fluctuations is (the correlation radius of fluctuations r_c grows). 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).

4.2.2 Impossibility of phase coexistence in one dimension

It is physically natural that fluctuations has much influence in one dimension: it is enough for one spin to flip to loose the information of the preferred orientation. It is thus not surprising that phase transitions are impossible in one-dimensional systems with short-range interaction. Another way to understand that the ferromagnetism is possible only starting from two dimensions is to consider the spin lattice at low temperatures. The state of lowest energy has all spins parallel. The first excited state correspond to one spin flip and has an energy higher by $\Delta E = \gamma J$, the concentration of such opposite spins is proportional to $\exp(-\gamma J/T)$ and must be low at low temperatures so that the magnetization is close to μN and $\eta \approx 1$. 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 $N\exp(-J/T)$ and in sufficiently long chain this number is larger than unity i.e. the mean magnetization is zero. Note that short pieces with $N < \exp(J/T)$ are magnetized.

That argument can be generalized for arbitrary systems with the short-range interaction in the following way (Landau, 1950): assume we have n contact points of two different phases. Those points add $n\epsilon - T\ln S$ to the thermodynamic potential. The entropy is 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 $\epsilon - T\ln(eL/n)$ 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. Landau & Lifshitz, Sect. 163.

4.2.3 Equivalent models

The anti-ferromagnetic case has $J < 0$ and the ground state at $T = 0$ corresponds to the alternating spins i.e. to two sublattices. Without an external magnetic field, the magnetization of every sublattice is the same as for Ising model with $J > 0$ which 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 (91). 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

$$Z(N_a, T) = \sum^{a} \exp(2JN_{aa}/T) \,, \tag{123}$$

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,

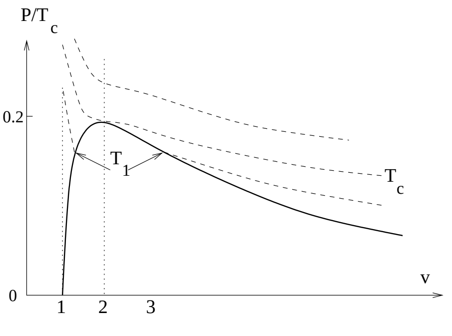
$$\mathcal{Z}(z, V, T) = \sum_{N_a}^{\infty} z^{N_a} Z(N_a, T)) \,, \tag{124}$$

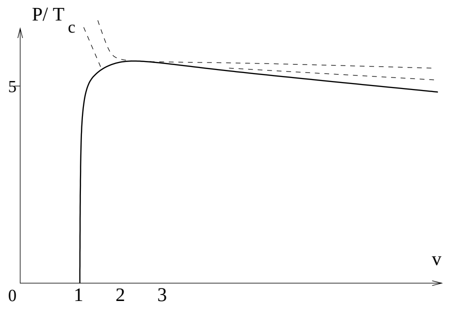
gives the equation of state in the implicit form (like in Sect. 4.1.1): $P = T \ln \mathcal{Z}/N$ and $1/v = (z/V)\partial \ln \mathcal{Z}/\partial z$. The correspondence with the Ising model can be established by saying that an occupied site has $\sigma = 1$ and unoccupied one has $\sigma = -1$. Then $N_a = N_+$ 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 (124) 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 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 η 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 (119) we get (denoting $B = \mu H$)

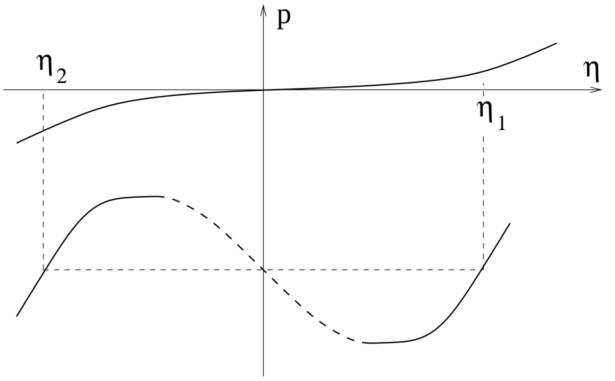
$$
\begin{aligned}
&P = B - \frac{\gamma J}{4}(1+\eta^2) - \frac{T}{2}\ln\frac{1-\eta^2}{4} \qquad B = \frac{\gamma J}{2} - \frac{T}{2}\ln z\,,\\
&v = \frac{2}{1+\eta}\,, \qquad \eta = \tanh\left(\frac{B}{T} + \frac{\gamma J\eta}{2T}\right)\,. \qquad (125)
\end{aligned}
$$

On the figure, the solid line corresponds to the solution with $B = 0$ at $T < T_c$, other isotherms are shown by broken lines. The right figure gives the exact two-dimensional solution.





The mean-field approximation (125) 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, η the isotherms at $t = 0$ (upper curve) and $t < 0$ (lower curve) look as follows:



The densities of the two phases in equilibrium, η_1, η_2 are given by the condition

$$\int\limits_1^2 v\,dp = 0 \;\Rightarrow \int\limits_1^2 \eta\,dp = \int\limits_{\eta_1 1}^{\eta_2} \eta \left(\frac{\partial p}{\partial \eta}\right)_t d\eta = \int\limits_{\eta_1}^{\eta_2} \eta\Big(2at + 12C\eta^2\Big)\,d\eta = 0\,, \quad (126)$$

where we have used $v = n^{-1} \sim n_c^{-1} - \eta n_c^{-2}$. We find from (126) $\eta_1 = -\eta_2 = (-at/2C)^{1/2}$. According to Clausius-Clapeyron equation (108) 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.

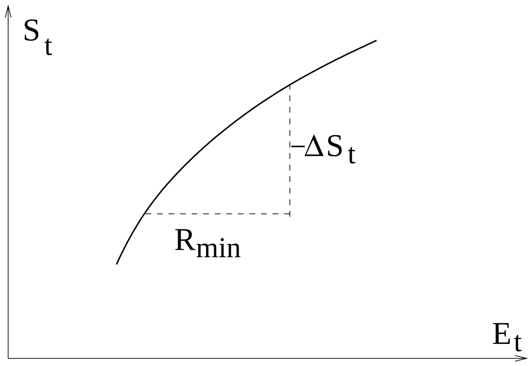
Another equivalent model is that of the binary alloy that is consisting of two types of atoms. X-ray scattering shows that below some transition temperature there are two crystal sublattices while there is only one lattice at higher temperatures. Here we need the three different energies of inter-atomic 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 5.3.

5 Fluctuations

5.1 Thermodynamic fluctuations

Consider fluctuations of energy and volume of a given (small) subsystem. The probability of a fluctuation is determined by the entropy change of the whole system $w \propto \exp(\Delta S_t)$ which is determined by the minimal work needed for a reversible creation of such a fluctuation: $T\Delta S_t = -R_{min} = T\Delta S - \Delta E - P\Delta V$ where $\Delta S, \Delta E, \Delta V$ relate to the subsystem.



