

Statistical physics

lecture notes

P. Ziherl

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

June 9, 2016

Contents

Foreword	7
1 Thermodynamics	9
1.1 System in thermostat	9
1.2 Ideal gas	11
1.3 Gibbs paradox	14
2 Ensemble theory	17
2.1 Phase space	17
2.2 Liouville theorem	18
2.3 Microcanonical ensemble	20
2.4 Canonical ensemble	22
2.5 Ideal gas	23
2.6 Canonical vs. microcanonical ensemble	24
2.7 Equipartition and virial theorems	26
2.8 Harmonic oscillator	28
2.9 Paramagnetism	30
2.10 Grand-canonical ensemble	31
3 Quantum statistics	35
3.1 Density matrix	35
3.2 Microcanonical and canonical ensemble	38
3.3 Particle in a box	39
3.4 Systems of indistinguishable particles	42
4 Ideal systems	45
4.1 Ideal Bose gas	45
4.2 Bose-Einstein condensation	47
4.3 Blackbody radiation	52
4.4 Ideal Fermi gas	57

4.5	White dwarfs	60
5	Interacting systems	65
5.1	Cluster expansion	65
5.2	Virial expansion	72
5.3	Second virial coefficient	75
5.4	Interacting quantum systems	83
6	Classical density functional theory	85
6.1	Grand potential as a functional	87
6.2	Uniqueness of density profile	89
6.3	Intrinsic free energy	89
6.4	Ornstein-Zernike equation	91
7	Phase transitions	95
7.1	Critical exponents	95
7.2	Landau theory	97
7.3	Thermodynamic inequalities	101
7.4	Ginzburg criterion	104
7.5	Exactly solvable models	106
7.6	Kosterlitz-Thouless transition	110
7.7	Renormalization group	119
8	Fluctuations	131
8.1	Brownian motion	135
8.2	Fokker-Planck equation	142
8.3	Fluctuation-dissipation theorem	146
8.4	Onsager relations	148
8.5	Kubo formula	153
8.6	Time-dependent Ginzburg-Landau theory	156
9	Stochastic thermodynamics	159
9.1	Dissipated work	160
9.2	Jarzynski relation	162
9.3	Crooks (fluctuation) theorem	163
	Bibliography	165

Foreword

These are my notes for the MS course in advanced statistical physics held at the Department of Physics at the University of Ljubljana. The notes are based primarily on the textbook of Pathria and Beale [1] with the material covered in the undergraduate statistical thermodynamics course left out. Here we first revisit the postulates of statistical physics and then turn to quantum statistical physics, focusing first on ideal Bose and Fermi systems. Then we discuss interacting systems, which paves the way to the theoretical understanding of phase transitions. The last large chapter deals with fluctuations and small deviations from equilibrium. Finally we briefly turn to stochastic thermodynamics.

Apart from Pathria and Beale's textbook, the reader may want to consult other contemporary account on the topic, say Plischke and Bergersen's, Schwabl's, Van Vliet's, or Landau and Lifshitz's textbooks. The notation is a bit of an issue. I tried to stick to the symbols used in Pathria and Beale's textbook but this is not always possible. I still trust that there are few ambiguities as per what a given symbol means. I will do my best to correct any typos or other errors that may remain in the text.

Primož Ziherl

Chapter 1

Thermodynamics

Let us begin by reviewing the statistical-mechanical fundamentals of thermodynamics, recalling the main elements of both fields. The macroscopic state of a system is described by a small number of thermodynamic variables, say the number of particles N , the volume V , and the internal energy E . In a non-interacting system,

$$E = \sum_i n_i \epsilon_i \quad (1.1)$$

and

$$N = \sum_i n_i, \quad (1.2)$$

where n_i is the number of particles with an energy ϵ_i . Each of the different ways of arranging the N particles in the single-particle states is a microstate, which may be viewed as independent solutions of the Schrödinger equation for the whole system. A natural assumption is that the system is equally likely in any of these microstates, which is sometimes referred to as the postulate of equal *a priori* probabilities.

The number of possible microstates depends on N , V , and E : $\Omega = \Omega(N, V, E)$. The form of this function is determined by the nature of the system, and will be illustrated by a few examples.

1.1 System in thermostat

One of the main concepts of thermodynamics is temperature, and this concept can be defined by considering two systems in thermal contact, that is, two systems that can exchange energy. System 1 is described by N_1 , V_1 ,

and E_1 and system 2 is described by N_2 , V_2 , and E_2 . If we assume that the systems cannot exchange energy with a third body, then their combined energy is fixed

$$E^{(0)} = E_1 + E_2 = \text{const.} \quad (1.3)$$

At any given E_1 , the number of microstates of system 1 is $\Omega_1 = \Omega(N_1, V_1, E_1)$; the number of microstates of system 2 is $\Omega_2 = \Omega(N_2, V_2, E_2)$. If there is no interaction between the systems then the total number of microstates is given by the product

$$\Omega^{(0)}(E^{(0)}, E_1) = \Omega_1(E_1)\Omega_2(E_2), \quad (1.4)$$

where we omitted the dependence of Ω s on N s and V s because the systems only exchange energy. Now $\Omega^{(0)}$ depends on both total energy $E^{(0)}$ and the way it is shared by systems 1 and 2. The question that now arises is how the energy is partitioned, that is what is E_1 and thus $E_2 = E^{(0)} - E_1$. We assume that this is done such that the total number of microstates is maximized, that is $\Omega^{(0)}(E^{(0)}, E_1) = \text{max.}$ This means that

$$\frac{\partial \Omega^{(0)}}{\partial E_1} = 0 \quad (1.5)$$

or, since $\Omega^{(0)}(E^{(0)}, E_1) = \Omega_1(E_1)\Omega_2(E^{(0)} - E_1)$,

$$\frac{\partial \Omega_1(E_1)}{\partial E_1} \Omega_2(E^{(0)} - E_1) + \Omega_1(E_1) \frac{\partial \Omega_2(E_2)}{\partial E_2} \underbrace{\frac{\partial E_2}{\partial E_1}}_{=-1} = 0. \quad (1.6)$$

On dividing by $\Omega_1(E_1)\Omega_2(E_2)$ we find that

$$\frac{\partial \ln \Omega_1(E_1)}{\partial E_1} = \frac{\partial \ln \Omega_2(E_2)}{\partial E_2} = -\beta. \quad (1.7)$$

This condition determines the equilibrium values of energies of the two systems \bar{E}_1 and \bar{E}_2 . The left- and the right-hand side only depend on E_1 and E_2 , respectively, which means that they must be equal to a constant commonly denoted by β .

We expect that β is related to the thermodynamic temperature T . The first law of thermodynamics states that $dE = TdS - PdV$ so that

$$\left(\frac{\partial S}{\partial E} \right)_{N,V} = \frac{1}{T}. \quad (1.8)$$

By comparing this formula to Eq. (1.7) we conclude that

$$\frac{\Delta S}{\Delta \ln \Omega} = \frac{1}{\beta T} = \text{const.} \quad (1.9)$$

and so

$$S = k_B \ln \Omega. \quad (1.10)$$

1.2 Ideal gas

In a system of non-interacting particles in a container of volume V , the number of microstates for a single particle must be proportional to V and thus

$$\Omega(N, E, V) \propto V^N. \quad (1.11)$$

Now the first law

$$dE = TdS - PdV \quad (1.12)$$

can be recast as

$$dS = \frac{1}{T}dE + \frac{P}{T}dV \quad (1.13)$$

so that

$$\left(\frac{\partial S}{\partial V} \right)_E = \frac{p}{T}. \quad (1.14)$$

This means that

$$\frac{P}{T} = k_B \left(\frac{\partial \ln \Omega(N, E, V)}{\partial V} \right)_E = k_B \frac{N}{V}. \quad (1.15)$$

This is the equation of state for the ideal gas.

By considering the possible functional form of entropy as a function of V and E , we can derive the equation describing a reversible adiabatic process. In an ideal gas, the number of states corresponding to a fixed energy ϵ of a single particle in a cubical box of edge length L is given by the number of solutions of

$$\frac{h^2}{2mL^2}(n_x^2 + n_y^2 + n_z^2) = \epsilon. \quad (1.16)$$

Here n_x, n_y , and n_z are non-negative integers. For N particles, the total energy thus reads

$$E = \frac{h^2}{2mL^2} \sum_{r=1}^{3N} n_r^2 \quad (1.17)$$

so that

$$\sum_{r=1}^{3N} n_r^2 = \frac{2mL^2E}{h^2} = \frac{2mV^{2/3}E}{h^2}. \quad (1.18)$$

Without actually calculating the number of states, we immediately see that it must depend on the product $V^{2/3}E$ rather than on V and E separately. This also means that

$$S_{id}(N, V, E) = S_{id}(N, V^{2/3}E) \quad (1.19)$$

so that in an adiabatic process at fixed N where $S = \text{const.}$

$$V^{2/3}E = \text{const.} \quad (1.20)$$

too. The first law [Eq. (1.12)] states that

$$\left(\frac{\partial E}{\partial V}\right)_{N,S} = -P \quad (1.21)$$

and thus from $V^{2/3}dE + (2/3)V^{-1/3}EdV = 0$ it follows that

$$P = \frac{2E}{3V}. \quad (1.22)$$

In turn

$$P \propto \frac{V^{-2/3}}{V} = V^{-5/3} \quad (1.23)$$

or

$$PV^{5/3} = \text{const.} \quad (1.24)$$

These results hold both for quantum and classical ideal gases.

To calculate the entropy as a function of energy, we should count the number of microstates where, consistent with Eq. (1.18),

$$\sum_{r=1}^{3N} n_r^2 = \mathcal{N}^2 = \text{const.}, \quad (1.25)$$

that is, the number of different sets of $n_r > 0$ that satisfy this constraint. This cannot be done easily. But we can employ a trick which works very well in the thermodynamic limit $N \rightarrow \infty$. To this end, compare the area of a 90° pie of radii R and $R + \Delta R$: Their ratio is $1/(1 + \Delta R/R)^2$. (90° pie corresponding to the first quadrant in the n_r -space for 2 degrees of freedom, that is 1 particle in 2D or 2 particles in 1D.) In half a 90° -spherical

wedge corresponding to the first octant in the 3D n_r -space, the ratio of the volumes is $1/(1 + \Delta R/R)^3$ and in $3N$ dimensional n_r -space corresponding to N particles in 3D space we see that the ratio of the volume of the smaller "half-wedge" is essentially negligible compared to the volume of the larger half-wedge, their ratio being $1/(1 + \Delta R/R)^{3N}$. For $N \rightarrow \infty$, this is so even if $\Delta R/R$ is small.

In turn, this means that the volume of the $3N$ -dimensional half-wedge of a given radius is completely dominated by the contribution of the shell, and thus the volumes of the shell may be identified with the volume of the half wedge. This is convenient because the latter is much easier to calculate — and the former is needed in Eq. (1.25).

The volume of a t -dimensional sphere of radius R is given by

$$\frac{\pi^{t/2} R^t}{\Gamma(t/2 + 1)}, \quad (1.26)$$

where Γ is the gamma function. This means that the volume of the $3N$ -dimensional half-wedge of radius \mathcal{N} is

$$\frac{1}{2^{3N}} \frac{\pi^{3N/2} \mathcal{N}^{3N}}{(3N/2)!}, \quad (1.27)$$

where we have approximated the gamma function by the factorial. In our case,

$$\mathcal{N} = \left(\frac{8mV^{2/3}E}{h^2} \right)^{1/2} \quad (1.28)$$

so that

$$\begin{aligned} S(N, V, E) &= k_B \ln \Omega(N, V, E) \\ &= k_B \ln \left(\frac{1}{2^{3N}} \frac{\pi^{3N/2}}{(3N/2)!} \left(\frac{8mV^{2/3}E}{h^2} \right)^{3N/2} \right) \\ &= k_B \ln \left(\left(\frac{V}{h^3} \right)^N (2\pi mE)^{3N/2} \right) - k_B \ln (3N/2)! \\ &= Nk_B \ln \left(\frac{V}{h^3} (2\pi mE)^{3/2} \right) - k_B \frac{3N}{2} \ln \frac{3N}{2} + \frac{3}{2} Nk_B \\ &= Nk_B \ln \left(\frac{V}{h^3} (2\pi mE)^{3/2} \right) - k_B \ln \left(\frac{3N}{2} \right)^{3N/2} + \frac{3}{2} Nk_B \\ &= Nk_B \ln \left(\frac{V}{h^3} \left(\frac{4\pi mE}{3N} \right)^{3/2} \right) + \frac{3}{2} Nk_B. \end{aligned} \quad (1.29)$$

Having calculated S as a function of E , we can invert it so as to spell out

$$E = E(S, V, N) = \frac{3h^2N}{4\pi mV^{2/3}} \exp\left(\frac{2S}{3Nk_b} - 1\right) \quad (1.30)$$

and then calculate

$$T = \left(\frac{\partial E}{\partial S}\right)_{N,V} = \frac{2}{3Nk_b} \frac{3h^2N}{4\pi mV^{2/3}} \exp\left(\frac{2S}{3Nk_b} - 1\right) = \frac{2}{3Nk_b} E \quad (1.31)$$

so that

$$E = \frac{3}{2} N k_B T. \quad (1.32)$$

1.3 Gibbs paradox

The entropy of the ideal gas [Eq. (1.29)] is not additive because if the number of particles and the volume are both increased by the same factor the entropy does not. Notice that Eq. (1.29) contains the ratio E/N , which is OK, and the prefactor in two terms is N , which is OK too, but the volume V is present by itself rather than as the ratio V/N .

The consequence of this anomaly is best illustrated by considering mixing of two gases containing N_1 and N_2 molecules of masses m_1 and m_2 , respectively, in volumes V_1 and V_2 ; assume that the gases are at the same temperature T (Fig. 1.1). Before mixing the entropy of either gas is

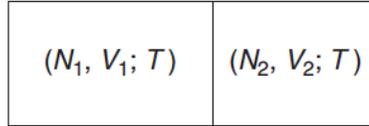


Figure 1.1: Mixing of two ideal gases [1].

$$S_i^I = N_i k_B \ln V_i + \frac{3}{2} N_i k_B \left[1 + \ln \left(\frac{2\pi m_i k_B T}{h^2} \right) \right], \quad (1.33)$$

where we have used Eq. (1.32) to replace E in Eq. (1.29) by T and we emphasized the dependence on V . After mixing the two gases occupy the same volume $V = V_1 + V_2$ and the entropy of either gas is

$$S_i^F = N_i k_B \ln(V_1 + V_2) + \frac{3}{2} N_i k_B \left[1 + \ln \left(\frac{2\pi m_i k_B T}{h^2} \right) \right]. \quad (1.34)$$

The total difference of entropies, the so-called entropy of mixing, is given by

$$\Delta S = \underbrace{(S_1^F + S_2^F)}_{\text{final}} - \underbrace{(S_1^I + S_2^I)}_{\text{initial}} = k_B \left(N_1 \ln \frac{V_1 + V_2}{V_1} + N_2 \ln \frac{V_1 + V_2}{V_2} \right). \quad (1.35)$$

As $V_1 + V_2 > V_1, V_2$, both terms are positive and thus the mixing entropy is positive too as it should be. Furthermore, for gases at the same initial densities where $N_1/V_1 = N_2/V_2$

$$\Delta S = k_B \left(N_1 \ln \frac{N_1 + N_2}{N_1} + N_2 \ln \frac{N_1 + N_2}{N_2} \right). \quad (1.36)$$

This result should also apply to mixing of two parts of identical gas at the same density and same temperature, e.g., by removing a partition that divides a vessel containing the same gas at same T and $n = N/V$ into two compartments. This is incorrect because macroscopically, nothing changes on removing the partition, and by reinserting we can recreate the initial state in a reversible fashion. In this case, the entropy of mixing should be zero.

To resolve the problem, Gibbs conjectured that the number of microstates in a system of N particles must be reduced by a factor of $N!$. To see how this works, notice that Eq. (1.36) can be recast as

$$\begin{aligned} \Delta S &= k_B [(N_1 + N_2) \ln(N_1 + N_2) - N_1 \ln N_1 - N_2 \ln N_2] \\ &= k_B [(N_1 + N_2) \ln(N_1 + N_2) - (N_1 + N_2) \\ &\quad - N_1 \ln N_1 + N_1 - N_2 \ln N_2 + N_2]. \end{aligned} \quad (1.37)$$

In each of the terms in the square bracket we recognize the Stirling approximation for the logarithm of the factorial, $\ln N! \approx N \ln N - N$ so we conjecture that this result really represents

$$\Delta S = k_B [\ln(N_1 + N_2)! - (\ln N_1! + \ln N_2!)], \quad (1.38)$$

the first term corresponding to the entropy of the final state and the second one to the two contributions to the entropy of the initial state. We see that if $k_B(N \ln N - N)$ were subtracted from Eq. (1.29) the mixing entropy of identical gases at same temperature and density would be zero. This leads to

$$S(N, V, E) = Nk_B \ln \left(\frac{V}{Nh^3} \left(\frac{4\pi mE}{3N} \right)^{3/2} \right) + \frac{5}{2}Nk_B. \quad (1.39)$$

This remedy was introduced by Gibbs well before quantum mechanics, and it amounts to dividing the number of microstates given by the volume of the phase space alone by $N!$. This is due to indistinguishability of the particles which requires that the many-body wave function be either symmetric or antisymmetric to particle exchange. In other words, the macroscopic state is described only by the number of particles in a given state but not their identity. If all particles are in different states then the phase space is reduced by a factor of $N!$. For example, putting 5 labeled balls in 5 drawers can be done in $5! = 125$ ways but if the balls are indistinguishable then all of these arrangements represent the same physical state, hence division by $5!$. If, on the other hand, some of the balls are in the same drawer the number of arrangements is somewhat smaller; specifically, it is given by the multinomial coefficient

$$\frac{N!}{n_1!n_2!\dots}, \quad (1.40)$$

where n_i are the numbers of balls in each drawer and $\sum_i n_i = N$. The Gibbs' remedy takes care of the numerator and works well if the average occupation number of each state n_i is much smaller than 1 (that is, mostly 0, some of them 1, and only a few of them larger than 1) so that $n_i! = 1$. This is referred to as the classical limit.

Chapter 2

Ensemble theory

The statistical-mechanical system evolves in time as dictated by the equations of motion, passing from one microstate to another. If monitored for a long enough time, the time-averaged value of any quantity of interest is equal to the average value averaged over all equiprobable microstates that correspond to the same macrostate. The collection of these microstates is referred to as the ensemble.

2.1 Phase space

The microstate of N point particles at any time t is specified by $3N$ coordinates q_i ($i = 1, \dots, 3N$) and $3N$ linear momenta p_i constituting a $6N$ -dimensional phase space. Particle motion is governed by the classical equations of motion

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad (2.1)$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad (2.2)$$

$H = H(q_i, p_i)$ being the Hamiltonian. The probability of finding the system in a certain part of the phase space is determined by the density function $\rho(q_i, p_i; t)$ and given by $\rho(q_i, p_i; t) d^{3N} q_i d^{3N} p_i$. The average of a quantity $f(q_i, p_i)$ is given by

$$\langle f \rangle = \frac{\int f(q_i, p_i) \rho(q_i, p_i; t) d^{3N} q_i d^{3N} p_i}{\int \rho(q_i, p_i; t) d^{3N} q_i d^{3N} p_i}. \quad (2.3)$$

The ensemble is stationary if

$$\frac{\partial \rho}{\partial t} = 0. \quad (2.4)$$

In this case, the system is in equilibrium because no physical quantity averaged over the ensemble depends on time.

2.2 Liouville theorem

Consider a volume ω in phase space. The probability that the system be in this volume is given by $\int \rho(q_i, p_i; t) d\omega$ where $d\omega = d^{3N}q_i d^{3N}p_i$. Much like mass in hydrodynamics, the probability of finding the system in this volume may either vary explicitly with time which is described by

$$\int \frac{\partial \rho}{\partial t} d\omega \quad (2.5)$$

or flow out of the volume which is described by

$$\oint_{\delta\omega} \rho \mathbf{v} \cdot d\boldsymbol{\sigma} \quad (2.6)$$

where $\delta\omega$ is the boundary of the region, \mathbf{v} is the velocity in the $6N$ -dimensional phase space (that is, a vector containing all q_i and p_i), and $d\boldsymbol{\sigma}$ is its surface element. Using Gauss theorem he find

$$\oint_{\delta\omega} \rho \mathbf{v} \cdot d\boldsymbol{\sigma} = \int \nabla \cdot (\rho \mathbf{v}) d\omega, \quad (2.7)$$

where $\nabla \cdot$ is the divergence in the $6N$ -dimensional phase space, that is, a vector containing all $\partial/\partial q_i$ and $\partial/\partial p_i$.

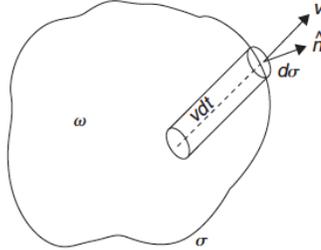


Figure 2.1: Liouville theorem [1].

The sum of the two terms must be zero because the state of the system can only move from one part of the phase space to another but cannot appear or disappear. Thus

$$\int \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) \right] d\omega = 0. \quad (2.8)$$

Since this must be true for any volume, we conclude that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad (2.9)$$

If spelled out,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) + \sum_{i=1}^{3N} \left(\rho \frac{\partial \dot{q}_i}{\partial q_i} + \rho \frac{\partial \dot{p}_i}{\partial p_i} \right) = 0 \quad (2.10)$$

Equations (2.2) imply that

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial}{\partial q_i} \dot{q}_i = \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} \quad (2.11)$$

and

$$\frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial}{\partial p_i} \dot{p}_i = \frac{\partial}{\partial p_i} \left(-\frac{\partial H}{\partial q_i} \right) \quad (2.12)$$

so that the second sum in Eq. (2.10) vanishes term by term. We find that

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = \frac{\partial \rho}{\partial t} + [\rho, H] = 0, \quad (2.13)$$

where we switched to the Poisson bracket notation. This is the Liouville theorem: points in phase space move through it like an incompressible fluid.

If $\partial \rho / \partial t = 0$, then

$$\sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = [\rho, H] = 0 \quad (2.14)$$

which means that $\rho = \text{const.}$ so that all microstates are equally probable. This is referred to as the microcanonical ensemble.

If the density function only depends on phase-space coordinates through the Hamiltonian

$$\rho(q_i, p_i) = \rho(H(q_i, p_i)) \quad (2.15)$$

then

$$\frac{\partial \rho}{\partial q_i} = \frac{\partial \rho}{\partial H} \frac{\partial H}{\partial q_i} \quad (2.16)$$

and

$$\frac{\partial \rho}{\partial p_i} = \frac{\partial \rho}{\partial H} \frac{\partial H}{\partial p_i} \quad (2.17)$$

so that

$$\sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial H} \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial H} \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} \right) = 0. \quad (2.18)$$

This is referred to as the canonical ensemble.

2.3 Microcanonical ensemble

Here the total energy of the system is fixed so that it lies within a narrow range centered around E , that is, between $E - \Delta/2$ and $E + \Delta/2$. All microstates of the whole system whose energies are in this range are equiprobable:

$$\rho(q_i, p_i) = \begin{cases} \text{const.} & \text{for } E - \Delta/2 < H(q_i, p_i) < E + \Delta/2 \\ 0 & \text{otherwise} \end{cases}. \quad (2.19)$$

If we denote the volume of the phase space occupied by the hypershell containing all microstates satisfying this constraint by ω , the entropy of the ensemble is given by

$$S = k_B \ln \left(\frac{\omega}{\omega_0} \right), \quad (2.20)$$

where ω_0 is the volume of a single microstate.

The volume of the microstate can be derived in several ways. One derivation relies on the microcanonical analysis of the ideal gas [1]. Here we reiterate the harmonic-oscillator derivation. The Hamiltonian of a single 1D harmonic oscillator is

$$H = \frac{p^2}{2m} + \frac{kq^2}{2}, \quad (2.21)$$

or, in terms of the frequency $\omega = \sqrt{k/m}$,

$$H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2}. \quad (2.22)$$

The constant-energy hypershell in the phase space is an ellipse of semiaxes $\sqrt{2mE}$ and $\sqrt{2E/(m\omega^2)}$:

$$\frac{p^2}{2mE} + \frac{q^2}{2E/(m\omega^2)} = 1. \quad (2.23)$$

The 2D volume, that is, the area in an annular region around energy E is given by the difference of areas of ellipses corresponding to energies $E - \Delta/2$ and $E + \Delta/2$. The area of an ellipse of semiaxes a and b is πab ; in our case $2\pi E/\omega$. This gives

$$\int_{E-\Delta/2 < H(q,p) < E+\Delta/2} dqdp = \frac{2\pi}{\omega} [(E + \Delta/2) - (E - \Delta/2)] = \frac{2\pi\Delta}{\omega}. \quad (2.24)$$

The spectrum of the harmonic oscillator is discrete

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad (2.25)$$

where $n = 0, 1, 2, \dots$. This means that the width of the energy shell corresponding to a single state is $\Delta = \hbar\omega$. Thus the volume of the phase space corresponding to a single state is $\omega_0 = 2\pi\hbar = h$; in a system of N oscillators, $\omega_0 = h^N$.

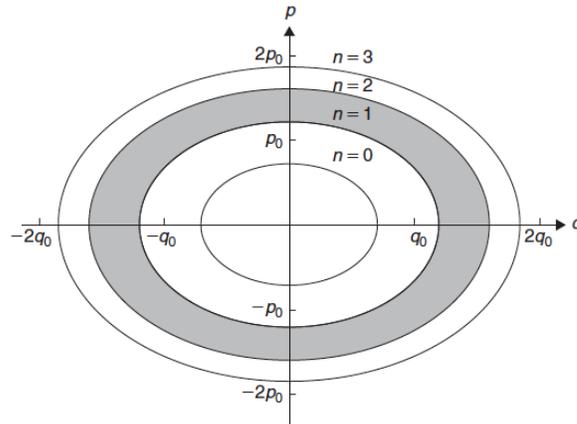


Figure 2.2: Illustration of the Bohr-Sommerfeld rule using the harmonic oscillator [1].

This argument is a mixture of classical-physics and quantum-mechanical ideas, and thus not quite self-consistent. But it works well — increasingly better as energy is increased. Yet another perspective of the volume of a single microstate may be obtained by observing that the Heisenberg uncertainty principle which states that

$$\Delta q \Delta p \gtrsim \hbar. \quad (2.26)$$

This too suggests that the location of a state in the phase space can be determined with a (volumetric) accuracy of the order of \hbar , and thus the volume of phase space that can be associated with a microstate in 1D is about \hbar ; for N particles in 3D, \hbar^{3N} . Up to a numerical factor, this agrees with the above result. Also available are additional arguments by Tetrode and Sackur, Bose etc. [1].

2.4 Canonical ensemble

In the argument concerning the equilibrium of two systems in thermal contact, we showed that in equilibrium

$$\frac{\partial \ln \Omega_1(E_1)}{\partial E_1} = \frac{\partial \ln \Omega_2(E_2)}{\partial E_2} = \beta. \quad (2.27)$$

Now we would like to apply this result to understand the probability that a small system in contact with a large heat reservoir have a *specified* state and a *specified* energy E_r . The probability that this state be reached depends on the number of microstates $\Omega'(E'_r)$ available to the reservoir given the energy of the small system E_r :

$$P_r \propto \Omega'(E'_r) = \Omega'(E^{(0)} - E_r), \quad (2.28)$$

where $E'_r = E^{(0)} - E_r$ is the energy of the reservoir. By expanding $\ln \Omega'(E'_r)$ around $E'_r = 0$ we find

$$\begin{aligned} \ln \Omega'(E'_r \ll E^{(0)}) &\approx \ln \Omega'(E^{(0)}) + \frac{\partial \ln \Omega'}{\partial E'}(E_r - E^{(0)}) + \dots \\ &= \ln \Omega'(E^{(0)}) - \beta(E_r - E^{(0)}) + \dots \end{aligned} \quad (2.29)$$

where we used Eq. (1.7). This means that

$$\Omega'(E'_r \ll E^{(0)}) = \Omega'(E^{(0)}) \exp\left(-\beta(E_r - E^{(0)})\right) \quad (2.30)$$

or

$$P_r \propto \exp(-\beta E_r). \quad (2.31)$$

The probability that the system be in a state of energy E_r is given by the normalized canonical distribution

$$P_r = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}, \quad (2.32)$$

where $\beta = 1/(k_B T)$ and the sum is over all states. If the sum goes over all energy levels, then an additional factor accounting for their degeneracy must be included.

The thermodynamics of the system is determined by the partition function

$$Q_N(V, T) = \sum_r \exp(-E_r/(k_B T)). \quad (2.33)$$

The internal energy U , the Helmholtz free energy A , and the pressure P are all obtained from $Q_N(V, T)$, and entropy is given by

$$S_N(V, T) = -k_B \sum_r P_r \ln P_r \quad (2.34)$$

which follows from the definition of the free energy $A = U - TS$. This formula applies equally well to the microcanonical ensemble where all of the Ω states are equally likely to occur so that their probability is $1/\Omega$. This gives

$$S_N(V, T) = -k_B \sum_{r=1}^{\Omega} \frac{1}{\Omega} \ln \frac{1}{\Omega} = k_B \ln \Omega. \quad (2.35)$$

2.5 Ideal gas

The partition function for an ideal gas is

$$Q_N(V, T) = \frac{1}{N! h^{3N}} \int \exp(-\beta H(q_i, p_i)) d\omega, \quad (2.36)$$

where

$$H(q_i, p_i) = \sum_{i=1}^N \frac{p_i^2}{2m} \quad (2.37)$$

and $d\omega$ is the volume element of the phase space. The result is

$$Q_N(V, T) = \frac{1}{N!} \left[\frac{V}{h^3} (2\pi m k_B T)^{3/2} \right]^N = \frac{1}{N!} [Q_1(V, T)]^N, \quad (2.38)$$

showing that it can be seen as the product of N partition functions of noninteracting single particles and the Gibbs correction factor $1/N!$.

When studying the entropy of the microcanonical ensemble [Eq. (1.29)], we approximated the number of states at a given energy $\Omega(N, V, E)$ by the number of *all* states $\Sigma(N, V, E)$ of energy smaller or equal to E . Now we can reuse this result to calculate the density of states, that is, the number of states within a narrow energy interval of width dE . To this end, we note that

$$\Sigma(N, V, E) = \frac{V^N}{N! h^{3N}} \frac{(2\pi m E)^{3N/2}}{(3N/2)!}, \quad (2.39)$$

where the Gibbs factor has been included. Then the density of states is

$$g(E) = \frac{\partial \Sigma(N, V, E)}{\partial E} = \frac{V^N}{N!} \left(\frac{2\pi m}{h^2} \right)^{3N/2} \frac{E^{3N/2-1}}{(3N/2-1)!}, \quad (2.40)$$

2.6 Canonical vs. microcanonical ensemble

The internal energy is given by

$$U = \langle E \rangle = \frac{\sum_r E_r \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}. \quad (2.41)$$

and its derivative with respect to β is

$$\frac{\partial U}{\partial \beta} = -\frac{\sum_r E_r^2 \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} + \left[\frac{\sum_r E_r^2 \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} \right]^2. \quad (2.42)$$

which is by definition equal to

$$-\langle E^2 \rangle + \langle E \rangle^2 \quad (2.43)$$

so that

$$(\Delta E)^2 = \langle E^2 \rangle - \langle E \rangle^2 = -\frac{\partial U}{\partial \beta} = k_B T \frac{\partial U}{\partial T} = k_B T^2 C_V \quad (2.44)$$

where C_V is the heat capacity of the system. The ratio of magnitude of energy fluctuations and mean energy, $\Delta E / \langle E \rangle$, is thus given by $\sqrt{k_B T^2 C_V / U} \propto N^{1/2} / N = N^{-1/2}$ and goes to 0 in the thermodynamic limit. Thus the energy of the canonical system fluctuates very little, and in practice the canonical ensemble is not very different from the microcanonical ensemble.

The probability that states in a given energy interval of width dE be occupied is given by the product of the Boltzmann factor and the density of states:

$$P(E)dE \propto \exp(-\beta E)g(E)dE. \quad (2.45)$$

We now ask what is the most probable energy E^* . On differentiating Eq. (2.45), we have

$$\left. \frac{\partial[\exp(-\beta E)g(E)]}{\partial E} \right|_{E=E^*} = 0 \quad (2.46)$$

or

$$\left[-\beta \exp(-\beta E)g(E) + \exp(-\beta E) \frac{\partial g(E)}{\partial E} \right]_{E=E^*} = 0. \quad (2.47)$$

On dividing by $\exp(-\beta E)$ we find that

$$\left. \frac{d \ln g(E)}{dE} \right|_{E=E^*} = \beta. \quad (2.48)$$

But since $S = k_B \ln g$ and $(\partial S/\partial E)_{E=U} = 1/T$ as implied by the first law, we conclude that $E^* = U$. The most probable energy of the system is equal to its mean energy.

Now we expand $\ln P(E)$ around $E = E^*(= U)$; the motivation for this is that $P(E^*) = \exp(-\beta E^*)g(E^*) = \exp(-\beta U) \exp(S/k_B)$ and so $\ln P(E^*)$ has a clear thermodynamic meaning: $-\beta(U - TS)$. On recalling that the magnitude of energy fluctuations is given Eq. (2.44), this leads to

$$\begin{aligned} \ln(\exp(-\beta E)g(E)) &= \ln(\exp(-\beta E^*)g(E^*)) \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial E^2} \ln(\exp(-\beta E)g(E)) \Big|_{E=E^*} (E - E^*)^2 + \dots \\ &= -\beta(U - TS) - \frac{1}{2(\Delta E)^2} (E - U)^2 + \dots \\ &= -\beta(U - TS) - \frac{1}{2k_B T^2 C_V} (E - U)^2 + \dots \end{aligned} \quad (2.49)$$

so that the probability $P(E)$ is essentially Gaussian:

$$P(E) \propto \exp(-\beta E)g(E) \approx \exp(-\beta(U - TS)) \exp\left(-\frac{(E - U)^2}{2k_B T^2 C_V}\right). \quad (2.50)$$

As stated above, the width of the distribution is much smaller than the typical energy scale so that the Gaussian is extremely narrow (Fig. 2.3).

This result can now be used to reiterate the connection between the Helmholtz free energy and the partition function. To this end, note that the latter can be written as

$$\begin{aligned} Q_N(V, T) &\approx \int_0^\infty \exp(-\beta(U - TS)) \exp\left(-\frac{(E - U)^2}{2k_B T^2 C_V}\right) dE \\ &= \exp(-\beta(U - TS)) \int_0^\infty \exp\left(-\frac{(E - U)^2}{2k_B T^2 C_V}\right) dE \\ &= \exp(-\beta(U - TS)) \frac{1}{2} \sqrt{2k_B T^2 C_V} \underbrace{\int_{-\infty}^\infty \exp(-x^2) dx}_{=\sqrt{\pi}} \\ &= \exp(-\beta(U - TS)) \frac{\sqrt{\pi}}{2} \sqrt{2k_B T^2 C_V}. \end{aligned} \quad (2.51)$$

Now

$$A = -k_B T \ln Q_N(V, T) = U - TS - \frac{1}{2} \ln(2\pi k_B T C_V). \quad (2.52)$$

Both U and S are proportional to N but the third term scales as $\ln N$ and is thus negligible.

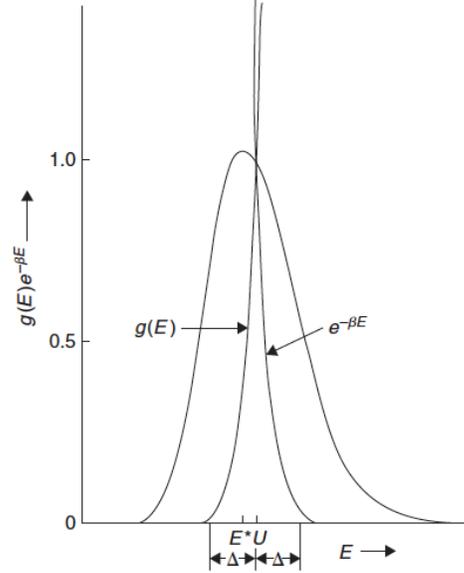


Figure 2.3: Probability that the energy of a canonical ensemble is E resolved into the Boltzmann factor $\exp(-\beta E)$ and the density of states $g(E)$ [1].

2.7 Equipartition and virial theorems

Consider the expectation value of a quantity $x_i(\partial H/\partial x_j)$ where x_i and x_j are either a generalized coordinate or a generalized momentum. The expectation value is defined by

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \frac{\int x_i (\partial H / \partial x_j) \exp(-\beta H) d\omega}{\int \exp(-\beta H) d\omega}. \quad (2.53)$$

Now integrate the numerator by parts over x_j , where $u = x_i$ and $dv = -\beta^{-1} \partial H / \partial x_j \exp(-\beta H)$ so that $v = -\beta^{-1} \exp(-\beta H)$:

$$\begin{aligned} & \int x_i (\partial H / \partial x_j) \exp(-\beta H) dx_j \\ &= -\beta^{-1} x_i \exp(-\beta H) \Big|_{(x_j)_1}^{(x_j)_2} + \beta^{-1} \int \underbrace{\frac{\partial x_i}{\partial x_j}}_{=\delta_{ij}} \exp(-\beta H) dx_j. \end{aligned} \quad (2.54)$$

The upper and the lower boundaries of the integral are such that the first term vanishes because the Boltzmann factor goes to zero (e.g., because the

kinetic energy diverges if $p \rightarrow \infty$ or the potential energy diverges because the coordinate of the particle is within the wall). Thus

$$\left\langle x_i \frac{\partial H}{\partial x_j} \right\rangle = \delta_{ij} k_B T. \quad (2.55)$$

Now if the Hamiltonian is a quadratic function

$$H = \sum_j A_j Q_j^2 + \sum_j B_j P_j^2 \quad (2.56)$$

then $\partial H / \partial Q_j = 2A_j Q_j$ and $\partial H / \partial P_j = 2B_j P_j$ so that

$$\sum_j \left(Q_j \frac{\partial H}{\partial Q_j} + P_j \frac{\partial H}{\partial P_j} \right) = 2H. \quad (2.57)$$

On the other hand,

$$\left\langle \sum_j \left(Q_j \frac{\partial H}{\partial Q_j} + P_j \frac{\partial H}{\partial P_j} \right) \right\rangle = f k_B T \quad (2.58)$$

where f is the number of the quadratic degrees of freedom. But the left-hand side is also equal to $\langle 2H \rangle$ and we conclude that

$$\langle 2H \rangle = \frac{f}{2} k_B T. \quad (2.59)$$

This is the equipartition theorem.