For small fluctuations we can expand ΔE up to the first non-vanishing terms (quadratic):

$$R_{min} = \Delta E + P\Delta V - T\Delta S = [E_{SS}(\Delta S)^2 + 2E_{SV}\Delta S\Delta V + E_{VV}(\Delta V)^2]/2$$
$$= (1/2)(E_S\Delta E_S + \Delta V\Delta E_V) = (1/2)(\Delta S\Delta T - \Delta P\Delta V)\,. \qquad (127)$$

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 cancelled because of Maxwell identities like $(\partial P/\partial T)_V = (\partial S/\partial V)_T$. That means the absence of cross-correlation i.e. respective quantities fluctuate independently[9]: $\langle \Delta T\Delta V\rangle = \langle \Delta P\Delta S\rangle = 0$. Indeed, choosing T and V as independent variables we must express

$$\Delta S = \left(\frac{\partial S}{\partial T}\right)_V \Delta T + \left(\frac{\partial S}{\partial V}\right)_T \Delta V = \frac{C_v}{T}\Delta T + \left(\frac{\partial P}{\partial T}\right)_V \Delta V \qquad (128)$$

---

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

and obtain

$$w \propto \exp\left[-\frac{C_v}{2T^2}(\Delta T)^2 + \frac{1}{2T}\left(\frac{\partial P}{\partial V}\right)_T (\Delta V)^2\right] . \tag{129}$$

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 given volume:

$$\langle(\Delta N)^2\rangle = -T\frac{N^2}{V^2}\left(\frac{\partial V}{\partial P}\right)_T . \tag{130}$$

For a classical **ideal gas** with $V = NT/P$ it gives $\langle(\Delta N)^2\rangle = N$. In this case, we can do more than considering small fluctuations (or large volumes). Namely, we can find the probability of fluctuations comparable to the mean value $\bar{N} = N_0 V/V_0$. The probability for N (noninteracting) particles to be inside some volume V out of the total volume V_0 is

$$\begin{aligned} w_N &= \frac{N_0!}{N!(N-N_0)!}\left(\frac{V}{V_0}\right)^N \left(\frac{V-V_0}{V_0}\right)^{N_0-N} \\ &\approx \frac{\bar{N}^N}{N!}\left(1-\frac{\bar{N}}{N_0}\right)^{N_0} \approx \frac{\bar{N}^N \exp(-\bar{N})}{N!} . \end{aligned} \tag{131}$$

This is the Poisson distribution which takes place for independent events. Mean squared fluctuation is the same as for small fluctuations:

$$\begin{aligned} \langle(\Delta N)^2\rangle = \langle N^2\rangle - \bar{N}^2 = \exp(-\bar{N})\sum_{N=1}\frac{\bar{N}^N N}{(N-1)!} - \bar{N}^2 \\ = \exp(-\bar{N})\left[\sum_{N=2}\frac{\bar{N}^N}{(N-2)!} + \sum_{N=1}\frac{\bar{N}^N}{(N-1)!}\right] - \bar{N}^2 = \bar{N} . \end{aligned} \tag{132}$$

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

5.2 Spatial correlation of fluctuations

We now consider systems with interaction and discuss a spatial correlation of fluctuations of concentration $n = N/V$, which is particularly interesting near the critical point. 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

$$w \propto \exp(-\Delta F/T)\,, \quad \Delta F = \frac{1}{2}\int \phi(r_{12})\Delta n(\mathbf{r}_1)\Delta n(\mathbf{r}_2)\, dV_1 dV_2\,. \tag{133}$$

Here ϕ is the second (variational) derivative of F with respect to $n(r)$. After Fourier transform,

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

the free energy change takes the form

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

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

$$\langle |\Delta n_k|^2\rangle = \frac{T}{V\phi(k)}\,. \tag{134}$$

Usually, the largest fluctuations correspond to small k where we can use the expansion called the Ornshtein-Zernicke approximation

$$\phi(k) \approx \phi_0 + 2gk^2\,.$$

From the previous section, $\phi_0(T) = n^{-1}(\partial P/\partial n)_T$.

Making the inverse Fourier transform we find (the large-scale part of) the the pair correlation function of the concentration:

$$\langle \Delta n(0)\Delta n(r)\rangle = \sum_k |\Delta n_k|^2 e^{i\mathbf{kr}} = \int |\Delta n_k|^2 e^{i\mathbf{kr}} \frac{V d^3k}{(2\pi)^3} = \frac{T\exp(-r/r_c)}{8\pi g r}\,. \tag{135}$$

It is a general form of the correlation function at long distances. We defined the correlation radius of fluctuations $r_c = [2g(T)/\phi_0(T)]^{1/2}$. Far from any phase transition, the correlation radius is typically the mean distance between molecules.

Near the critical point, $\phi_0(T) \propto T - T_c$ and the correlation radius increases, $r_c \propto (T - T_c)^{-1/2}$, so that the correlation function approaches power law $1/r$. Of course, those scalings are valid under the condition that the criterium (114) is satisfied that is not very close to T_c . As we have seen from the exact solution of 2d Ising model, the true scaling laws 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, another at the density of gas. As we approach the critical point and the distance between peaks is getting comparable to the 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). Landau & Lifshitz, Sects. 116, 152.

5.3 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 type of symmetry which 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, according to the *scaling hypothesis*, must be scale invariant, that is to transform under a re-scaling of arguments as follows: $F(\lambda^a t, \lambda^b h) = \lambda F(t, h)$. This is a very powerful statement which tells that this is the function of one argument (rather than two), for instance,

$$F(t,h)=t^{1/a}g(h/t^{b/a})\ . \tag{136}$$

One can now express $\beta=(1-b)/a$ etc.

A general formalism which describes how to make a coarse-graining to the description to keep only most salient features 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. For Ising model, it is achieved with the help of a *block spin transformation* that is dividing 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 σ' which is ± 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[10]) 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 σ_3,σ_4 keeping $\sigma'_1=\sigma_2$ and $\sigma'_2=\sigma_5$ fixed. The respective factors in the partition function can be written as follows: $\exp[K\sigma_3\sigma_4]=\cosh K+\sigma_3\sigma_4\sinh K$. Denote $x=\tanh K$. Then only the terms with even powers of σ_3,σ_4 contribute and

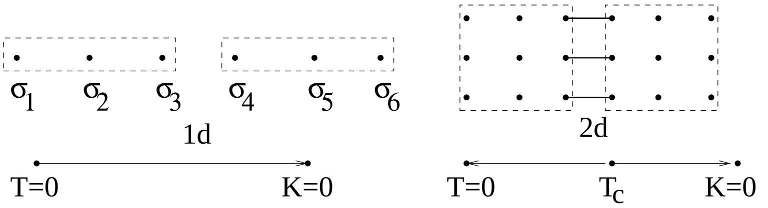
$$\cosh^3 K\sum_{\sigma_3,\sigma_4=\pm1}(1+x\sigma'_1\sigma_3)(1+x\sigma_4\sigma_3)(1+x\sigma'_2\sigma_4)=4\cosh^3 K(1+x^3\sigma'_1\sigma'_2)$$

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 $x'=x^3$. Note that $T\to\infty$ 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, the first one stable and the second one unstable. Indeed, after iterating the process 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.

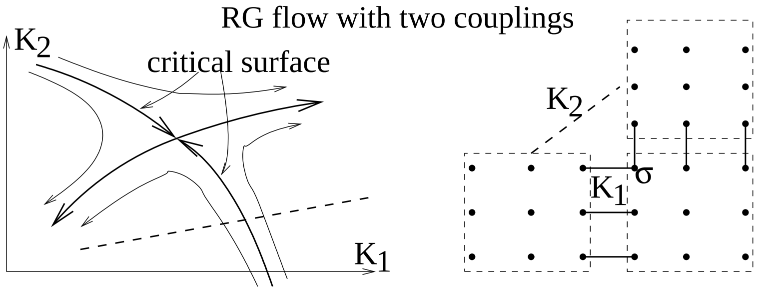
---

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

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 $x \to 1$ or $T \to 0$ as $\exp(2K) = \exp(J/T)$.



The picture of RG flow is different in higher dimensions. Indeed, in 1d in the low-temperature region $(x \approx 1, K \to \infty)$ the interaction constant K is not changed upon renormalization: $K' \approx K\langle\sigma_3\rangle_{\sigma_2=1}\langle\sigma_4\rangle_{\sigma_5=1} \approx K$. This is clear because the interaction between k-blocks is mediated by their boundary spins (that all look at the same direction). In d dimensions, there are k^{d-1} spins at the block side so that $K' \propto k^{d-1}K$ as $K \to \infty$. 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 $K = K_c$. Since $kr_c(K') = r_c(K)$ then $\nu = 1/y$. 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.


RG flow with two couplings

Already in 2d, summing over corner spin σ 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 critical behavior in a sense that systems in different regions of the parameter $\mathbf{K}$-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, by 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 ferromagnet. See Cardy, Sect 3 and http://www.weizmann.ac.il/home/fedomany/

5.4 Response and fluctuations

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. One can formulate a general relation. 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

$$\chi \equiv \frac{\partial \bar{B}}{\partial f} = \frac{\langle Bx \rangle - \bar{B}\bar{x}}{T} \equiv \frac{\langle Bx \rangle_c}{T} \ . \tag{137}$$

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.

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 (137) 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 $\epsilon(\mathbf{r})$:

$$\begin{aligned} & T\partial E/\partial T = TC_v = \langle (\Delta E)^2 \rangle / T \\ & = \frac{1}{T}\int \langle \epsilon(\mathbf{r})\epsilon(\mathbf{r}') \rangle_c \, d\mathbf{r} d\mathbf{r}' = \frac{V}{T}\int \langle \epsilon(\mathbf{r})\epsilon(0) \rangle_c \, d\mathbf{r} \end{aligned}$$

2. If $f = h$ is a magnetic field then the coordinate $x = M$ is the magnetization and (137) 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 χ near Curie point means the growth of correlations between distant spins i.e. the growth of correlation length.

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 can be written as an integral with the response (Green) function:

$$\bar{a}(\mathbf{r}) = \int G(\mathbf{r}-\mathbf{r}')f(\mathbf{r}')\,d\mathbf{r}' , \quad \bar{a}_k = G_k f_k \ . \tag{138}$$

One relates the Fourier components of the Green function and the pair correlation function of the coordinate fluctuations choosing $B = a_k$ in (137):

$$\begin{aligned} VG_k &= \frac{1}{T}\int \langle a(\mathbf{r})a(\mathbf{r}')\rangle_c e^{i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})}\,d\mathbf{r}d\mathbf{r}' = \frac{V}{T}\int \langle a(\mathbf{r})a(0)\rangle_c e^{-i\mathbf{k}\cdot\mathbf{r}}\,d\mathbf{r} \ . \\ TG_k &= (a^2)_k \ . \end{aligned} \tag{139}$$

4. If $B = x = N$ then f is the chemical potential μ:

$$\left(\frac{\partial N}{\partial \mu}\right)_{T,V} = \frac{\langle N^2\rangle_c}{T} = \frac{\langle(\Delta N)^2\rangle}{T} = \frac{V}{T}\int \langle n(\mathbf{r})n(0)\rangle_c\,d\mathbf{r} \ .$$

This formula coincides with (130) if one accounts for

$$\begin{aligned} -n^2\left(\frac{\partial V}{\partial P}\right)_{T,N} &= N\left(\frac{\partial n}{\partial P}\right)_{T,N} = n\left(\frac{\partial N}{\partial P}\right)_{T,V} \\ &= \left(\frac{\partial P}{\partial \mu}\right)_{T,V}\left(\frac{\partial N}{\partial P}\right)_{T,V} = \left(\frac{\partial N}{\partial \mu}\right)_{T,V} \ . \end{aligned}$$

Hence the response of the density to the pressure is related to the density fluctuations.

Shang-Keng Ma, Statistical Mechanics, Sect. 13.1

5.5 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[11] because one must find the non-equilibrium time-dependent probability density in the phase space solving the Liouville equation

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial x}\frac{\partial \mathcal{H}}{\partial p} - \frac{\partial \rho}{\partial p}\frac{\partial \mathcal{H}}{\partial x} \equiv \{\rho, \mathcal{H}\} \ , \tag{140}$$

or the respective equation for the density matrix in the quantum case. Here p is the canonical momentum conjugated to the coordinate x. One can solve the equation (140) perturbatively in f starting from $\rho_0 = Z^{-1}\exp(-\beta\mathcal{H}_0)$ and then solving

$$\frac{\partial \rho_1}{\partial t} + \mathcal{L}\rho_1 = -f\beta\frac{\partial \mathcal{H}_0}{\partial p}\rho_0 \ . \tag{141}$$

Here we denoted the linear operator $\mathcal{L}\rho_1 = \{\rho_1, \mathcal{H}_0\}$. Recall now that $\partial\mathcal{H}_0/\partial p = \dot{x}$ (calculated at $f = 0$). If t_0 is the time when we switched on the force $f(t)$ then the formal solution of (141) is written as follows

$$\rho_1 = \beta\rho_0 \int_{t_0}^{t} e^{(\tau - t)\mathcal{L}}\dot{x}(\tau)f(\tau)\, d\tau = \beta\rho_0 \int_{t_0}^{t} \dot{x}(\tau - t)f(\tau)\, d\tau \ . \tag{142}$$

We used the fact that $\exp(t\mathcal{L})$ is a time-displacement (or evolution) operator that moves any function of phase variables forward in time by t as it follows from the fact that for any function on the phase space $dA(p,x)/dt = \mathcal{L}A$. In (142), the function of phase space variables is $\dot{x}[p(\tau), x(\tau)] = \dot{x}(\tau)$.

We now use (142) 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

$$\langle x(t)\rangle \equiv \int_{-\infty}^{t} \alpha(t,t')f(t')\, dt' = \int x dx \rho_1(x,t)\, , \tag{143}$$

which defines generalized susceptibility (also called response or Green function) $\alpha(t,t') = \alpha(t-t') \equiv \delta\langle x(t)\rangle/\delta f(t')$. From (142,143) we can now obtain the fluctuation-dissipation theorem

$$\frac{\partial}{\partial t'}\langle x(t)x(t')\rangle = T\alpha(t,t') \ . \tag{144}$$

---

[11] As the poet (Brodsky) said, "Time is bigger than space: space is an entity, time is in essence a thought of an entity."

It relates quantity in equilibrium (the decay rate of correlations) to the weakly non-equilibrium quantity (response to a small perturbation). To understand it better, introduce the spectral decomposition of the fluctuations:

$$x_\omega = \int_{-\infty}^{\infty} x(t)e^{\imath\omega t}dt\,, \quad x(t) = \int_{-\infty}^{\infty} x_\omega e^{-\imath\omega t}\frac{d\omega}{2\pi}\ . \tag{145}$$

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

$$\alpha(\omega) = \int_0^\infty \alpha(t)\,dt = \alpha + \imath\alpha''$$

is analytic in the upper half-plane of complex ω and $\alpha(-\omega^*) = \alpha^*(\omega)$. Let us show that the imaginary part α'' determines the energy dissipation,

$$\frac{dE}{dt} = \overline{\frac{d\mathcal{H}}{dt}} = \overline{\frac{\partial\mathcal{H}}{\partial t}} = -\overline{\frac{\partial\mathcal{H}}{\partial f}}\frac{df}{dt} = -\bar{x}\frac{df}{dt} \tag{146}$$

For purely monochromatic perturbation, $f(t) = f_\omega\exp(-\imath\omega t) + f_\omega^*\exp(\imath\omega t)$, $2\bar{x} = \alpha(\omega)f_\omega\exp(-\imath\omega t) + \alpha(-\omega)f_\omega^*\exp(\imath\omega t)$, the dissipation averaged over a period is as follows:

$$\overline{\frac{dE}{dt}} = \int_0^{2\pi/\omega}\frac{\omega dt}{2\pi}[\alpha(-\omega)-\alpha(\omega)]\imath\omega|f_\omega|^2 = 2\omega\alpha''_\omega|f_\omega|^2\ . \tag{147}$$