The virial theorem refers to the sum of products of particle positions and forces on them:

$$\sum_i q_i \dot{p}_i, \quad (2.60)$$

which is known as the virial \mathcal{V} ; the sum is over all particles. Given that Newton's law states that $\dot{p}_i = -\partial H / \partial q_i$, the virial can be written as

$$-\sum_i q_i \frac{\partial H}{\partial q_i}, \quad (2.61)$$

which is, upon averaging, yet another case of the general relation Eq. (2.55). Since there are $3N$ degrees of freedom we have

$$\langle \mathcal{V} \rangle = -3N k_B T, \quad (2.62)$$

To illustrate the virial theorem, we consider first the ideal gas. Here the forces on particles are due to pressure integrated over the wall of the vessel:

$$\mathcal{V}_0 = \sum_i q_i \dot{p}_i = -P \oint \mathbf{r} \cdot d\mathbf{S} = -P \int \nabla \cdot \mathbf{r} dV, \quad (2.63)$$

where we have used the Gauss theorem; the minus sign is because the force due to pressure points inward. Since $\nabla \cdot \mathbf{r} = 3$ we have

$$\mathcal{V}_0 = -3PV, \quad (2.64)$$

which gives $PV = Nk_B T$. — More illuminating applications, of course, involve interacting particles where the virial theorem leads to

$$\frac{PV}{Nk_B T} = 1 + \frac{1}{3Nk_B T} \left\langle \sum_{i<j} \mathbf{F}(\mathbf{r}_{ij}) \cdot \mathbf{r}_{ij} \right\rangle = 1 - \frac{1}{3Nk_B T} \left\langle \sum_{i<j} \frac{\partial u(r_{ij})}{\partial r_{ij}} r_{ij} \right\rangle, \quad (2.65)$$

where $\mathbf{F}(\mathbf{r}_{ij}) = -\nabla u(r_{ij})$ is the force between particles i and j and u is the pair potential.

2.8 Harmonic oscillator

So far we have developed considerable experience in comparing microcanonical and canonical ensembles. The harmonic oscillator is a very important model used to describe vibrations in molecules and solids (Debye model) and excitations of the electromagnetic field (blackbody radiation). Here the starting point is the spectrum of an oscillator

$$\epsilon_n = \left(n + \frac{1}{2} \right) \hbar\omega, \quad (2.66)$$

where $n = 0, 1, 2, \dots$. The partition function for N distinguishable oscillators is

$$Q_N(\beta) = Q_1^N(\beta) = \left[\sum_{n=0}^{\infty} \exp \left(-\beta \left(n + \frac{1}{2} \right) \hbar\omega \right) \right]^N = [2 \sinh(\beta\hbar\omega/2)]^{-N}. \quad (2.67)$$

The internal energy, entropy, specific heat, and Helmholtz free energy can all be derived in a straightforward way. Here we will merely sketch a microcanonical analysis. — Consider N independent distinguishable oscillators

with the same frequency, their combined energy being E . The energy can be written as

$$E = \left(R + \frac{N}{2} \right) \hbar\omega \quad (2.68)$$

so that each oscillator has at least the zero-point energy $\hbar\omega/2$ as it must have. The part of energy that can be distributed between the oscillators is $R\hbar\omega$. This combinatorial problem is best presented by examining a row of R balls to be partitioned into N boxes. The balls are all indistinguishable and thus their identity is inaccessible; all that we can vary is the position of $N - 1$ walls between balls. There are

$$\frac{(R + N - 1)!}{R!(N - 1)!}. \quad (2.69)$$

The entropy of these different arrangements is

$$S = k_B [\ln(R + N)! - \ln R! - \ln N!] \quad (2.70)$$

(we have neglected "−1" as R and N are both very large). Again, on using the Stirling formula,

$$S = k_B [(N + R) \ln(R + N) - R \ln R! - N \ln N]. \quad (2.71)$$

Temperature is given by

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial E} \right)_N = \left(\frac{\partial S}{\partial R} \right)_N \underbrace{\frac{dR}{dE}}_{=1/(\hbar\omega)} = \frac{k_B}{\hbar\omega} \ln \left(\frac{R + N}{R} \right) \\ &= \frac{k_B}{\hbar\omega} \ln \left(\frac{E + N\hbar\omega/2}{E - N\hbar\omega/2} \right). \end{aligned} \quad (2.72)$$

This result can be inverted so as to express E :

$$\frac{E}{N} = \frac{\hbar\omega \exp(\hbar\omega/(k_B T)) + 1}{2 \exp(\hbar\omega/(k_B T)) - 1}. \quad (2.73)$$

This result is identical to the internal energy obtained within the canonical formalism:

$$U = N \left[\frac{\hbar\omega}{2} + \frac{\hbar\omega}{\exp(\hbar\omega/(k_B T)) - 1} \right]. \quad (2.74)$$

One may also evaluate the $T \rightarrow \infty$ limit of Eq. (2.73) by Taylor-expanding the exponential in the denominator; the result is $E = Nk_B T$.

2.9 Paramagnetism

This system too was already studied in some detail including the Curie law. Some of its aspects are also included in the two-level system. Among the leftover issues, the most important is the specific heat. Here we will study it using a shorthand notation suitable for $j = 1/2$ case. Here one only has two energy levels, the energies being $\pm p_B B = \pm \epsilon$, where $p_B = e\hbar/(2m_e)$ is Bohr magneton and B ; m_e is the mass of electron. The partition function is

$$Q_N(\beta) = [\exp(\beta\epsilon) + \exp(-\beta\epsilon)]^N = [2 \cosh(\beta\epsilon)]^N. \quad (2.75)$$

The Helmholtz free energy is

$$A = -Nk_B T \ln(2 \cosh(\beta\epsilon)). \quad (2.76)$$

The internal energy can be spelled out simply by weighing the two energies by the Boltzmann factor and is

$$U = -N\epsilon \tanh\left(\frac{\epsilon}{k_B T}\right) \quad (2.77)$$

and the heat capacity is

$$C = \left(\frac{\partial U}{\partial T}\right)_B = N \left(\frac{\epsilon}{k_B T}\right)^2 \frac{1}{\cosh^2(\epsilon/(k_B T))}. \quad (2.78)$$

This interesting result is plotted in Fig. 2.4. At low temperatures, the heat capacity increases exponentially:

$$C(T \rightarrow 0) \approx 4N \left(\frac{\epsilon}{k_B T}\right)^2 \exp\left(-\frac{2\epsilon}{k_B T}\right) \quad (2.79)$$

whereas at high temperatures it decreases as a power law:

$$C(T \rightarrow \infty) \approx N \left(\frac{\epsilon}{k_B T}\right)^2. \quad (2.80)$$

This implies that C must have a peak at some finite temperature; such a peak is known as the Schottky anomaly and is typical for systems with a gap in the spectrum.

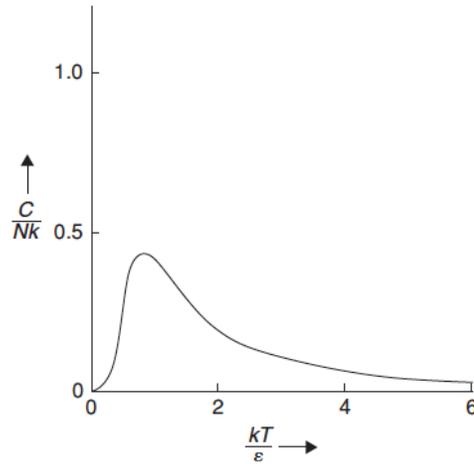


Figure 2.4: Specific heat of a two-level system [1].

2.10 Grand-canonical ensemble

The canonical statistical-mechanical formalism can be generalized so as to allow the exchange of particles with a particle reservoir. This is done completely analogously as in the case of canonical ensemble, and the resulting probability that the system be in a microstate with N_r particles and an energy of E_s is exponential:

$$P_{s,r} \propto \exp(\beta\mu N_r - \beta E_s). \quad (2.81)$$

Here μ is the chemical potential. A properly normalized $P_{s,r}$ is

$$P_{s,r} = \frac{\exp(\beta\mu N_r - \beta E_s)}{\sum_{r,s} \exp(\beta\mu N_r - \beta E_s)}. \quad (2.82)$$

In the grand-canonical formalism, the partition function $Q_N(V, T)$ is replaced by the grand partition function

$$\mathcal{Q}(z, V, T) = \sum_{N_r=0}^{\infty} z^{N_r} Q_{N_r}(V, T), \quad (2.83)$$

where

$$z = \exp(\beta\mu) \quad (2.84)$$

is known as the *fugacity*.

Like the mean energy in the canonical ensemble, the mean number of particles in the grand canonical ensemble is very well defined. To see this, recall that

$$\langle N \rangle = \frac{\sum_{r,s} N_r \exp(\beta\mu N_r - \beta E_S)}{\sum_{r,s} \exp(\beta\mu N_r - \beta E_S)}. \quad (2.85)$$

Like in Eq. (2.42), we have

$$\begin{aligned} \left(\frac{\partial \langle N \rangle}{\partial(\beta\mu)} \right)_{\beta, E_S} &= \frac{\sum_{r,s} N_r^2 \exp(\beta\mu N_r - \beta E_S)}{\sum_{r,s} \exp(\beta\mu N_r - \beta E_S)} \\ &\quad - \left[\frac{\sum_{r,s} N_r \exp(\beta\mu N_r - \beta E_S)}{\sum_{r,s} \exp(\beta\mu N_r - \beta E_S)} \right]^2 \\ &= \langle N^2 \rangle - \langle N \rangle^2 \end{aligned} \quad (2.86)$$

so that the magnitude of fluctuations is given by

$$\langle (\Delta N)^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2 = k_B T \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V} \quad (2.87)$$

This result can be used to calculate the fluctuations of particle density $n = N/V$. On dividing by $\langle N \rangle^2$, we find that

$$\frac{\langle (\Delta N)^2 \rangle}{\langle N \rangle^2} = \frac{\langle (\Delta n)^2 \rangle}{\langle n \rangle^2} = \frac{k_B T}{\langle N \rangle^2} \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V}. \quad (2.88)$$

This can be rewritten using the specific volume $v = V/\langle N \rangle$

$$\frac{\langle (\Delta n)^2 \rangle}{\langle n \rangle^2} = \frac{k_B T v^2}{V^2} \left(\frac{\partial(V/v)}{\partial \mu} \right)_{T,V}. \quad (2.89)$$

Since the derivative is to be taken at constant V , this is equal to

$$\frac{\langle (\Delta n)^2 \rangle}{\langle n \rangle^2} = \frac{k_B T v^2}{V} \left(\frac{\partial(1/v)}{\partial \mu} \right)_{T,V} = -\frac{k_B T}{V} \left(\frac{\partial v}{\partial \mu} \right)_T. \quad (2.90)$$

The differential of the chemical potential reads $d\mu = -sdT + vdP$ so that at $T = \text{const.}$, $d\mu = vdP$ and

$$\left(\frac{\partial v}{\partial \mu} \right)_{T,V} = \left(\frac{\partial v}{\partial P} \right)_T \underbrace{\left(\frac{\partial P}{\partial \mu} \right)_T}_{=1/v}. \quad (2.91)$$

Since $-(1/v) (\partial P / \partial \mu)_T = \chi_T$, the square of relative density fluctuations are proportional to the isothermal compressibility of the system:

$$\frac{\langle (\Delta n)^2 \rangle}{\langle n \rangle^2} = \frac{k_B T}{V} \chi_T. \quad (2.92)$$

The root of the right-hand side of this expression, that is the magnitude of relative fluctuations, is inversely proportional to $V^{1/2} \propto N^{1/2}$. Thus the magnitude of fluctuations is usually small but in cases where compressibility is large it does increase. This happens at phase transitions.

Chapter 3

Quantum statistics

3.1 Density matrix

Consider an ensemble of $\mathcal{N} \gg 1$ identical systems characterized by a Hamiltonian \hat{H} . Denote the physical state of k th system by $\psi^k(\mathbf{r}_i, t)$. The evolution of this state is governed by the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi. \quad (3.1)$$

The wavefunction may be expanded in terms of an orthonormal set ϕ_n such that

$$\psi^k(t) = \sum_n a_n^k(t) \phi_n, \quad (3.2)$$

where the coefficients

$$a_n^k(t) = \int \phi_n^* \psi^k(t) d\tau. \quad (3.3)$$

only depend on time and ϕ_n only depend on the spatial coordinates; $d\tau$ is the volume element. Equation of motion for a_n^k reads

$$\begin{aligned} i\hbar \frac{\partial a_n^k}{\partial t} &= i\hbar \frac{\partial}{\partial t} \int \phi_n^* \psi^k(t) d\tau = \int \phi_n^* \hat{H} \psi^k(t) d\tau \\ &= \int \phi_n^* \hat{H} \left[\sum_m a_m^k(t) \phi_m \right] d\tau = \sum_m H_{nm} a_m^k(t), \end{aligned} \quad (3.4)$$

where

$$H_{nm} = \int \phi_n^* \hat{H} \phi_m d\tau. \quad (3.5)$$

is the matrix element.

The meaning of $a_n^k(t)$ is as follows: $|a_n^k(t)|^2$ is the probability that the k th systems be in state n at a time t . Obviously

$$\sum_n |a_n^k(t)|^2 = 1. \quad (3.6)$$

Density operator

The density operator $\hat{\rho}(t)$ of the statistical ensemble is defined by the average of $a_m^k(t)a_n^{k*}(t)$ over all systems:

$$\rho_{mn}(t) = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} a_m^k(t)a_n^{k*}(t). \quad (3.7)$$

(This is referred to as the density matrix.) The diagonal elements of the density operator $\rho_{nn}(t) = |a_n(t)|^2$ are the ensemble averages of the probability that a randomly chosen system (rather than a given system k) be in state ϕ_n . Obviously

$$\sum_{n=1} \rho_{nn} = 1. \quad (3.8)$$

Let us derive the equation of motion for the density matrix:

$$i\hbar \frac{\partial \rho_{mn}(t)}{\partial t} = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} i\hbar \left[\frac{\partial a_m^k(t)}{\partial t} a_n^{k*}(t) + a_m^k(t) \frac{\partial a_n^{k*}(t)}{\partial t} \right]. \quad (3.9)$$

Using Eq. (3.4), we have

$$\begin{aligned} i\hbar \frac{\partial \rho_{mn}(t)}{\partial t} &= \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \left\{ \left[\sum_l H_{ml} a_l^k(t) \right] a_n^{k*}(t) - a_m^k(t) \left[\sum_l \underbrace{H_{nl}^*}_{=H_{ln}} a_l^{k*}(t) \right] \right\} \\ &= \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \left\{ \left[\sum_l H_{ml} a_l^k(t) \right] a_n^{k*}(t) - a_m^k(t) \left[\sum_l H_{ln} a_l^{k*}(t) \right] \right\} \\ &= \sum_l \left[H_{ml} \underbrace{\frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} a_l^k(t) a_n^{k*}(t)}_{=\rho_{ln}} - \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} a_m^k(t) a_l^{k*}(t) H_{ln} \right] \\ &= \left(\hat{H} \hat{\rho} - \hat{\rho} \hat{H} \right)_{mn}. \end{aligned} \quad (3.10)$$

In other words,

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = \left[\hat{H}, \hat{\rho} \right] \quad (3.11)$$

so that if $\hat{\rho}$ commutes with \hat{H} , $\partial\hat{\rho}/\partial t = 0$. This is the case $\hat{\rho}$ is a function of \hat{H} so that it can be expanded in power series and then it is evident that it does commute with \hat{H} . Assuming that \hat{H} itself does not depend on time, such a $\hat{\rho}$ then means that the density matrix ρ_{mn} does not depend on time either and the system is stationary.

The form of ρ_{mn} depends on the choice of basis functions. If ϕ_n are the eigenfunctions of \hat{H} , $\hat{H}\phi_n = E_n\phi_n$ then H_{mn} [Eq. (3.5)] is diagonal:

$$H_{nm} = E_n\delta_{mn}. \quad (3.12)$$

This is referred to as the energy representation. In this case,

$$i\hbar\frac{\partial a_n^k}{\partial t} = E_n a_n^k(t) \quad (3.13)$$

and

$$a_n^k(t) = \exp(-iE_n t/\hbar + i\delta_n). \quad (3.14)$$

Thus the coefficients $a_m^k(t)$ and $a_N^{k*}(t)$ in Eq. (3.7) generally contain random phases unless $m = n$ so that all off-diagonal elements of the density matrix vanish:

$$\rho_{mn} = \rho_n\delta_{mn}. \quad (3.15)$$

In this representation, the density operator reads

$$\hat{\rho} = \sum_n |\phi_n\rangle\rho_n\langle\phi_n|. \quad (3.16)$$

Irrespective of representation, ρ_{mn} must be symmetric because of detailed balance. In equilibrium, the probability for a transition from a given initial state to a given final state must be the same as the probability for the reverse transition.

Expectation values

The expectation value of a quantity G represented by the operator \hat{G} is given by

$$\langle G \rangle = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \int \psi^{k*} \hat{G} \psi^k d\tau. \quad (3.17)$$

On inserting Eq. (3.2), we find

$$\langle G \rangle = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} \left[\underbrace{\sum_{m,n} a_m^{k*} a_n^k}_{=\rho_{mn}} G_{nm} \right] \quad (3.18)$$

with

$$G_{nm} \int \phi_n^* \hat{G} \phi_m d\tau. \quad (3.19)$$

We recognize that this form of $\langle G \rangle$ contains $\hat{\rho}$ so that

$$\langle G \rangle = \sum_{m,n} \rho_{mn} G_{nm} = \sum_m \left(\hat{\rho} \hat{G} \right)_{mm} = \text{tr} \left(\hat{\rho} \hat{G} \right). \quad (3.20)$$

This result also shows that $\text{tr} \hat{\rho} = 1$ as it should; this is seen by choosing $\hat{G} = 1$.

3.2 Microcanonical and canonical ensemble

In the energy representation of the microcanonical ensemble,

$$\rho_n = 1/\Gamma \quad (3.21)$$

is the same for all states such that the sum of their energies lies within a narrow interval of energies from $E - \Delta/2$ to $E + \Delta/2$. Here Γ is the number of accessible states. Entropy reads

$$S = k_B \ln \Gamma. \quad (3.22)$$

We expect that in some other representation, that is, in a representation based on wavefunctions that are not eigenfunctions of the Hamiltonian, the overall form of the density matrix should be the same as in the energy representation in the sense that the off-diagonal terms must be zero and the diagonal terms are identical. This is ensured by assuming that the probability amplitudes a_n^k have random phases so that the wavefunction ψ^k is an incoherent superposition of all ϕ_n^k . For $a_n^k = a \exp(i\theta_n^k)$ Eq. (3.7) reduces to

$$\begin{aligned} \rho_{mn}(t) &= \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} a_m^k(t) a_n^{k*}(t) = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} |a|^2 \exp(i(\theta_m^k - \theta_n^k)) \\ &= c \langle \exp(i(\theta_m^k - \theta_n^k)) \rangle = c \delta_{mn}. \end{aligned} \quad (3.23)$$

In the canonical ensemble

$$\rho_n = C \exp(-\beta E_n), \quad (3.24)$$

where C is the normalization constant given by the condition that $Q_N(\beta) = \sum_n \rho_n = 1$; $q_N(\beta)$ is the partition function.

The density operator is then given by

$$\hat{\rho} = \sum_n |\phi_n\rangle \frac{\exp(-\beta E_n)}{Q_N(\beta)} \langle \phi_n|. \quad (3.25)$$

Now note that $|\phi_n\rangle \exp(-\beta E_n) = \exp(-\beta E_n)|\phi_n\rangle = \exp(-\beta \hat{H})|\phi_n\rangle$; this is evident if one sees $\exp(-\beta \hat{H})$ as a Taylor series. The sum now only applies to $|\phi_n\rangle \langle \phi_n|$:

$$\hat{\rho} = \frac{\exp(-\beta \hat{H})}{Q_N(\beta)} \underbrace{\sum_n |\phi_n\rangle \langle \phi_n|}_{=1} = \frac{\exp(-\beta \hat{H})}{\text{tr} \exp(-\beta \hat{H})}. \quad (3.26)$$

The expectation value of a quantity G is

$$\langle G \rangle = \text{tr} (\hat{\rho} \hat{G}) = \frac{1}{Q_N(\beta)} \text{tr} (\hat{G} \exp(-\beta \hat{H})) = \frac{\text{tr} (\hat{G} \exp(-\beta \hat{H}))}{\text{tr} \exp(-\beta \hat{H})}. \quad (3.27)$$

3.3 Particle in a box

Let us now use this apparatus to study a very familiar and transparent example of a single particle in a box of side L . The Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m} \nabla^2 \quad (3.28)$$

and we choose periodic boundary conditions so as to simplify the analysis:

$$\psi(x, y, z) = \psi(x + L, y, z) \quad (3.29)$$

etc. The allowed wavevectors are

$$\mathbf{k} = \frac{2\pi}{L} (n_x, n_y, n_z), \quad (3.30)$$

where n_x, n_y , and n_z are integers $0, \pm 1, \pm 2 \dots$ and the corresponding wavefunctions are plane waves:

$$\phi_E(\mathbf{r}) \equiv |\mathbf{k}\rangle = \frac{1}{L^{3/2}} \exp(i\mathbf{k} \cdot \mathbf{r}) \quad (3.31)$$

where $E = \hbar^2 k^2 / (2m)$.

We now calculate the density matrix in the coordinate representation where the basis functions are delta functions of position:

$$|\mathbf{r}_1\rangle = \delta(\mathbf{r} - \mathbf{r}_1). \quad (3.32)$$

Up to a normalization factor, the density matrix is given by

$$\langle \mathbf{r}_1 | \exp(-\beta \hat{H}) | \mathbf{r}_2 \rangle = \langle \mathbf{r}_1 | \left(\sum_{\mathbf{k}} |\mathbf{k}\rangle \exp(-\beta E_{\mathbf{k}}) \langle \mathbf{k}| \right) | \mathbf{r}_2 \rangle. \quad (3.33)$$

Since

$$\langle \mathbf{r}_1 | \mathbf{k} \rangle = \frac{1}{L^{3/2}} \int \delta(\mathbf{r} - \mathbf{r}_1) \exp(i\mathbf{k} \cdot \mathbf{r}) \, d\mathbf{r} = \frac{1}{L^{3/2}} \exp(i\mathbf{k} \cdot \mathbf{r}_1), \quad (3.34)$$

$$\begin{aligned} \langle \mathbf{r}_1 | \exp(-\beta \hat{H}) | \mathbf{r}_2 \rangle &= \frac{1}{L^3} \sum_{\mathbf{k}} \exp\left(-\beta \frac{\hbar^2 k^2}{2m}\right) \exp(i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)) \\ &= \frac{1}{(2\pi)^3} \int \exp\left(-\beta \frac{\hbar^2 k^2}{2m} + i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)\right) \, d\mathbf{k} \end{aligned} \quad (3.35)$$

To calculate this integral, we rotate the coordinate system such that k_x points along $\mathbf{r}_1 - \mathbf{r}_2$; the imaginary term is of type $\int \exp(-k^2) \sin(kx) \, dk$ and vanishes because of symmetry. The right-hand side reads

$$\begin{aligned} &\frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \exp\left(-\beta \frac{\hbar^2 k_x^2}{2m}\right) \cos(k_x |\mathbf{r}_1 - \mathbf{r}_2|) \, dk_x \\ &\times \int_{-\infty}^{\infty} \exp\left(-\beta \frac{\hbar^2 k_y^2}{2m}\right) \, dk_y \times \int_{-\infty}^{\infty} \exp\left(-\beta \frac{\hbar^2 k_z^2}{2m}\right) \, dk_z \\ &= \frac{1}{(2\pi)^3} \sqrt{\frac{2\pi m}{\beta \hbar^2}} \exp\left(-\frac{m}{2\beta \hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2\right) \times \sqrt{\frac{2\pi m}{\beta \hbar^2}} \times \sqrt{\frac{2\pi m}{\beta \hbar^2}} \\ &= \left(\frac{m}{2\pi \beta \hbar^2}\right)^{3/2} \exp\left(-\frac{m}{2\beta \hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2\right). \end{aligned} \quad (3.36)$$

To normalize the density matrix, we also need to evaluate $\text{tr} \exp(-\beta \hat{H})$: From Eq. (3.36) it follows that

$$\text{tr} \exp(-\beta \hat{H}) = \int \langle \mathbf{r} | \exp(-\beta \hat{H}) | \mathbf{r} \rangle \, d\mathbf{r} = \left(\frac{m}{2\pi \beta \hbar^2}\right)^{3/2} V \quad (3.37)$$

so that finally

$$\langle \mathbf{r}_1 | \hat{\rho} | \mathbf{r}_2 \rangle = \frac{1}{V} \exp \left(-\frac{m}{2\beta\hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2 \right). \quad (3.38)$$

This result is symmetric with respect to $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$. If $\mathbf{r}_1 = \mathbf{r}_2$, the element of the density matrix represents the probability density that a particle be located at a given point \mathbf{r}_1 , and this probability density is $1/V$ independent of position so that it gives 1 upon integration over \mathbf{r} .

On the other hand, if \mathbf{r}_1 and \mathbf{r}_2 are different then the off-diagonal element of the density matrix is the probability density that a particle at \mathbf{r}_1 spontaneously moves to \mathbf{r}_2 . In other words, if the particle is to be viewed as a wave packet centered at \mathbf{r}_1 , $\langle \mathbf{r}_1 | \hat{\rho} | \mathbf{r}_2 \rangle$ can be used to estimate the size of the packet, which is of the order of $\hbar/\sqrt{mk_B T}$. (This is also the thermal wavelength of the packet.)

Let us use the density matrix to calculate the expectation value of energy:

$$\langle H \rangle = \text{tr} \left(\hat{H} \hat{\rho} \right) = \frac{1}{V} \int \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \exp \left(-\frac{m}{2\beta\hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2 \right)_{\mathbf{r}_2=\mathbf{r}_1} d\mathbf{r}_1. \quad (3.39)$$

In taking the second derivatives, notice that $(\partial/\partial x)^2 \exp(-\alpha x^2) = -2\alpha(1 - 2\alpha x^2) \exp(-\alpha x^2)$ so that

$$\begin{aligned} & \nabla^2 \exp \left(-\frac{m}{2\beta\hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2 \right) \\ &= -\frac{m}{\beta\hbar^2} \left[3 - \frac{m}{\beta\hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2 \right] \exp \left(-\frac{m}{2\beta\hbar^2} |\mathbf{r}_1 - \mathbf{r}_2|^2 \right), \end{aligned} \quad (3.40)$$

where we have re-collected the squares of all three relative coordinates. For $\mathbf{r}_1 = \mathbf{r}_2$, this reduces to a very simple result

$$-\frac{3m}{\beta\hbar^2} \quad (3.41)$$

so that finally

$$\langle H \rangle = \frac{1}{V} \frac{\hbar^2}{2m} \frac{3m}{\beta\hbar^2} \times \underbrace{\int d\mathbf{r}}_{=V} = \frac{3k_B T}{2}. \quad (3.42)$$

This is a well-known result.

3.4 Systems of indistinguishable particles

In systems of particles we can construct the many-body wavefunction as a product of single-particle wavefunctions:

$$\psi(\mathbf{q}) = \prod_{m=1}^{n_1} u_1(m) \prod_{m=n_1+1}^{n_1+n_2} u_2(m) \dots, \quad (3.43)$$

where \mathbf{q} is the vector of coordinates of all N particles and $u_i(m)$ are shorthands for single-particle wavefunctions of m th particle $u_i(q_m)$. This form is not suitable for the description of systems of indistinguishable particles because here a permutation of two or more particles must not change the physical state of the system. For example, simultaneously replacing $u_1(q_2)$ by $u_1(q_5)$ and $u_2(q_5)$ by $u_2(q_1)$ should not make any difference. Since the physical observable is the probability density $|\psi|^2$, we find that

$$|P\psi|^2 = |\psi|^2, \quad (3.44)$$

where $P\psi$ is the wavefunction obtained from ψ after a permutation. There are two possibilities. The first one is

$$P\psi = \psi \quad (3.45)$$

for all P s, which means that ψ is symmetric in all arguments. Particles obeying this rule are referred to as bosons (particles with integer spin, e.g., photons, phonons, He^4 atoms...) and the corresponding statistics is Bose-Einstein statistics. The second possibility is that

$$P\psi = \begin{cases} \psi, & P \text{ even} \\ -\psi, & P \text{ odd} \end{cases}. \quad (3.46)$$

This applies to fermions (particles with half-integer spin, e.g., electrons, protons, He^3 atoms...) described by Fermi-Dirac statistics. It is also possible that the wavefunction changes by a complex phase factor $\exp(i\theta)$ on permutation. Certain quasiparticles are indeed described by this rule; they are called anyons.

The single-particle density-matrix formalism introduced above can be readily generalized to many-particle systems but we will not pursue it here. As an illuminating result, we only cite the diagonal element of the two-body density matrix for two free particles:

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \hat{\rho} | \mathbf{r}_1, \mathbf{r}_2 \rangle = \frac{1}{V^2} \left[1 \pm \exp\left(-\frac{2\pi r_{12}^2}{\lambda^2}\right) \right], \quad (3.47)$$

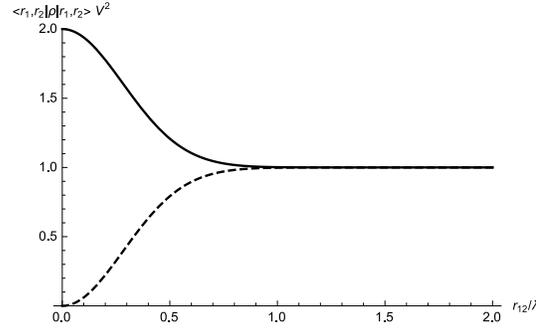


Figure 3.1: Two-body density matrix vs. distance: Fermi-Dirac (solid line) and Bose-Einstein (dashed line).

where r_{12} is the separation of the particles and

$$\lambda = \frac{h}{\sqrt{2\pi mk_B T}} \quad (3.48)$$

is the thermal de Broglie wavelength. In this result, + sign corresponds to bosons and – sign to fermions.

Equation (3.47) shows that the free quantum-mechanical particles interact with each other if the distance between them is smaller than λ . In fermions, the probability that they be at the same position decreases to 0 as $r_{12} \rightarrow 0$ whereas in bosons it increases to 2 (Fig. 3.1). The effect of boson bunching was observed, e.g., in the 1956 interferometric experiment by Hanbury Brown and Twiss where light from Sirius was collected by two photomultiplier tubes separated by 6 m, and positive correlation between signals was seen.

Note that at $r_{12} \gg \lambda$ Eq. (3.47) reduces to the classical result $1/V^2$.

Chapter 4

Ideal systems

4.1 Ideal Bose gas

In bosons, the mean occupation number of a state of energy ε is

$$\langle n_\varepsilon \rangle = \frac{1}{\exp((\varepsilon - \mu)/k_B T) + a}, \quad (4.1)$$

where $a = 0, -1,$ and 1 in classical, Bose-Einstein, and Fermi-Dirac systems, respectively. Clearly the effects of quantum statistics related to the indistinguishability of the particles are not important if the occupation number is small, that is, if

$$\exp((\varepsilon - \mu)/k_B T) \gg 1. \quad (4.2)$$

In this limit, the three types of occupation numbers are essentially the same, and this regime is referred to as *weak degeneracy*. (A degenerate system is a system that cannot be described classically; that is, a system where quantum effects are prominent.) In free particles where ε consists solely of kinetic energy and $\varepsilon > 0$, this regime is reached if $\mu < 0$ and large in magnitude. In turn, $\mu < 0$ means that the fugacity¹ defined by $z = \exp(\mu/k_B T)$ must be much smaller than 1:

$$\exp(\mu/k_B T) \ll 1. \quad (4.3)$$

¹Fugacity is a concept introduced to interpret the Gibbs free energy of real gases and liquids. At constant temperature, $dG = VdP$ and for ideal gas, $V = Nk_B T/P$ so that $dG^{\text{id}} = Nk_B T dP/P$ and $G^{\text{id}} = Nk_B T \ln P$. We can write the Gibbs free energy using the same form if P is replaced by fugacity: $G^{\text{real}} = Nk_B T \ln f$. This can be viewed as the definition of fugacity.

This condition is directly related to particle density. In a classical gas, the grand-canonical partition function reads

$$\mathcal{Q}(z, V, T) = \sum_{n_r=0}^{\infty} z^{N_r} Q_{N_r}(V, T) \quad (4.4)$$

where

$$Q_{N_r}(V, T) = \frac{1}{N_r!} \left[\frac{V}{h^3} \left(\frac{2\pi m}{\beta} \right)^{3/2} \right]^{N_r}. \quad (4.5)$$

Equation (4.4) is a Taylor series of an exponential so that the result is

$$\mathcal{Q}(z, V, T) = \exp \left(z \frac{V}{h^3} \left(\frac{2\pi m}{\beta} \right)^{3/2} \right) = \exp \left(z \frac{V}{\lambda^3} \right) \quad (4.6)$$

where $\lambda = h/\sqrt{2\pi m k_B T}$ is the de Broglie wavelength. Since

$$\langle N \rangle = \left(\frac{\partial \ln \mathcal{Q}}{\partial (\beta \mu)} \right)_{T, V} = \left(\frac{\partial \ln \mathcal{Q}}{\partial z} \right)_{T, V} \underbrace{\frac{\partial z}{\partial (\beta \mu)}}_{=\exp(\beta \mu)=z} = z \left(\frac{\partial \ln \mathcal{Q}}{\partial z} \right)_{T, V} \quad (4.7)$$

we have

$$\langle N \rangle = z \frac{V}{\lambda^3} \quad (4.8)$$

or

$$z = \frac{\langle N \rangle \lambda^3}{V} = n \lambda^3. \quad (4.9)$$

The condition $z \ll 1$ then means that

$$n \lambda^3 \ll 1. \quad (4.10)$$

This result is consistent with the two-particle density matrix [Eq. (3.47)] presented in Fig. 3.1 which shows that at distances smaller than λ quantum effects are important.

Let us estimate the de Broglie wavelength. The mass of the molecule is $m = M/N_A$ which is $(29 \text{ kg/kmol})/(6.02 \times 10^{26} / \text{kmol}) = 4.8 \times 10^{-26} \text{ kg}$ for air so that at $T = 300 \text{ K}$, $\lambda = 1.9 \times 10^{-11} \text{ m}$ and at $T = 3 \text{ K}$, $\lambda = 1.9 \times 10^{-10} \text{ m}$, which is comparable to the size of the molecules; at $T = 0.03 \text{ K}$, $\lambda = 1.9 \times 10^{-9} \text{ m}$. This is the reason why in cold gases quantum effects are prominent even at moderate densities.

4.2 Bose-Einstein condensation

In a Bose gas, the grand-canonical partition function reads

$$\mathcal{Q} = \prod_r \sum_{N_r=0}^{\infty} \exp(\beta N_r (\mu - \varepsilon_r)) = \prod_r \frac{1}{1 - \exp(\beta(\mu - \varepsilon_r))} \quad (4.11)$$

and since $\ln \mathcal{Q} = \beta PV$ we have

$$\frac{PV}{k_B T} = - \sum_r \ln(1 - z \exp(-\beta \varepsilon_r)). \quad (4.12)$$

The number of particles is

$$N = \sum_r \frac{1}{\exp(\beta(\varepsilon_r - \mu)) - 1} = \sum_r \frac{1}{z^{-1} \exp(\beta \varepsilon_r) - 1}. \quad (4.13)$$

The sums on r can be replaced by an integral. Recall that for non-relativistic particles with $\varepsilon = p^2/(2m)$ the single-particle density of states is given by

$$\frac{V}{h^3} 4\pi p^2 dp = \frac{2\pi V (2m)^{3/2}}{h^3} \varepsilon^{1/2} d\varepsilon. \quad (4.14)$$

This step is justified for all states except for the $\mathbf{p} = 0, \varepsilon = 0$ state which would have a density of 0 according to the above formula, and this would be incorrect. Given that there may be any number of particles in this state, this detail does matter; for fermions the $\mathbf{p} = 0, \varepsilon = 0$ state may be safely disregarded as there may be at most 2 particles in this state.

Thus we have

$$\begin{aligned} \frac{P}{k_B T} &= - \int_0^{\infty} \frac{2\pi(2m)^{3/2}}{h^3} \ln(1 - z \exp(-\beta \varepsilon_r)) \varepsilon^{1/2} d\varepsilon - \frac{1}{V} \ln(1 - z) \\ &= - \int_0^{\infty} \frac{2\pi(2mk_B T)^{3/2}}{h^3} \ln(1 - z \exp(-x)) x^{1/2} dx - \frac{1}{V} \ln(1 - z) \\ &= - \frac{2}{\sqrt{\pi} \lambda^3} \int_0^{\infty} \ln(1 - z \exp(-x)) x^{1/2} dx - \frac{1}{V} \ln(1 - z) \end{aligned} \quad (4.15)$$

and

$$\begin{aligned} \frac{N}{V} &= \int_0^{\infty} \frac{2\pi(2m)^{3/2}}{h^3} \frac{1}{z^{-1} \exp(\beta \varepsilon_r) - 1} \varepsilon^{1/2} d\varepsilon + \frac{1}{V} \frac{1}{z^{-1} - 1} \\ &= \int_0^{\infty} \frac{2\pi(2mk_B T)^{3/2}}{h^3} \frac{1}{z^{-1} \exp(x) - 1} x^{1/2} dx + \frac{1}{V} \frac{1}{z^{-1} - 1} \\ &= \frac{2}{\sqrt{\pi} \lambda^3} \int_0^{\infty} \frac{x^{1/2} dx}{z^{-1} \exp(x) - 1} + \frac{1}{V} \underbrace{\frac{z}{1 - z}}_{=N_0}. \end{aligned} \quad (4.16)$$

In the last equation, we emphasized that the number of particles in the ground state is given by

$$N_0 = \frac{z}{1-z}. \quad (4.17)$$

In turn, this means that

$$z = \frac{N_0}{N_0 + 1} \quad (4.18)$$

so that z is small if $N_0 \ll 1$. In turn, we may expect that if quantum effects are important and the $z \ll 1$ condition is not fulfilled, N_0 is much larger than 1. Such a state is known as the *Bose-Einstein condensation*.