We can now calculate the average dissipation using (142)

$$\overline{\frac{dE}{dt}} = -\int x\dot{f}\rho_1\,dpdx = \beta\omega^2|f_\omega|^2(x^2)_\omega\ , \tag{148}$$

where the spectral density of the fluctuations is calculated with ρ_0 (i.e. at unperturbed equilibrium). Comparing (147) and (148) or directly from (144) we obtain the spectral form of the fluctuation-dissipation theorem (Callen and Welton, 1951):

$$2T\alpha''(\omega) = \omega(x^2)_\omega\ . \tag{149}$$

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 (149) over frequencies we get

$$\langle x^2\rangle = \int_{-\infty}^{\infty}(x^2)_\omega \frac{d\omega}{2\pi} = \frac{T}{\pi}\int_{-\infty}^{\infty}\frac{\alpha''(\omega)d\omega}{\omega} = \frac{T}{\imath\pi}\int_{-\infty}^{\infty}\frac{\alpha(\omega)d\omega}{\omega} = T\alpha(0)\ . \quad (150)$$

The spectral density has a universal 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 $\bar{x}$:

$$\dot{x} = -\lambda(x - \bar{x})\ . \quad (151)$$

For harmonics, $\alpha(\omega) = \alpha(0)\lambda(\lambda - i\omega)^{-1}$ and $\alpha''(\omega) = \alpha(0)\omega(\lambda^2 + \omega^2)^{-1}$. The spectral density of such (so-called quasistationary) fluctuations is as follows:

$$(x^2)_\omega = \langle x^2\rangle \frac{2\lambda}{\lambda^2 + \omega^2}\ . \quad (152)$$

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 (135). Non-smooth behavior at zero is an artefact of the long-time approximation, consistent consideration would give zero derivative at $t = 0$.

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)

$$\dot{x}_i = -\lambda_{ij}x_j\ . \quad (153)$$

Single-time probability distribution of small fluctuations is Gaussian $w(\mathbf{x}) \sim \exp(\Delta S) \approx \exp(-\beta_{jk}x_jx_k)$. Introduce forces $X_j = \partial S/\partial x_j = \beta_{ij}x_j$ so that $\dot{x}_i = \gamma_{ij}X_j$, $\gamma_{ij} = \lambda_{ik}(\hat{\beta}^{-1})_{kj}$ with $\langle x_iX_j\rangle = \delta_{ij}$, $\langle X_jX_j\rangle = \beta_{ij}$ and $\langle x_jx_k\rangle = (\hat{\beta}^{-1})_{jk}$. 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 crystals without magnetic field. Also, a temperature difference produces the same electric current as the heat current produced by a voltage.

See Landay & Lifshitz, Sect. 119-120 for the details and Sect. 124 for the quantum case. Also Kittel 33-34.

5.6 Brownian motion

The momentum of a particle in a fluid, $\mathbf{p} = M\mathbf{v}$, changes because of collisions with the molecules. When the particle is much heavier than the molecules then its velocity is small comparing to the typical velocities of the molecules. Then one can write the force acting on it as Taylor expansion with the parts independent of $\mathbf{p}$ and linear in $\mathbf{p}$:

$$\dot{\mathbf{p}} = -\alpha\mathbf{p} + \mathbf{f}\ . \tag{154}$$

Here, $\mathbf{f}(t)$ is a random function which makes (154) Langevin equation. Its solution

$$\mathbf{p}(t) = \int_{-\infty}^{t} \mathbf{f}(t')e^{\alpha(t'-t)}dt'\ . \tag{155}$$

We now 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 τ which is much smaller than α^{-1}. Since the integration time in (155) is of order α^{-1} then the condition $\alpha\tau \ll 1$ means that the momentum of a Brownian particle can be considered as a sum of many independent random numbers (integrals over intervals of order τ) and so it must have a Gaussian statistics $\rho(\mathbf{p}) = (2\pi\sigma^2)^{-3/2}\exp(-p^2/2\sigma^2)$ where

$$\begin{aligned}\sigma^2 &= \langle p_x^2\rangle = \langle p_y^2\rangle = \langle p_z^2\rangle = \int_0^\infty C(t_1-t_2)e^{-\alpha(t_1+t_2)}dt_1dt_2 \\ &\approx \int_0^\infty e^{-2\alpha t}\,dt\int_{-2t}^{2t} C(t')\,dt' \approx \frac{1}{2\alpha}\int_{-\infty}^{\infty} C(t')\,dt'\ .\end{aligned} \tag{156}$$

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):

$$\alpha = \frac{1}{2TM}\int_{-\infty}^{\infty} C(t')\,dt'\ . \tag{157}$$

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 $\langle\mathbf{v}(t)\cdot\mathbf{v}(0)\rangle = (3T/M)\exp(-\alpha t)$ which can be obtained from (155)[12]. That gives $\langle(\Delta\mathbf{r})^2\rangle = 6Tt'/M\alpha$ and the probability distribution of displacement, $\rho(\Delta\mathbf{r}, t') = (4\pi Dt')^{-3/2}\exp[-(\Delta\mathbf{r})^2/4Dt'$ that satisfies the diffusion equation $\partial\rho/\partial t' = D\nabla^2\rho$ with the diffusivity $D = T/M\alpha$ — the Einstein relation.

Ma, Sect. 12.7

---

[12] Note that the friction makes velocity correlated on a longer timescale than the force.

6 Kinetics

Here we consider non-equilibrium behavior of a rarefied classical gas.

6.1 Boltzmann equation

In kinetics, the probability distribution in the phase space is traditionally denoted $f(\mathbf{r}(t),\mathbf{p}(t),t)$ (reserving ρ for the mass density in space). We write the equation for the distribution in the following form

$$\frac{\partial f}{\partial t}+\frac{\partial f}{\partial \mathbf{r}}\frac{\partial \mathbf{r}}{\partial t}+\frac{\partial f}{\partial \mathbf{v}}\frac{\partial \mathbf{v}}{\partial t}=\frac{\partial f}{\partial t}+\mathbf{v}\frac{\partial f}{\partial \mathbf{r}}+\frac{\mathbf{F}}{m}\frac{\partial f}{\partial \mathbf{v}}=I\,, \tag{158}$$

where F is the force acting on the particle of mass m while I represent the interaction with other particles that are assumed to be only binary collisions. The number of collisions (per unit time per unit volume) that change velocities of two particles from $\mathbf{v}$, $\mathbf{v}_1$ to $\mathbf{v}'$, $\mathbf{v}_1'$ is written as follows

$$w(\mathbf{v},\mathbf{v}_1;\mathbf{v}',\mathbf{v}_1')ff_1\,d\mathbf{v}d\mathbf{v}_1d\mathbf{v}'d\mathbf{v}_1'\ . \tag{159}$$

Note that we assumed here that the particle velocity is independent of the position and that the two particles are statistically independent that is the probability to find two particles simultaneously is the product of single-particle probabilities. This sometimes is called *the hypothesis of molecular chaos* and has been proved only for few simple cases. We *believe* that (159) must work well when the distribution function evolves on a time scale much longer than that of a single collision. Since $w\propto|\mathbf{v}-\mathbf{v}_1|$ then one may introduce the scattering cross-section $d\sigma=wd\mathbf{v}'d\mathbf{v}_1'/|\mathbf{v}-\mathbf{v}_1|$ which in principle can be found for any given law of particle interaction by solving a kinematic problem. Here we describe the general properties. Since mechanical laws are time reversible then

$$w(-\mathbf{v},-\mathbf{v}_1;-\mathbf{v}',-\mathbf{v}_1')=w(\mathbf{v}',\mathbf{v}_1';\mathbf{v},\mathbf{v}_1)\ . \tag{160}$$

If, in addition, the medium is invariant with respect to inversion $\mathbf{r}\to-\mathbf{r}$ then we have the *detailed equilibrium*:

$$w\equiv w(\mathbf{v},\mathbf{v}_1;\mathbf{v}',\mathbf{v}_1')=w(\mathbf{v}',\mathbf{v}_1';\mathbf{v},\mathbf{v}_1)\equiv w'\ . \tag{161}$$