Before proceeding, we note that while the $\varepsilon = 0$ state does matter very much in Eq. (4.16), it does not matter in Eq. (4.15) because $-\ln(1-z) = -\ln(1 - N_0/(N_0 + 1)) = \ln(N_0 + 1) = \mathcal{O}(\ln N)$ and so the $\varepsilon = 0$ -state contribution to pressure scales as

$$-\frac{1}{V} \ln(1-z) \propto \frac{1}{N} \ln N, \quad (4.19)$$

which is small for $N \rightarrow \infty$. The second term in Eq. (4.15) can therefore be neglected. Furthermore, if we integrate the integral in Eq. (4.15) by parts

$$\begin{aligned} & \int_0^\infty \ln(1 - z \exp(-x)) x^{1/2} dx \\ &= \underbrace{\ln(1 - z \exp(-x)) \frac{2}{3} x^{3/2}}_{=0} \Big|_0^\infty - \frac{2}{3} \int_0^\infty \frac{x^{3/2} dx}{z^{-1} \exp(x) - 1} \end{aligned} \quad (4.20)$$

we have

$$\frac{P}{k_B T} = \frac{4}{3\sqrt{\pi}\lambda^3} \int_0^\infty \frac{x^{3/2} dx}{z^{-1} \exp(x) - 1} = \frac{1}{\lambda^3} g_{5/2}(z) \quad (4.21)$$

and

$$\frac{N - N_0}{V} = \frac{2}{\sqrt{\pi}\lambda^3} \int_0^\infty \frac{x^{1/2} dx}{z^{-1} \exp(x) - 1} = \frac{1}{\lambda^3} g_{3/2}(z). \quad (4.22)$$

Here we replaced $\sqrt{\pi}/2$ by $\Gamma(3/2)$ and $3\sqrt{\pi}/4$ by $\Gamma(5/2)$ and we defined the Bose-Einstein function (Fig. 4.1)

$$g_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1} dx}{z^{-1} \exp(x) - 1}. \quad (4.23)$$

Figure 4.1 shows that at small z , that is in the classical regime, $g_{3/2}$ and $g_{5/2}$ are almost identical to each other. This is even better seen from the series expansion

$$g_\nu(z) = z + \frac{z^2}{2^\nu} + \frac{z^3}{3^\nu} + \dots \quad (4.24)$$

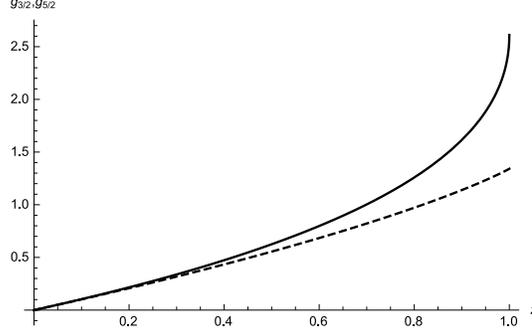


Figure 4.1: Bose-Einstein function vs. fugacity: $g_{3/2}$ (solid line) and $g_{5/2}$ (dashed line). Recall that $N - N_0$ is proportional to $g_{3/2}$ and pressure is proportional to $g_{5/2}$.

This means that at $z \rightarrow 0$

$$\frac{P}{k_B T} \approx \frac{N}{V}, \quad (4.25)$$

where we used the fact that $N_0 = z/(1-z) \approx z$ is also very small for $z \rightarrow 0$. This is the equation of state for ideal gas.

On the other hand, one may see from Fig. 4.1 that the number of particles in all states with $\varepsilon > 0$

$$N_e = N - N_0 = \frac{V}{\lambda^3} g_{3/2}(z) \quad (4.26)$$

remains finite as $z \rightarrow 1$. The largest possible value of N_e corresponds to $z = 1$ where $g_{3/2}(1) = \zeta(3/2) = 2.612\dots$, $\zeta(s)$ being the Riemann zeta function. If the actual number of particles is larger than N_e , then we expect that the $\varepsilon > 0$ states will accept as many particles as possible, leaving the remaining $N - N_e$ particles in the ground state. Note that N_e is largest at $z = 1$ which means that the chemical potential $\mu = 0$.

Since λ depends on temperature, so does N_e and thus

$$\begin{aligned} N_0(T) &= N - \frac{V}{\lambda^3} \zeta(3/2) = N - \frac{V(2\pi m k_B T)^{3/2}}{h^3} \zeta(3/2) \\ &= N \left[1 - \left(\frac{T}{T_c} \right)^{3/2} \right]. \end{aligned} \quad (4.27)$$

Here we have introduced the critical temperature

$$T_c = \frac{h^2}{2\pi m k_B} \left[\frac{N}{V \zeta(3/2)} \right]^{2/3}. \quad (4.28)$$

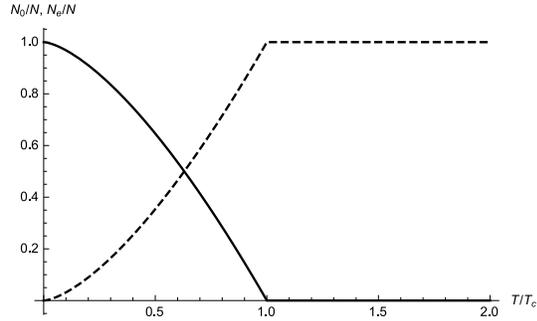


Figure 4.2: N_0/N (solid line) and N_e/N (dashed line) vs. reduced temperature.

In Fig. 4.2 we plot the fraction of particles in the ground state and the fraction of particles in states with $\varepsilon > 0$ vs. temperature.

Isotherms

In principle, the equation of state is obtained by eliminating λ^3 simultaneously from Eqs. (4.21) and (4.22). For $T > T_c$, N_0/N is very small and we have

$$P = \frac{Nk_B T}{V} \frac{g_{5/2}(z)}{g_{3/2}(z)} \quad (4.29)$$

where z itself depends on density as described by Eq. (4.22) with $N_0 = 0$.

At $T < T_c$, fugacity $z = 1$ and Eq. (4.21) then suggests that pressure does not depend on volume

$$P = P(T) = \frac{k_B T}{\lambda^3} \zeta(5/2) \propto T^{5/2} \quad (4.30)$$

which means that the compressibility of the gas is infinite.

An isotherm is shown in Fig. 4.3; transition between the two regimes is at a volume where $N_e = N$ or

$$\frac{N}{V} = \frac{1}{\lambda^3} \zeta(3/2). \quad (4.31)$$

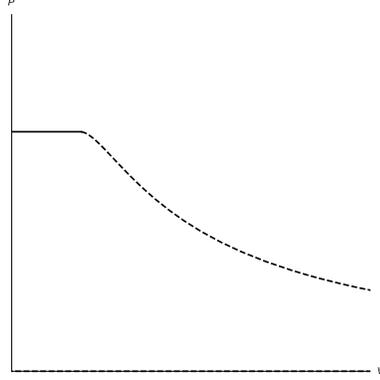


Figure 4.3: Isotherms of Bose-Einstein condensate: Condensed state (solid line) vs. normal state (dashed line).

Specific heat

These results may be used to calculate the specific heat of the system. To this end, recall that

$$U = - \left(\frac{\partial \ln \mathcal{Q}}{\partial \beta} \right)_{z, \beta} = k_B T^2 \left(\frac{\partial (PV/k_B T)}{\partial T} \right)_{z, \beta}, \quad (4.32)$$

where we used the relation $\ln \mathcal{Q} = PV/(k_B T)$. At $T < T_c$, Eq. (4.30) gives

$$\frac{PV}{k_B T} = \frac{V}{\lambda^3} \zeta(5/2) \propto T^{3/2} \quad (4.33)$$

so that $U \propto T^{5/2}$ and thus

$$c_V(T < T_c) \propto T^{3/2}. \quad (4.34)$$

We can derive the same result by starting with Eq. (4.16). We keep only the first term which represents the particles in states with $\varepsilon > 0$. We insert energy $\varepsilon = x/\beta$ in the integral and then the result is just the total energy of the gas:

$$\frac{U}{V} = \frac{2}{\sqrt{\pi} \lambda^3} \int_0^\infty \frac{\varepsilon x^{1/2} dx}{\exp(x) - 1} = \frac{2k_B T}{\sqrt{\pi} \lambda^3} \int_0^\infty \frac{x^{3/2} dx}{\exp(x) - 1} = \frac{3k_B T}{2\lambda^3} g_{5/2}(1) \quad (4.35)$$

We also replaced z by 1. Since $g_{5/2}(1) = \zeta(5/2) = 1.34149$ and since Eq. (4.28) implies that

$$\frac{V}{\lambda^3} = \frac{N}{\zeta(3/2)} \left(\frac{T}{T_c} \right)^{3/2} \quad (4.36)$$

we have

$$U = \frac{3\zeta(5/2)}{2\zeta(3/2)} Nk_B \frac{T^{5/2}}{T_c^{3/2}} = 0.770269 Nk_B \frac{T^{5/2}}{T_c^{3/2}}. \quad (4.37)$$

Thus

$$c_V(T < T_c) = 1.92567 \left(\frac{T}{T_c} \right)^{3/2}. \quad (4.38)$$

At $T \gg T_c$, the system approaches the classical limit and so

$$C_V(T > T_c) = \frac{3Nk_B}{2}. \quad (4.39)$$

This characteristic behavior is plotted in Fig. 4.4. A very similar specific-heat profile is observed in He^4 at 2.16 K undergoing the normal-superfluid transition. The $\varepsilon = 0$ particles may be identified with the superfluid component with zero viscosity whereas the rest of the particles are the normal component. (NB: Not all superfluids are Bose-Einstein condensates neither are all Bose-Einstein condensates superfluids.)

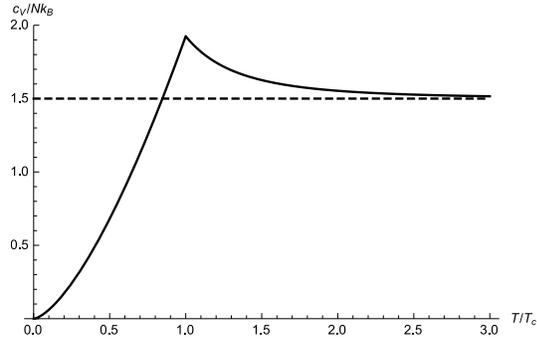


Figure 4.4: Specific heat of a Bose-Einstein condensate vs. reduced temperature; the $T < T_c$ part is proportional to $T^{3/2}$ and the high-temperature limit is classical, $3Nk_B/2$. The $T > T_c$ part of the curve connecting the peak at T_c with the high-temperature limit is semiquantitative.

4.3 Blackbody radiation

In Planck's interpretation of blackbody radiation, radiation is viewed as an ensemble of harmonic oscillators with a continuous range of frequencies from 0 to ∞ . At a finite temperature, the average energy of an oscillator is

$$\langle \varepsilon \rangle = \hbar\omega_s \left[\frac{1}{2} + \frac{1}{\exp(\hbar\omega/k_B T) - 1} \right]. \quad (4.40)$$

The zero-point energy $\hbar\omega/2$ may be omitted because it even when summed over all oscillators it is merely a constant offset; from now on, we use

$$\langle \varepsilon \rangle = \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1}. \quad (4.41)$$

The number of normal modes in a volume V is given by

$$2 \times \frac{V4\pi p^2 dp}{h^3} \quad (4.42)$$

The factor of 2 accounts for two polarizations, the momentum $\mathbf{p} = \hbar\mathbf{k}$, and $\omega = ck$ so that

$$2 \times \frac{V4\pi p^2 dp}{h^3} = \frac{V\omega^2 d\omega}{\pi^2 c^3}. \quad (4.43)$$

The energy density is then given by

$$\frac{dw}{d\omega} = \frac{V\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega/k_B T) - 1}. \quad (4.44)$$

This is Planck law (1900). If we redo the calculation by treating the oscillators classically which means that their average energy is $2 \times k_B T/2 = k_B T$, we obtain the Rayleigh-Jeans law (1900, 1905):

$$\left. \frac{dw}{d\omega} \right|_{\text{Rayleigh-Jeans}} = \frac{V k_B T \omega^2}{\pi^2 c^3}. \quad (4.45)$$

These two results may be conveniently written using the dimensionless frequency $x = \hbar\omega/k_B T$:

$$\frac{dw}{dx} = \frac{V(k_B T)^4}{\hbar^3 \pi^2 c^3} \frac{x^3}{\exp(x) - 1} \quad (4.46)$$

and

$$\left. \frac{dw}{dx} \right|_{\text{Rayleigh-Jeans}} = \frac{V(k_B T)^4}{\hbar^3 \pi^2 c^3} x^2 \quad (4.47)$$

and plotted against x (Fig. 4.5). It is easily seen that the Rayleigh-Jeans law does not converge to a finite total energy of the radiation because of the contribution of the high-frequency modes. This is known as the ultraviolet catastrophe.

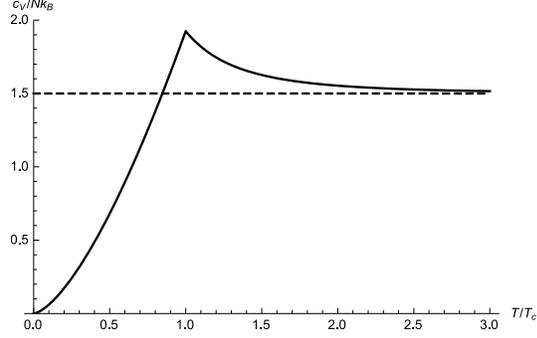


Figure 4.5: Planck law plotted against dimensionless frequency $x = \hbar\omega/k_B T$. Also plotted are Rayleigh-Jeans and Wien law.

Wien law

The maximum of dw/dx gives the Wien law:

$$\frac{d}{dx} \frac{dw}{dx} = \frac{V(k_B T)^4}{\hbar^3 \pi^2 c^3} \frac{3x^2 [\exp(x) - 1] - x^3 \exp(x)}{[\exp(x) - 1]^2} \propto \frac{x^2 [(3 - x) \exp(x) - 3]}{[\exp(x) - 1]^2} \quad (4.48)$$

and the maximum is determined by

$$(3 - x) \exp(x) - 3 = 0 \quad (4.49)$$

which gives $x_{\max} = 2.82144$. The Wien law thus states that

$$\omega_{\max} T = \frac{2.82144 k_B}{\hbar} = 3.708 \times 10^{11} \frac{\text{K}}{\text{s}}. \quad (4.50)$$

Alternatively, one may consider $dw/d\lambda$ instead of $dw/d\omega$ which gives the more familiar version of this result $\lambda_{\max} T = 2896 \text{ K}\mu\text{m}$.

The Rayleigh-Jeans is not the only alternative model for the blackbody radiation. Wien (1896) proposed a formula

$$\left. \frac{dw}{dx} \right|_{\text{Wien}} = \frac{V(k_B T)^4}{\hbar^3 \pi^2 c^3} x^3 \exp(-x) \quad (4.51)$$

which does not suffer from the ultraviolet catastrophe and is a good approximation to the Planck law at large frequencies.

Stefan law

The total energy of the radiation is

$$U = \int_0^\infty \frac{dw}{dx} dx = \frac{V(k_B T)^4}{\hbar^3 \pi^2 c^3} \underbrace{\int_0^\infty \frac{x^3}{\exp(x) - 1} dx}_{=\pi^4/15}. \quad (4.52)$$

Doing the integral is easier than one may think:

$$\begin{aligned} \int_0^\infty \frac{x^3}{\exp(x) - 1} dx &= \int_0^\infty \frac{x^3 \exp(-x)}{1 - \exp(-x)} dx \\ &= \sum_{n=0}^\infty \int_0^\infty x^3 \exp(-(n+1)x) dx \\ &= \sum_{n=1}^\infty \frac{6}{n^4} = 6\zeta(4) \end{aligned} \quad (4.53)$$

The value of Riemann zeta function $\zeta(4) = \pi^4/90$ so that the result is indeed $\pi^4/15$. Finally

$$U = \frac{\pi^2 V (k_B T)^4}{15 \hbar^3 c^3}. \quad (4.54)$$

The energy flux density is given by $j = (U/V)c/4$ (this result we borrow from the kinetic theory of gases) so that

$$j = \frac{\pi^2 (k_B T)^4}{60 \hbar^3 c^2} = \sigma T^4 \quad (4.55)$$

where

$$\sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2} = 5.67 \times 10^{-8} \frac{\text{W}}{\text{m}^2 \text{K}^4} \quad (4.56)$$

is the Stefan constant.

Thermodynamics of blackbody radiation

To study the thermodynamics, we switch to the particle picture of the radiation, that is, to the photon-gas picture. Note that in photons, $\varepsilon = cp$ where c is the speed of light and p is the momentum.

We first calculate the pressure. Starting from $PV/k_B T = \ln Q$ we have

$$\frac{PV}{k_B T} = - \sum_r \ln(1 - \exp(-\beta \varepsilon_r)) \quad (4.57)$$

like in Eq. (4.12) with $z = 1$, that is, $\mu = 0$. We replace the sum by the integral, again accounting for two polarizations. We first recast Eq. (4.43) as

$$\frac{V\varepsilon^2 d\varepsilon}{\pi^2 \hbar^3 c^3}. \quad (4.58)$$

so that after replacing the sum by the integral we have

$$\frac{PV}{k_B T} = -\frac{V}{\pi^2 \hbar^3 c^3} \int_0^\infty \ln(1 - \exp(-\beta\varepsilon_r)) \varepsilon^2 d\varepsilon. \quad (4.59)$$

We introduce $x = \beta\varepsilon$ and like in Eq. (4.20), we integrate by parts:

$$\begin{aligned} & \int_0^\infty \ln(1 - \exp(-x)) x^2 dx \\ &= \underbrace{\ln(1 - \exp(-x)) \frac{1}{3} x^3 \Big|_0^\infty}_{=0} - \frac{1}{3} \int_0^\infty \frac{x^3 dx}{\exp(x) - 1}. \end{aligned} \quad (4.60)$$

The result is

$$\frac{PV}{k_B T} = \frac{V(k_B T)^3}{3\pi^2 \hbar^3 c^3} \underbrace{\int_0^\infty \frac{x^3 dx}{\exp(x) - 1}}_{=\pi^4/15} \quad (4.61)$$

and thus

$$P = \frac{\pi^2 (k_B T)^4}{45 \hbar^3 c^3} = \frac{U}{3V}. \quad (4.62)$$

Using the first law, $dU = TdS - PdV$, we have

$$dS = \frac{dU}{T} + \frac{PdV}{T}. \quad (4.63)$$

Now $dU = d[Vu(T)] = u(T)dV + V[du(T)/dT]dT$. Since $u(T) = 4\sigma T^4$ this gives

$$\begin{aligned} dS &= \frac{4\sigma T^4 dV}{T} + \frac{16V\sigma T^3 dT}{T} + \frac{4\sigma T^4 dV}{3T} \\ &= 16\sigma \left(\frac{1}{3} T^3 dV + VT^2 dT \right). \end{aligned} \quad (4.64)$$

This is a total differential and so

$$S(T, V) = \frac{16}{3} \sigma T^3 V. \quad (4.65)$$

From Eq. (4.64) it also follows that during an adiabatic change

$$\frac{dV}{3V} = \frac{dT}{T} \quad (4.66)$$

or $VT^3 = \text{const.}$; given that $P \propto T^4$ this can also be written as $pV^{4/3} = \text{const.}$

4.4 Ideal Fermi gas

In this case, the occupation number of any state can be either 0 or 1. The grand-canonical partition function reads

$$\begin{aligned} \mathcal{Q} &= \prod_r \sum_{N_r=0}^1 \exp(\beta N_r (\mu - \varepsilon_r)) = \prod_r [1 + \exp(\beta(\mu - \varepsilon_r))] \\ &= \prod_r [1 + z \exp(-\beta \varepsilon_r)] \end{aligned} \quad (4.67)$$

and since $\ln \mathcal{Q} = \beta PV$ we have

$$\frac{PV}{k_B T} = \sum_r \ln(1 + z \exp(-\beta \varepsilon_r)). \quad (4.68)$$

The number of particles is

$$N = \sum_r \frac{1}{\exp(\beta(\varepsilon_r - \mu)) + 1} = \sum_r \frac{1}{z^{-1} \exp(\beta \varepsilon_r) + 1}. \quad (4.69)$$

Like in the Bose gas, we replace the sum over states by an integral over energy, the single-particle density of states being given by

$$\frac{V}{h^3} 4\pi p^2 dp = \frac{2\pi V (2m)^{3/2}}{h^3} \varepsilon^{1/2} d\varepsilon. \quad (4.70)$$

Thus we have

$$\begin{aligned} \frac{P}{k_B T} &= \int_0^\infty \frac{2\pi (2m)^{3/2}}{h^3} \ln(1 + z \exp(-\beta \varepsilon)) \varepsilon^{1/2} d\varepsilon \\ &= \int_0^\infty \frac{2\pi (2mk_B T)^{3/2}}{h^3} \ln(1 + z \exp(-x)) x^{1/2} dx \\ &= \frac{2}{\sqrt{\pi} \lambda^3} \int_0^\infty \ln(1 + z \exp(-x)) x^{1/2} dx \end{aligned} \quad (4.71)$$

and

$$\begin{aligned}
\frac{N}{V} &= \int_0^\infty \frac{2\pi(2m)^{3/2}}{h^3} \frac{1}{z^{-1} \exp(\beta\varepsilon_r) + 1} \varepsilon^{1/2} d\varepsilon \\
&= \int_0^\infty \frac{2\pi(2mk_B T)^{3/2}}{h^3} \frac{1}{z^{-1} \exp(x) + 1} x^{1/2} dx \\
&= \frac{2}{\sqrt{\pi}\lambda^3} \int_0^\infty \frac{x^{1/2} dx}{z^{-1} \exp(x) + 1}. \tag{4.72}
\end{aligned}$$

Like in the ideal Bose gas, Eq. (4.71) can be integrated by part to read

$$\frac{P}{k_B T} = \frac{4}{3\sqrt{\pi}\lambda^3} \int_0^\infty \frac{x^{3/2} dx}{z^{-1} \exp(x) + 1}. \tag{4.73}$$

By introducing Fermi-Dirac functions

$$f_\nu(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1} dx}{z^{-1} \exp(x) + 1}, \tag{4.74}$$

we can write the pressure and the number of particles as

$$\frac{P}{k_B T} = \frac{g}{\lambda^3} f_{5/2}(z) \tag{4.75}$$

and

$$\frac{N}{V} = \frac{g}{\lambda^3} f_{3/2}(z), \tag{4.76}$$

respectively. Here g is the spin degeneracy $2j + 1$. These two functions are plotted in Fig. 4.6.

The equation of state is obtained by eliminating fugacity z from these equations. At small z , the power series of f_ν is useful:

$$f_\nu(z) = z - \frac{z^2}{2^\nu} + \frac{z^3}{3^\nu} - \dots \tag{4.77}$$

To lowest order, $f_\nu = z$ so that for $z \rightarrow 0$

$$\frac{P}{k_B T} \approx \frac{N}{V}. \tag{4.78}$$

This is the equation of state for ideal gas. Like in the ideal Bose gas, in the regime of weak degeneracy (that is, at small fugacity/small densities), the isotherms are the same as in ideal gas.

At finite fugacities, the isotherms depart from the $1/V$ law (Fig. 4.7).

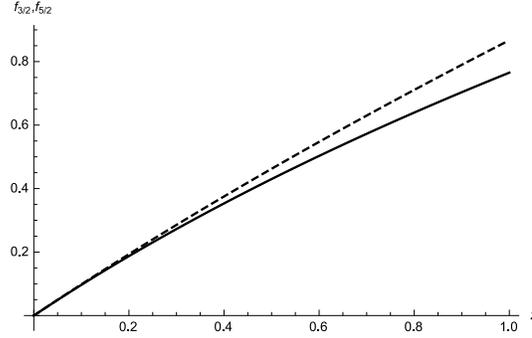


Figure 4.6: Ideal Fermi gas: $f_{3/2}$ (solid line) and $f_{5/2}$ (dashed line). Recall that N/V is proportional to $f_{3/2}$ and pressure is proportional to $f_{5/2}$.

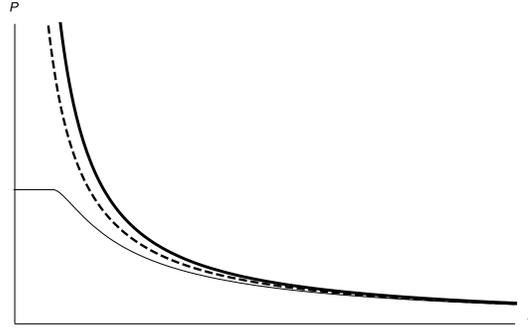


Figure 4.7: Isotherms of the ideal Fermi gas and the ideal Bose gas (thick and thin line, respectively) compared to the ideal-gas isotherm (dashed line).

The Fermi-Dirac functions can also be used to evaluate the specific heat. Note that

$$U = - \left(\frac{\partial \ln \mathcal{Q}}{\partial \beta} \right)_{z,V} = - \frac{dT}{d\beta} \left(\frac{\partial \ln \mathcal{Q}}{\partial T} \right)_{z,V} = k_B T^2 \left(\frac{\partial \ln \mathcal{Q}}{\partial T} \right)_{z,V}. \quad (4.79)$$

Now $\ln \mathcal{Q} = PV/(k_B T) = gV f_{5/2}(z)/\lambda^3$. All temperature dependence is hidden in the de Broglie wavelength so that

$$\begin{aligned} U &= k_B T^2 gV f_{5/2}(z) \frac{d}{dT} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} = \frac{3}{2} k_B T \frac{gV}{\lambda^3} f_{5/2}(z) \\ &= \frac{3}{2} k_B T \frac{gV}{\lambda^3} f_{5/2}(z). \end{aligned} \quad (4.80)$$

Note that $gV/\lambda^3 = Nf_{3/2}(z)$ so that

$$U = \frac{3}{2}Nk_B T \frac{f_{5/2}(z)}{f_{3/2}(z)}. \quad (4.81)$$

To calculate the heat capacity, we differentiate U with respect to T ; the final result is

$$C_v = Nk_B \left[\frac{15f_{5/2}(z)}{4f_{3/2}(z)} - \frac{9f_{3/2}(z)}{4f_{1/2}(z)} \right]. \quad (4.82)$$

Recall that z depends on T through the de Broglie wavelength [Eq. (4.76)]. In Fig. 4.8, heat capacity is plotted vs. temperature; the low-temperature behavior is linear, $C_V \propto T$ and the high-temperature limit is the classical result $3Nk_B/2$.



Figure 4.8: Heat capacity of ideal Fermi gas.

4.5 White dwarfs

White dwarfs are old stars that have burnt most of their hydrogen. They contain about 10^{30} kg of helium at density of about 10^{10} kg/m³ and a temperature of about 10^7 K. At this temperature, the thermal energy $k_B T \sim 10^3$ eV which is much more than the ionization energy of helium atoms — thus all helium is essentially ionized, and the dwarf consists of electrons and nuclei. Let N be the number of electrons; then the mass of the star is to a good approximation given by the mass of nuclei $M = (N/2)4m_p = 2Nm_p$, where m_p is the proton mass. Given that the nuclei are much heavier than the electrons, they may be viewed merely as a static background, and the thermodynamic properties of the star such as pressure may be associated solely with electrons.

The number density of electrons is

$$\frac{N}{V} = \frac{M/(2m_p)}{V} = \frac{\rho}{2m_p} \sim 10^{37}/\text{m}^3. \quad (4.83)$$

The corresponding Fermi momentum of electrons is

$$p_F \sim h \left(\frac{N}{V} \right)^{1/3} = 6.65 \times 10^{-34} \text{ Js} \times 2 \times 10^{12} \text{ m}^{-1} \sim 10^{-22} \text{ kgm/s} \quad (4.84)$$

and is very similar to their relativistic momentum $m_e c = 9.1 \times 10^{-31} \text{ kg} \times 3 \times 10^8 \text{ m/s} \sim 10^{22} \text{ kgm/s}$. This means that the electrons are essentially relativistic. At the same time, their Fermi temperature is

$$T_F = \frac{p_F^2}{2m_e k_B} \sim 10^9 \text{ K} \quad (4.85)$$

which is 100 times larger than the actual temperature of the star; thus the electrons are completely degenerate.

Let us now calculate the pressure P at absolute zero. The pressure is obtained from

$$\frac{PV}{k_B T} = \sum_r \ln(1 + z \exp(-\beta \varepsilon_r)) = \frac{8\pi V}{h^3} \int_0^\infty \ln(1 + z \exp(-\beta \varepsilon(p))) p^2 dp. \quad (4.86)$$

Integrating by parts gives

$$\begin{aligned} \frac{P}{k_B T} &= \frac{8\pi\beta}{h^3} \int_0^\infty \frac{1}{3} \frac{z \exp(-\beta \varepsilon(p))}{1 + z \exp(-\beta \varepsilon(p))} \frac{d\varepsilon}{dp} p^3 dp \\ &= \frac{8\pi\beta}{3h^3} \int_0^\infty \frac{p^3 dp}{z^{-1} \exp(\beta \varepsilon(p)) + 1} \frac{d\varepsilon}{dp}. \end{aligned} \quad (4.87)$$

Here we took into account that the uv term in the *per partes* formula vanishes. Now we recognize that the Fermi-Dirac mean occupancy is 1 for $p < p_F$ and 0 for $p > p_F$ so that

$$P = \frac{8\pi}{3h^3} \int_0^{p_F} p^3 dp \frac{d\varepsilon}{dp}. \quad (4.88)$$

The energy of a relativistic particle is

$$\varepsilon(p) = mc^2 \left[\sqrt{1 + \left(\frac{p}{mc} \right)^2} - 1 \right] \quad (4.89)$$

so that its speed is

$$\frac{d\varepsilon}{dp} = \frac{p/m}{\sqrt{1 + [p/(mc)]^2}}. \quad (4.90)$$

Thus

$$P = \frac{8\pi}{3h^3m} \int_0^{p_F} \frac{p^4 dp}{\sqrt{1 + [p/(mc)]^2}} \quad (4.91)$$

The number of electrons is

$$N = \frac{8\pi V}{h^3} \int_0^{p_F} p^2 dp = \frac{8\pi V p_F^3}{3h^3}. \quad (4.92)$$

In view of the root in Eq. (4.91), it is convenient to introduce a new variable such that

$$p = mc \sinh \theta. \quad (4.93)$$

Then the integrand in Eq. (4.91) is given by $(mc)^5 \sinh^4 \theta d\theta$ and

$$P = \frac{8\pi m^4 c^5}{3h^3} \int_0^{\theta_F} \sinh^4 \theta d\theta = \frac{\pi m^4 c^5}{3h^3} A(x), \quad (4.94)$$

where $x = \sinh \theta_F$ and

$$A(x) = x(x^2 + 1)^{1/2}(2x^2 - 3) + 3 \sinh^{-1} x. \quad (4.95)$$

In terms of x ,

$$N = \frac{8\pi V m^3 c^3}{3h^3} x^3. \quad (4.96)$$

This result is needed to determine θ_F as a function of N/V and thus as function of the ratio M/R^3 through Eq. (4.83).

The star is isolated, $dE = -PdV$, and held together by gravitational attraction; the sum of internal and gravitational energy is constant, $E + E_g = \text{const.}$ so that $dE = -dE_g$. The mechanical work done on expansion equals minus the change of gravitational energy:

$$-dE_g = -PdV = -P(R) \times 4\pi R^2 dR, \quad (4.97)$$

where R is the radius of the star and density is assumed to be uniform. The self-gravitational energy of a homogeneous sphere of mass M is

$$E_g = -\frac{3GM^2}{5R} \quad (4.98)$$

so that

$$dE_g = \left(\frac{dE_g}{dR} \right) dR = -\alpha \frac{GM^2}{R^2} dR. \quad (4.99)$$

Here α is a numerical constant of order 1; technically, $3/5$ but in practice somewhat different due to the variation of density within the star. Thus

$$P(R) = \frac{\alpha}{4\pi} \frac{GM^2}{R^4}. \quad (4.100)$$

Here R is the equilibrium radius of the star.

This result should be equal to the pressure due to electrons [Eq. (4.94)], providing a relation between mass and radius of the star. To understand this relationship, first recall that $x = \sinh \theta_F$ so that $p_F = mcx$. Thus $x \ll 1$ corresponds to the non-relativistic regime where $p \ll mc$ whereas $x \gg 1$ corresponds to the ultra-relativistic regime. According to Eq. (4.96),

$$x = \left(\frac{3N}{8\pi V} \right)^{1/3} \frac{h}{mc} = \left(\frac{9N}{32\pi^2} \right)^{1/3} \frac{h}{mcR} \quad (4.101)$$

Since $M = 2Nm_p$, we have

$$x = \left(\frac{9M}{64\pi^2 m_p} \right)^{1/3} \frac{h}{mcR}. \quad (4.102)$$

The pressure balance now reads

$$\frac{\pi m^4 c^5}{3h^3} A \left(\left(\frac{9\pi M}{8m_p} \right)^{1/3} \frac{\hbar}{mcR} \right) = \frac{\alpha}{4\pi} \frac{GM^2}{R^4} \quad (4.103)$$

or

$$A \left(\left(\frac{9M}{64\pi^2 m_p} \right)^{1/3} \frac{\lambda_C}{R} \right) = \frac{3\alpha}{4\pi^2} \left(\frac{\lambda_C}{R} \right)^3 \frac{GM^2/R}{mc^2}. \quad (4.104)$$

This result has an appealing form. Apart from numerical constants, the argument of function A is spelled out as a product of the cubic root of ratio of star mass and mass of proton $(M/m_p)^{1/3}$ and the ratio of the Compton wavelength $\lambda_C = h/(mc)$ and star radius λ_C/R . The latter also appears on the right-hand side of the equation as $(\lambda_C/R)^3$, multiplied by the ratio of gravitational energy and rest energy of the electron $GM^2 R^{-1}/(mc^2)$.

In general, this equation cannot be solved explicitly for $R = R(M)$. But for large values of masses/small radii, the argument of function $A(x)$ is very large and in this limit $A(x) \approx 2x^2(x^2 - 1)$. In this case,

$$R \approx \frac{(9\pi)^{1/3}}{2} \lambda_C \left(\frac{M}{m_p} \right)^{1/3} \sqrt{1 - \left(\frac{M}{M_0} \right)^{2/3}}, \quad (4.105)$$

where

$$M_0 = \frac{9}{24} \left(\frac{3\pi}{\alpha^3} \right)^{1/3} \frac{(\hbar c G)^{3/2}}{m_p^2} \quad (4.106)$$

is the limiting mass also known as the Chandrasekhar limit. M_0 about $1.44M_\odot$ where M_\odot is the mass of Sun. The numerically obtained relationship $R = R(M)$ shown in Fig. 4.9 is a decreasing function characterized by zero radius at $M = M_0$, which means that the white dwarf of a large enough mass will collapse due to gravitational attraction, the pressure of the electron gas not being able to support the star.

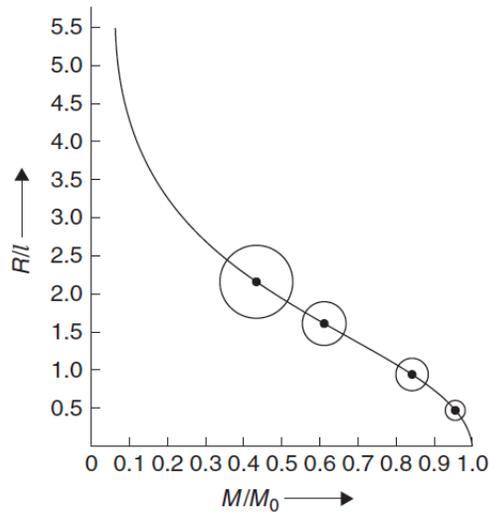


Figure 4.9: White dwarf radius vs. mass in units of M_\odot . The Chandrasekhar limit is about $1.44M_\odot$ [1].

Chapter 5

Interacting systems

All of the above results pertain to ideal, that is, noninteracting systems. Interparticle interactions change the formalism dramatically because they couple the degrees of freedom; typically positions of particles as the interactions depend on their relative positions. The thermodynamics of interacting systems can be dealt with systematically in the low-density limit by resorting to a suitable series expansion such as the virial expansion. The most general approach of this type is the cluster expansion due to Mayer.

5.1 Cluster expansion

The Hamiltonian of an interacting system consists of the kinetic energy and of the pair interactions:

$$H = \sum_i \frac{\mathbf{p}^2}{2m} + \sum_{i<j} u_{ij}, \quad (5.1)$$

where $i, j, = 1, 2, \dots N..$ The pair potential u_{ij} usually depends only on the distance between particles $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The canonical partition function

$$Q_N(V, T) = \frac{1}{N!h^{3N}} \int \exp \left(-\beta \left[\sum_i \frac{\mathbf{p}^2}{2m} + \sum_{i<j} u_{ij} \right] \right) d^{3N}p d^{3N}r \quad (5.2)$$

can be readily integrated over momenta

$$Q_N(V, T) = \frac{1}{N!\lambda^{3N}} \int \exp \left(-\beta \sum_{i<j} u_{ij} \right) d^{3N}r \quad (5.3)$$

which can also be written as

$$Q_N(V, T) = \frac{1}{N! \lambda^{3N}} Z_N(V, T) \quad (5.4)$$

where

$$Z_N(V, T) = \int \exp \left(-\beta \sum_{i < j} u_{ij} \right) d^{3N} r = \int \prod_{i < j} \exp(-\beta u_{ij}) d^{3N} r \quad (5.5)$$

is the configuration integral. In a noninteracting system, $Z_N^{(0)}(V, T) = V^N$.

The key quantity of the cluster expansion is the Mayer function

$$f_{ij} = \exp(-\beta u_{ij}) - 1, \quad (5.6)$$

which is 0 in a noninteracting system. For a typical interparticle potential consisting of short-range repulsion and a van der Waals type attraction, the Mayer function essentially equals -1 at small separations and is positive within a limited range comparable to the diameter of the molecules (Fig. 5.1). In terms of the Mayer function, the configuration integral can be

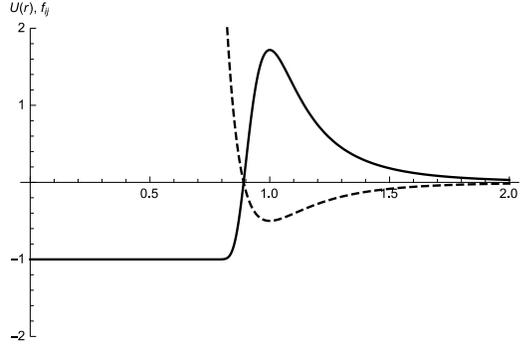


Figure 5.1: Mayer function (solid line) for a typical 6 – 12 Lennard-Jones potential (dashed line).

written as

$$Z_N(V, T) = \int \prod_{i < j} (1 + f_{ij}) d^3 r_1 \dots d^3 r_N. \quad (5.7)$$

The product can now be spelled out

$$Z_N(V, T) = \int \left(1 + \sum f_{ij} + \sum f_{ij} f_{kl} + \dots \right) d^3 r_1 \dots d^3 r_N. \quad (5.8)$$

The idea of cluster expansion is in a systematic treatment of the different terms in this expression. This is best done by viewing the terms as graphs. For example, consider this $N = 8$ term:

$$t_A = \int f_{12}f_{14}f_{67}d^3r_1 \dots d^3r_8 \quad (5.9)$$

which can be factorized as

$$\begin{aligned} t_A &= \int d^3r_3 \int d^3r_5 \int d^3r_8 \int f_{67}d^3r_6 d^3r_7 \int f_{12}f_{14}d^3r_1 d^3r_2 d^3r_4 \\ &= [3].[5].[8].[6-7][2-1-4]. \end{aligned} \quad (5.10)$$

This form emphasizes that the graph representing the term can be decomposed into 5 disconnected graphs, 3 of which correspond to a single particle whereas the remaining 2 couple two and three particles, respectively. In other words, the graph represents a configuration of 3 one-particle clusters, 1 two-particle cluster and 1 three-particle clusters.