Another condition is the probability normalization which states the sum of transition probabilities over all possible states, either final or initial, is unity and so the sums are equal to each other:

$$\int w(\mathbf{v},\mathbf{v}_1;\mathbf{v}',\mathbf{v}_1')\,d\mathbf{v}'d\mathbf{v}_1' = \int w(\mathbf{v}',\mathbf{v}_1';\mathbf{v},\mathbf{v}_1)\,d\mathbf{v}'d\mathbf{v}_1' \ . \tag{162}$$

We can now write the collision term as the difference between the number of particles coming and leaving the given region of phase space around $\mathbf{v}$:

$$\begin{aligned} I &= \int (w'f'f_1' - wff_1)\,d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1' \\ &= \int w'(f'f_1' - ff_1)\,d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1' \,. \end{aligned} \tag{163}$$

Here we used (162) in transforming the second term. We can now write the famous *Boltzmann kinetic equation* (1872)

$$\frac{\partial f}{\partial t} + \mathbf{v}\frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m}\frac{\partial f}{\partial \mathbf{v}} = \int w'(f'f_1' - ff_1)\,d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1' \,, \tag{164}$$

6.2 H-theorem

The entropy of the ideal classical gas can be derived for an arbitrary (not necessary equilibrium) distribution in the phase. Consider an element $d\mathbf{p}d\mathbf{r}$ which has $G_i = d\mathbf{p}d\mathbf{r}/h^3$ states and $N_i = fG_i$ particles. The entropy of the element is $S_i = \ln(G_i^{N_i}/N_i!) \approx N_i \ln(eG_i/N_i) = f\ln(e/f)d\mathbf{p}d\mathbf{r}/h^3$. We write the total entropy up to the factor M/h^3: $S = \int f\ln(e/f)\,d\mathbf{r}d\mathbf{v}$. Let us look at the evolution of the entropy

$$\frac{dS}{dt} = -\int \frac{\partial f}{\partial t}\ln f\,d\mathbf{r}d\mathbf{v} = -\int I\ln f\,d\mathbf{r}d\mathbf{v} \ , \tag{165}$$

since

$$\int \ln f\left(\mathbf{v}\frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m}\frac{\partial f}{\partial \mathbf{v}}\right)d\mathbf{r}d\mathbf{v} = \int\left(\mathbf{v}\frac{\partial}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m}\frac{\partial}{\partial \mathbf{v}}\right)f\ln\frac{f}{e}\,d\mathbf{r}d\mathbf{v} = 0 \ .$$

The integral (165) contains the integrations over all velocities so we may exploit two interchanges, $\mathbf{v}_1 \leftrightarrow \mathbf{v}$ and $\mathbf{v},\mathbf{v}_1 \leftrightarrow \mathbf{v}',\mathbf{v}_1'$:

$$\begin{aligned} \frac{dS}{dt} &= \int w'\ln f(ff_1 - f'f_1')\,d\mathbf{v}d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1'\,d\mathbf{r} \\ &= \frac{1}{2}\int w'\ln ff_1(ff_1 - f'f_1')\,d\mathbf{v}d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1'\,d\mathbf{r} \\ &= \frac{1}{2}\int w'\ln(ff_1/f'f_1')ff_1\,d\mathbf{v}d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1'\,d\mathbf{r} \geq 0 \ , \end{aligned} \tag{166}$$

Here we may add the integral $\int w'(ff_1 - f'f_1')\, d\mathbf{v}d\mathbf{v}_1 d\mathbf{v}'d\mathbf{v}_1'\, d\mathbf{r}/2 = 0$ and then use the inequality $x \ln x - x + 1 \geq 0$ with $x = ff_1/f'f_1'$. Note that entropy production is positive in every element $d\mathbf{r}$.

Even though we use scattering cross-sections obtained from mechanics reversible in time, our use of molecular chaos hypothesis made the kinetic equation irreversible. The reason for irreversibility is coarse-graining that is finite resolution in space and time, as was explained in Sect. 1.5.

Equilibrium distribution realizes the entropy maximum and so must be a steady solution of the Boltzmann equation. Indeed, the equilibrium distribution depends only on the integrals of motion. For any function of the conserved quantities, the left-hand-side of (164) (which is a total time derivative) is zero. Also the collision integral turns into zero by virtue of $f_0(\mathbf{v})f_0(\mathbf{v}_1) = f_0(\mathbf{v}')f_0(\mathbf{v}_1')$ since $\ln f_0$ is the linear function of the integrals of motion as was explained in Sect. 1.1. Note that all this is true also for the inhomogeneous equilibrium in the presence of an external force.

6.3 Conservation laws

Conservation of energy and momentum in collisions unambiguously determine $\mathbf{v}', \mathbf{v}_1'$ so we can also write the collision integral via the cross-section which depends only on the relative velocity:

$$I = \int |\mathbf{v} - \mathbf{v}_1|(f'f_1' - ff_1)\, d\sigma d\mathbf{v}_1 \ .$$

We considered collisions as momentary acts that happen in a point so that we do not resolve space regions compared with molecule sizes d and time intervals comparable with the collision time d/v. The collision integral can be roughly estimated via the mean free path between collisions, $l \simeq 1/n\sigma \simeq 1/nd^2 = d/(nd^3)$. Since we assume the gas dilute, that is $nd^3 \ll 1$ then $d \ll n^{-1/3} \ll l$. The mean time between collisions can be estimated as $\tau \simeq l/\bar{v}$ and the collision integral in the so-called τ-approximation is estimated as follows: $I \simeq (f - f_0)/\tau = \bar{v}(f - f_0)/l$. If the scale of f change (imposed by external fields) is L then the left-hand side of (164) can be estimates as $\bar{v}f/L$, comparing this to the collision integral estimate in the τ-approximation one gets $\delta f/f \sim l/L$. When this ratio is small one can derive macroscopic description assuming f to be close to f_0.

One uses conservation properties of the Boltzmann equation to derive such macroscopic (hydrodynamic) equations. Define the local density $\rho(\mathbf{r}, t) =$

$m\int f(\mathbf{r},\mathbf{v},t)\,d\mathbf{v}$ and velocity $\mathbf{u}=\int \mathbf{v}f\,d\mathbf{v}/\int f\,d\mathbf{v}$. Collisions do not change total number of particles, momentum and energy so that if we multiply (164) respectively by m, mv_α, ϵ and integrate over $d\mathbf{v}$ we get three conservation laws (mass, momentum and energy):

$$\frac{\partial\rho}{\partial t}+div\,\rho\mathbf{u}=0\,, \tag{167}$$

$$\frac{\partial\rho u_\alpha}{\partial t}=nF_\alpha-\frac{\partial}{\partial x_\beta}\int mv_\alpha v_\beta f\,d\mathbf{v}\equiv nF_\alpha-\frac{\partial P_{\alpha\beta}}{\partial x_\beta}\,, \tag{168}$$

$$\frac{\partial n\bar\epsilon}{\partial t}=n(\mathbf{F}\cdot\mathbf{u})-div\int\epsilon\mathbf{v}f\,d\mathbf{v}\equiv n(\mathbf{F}\cdot\mathbf{u})-div\,\mathbf{q}\,, \tag{169}$$

While the form of those equations is suggestive, to turn them into the hydrodynamic equations ready to use practically, one needs to find f and express the tensor of momentum flux $P_{\alpha\beta}$ and the vector of the energy flux $\mathbf{q}$ via the macroscopic quantities $\rho,\mathbf{u},n\bar\epsilon$. Since we consider situations when ρ and $\mathbf{u}$ are both inhomogeneous then the system is clearly not in equilibrium. Closed macroscopic equations can be obtained when those inhomogeneities are smooth so that in every given region (much larger than the mean free path but much smaller than the scale of variations in ρ and $\mathbf{u}$) the distribution is close to equilibrium.