Thus we may introduce a N -particle graph as a "collection of N circles numbered $1, 2, \dots, N$ with lines linking some or all of the circles". Then the configuration integral Z_N may be viewed as

$$Z_N(V, T) = \text{sum of all distinct } N\text{-particle graphs.} \quad (5.11)$$

Here "distinct" refers to graphs that differ in the links between circles in any way; that is, either in the number of links or the way they connect the different circles or in the labels of circles that they connect. For example,

$$t_A = \int f_{12}f_{14}f_{67}d^3r_1 \dots d^3r_8 \quad (5.12)$$

and

$$t'_A = \int f_{12}f_{14}f_{78}d^3r_1 \dots d^3r_8 \quad (5.13)$$

are distinct because in t_A , particle 7 is coupled to particle 6 whereas in t'_A it is coupled to particle 8.

Given that some terms may be factorized (and the corresponding graphs decomposed into disconnected clusters) it is convenient to introduce l -clusters as " l -particle graphs where every circle is directly or indirectly connected to every other circle". An example of a 5-cluster is shown in Fig. 5.2. The clusters cannot be decomposed into smaller graphs. The value of some clusters is the same: For example, these chain-like clusters of three particles

$$t_A^3 = \int f_{12}f_{23}d^3r_1 d^3r_2 d^3r_3, \quad (5.14)$$

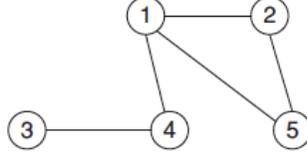


Figure 5.2: A 5-cluster [1].

$$t_B^3 = \int f_{12}f_{13}d^3r_1d^3r_2d^3r_3, \quad (5.15)$$

and

$$t_C^3 = \int f_{13}f_{23}d^3r_1d^3r_2d^3r_3 \quad (5.16)$$

differ only in the labels of the particles and are evidently identical. This one

$$t_D^3 = \int f_{12}f_{23}f_{31}d^3r_1d^3r_2d^3r_3 \quad (5.17)$$

is not of chain type and is different from the above three clusters.

We now introduce the cluster integral b_l by

$$b_l(V, T) = \frac{1}{l!\lambda^{3(l-1)}V} \times \text{sum of all possible } l - \text{clusters.} \quad (5.18)$$

By including a factor of V in the denominator, we take into account that any cluster is translationally invariant: If we place one of the particles (say particle 1) in the origin such that all other particle positions may be measured relative to its location, the integrand then depends only on these relative positions. We first carry out the integrals over particle 2, 3, etc., which amount to a certain quantity with the dimension (but not the magnitude!) of $(\text{volume})^{l-1}$; the size of the volume is determined by the range of the interparticle potential. The remaining integral over the location of particle 1 gives a factor of V . Thus by dividing the cluster by $\lambda^{3(l-1)}V$ we ensure that the cluster integral is dimensionless.

The additional factor of $l!$ in the denominator accounts for permutations of particles between circles in a given cluster. Finally note that as $V \rightarrow \infty$, $b_l(V, T)$ approach a finite value independent of volume $\bar{b}_l(T)$.

The simplest cluster integrals are

$$b_1 = \frac{1}{V}[1] = \frac{1}{V} \int d^3r_1 = 1 \quad (5.19)$$

and

$$\begin{aligned} b_2 &= \frac{1}{2\lambda^3 V} [1 - 2] = \frac{1}{2\lambda^3 V} \int f_{12} d^3 r_1 d^3 r_2 = \frac{1}{2\lambda^3} \int f_{12} d^3 r_{12} \\ &= \frac{2\pi}{\lambda^3} \int f(r) r^2 dr = \frac{2\pi}{\lambda^3} \int_0^\infty [\exp(-\beta u(r)) - 1] r^2 dr \end{aligned} \quad (5.20)$$

where we took into account that in 1 and in 2 particles, there exist a single cluster. On the other hand, there exist 2 types of 3-clusters: three chain-type clusters and one circle-type cluster. Thus

$$\begin{aligned} b_3 &= \frac{1}{6\lambda^6 V} \{[1 - 2 - 3] + [2 - 1 - 3] + [2 - 3 - 1] + \text{circle}\} \\ &= \frac{1}{6\lambda^6 V} \int (f_{12} f_{23} + f_{12} f_{13} + f_{23} f_{13} + f_{12} f_{23} f_{13}) d^3 r_1 d^3 r_2 d^3 r_3 \\ &= \frac{1}{6\lambda^6 V} \left(3V \int f_{12} f_{13} d^3 r_{12} d^3 r_{13} + V \int f_{12} f_{23} f_{13} d^3 r_{12} d^3 r_{13} \right) \\ &= \frac{1}{6\lambda^6 V} \left(3V \int f_{12} d^3 r_{12} \times \int f_{13} d^3 r_{13} + V \int f_{12} f_{23} f_{13} d^3 r_{12} d^3 r_{13} \right) \\ &= 2b_2^2 + \frac{1}{6\lambda^6} \int f_{12} f_{23} f_{13} d^3 r_{12} d^3 r_{13}. \end{aligned} \quad (5.21)$$

The sum over all distinct N -particle graphs in the configuration integral can be viewed as a sum over the different clusters. The numbers of different clusters, m_l , in a N -particle graphs must satisfy the constraint

$$\sum_{l=1}^N l m_l = N, \quad m_l = 0, 1, 2 \dots N. \quad (5.22)$$

A given set m_l does not specify a single graph; instead, it represents a collection of graphs $S\{m_l\}$ differing in the way the particles are *assigned* to the clusters and in the way the clusters are *formed* from a given set of particles. The configuration integral can now be written as

$$Z_N(V, T) = \sum_{\{m_l\}}' S\{m_l\} \quad (5.23)$$

where the primed sum is over all sets subject to the condition $\sum_{l=1}^N l m_l = N$.

The number of different assignments of particles to the clusters is given by the multinomial symbol

$$\# \text{ assignments} = \frac{N!}{(1!)^{m_1} (2!)^{m_2} \dots} = \frac{N!}{\prod_l (l!)^{m_l}}. \quad (5.24)$$

The number of ways a cluster may be formed is best understood by first assuming that the clusters are unique, that is, there exists a single way of forming an l -cluster. If so,

$$\tilde{S}\{m_l\} = \prod_l \frac{(\text{value of } l\text{-cluster})^{m_l}}{m_l!}. \quad (5.25)$$

Here $\tilde{S}\{m_l\}$ is the sum of all graphs for a *given* assignment of particles to clusters.

The numerator states that the graphs are products of clusters, and the denominator accounts for the fact that if all particles in a given l -cluster are replaced by all particles of another l -cluster, the value of the graph does not change. The most trivial example are two 2-clusters of $N = 4$ particles. ("Trivial" because there is indeed a single way of forming 2-clusters using a given pair of particles.) The complete $S\{m_1 = 0, m_2 = 2, m_3 = 0, m_4 = 0\}$ graph can be written in two ways

$$\tilde{S}\{m_1 = 0, m_2 = 2, m_3 = 0, m_4 = 0\}_A = \underbrace{\int f_{12}d^3r_1d^3r_2}_{\text{1st 2-cluster}} \times \underbrace{\int f_{34}d^3r_3d^3r_4}_{\text{2nd 2-cluster}}, \quad (5.26)$$

and

$$\tilde{S}\{m_1 = 0, m_2 = 2, m_3 = 0, m_4 = 0\}_B = \underbrace{\int f_{34}d^3r_3d^3r_4}_{\text{1st 2-cluster}} \times \underbrace{\int f_{12}d^3r_1d^3r_2}_{\text{2nd 2-cluster}}. \quad (5.27)$$

These two ways are evidently equivalent simply because multiplication is commutative and should not be counted as distinct, hence division by $m_l!$ in Eq. (5.25).

In addition, except for $l = 2$ there exist more than a single way of forming an l -cluster. For example, 3-clusters in Eqs. (5.14)-(5.16) are equivalent but that in Eq. (5.17) is different. This is taken care of by replacing "(value of l -cluster)" in Eq. (5.25) by "(sum of values of all possible l -clusters)". To appreciate this, recall that the original configuration integral Eq. (5.8) contains a sum of all l -clusters of a given set of particles, the different types of clusters being included in the different partial sums. For example, 3-clusters of type Eqs. (5.14) are included in the partial sum containing two bonds $f_{ij}f_{kl}$ whereas the ring cluster of type Eq. (5.17) is included in the partial sum containing three bonds $f_{ij}f_{kl}f_{mn}$. Thus we con-

clude that

$$\tilde{S}\{m_l\} = \prod_l \frac{(\text{sum of values of all possible } l\text{-clusters})^{m_l}}{m_l!}. \quad (5.28)$$

The (sum of values of all possible l -clusters) can be expressed using cluster integrals [Eq. (5.18)]. If we also include the number of different assignments of particles to the different clusters [Eq. (5.24)] we finally have

$$S\{m_l\} = N! \prod_l \frac{[b_l l! \lambda^{3(l-1)} V]^{m_l}}{m_l!} \frac{1}{l!^{m_l}}. \quad (5.29)$$

Now the $l!^{m_l}$ factors in the numerator and the denominator cancel and since $\sum_l l m_l = N$ we have

$$S\{m_l\} = N! \lambda^{3N} \prod_l \left(\frac{b_l V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!}. \quad (5.30)$$

Note that the $N! \lambda^{3N}$ factor is exactly the inverse of the ratio of the canonical partition function and the configuration integral. Thus the canonical partition function is

$$Q_N(V, T) = \sum'_{\{m_l\}} \left[\prod_l \left(\frac{b_l V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} \right]. \quad (5.31)$$

Recall that the primed sum is over all sets $\{m_l\}$ subject to the condition $\sum_{l=1}^N l m_l = N$. Because of this constraint, doing the sum is technically quite demanding; it is a lot easier to consider the grand-canonical partition function instead.

$$\mathcal{Q}(z, V, T) = \sum_{N=0}^{\infty} z^N Q_N(V, T). \quad (5.32)$$

Here the constraint is lifted and we sum over all possible values of m_l rather than over all N and all sets $\{m_l\}$ for a given N . Note that $z^N = z^{\sum_l l m_l}$; this eventually gives

$$\begin{aligned} \mathcal{Q}(z, V, T) &= \sum_{m_1, m_2, \dots = 0}^{\infty} \prod_{l=1}^{\infty} \left(\frac{b_l z^l V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} = \prod_{l=1}^{\infty} \sum_{m_l=0}^{\infty} \left(\frac{b_l z^l V}{\lambda^3} \right)^{m_l} \frac{1}{m_l!} \\ &= \prod_{l=1}^{\infty} \exp \left(\frac{b_l z^l V}{\lambda^3} \right) = \exp \left(V \sum_{l=1}^{\infty} \frac{b_l z^l}{\lambda^3} \right). \end{aligned} \quad (5.33)$$

Thus

$$\frac{1}{V} \ln \mathcal{Q} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l. \quad (5.34)$$

As $V \rightarrow \infty$, $b_l \rightarrow \bar{b}_l$ and thus

$$\frac{P}{k_B T} = \frac{\ln \mathcal{Q}}{V} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \bar{b}_l z^l \quad (5.35)$$

and

$$\frac{N}{V} = \frac{1}{V} \frac{\partial \ln \mathcal{Q}}{\partial \ln z} = \frac{1}{V} \underbrace{\frac{\partial z}{\partial \ln z}}_{=z} \frac{\partial \ln \mathcal{Q}}{\partial z} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l \bar{b}_l z^l. \quad (5.36)$$

These two equations are known as the cluster expansions of the Mayer-Ursell formalism. Note that in a non-interacting system, all \bar{b}_l except $\bar{b}_1 = 1$ vanish and in this case, the equations reduce to those corresponding to the ideal gas.

5.2 Virial expansion

The Mayer-Ursell equations give $P/(k_B T)$ and N/V as a power series of z . On the other hand, the virial expansion

$$\frac{PV}{Nk_B T} = \sum_{l=1}^{\infty} B_l(T) \rho^{l-1} \quad (5.37)$$

where $B_1 = 1$ may also be defined as

$$\frac{P}{k_B T} = \frac{1}{v} \sum_{l=1}^{\infty} a_l(T) \left(\frac{\lambda^3}{v} \right)^{l-1} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} a_l(T) \left(\frac{\lambda^3}{v} \right)^l, \quad (5.38)$$

where $v = V/N$ is the specific volume per particle and divided by v . To relate the virial equation of state to the cluster integrals, we Taylor-expand z as a function of λ^3/v :

$$z = \sum_{w=1}^{\infty} c_w \left(\frac{\lambda^3}{v} \right)^w \quad (5.39)$$

(we excluded the $w = 0$ term without any loss of generality) and plug it back into Eq. (5.36) multiplied by λ^3/v :

$$\frac{N}{V} = \frac{1}{v} = \sum_{l=1}^{\infty} l \bar{b}_l \left[\sum_{w=1}^{\infty} c_w \left(\frac{\lambda^3}{v} \right)^w \right]^l. \quad (5.40)$$

The only nonzero term on the right-hand side is the $w = 1, l = 1$ term. The coefficient of this term must be 1 because λ^3/v appears on the left-hand side with a prefactor of 1. Thus $c_1 = 1/\bar{b}_1 = 1$. There are two $(\lambda^3/v)^2$ terms, one with $l = 1, w = 2$ and the other one with $l = 2, w = 1$. These two terms must cancel so that $1\bar{b}_1c_2 + 2\bar{b}_2c_1^2 = 0$. Since $c_1 = 1/\bar{b}_1 = 1$, this gives $c_2 = -2\bar{b}_2/\bar{b}_1^3$ etc.:

$$z = \frac{1}{\bar{b}_1} \left(\frac{\lambda^3}{v} \right) - \frac{2\bar{b}_2}{\bar{b}_1^3} \left(\frac{\lambda^3}{v} \right)^2 + \dots \quad (5.41)$$

Plugging this into Eq. (5.35) gives

$$\begin{aligned} \frac{P}{k_B T} = \frac{1}{\lambda^3} \left\{ \bar{b}_1 \left[\frac{1}{\bar{b}_1} \left(\frac{\lambda^3}{v} \right) - \frac{2\bar{b}_2}{\bar{b}_1^3} \left(\frac{\lambda^3}{v} \right)^2 + \dots \right] \right. \\ \left. + \bar{b}_2 \left[\frac{1}{\bar{b}_1} \left(\frac{\lambda^3}{v} \right) - \frac{2\bar{b}_2}{\bar{b}_1^3} \left(\frac{\lambda^3}{v} \right)^2 + \dots \right]^2 + \dots \right\}. \end{aligned} \quad (5.42)$$

Summing up all terms of same order in λ^3/v we find that

$$\begin{aligned} a_1 &= \bar{b}_1 = 1, \\ a_2 &= -\frac{\bar{b}_2}{\bar{b}_1^2} = -\bar{b}_2, \\ a_3 &= \dots = 4\bar{b}_2^2 - 2\bar{b}_3, \end{aligned} \quad (5.43)$$

etc.

It can be shown that the virial coefficients are proportional to the so-called *irreducible cluster integrals*:

$$a_l = -\frac{l-1}{l} \beta_{l-1}, \quad (5.44)$$

where

$$\beta_{l-1} = \frac{1}{(l-1)! \lambda^{3(l-1)} V} \times (\text{sum of all irreducible } l\text{-clusters}). \quad (5.45)$$

The irreducible cluster integrals are all graphs where any two circles are connected by at least 2 independent nonintersecting paths. Of the four 3-clusters [Eqs. (5.14)-(5.17)], only the loop is irreducible. Figure 5.3 shows irreducible clusters for $l = 2, 3$, and 4.

This framework can be used to derive various virial equations of state. Often this is done at the level of the second virial coefficient only, the most

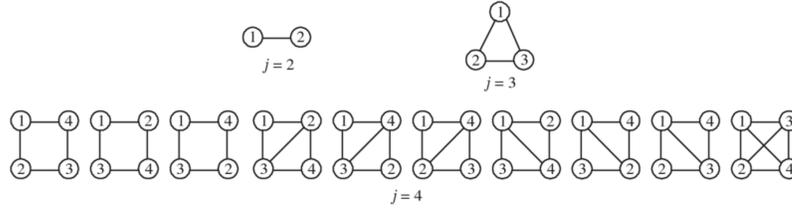


Figure 5.3: Irreducible clusters for $l = 2, 3$, and 4 **REFERENCE MISSING!**

remarkable result being that based on the Lennard-Jones pair interaction or its suitable simplification such as the hard-core repulsion of range σ combined by a long-range attraction of type $-u_0(\sigma/r)^6$. Here

$$a_2 = -\frac{2\pi}{\lambda^3} \left\{ -\int_0^\sigma r^2 dr - \int_\sigma^\infty \left[\exp\left(-\frac{u_0}{k_B T} \left(\frac{\sigma}{r}\right)^6\right) - 1 \right] r^2 dr \right\}. \quad (5.46)$$

This is easily evaluated at large temperatures where

$$a_2 = \frac{\pi\sigma^3}{14\lambda^3} \left(1 - \frac{u_0}{k_B T} \right), \quad (5.47)$$

which agrees with the functional form of the van der Waals equation of state at large volumes.

Another important example where the virial expansion proved very useful is the equation of state for a hard-sphere liquid. In hard spheres, the second virial coefficient reads

$$a_2 = \frac{2\pi\sigma^3}{3\lambda^3} \quad (5.48)$$

and the third virial coefficient, which is more difficult to evaluate, is

$$a_3 = \frac{5\pi\sigma^6}{18\lambda^6}. \quad (5.49)$$

The fourth virial coefficient can also be calculated and coefficients up to 10th order were computed numerically. It is convenient to express them in reduced units obtained by replacing specific volume in the virial expansion by the packing fraction

$$\eta = \frac{\pi\sigma^3}{6v}; \quad (5.50)$$

then the reduced $\tilde{a}_2 = 4$, $\tilde{a}_3 = 10$, $\tilde{a}_4 = 18.365$, $\tilde{a}_5 = 28.22$, $\tilde{a}_6 = 39.8$, $\tilde{a}_7 = 53.3$ etc. Carnahan and Starling have replaced these coefficients by integers

close to the actual values and proposed an analytical formula for them:

$$\tilde{a}_l = l^2 + 3l \quad (5.51)$$

and then assumed that the formula is valid for all l . Then the expansion can be summed and the result is

$$\frac{Pv}{k_B T} = \frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3}, \quad (5.52)$$

which agrees very well with the numerically obtained pressure within the whole range of packing fractions where the liquid state is stable, that is up to $\eta = 0.494$.

5.3 Second virial coefficient

The second virial coefficient representing interactions of pairs of particles can be calculated exactly within the quantum-mechanical formalism. To this end, first note that the grand-canonical partition function

$$\mathcal{Q}(z, V, T) = \sum_{N=0}^{\infty} Q_N(V, T) z^N = \sum_{N=0}^{\infty} \frac{Z_N(V, T)}{N!} \left(\frac{z}{\lambda^3}\right)^N. \quad (5.53)$$

Since $P/(k_B T) = V^{-1} \ln \mathcal{Q}$ we have

$$\begin{aligned} \frac{P}{k_B T} &= \frac{1}{V} \ln \left(1 + \frac{Z_1(V, T)}{1!} \left(\frac{z}{\lambda^3}\right)^1 + \frac{Z_2(V, T)}{2!} \left(\frac{z}{\lambda^3}\right)^2 \right. \\ &\quad \left. + \frac{Z_3(V, T)}{6!} \left(\frac{z}{\lambda^3}\right)^3 + \dots \right). \end{aligned} \quad (5.54)$$

By Taylor-expanding this expression

$$\begin{aligned} \frac{P}{k_B T} &= \frac{1}{V} \left[Z_1 \left(\frac{z}{\lambda^3}\right)^1 + \frac{1}{2} (Z_2 - Z_1^2) \left(\frac{z}{\lambda^3}\right)^2 \right. \\ &\quad \left. + \frac{1}{6} (Z_3 - 3Z_2 Z_1 + 2Z_1^3) \left(\frac{z}{\lambda^3}\right)^3 + \dots \right] \end{aligned} \quad (5.55)$$

and comparing it with Eq. (5.35) we find that

$$\begin{aligned} \bar{b}_1 &= \frac{1}{V} Z_1 = 1 \\ \bar{b}_2 &= \frac{1}{2! \lambda^3 V} (Z_2 - Z_1^2) \\ \bar{b}_3 &= \frac{1}{2! \lambda^6 V} (Z_3 - 3Z_2 Z_1 + 2Z_1^3) \\ &\dots \end{aligned} \quad (5.56)$$

Let us now focus on \bar{b}_2 . Recall that the configuration integral is equal to V both in the interacting and in the non-interacting system [$Z_1 = Z_2^{(0)} = V$] so that the difference of the corresponding second virial coefficients is

$$\bar{b}_2 - \bar{b}_2^{(0)} = \frac{1}{2\lambda^3 V} [Z_2 - Z_2^{(0)}] \quad (5.57)$$

and since

$$Q_2 = \frac{1}{2\lambda^6} Z_2 \quad (5.58)$$

we have

$$\bar{b}_2 - \bar{b}_2^{(0)} = \frac{\lambda^3}{V} [Q_2 - Q_2^{(0)}] = \frac{\lambda^3}{V} \text{tr} \left(\exp(-\beta \hat{H}_2) - \exp(-\beta \hat{H}_2^{(0)}) \right). \quad (5.59)$$

To evaluate the trace we need to solve the Schrödinger equation for two particles

$$\hat{H}_2 \psi_\alpha(\mathbf{r}_1, \mathbf{r}_2) = E_\alpha \psi_\alpha(\mathbf{r}_1, \mathbf{r}_2), \quad (5.60)$$

where

$$\hat{H}_2 = -\frac{\hbar^2}{2m} (\nabla_1^2 + \nabla_2^2) + u(r_{12}). \quad (5.61)$$

This is a two-body problem; the center of mass $\mathbf{R} = (\mathbf{r}_1 + \mathbf{r}_2)/2$ moves at a steady velocity, its momentum being \mathbf{P} , and the particles oscillate about it, their relative coordinate being $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$:

$$\psi_\alpha(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{V^{1/2}} \exp(i\mathbf{P}_j \cdot \mathbf{R}) \psi_n(\mathbf{r}). \quad (5.62)$$

The total energy thus consists of the translational kinetic energy and the energy associated with relative motion:

$$E_\alpha = \frac{\mathbf{P}_j^2}{2(2m)} + \varepsilon_n. \quad (5.63)$$

(Note that since the particles are identical, their combined mass is $2m$ and their reduced mass is $m/2$.) The relative motion is described by

$$\left[-\frac{\hbar^2}{2m} \nabla_{\mathbf{r}}^2 + u(r) \right] \psi_n(\mathbf{r}) = \varepsilon_n \psi_n(\mathbf{r}). \quad (5.64)$$

Given that the $1/V^{1/2}$ in Eq. (5.62) takes care of the normalization of the wavefunction describing translational motion, $\psi_n(\mathbf{r})$ must be normalized it-

self. In the trace

$$\begin{aligned}\bar{b}_2 - \bar{b}_2^{(0)} &= \frac{\lambda^3}{V} \sum_{\alpha} \left[\exp(-\beta \hat{H}_2) - \exp(-\beta \hat{H}_2^{(0)}) \right] \\ &= \frac{\lambda^3}{V} \sum_j \exp\left(-\beta \frac{\mathbf{P}_j^2}{4m}\right) \sum_n \left[\exp(-\beta \varepsilon_n) - \exp(-\beta \varepsilon_n^{(0)}) \right]\end{aligned}\quad (5.65)$$

we can evaluate the first sum by replacing it by the integral over \mathbf{P} :

$$\sum_j \exp\left(-\beta \frac{\mathbf{P}_j^2}{4m}\right) = \frac{4\pi V}{h^3} \int_0^\infty \exp\left(-\beta \frac{\mathbf{P}_j^2}{4m}\right) P^2 dP = \frac{V}{\lambda^3} 2^{3/2}; \quad (5.66)$$

the additional factor of $2^{3/2}$ arises because the mass of two particles is twice the mass of a single particle, and λ is defined for a single particle. Thus

$$\bar{b}_2 - \bar{b}_2^{(0)} = 2^{3/2} \sum_n \left[\exp(-\beta \varepsilon_n) - \exp(-\beta \varepsilon_n^{(0)}) \right]. \quad (5.67)$$

To evaluate this difference, we need to know the spectrum of the interacting and the non-interacting pair and the corresponding densities of states; this is required so that the sums can be replaced by integrals. The spectrum of the non-interacting pair contains only unbound states

$$\varepsilon_n^{(0)} = \frac{\hbar^2 k^2}{m} \quad (5.68)$$

with the standard density of states $g^{(0)}(k)$. Unbound states are also present in the interacting system, albeit with a different density of states $g(k)$. By taking this into account, we have

$$\bar{b}_2 - \bar{b}_2^{(0)} = 2^{3/2} \sum_B \exp(-\beta \varepsilon_B) + 2^{3/2} \int_0^\infty \exp\left(-\beta \frac{\hbar^2 k^2}{m}\right) \left[g(k) - g^{(0)}(k) \right] dk. \quad (5.69)$$

Here ε_B are the bound states, and the sum is on them.

Now we turn to the density of states. We assume that the potential is central, $u = u(r)$, and thus the wavefunction may be factorized as follows:

$$\psi_{klm}(\mathbf{r}) = A_{klm} \frac{\chi_{kl}(r)}{r} Y_{lm}(\theta, \phi) \quad (5.70)$$

so that the radial Schrödinger equation is

$$-\frac{\hbar^2}{mr^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) \frac{\chi_{kl}(r)}{r} + \left[\frac{\hbar^2 l(l+1)}{mr^2} + u(r) \right] \frac{\chi_{kl}(r)}{r} = \varepsilon \frac{\chi_{kl}(r)}{r} \quad (5.71)$$

or

$$-\frac{\hbar^2}{m} \frac{1}{r} \frac{\partial^2 \chi_{kl}(r)}{\partial r^2} + \left[\frac{\hbar^2 l(l+1)}{mr^2} + u(r) \right] \frac{\chi_{kl}(r)}{r} = \varepsilon \frac{\chi_{kl}(r)}{r} \quad (5.72)$$

and

$$-\frac{\partial^2 \chi_{kl}(r)}{\partial r^2} + \left[\frac{l(l+1)}{r^2} + \frac{m}{\hbar^2 k^2} u(r) \right] \chi_{kl}(r) = k^2 \chi_{kl}(r), \quad (5.73)$$

where we have inserted the free particle energy and divided by \hbar^2/m . At large r , the centrifugal term is small and we have

$$\chi_{kl}(r) \approx \sin(kr - \pi l/2 + \eta_l(k)). \quad (5.74)$$

This asymptotic behavior including $\pi l/2$ follows from the expansion of the spherical Bessel function $j_l(kr) = \chi_{kl}$; the phase shift of $\eta_l(k)$ is due to the potential.

The boundary condition for χ_{kl} is that at some large $r = R_0 \rightarrow \infty$ be equal to 0. We now use this boundary condition to determine the density of states. To this end, note that the phases corresponding to two consecutive states n and $n+1$ differ by π and so

$$\left[R_0 + \frac{d\eta_l(k)}{dk} \right] \Delta k = \pi \quad (5.75)$$

so that

$$g_l(k) = \frac{2l+1}{\Delta k} = \frac{2l+1}{\pi} \left[R_0 + \frac{d\eta_l(k)}{dk} \right], \quad (5.76)$$

where the degeneracy of orbital states has been taken into account by the factor of $2l+1$.

At this stage, we have to distinguish between bosons and fermions with an even $\psi(\mathbf{r}) = \psi(-\mathbf{r})$ and an odd $\psi(\mathbf{r}) = -\psi(-\mathbf{r})$ wavefunction, respectively. [Note that \mathbf{r} is a relative coordinate of the two particles so that inversion really corresponds to particle exchange.] The parity of the wavefunction is given by l so that the k -dependent density of states obtained after summing over all l must include only $l = 0, 2, 4, \dots$ for bosons and only $l = 1, 3, 5, \dots$ for fermions, which is indicated by the prime:

$$g(k) = \sum_l' g_l(k) = \sum_l' \frac{2l+1}{\pi} \left[R_0 + \frac{d\eta_l(k)}{dk} \right], \quad (5.77)$$

In the non-interacting system, $u(r) = 0$ and thus the phase shift of $\eta_l(k)$ in Eq. (5.74) vanishes:

$$g^{(0)}(k) = \sum_l' \frac{2l+1}{\pi} R_0, \quad (5.78)$$

which means that the *difference* in the density of states in Eq. (5.69) reads

$$g(k) - g^{(0)}(k) = \sum_l' \frac{2l+1}{\pi} \frac{d\eta_l(k)}{dk}. \quad (5.79)$$

This leads to the final form of Eq. (5.69):

$$\begin{aligned} \bar{b}_2 - \bar{b}_2^{(0)} &= 2^{3/2} \sum_B \exp(-\beta\varepsilon_B) \\ &+ \frac{2^{3/2}}{\pi} \sum_l' (2l+1) \int_0^\infty \exp\left(-\beta \frac{\hbar^2 k^2}{m}\right) \frac{d\eta_l(k)}{dk} dk. \end{aligned} \quad (5.80)$$

Both the first and the second term depend on the pair potential as they should; the former through the spectrum of bound states and the latter through the phase shift $\eta_l(k)$.

Non-interacting systems

Before proceeding, let us first determine the second virial coefficient for non-interacting bosons and fermions, which is entirely a quantum-mechanical effect. These can be readily obtained by comparing series expansions of $P/(k_B T)$ and N/V as functions of z for both types of particles. In case of bosons, recall that

$$\frac{P}{k_B T} = \frac{1}{\lambda^3} g_{5/2}(z) = \frac{1}{\lambda^3} \left(z + \frac{z^2}{2^{5/2}} + \dots \right) \quad (5.81)$$

and

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z) = \frac{1}{\lambda^3} \left(z + \frac{z^2}{2^{3/2}} + \dots \right); \quad (5.82)$$

in the latter expression, we assumed that the system is far from condensation so that $N - N_0$ in Eq. (4.21) can be replaced by N . Again we resort to series inversion and write z as a Taylor series of powers of $N\lambda^3/V = \lambda^3/v$. After determining the coefficients from Eq. (5.82) we insert $z(\lambda^3/v)$ into Eq. (5.81) to find that for bosons,

$$\frac{PV}{Nk_B T} = 1 - \frac{1}{2^{5/2}} \frac{\lambda^3}{v} + \dots \quad (5.83)$$

At the level of the second virial coefficient, this result may also be obtained by truncating the two series after z^2 terms, solving Eq. (5.82) as a quadratic

equation for z , inserting it in Eq. (5.81) and expanding the result as a Taylor series.

The negative second virial coefficient $a_{2\text{BE}}^{(0)} = -1/2^{5/2} = -\bar{b}_2^{(0)}$ states that at densities where λ^3/v is sufficiently large (that is, comparable of larger than 1), bosons experience an attractive interaction.

For fermions, we proceed along the same steps, using Eqs. (4.75) and (4.76) to find that here

$$\frac{PV}{Nk_B T} = 1 + \frac{1}{2^{5/2}} \frac{\lambda^3}{v} + \dots \quad (5.84)$$

and $a_{2\text{FD}}^{(0)} = +1/2^{5/2}$ consistent with the repulsion between fermions.

Note that $a_2^{(0)} = -\bar{b}_2^{(0)}$ so that

$$\bar{b}_2^{(0)} = \pm \frac{1}{2^{5/2}}, \quad (5.85)$$

where $+$ and $-$ correspond to bosons and to fermions, respectively.

This is consistent with the behavior of the two-particle density matrix, and may be calculated from the statistical potential $u_{\text{eff}}(r)$ extracted from the density matrix Eq. (3.47)

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \hat{\rho} | \mathbf{r}_1, \mathbf{r}_2 \rangle = \frac{1}{V^2} \left[1 \pm \exp \left(-\frac{2\pi r^2}{\lambda^2} \right) \right]. \quad (5.86)$$

For classical particles, $\langle \mathbf{r}_1, \mathbf{r}_2 | \hat{\rho} | \mathbf{r}_1, \mathbf{r}_2 \rangle = 1/V^2$ and so the expression in square brackets may be viewed as $\exp(-u_{\text{eff}}(r)/(k_B T))$. The effective statistical potentials for bosons and fermions are plotted in Fig.5.4.

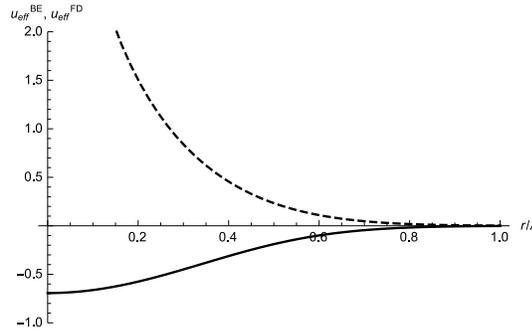


Figure 5.4: Effective statistical pair potential between bosons (solid line) and fermions (dashed line) vs. reduced distance r/λ .

Evidently the non-interacting Mayer functions read

$$f(r) = \exp(-u_{\text{eff}}(r)/(k_B T)) - 1 = \pm \exp\left(-\frac{2\pi r^2}{\lambda^2}\right) \quad (5.87)$$

where + and - correspond to bosons and to fermions, respectively. Now

$$\bar{b}_2 = \frac{2\pi}{\lambda^3} \int_0^\infty f(r) r^2 dr = \pm \frac{2\pi}{\lambda^3} \int_0^\infty \exp\left(-\frac{2\pi r^2}{\lambda^2}\right) r^2 dr = \pm \frac{1}{2^{5/2}}. \quad (5.88)$$

Hard-core repulsion

In case of hard-core repulsion, $\chi_{kl}(r)$ must be zero for all $r < \sigma$. In this case, the scattering phase shift is

$$\eta_l(k) = \tan^{-1} \frac{j_l(k\sigma)}{n_l(k\sigma)}, \quad (5.89)$$

where j_l and n_l are spherical Bessel and Neumann functions, respectively. This gives

$$\begin{aligned} \eta_0(k) &= -k\sigma \\ \eta_1(k) &= -[k\sigma - \tan^{-1}(k\sigma)] = -\frac{(k\sigma)^3}{3} + \frac{(k\sigma)^5}{5} + \dots \\ \eta_2(k) &= -\left[k\sigma - \tan^{-1}\left(\frac{3k\sigma}{3 - (k\sigma)^2}\right)\right] = -\frac{(k\sigma)^5}{45} + \dots \\ &\dots \end{aligned} \quad (5.91)$$

Now we return to Eq. (5.80). We note that for a repulsive potential, there exist no bound states and so first term in Eq. (5.80) is absent. The integral can be integrated by parts. We immediately recognize that the uv term is zero because $\eta_l(0) = 0$ (the lower boundary) and $\exp(-\beta\hbar^2 k^2/m)|_{k \rightarrow \infty} = 0$. Thus

$$\begin{aligned} \bar{b}_2 - \bar{b}_2^{(0)} &= \frac{2^{5/2} \beta \hbar^2}{\pi m} \sum_l' (2l+1) \int_0^\infty \exp(-\beta\hbar^2 k^2/m) \eta_l(k) k dk \\ &= 2^{7/2} \lambda^2 \sum_l' (2l+1) \int_0^\infty \exp(-\beta\hbar^2 k^2/m) \eta_l(k) k dk. \end{aligned} \quad (5.92)$$

Finally, we obtain the second virial coefficients $a_2 = -\bar{b}_2$ for hard-core bosons

$$a_2^{\text{BE}} = -\frac{1}{2^{5/2}} + 2 \left(\frac{\sigma}{\lambda}\right)^1 + \frac{10\pi^2}{3} \left(\frac{\sigma}{\lambda}\right)^5 - \dots \quad (5.93)$$

and for hard-core fermions

$$a_2^{\text{FD}} = +\frac{1}{2^{5/2}} + 6\pi \left(\frac{\sigma}{\lambda}\right)^3 - 18\pi^2 \left(\frac{\sigma}{\lambda}\right)^5 - \dots \quad (5.94)$$

Recall that for classical hard spheres

$$a_2^{\text{MB}} = \frac{2\pi}{3} \left(\frac{\sigma}{\lambda}\right)^3. \quad (5.95)$$

The two coefficients differ at small σ/λ but become increasingly more similar as σ/λ is increased. The behavior shown in Fig. 5.5 should be viewed *cum grano salis* because we have only plotted the three lowest-order terms included in Eqs. (5.94) and (5.95). Qualitatively, we expect that a_2 must increase monotonically both in bosons and in fermions so that the results presented in Fig. 5.5 are probably valid only up to $\sigma/\lambda \approx 0.3$.

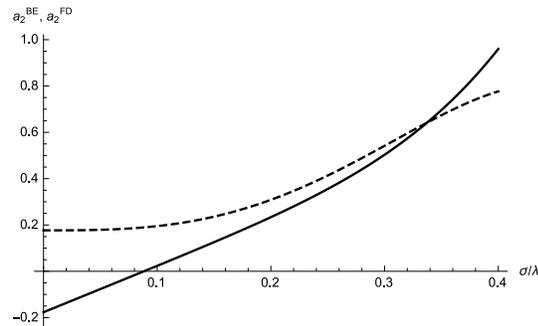


Figure 5.5: Second virial coefficient for the hard-core bosons (solid line) and hard-core fermions (dashed line) vs. reduced hard-core diameter σ/λ .

An interesting feature of Fig. 5.5 is the change of sign of the virial coefficient for bosons, which is negative at small σ/λ and positive at large σ/λ . This means that for large enough σ , the attractive statistical potential between bosons is overpowered by the hard-core repulsion, and we may expect that in this regime Bose-Einstein condensation does not take place. In a similar vein, one would expect that an attractive potential between fermions can make the fermionic a_2 negative, and thus lead to condensation of some kind. This indeed takes place in a superconductor. However, the analysis of this transition is far more transparent in the case presented here because there exist no bound states.