At the first step we assume that $f=f_0$ which (as we shall see) means neglecting dissipation and obtaining so-called *ideal hydrodynamics*. Equilibrium in the piece moving with the velocity $\mathbf{u}$ just correspond to the changes $\mathbf{v}=\mathbf{v}'+\mathbf{u}$ and $\epsilon=\epsilon'+m(\mathbf{u}\cdot\mathbf{v}')+mu^2/2$ where primed quantities relate to the co-moving frame where the distribution is isotropic and $\langle v'_\alpha v'_\beta\rangle=\langle v^2\rangle\delta_{\alpha\beta}/3$. The fluxes are thus

$$P_{\alpha\beta}=\rho\langle v_\alpha v_\beta\rangle=\rho(u_\alpha u_\beta+\langle v'_\alpha v'_\beta\rangle)=\rho u_\alpha u_\beta+P\delta_{\alpha\beta}\,, \tag{170}$$

$$\mathbf{q}=n\langle\epsilon\mathbf{v}\rangle=n\mathbf{u}\left(\frac{mu^2}{2}+\frac{m}{3}\langle v'^2\rangle+\bar\epsilon'\right)=\mathbf{u}\left(\frac{\rho u^2}{2}+\mathcal{W}\right)\,. \tag{171}$$

Here P is pressure and $\mathcal{W}=P+n\bar\epsilon'$ is the enthalpy per unit volume. Along $\mathbf{u}$ there is the flux of parallel momentum $P+\rho u^2$ while perpendicular to $\mathbf{u}$ the momentum component is zero and the flux is P. For example, if we direct the x-axis along velocity at a given point then $P_{xx}=P+v^2$, $P_{yy}=P_{zz}=P$ and all the off-diagonal components are zero. Note that the energy flux *is not* $\mathbf{u}n\bar\epsilon$ i.e. the energy is not a passively transported quantity. Indeed, to calculate the energy change in any volume we integrate $div\,\mathbf{q}$ over the volume which turns into surface integrals of two terms. One is $\mathbf{u}n\bar{\epsilon}$ which is the energy brought into the volume, another is the pressure term $P\mathbf{u}$ which gives the work done. The closed first-order equations (167-171) constitute ideal hydrodynamics. While we derived it only for a dilute gas they are used for liquids as well which can be argued heuristically.

6.4 Transport and dissipation

To describe the transport and dissipation of momentum and energy (i.e. viscosity and thermal conductivity), we now account for the first non-equilibrium correction to the distribution function which we write as follows:

$$\delta f = f - f_0 \equiv -\frac{\partial f_0}{\partial \epsilon}\chi(\mathbf{v}) = \frac{f_0}{T}\chi\ . \tag{172}$$

The linearized collision integral takes the form $(f_0/T)I(\chi)$ with

$$I(\chi) = \int w' f_0(\mathbf{v}_1)(\chi' + \chi_1' - \chi - \chi_1)\, d\mathbf{v}_1 d\mathbf{v}' d\mathbf{v}_1'\ . \tag{173}$$

This integral is turned into zero by three functions, χ =const, $\chi = \epsilon$ and $\chi = \mathbf{v}$, which correspond to the variation of the three parameters of the equilibrium distribution. Indeed, varying the number of particles we get $\delta f = \delta N \partial f_0/\partial N = \delta N f_0/N$ while varying the temperature we get $\delta f = \delta T \partial f_0/\partial T$ which contains ϵf_0. The third solution is obtained by exploiting the Galilean invariance (in the moving reference frame the equilibrium function must also satisfy the kinetic equation). In the reference frame moving with $\delta\mathbf{u}$ the change of f is $\delta\mathbf{u}\cdot\partial f_0/\partial\mathbf{v} = -(\delta\mathbf{u}\cdot\mathbf{p})f_0/T$. We define the parameters (number of particles, energy and momentum) by f_0 so that the correction must satisfy the conditions,

$$\int f_0\chi\, d\mathbf{v} = \int \mathbf{v} f_0\chi\, d\mathbf{v} = \int \epsilon f_0\chi\, d\mathbf{v} = 0\ , \tag{174}$$

which eliminates the three homogeneous solutions.

Deviation of f from f_0 appears because of spatial and temporal inhomogeneities. In other words, in the first order of perturbation theory, the collision integral (173) is balanced by the left-hand side of (164) where we substitute the Boltzmann distribution with inhomogeneous $u(\mathbf{r},t), T(\mathbf{r},t)$ and $P(\mathbf{r},t)$ [and therefore $\mu(\mathbf{r},t)$]:

$$f_0 = \exp\left(\frac{\mu - \epsilon_i}{T} - \frac{m(\mathbf{v}-\mathbf{u})^2}{2T}\right)\ . \tag{175}$$

We split the energy of a molecule into kinetic and internal: $\epsilon = \epsilon_i + mv^2/2$. Having in mind both viscosity and thermal conductivity we assume all macroscopic parameters to be functions of coordinates and put $\mathbf{F} = 0$ zero. We can simplify calculations doing them in the point with $\mathbf{u} = 0$ because the answer must depend only on velocity gradients. Differentiating (175) one gets

$$\begin{aligned}
\frac{T}{f_0}\frac{\partial f_0}{\partial t} &= \left[\left(\frac{\partial \mu}{\partial T}\right)_T - \frac{\mu-\epsilon}{T}\right]\frac{\partial T}{\partial t} + \left(\frac{\partial \mu}{\partial P}\right)_T \frac{\partial P}{\partial t} + m\mathbf{v}\frac{\partial \mathbf{u}}{\partial t} \\
&= \frac{\epsilon - w}{T}\frac{\partial T}{\partial t} + \frac{1}{n}\frac{\partial P}{\partial t} + m\mathbf{v}\frac{\partial \mathbf{u}}{\partial t}\,. \\
\frac{T}{f_0}\mathbf{v}\nabla f_0 &= \left[\left(\frac{\partial \mu}{\partial T}\right)_T - \frac{\mu-\epsilon}{T}\right]\mathbf{v}\nabla\, T + \frac{1}{n}\mathbf{v}\nabla\, P + mv_a v_b u_{ab}\,.
\end{aligned}$$

Here $u_{ab} = (\partial u_a/\partial x_b + \partial u_b/\partial x_a)/2$. We now add those expressions and substitute time derivatives from the ideal expressions (167-171), $\partial \mathbf{u}/\partial t = -\rho^{-1}\nabla P$, $\rho^{-1}\partial\rho/\partial t = (T/P)\partial(P/T)/\partial t = -div\,\mathbf{u}$,

$\partial s/\partial t = (\partial s/\partial T)_P \partial T/\partial t + (\partial s/\partial P)_T \partial P/\partial t = (c_p/T)\partial T/\partial t - P^{-1}\partial P/\partial t$, etc.

After some manipulations one gets the kinetic equation (for the classical gas with $w = c_pT$) in the following form:

$$(\epsilon/T - c_p)\mathbf{v}\nabla\, T + (mv_a v_b - \delta_{ab}\epsilon/c_v)u_{ab} = I(\chi)\ . \tag{176}$$

The expansion in gradients or in the parameter l/L where l is the mean-free path and L is the scale of velocity and temperature variations is called Chapman-Enskog method (1917). Note that the pressure gradient cancel out which means that it does not lead to the deviations in the distribution (and to dissipation).

Thermal conductivity. Put $u_{ab} = 0$. The solution of the linear integral $(\epsilon - c_pT)\mathbf{v}\nabla\, T = TI(\chi)$ has the form $\chi(\mathbf{r},\mathbf{v}) = \mathbf{g}(\mathbf{v}) \cdot \nabla T(\mathbf{r})$. One can find $\mathbf{g}$ specifying the scattering cross-section for any material. In the simplest case of the τ-approximation, $\mathbf{g} = \mathbf{v}(mv^2/2T - 5/2)\tau$ [13]. And generally, one can estimate $g \simeq l$ and obtain the applicability condition for the Chapman-Enskog expansion: $\chi \ll T \Rightarrow l \ll L \equiv T/|\nabla T|$.