Also plausible is that the bosonic and the fermionic virial coefficients approach each other at $\sigma/\lambda > 0.2$ or so, which suggests that at large σ/λ

the hard-core repulsion is dominant. Note that in this regime a_2 does not reduce to the classical value.

There exists a quantum-mechanical variant of the classical cluster expansion which eventually leads to formally identical expressions for pressure and number of particles except that the cluster integrals \bar{b}_l are considerably more complicated and that the $\bar{b}_{l \geq 2}$ are non-zero even in the case of non-interacting particles — as illustrated by the above example of hard-core repulsion.

5.4 Interacting quantum systems

Understanding of interactions in quantum-mechanical systems is very important as it leads to a range of fascinating phenomena, some of which are of practical importance:

- **Ultracold atomic BE condensates** Here a cold atomic gas such as is confined by a magnetic trap, and one needs to account for the interparticle interactions as well as the effect of the external potential.
- **Excitations in liquid helium** include rotons, low-lying excitations at a finite wavevector.
- **Fermi liquid** denotes a model of interacting fermions and corresponds to a normal metal.
- **Fermi condensates** In the presence of attraction between fermions, they may form bosonic pairs and then undergo condensation. Examples of this process include superconductivity, superfluidity in helium 3, and condensation in ultracold atomic Fermi gases.

These phenomena are more readily studied using the method of second quantization rather than in terms of cluster expansion. This approach is discussed in the advanced quantum mechanics course.

Chapter 6

Classical density functional theory

The standard way of studying quantum-mechanical problems is by solving the Schrödinger equation. In a many body system this is generally a difficult task, and usually a suitable approximation is made. Once the wavefunction is calculated or computed, one can then determine the spatial profile of the electron density $\rho(\mathbf{r})$. An alternative way of looking at the problem is the Thomas-Fermi theory where the energy of a continuous distribution of electrons consisting of the kinetic term, the electrostatic interaction, and the interaction with external field is spelled out as a functional of $\rho(\mathbf{r})$

$$E[\rho(\mathbf{r})] = \underbrace{C_F \int \rho^{5/3}(\mathbf{r}) d\mathbf{r}}_{\text{kinetic energy}} + \underbrace{e_0 \int \rho(\mathbf{r}) V_{\text{ext}} d\mathbf{r}}_{\text{electron-nucleus attraction}} + \underbrace{\frac{e_0^2}{8\pi\epsilon_0} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'}_{\text{electron-electron repulsion}} \quad (6.1)$$

Here

$$C_F = \frac{3h^2}{10m_e} \left(\frac{3}{8\pi} \right)^{2/3} \quad (6.2)$$

and V_{ext} is the Coulombic electron-nucleus interaction $\propto Z/r$. — The origin of the kinetic term is as follows: We assume that the number of electrons in a volume V is N and thus at $T = 0$, the Fermi level is determined by the condition

$$\frac{2 \times 4\pi V P_F^3}{3h^3} = N \quad (6.3)$$

or

$$P_F = \left(\frac{3h^3 \rho}{8\pi} \right)^{1/3} \quad (6.4)$$

so that the kinetic energy reads

$$E_{\text{kin}} = \int_0^{P_F} \frac{P^2}{2m} \frac{2 \times 4\pi V P^2 dP}{3h^3} \propto P_F^5 \propto \rho^{5/3}. \quad (6.5)$$

Once the energy $E[\rho(\mathbf{r})]$ is known, the equilibrium density profile is found variationally at constant total number of electrons $N = \int \rho(\mathbf{r}) d\mathbf{r}$.

This view of treating many-body quantum-mechanical problems has later evolved into the density functional theory, primarily due to Kohn. The theoretical basis of the density-functional theory is in the Hohenberg-Kohn theorems, the first of which states that the properties of the many-body electron system are uniquely determined by the electron density $\rho(\mathbf{r})$.

Here we will outline the classical density functional theory used to study the structure of non-uniform fluids [2]. The main object of this theory is the time and ensemble average of the particle density

$$\rho(\mathbf{r}) = \lim_{\Delta V \rightarrow 0} \frac{\Delta N(\mathbf{r})}{\Delta V}. \quad (6.6)$$

(Also possible is a generalization of this approach to dynamical problems.)

The Hamiltonian of a N -body system reads

$$H_N = \sum_i \frac{\mathbf{p}_i^2}{2m} + U(\mathbf{r}_1, \mathbf{r}_2, \dots) + \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i). \quad (6.7)$$

The grand-canonical partition function expressed in terms of the chemical potential $\mu = k_B T \ln z$ rather than in terms of fugacity z is

$$\mathcal{Q}(\mu, V, T) = \sum_{N=0}^{\infty} \frac{1}{N! h^{3N}} \int d\mathbf{r}^N d\mathbf{p}^N \exp(-\beta(H_N - \mu N)). \quad (6.8)$$

By definition, the expectation value of any operator is given by

$$\begin{aligned} \langle \hat{O}(\mathbf{r}^N, \mathbf{p}^N) \rangle &= \frac{1}{\mathcal{Q}(\mu, V, T)} \sum_{N=0}^{\infty} \frac{1}{N! h^{3N}} \\ &\quad \times \int d\mathbf{r}^N d\mathbf{p}^N \hat{O}(\mathbf{r}^N, \mathbf{p}^N) \exp(-\beta(H_N - \mu N)) \\ &= \underbrace{\sum_{N=0}^{\infty} \frac{1}{N! h^{3N}} \int d\mathbf{r}^N d\mathbf{p}^N \hat{O}(\mathbf{r}^N, \mathbf{p}^N)}_{=\text{tr}} \underbrace{\frac{\exp(-\beta(H_N - \mu N))}{\mathcal{Q}(\mu, V, T)}}_{=f_0} \\ &= \text{tr} \left(f_0 \hat{O}(\mathbf{r}^N, \mathbf{p}^N) \right). \end{aligned} \quad (6.9)$$

Here f_0 is the *equilibrium* phase-space probability density. [This is just another way of spelling Eq. (3.20).]

The one-particle density operator is

$$\hat{\rho}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \quad (6.10)$$

and thus the one-particle density defined as the ensemble average reads

$$\rho(\mathbf{r}) = \langle \hat{\rho}(\mathbf{r}) \rangle = \text{tr} \left(f_0 \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right). \quad (6.11)$$

Given that the equilibrium phase-space probability density depends on both interparticle interactions $U(\mathbf{r}_1, \mathbf{r}_2 \dots)$ and the external potential $V_{\text{ext}}(\mathbf{r})$, so does the equilibrium one-particle density $\rho_0(\mathbf{r})$. The power of density-functional theory is in the fact that the converse also holds. The external potential and the interparticle interactions are uniquely determined by a given equilibrium $\rho_0(\mathbf{r})$. In turn, this means that the grand potential

$$q = -k_B T \ln \mathcal{Q} \quad (6.12)$$

may be viewed as a functional of $\rho(\mathbf{r})$ to be varied so as to find the equilibrium state.

6.1 Grand potential as a functional

To show this, let us begin with thermodynamics. Note that $q = q(\mu, V, T) = F - \mu N = U - TS - \mu \langle N \rangle$ where $U = \langle H_N \rangle$ and $S = -k_B \langle \ln f_0 \rangle$ (the Gibbs formula). Thus

$$q[f_0] = \text{tr} \left(f_0 (H_N - \mu N + \beta^{-1} \ln f_0) \right). \quad (6.13)$$

This expression depends on the equilibrium phase-space probability density f_0 which contains the partition function. If the partition function can be calculated then q is known.

Now we consider the grand potential in Eq. (6.13) as a functional of an *unknown* probability density $f(\mathbf{r}^N, \mathbf{p}^N)$ such that $\text{tr} f = 1$:

$$q[f] = \text{tr} \left(f (H_N - \mu N + \beta^{-1} \ln f) \right). \quad (6.14)$$

We subtract Eq. (6.14) from Eq. (6.13):

$$\begin{aligned}
q[f] - q[f_0] &= \text{tr} \left(f (H_N - \mu N) \right) - \text{tr} \left(f_0 (H_N - \mu N) \right) \\
&\quad + \beta^{-1} \text{tr} \left(f \ln f - f_0 \ln f_0 \right) \\
&= \text{tr} \left((f - f_0) (H_N - \mu N) \right) \\
&\quad + \beta^{-1} \text{tr} \left(f \ln f - f \ln f_0 \right) \\
&\quad + \beta^{-1} \text{tr} \left(\underbrace{f \ln f_0 - f_0 \ln f_0}_{=(f-f_0) \ln f_0} \right). \tag{6.15}
\end{aligned}$$

[Since $\text{tr} f = \text{tr} f_0 = 1$ we could already have cancelled the $\text{tr}(f\mu N)$ and $\text{tr}(f_0\mu N)$ terms.] Now note that

$$\ln f_0 = -\beta(H_N - \mu N) - \ln \mathcal{Q}(\mu, V, T) \tag{6.16}$$

so that the last term in Eq. (6.15) reads

$$\begin{aligned}
\beta^{-1} \text{tr} \left((f - f_0) \ln f_0 \right) &= -\text{tr} \left((f - f_0) (H_N - \mu N) \right) \\
&\quad - \beta^{-1} \text{tr} \left((f - f_0) \ln \mathcal{Q}(\mu, V, T) \right). \tag{6.17}
\end{aligned}$$

The first term on the right-hand side cancels with the first term in Eq. (6.15) and the last term is zero because $\ln \mathcal{Q}(\mu, V, T)$ is unaffected by the trace and $\text{tr} f = \text{tr} f_0 = 1$. Thus we find that

$$q[f] = q[f_0] + \beta^{-1} \text{tr} (f \ln f - f \ln f_0). \tag{6.18}$$

Now note that

$$\ln x \leq x - 1 \tag{6.19}$$

(this is known as the Gibbs inequality) or $-\ln x \geq 1 - x$ for any $x > 0$. We use this in $f \ln f - f \ln f_0 = -f \ln(f_0/f) \geq f(1 - f_0/f) = f - f_0$ to find that

$$q[f] \geq q[f_0] + \beta^{-1} \underbrace{\text{tr} (f - f_0)}_{=0} \tag{6.20}$$

and finally

$$q[f] \geq q[f_0]. \tag{6.21}$$

This means that the grand potential is minimized by the equilibrium probability density; conversely, the equilibrium probability density is determined by the minimum of the grand potential. Equality holds for $f = f_0$.

6.2 Uniqueness of density profile

To show that two different external potentials V_{ext} and V'_{ext} must give rise to two different one-particle densities $\rho_0(\mathbf{r})$, we use *reductio ad absurdum*. First we note that the two grand potentials differ by

$$q' - q = \int d\mathbf{r}^N [\rho'_0(\mathbf{r})V'_{\text{ext}}(\mathbf{r}) - \rho_0(\mathbf{r})V_{\text{ext}}(\mathbf{r})] \quad (6.22)$$

so that

$$q' = q + \int d\mathbf{r}^N [\rho'_0(\mathbf{r})V'_{\text{ext}}(\mathbf{r}) - \rho_0(\mathbf{r})V_{\text{ext}}(\mathbf{r})]. \quad (6.23)$$

Now if the equilibrium ρ'_0 is replaced by ρ_0

$$q' < q + \int d\mathbf{r}^N \rho_0(\mathbf{r}) [V'_{\text{ext}}(\mathbf{r}) - V_{\text{ext}}(\mathbf{r})]. \quad (6.24)$$

At the same time, Eq (6.23) also means that

$$q = q' + \int d\mathbf{r}^N [\rho_0(\mathbf{r})V_{\text{ext}}(\mathbf{r}) - \rho'_0(\mathbf{r})V'_{\text{ext}}(\mathbf{r})] \quad (6.25)$$

As before, if we replace the equilibrium ρ_0 by ρ'_0

$$q < q' + \int d\mathbf{r}^N \rho'_0(\mathbf{r}) [V'_{\text{ext}}(\mathbf{r}) - V_{\text{ext}}(\mathbf{r})] \quad (6.26)$$

and if the two equilibrium densities were the same then

$$q < q' + \int d\mathbf{r}^N \rho_0(\mathbf{r}) [V'_{\text{ext}}(\mathbf{r}) - V_{\text{ext}}(\mathbf{r})]. \quad (6.27)$$

We now compare Eqs. (6.24) and (6.27) and see that they contradict each other — as can be verified by adding them, which gives $q' + q < q + q'$ because the integrals cancel.

6.3 Intrinsic free energy

By combining the internal energy and the entropic term in Eq. (6.13), we find that the grand potential can be written as

$$q[\rho_0(\mathbf{r})] = F[\rho_0(\mathbf{r})] + \int d\mathbf{r}^N [V_{\text{ext}}(\mathbf{r}) - \mu] \rho_0(\mathbf{r}), \quad (6.28)$$

where we employed the fact that $\int d\mathbf{r}^N \rho_0(\mathbf{r}) = N$ and where $F[\rho_0(\mathbf{r})]$ is the *intrinsic free energy*. In equilibrium, the functional derivative of q with respect to $\rho(\mathbf{r})$ must vanish:

$$\left. \frac{\delta q[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right|_{\rho(\mathbf{r})=\rho_0(\mathbf{r})} = 0 \quad (6.29)$$

or explicitly

$$\left. \frac{\delta F[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right|_{\rho(\mathbf{r})=\rho_0(\mathbf{r})} = \mu - V_{\text{ext}}(\mathbf{r}). \quad (6.30)$$

This equation defines the intrinsic functional $F[\rho]$ for a given interparticle interaction through an arbitrary external potential. This is even more evident if written as

$$\delta F[\rho(\mathbf{r})] = [\mu - V_{\text{ext}}(\mathbf{r})] \delta \rho(\mathbf{r}). \quad (6.31)$$

The intrinsic free energy $F[\rho(\mathbf{r})]$ is the main ingredient of the density-functional theory. For non-interacting particles (i.e., classical ideal gas),

$$F_{\text{id}}[\rho] = k_B T \int d^3r \rho(\mathbf{r}) [\ln(\rho(\mathbf{r})\Lambda^3) - 1], \quad (6.32)$$

which in a uniform system where $\rho(\mathbf{r}) = \rho$ reduces to $Nk_B T [\ln(\rho\Lambda^3) - 1]$ consistent with $\exp(-\beta F_{\text{id}}) = V^N / (N! \Lambda^{3N})$ for a uniform ideal gas. In an ideal gas, Eq. (6.30) immediately gives

$$\rho(\mathbf{r}) = \frac{1}{\Lambda^3} \exp(\beta\mu) \exp(-\beta V_{\text{ext}}(\mathbf{r})). \quad (6.33)$$

This is a generalized barometric formula.

In interacting systems, $F[\rho(\mathbf{r})]$ is often divided into the ideal and the excess term

$$F[\rho] = F_{\text{id}}[\rho] + F_{\text{ex}}[\rho]. \quad (6.34)$$

The excess free energy can be used to generate direct correlation functions of n -th order

$$c^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = - \frac{\delta^n \beta F_{\text{ex}}}{\delta \rho(\mathbf{r}_1) \delta \rho(\mathbf{r}_1) \dots \delta \rho(\mathbf{r}_n)}. \quad (6.35)$$

In ideal gas, $F_{\text{ex}} = 0$ and all correlation functions vanish. If these derivatives are known, then we may calculate the free energy using the expansion around

a suitable reference density field $\rho_i(\mathbf{r})$:

$$F[\rho] = F[\rho_i] + \sum_{n=1}^{\infty} \frac{1}{n!} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \int d\mathbf{r}_n \frac{\delta^n F_{\text{ex}}}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_1)\dots\delta\rho(\mathbf{r}_n)} \times \Delta\rho(\mathbf{r}_1)\Delta\rho(\mathbf{r}_2)\dots\Delta\rho(\mathbf{r}_n), \quad (6.36)$$

where $\Delta\rho(\mathbf{r}_k) = \rho(\mathbf{r}_k) - \rho_i(\mathbf{r}_k)$. This expansion is useful if truncated, and this usually done by disregarding all $n > 2$ terms.

6.4 Ornstein-Zernike equation

We first introduce the (dimensionless) extended external field

$$u(\mathbf{r}) = \beta[\mu - V_{\text{ext}}(\mathbf{r})] \quad (6.37)$$

and split the Hamiltonian into \tilde{H}_N including the kinetic and interparticle terms and V_{ext} . Recall that the density operator is defined by Eq. (6.10) and thus

$$\mu N - \sum_i V_{\text{ext}} = \int \underbrace{(\mu - V_{\text{ext}})}_{=u(\mathbf{r})} \hat{\rho}(\mathbf{r}) d\mathbf{r}. \quad (6.38)$$

Then the grand partition function may be viewed as a functional of $u(\mathbf{r})$

$$\mathcal{Q}[u] = \sum_{N=0}^{\infty} \frac{1}{N!h^{3N}} \int d\mathbf{r}^N d\mathbf{p}^N \exp(-\beta\tilde{H}_N) \times \exp\left(\int u(\mathbf{r})\hat{\rho}(\mathbf{r})d\mathbf{r}\right). \quad (6.39)$$

As a result,

$$\frac{\delta \ln \mathcal{Q}[u]}{\delta u(\mathbf{r})} = \rho_0(\mathbf{r}), \quad (6.40)$$

which means that the equilibrium density is given to minus the functional derivative of the grand potential with respect to the extended external potential. Specifically, if we choose to observe the system at \mathbf{r}_1 ,

$$\frac{\delta \ln \mathcal{Q}[u]}{\delta u(\mathbf{r}_1)} = \rho_0(\mathbf{r}_1). \quad (6.41)$$

Now we would like to relate the derivative of the density at \mathbf{r}_1 with respect to the extended external potential applied at \mathbf{r}_3

$$\frac{\delta \rho_0(\mathbf{r}_1)}{\delta u(\mathbf{r}_3)}. \quad (6.42)$$

This derivative is equal to the density correlation function defined by

$$\begin{aligned} G(\mathbf{r}_1, \mathbf{r}_3) &= \langle \Delta\rho(\mathbf{r}_1)\Delta\rho(\mathbf{r}_3) \rangle \\ &= \langle [\hat{\rho}(\mathbf{r}_1) - \rho_0(\mathbf{r}_1)][\hat{\rho}(\mathbf{r}_3) - \rho_0(\mathbf{r}_3)] \rangle \\ &= \langle \hat{\rho}(\mathbf{r}_1)\hat{\rho}(\mathbf{r}_3) \rangle - \rho_0(\mathbf{r}_1)\rho_0(\mathbf{r}_3). \end{aligned} \quad (6.43)$$

The equality

$$\frac{\delta\rho_0(\mathbf{r}_1)}{\delta u(\mathbf{r}_3)} = \langle \Delta\rho(\mathbf{r}_1)\Delta\rho(\mathbf{r}_3) \rangle = G(\mathbf{r}_1, \mathbf{r}_3) \quad (6.44)$$

may be viewed as a generalization of the formula for the average square fluctuations of particle number, which is given by the derivative of the average particle number with respect to $\beta\mu$:

$$\frac{\partial\langle N \rangle}{\partial(\beta\mu)} = -\frac{\partial^2(\beta q)}{\partial(\beta\mu)^2} = \langle (\Delta N)^2 \rangle. \quad (6.45)$$

Equation (6.43) can be recast by using the two-particle density operator defined by

$$\hat{\rho}^{(2)}(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^N \sum_{j=1}^N \delta(\mathbf{r} - \mathbf{r}_i)\delta(\mathbf{r}' - \mathbf{r}_j) - \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)\delta(\mathbf{r}' - \mathbf{r}_i). \quad (6.46)$$

The first term on the right-hand side is product of single-particle densities appearing in Eq. (6.43) and so

$$\langle \hat{\rho}(\mathbf{r}_1)\hat{\rho}(\mathbf{r}_3) \rangle = \rho_0^{(2)}(\mathbf{r}_1, \mathbf{r}_3) + \delta(\mathbf{r}_1 - \mathbf{r}_3)\rho_0(\mathbf{r}_1). \quad (6.47)$$

Finally

$$\frac{\delta\rho_0(\mathbf{r}_1)}{\delta u(\mathbf{r}_3)} = \rho_0^{(2)}(\mathbf{r}_1, \mathbf{r}_3) + \delta(\mathbf{r}_1 - \mathbf{r}_3)\rho_0(\mathbf{r}_1) - \rho_0(\mathbf{r}_1)\rho_0(\mathbf{r}_3). \quad (6.48)$$

Recall that the extended external potential $u(\mathbf{r})$ also appears in Eq. (6.30) if the latter is multiplied by β . The intrinsic free energy F , in turn, consist of the ideal [Eq. (6.32)] and the excess term. The functional derivative of the ideal term multiplied by β is $\delta(\beta F_{\text{id}})/\delta\rho(\mathbf{r}) = \ln(\rho_0(\mathbf{r})\Lambda^3)$ and the functional derivative of βF_{ex} with respect to $\rho(\mathbf{r})$ is $c^{(1)}(\mathbf{r})$ by definition. Thus

$$u(\mathbf{r}) = \ln(\rho_0(\mathbf{r})\Lambda^3) - c^{(1)}(\mathbf{r}). \quad (6.49)$$

Now we differentiate this result at $\mathbf{r} = \mathbf{r}_3$ with respect to $\delta\rho_0(\mathbf{r}_2)$:

$$\frac{\delta u(\mathbf{r}_3)}{\delta\rho_0(\mathbf{r}_2)} = \frac{\delta(\mathbf{r}_2 - \mathbf{r}_3)}{\rho_0(\mathbf{r}_2)} - \frac{\delta c^{(1)}(\mathbf{r}_3)}{\delta\rho_0(\mathbf{r}_2)} = \frac{\delta(\mathbf{r}_2 - \mathbf{r}_3)}{\rho_0(\mathbf{r}_2)} - c^{(2)}(\mathbf{r}_2, \mathbf{r}_3). \quad (6.50)$$

The last step is to consider the following identity:

$$\frac{\delta\rho_0(\mathbf{r}_1)}{\delta\rho_0(\mathbf{r}_2)} = \int d\mathbf{r}_3 \frac{\delta\rho_0(\mathbf{r}_1)}{\delta u(\mathbf{r}_3)} \frac{\delta u(\mathbf{r}_3)}{\delta\rho_0(\mathbf{r}_2)} = \delta(\mathbf{r}_1 - \mathbf{r}_2). \quad (6.51)$$

On inserting Eqs. (6.48) and (6.50), we can derive the generalized Ornstein-Zernike equation for inhomogeneous systems. The result is less illuminating than the homogeneous version obtained by replacing $\rho_0(\mathbf{r}_i)$ by ρ in Eqs. (6.48) and (6.50), by expressing the two-particle density in terms of the radial distribution function $g(r)$ and then by the total correlation function $h(r) = g(r) - 1$

$$\rho_0^{(2)}(\mathbf{r}_1, \mathbf{r}_3) = \rho^2 [h(|\mathbf{r}_1 - \mathbf{r}_3|) + 1], \quad (6.52)$$

and by using $c(r)$ instead of $c^{(2)}(\mathbf{r}_2, \mathbf{r}_3)$. With these manipulations, Eq. (6.51) reads

$$\int d\mathbf{r}_3 \{ \rho^2 [h(|\mathbf{r}_1 - \mathbf{r}_3|) + 1] + \delta(\mathbf{r}_1 - \mathbf{r}_3)\rho - \rho^2 \} \times [\rho^{-1}\delta(\mathbf{r}_2 - \mathbf{r}_3) - c(|\mathbf{r}_2 - \mathbf{r}_3|)] = \delta(\mathbf{r}_1 - \mathbf{r}_2). \quad (6.53)$$

The ρ^2 terms in the curly bracket cancel and so

$$\begin{aligned} & \rho \int d\mathbf{r}_3 [\rho h(|\mathbf{r}_1 - \mathbf{r}_3|) + \delta(\mathbf{r}_1 - \mathbf{r}_3)] \times [\rho^{-1}\delta(\mathbf{r}_2 - \mathbf{r}_3) - c(|\mathbf{r}_2 - \mathbf{r}_3|)] \\ &= \rho \int d\mathbf{r}_3 [h(|\mathbf{r}_1 - \mathbf{r}_3|)\delta(\mathbf{r}_2 - \mathbf{r}_3) + \rho^{-1}\delta(\mathbf{r}_1 - \mathbf{r}_3)\delta(\mathbf{r}_2 - \mathbf{r}_3) \\ & \quad - h(|\mathbf{r}_1 - \mathbf{r}_3|)c(|\mathbf{r}_2 - \mathbf{r}_3|) + \delta(\mathbf{r}_1 - \mathbf{r}_3)c(|\mathbf{r}_2 - \mathbf{r}_3|)] \\ &= h(|\mathbf{r}_1 - \mathbf{r}_2|) + \delta(\mathbf{r}_1 - \mathbf{r}_2) \\ & \quad - \rho \int d\mathbf{r}_3 h(|\mathbf{r}_1 - \mathbf{r}_3|)c(|\mathbf{r}_2 - \mathbf{r}_3|) + c(|\mathbf{r}_2 - \mathbf{r}_1|) \\ &= \delta(\mathbf{r}_1 - \mathbf{r}_2). \end{aligned} \quad (6.54)$$

On rearrangement, this result reads

$$h(|\mathbf{r}_1 - \mathbf{r}_2|) = c(|\mathbf{r}_1 - \mathbf{r}_2|) + \rho \int d\mathbf{r}_3 h(|\mathbf{r}_1 - \mathbf{r}_3|)c(|\mathbf{r}_2 - \mathbf{r}_3|), \quad (6.55)$$

which is known as the Ornstein-Zernike equation.

The classical density-functional theory is used in many contexts, especially in various problems involving inhomogeneous systems. One example is the fluid-fluid interface and wetting in a binary mixture of polymer chains

where the chains are treated as ultrasoft particles [2]. Figure 6.1 shows the profiles of the two polymer species at a wetting wall favoring species 1 at different concentrations of species 2. The other example is melting and relative stability of soft-particle crystals. Here the solid-fluid phase transition can be estimated using the Lindemann criterion 6.2.

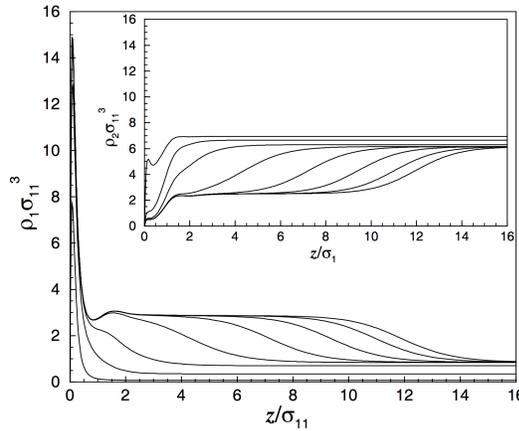


Figure 6.1: Wetting in a binary mixture of polymer chains: Density profiles of species 1 at a wall at different concentrations of species 2 decreasing from curve to curve from left to right. Inset shows concentration of species 2 [2].

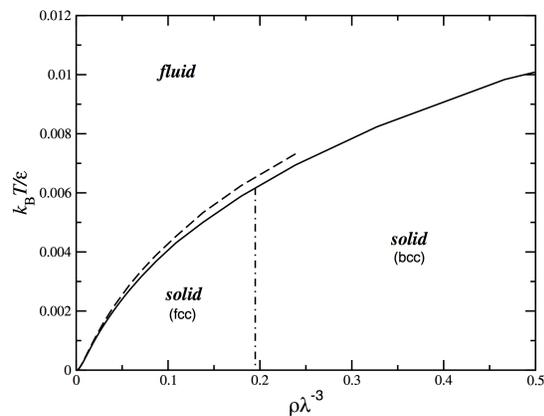


Figure 6.2: Phase diagram of a system of soft particles characterized by the Yukawa pair interaction; the melting line was computed using density-functional theory and the Lindemann criterion [2].

Chapter 7

Phase transitions

In the vicinity of the phase transition, the various response parameters and other quantities describing the phase transition consist of a regular part, which is analytic except for a jump in some cases, and of a singular part. The singular part typically varies as power laws of a suitable variable such as temperature, external field etc. Each of these power laws is characterized by a given critical exponent. The basic variables of interest are the

- order parameter m (magnetization \mathbf{M} in magnetic materials, the difference of densities of the liquid and the gas phase $\rho_l - \rho_g$ in the liquid-gas transition etc.),
- external field h (magnetic field strength \mathbf{H} in magnetic materials, external pressure relative to critical pressure $P - P_c$ in the liquid-gas transition etc.), and
- reduced temperature

$$t = \frac{T - T_c}{T_c}. \quad (7.1)$$

7.1 Critical exponents

For lexicographical reasons, we first introduce the exponent α describing the behavior of specific heat close to the critical temperature $t = 0$

$$c_V \propto |t|^{-\alpha}. \quad (7.2)$$

This definition includes the $-$ sign because c_V often has a peak at the transition. Generally, one can introduce two different exponents for $t > 0$ and $t < 0$ but they turn out to be the same.

The exponent β refers to the temperature dependence of the order parameter m in absence of external field:

$$m(h = 0, t < 0) \propto |t|^\beta. \quad (7.3)$$

This exponent can only be defined at $t < 0$ since $m = 0$ at $t > 0$. On the other hand, the divergence of generalized susceptibility (magnetic susceptibility in magnetic materials, isothermal compressibility in liquids and gases etc.

$$\chi = \left(\frac{\partial m}{\partial h} \right)_{t, h \rightarrow 0} \propto |t|^{-\gamma} \quad (7.4)$$

may in principle again be described by two different exponents γ and γ' on above and below the transition but again the exponents are identical.

The next exponent relates the order parameter and the external field at t , that is, the critical isotherm:

$$m(t = 0, h \rightarrow 0) \propto h^{1/\delta} \quad (7.5)$$

or

$$h \propto m^\delta(t = 0, h \rightarrow 0). \quad (7.6)$$

The power-law ansatz for the functional dependence of these quantities is suitable as long as the exponent is finite; but in case of specific heat, it is usually close to 0. A more general ansatz reads

$$f(t) \sim \frac{|t|^{-\lambda} - 1}{\lambda}, \quad (7.7)$$

where $t = (T - T_c)/T_c$. In this expression, "1" can be neglected if $\lambda > 0$ and so the power law is recovered but if $\lambda \rightarrow 0$ its limit reads

$$f(t) \sim \lim_{\lambda \rightarrow 0} \frac{|t|^{-\lambda} - 1}{\lambda} = -\ln |t|. \quad (7.8)$$

In this sense, $\lambda = 0$ represents a logarithmic rather than a power-law divergence.

We also introduce two other critical exponents pertaining to the two-point correlation function or the response function to be introduced a little later. The correlation length diverges as

$$\xi \propto |t|^{-\nu} \quad (7.9)$$

and the power-law decay of the response function is characterized by

$$r^{-p}, \text{ where } p = d - 2 + \eta. \quad (7.10)$$

system	α	β	γ	δ	ν	η
gas-liquid	0.1-0.2	0.32-0.35	1.2-1.3	4.6-5.0	0.57	—
binary mixt.	0.05-0.15	0.30-0.34	1.2-1.4	4.0-5.0	—	—
binary alloy	—	0.305	1.23-1.24	—	0.65	0.03
magnetic	0-0.2	0.30-0.36	1.0-1.4	4.2-4.8	0.62-0.68	0.03-0.15
ferroelectric	—	0.33-0.34	1.0-1.2	—	0.5-0.8	—
mean-field	0	1/2	1	3	1/2	0
Ising 2D	0	1/8	7/4	15	1	1/4
Ising 3D	1/8	5/16	5/4	5	0.64	0.05
Heisenberg 3D	1/16	5/16	21/16	—	0.7	0.04

Table 7.1: Critical exponents: Experimental data vs. exact models (mean-field, Ising 2D) vs. approximate models (Ising 3D, Heisenberg 3D) [1, 3].

Here d is the dimensionality of the system.

The experimental, theoretical, and computational data for the critical exponents are summarized in Table 7.1. Note the small variation of the values across different systems as well as the rather good agreement of the experimental values and those obtained in 3D Ising and Heisenberg models; the exactly solvable models (mean-field and 2D Ising) depart from the experimental values.

The universality of the exponents is nicely shown in Fig. 7.1 where the reduced coexisting densities of liquid and gas phase is plotted for 8 different fluids; the temperature range covered extends from $\approx T_c/2$ to T_c .

The critical exponents, and thus the nature of the phase transition, depends on a small number of parameters including

- the dimensionality of the embedding space (d),
- the number of components of the order parameter (n), and
- the range of microscopic interactions.

These parameters define the so-called *universality classes*: All systems and models with the same d , n , and range of interactions belong to the same universality class.

7.2 Landau theory

The Landau theory is a mean-field framework where one assumes that in the vicinity of the phase transition, the free energy density is an analytic

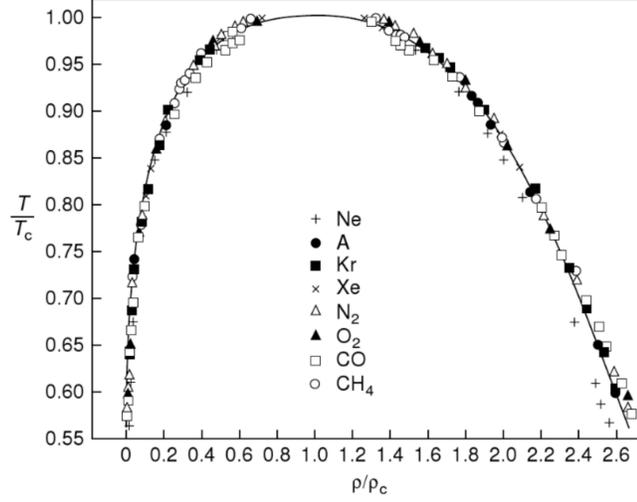


Figure 7.1: Reduced densities of coexisting liquid and gas phase, ρ_l/ρ_c and ρ_g/ρ_c , for different fluids [3].

function of the order parameter. Thus it can be expanded in powers of the order parameter $m(\mathbf{r})$:

$$f = a_0(t) + \frac{1}{2}a_2(t)m^2(\mathbf{r}) + \frac{1}{4}a_4(t)m^4(\mathbf{r}) + \dots \quad (7.11)$$

The linear term should not be included or else $m = 0$ would not be a possible solution; odd terms are absent because $m(\mathbf{r})$ and $-m(\mathbf{r})$ are physically equivalent states of the system. The constant term $a_0(t)$ represents the free energy of the high-temperature phase and the temperature dependence of $a_4(t)$ is usually not important as the main role of this term is to stabilize the theory. The key temperature dependence of the coefficients is

$$a_2(t) = at, \quad (7.12)$$

where a is a positive constant. The expansion is usually truncated after the fourth-order term.

Now we consider the Landau theory as a density functional theory, $m(\mathbf{r})$ playing the role of density $\rho(\mathbf{r})$, and allow for (i) inhomogeneities by including the elastic term and (ii) coupling to external field. In this case,

$$F[m(\mathbf{r})] = \int d\mathbf{r} \left\{ \frac{1}{2}atm^2(\mathbf{r}) + \frac{1}{4}a_4m^4(\mathbf{r}) + \frac{1}{2}b[\nabla m(\mathbf{r})]^2 - m(\mathbf{r})h(\mathbf{r}) \right\}. \quad (7.13)$$

In equilibrium, $\delta F[m(\mathbf{r})]/\delta m(\mathbf{r}) = 0$ or

$$atm(\mathbf{r}) + a_4m^3(\mathbf{r}) - b\nabla^2m(\mathbf{r}) - h(\mathbf{r}) = 0. \quad (7.14)$$

[We have disposed of the constant term $a_0(t)$.] Thus in a zero-field bulk homogeneous system,

$$m_0(t) = \begin{cases} \sqrt{a_2(t)/a_4} = \sqrt{-at/a_4}, & t < 0 \\ 0, & t > 0 \end{cases}. \quad (7.15)$$

The critical exponent $\beta = 1/2$. If $t > 0$ and we turn on a homogeneous $h \neq 0$ so that the equilibrium m is finite solely because of the presence of the field, then Eq. (7.14) reduces to $atm(h > 0) - h = 0$. Thus $m = h/(at)$ and the susceptibility is $1/(at)$ which means that the critical exponent $\gamma = 1$. In a similar fashion, one can show that $\gamma' = 1$ which demonstrates that the two exponents are the same. The exponent δ can be read from Eq. (7.14) immediately: At $t = 0$, $a_4m^3(h > 0) - h = 0$ so that $\delta = 3$. To calculate α , we insert Eq. (7.15) in the free energy to find that F contains two terms quadratic in t . Since $S = -dF/dT$ and $c_V \propto T(\partial S/\partial T)_V \propto -T(\partial^2 F/\partial T^2)_V = -(T/T_c^2)\partial^2 F/\partial t^2$, we find that $\alpha = 0$.

Response function

Now we introduce the response function

$$\varphi(\mathbf{r}) = \frac{\delta m(\mathbf{r})}{\delta h(\mathbf{r})} \quad (7.16)$$

which can be obtained from Eq. (7.14) by examining the response of the order parameter to a delta-function perturbation. In this case, the external field is described by

$$h(\mathbf{r}) = h_0(\mathbf{r}) + \underbrace{h_1\delta(\mathbf{r} - \mathbf{r}')}_{=\Delta h} \quad (7.17)$$

and the order parameter is

$$m(\mathbf{r}) = m_0(\mathbf{r}) + \underbrace{\int d\mathbf{r}' \varphi(\mathbf{r}, \mathbf{r}') \Delta h(\mathbf{r}, \mathbf{r}')}_{=\Delta m(\mathbf{r})} = m_0(\mathbf{r}) + h_1\varphi(\mathbf{r}, \mathbf{r}' = \mathbf{r}). \quad (7.18)$$

We insert these expressions in Eq. (7.14) and we expand the third-order term in $m(\mathbf{r})$ and keep only the $m_0^2\Delta m$ term linear in Δm . Note that all terms

containing only m_0 and h_0 cancel because m_0 represents the equilibrium state. On dividing by h_1 and b we have

$$\frac{at + 3a_4m_0^2}{b}\varphi - \nabla^2\varphi - b^{-1}\delta(\mathbf{r} - \mathbf{r}') = 0. \quad (7.19)$$

Now we assume that the equilibrium state is uniform so that m_0 is given by Eq. (7.15). In this case,

$$\frac{at + 3a_4m_0^2}{b} = \begin{cases} -2at/b, & t < 1 \\ at/b, & t > 1 \end{cases}. \quad (7.20)$$

This allows us to define the correlation length ξ by

$$\xi(t) = \begin{cases} \sqrt{-b/(2at)}, & t < 0 \\ \sqrt{b/(at)}, & t > 0 \end{cases}. \quad (7.21)$$

Then Eq. (7.19) reads

$$\nabla^2\varphi - \xi^{-2}\varphi = -b^{-1}\delta(\mathbf{r} - \mathbf{r}'). \quad (7.22)$$

To solve it, recall that $\nabla^2(1/r) = -4\pi\delta(\mathbf{r})$ so that

$$\varphi(\mathbf{R}) = \frac{\exp(-R/\xi)}{4\pi b R} \quad (7.23)$$

where $\mathbf{R} = \mathbf{r} - \mathbf{r}'$.