The correction χ to the distribution makes for the correction to the energy flux (which for $\mathbf{u} = 0$ is the total flux):

$$\mathbf{q} = -\kappa\nabla T\,, \quad \kappa = -\frac{1}{3T}\int f_0\epsilon(\mathbf{v}\cdot\mathbf{g})\,d\mathbf{v} \,\simeq\, l\bar{v} \simeq \frac{\bar{v}}{n\sigma} \simeq \frac{1}{n\sigma}\sqrt{\frac{T}{m}}\ . \tag{177}$$

---

[13] check that it satisfies (174) i.e. $\int f_0(\mathbf{v}\cdot\mathbf{g})\,d\mathbf{v} = 0$.

Note that the thermal conductivity κ does not depend on on the gas density (or pressure). This is because we accounted only for binary collisions which is OK for a dilute gas.

Viscosity. We put $\nabla T = 0$ and separate the compressible part $\mathrm{div}\,\mathbf{u}$ from other derivatives which turns (176) into

$$mv_a v_b(u_{ab} - \delta_{ab}\,\mathrm{div}\,\mathbf{u}/3) + \left(\frac{mv^2}{3} - \frac{\epsilon}{c_v}\right)\mathrm{div}\,\mathbf{u} = I(\chi)\ . \tag{178}$$

The two terms in the left-hand side give $\chi = g_{ab}u_{ab} + g'\mathrm{div}\,\mathbf{u}$. that give the following viscous contributions into the momentum flux P_{ab}:

$$2\eta(u_{ab} - \delta_{ab}\,\mathrm{div}\,\mathbf{u}/3) + \zeta\delta_{ab}\mathrm{div}\,\mathbf{u}\ . \tag{179}$$

They correspond respectively to the so-called first viscosity

$$\eta = -\frac{m}{10T}\int v_a v_b g_{ab} f_o\, d\mathbf{v} \simeq mn\bar{v}l \simeq \frac{\sqrt{mT}}{\sigma}\ ,$$

and the second viscosity ζ [14]. One can estimate the viscosity saying that the flux of particles through the plane (perpendicular to the velocity gradient) is $n\bar{v}$, they come from a layer of order l, have velocity difference $l\nabla u$ which causes momentum flux $mn\bar{v}l\nabla u \simeq \eta\nabla u$. Notice that the viscosity is independent of density (at a given T) because while the fluxes grow with n so does the momentum so the momentum transfer rate does not change. Viscosity increases when molecules are smaller (i.e. σ decreases) because of the increase of the mean free path l.

Note that the kinematic viscosity $\nu = \eta/mn$ is the same as thermal conductivity because the same molecular motion is responsible for transports of both energy and momentum (the diffusivity is of the same order too).

Lifshitz & Pitaevsky, Physical Kinetics, Sects. 1-8. Huang, Sects. 3.1-4.2 and 5.1-5.7.

---

[14] $\zeta = 0$ for mono-atomic gases which have $\epsilon = mv^2/2, c_v = 3/2$ so that the second term in the lhs of (178) turns into zero

7 Conclusion: information theory approach

Here I briefly re-tell the story of statistical physics using a different language. 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 do not know? 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 few 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)

$$I(n) = I(e) \ln n = k \ln n \tag{180}$$

We can easily generalize this definition 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.

If we have an alphabet with n symbols then the message of the length N can potentially be one of n^N possibilities so that it brings the information $kN \ln n$ or $k \ln n$ per symbol. In reality though we know that letters are used with different frequencies. Consider now the situation when there is a probability w_i assigned to each letter (or box) $i = 1, \ldots, n$. Now if we want to evaluate the missing information (or, the information that one symbol brings us on average) we ought to think about repeating our choice 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 missing information is

$$I_N = k \ln\Big(N!/\Pi_i(Nw_i)!\Big) \approx -Nk \sum_i w_i \ln w_i + O(lnN) \ . \tag{181}$$

The missing information per problem (or per symbol in the language) coincides with the entropy (18):

$$I = \lim_{N\to\infty} I_N/N = -k \sum_{i=1}^{n} w_i \ln w_i \ . \tag{182}$$

Note that when $n \to \infty$ then (180) diverges while (182) may well be finite.

We can generalize this 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 enters. We use canonical coordinates in the phase space and write the missing information in terms of the density which may also depend on time:

$$I(t) = -\int \rho(p,q,t)\ln[\rho(p,q,t)]\,dpdq\ . \tag{183}$$

If the density of the discrete points in the continuous limit is inhomogeneous, say $m(x)$, then the proper generalization is

$$I(t) = -\int \rho(x)\ln[\rho(x)/m(x)]\,dx\ . \tag{184}$$

Note that (184) is invariant with respect to an arbitrary change of variables $x \to y(x)$ since $\rho(y) = \rho(x)dy/dx$ and $m(y) = m(x)dy/dx$ while (183) 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.

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

$$I - \sum_j \lambda_j \langle R_j(p,q,t)\rangle = \int \rho(p,q,t)\{\ln[\rho(p,q,t)] - \sum_j \lambda_j \langle R_j(p,q,t)\}\,dpdq\ ,$$

we obtain the distribution

$$\rho(p,q,t) = Z^{-1}\exp\Big[-\sum_j \lambda_j R_j(p,q,t)\Big]\ , \tag{185}$$

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

$$\frac{\partial \ln Z}{\partial \lambda_i} = -r_i \ . \tag{186}$$

For example, consider our initial "candy-in-the-box" problem (think of an impurity atom in a lattice if you prefer physics). Let us denote the number of the box with the candy j. Different attempts give different j (for impurity, think of X-ray 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 (186) 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 (183,185) 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. 1.5. 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. Also, the concept of entropy as missing information[15] allows

---

[15] that entropy is not a property of the system but of our knowledge about the system

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 (Szilard 1929).

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 (185) takes the canonical form (2,3) in equilibrium. 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 (185) deviates from equilibrium. Examples are currents, velocity or temperature gradients like considered in kinetics.

Ar the end, mention briefly the communication theory which studies transmissions through imperfect channels. Here, the message (measurement) A we receive gives the information about the event B as follows: $I(A,B) = \ln P(B|A)/P(B)$, where $P(B|A)$ is the so-called conditional probability (of B in the presence of A). Summing over all possible $B_1, \dots, B_n$ and $A_1, \dots, A_m$ we obtain Shannon's "mutual information" used to evaluate the quality of communication systems

$$\begin{aligned} I(A,B) &= \sum_{i=1}^{m}\sum_{j=1}^{n} P(A_j, B_j)\ln[P(B_j|A_i)/P(B_j)] \\ \to I(Z,Y) &= \int dzdyp(z,y)\ln[p(z|y)/p(y)]\,. \end{aligned} \tag{187}$$

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 (187) is the particular case of multidimensional (184) where one takes $x = (y,z)$, $m = p(z)p(y)$ and uses $p(z,y) = p(z|y)p(y)$.

More details can be found in Katz, Chapters 2-5 and Sethna Sect. 5.3.

Basic books

L. D. Landau and E. M. Lifshitz, Statistical Physics Part 1, 3rd edition (Course of Theor. Phys, Vol. 5).

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.

E. M. Lifshitz and L.P. Pitaevsky, Physical Kinetics.

A. Katz, Principles of Statistical Mechanics.

J. Cardy, Scaling and renormalization in statistical physics.

M. Kardar, Statistical Physics of Particles, Statistical Physics of Fields.

J. Sethna, Entropy, Order Parameters and Complexity.

Exam, Feb 22, 2007

1. A lattice in one dimension has N cites and is at temperature T. At each cite there is an atom which can be in either of two energy states: $E_i = \pm\epsilon$. When L consecutive atoms are in the $+\epsilon$ state, we say that they form a cluster of length L (provided that the atoms adjacent to the ends of the cluster are in the state $-\epsilon$). In the limit $N \to \infty$,

a) Compute the probability $\mathcal{P}_L$ that a given cite belongs to a cluster of length L (don't forget to check that $\sum_{L=0}^{\infty} \mathcal{P}_L = 1$);

b) Calculate the mean length of a cluster $\langle L \rangle$ and determine its low- and high-temperature limits.

2. Consider a box containing an ideal classical gas at pressure P and temperature T. The walls of the box have N_0 absorbing sites, each of which can absorb at most two molecules of the gas. Let $-\epsilon$ be the energy of an absorbed molecule. Find the mean number of absorbed molecules $\langle N \rangle$. The dimensionless ratio $\langle N \rangle / N_0$ must be a function of a dimensionless parameter. Find this parameter and consider the limits when it is small and large.

3. Consider the spin-1 Ising model on a cubic lattice in d dimensions, given by the Hamiltonian

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j - \Delta \sum_i S_i^2 - h \sum_i Si \ ,$$

where $S_i = 0, \pm 1$, $\sum_{<ij>}$ denote a sum over z nearest neighbor sites and $J, \Delta > 0$.

(a) Write down the equation for the magnetization $m = \langle S_i \rangle$ in the mean-field approximation.

(b) Calculate the transition line in the (T, Δ) plane (take $h = 0$) which separates the paramagnetic and the ferromagnetic phases. Here T is the temperature.

(c) Calculate the magnetization (for $h = 0$) in the ferromagnetic phase near the transition line, and show that to leading order $m \sim \sqrt{T_c - T}$, where T_c is the transition temperature.

(d) Show that the zero-field ($h = 0$) susceptibility χ in the paramagnetic phase is given by

$$\chi = \frac{1}{k_B T} \frac{1}{1 + \frac{1}{2} e^{-\beta\Delta} - \frac{Jz}{k_B T}} \ .$$

4. Compare the decrease in the entropy of a reader's brain with the increase in entropy due to illumination. Take, for instance, that it takes $t = 100$ seconds to read one page with 3000 characters written by the alphabet that uses 32 different characters (letters and punctuation marks). At the same time, the illumination is due to a 100 Watt lamp (which emits $P = 100\text{J/s}$). Take $T = 300\text{K}$ and use the Boltzmann constant $k = 1.38 \cdot 10^{-23}\text{J/K}$.

Answers

Problem 1.

a) Probabilities of any cite to have energies $\pm\epsilon$ are

$$\mathcal{P}_\pm = e^{\pm\beta\epsilon}(e^{\beta\epsilon} + e^{-\beta\epsilon})^{-1} \ .$$

The probability for a given cite to belong to an L-cluster is $\mathcal{P}_L = L\mathcal{P}_+^L \mathcal{P}_-^2$ for $L \geq 1$ since cites are independent and we also need two adjacent cites to have $-\epsilon$. The cluster of zero length corresponds to a cite having $-\epsilon$ so that $\mathcal{P}_L = \mathcal{P}_-$ for $L = 0$. We ignore the possibility that a given cite is within L of the ends of the lattice, it is legitimate at $N \to \infty$.

$$\begin{aligned}
\sum_{L=0}^{\infty} \mathcal{P}_L &= \mathcal{P}_- + \mathcal{P}_-^2 \sum_{L=1}^{\infty} L\mathcal{P}_+^L = \mathcal{P}_- + \mathcal{P}_-^2 \mathcal{P}_+ \frac{\partial}{\partial \mathcal{P}_+} \sum_{L=1}^{\infty} \mathcal{P}_+^L \\
&= \mathcal{P}_- + \frac{\mathcal{P}_-^2 \mathcal{P}_+}{(1 - \mathcal{P}_+)^2} = \mathcal{P}_- + \mathcal{P}_+ = 1 \ .
\end{aligned}$$

b)

$$\langle L \rangle = \sum_{L=0}^{\infty} L\mathcal{P}_L = \mathcal{P}_-^2 \mathcal{P}_+ \frac{\partial}{\partial \mathcal{P}_+} \sum_{L=1}^{\infty} L\mathcal{P}_+^L = \frac{\mathcal{P}_+(1 + \mathcal{P}_+)}{\mathcal{P}_-} = e^{-2\beta\epsilon} \frac{e^{\beta\epsilon} + 2e^{-\beta\epsilon}}{e^{\beta\epsilon} + e^{-\beta\epsilon}} \ .$$

At $T = 0$ all cites are in the lower level and $\langle L \rangle = 0$. As $T \to \infty$, the probabilities $\mathcal{P}_+$ and $\mathcal{P}_-$ are equal and the mean length approaches its maximum $\langle L \rangle = 3/2$.

Problem 2.

Since each absorbing cite is in equilibrium with the gas, then the cite and the gas must have the same chemical potential μ and the same temperature T. The fugacity of the gas $z = \exp(\beta\mu)$ can be expressed via the pressure from the grand canonical partition function

$$\begin{aligned} \mathcal{Z}_g(T, V, \mu) &= \exp[zV(2\pi mT)^{3/2}h^{-3}] \,, \\ PV = \Omega = T \ln \mathcal{Z}_g &= zVT^{5/2}(2\pi m)^{3/2}h^{-3} \ . \end{aligned}$$

The grand canonical partition function of an absorbing cite $\mathcal{Z}_{cite} = 1 + ze^{\beta\epsilon} + z^2 e^{2\beta\epsilon}$ gives the average number of absorbed molecules per cite:

$$\frac{\langle N \rangle}{N_0} = z\frac{\partial \mathcal{Z}_{cite}}{\partial z} = \frac{x + 2x^2}{1 + x + x^2}$$

where the dimensionless parameter is $x = PT^{-5/2}e^{\beta\epsilon}h^3(2\pi m)^{-3/2}$. The limits are $\langle N \rangle/N_0 \to 0$ as $x \to 0$ and $\langle N \rangle/N_0 \to 2$ as $x \to \infty$.

Problem 3.

a) $\mathcal{H}_{eff}(S) = -JmzS - \Delta S^2 - hS$, $S = 0, \pm 1$.

$$m = e^{\beta\Delta}\frac{e^{\beta(Jzm+h)} - e^{-\beta(Jzm+h)}}{1 + e^{\beta\Delta}\left[e^{\beta(Jzm+h)} + e^{-\beta(Jzm+h)}\right]} \ .$$

b) $h = 0$,

$$m \approx e^{\beta\Delta}\frac{2\beta Jzm + (\beta Jzm)^3/3}{1 + 2e^{\beta\Delta}[1 + (\beta Jzm)^2/2]} \ .$$

Transition line $\beta_c Jz = 1 + \frac{1}{2}e^{-\beta_c\Delta}$. At $\Delta \to \infty$ it turns into Ising.

c)

$$m^2 = \frac{(\beta - \beta_c)Jz}{(\beta_c Jz)^2/2 - (\beta_c Jz)^3/6} \ .$$

d)

$$m \approx e^{\beta\Delta}\frac{2\beta Jzm + 2\beta h}{1 + 2e^{\beta\Delta}} \,, \ m \approx 2\beta h(2 + e^{-\beta\Delta} - \beta Jz)^{-1} \,, \ \chi = \partial m/\partial h \ .$$

Problem 4

Since there are $2^5 = 32$ different characters then every character brings 5 bits and the entropy decrease is $5 \times 3000/\log_2 e$. The energy emitted by the lamp Pt brings the entropy increase Pt/kT which is $100 \times 100 \times 10^{23} \times \log_2 e/1.38 \times 300 \times 5 \times 3000 \simeq 10^{20}$ times larger.