Note that the correlation length diverges at the critical temperature $t = 1$ and that the critical exponent $\nu = 1/2$ defined by Eq. (7.9). Thus at $t = 1$, the response function is characterized by R^{-1} power law decay so that the critical exponent η defined by Eq. (7.10) is $\nu = 0$.

Equation (7.23) is the response function obtained within a mean-field theory. Other exactly solvable models predict a more generalized functional form of the response function where the power-law decay is characterized by an exponent different from 1.

Two-point correlation function

The two-point correlation function (in some contexts referred to as the spin-spin correlation function) measures the persistence of spatial order and is defined by

$$\begin{aligned} g(\mathbf{r}, \mathbf{r}') &= \langle [m(\mathbf{r}) - \langle m(\mathbf{r}) \rangle][m(\mathbf{r}') - \langle m(\mathbf{r}') \rangle] \rangle \\ &= \langle m(\mathbf{r})m(\mathbf{r}') \rangle - \langle m(\mathbf{r}) \rangle \langle m(\mathbf{r}') \rangle. \end{aligned} \quad (7.24)$$

We will show later that

$$g(\mathbf{r}, \mathbf{r}') = g(R = |\mathbf{r} - \mathbf{r}'|) = k_B T \underbrace{\frac{\exp(-R/\xi)}{4\pi b R}}_{\varphi(\mathbf{r}, \mathbf{r}')}. \quad (7.25)$$

This is the equilibrium variant of the *fluctuation-dissipation theorem* (to be discussed later) stating that the correlation function and thus the fluctuations are connected to the response function and thus to the susceptibility.

7.3 Thermodynamic inequalities

The critical exponents are not independent of each other; there exist four relations connecting them. These relations read

$$\alpha + 2\beta + \gamma = 2 \quad (\text{Rushbrooke}), \quad (7.26)$$

$$\beta(\delta - 1) = \gamma \quad (\text{Widom}), \quad (7.27)$$

$$\alpha = 2 - \nu d \quad (\text{Josephson}), \text{ and} \quad (7.28)$$

$$\gamma = \nu(2 - \eta) \quad (\text{Fisher}). \quad (7.29)$$

All of these relations are obeyed in all known models of phase transitions as well as by experiments; the first two are referred to as the *scaling relations* and the third and the fourth are the *hyperscaling relations*. "Hyperscaling" refers to the fact that they involve exponents describing spatial correlations, and thus cannot be derived without considering the volume of the domains or some related quantity.

Rushbrooke inequality

This inequality is usually derived for a magnetic system. Consider the difference between the specific heat at constant field c_H and the specific heat at constant magnetization c_M defined by

$$c_H = \frac{T}{m} \left(\frac{\partial S}{\partial T} \right)_H \quad \text{and} \quad c_M = \frac{T}{m} \left(\frac{\partial S}{\partial T} \right)_M, \quad (7.30)$$

respectively. Their difference is obtained from the derivative

$$\left(\frac{\partial S}{\partial T} \right)_H = \left(\frac{\partial S}{\partial T} \right)_M + \left(\frac{\partial S}{\partial M} \right)_T \left(\frac{\partial M}{\partial T} \right)_H. \quad (7.31)$$

For a magnetic system where $dF = -SdT + \mu_0 V H dM$, the first Maxwell relation reads

$$\left(\frac{\partial S}{\partial M}\right)_T = -\mu_0 V \left(\frac{\partial H}{\partial T}\right)_M \quad (7.32)$$

so that

$$c_H - c_M = -\frac{T}{m} \mu_0 V \left(\frac{\partial H}{\partial T}\right)_M \left(\frac{\partial M}{\partial T}\right)_H. \quad (7.33)$$

Now recall that

$$dM = \left(\frac{\partial M}{\partial T}\right)_H dT + \left(\frac{\partial M}{\partial H}\right)_T dH \quad (7.34)$$

which means that

$$\left(\frac{\partial H}{\partial T}\right)_M = -\frac{(\partial M/\partial T)_H}{(\partial M/\partial H)_T}. \quad (7.35)$$

Thus

$$c_H - c_M = \frac{T}{m} \mu_0 V \left[\left(\frac{\partial M}{\partial T}\right)_H\right]^2 \left[\left(\frac{\partial M}{\partial H}\right)_H\right]^{-1}. \quad (7.36)$$

Given that $c_M > 0$, this means that

$$c_H > \frac{T}{m} \mu_0 V \left[\left(\frac{\partial M}{\partial T}\right)_H\right]^2 \left[\left(\frac{\partial M}{\partial H}\right)_H\right]^{-1}. \quad (7.37)$$

Recall that $c_H \propto |t|^{-\alpha}$, $(\partial M/\partial T)_H \propto |t|^{\beta-1}$ [because $m(t) \propto |t|^\beta$], and $(\partial M/\partial H)_H \propto |t|^{-\gamma}$ so that Eq. (7.37) reads

$$D_1 |t|^{-\alpha} > D_2 |t|^{2(\beta-1)} |t|^\gamma, \quad (7.38)$$

where D_1 and D_2 are positive constants. In other words,

$$|t|^{2-\alpha-2\beta-\gamma} > D_2/D_1. \quad (7.39)$$

Since $|t|$ can be arbitrarily small, $2 - \alpha - 2\beta - \gamma$ must be positive:

$$\alpha + 2\beta + \gamma \geq 2. \quad (7.40)$$

In fact, this inequality is actually an equality, but the proof is more involved. — The other equalities can be obtained using similar arguments.

Dimensional analysis

The scaling relations can also be substantiated by dimensional arguments and the *scaling hypothesis* stating that close to the phase transition, any characteristic length L scales with the correlation length $\xi \propto |t|^{-\nu}$ [3]. Note that this hypothesis is an assumption!

Close to the critical point, the order parameter fluctuates considerably, and is correlated across distances given by the correlation length. The magnitude of the order parameter can be estimated from the magnitude of its fluctuations, in turn related to the two-point correlation function. The correlation function scales as

$$[g] \sim L^{2-d-\eta} \quad (7.41)$$

and the magnitude of order parameter itself can be estimated by the magnitude of fluctuations of the order parameter. Thus $[m] \sim \sqrt{[g]}$

$$[m] \sim L^{(2-d-\eta)/2} \sim t^{-\nu(2-d-\eta)/2}, \quad (7.42)$$

where we have inserted $\xi \propto |t|^{-\nu}$. The exponents on the left- and the right-hand side must be the same which means that

$$\beta = -\nu \frac{2-d-\eta}{2}. \quad (7.43)$$

Next we consider the susceptibility defined by

$$\chi = \int d\mathbf{r} \varphi(\mathbf{r}) \quad (7.44)$$

(so that we can distinguish between the local response function and the response of a finite volume of matter). Then

$$[\chi] \sim L^d L^{2-d-\eta} = L^{2-\eta} \quad (7.45)$$

and like above we find

$$\gamma = \nu(2-\eta), \quad (7.46)$$

which is the Fisher equality. Similar arguments can be employed to obtain other relationships wherefrom one can then extract the equalities by addition and subtraction.

7.4 Ginzburg criterion

The problem of the mean-field theories is that they do not include fluctuations. One must then ask how important are fluctuations compared to the average value of the order parameter — should they be small the theory is valid but otherwise it is not. At the critical temperature, the two-point correlation function becomes a power law because the correlation length diverges. This means that at the phase transition, the system is characterized by long-range order even if the microscopic interactions are short-range, and thus neglecting fluctuations is a serious drawback to the point that the very premises of the mean-field theories such as the Landau theory are questionable.

Let us now examine the fluctuations in more detail. When discussing the grand canonical ensemble, we showed that the ratio of density fluctuations and density itself

$$\frac{\langle(\Delta\rho)^2\rangle}{\langle\rho\rangle^2} = \frac{k_B T}{V} \chi_T. \quad (7.47)$$

[Eq. (2.92) rewritten using ρ for particle density.] This expression is a special case of a more general statement connecting fluctuations and susceptibility known as the fluctuation-dissipation theorem. Consider an ensemble where a microscopic variable m (here denoted by the same symbol as the order parameter on purpose; still not that the microscopic variable is an extensive quantity so that mh has the dimension of energy) coupled to an external field h so that the energy reads

$$E = \varepsilon - mh, \quad (7.48)$$

where ε contains all terms except for the coupling with external field. Then $\langle m \rangle$ is given by

$$\langle m \rangle = \frac{\int m \exp(-\beta(\varepsilon - mh)) d\omega}{\int \exp(-\beta(\varepsilon - mh)) d\omega} = \frac{\partial \ln Q_N(h, T)}{\partial(\beta h)}. \quad (7.49)$$

In turn,

$$\begin{aligned} \frac{\partial^2 \ln Q_N(h, T)}{\partial(\beta h)^2} &= \frac{\partial}{\partial(\beta h)} \left[\frac{1}{Q_N} \frac{\partial Q_N(h, T)}{\partial(\beta h)} \right] \\ &= \frac{1}{Q_N} \frac{\partial^2 Q_N(h, T)}{\partial(\beta h)^2} - \left[\frac{1}{Q_N} \frac{\partial Q_N(h, T)}{\partial(\beta h)} \right]^2 \\ &= \langle m^2 \rangle - \langle m \rangle^2. \end{aligned} \quad (7.50)$$

At the same time, the left-hand side of this result is also equal to

$$\frac{\partial \langle m \rangle}{\partial(\beta h)} = \beta^{-1} \frac{\partial \langle m \rangle}{\partial h} = \beta^{-1} \chi, \quad (7.51)$$

where χ is the susceptibility now defined as the derivative of the extensive generalized coordinate with respect to the external force. By combining Eqs. (7.50) and (7.51) we find that

$$\chi = \frac{\langle m^2 \rangle - \langle m \rangle^2}{k_B T}. \quad (7.52)$$

Should the mean-field theory be valid, $\langle (\Delta m)^2 \rangle$ must be much smaller than $\langle m \rangle^2$ or

$$k_B T \chi \ll \langle m \rangle^2. \quad (7.53)$$

This is known as the Ginzburg criterion.

We now show how this criterion depends on the dimensionality of the system. For a domain of volume ξ^d , the Ginzburg criterion reads

$$k_B T_c A \xi^d |t|^{-\gamma} \ll (B \xi^d |t|^\beta)^2, \quad (7.54)$$

where we took into account that χ scales with volume and is also proportional to $|t|^{-\gamma}$; also proportional to volume is the extensive variable m . A and B are positive constants and on the left-hand side we inserted T_c for T . The correlation length $\xi = a|t|^{-\nu}$ where a is a constant and so

$$|t|^{\nu d - \gamma - 2\beta} \ll \frac{B^2 a^d}{A k_B T_c} = D, \quad (7.55)$$

where $D > 0$. Using the Rushbrooke equality $\alpha + 2\beta + \gamma = 2$ this can be cast as

$$|t|^{\nu d - 2 + \alpha} \ll D. \quad (7.56)$$

In the mean-field theory, $\nu = 1/2$ and $\alpha = 0$ so that

$$|t|^{(d-4)/2} \ll D. \quad (7.57)$$

Should this condition be fulfilled for arbitrarily small t and thus the power must be positive, which is true for $d > 4$. In turn, for $d > 4$ the fluctuations are small so that the mean-field theory is valid. At the same time, the above condition was derived for a scalar order parameter $n = 1$. Given that fluctuations are expected to decrease with increasing number of components of the order parameter n , mean-field theory should be valid in $d > 4$ in all

models irrespective of n . Thus $d = 4$ is referred to as the *upper critical dimension*.

Equation (7.56) can also be interpreted in another way. At the phase transition, we may expect that the magnitude of fluctuations of the order parameter is comparable to the order parameter itself. This also means that the left-hand side of Eq. (7.56) is of order 1 which implies that the power must be 0. This gives the Josephson hyperscaling equality

$$\nu d - 2 + \alpha = 0, \quad (7.58)$$

which is indeed obeyed in $d < 4$ as shown in Table 7.1, except, of course, by the mean-field theory.

7.5 Exactly solvable models

Here we briefly review the main features of the Tonks gas and the 1D and the 2D Ising models, showing that order is indeed controlled by the dimensionality of the system. Neither of the two 1D models has a phase transition but the 2D Ising model does have an ordered phase at low enough temperatures.

Tonks gas

This is a 1D model of a fluid. Various nearest-neighbors pair potentials can be explored, the simplest one being the hard-sphere repulsion. Consider N spheres of diameter σ arranged on a ring of perimeter L ; periodic boundary conditions are assumed. Particle 1 is located at $x_1 = 0$ and the positions of all other particles are restricted by their respective neighbors such that

$$x_{j-1} + \sigma < x_j < L - (N - j + 1)\sigma. \quad (7.59)$$

The configuration integral reads

$$Z_N(L) = \frac{L}{N} \int_{\sigma}^{L-(N-1)\sigma} dx_2 \int_{x_2+\sigma}^{L-(N-2)\sigma} dx_3 \dots \int_{x_{N-1}+\sigma}^{L-\sigma} dx_N. \quad (7.60)$$

Here the prefactor L arises from integrating x_1 over the perimeter, and $1/N$ accounts for indistinguishability of particles in that any particle may be chosen as the first one. The integral over x_{N-1} is $L - 2\sigma - x_{N-1}$, the integral over x_{N-2} gives $(L - 2\sigma)(L - 3\sigma - x_{N-2}) + (1/2)[(L - 2\sigma)^2 - (x_{N-2} + \sigma)^2]$ etc. and the final result is

$$Z_N(L) = \frac{L(L - N\sigma)^{N-1}}{N!}. \quad (7.61)$$

If N is large, the numerator in Z_N can be written as $(L - N\sigma)^N$ and then the free energy reads

$$F(N, L, T) = -Nk_B T \ln \left(\frac{L - N\sigma}{N\Lambda} \right) - Nk_B T. \quad (7.62)$$

The pressure

$$p = - \left(\frac{\partial F}{\partial L} \right)_{T, N} = \frac{Nk_B T}{L - N\sigma} \quad (7.63)$$

corresponds to the ideal gas confined to a volume $L - N\sigma$ representing the total volume less the intrinsic volume of the spheres. This isotherm is strictly monotonic and so there is no phase transition in this model.

1D Ising model

Here too one considers spins on a ring. The Hamiltonian including the spin-spin exchange interaction and the coupling to external magnetic field B reads

$$H_N\{\sigma_j\} = -J \sum_{\text{n.n.}} \sigma_i \sigma_j - \mu B \sum_{i=1}^N \sigma_i. \quad (7.64)$$

Here $\sigma_i = \pm 1$ and μ is the Bohr magneton. The partition function can be evaluated using the transfer matrix method. The result is

$$Q_N(B, T) = \left\{ \exp(\beta J) \cosh(\beta \mu B) + [\exp(-2\beta J) + \exp(2\beta J) \sinh^2(\beta \mu B)]^{1/2} \right\}^N \quad (7.65)$$

and the free energy is

$$F(B, T) = -NJ - Nk_B T \ln \left(\cosh(\beta \mu B) + [\exp(-4\beta J) + \sinh^2(\beta \mu B)]^{1/2} \right). \quad (7.66)$$

The magnetization reads

$$\overline{M}(B, T) = \frac{N\mu \sinh(\beta \mu B)}{[\exp(-4\beta J) + \sinh^2(\beta \mu B)]^{1/2}}, \quad (7.67)$$

which shows that in zero field, $\overline{M} = 0$ so that there is no spontaneous magnetization at any finite temperature, and thus no phase transition.

2D Ising model

The 2D Ising model can too be solved using the transfer matrix method, which was first applied to this problem by Onsager in 1944. The solution is extremely complicated and we only list a few key results. The main qualitative conclusion is that in 2D, the Ising model does have a phase transition. The zero-field partition function for the square lattice is given by

$$\frac{1}{N} \ln Q(T) = \ln \left(2^{1/2} \cosh(2\beta J) \right) + \frac{1}{\pi} \int_0^{\pi/2} d\phi \ln \left(1 + \sqrt{1 - \kappa^2 \sin^2 \phi} \right), \quad (7.68)$$

where

$$\kappa = \frac{2 \sinh(2\beta J)}{\cosh^2(2\beta J)}. \quad (7.69)$$

Note that the second term in Eq. (7.68) has a singularity at $\kappa = 1$; this is where the phase transition takes place. Specifically, the phase transition is located at

$$(\beta J)_c = \frac{1}{2} \ln \left(\sqrt{2} + 1 \right) \approx 0.4407 \quad (7.70)$$

or

$$T_c \approx 2.269J/k_B. \quad (7.71)$$

An important result obtained by Onsager is the critical behavior of the specific heat

$$\begin{aligned} \frac{C(T)}{Nk_B} &\approx \frac{8}{\pi} [(\beta J)_c]^2 \left\{ -\ln \left| 1 - \frac{T}{T_c} \right| + \left[\ln \left(\frac{\sqrt{2}}{(\beta J)_c} \right) - 1 - \frac{\pi}{4} \right] \right\} \\ &\approx -0.4945 \ln \left| 1 - \frac{T}{T_c} \right|, \end{aligned} \quad (7.72)$$

which diverges logarithmically so that $\alpha = \alpha' = 0$. This exact result is qualitatively different from the predictions of the Bethe approximation¹ and the mean-field Bragg-Williams approximation, which both suggest that the specific heat has a jump at the transition (Fig. 7.2). Note that the critical temperatures predicted by these two approximations are different from the correct one; specifically, the Bethe approximation gives $T_c = 2.885J/k_B$ and the Bragg-Williams approximation gives $T_c = 4J/k_B$.

Finally, the 2D Ising model can also be studied by considering domains

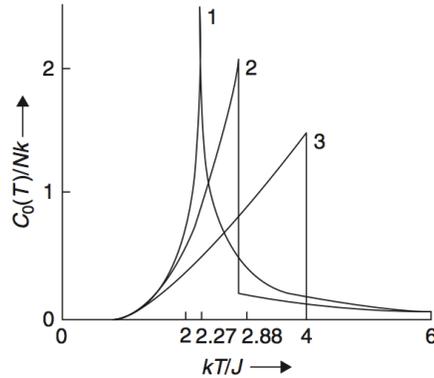


Figure 7.2: Specific heat of the 2D Ising model obtained using the Onsager solution (1) and the Bethe and the Bragg-Williams approximations (2 and 3, respectively) [1].

pointing in the opposite direction as the ground state (Fig. 7.3). The energy of a domain is proportional to the length of the domain wall, and its entropy depends on the number of arrangements of a domain wall of a given length. In the simplest approximation we assume that from a given site, the domain wall may be continued in 3 different directions. Thus we find that the entropy is roughly given by $S(L) \approx k_B \ln 3^L$ where L is the length of the domain wall. Here many details of the possible domain configurations are neglected, especially the fact that the domain wall must be a closed loop. — Since the energy of the wall is $2JL$, the free energy of a domain

$$\Delta F_{2D}(L) = (2J - k_B T \ln 3) L, \quad (7.73)$$

which suggests that the critical temperature is

$$T_c = \frac{2}{\ln 3} J/k_B = 1.820 J/k_B. \quad (7.74)$$

This value is too low; we have overestimated the entropy. Obviously this result can be refined.

A similar argument shows that in the 1D Ising model there is no phase transition, The energy of the domain wall is $2 \times 2J$ irrespective of domain size, and the entropy can be estimated by noting that in a ring of N spins,

¹Here the interaction of a given spin with its nearest neighbors is taken into account exactly whereas the interaction of its neighbors with their own neighbors is approximated in a mean-field fashion.

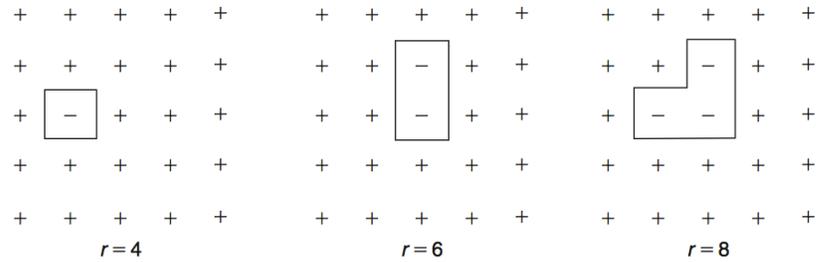


Figure 7.3: Domains in the 2D Ising models: The smallest domain containing only 1 unlike spin can only be constructed in 1 way, whereas the phase space of domains containing many unlike spins is much larger. Shown here are domains of length 4, 6, and 8; the latter also exists in the elongated configuration [1].

the first domain wall can be placed at $N - 1$ positions whereas the second one can be placed at $N - 2$ positions (because one position is already occupied). Thus one concludes that there are $(N - 1)(N - 2)/2$ arrangements of domain walls with a given energy, the additional factor of $1/2$ accounting for the double counting of domains due to indistinguishability of the two walls. This number is of the order of $\sim N^2$ so that the free energy of a domain in 1D is

$$\Delta F_{1D} = 4J - 2k_B T \ln N. \quad (7.75)$$

In thermodynamic limit, ΔF_{1D} is negative at any finite T .

7.6 Kosterlitz-Thouless transition

The phase transition in planar systems with continuous symmetry (rather than a discrete symmetry like in the Ising model $m \rightarrow -m$) is special because in this case the low-temperature phase lacks long-range order. This kind of transition is known as the Kosterlitz-Thouless transition, and is claimed to be seen in some planar magnets, films of liquid helium, thin superconducting films, liquid-crystal films, etc. [4]. Essential to the understanding of this transition is to see how and why the spatial dimension of 2 separates the non-ordering, transition-less 1D systems from 3D systems with critical behavior.

To this end, we consider a 2D XY model with the Hamiltonian

$$H = -J \sum_{i,j \text{ neighbors}} (S_{ix}S_{jx} + S_{iy}S_{jy}) = -JS^2 \sum_{i,j \text{ neighbors}} \cos(\phi_i - \phi_j). \quad (7.76)$$

Here spins are classical 2D vectors of magnitude S lying on a d -dimensional hypercubic lattice, and ϕ_i is their azimuthal angle. This Hamiltonian has a continuous rotational symmetry because it is left unchanged by the transformation $\phi_i \rightarrow \phi_i + \phi_0$ with an arbitrary ϕ_0 . The ground state of this system is fully aligned with a constant $\phi_i = \phi$ for all i . Now assume that at a low enough temperature, the orientations of the nearest neighbors are not too different so that $\phi_i - \phi_j \ll 2\pi$ for i, j nearest neighbors. Then the cosine may be expanded so that the Hamiltonian reads

$$\begin{aligned} H &= -\frac{qNJS^2}{2} + \frac{JS^2}{2} \sum_{i,j \text{ neighbors}} (\phi_i - \phi_j)^2 \\ &= E_0 + \frac{JS^2}{4} \sum_{\mathbf{r}, \mathbf{a}} [\phi(\mathbf{r} + \mathbf{a}) - \phi(\mathbf{r})]^2. \end{aligned} \quad (7.77)$$

Here N is the number of spins, q is the coordination number, \mathbf{r} is a lattice vector, and \mathbf{a} is a vector pointing to nearest neighbors of \mathbf{r} . Notice that in this form, each pair is counted twice, hence an additional factor of $1/2$. The last version of the Hamiltonian can be cast in a continuum form by replacing the differences $\phi(\mathbf{r} + \mathbf{a}) - \phi(\mathbf{r})$ by differentials $a\nabla\phi(\mathbf{r})$ (here a is the lattice spacing) and the sum on \mathbf{r} by the integral $\int d^d\mathbf{r}/a^d$:

$$H = E_0 + \frac{JS^2}{2a^{d-2}} \int d^d\mathbf{r} [\nabla\phi(\mathbf{r})]^2. \quad (7.78)$$

Finally, we dispose of the restriction that ϕ must lie in the range $[0, 2\pi]$ and treat it as an unbound variable. This is not expected to have a dramatic effect in the ground state where ϕ varies very little.

Spin-spin correlation function

We now calculate the spin-spin correlation function defined by

$$g(r) = \langle \exp(i[\phi(\mathbf{r}) - \phi(0)]) \rangle. \quad (7.79)$$

To this end, we expand the angle in Fourier series

$$\phi(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \phi_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r}) \quad (7.80)$$

so that

$$\phi(\mathbf{r}) - \phi(0) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} \phi_{\mathbf{k}} [\exp(i\mathbf{k} \cdot \mathbf{r}) - 1]. \quad (7.81)$$

The ensemble average is done by first writing the Hamiltonian in terms of the Fourier components:

$$H = E_0 + \frac{JS^2\Omega}{2a^{d-2}N} \sum_{\mathbf{k}} k^2 |\phi_{\mathbf{k}}|^2. \quad (7.82)$$

Here Ω is the volume of the system. Clearly $\Omega = Na^d$ so that

$$H = E_0 + \frac{JS^2a^2}{2} \sum_{\mathbf{k}} k^2 |\phi_{\mathbf{k}}|^2. \quad (7.83)$$

In evaluating the average in Eq. (7.79), we may neglect the constant term in the energy. Temporarily we introduce a shorthand $K = JS^2a^2$:

$$g(r) = \frac{\int \prod_{\mathbf{k}} d\phi_{\mathbf{k}} \exp(-(\beta K/2) \sum_{\mathbf{k}} \{k^2 |\phi_{\mathbf{k}}|^2 + i\phi_{\mathbf{k}} [\exp(i\mathbf{k} \cdot \mathbf{r}) - 1]\})}{\int \prod_{\mathbf{k}} d\phi_{\mathbf{k}} \exp(-(\beta K/2) \sum_{\mathbf{k}} k^2 |\phi_{\mathbf{k}}|^2)}. \quad (7.84)$$

The Fourier components of ϕ are complex, $\phi_{\mathbf{k}} = \alpha_{\mathbf{k}} + i\gamma_{\mathbf{k}}$. Since ϕ is a real quantity $\phi_{\mathbf{k}} = \phi_{-\mathbf{k}}$ which means that $\alpha_{\mathbf{k}}$ is even and $\gamma_{\mathbf{k}}$ is odd. Each term in the argument of the exponential in the numerator then reads

$$\begin{aligned} & -\frac{\beta K k^2}{2} (\alpha_{\mathbf{k}}^2 + \gamma_{\mathbf{k}}^2) + i\alpha_{\mathbf{k}} [\cos(\mathbf{k} \cdot \mathbf{r}) - 1] - \underbrace{\alpha_{\mathbf{k}} \sin(\mathbf{k} \cdot \mathbf{r})}_{\text{odd}} \\ & - \underbrace{\gamma_{\mathbf{k}} [\cos(\mathbf{k} \cdot \mathbf{r}) - 1]}_{\text{odd}} - i\gamma_{\mathbf{k}} \sin(\mathbf{k} \cdot \mathbf{r}). \end{aligned} \quad (7.85)$$

The two odd terms will vanish on summation over \mathbf{k} and we can drop them at this stage already. Now we combine the square and linear terms in $\alpha_{\mathbf{k}}$ and $\gamma_{\mathbf{k}}$, respectively, and complete the squares. After this, Eq. (7.85) reads

$$\begin{aligned} & -\frac{\beta K k^2}{2} \left(\left\{ \alpha_{\mathbf{k}} - \frac{i [\cos(\mathbf{k} \cdot \mathbf{r}) - 1]}{\beta K k^2} \right\}^2 + \left\{ \gamma_{\mathbf{k}} + \frac{i \sin(\mathbf{k} \cdot \mathbf{r})}{\beta K k^2} \right\}^2 \right) \\ & -\frac{1}{2\beta K k^2} \{ [\cos(\mathbf{k} \cdot \mathbf{r}) - 1]^2 + \sin(\mathbf{k} \cdot \mathbf{r})^2 \} \\ = & -\frac{\beta K k^2}{2} \left(\left\{ \alpha_{\mathbf{k}} - \frac{i [\cos(\mathbf{k} \cdot \mathbf{r}) - 1]}{\beta K k^2} \right\}^2 + \left\{ \gamma_{\mathbf{k}} + \frac{i \sin(\mathbf{k} \cdot \mathbf{r})}{\beta K k^2} \right\}^2 \right) \\ & -\frac{1 - \cos(\mathbf{k} \cdot \mathbf{r})}{\beta K k^2}. \end{aligned} \quad (7.86)$$

Note that $\alpha_{\mathbf{k}}$ and $\gamma_{\mathbf{k}}$ are only contained in the first term but not in the second term $[1 - \cos(\mathbf{k} \cdot \mathbf{r})]/(\beta K k^2)$. At each \mathbf{k} , the first term thus results in

two shifted Gaussians which exactly cancel with the respective non-shifted Gaussians in the denominator. (The shift is immaterial since the integrals go from $-\infty$ to $+\infty$.) The result is thus

$$g(r) = \exp \left(-\frac{k_B T}{JS^2 a^2} \sum_{\mathbf{k}} \frac{1 - \cos(\mathbf{k} \cdot \mathbf{r})}{k^2} \right). \quad (7.87)$$

Now the sum on \mathbf{k} is written as an integral

$$\sum_{\mathbf{k}} = \frac{a^d}{(2\pi)^d} \int d^d \mathbf{k}, \quad (7.88)$$

which gives

$$g(r) = \exp \left(-\frac{k_B T a^{d-2}}{(2\pi)^d JS^2} \int d^d \mathbf{k} \frac{1 - \cos(\mathbf{k} \cdot \mathbf{r})}{k^2} \right). \quad (7.89)$$

In 2D, the remaining integral can be done in polar coordinates. Recall that

$$\int_0^{2\pi} d\theta \cos(kr \cos \theta) = 2\pi J_0(kr), \quad (7.90)$$

where J_0 is the zero-order Bessel function. Thus

$$g(r) = \exp \left(-\frac{k_B T}{2\pi JS^2} \int_0^{\pi/a} dk \frac{1 - J_0(kr)}{k} \right). \quad (7.91)$$

Here π/a is the lattice cutoff.

In the remaining integral, the presence of the Bessel function removes the singularity at the lower boundary because $J_0(0) = 1$. For large k , the Bessel function is much smaller than 1 so that we can neglect it to find that the result is essentially

$$\frac{k_B T}{2\pi JS^2} \int_0^{\pi r/a} dx \frac{1 - J_0(x)}{x} \approx \frac{k_B T}{2\pi JS^2} \ln \left(\frac{\pi r}{a} \right). \quad (7.92)$$

In passing, we have introduced a new variable $x = kr$. Thus the correlation function is a power law at all temperatures:

$$g(r) = \exp \left(-\frac{k_B T}{2\pi JS^2} \ln \left(\frac{\pi r}{a} \right) \right) = \left(\frac{\pi r}{a} \right)^{-\eta(T)}, \quad (7.93)$$

where

$$\eta(T) = \frac{k_B T}{2\pi JS^2} \quad (7.94)$$

is the temperature-dependent exponent.

This result is truly remarkable in that it predicts an algebraic decay of order at all temperatures, which means that there is no long-range order in this system at any temperature. Note that this is the consequence of 2 spatial dimensions: In 3D, Eq. (7.89) includes $d^3\mathbf{k} \sim k^2 dk$ so that the k^2 in the denominator cancels with the k^2 from the phase space. As a result, the integral is a constant rather than a logarithm and thus $g(r) = \text{const.}$ for large r as expected in an ordered system.

Physically, order is destroyed by long-wavelength spin fluctuations related to the continuous rotational symmetry of the system which implies that there must exist elementary excitations of Goldstone type. The frequency of these excitations decreases to 0 as wavelength approaches infinity, which means that their amplitude is large at large wavelengths. As a result, the long-range orientational order is destroyed.

Yet another aspect of the power-law decay of correlations is the following. Recall that the Landau theory predicts an exponential decay of the correlation function and a power law only at the critical temperature; this would mean that the 2D Heisenberg model is critical at all temperatures. This does not seem plausible, and Kosterlitz and Thouless that there must exist another type of excitations which destroy the low-temperature phase and lead to a disordered high-temperature phase.

Vortex unbinding

Kosterlitz and Thouless proposed that these excitations be vortices which appear in pairs at low temperatures but unbind at high temperatures. Now we calculate the energy of a single isolated vortex described by the orientation of the spin at $\phi(r, \theta)$. Assume that $\phi(r, \theta) = n\theta$ where $n \in \mathbb{Z}$ is the strength of vortex such that

$$\oint \nabla\phi \cdot d\mathbf{l} = 2\pi n, \quad (7.95)$$

where $\nabla\phi = (n/r)\mathbf{e}_\theta$. The vortex energy is then

$$E_{\text{vortex}} = \frac{JS^2}{2} \int d^2\mathbf{r} [\nabla\phi(\mathbf{r})]^2 = \pi JS^2 n^2 \int_a^L \frac{dr}{r} = \pi JS^2 n^2 \ln \frac{L}{a}. \quad (7.96)$$

Here L is the size of the system. The entropy of the vortex is associated with its position. In a 2D lattice model, the phase space is $(L/a)^2$ so that $S = k_B \ln(L/a)^2$ and then the free energy of the vortex is

$$\Delta F = (\pi JS^2 n^2 - 2k_B T) \ln \frac{L}{a}. \quad (7.97)$$

This result shows that for $T < \pi JS^2 n^2 / (2k_B)$, isolated vortices are not stable. The critical temperature depends on vortex strength n and is lowest for $n = \pm 1$ which allows us to define the Kosterlitz-Thouless temperature

$$T_{KT} = \frac{\pi JS^2}{2k_B}, \quad (7.98)$$

above which the system contains unbound vortices.

Below T_{KT} , vortices appear in pairs. This is most easily seen by calculating the energy of a single vortex pair. To this end, recall that the equilibrium spin configuration is determined by the minimum of Eq. (7.78) which in 2D reads

$$H = \frac{JS^2}{2} \int d^2\mathbf{r} [\nabla\phi(\mathbf{r})]^2. \quad (7.99)$$

In equilibrium, $\delta E = 0$ which implies that

$$\nabla^2\phi(\mathbf{r}) = 0. \quad (7.100)$$

This Laplace equation must be solved together with boundary conditions requiring that the circulation around each of the vortices is nonzero as specified by the strength of the vortices n_1 and n_2 . This problem is completely analogous to 2D magnetism, the spins playing the role of magnetic field around two straight wires perpendicular to the plane. By borrowing the result from electromagnetism, we find that the energy of a vortex pair is

$$E_{\text{pair}}(\mathbf{r}_1, \mathbf{r}_2) = -2\pi JS^2 n_1 n_2 \ln \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{a}. \quad (7.101)$$

(There also exists an analogy with 2D electrostatics.) This result shows that two oppositely charged vortices occur in a tightly bound pair: The prefactor is positive as $n_1 n_2 < 0$ but the logarithm $\rightarrow \infty$ at $|\mathbf{r}_1 - \mathbf{r}_2| \rightarrow 0$.

Given that the energy of a pair does not depend on the size of the system while its entropy increases with the size of the system, we expect that the low-temperature state will contain a finite density of bound vortex pairs, the equilibrium density being determined by the pair-pair interaction (not considered here). In the high-temperature state, the vortex pairs unbind (Fig. 7.4). It can be shown that vortex unbinding leads to an exponential rather than power-law correlation function.

Other 2D systems with continuous symmetry

The above discussion suggests that there exists no long-range orientational order in 2D spin system. A similar argument applies to the positional order

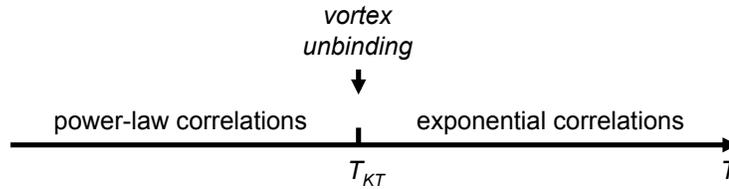


Figure 7.4: Kosterlitz-Thouless transition: In the low-temperature phase, the spin-spin correlation function is a power law whereas in the high-temperature phase it is exponential.

of, e.g., particles in a floating monolayer adhering to the substrate but not strongly perturbed by it so that the structure of the monolayer is determined by the interparticle interactions rather than by the interactions with the substrate. Here the continuous symmetry refers to in-plane translation, and absence of long-range translational order is seen in the diffraction pattern where the profile of the peaks is a power-law rather than a δ -function like in a true crystal.

Despite the lack of the long-range positional order, 2D crystals may have long-range orientational order of bonds between nearest neighbors. The combination of quasi-long-range positional order and long-range orientational order is characteristic of the 2D hexagonal solid phase. On heating, this phase first undergoes the dislocation unbinding transition where pairs of dislocations dissociate (Fig. 7.5c-e). This process transforms the hexagonal phase into the hexatic phase of short-range positional order and quasi-long-range orientational order. At a higher temperature still, disclinations unbind too. Thus isolated dislocations are formed and the hexatic phase melts into the isotropic phase. This behavior is seen in a system as simple as hard disks. Figure 7.5 shows a minus and a plus disclination, a dislocation and the process of pair formation and dissociation, and Fig. 7.6 shows the hexagonal, hexatic, and isotropic phase obtained in simulations of hard spheres forming a monolayer.

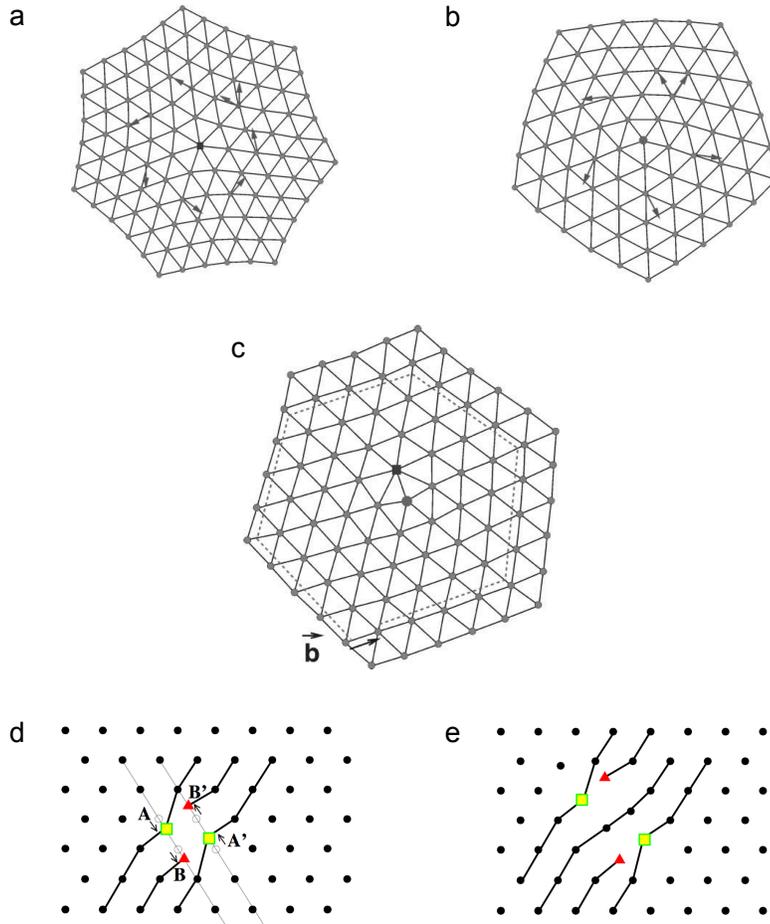


Figure 7.5: Disclinations and dislocations on a hexagonal lattice. A minus (a) and a plus disclination (b) break the bond orientational order of the lattice. A pair of a plus and a minus disclinations constitute a dislocation (c) which breaks the translational order [5]. Formation (d) and dissociation (e) of dislocation pairs [6].

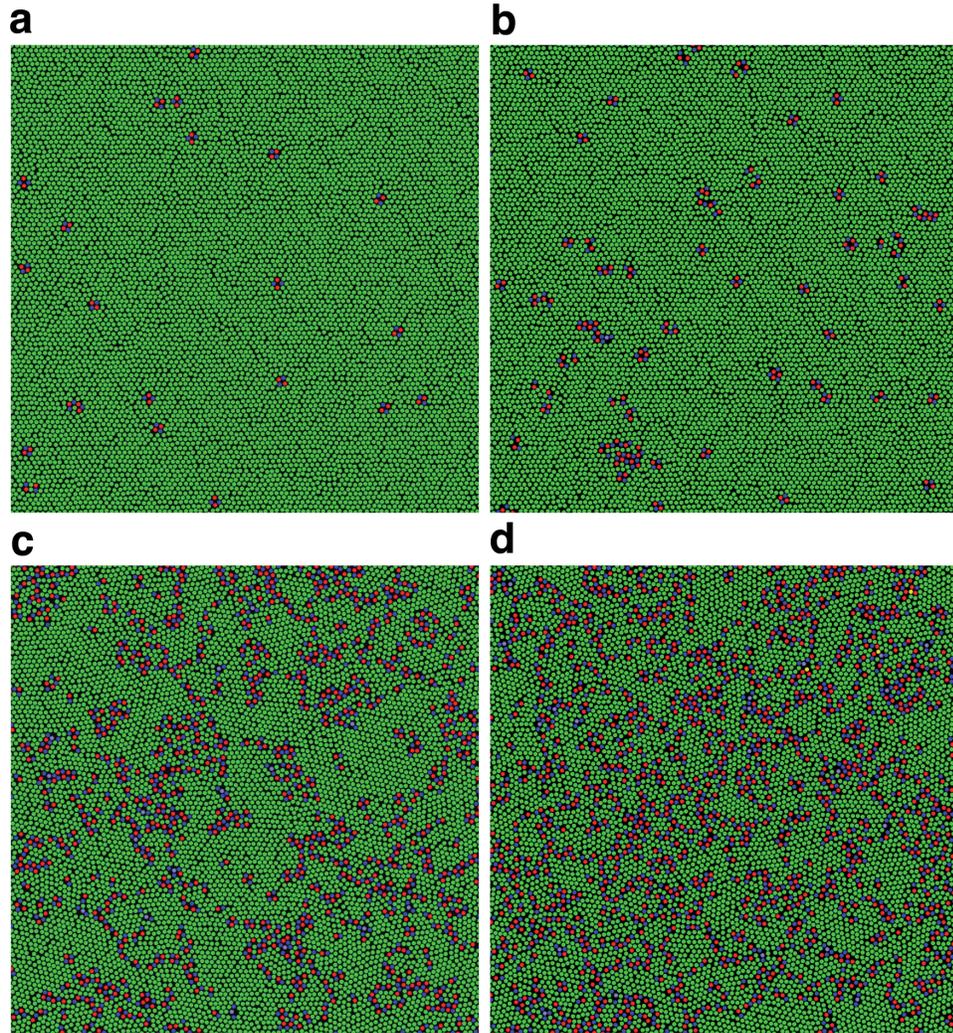


Figure 7.6: Hexagonal (a), hexatic (b and c, the latter at the coexistence packing fraction), and isotropic phase (d; at the coexistence packing fraction) of a monolayer of hard spheres [7]. Blue, green, and red particles are five-, six-, and seven-coordinated, respectively.

7.7 Renormalization group

By now, we have realized that virtually any analysis of the phase transition is rather complicated. There exist only a few exactly solvable models and approximations such as the mean-field approximation usually miss the main features even at a qualitative level. The most convincing theory of the phase transitions proposed by now relies on the behavior of the system that plagues all approximate theories: The absence of a characteristic lengthscale at the critical point. Kadanoff employed this property to view the critical point as a state that is insensitive to a change of the spatial scale.

The change of scale can be illustrated by considering Ising spins (Fig. 7.7). During the change of scale, a l^2 group of spins (if $l = 2$ in $d = 2$ dimensions, a 2×2 block of 4 spins) is replaced by a single spin. The transformed lattice with a larger lattice constant

$$a' = la \quad (7.102)$$

contains

$$N' = l^{-d}N \quad (7.103)$$

spins. In order to keep the density of the degrees of freedom (whose num-

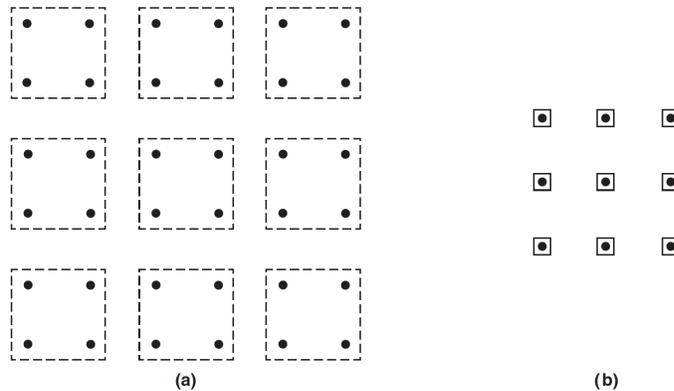


Figure 7.7: A square lattice of Ising spins undergoing a change of scale: The initial lattice consists of 36 spins (a) grouped into sets of 4 and rescaled to 9 spins (b).

ber is decreases during the above procedure referred to as *decimation*) unchanged, the distances are rescaled according to

$$\mathbf{r}' = l^{-1}\mathbf{r}. \quad (7.104)$$

During decimation, the partition function of the system

$$Q_N = \sum_{\{\sigma_j\}} \exp(-\beta H_N\{\sigma_j\}) \quad (7.105)$$

should remain formally unchanged. The singular, diverging part of the free energy $Nf^{(s)}$ should also remain unchanged:

$$N'f^{(s)}(t', h') = Nf^{(s)}(t, h), \quad (7.106)$$

where f is the free energy per spin; t and h are the reduced temperature and the reduced external field, respectively. This implies that

$$f^{(s)}(t, h) = l^{-d} f^{(s)}(t', h'). \quad (7.107)$$

Note that by replacing a l^d block of spins by a single spin, the magnitude of the interaction between the spins is changed, and thus neither the reduced temperature nor the reduced field remain the same. Close to the critical point, t is small and so should be t' ; we assume that

$$t' = l^{y_t} t, \quad (7.108)$$

where y_t is unknown at this stage. In a similar fashion

$$h' = l^{y_h} h. \quad (7.109)$$

This means that Eq. (7.107) reads

$$f^{(s)}(t, h) = l^{-d} f^{(s)}(l^{y_t} t, l^{y_h} h). \quad (7.110)$$

We expect that the singular part of the free energy should be insensitive to the change of scale, that is, we should be able to eliminate l from Eq. (7.110). This can be done if $f^{(s)}(t, h)$ is of the form

$$f^{(s)}(t, h) \propto \bar{f}(h/t^\Delta) \quad (7.111)$$

so that \bar{f} does not depend separately on t and on h but only on a combination h/t^Δ . By eliminating l from Eqs. (7.108) and (7.109) we find that $l = (t'/t)^{1/y_t} = (h'/h)^{1/y_h}$ so that the variable that \bar{f} may depend on is

$$\frac{h}{|t|^{y_h/y_t}} = \frac{h'}{|t'|^{y_h/y_t}}; \quad (7.112)$$

Δ is obviously equal to y_h/y_t . The proportionality coefficient in Eq. (7.111) must accommodate the factor of l^{-d} in Eq. (7.110); this is taken care of by setting

$$f^{(s)}(t, h) = |t|^{d/y_t} \bar{f}(h/|t|^\Delta). \quad (7.113)$$

Then, indeed,

$$\begin{aligned} f^{(s)}(t, h) &= |t|^{d/y_t} \bar{f}(h/|t|^\Delta) = l^{-d} \underbrace{|t'|^{d/y_t} \bar{f}(h'/|t'|^\Delta)}_{=f^{(s)}(t', h')} \\ &= \underbrace{l^{-d} |t'|^{d/y_t}}_{=|t|^{d/y_t}} \underbrace{\bar{f}(h'/|t'|^\Delta)}_{=h/|t|^\Delta} = |t|^{d/y_t} \bar{f}(h/|t|^\Delta). \end{aligned} \quad (7.114)$$

Now we can calculate the specific heat defined by $C^{(s)}(t, h = 0) = -\partial^2 f^{(s)}/\partial t^2$ (recall that in dimensional units, $c = (T/m)\partial S/\partial T = -(T/m) \times (\partial^2 F/\partial T^2)$). The dependence of $f^{(s)}$ on t is in the prefactor $|t|^{d/y_t}$ and in the argument of \bar{f} but note that

$$\left. \frac{\partial}{\partial t} \frac{h}{|t|} \right|_{h=0} = - \left. \frac{h}{|t|^2} \right|_{h=0} = 0 \quad (7.115)$$

and the same for the second derivative. Thus we conclude that

$$C^{(s)}(t, h = 0) \sim |t|^{d/y_t - 2} \underbrace{\bar{f}(h/|t|^\Delta = 0)}_{=const.} \quad (7.116)$$

so that

$$\alpha = 2 - d/y_t. \quad (7.117)$$

In a similar fashion, we can obtain the other critical exponents; the order parameter m , e.g., is given by $m = -\partial f^{(s)}/\partial h$ or

$$m \sim |t|^{d/y_t} |t|^{-\Delta} \bar{f}'(h/|t|^\Delta). \quad (7.118)$$

For $h = 0$, $\bar{f}'(h/|t|^\Delta)$ is some constant and we conclude that

$$\beta = d/y_t - \Delta = (d - y_h)/y_t. \quad (7.119)$$

The next two exponents we list without derivation:

$$\gamma = (2y_h - d)/y_t \quad (7.120)$$

and

$$\delta = y_h/(d - y_h). \quad (7.121)$$

In turn, the rescaled and the original correlation lengths ξ' and ξ are related by

$$\xi' = l^{-1}\xi. \quad (7.122)$$

Since the two lengths scale as $|t'|^{-\nu}$ and $|t|^{-\nu}$, respectively, we find that

$$\frac{\xi'}{\xi} = \left(\frac{t'}{t}\right)^{-\nu} = (l^{y_t})^{-\nu} = l^{-\nu y_t}. \quad (7.123)$$

Comparing this result to Eq. (7.122) gives

$$\nu = 1/y_t. \quad (7.124)$$

If this result is combined with Eq. (7.117) we obtain the Josephson equality $\alpha = 2 - d\nu$, which suggests that the scaling arguments used must be sound. More importantly, they allow us to express all critical exponents in terms of just two unknown parameters y_h and y_t together with the dimensionality of the system.

As a final demonstration, we consider the correlation function. At the critical point, it exhibits a power-law dependence so that

$$g(\mathbf{r}_1, \mathbf{r}_2) = \langle \sigma(\mathbf{r}_1)\sigma(\mathbf{r}_2) \rangle \sim r^{-(d-2+\eta)} \quad (7.125)$$

and after rescaling

$$g(\mathbf{r}'_1, \mathbf{r}'_2) = \langle \sigma'(\mathbf{r}'_1)\sigma'(\mathbf{r}'_2) \rangle \sim (r')^{-(d-2+\eta)}. \quad (7.126)$$

On rescaling, $r \rightarrow r' = l^{-1}r$ which implies that the spins themselves must transform as

$$\sigma'(\mathbf{r}') = l^{(d-2+\eta)/2}\sigma(\mathbf{r}) \quad (7.127)$$

so as to ensure that the correlation function remains unchanged. The exponent η may be obtained from the Fisher equality $\gamma = \nu(2 - \eta)$:

$$\eta = d + 2 - 2y_h. \quad (7.128)$$

1D Ising model

Let us now apply these ideas to the Ising model on a loop containing N spins (Fig. 7.8). To simplify the notation, we write the Hamiltonian H in reduced units so that

$$\begin{aligned} -\beta H\{\sigma_i\} &= -\beta \sum_{i=1}^N (-J\sigma_i\sigma_{i+1} - \mu B\sigma_i) \\ &= \sum_{i=1}^N \left[K_0 + K_1\sigma_i\sigma_{i+1} + \frac{1}{2}K_2(\sigma_i + \sigma_{i+1}) \right]. \end{aligned} \quad (7.129)$$

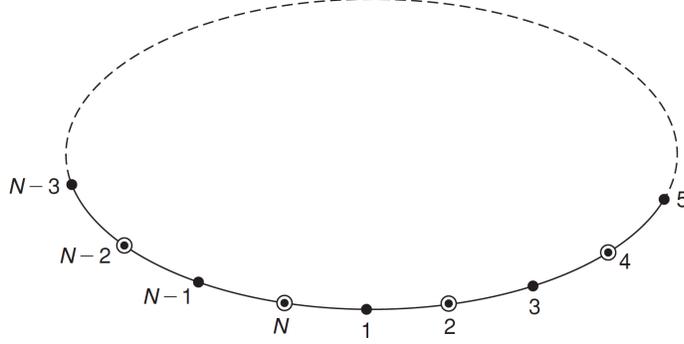


Figure 7.8: A loop of N spins constituting the 1D Ising model [1].

Here $K_0 = 0$, $K_1 = \beta J$, and $K_2 = \beta \mu B$; the coupling with the external field has been spelled out in a symmetrical form where each spin is counted twice, hence the factor of $1/2$. Now we sum the partition function

$$\begin{aligned} Q_N(B, T) &= \sum_{\{\sigma_i\}} \exp \left(\sum_{i=1}^N \left[K_0 + K_1 \sigma_i \sigma_{i+1} + \frac{1}{2} K_2 (\sigma_i + \sigma_{i+1}) \right] \right) \\ &= \sum_{\{\sigma_i\}} \prod_{i=1}^N \exp \left(K_0 + K_1 \sigma_i \sigma_{i+1} + \frac{1}{2} K_2 (\sigma_i + \sigma_{i+1}) \right) \end{aligned} \quad (7.130)$$

over all spins with even indices (highlighted by circles in Fig. 7.8). This is done by recasting the summand as follows:

$$\begin{aligned} &\prod_{i=1}^N \exp \left(K_0 + K_1 \sigma_i \sigma_{i+1} + \frac{1}{2} K_2 (\sigma_i + \sigma_{i+1}) \right) \\ &= \prod_{j=1}^{N/2} \exp \left(2K_0 + K_1 (\sigma_{2j-1} \sigma_{2j} + \sigma_{2j} \sigma_{2j+1}) \right. \\ &\quad \left. + \frac{1}{2} K_2 (\sigma_{2j-1} + 2\sigma_{2j} + \sigma_{2j+1}) \right). \end{aligned} \quad (7.131)$$

(Note that we have introduced j as the new index.) In the partition function, we sum over all spin configurations, that is, $\sigma_i = \pm 1$ and summation over all even spins can now be readily done. The factor affected by summation is

$$\begin{aligned} &\exp (K_1 (\sigma_{2j-1} \sigma_{2j} + \sigma_{2j} \sigma_{2j+1}) + K_2 \sigma_{2j}) \\ &= \exp ([K_1 (\sigma_{2j-1} + \sigma_{2j+1}) + K_2] \sigma_{2j}) \end{aligned} \quad (7.132)$$

and the result is obviously

$$\prod_{j=1}^{N/2} \exp(2K_0) 2 \cosh(K_1(\sigma_{2j-1} + \sigma_{2j+1}) + K_2) \exp\left(\frac{1}{2}K_2(\sigma_{2j-1} + \sigma_{2j+1})\right). \quad (7.133)$$

We relabel $\sigma_{2j-1} \rightarrow \sigma'_j$ etc. and $N/2 \rightarrow N'$ so that the partition function reads

$$Q_{N'}(B, T) = \sum_{\{\sigma'_j\}} \prod_{j=1}^{N'} \exp(2K_0) 2 \cosh(K_1(\sigma'_j + \sigma'_{j+1}) + K_2) \times \exp\left(\frac{1}{2}K_2(\sigma'_j + \sigma'_{j+1})\right) \quad (7.134)$$

and the key point of the procedure is to rewrite this result such that it is formally identical to the original partition function [Eq. (7.130)], that is

$$Q_{N'}(B, T) = \sum_{\{\sigma'_j\}} \prod_{j=1}^{N'} \exp\left(\left[K'_0 + K'_1\sigma'_j\sigma'_{j+1} + \frac{1}{2}K'_2(\sigma'_j + \sigma'_{j+1})\right]\right). \quad (7.135)$$

Note that the coefficients K'_0 , K'_1 , and K'_2 are rescaled to compared to the original ones.

The two expressions should be identical for any choice of σ'_j and σ'_{j+1} which implies that

$$\begin{aligned} & \exp\left(\left[K'_0 + K'_1\sigma'_j\sigma'_{j+1} + \frac{1}{2}K'_2(\sigma'_j + \sigma'_{j+1})\right]\right) \\ &= \exp(2K_0) 2 \cosh(K_1(\sigma'_j + \sigma'_{j+1}) + K_2) \exp\left(\frac{1}{2}K_2(\sigma'_j + \sigma'_{j+1})\right). \end{aligned} \quad (7.136)$$

Depending on the choice of σ'_j and σ'_{j+1} , this condition leads to three equations. Note that the spin-spin coupling term containing the product $\sigma'_j\sigma'_{j+1}$ can have only two values, ± 1 , whereas $\sigma'_j + \sigma'_{j+1}$ in the coupling to the external field can be either $-2, 0$, or $+2$. Thus we obtain three equations for

three unknowns K'_0, K'_1 , and K'_2 :

$$\begin{aligned}
\sigma'_j = \sigma'_{j+1} = -1 : \exp(K'_0 + K'_1 - K'_2) &= \exp(2K_0)2 \cosh(2K_1 - K_2) \exp(-K_2) \\
\sigma'_j = -\sigma'_{j+1} = \pm 1 : \exp(K'_0 - K'_1) = \exp(2K_0)2 \cosh(K_2) & \\
\sigma'_j = \sigma'_{j+1} = +1 : \exp(K'_0 + K'_1 + K'_2) &= \exp(2K_0)2 \cosh(2K_1 + K_2) \exp(K_2).
\end{aligned} \tag{7.137}$$

These equations can be solved for K'_0, K'_1 , and K'_2 . The fourth root of the product of the first, the second squared, and the last equation gives K'_0 :

$$\exp(K'_0) = 2 \exp(K_0) [\cosh(2K_1 - K_2) \cosh^2(K_2) \cosh(2K_1 + K_2)]^{1/4}. \tag{7.138}$$

K'_1 is obtained by dividing the product of the first and the last equation by the square of the second one:

$$\exp(K'_1) = \left[\frac{\cosh(2K_1 - K_2) \cosh(2K_1 + K_2)}{\cosh^2(K_2)} \right]^{1/4}. \tag{7.139}$$

By dividing the last equation by the first one we find that

$$\exp(K'_2) = \exp(K_2) \sqrt{\frac{\cosh(2K_1 + K_2)}{\cosh(2K_1 - K_2)}}. \tag{7.140}$$

At this stage, we realize that although apparently unnecessary, the K_0 term was needed because after summing over the even spins, the quadratic terms in σ'_j are partly demoted to linear terms and the linear terms originally only arising from the coupling to the external field are partly demoted to terms independent of σ'_j . To ensure that the structure of the Hamiltonian remains the same after rescaling, K_0 must be included from the beginning.

There also exists a more physical interpretation for the need for $K'_0 \neq 0$. To this end, we set $K_0 = 0$ and express $Q_N(K_1, K_2)$ [Eq. (7.130)] in terms of $Q_{N'}(K'_1, K'_2)$ [Eq. (7.135)]. In the latter, the factor involving K'_0 is trivial and we find

$$Q_N(K_1, K_2) = \exp(N'K'_0)Q_{N'}(K'_1, K'_2) \tag{7.141}$$

so that the free energy of the whole system reads

$$F_N(K_1, K_2) = -N'k_BTK'_0 + F_{N'}(K'_1, K'_2). \quad (7.142)$$

The free energy per spin $f = F/N$ is

$$f(K_1, K_2) = -\frac{1}{2}k_BTK'_0 + \frac{1}{2}f(K'_1, K'_2). \quad (7.143)$$

As $T \rightarrow \infty$, K_1 and K_2 as well as K'_1 and K'_2 go to 0. At the same time, Eq. (7.138) suggests that for $K_0 = 0$ $K'_0 = \ln 2$ so that

$$f(0, 0) = -2\frac{1}{2}k_BTK'_0 = -k_BT \ln 2. \quad (7.144)$$

This results represents the entropic part of the free energy per spin: Since there are two states, the entropy per spin is $k_B \ln 2$.

Zero field If $K_2 = 0$ then $K'_0 = \ln \left(2\sqrt{\cosh(2K_1)} \right)$, $K'_1 = \ln \sqrt{\cosh(2K_1)}$, and $K'_2 = 0$. This means that

$$f(K_1, 0) = -\frac{1}{2}k_BT \ln \left(2\sqrt{\cosh(2K_1)} \right) + \frac{1}{2}f \left(\ln \sqrt{\cosh(2K_1)}, 0 \right). \quad (7.145)$$

and the solution is

$$f(K_1, 0) = -k_BT \ln(2 \cosh K_1) = -k_BT \ln \left(2 \cosh(\beta J) \right), \quad (7.146)$$

which agrees with Eq. (7.66).

Paramagnet It is instructive to see what these results mean. If we set $J = 0$ then $K_1 = 0$ and there is no coupling between spins; this is the paramagnet. In this case, $K'_0 = \ln(2 \cosh K_2)$, $K'_1 = 0$, and $K'_2 = K_2$ so that the free energy per spin is

$$f(0, K_2) = -\frac{1}{2}k_BT \ln(2 \cosh K_2) + \frac{1}{2}f(0, K'_2). \quad (7.147)$$

so that

$$f(0, K_2) = -k_BT \ln(2 \cosh K_2) = -k_BT \ln \left(2 \cosh(\beta \mu B) \right). \quad (7.148)$$

This is a known result from the standard theory of paramagnetism.

Flow in parameter space and fixed points

The above example illustrates how one parametrizes the Hamiltonian using K_0, K_1 , and K_2 and how do these parameters change on decimation to K'_0, K'_1 , and K'_2 . Let us now look at this transformation from a more abstract perspective.

Consider a system with a Hamiltonian $H_{\{\sigma_l\}}(\{K_\alpha\})$ parametrized by a set of parameters $\{K_\alpha\}$. The free energy is determined by the partition function

$$\exp(-\beta F) = \sum_{\{\sigma_l\}} \exp(-\beta H_{\{\sigma_l\}}(\{K_\alpha\})). \quad (7.149)$$

In one decimation step, the number of degrees of freedom is reduced from N to N' such that $N' = l^{-d}N$, $l > 1$ being the change of scale and d being the dimension of the system. The correlation length decreases from ξ to $\xi' = l^{-1}\xi$. In this transformation, the parameters K_0, K_1 , and K_2 are replaced by K'_0, K'_1 , and K'_2 and the partition function reads

$$\exp(-\beta F) = \exp(N' K'_0) \sum_{\{\sigma'_l\}} \exp(-\beta H_{\{\sigma'_l\}}(\{K'_\alpha\})). \quad (7.150)$$

The free energy per spin, F/N , is given by

$$f(\{K_\alpha\}) = l^{-d} [-k_B T K'_0 + f(\{K'_\alpha\})]. \quad (7.151)$$

On decimation, the vector of parameters \mathbf{K} changes to \mathbf{K}' , which may be viewed as flow in the parameter space (Fig. 7.9). This flow can be concisely encoded by

$$\mathbf{K}' = \mathcal{R}_l(\mathbf{K}), \quad (7.152)$$

where \mathcal{R}_l is the renormalization group (RG) operator, the subscript denoting the change of scale. The name "renormalization group" was introduced because

$$\mathcal{R}_l(\mathcal{R}_l(\mathbf{K})) = \mathcal{R}_{l^2}(\mathbf{K}). \quad (7.153)$$

Speaking of a group is however not entirely sound because there exists no inverse operation \mathcal{R}_l^{-1} .

Repeated application of the RG operator produces a sequence $\mathbf{K}, \mathbf{K}', \mathbf{K}'', \dots$ where

$$\mathbf{K}^{(n)} = \mathcal{R}_l(\mathbf{K}^{(n-1)}) = \mathcal{R}_l^n(\mathbf{K}). \quad (7.154)$$

The corresponding correlation length and free energy per spin read

$$\xi^{(n)} = l^{-n} \xi^{(0)} \quad (7.155)$$

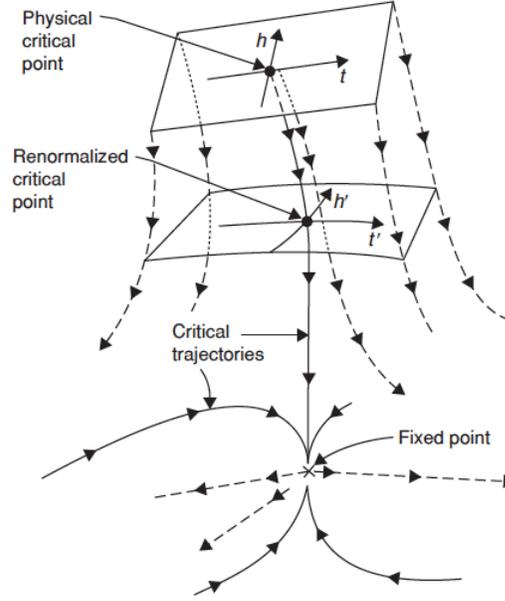


Figure 7.9: Schematic representation of flow in parameter space under successive RG transformations [1].

and

$$[f^{(s)}]^{(n)} = l^{-nd} [f^{(s)}]^{(0)}. \quad (7.156)$$

The transformation Eq. (7.152) may have a fixed point \mathbf{K}^* such that

$$\mathbf{K}^* = \mathcal{R}_l(\mathbf{K}^*). \quad (7.157)$$

This also means that in the critical point,

$$\xi(\mathbf{K}^*) = l^{-1}\xi(\mathbf{K}^*), \quad (7.158)$$

which is solved by $\xi = 0$ or by $\xi \rightarrow \infty$, the latter case corresponding to the critical point. By determining the fixed point of the RG transformation, we thus locate the phase transition.

Vectors in the parameter space may be expanded in terms of suitable basis vectors; in the vicinity of the fixed point, the most natural choice are the eigenvectors of \mathcal{A} given by \mathcal{R} linearized at \mathbf{K}^* . The coefficients of \mathbf{K} close to \mathbf{K}^* in this basis are commonly referred to as the *scaling fields*. Rather naturally, not all scaling fields play the same role:

- Those driving the system away from the fixed point are *relevant variables*; in the Ising model, t and h are relevant variables.
- Those driving the system towards the fixed point are *irrelevant variables*.
- Those that do not drive the system away nor towards the fixed point are *marginal variables*.

There exist a continuum of trajectories leading to the fixed point; they lie on the *critical surface*.

Fixed points of 1D Ising model

The RG transformation is explicitly given by

$$K'_1 = \frac{1}{4} \ln \frac{\cosh(2K_1 + K_2) \cosh(2K_1 - K_2)}{\cosh^2 K_2}, \quad (7.159)$$

$$K'_2 = K_2 + \frac{1}{2} \ln \frac{\cosh(2K_1 + K_2)}{\cosh(2K_1 - K_2)}. \quad (7.160)$$

(Note that K'_0 is determined by K_0, K_1 , and K_2 whereas K'_1 and K'_2 only depend on K_1 and K_2 but not on K_0 . Thus the flow in the parameter space is determined by K_1 and K_2 alone.)

The flow in the parameter space is shown in Fig. 7.10 which suggests that there exists a line of trivial fixed points at $K_1 = 0$ (that is, $J = 0$ or $T \rightarrow \infty$) and arbitrary K_2 (that is, arbitrary B). These fixed points are trivial because in absence of coupling, the correlation length is zero. The other fixed point is at $K_1 \rightarrow \infty$, that is at $T = 0$, and $K_2 = 0$. In this case, the correlation length is infinite. Given that there exists no critical point at a finite K_1 , we conclude that the 1D Ising model does not exhibit a phase transition.

We can use these results further. In the vicinity of the stable fixed point $K_1 \rightarrow \infty$, the above transformation rules reduce to

$$K'_1 \approx K_1 - \frac{\ln 2}{2}, \quad (7.161)$$

$$K'_2 = 2K_2. \quad (7.162)$$

We are interested in the behavior of the system close to the fixed point but given that $K_1^* \rightarrow \infty$, we introduce a dimensionless temperature parameter by

$$t = \exp(-pK_1) = \exp(-p\beta J), \quad (7.163)$$

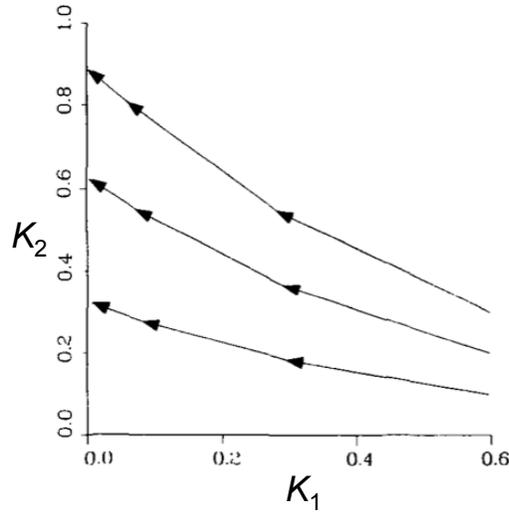


Figure 7.10: Flow in 1D Ising model [4]; the horizontal axis is the dimensionless exchange coupling βJ and the vertical axis is the dimensionless external field $\beta\mu B$.

where $p > 0$. This means that close to the fixed point $\exp(-pK'_1) = \exp(-pK_1 + p(\ln 2)/2)$ or

$$t' = 2^{p/2}t. \quad (7.164)$$

As the scale factor in the Ising decimation is 2, we find the the exponents introduced in Eqs. (7.108) and (7.109) are

$$y_1(\text{temperature}) = p/2 \quad \text{and} \quad y_2(\text{external field}) = 1. \quad (7.165)$$

Using results from the beginning of this Section we find

$$\alpha = 2 - 2/p, \quad \beta = 0, \quad \gamma = 2/p, \quad \delta \rightarrow \infty, \quad \eta = 1, \quad \text{and} \quad \nu = 2/p. \quad (7.166)$$

The inherent arbitrariness of these results lies in the way the dimensionless temperature parameter t was introduced.

Chapter 8

Fluctuations

When discussing the microcanonical and the canonical ensembles, we found that in the thermodynamic limit the fluctuations usually do not play an important role in the sense that the mean energy in the canonical ensemble is well-defined — and essentially the same as the energy in the microcanonical ensemble. However, at the critical point fluctuations are an essential feature of the system. Here we study the fluctuations in more detail, going from a more general discussion to Brownian motion and the fluctuation-dissipation theorem as examples of phenomena relying on fluctuations.

We begin by examining the fluctuations of extensive and intensive quantities in equilibrium. Consider two canonical systems in contact; one of the systems (labeled by "2") may be viewed as the reservoir. The total energy E and volume V are fixed and so are the numbers of particles in part 1 and 2 N_1 and N_2 , respectively. The total energy consists of equilibrium \bar{E}_1 and \bar{E}_2 determined by the condition that the parts are at the same temperature T^* . The same goes for the average volumes \bar{V}_1 and \bar{V}_2 which are such that the pressure in the two parts P^* is the same.

We denote the entropy of the equilibrium state by S_0 . Given that the whole system 1 + 2 is microcanonical, S_0 is maximal and any fluctuation must lead to a decrease of the entropy $\Delta S = S - S_0 = k_B \ln \Omega_f - k_B \ln \Omega_0$, where Ω_f and Ω_0 are the respective numbers of microstates. The probability that the fluctuation will take place is

$$p \propto \Omega_f \propto \exp(\Delta S/k_B). \quad (8.1)$$

The entropy increase may be divided into terms corresponding to the small

system "1" and to the reservoir "2":

$$\Delta S = \underbrace{\Delta S_1}_{\text{system}} + \underbrace{\Delta S_2}_{\text{reservoir}} = \Delta S_1 + \int_0^f \frac{dE_2 + P_2 dV_2}{T_2}, \quad (8.2)$$

where we used the first law in expressing dS_2 in terms of dE_2 and dV_2 . In principle, the temperature of the reservoir may be different from the common temperature T^* but the difference is tiny; this is because the reservoir is supposed to be much larger than system 1. The same goes for the pressure $P_2 = P^*$. At the same time, $dE_2 = -dE_1$ (because the whole system 1 + 2 is microcanonical) and $dV_2 = -dV_1$ (because the volume of the whole system is fixed). Thus we can express the entropy decrease in terms of quantities pertaining solely to system 1:

$$\Delta S = \Delta S_1 - \frac{\Delta E_1 + P^* \Delta V_1}{T^*}, \quad (8.3)$$

where we have integrated from the equilibrium state to the final state. This means that $p \propto \exp(-(\Delta E_1 - T^* \Delta S_1 + P^* \Delta V_1)/(k_B T^*))$. As this conclusion applies to any part of a large system at some given temperature T and pressure P , the indices can be disposed of and we find that

$$p \propto \exp\left(-\frac{\Delta E - T \Delta S + P \Delta V}{k_B T}\right). \quad (8.4)$$

The three extensive quantities appearing in this result are not independent. Let us express ΔE in terms of ΔS and ΔV — to second order, because in equilibrium the first-order terms $(\partial E / \partial S)_V \Delta S = T \Delta S$ and $(\partial E / \partial V)_S \Delta V = -P \Delta V$ determined by the first law cancel with the other two terms in Eq. (8.4). Thus

$$\begin{aligned} \Delta E &= \left(\frac{\partial E}{\partial S}\right)_V \Delta S + \left(\frac{\partial E}{\partial V}\right)_S \Delta V \\ &+ \frac{1}{2} \left[\left(\frac{\partial^2 E}{\partial S^2}\right)_V (\Delta S)^2 + 2 \left(\frac{\partial^2 E}{\partial S \partial V}\right) \Delta S \Delta V + \left(\frac{\partial^2 E}{\partial V^2}\right)_S (\Delta V)^2 \right] + \dots \end{aligned} \quad (8.5)$$

The second derivatives can be further rearranged:

$$\begin{aligned}
& \left(\frac{\partial^2 E}{\partial S^2} \right)_V (\Delta S)^2 + \left(\frac{\partial^2 E}{\partial S \partial V} \right) \Delta S \Delta V \\
&= \left[\left(\frac{\partial^2 E}{\partial S^2} \right)_V \Delta S + \left(\frac{\partial^2 E}{\partial S \partial V} \right) \Delta V \right] \Delta S \\
&= \left[\frac{\partial}{\partial S} \left(\frac{\partial E}{\partial S} \right)_V \Delta S + \frac{\partial}{\partial V} \left(\frac{\partial E}{\partial S} \right)_V \Delta V \right] \Delta S \\
&= \Delta \left(\frac{\partial E}{\partial S} \right)_V \Delta S = \Delta T \Delta S.
\end{aligned} \tag{8.6}$$

In a similar fashion, we find that

$$\left(\frac{\partial^2 E}{\partial V^2} \right)_S (\Delta V)^2 + \left(\frac{\partial^2 E}{\partial V \partial S} \right) \Delta V \Delta S = -\Delta P \Delta V \tag{8.7}$$

so that

$$p \propto \exp \left(-\frac{\Delta T \Delta S - \Delta P \Delta V}{2k_B T} \right). \tag{8.8}$$

The four thermodynamic variables appearing in this equation are not independent. If we choose T and V as independent variables, then

$$\Delta S = \left(\frac{\partial S}{\partial T} \right)_V \Delta T + \left(\frac{\partial S}{\partial V} \right)_T \Delta V = \frac{C_V}{T} \Delta T + \frac{\beta}{\chi_T} \Delta V \tag{8.9}$$

and

$$\Delta P = \left(\frac{\partial P}{\partial T} \right)_V \Delta T + \left(\frac{\partial P}{\partial V} \right)_T \Delta V = \frac{\beta}{\chi_T} \Delta T - \frac{1}{\chi_T V} \Delta V. \tag{8.10}$$

On inserting Eqs. (8.9) and (8.10) in Eq. (8.8) we find that the terms arising from the second term in Eq. (8.9) and first term in Eq. (8.10) cancel so that the result is

$$p \propto \exp \left(-\frac{C_V}{2k_B T^2} (\Delta T)^2 - \frac{1}{2k_B T \chi_T V} (\Delta V)^2 \right). \tag{8.11}$$

This means that the fluctuations of T and V are independent Gaussian variables so that $\overline{\Delta T \Delta V} = 0$; the averages

$$\overline{(\Delta T)^2} = \frac{k_B T^2}{C_V} \quad \text{and} \quad \overline{(\Delta V)^2} = k_B T \chi_T V. \tag{8.12}$$

Note that the root-mean-square fluctuation of temperature as an intensive variable is inversely proportional to the root of heat capacity as an extensive quantity and thus to the size of the system. On the other hand, the root-mean-square fluctuation of the volume as an extensive quantity is proportional to the root of the volume so that the relative root-mean-square fluctuations are again inversely proportional to the root of the volume, that is, to the size of the system.

Instead of T and V , we may choose S and P as independent variables. Then

$$p \propto \exp\left(-\frac{1}{2k_B C_P}(\Delta S)^2 - \frac{\chi_S V}{2k_B T}(\Delta P)^2\right). \quad (8.13)$$

and

$$\overline{(\Delta S)^2} = k_B C_P \quad \text{and} \quad \overline{(\Delta P)^2} = \frac{k_B T}{\chi_S V}. \quad (8.14)$$

Using Eqs. (8.12) we can readily calculate the energy fluctuations. We start with

$$\Delta E = \left(\frac{\partial E}{\partial T}\right)_V \Delta T + \left(\frac{\partial E}{\partial V}\right)_T \Delta V, \quad (8.15)$$

square it and evaluate the average; at this point recall that $\overline{\Delta T \Delta V} = 0$ and that $(\partial E/\partial T)_V = C_V$. Without proof we generalize the result to the grand-canonical distribution where the change of volume is due to a change of the number of particles. In this case,

$$\begin{aligned} \overline{(\Delta E)^2} &= C_V^2 \frac{k_B T^2}{C_V} + \left[\left(\frac{\partial E}{\partial V}\right)_T\right]^2 k_B T \chi_T V \\ &= C_V k_B T^2 + \left[\left(\frac{\partial E}{\partial N}\right)_T \left(\frac{\partial N}{\partial V}\right)_T\right]^2 k_B T \chi_T V \\ &= C_V k_B T^2 + \left[\left(\frac{\partial E}{\partial N}\right)_T \frac{N}{V}\right]^2 k_B T \chi_T V \\ &= C_V k_B T^2 + \left[\left(\frac{\partial E}{\partial N}\right)_T\right]^2 \frac{N^2}{V} k_B T \chi_T. \end{aligned} \quad (8.16)$$

From Eq. (2.92) we know how the fluctuations in number density are related to the number of particles itself. By identifying the average number of particles with the fixed number of particles in the canonical ensemble, we have

$$\overline{(\Delta N)^2} = \frac{N^2 k_B T \chi_T}{V}. \quad (8.17)$$

This means that

$$\overline{(\Delta E)^2} = C_V k_B T^2 + \left[\left(\frac{\partial E}{\partial N} \right)_T \right]^2 \overline{(\Delta N)^2}. \quad (8.18)$$

Equation (8.18) may be interpreted in various ways. In a canonical system with fixed N, V , and T , the second term is zero and $\overline{(\Delta E)^2} = k_B T^2 C_V$, which is a well-known result. In a grand-microcanonical system with fixed energy, $\Delta E = 0$ and thus Eq. (8.18) leads to

$$C_V^2 \overline{(\Delta T)^2} = - \left[\left(\frac{\partial E}{\partial N} \right)_T \right]^2 \overline{(\Delta N)^2}, \quad (8.19)$$

where we have reintroduced $\overline{(\Delta T)^2}$ from Eq. (8.12). This says that if the number of particles will fluctuate at fixed total energy, so will the temperature.

8.1 Brownian motion

One of the most important phenomena involving fluctuations is Brownian motion, that is the motion of a colloidal particle due to impacts of the molecules of the surrounding liquid. Here we present three different views of this phenomenon.

The first approach is due to Einstein who considered Brownian motion as a random walk. In the 1D simplified version, one is interested in the location of a particle initially at $x = 0$ at some finite time t . Imagine that the particle makes uncorrelated discrete steps of constant length l in time intervals of τ^* ; this is the random walk on a lattice. Steps along either direction of the x axis are equally probable so that Δx may be either $+l$ or $-l$. The probability that after $n = t/\tau^*$ steps the particle is m steps from the origin is determined by the phase space of paths with m more steps in the positive direction than in the negative direction, the former and the latter being $(n+m)/2$ and $(n-m)/2$, respectively. The probability that this be the case is given by the binomial distribution with $f = 1/2$

$$\begin{aligned} p_n(m) &= \binom{n}{(n+m)/2} f^{(n+m)/2} (1-f)^{n-(n+m)/2} \\ &= \frac{n!}{[(n+m)/2]! [(n-m)/2]!} \left(\frac{1}{2} \right)^n. \end{aligned} \quad (8.20)$$

This gives $\overline{m} = 0$ and $\overline{m^2} = n$ corresponding to $\overline{x(t)} = 0$ and

$$\overline{x^2(t)} = \frac{l^2 t}{\tau^*}, \quad (8.21)$$

respectively. (The last equation was obtained by multiplying $\overline{m^2} = n$ by l^2 and replacing n by t/τ^* .) The root-mean-square displacement is thus proportional to $t^{1/2}$.

Smoluchowski equation

An alternative description of this process was proposed by Smoluchowski who considered the conditional probability $p_n(x_0|x)$ that a particle initially at x_0 reached point x measured in units of l in n steps. Clearly

$$p_n(x_0|x) = \sum_{z=-\infty}^{\infty} p_{n-1}(x_0|z)p_1(z|x). \quad (8.22)$$

Note that $p_0(z|x) = \delta_{z,x}$ (making no steps does not take the particle anywhere) and that since steps in either direction are equiprobable, $p_1(z|x) = (\delta_{z,x-1} + \delta_{z,x+1})/2$. Using this result we recognize that Eq. (8.22) is a recurrence equation

$$p_n(x_0|x) = \frac{1}{2} [p_{n-1}(x_0|x-1) + p_{n-1}(x_0|x+1)]. \quad (8.23)$$

This states that the probability of making it in n steps to x is half of the combined probability of making it in $n-1$ steps to $x-1$ (wherefrom one can go to x in the n th step with probability of $1/2$) and of making it in $n-1$ steps to $x+1$ (wherefrom one can go to x in the n th step too). This equation can be solved using the generating function

$$Q_n(\xi) = \sum_{x=-\infty}^{\infty} p_n(x_0|x)\xi^{x-x_0}. \quad (8.24)$$

To express Eq. (8.23) in terms of $Q_n(\xi)$, we multiply it by ξ^{x-x_0}

$$\begin{aligned} p_n(x_0|x)\xi^{x-x_0} &= \frac{1}{2} [p_{n-1}(x_0|x-1)\xi^{x-x_0} + p_{n-1}(x_0|x+1)\xi^{x-x_0}] \\ &= \frac{1}{2} [\xi p_{n-1}(x_0|x-1)\xi^{x-1-x_0} + \xi^{-1} p_{n-1}(x_0|x+1)\xi^{x+1-x_0}] \end{aligned} \quad (8.25)$$

and sum over all x which gives

$$Q_n(\xi) = \frac{1}{2} [\xi + \xi^{-1}] Q_{n-1}(\xi). \quad (8.26)$$

Given that

$$Q_0(\xi) = \sum_{x=-\infty}^{\infty} p_0(x_0|x) \xi^{x-x_0} = \sum_{x=-\infty}^{\infty} \delta_{x_0,x} \xi^{x-x_0} = 1, \quad (8.27)$$

we find that $Q_1 = (1/2) [\xi + \xi^{-1}]$ and

$$Q_n(\xi) = \left(\frac{1}{2}\right)^n [\xi + \xi^{-1}]^n. \quad (8.28)$$

The idea of the generating function is that if known in closed form, it can be used to calculate the coefficients $p_n(x_0|x)$. To this end, we expand the binomial, obtaining a finite sum of powers of ξ with exponents from n to $-n$. By comparing this to the definition Eq. (8.24), we immediately see that $p_n(x_0|x) = 0$ if $|x - x_0| > n$ since these large powers are just not included in the expanded binomial. On the other hand, if $|x - x_0| \leq n$ then

$$p_n(x_0|x) = \left(\frac{1}{2}\right)^n \frac{n!}{[(n+x-x_0)/2]! [(n-x+x_0)/2]!} \quad (8.29)$$

which is consistent with Einstein's result Eq. (8.20).

At this point, we may work out the continuum limit. To this end, we calculate the logarithm of $p_n(m)$ [Eq. (8.20)] which allows us to approximate the factorials using the Stirling formula. This gives

$$\ln p_n(m) = n \ln n - \frac{n+m}{2} \ln \frac{n+m}{2} - \frac{n-m}{2} \ln \frac{n-m}{2} - n \ln 2. \quad (8.30)$$

If $m \ll n$,

$$\begin{aligned} \ln p_n(m) &\approx n \ln n - \frac{n+m}{2} \ln \frac{n}{2} - \frac{n+m}{2} \frac{m}{2n} - \frac{n-m}{2} \ln \frac{n}{2} + \frac{n-m}{2} \frac{m}{2n} \\ &\quad - n \ln 2 \\ &\approx -\frac{m^2}{2n}. \end{aligned} \quad (8.31)$$

Apart from the prefactor which requires a somewhat more detailed calculation (specifically, a more accurate version of the Stirling formula), this gives a Gaussian probability distribution

$$p_n(m) \approx \frac{2}{\sqrt{2\pi n}} \exp\left(-\frac{m^2}{2n}\right). \quad (8.32)$$

(Note that if n is odd, one cannot make it to an even m : For example, with $n = 3$ steps one can reach $m = 3, 1, -1$, and -3 but not $m = 2, 0$, and -2 simply because it takes a pair of steps in opposite direction to decrease m . Conversely, an even n means that odd m 's cannot be reached. Thus $\Delta m = 2$ rather than 1.)

Using $x/l = m$, $n = t/\tau^*$ we find that

$$\frac{m^2}{2n} = \frac{x^2}{2l^2t/\tau^*} = \frac{x^2}{4Dt} \quad (8.33)$$

so that

$$p(x)dx = \frac{dx}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right). \quad (8.34)$$

where

$$D = \frac{l^2}{2\tau^*} \quad (8.35)$$

is the diffusion constant.

This result suggests that Brownian motion may be viewed as diffusion. Indeed, by considering an ensemble of N particles described by their number density $n(\mathbf{r}, t)$ rather than a single particle, we find that the evolution of n is described by the diffusion equation

$$\frac{\partial n(\mathbf{r}, t)}{\partial t} = D\nabla^2 n(\mathbf{r}, t), \quad (8.36)$$

its Green function in 3D being

$$n(\mathbf{r}, t) = \frac{N}{\sqrt{4\pi Dt}^3} \exp\left(-\frac{r^2}{4Dt}\right); \quad (8.37)$$

This distribution is normalized to N and readily gives

$$\langle r^2 \rangle = \frac{1}{N} \int_0^\infty n(\mathbf{r}, t) 4\pi r^4 dr = 6Dt, \quad (8.38)$$

which agrees with previous results obtained in 1D.

Langevin theory

Yet another view of the Brownian motion is provided by the Langevin theory dealing with the deterministic Newton's equation of motion involving a random force $\mathbf{F}(t)$ such that $\overline{\mathbf{F}(t)} = 0$:

$$M \frac{d\mathbf{v}}{dt} + \frac{\mathbf{v}}{B} = \mathbf{F}(t). \quad (8.39)$$

Here M is the mass of the particle and B is its mobility. The homogeneous part of the equation is characterized by an exponential decay of the initial velocity

$$\mathbf{v}(t) = \mathbf{v}(0) \exp(-t/\tau), \quad (8.40)$$

where

$$\tau = MB \quad (8.41)$$

is the relaxation time. On dividing the equation of motion by M and introducing $\mathbf{A}(t) = \mathbf{F}(t)/M$, we have

$$\frac{d\mathbf{v}}{dt} + \frac{\mathbf{v}}{\tau} = \mathbf{A}(t). \quad (8.42)$$

Now we dot this equation with \mathbf{r} :

$$\mathbf{r} \cdot \frac{d\mathbf{v}}{dt} + \mathbf{r} \cdot \frac{\mathbf{v}}{\tau} = \mathbf{r} \cdot \mathbf{A}(t). \quad (8.43)$$

The first term is equal to $(d^2r^2/dt^2)/2 - v^2$, in the second term $\mathbf{r} \cdot \mathbf{v} = (dr^2/dt)/2$, and the last term vanishes on averaging over the ensemble. Thus

$$\frac{1}{2} \frac{d^2 \langle r^2 \rangle}{dt^2} - \langle v^2 \rangle + \frac{1}{2\tau} \frac{d \langle r^2 \rangle}{dt} = 0. \quad (8.44)$$

In thermal equilibrium, $\langle v^2 \rangle = 3k_B T/M$ so that the equation of motion for r^2 is

$$\frac{1}{2} \frac{d^2 \langle r^2 \rangle}{dt^2} + \frac{1}{2\tau} \frac{d \langle r^2 \rangle}{dt} = \frac{3k_B T}{M}. \quad (8.45)$$

On integration over t , we have

$$\langle r^2(t) \rangle = \frac{6k_B T \tau^2}{M} \left\{ \frac{t}{\tau} - \left[1 - \exp\left(-\frac{t}{\tau}\right) \right] \right\}. \quad (8.46)$$

If $t \ll \tau$, we expand the exponential to second order to find that

$$\langle r^2(t \ll \tau) \rangle = \frac{3k_B T t^2}{M} = \langle v^2 \rangle t^2. \quad (8.47)$$

In this regime, the motion of particles is ballistic and so $r = \sqrt{\langle v^2 \rangle} t$. On the other hand, for $t \gg \tau$

$$\langle r^2(t) \rangle = \frac{6k_B T \tau t}{M} = 6k_B T B t, \quad (8.48)$$

which is of the same form as the result of Einstein and Smoluchowski so that

$$D = k_B T B. \quad (8.49)$$

This is known as the Einstein relation.

In these results, the fluctuating force is not directly included because we have assumed that $\langle v^2 \rangle$ has reached the equilibrium value of $3k_B T/M$. At short times, this is not true and now we integrate the inhomogeneous Eq. (8.42) so as to find the dependence of $\mathbf{v}(t)$ on $\mathbf{A}(t)$. Formally, the solution includes the complementary integral and the particular integral:

$$\mathbf{v}(t) = \underbrace{\mathbf{v}(0) \exp(-t/\tau)}_{\text{complementary integral}} + \underbrace{\exp(-t/\tau) \int_0^t \exp(u/\tau) \mathbf{A}(u) du}_{\text{particular integral}}. \quad (8.50)$$

Since $\langle \mathbf{A}(u) \rangle = 0$, the average velocity $\langle \mathbf{v}(t) \rangle = \mathbf{v}(0) \exp(-t/\tau)$ does not depend on $\mathbf{A}(t)$ but the average square velocity does:

$$\begin{aligned} \langle v^2(t) \rangle &= v^2(0) \exp(-2t/\tau) + 2 \exp(-2t/\tau) \mathbf{v}(0) \cdot \int_0^t \exp(u/\tau) \langle \mathbf{A}(u) \rangle du \\ &\quad + \exp(-2t/\tau) \int_0^t \exp((u_1 + u_2)/\tau) \langle \mathbf{A}(u_1) \cdot \mathbf{A}(u_2) \rangle du_1 du_2. \end{aligned} \quad (8.51)$$

The second term vanishes because $\langle \mathbf{A}(u) \rangle = 0$ but the third term, which contains the autocorrelation function of the random force $\langle \mathbf{A}(u_1) \cdot \mathbf{A}(u_2) \rangle$, does not.

The evaluation of the integral in Eq. (8.51) involves both the generic properties of the autocorrelation function and the fact that for large times t , $\langle v^2(t) \rangle$ must tend to $3k_B T/M$. The final result is

$$\langle v^2(t) \rangle = v^2(0) + \left[\frac{3k_B T}{M} - v^2(0) \right] \left[1 - \exp\left(-\frac{2t}{\tau}\right) \right]. \quad (8.52)$$

Figure 8.1 shows how the mean square velocity approaches the limiting value of $3k_B T/M$.

On inserting this result into Eq. (8.44) and solving it, we find that

$$\begin{aligned} \langle r^2(t) \rangle &= v^2(0) \tau^2 \left[1 - \exp\left(-\frac{t}{\tau}\right) \right]^2 \\ &\quad - \frac{3k_B T}{M} \tau^2 \left[1 - \exp\left(-\frac{t}{\tau}\right) \right] \left[3 - \exp\left(-\frac{t}{\tau}\right) \right] + \frac{6k_B T \tau}{M} t. \end{aligned} \quad (8.53)$$

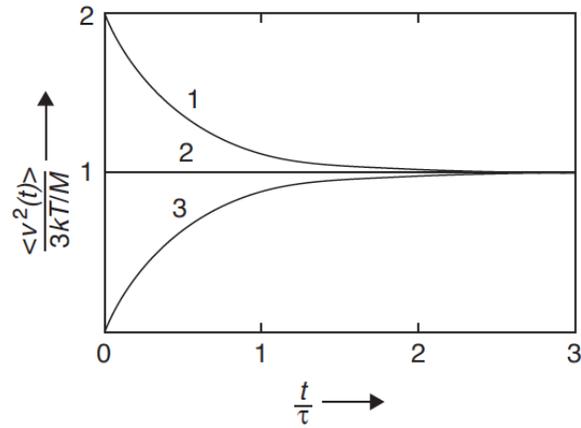


Figure 8.1: Mean square velocity of a particle executing Brownian motion within the Langevin theory for $v^2(0) = 6k_B T/M, 3k_B T/M$, and 0 [curves (1), (2), and (3), respectively] [1].

This result contains both the short-time ballistic behavior $\langle r^2 \rangle = v^2(0)t^2$ and the long-time diffusive regime where $\langle r^2 \rangle = 6k_B T B t$. The crossover between the two regimes is seen in Fig. 8.2.

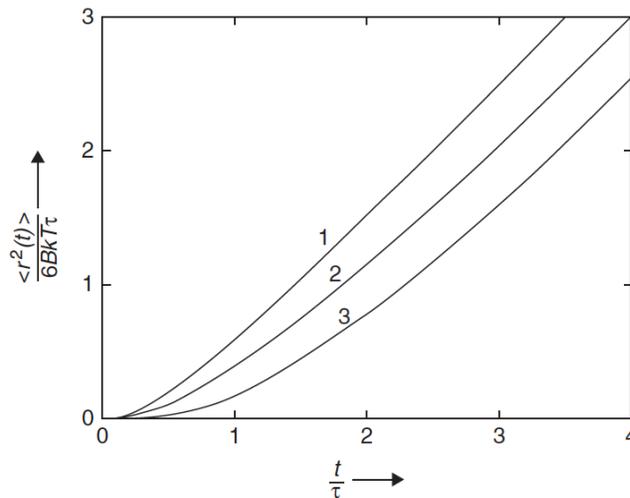


Figure 8.2: Mean square displacement of a particle executing Brownian motion within the Langevin theory for $v^2(0) = 6k_B T/M, 3k_B T/M$, and 0 [curves (1), (2), and (3), respectively] [1].

8.2 Fokker-Planck equation

When discussing the motion of an ensemble of particles undergoing Brownian motion we ended up with the diffusion equation as a framework describing the evolution of the distribution of the particles. A more general scheme governing the dynamics of the distribution is known as the Master Equation. We illustrate it using a 1D system. Here one considers the probability $f(x, t)$ that at a time t , the particle be at a given position between $x - dx/2$ to $x + dx/2$. This probability is normalized,

$$\int_{-\infty}^{\infty} f(x, t) dx = 1. \quad (8.54)$$

The Master Equation states that the change of the probability that the particle be at x is due to the probability that it move from some other position x' to x less the probability that it move from x to some other position x' :

$$\frac{\partial f}{\partial t} = \int_{-\infty}^{\infty} [-f(x, t)W(x, x') + f(x', t)W(x', x)] dx'. \quad (8.55)$$

Here $W(x, x')dx'dt$ is the probability that the particle move from x to an interval of width of dx' at x' during a period of dt .

We are interested in the way and non-equilibrium distribution approaches the equilibrium distribution f_0 where by definition $\partial f_0/\partial t = 0$. The simplest model describing this relaxation is of the form

$$\frac{\partial f}{\partial t} = -\frac{f - f_0}{\tau}. \quad (8.56)$$

Obviously, if $f = f_0$ the right-hand side vanishes and so the equation is satisfied. The minus sign on the right-hand side ensures that f decreases if $f - f_0 > 0$ so that f indeed approached f_0 , and τ is a phenomenological relaxation time. — At the level of the order parameter m in, e.g., a Landau theory, a similar model relaxation dynamics is known as the Landau-Khalatnikov equation

$$\frac{\partial m}{\partial t} = -\Gamma \frac{\delta F}{\delta m}, \quad (8.57)$$

where Γ is a suitable kinetic coefficient controlling the speed of approach to the equilibrium where $m = m_0$. Close to the equilibrium, $f \propto (m - m_0)^2$ so that $\delta F/\delta m \propto (m - m_0)$ which leads to an equation formally similar to Eq. (8.56).

The integral representation of the evolution of the probability distribution described by the Master Equation is the Chapman-Kolmogorov equation [8]:

$$f(x, t + \varepsilon) = \int_{-\infty}^{\infty} w(x, x'; \varepsilon) f(x', t) dx'. \quad (8.58)$$

Here ε is the discretized time step, and $w(x, x'; \varepsilon)$ is the probability that the particle moves from x' to x during this time interval. In general, a Brownian particle is not subject only to thermal diffusive motion but also to an external force leading to the drift of the particle in a given direction; in our 1D model, in either direction along the x axis. If the motion were strictly deterministic, then

$$x = x' + v(x')\varepsilon, \quad (8.59)$$

where $v(x')$ is the drift velocity, and

$$w(x, x'; \varepsilon) = \delta(x - x' - v\varepsilon). \quad (8.60)$$

But the motion of a Brownian particle is not deterministic and so

$$x = x' + v(x')\varepsilon + \eta_\varepsilon(t), \quad (8.61)$$

where $\eta_\varepsilon(t)$ is a stochastic displacement usually assumed to be Gaussian. This stochastic displacement means that $w(x, x'; \varepsilon)$ will not be a delta function but will be a finite-width peaked function centered at $x' - v(x')\varepsilon$:

$$w(x, x'; \varepsilon) = \frac{1}{\sqrt{4\pi D\varepsilon}} \exp\left(-\frac{[x - x' - v(x')\varepsilon]^2}{4D\varepsilon}\right). \quad (8.62)$$

Now the left-hand side of Eq. (8.58) can be expanded to first order, and on inserting the above $w(x, x'; \varepsilon)$ we have

$$\underbrace{f(x, t) + \frac{\partial f(x, t)}{\partial t} \varepsilon}_{=f(x, t+\varepsilon)} = \int_{-\infty}^{\infty} \frac{\exp(-[x - x' - v(x')\varepsilon]^2/(4D\varepsilon))}{\sqrt{4\pi D\varepsilon}} f(x', t) dx'. \quad (8.63)$$

One immediately sees that if $\varepsilon \rightarrow 0$, the Gaussian reduced to a delta function which cancels with $f(x, t)$ on the left-hand side; the second term on the left-hand side vanishes too. In the next step, we need to expand the right-hand side to first order in ε , and this is a somewhat tedious step. Then we can equate the linear term on the two sides, finding the Fokker-Planck equation:

$$\frac{\partial f(x, t)}{\partial t} = -\frac{\partial}{\partial x} [v(x)f(x, t)] + D \frac{\partial^2}{\partial x^2} f(x, t). \quad (8.64)$$

This variant of the Fokker-Planck equation where one considers the distribution of particle positions is also referred to as the Smoluchowski equation; in general, the Fokker-Planck equation may involve any distribution describing the state of a system, which may depend on several variables rather than just on one.

The second term in the Fokker-Planck equation describes diffusion, whereas the first one is best appreciated by setting $v = \text{const.}$ and $D = 0$. In this case,

$$\frac{\partial f(x, t)}{\partial t} = -v \frac{\partial f(x, t)}{\partial x}. \quad (8.65)$$

Since $v = dx/dt$, we see that this equation describes the motion of a traveling wave:

$$\frac{\partial f(x, t)}{\partial t} = -\frac{\partial f(x, t)}{\partial x} dx/dt. \quad (8.66)$$

The solution is of the form $f(x, t) = f(x - vt)$. This helps us appreciate the content of the Fokker-Planck equation, which combines both deterministic and stochastic evolution of particle distribution.

The drift velocity can be related to the external force $v(x) = BF(x) = -BdU/dx$, where B is the mobility and U is the potential. This gives

$$\frac{\partial f(x, t)}{\partial t} = B \frac{\partial}{\partial x} \left[\frac{dU}{dx} f(x, t) \right] + D \frac{\partial^2}{\partial x^2} f(x, t). \quad (8.67)$$

Let us now look at a simple example illustrating the workings of the Fokker-Planck equation. Here the restoring force is proportional to the displacement from the origin, $F = -\lambda x$, then the drift velocity is

$$v(x) = BF(x) = -B\lambda x. \quad (8.68)$$

The Fokker-Planck equation then reads

$$\frac{\partial f(x, t)}{\partial t} = B\lambda \frac{\partial}{\partial x} [xf(x, t)] + D \frac{\partial^2}{\partial x^2} f(x, t). \quad (8.69)$$

The first term on the right-hand side drives the particles towards the origin. If it were absent, the Fokker-Planck equation were a diffusion equation and the distribution would spread in time indefinitely. However, the restoring force curbs diffusion so that $f(x, t)$ eventually reaches a stationary state determined by

$$0 = B\lambda \frac{\partial}{\partial x} [xf(x, t)] + \underbrace{Bk_B T}_{=D} \frac{\partial^2}{\partial x^2} f(x, t). \quad (8.70)$$

(We have used the Einstein relation $D = Bk_B T$.) The solution of this equation is a Gaussian:

$$f_\infty = \left(\frac{\lambda}{2\pi k_B T} \right)^{1/2} \exp\left(-\frac{\lambda x^2}{2k_B T}\right) \quad (8.71)$$

The width of this distribution is $\sqrt{k_B T/\lambda}$:

$$\overline{x^2} = \frac{k_B T}{\lambda}. \quad (8.72)$$

This is consistent with the equipartition theorem; the mean potential energy $(1/2)\lambda\overline{x^2} = k_B T/2$. — For this case, the Fokker-Planck equation can be solved analytically, and the result is a Gaussian. The solution is intuitive: If the initial distribution is a delta function located at some $x_0 > 0$, the peak decays exponentially and the width of the distribution approaches the final value exponentially:

$$\bar{x} = x_0 \exp(-B\lambda t) \quad \text{and} \quad \overline{x^2} = x_0^2 \exp(-2B\lambda t) + \frac{k_B T}{\lambda} [1 - \exp(-2B\lambda t)]. \quad (8.73)$$

This behavior is shown in Fig. 8.3.

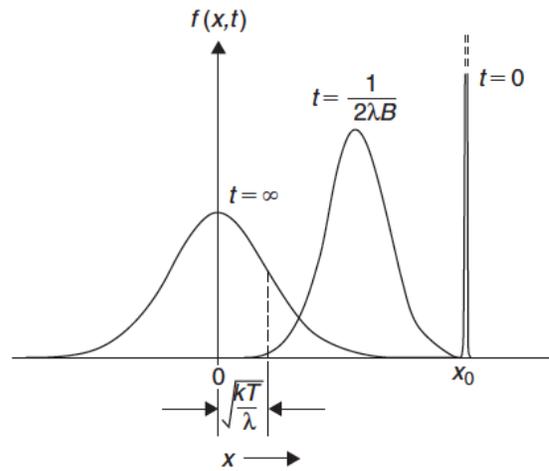


Figure 8.3: Evolution of the distribution of Brownian particles experiencing a restoring force pulling them towards the origin [1].

8.3 Fluctuation-dissipation theorem

When discussing the Langevin theory of Brownian motion, we already mentioned that the Newton's law may be viewed as the inhomogeneous first-order differential equation for velocity. Formally, the solution is given by Eq. (8.50), here rewritten in 1D for simplicity:

$$v(t) = v(0) \exp(-t/\tau) + \int_0^t \exp(-(t-u)/\tau) A(u) du. \quad (8.74)$$

Physically, the first term represents the exponential decay of the initial velocity and the second term contains the combined effect of reduced random forces ($A(t) = F(t)/M$) at times $u < t$, each weighted by an exponential $\exp(-(t-u)/\tau)$ depending on the time between the action of the force and time t .

At large t , the first term will decay to 0 and thus

$$\lim_{t \rightarrow \infty} v(t) = \int_0^\infty \exp(-(t-u)/\tau) A(u) du. \quad (8.75)$$

Note that we did extend the upper limit of the integral to ∞ but we retain t in the exponent in the integrand because $t-u$ is still finite for large u [9]. In other words, no matter how long the time t , the random force acting immediately before the observation time still carries a considerable weight.

Now we calculate the equilibrium kinetic energy of the particle. To this end, we square $\lim_{t \rightarrow \infty} v(t)$ and evaluate the thermal average:

$$\lim_{t \rightarrow \infty} \langle v^2(t) \rangle = \int_0^\infty \int_0^\infty \exp(-(2t-u-u')/\tau) \langle A(u)A(u') \rangle du du'. \quad (8.76)$$

It is reasonable to expect that the force autocorrelation function $\langle A(u)A(u') \rangle$ depends only on the difference $u'-u$ but not on u and u' itself. So we replace it by $\langle A(0)A(u'-u) \rangle$, $s' = u' - u$ clearly being the new variable of interest; we also introduce $s = t - u'$. Note that s' ranges from $-\infty$ to ∞ and s ranges from 0 (corresponding to times immediately before the observation time) to ∞ (corresponding to the very distant past). Then Eq. (8.76) becomes

$$\lim_{t \rightarrow \infty} \langle v^2(t) \rangle = \int_0^\infty ds \int_{-\infty}^\infty ds' \exp(-\underbrace{(2t-u-u')/\tau}_{=2s+s'}) \langle A(0)A(s') \rangle. \quad (8.77)$$

Note that now the two integrals are decoupled and we have

$$\begin{aligned}\lim_{t \rightarrow \infty} \langle v^2(t) \rangle &= \int_0^\infty ds \exp(-2s/\tau) \int_{-\infty}^\infty ds' \exp(-s'/\tau) \langle A(0)A(s') \rangle \\ &= \frac{\tau}{2} \int_{-\infty}^\infty ds' \exp(-s'/\tau) \langle A(0)A(s') \rangle.\end{aligned}\quad (8.78)$$

On multiplying this result by $M/2$ and recalling that $\tau = MB$, we find that

$$\lim_{t \rightarrow \infty} \frac{M}{2} \langle v^2(t) \rangle = \frac{M^2 B}{4} \int_{-\infty}^\infty ds' \exp(-s'/\tau) \langle A(0)A(s') \rangle.\quad (8.79)$$

According to the equipartition theorem, the left-hand side equals $k_B T/2$ and thus

$$\frac{2k_B T}{B} = M^2 \int_{-\infty}^\infty ds' \exp(-s'/\tau) \langle A(0)A(s') \rangle.\quad (8.80)$$

Recall that $MA(t) = F(t)$ and inverse mobility $1/B$ is equal to the friction coefficient γ so that

$$2\gamma k_B T = \int_{-\infty}^\infty ds' \exp(-s'/\tau) \langle F(0)F(s') \rangle.\quad (8.81)$$

Finally, if we assume that the autocorrelation function $\langle F(0)F(s') \rangle$ decays in a time shorter than the relaxation time, the exponential damping term in the integrand can be replaced by 1

$$2\gamma k_B T = \int_{-\infty}^\infty ds' \langle F(0)F(s') \rangle.\quad (8.82)$$

This relation between friction and fluctuations in the system is known as the fluctuation-dissipation theorem. In 3D we have

$$6\gamma k_B T = \int_{-\infty}^\infty ds' \langle \mathbf{F}(0)\mathbf{F}(s') \rangle.\quad (8.83)$$

The theorem states that the average friction force on the particle subsumed in the friction coefficient is directly related to the fluctuating part of the force on the particle. The magnitude of the autocorrelation function depends both on the magnitude of the force as well as on the time over which the force stays correlated. The larger the fluctuating force exerted on the particle by the surrounding molecules, the larger the friction coefficient.

At a more abstract level, the most impressive aspect of the fluctuation-dissipation theorem is that it establishes a connection between the equilibrium properties of the system (that is, the fluctuations and the temperature)

and the friction coefficient as a parameter pertaining to the behavior of the system when driven out of equilibrium by an external force. Thus it is not surprising that it can also be derived within the linear-response theory where one considers the generalized displacement x caused by a small generalized applied force f [10]. These two quantities are related by the generalized susceptibility by

$$x(t) = \hat{\alpha}f(t) = \int_{-\infty}^t \alpha(t-u)f(u)du. \quad (8.84)$$

$\hat{\alpha}$ is the susceptibility operator, and $\alpha(t)$ is referred to as the primary response function. — This relation is usually considered in Fourier space (physically corresponding to the response of the system to an oscillatory force) where

$$x(\omega) = \alpha(\omega)f(\omega) \quad (8.85)$$

with a complex

$$\alpha(\omega) = \int_0^{\infty} \alpha(t) \exp(i\omega t) dt = \alpha'(\omega) + i\alpha''(\omega). \quad (8.86)$$

(The lower limit of the integral is 0 because of causality; $\alpha(t < 0)$ must be 0.) The imaginary part of the susceptibility describes the out-of-phase response, that is, dissipation. We skip the derivation of the theorem and only spell out the classical version of the result:

$$\alpha''(\omega) = \frac{\omega}{2k_B T} S(\omega), \quad (8.87)$$

where

$$S(\omega) = \int_{-\infty}^{\infty} \langle x(0)x(t) \rangle \exp(i\omega t) dt \quad (8.88)$$

is the power spectrum of the equilibrium fluctuations of the displacement.

8.4 Onsager relations

In any non-equilibrium state, the physical system in question responds to non-zero generalized driving force¹ by a generalized current. For example, a

¹One needs to distinguish between the generalized driving force that generates a current in a non-equilibrium state and the generalized force that specifies the magnitude of the generalized coordinate in equilibrium. For example, a pressure difference drives flux of matter between two parts of the system of interest whereas the pressure itself determines the volume of the system as a whole.

temperature difference gives rise to heat flux, a concentration difference gives rise to the flux of matter etc. The system is usually characterized by several generalized coordinates x_i , each associated with the corresponding current \dot{x}_i ; e.g., mass m is associated with the mass flow $\Phi_m = dm/dt$. It is natural to expect that to lowest order, the vector of currents \dot{x}_i is proportional to the vector of forces X_j such that

$$\dot{x}_i = \gamma_{ij} X_j, \quad (8.89)$$

where γ_{ij} are the kinetic coefficients (here we use the Einstein summation convention). In addition to the diagonal (also referred to as direct) coefficients such as thermal conductivity and diffusion coefficient, the matrix of kinetic coefficients also contains non-zero off-diagonal coefficients pertaining to the cross transport phenomena such as the thermoelectric effect. Here we discuss the symmetry of the matrix and show that it must be symmetric so that $\gamma_{ij} = \gamma_{ji}$.

We consider the entropy of a system slightly out of equilibrium. The entropy is a suitable thermodynamic potential because we may assume that while equilibrium is being established, the system is isolated because the process is fast, and secondly because the currents resulting from any imbalance lead to a more disordered system. The entropy depends on the generalized coordinates x_i somewhat different from their respective equilibrium values \tilde{x}_i , and to second order

$$\begin{aligned} S(x_i) &= S(\tilde{x}_i) + \left(\frac{\partial S}{\partial x_i} \right)_{x_i=\tilde{x}_i} (x_i - \tilde{x}_i) \\ &\quad + \frac{1}{2} \left(\frac{\partial^2 S}{\partial x_i \partial x_j} \right)_{x_i=\tilde{x}_i, x_j=\tilde{x}_j} (x_i - \tilde{x}_i)(x_j - \tilde{x}_j). \end{aligned} \quad (8.90)$$

In equilibrium, $S(\tilde{x}_i)$ is maximal and so the first derivatives must all be 0 so that

$$\Delta S = S(x_i) - S(\tilde{x}_i) = -\frac{1}{2} \beta_{ij} (x_i - \tilde{x}_i)(x_j - \tilde{x}_j). \quad (8.91)$$

Here

$$\beta_{ij} = -\frac{1}{2} \left(\frac{\partial^2 S}{\partial x_i \partial x_j} \right)_{x_i=\tilde{x}_i, x_j=\tilde{x}_j} \quad (8.92)$$

must all be positive because by definition, any variation of any generalized coordinate leads to a decrease of entropy. Obviously the matrix of second derivatives is symmetric, $\beta_{ij} = \beta_{ji}$.

The generalized driving force can be defined as the derivative of entropy with respect to the generalized coordinate:

$$X_i = \frac{\partial S}{\partial x_i} \quad (8.93)$$

so that in equilibrium where S is maximal, the forces vanish. In turn, this means that

$$X_i = -\beta_{ij} (x_i - \tilde{x}_i). \quad (8.94)$$

Now recall that the probability of a given fluctuation is proportional to $\exp(-\Delta S/k_B)$. This means that the ensemble average of x_i equal to \tilde{x}_i is formally given by

$$\langle x_i \rangle = \frac{\int_{-\infty}^{\infty} x_i \exp(-(\beta_{ij}/k_B)(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)) dx_1 dx_2 \dots}{\int_{-\infty}^{\infty} \exp(-(\beta_{ij}/k_B)(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)) dx_1 dx_2 \dots} = \tilde{x}_i. \quad (8.95)$$

(We have tacitly extended the integration domain to all values of x_i .) On differentiating this equation with respect to \tilde{x}_j we obtain

$$\begin{aligned} & \frac{\int_{-\infty}^{\infty} x_i (-\beta_{ij}/k_B) (x_i - \tilde{x}_i) \exp(-(\beta_{ij}/k_B)(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)) dx_1 dx_2 \dots}{\int_{-\infty}^{\infty} \exp(-(\beta_{ij}/k_B)(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)) dx_1 dx_2 \dots} \\ &= -\frac{1}{k_B} \frac{\int_{-\infty}^{\infty} x_i \overbrace{\beta_{ij} (x_i - \tilde{x}_i)}^{=X_j} \exp(-(\beta_{ij}/k_B)(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)) dx_1 dx_2 \dots}{\int_{-\infty}^{\infty} \exp(-(\beta_{ij}/k_B)(x_i - \tilde{x}_i)(x_j - \tilde{x}_j)) dx_1 dx_2 \dots} \\ &= -\frac{1}{k_B} \langle x_i X_j \rangle = \delta_{ij}. \end{aligned} \quad (8.96)$$

This result is important because we now relate the force X_j to the current \dot{x}_i , which are in turn obtained by differentiating the respective generalized coordinate x_i with respect to time. Specifically, consider the correlation function which must be invariant to time reversal because the underlying microscopic phenomena are:

$$\langle x_i(0)x_j(t) \rangle = \langle x_i(0)x_j(-t) \rangle. \quad (8.97)$$

In addition, it is also invariant to time shift,

$$\langle x_i(0)x_j(-t) \rangle = \langle x_i(t)x_j(0) \rangle, \quad (8.98)$$

so that finally

$$\langle x_i(0)x_j(t) \rangle = \langle x_i(t)x_j(0) \rangle. \quad (8.99)$$

Now we subtract $\langle x_i(0)x_j(0) \rangle$ from Eq. (8.99) and divide by t to find that

$$\frac{\langle x_i(0)[x_j(t) - x_j(0)] \rangle}{t} = \frac{\langle [x_i(t) - x_i(0)]x_j(0) \rangle}{t}. \quad (8.100)$$

On taking the limit $t \rightarrow 0$ the differences become derivatives and

$$\langle x_i(0)\dot{x}_j(0) \rangle = \langle \dot{x}_i(0)x_j(0) \rangle. \quad (8.101)$$

The currents $\dot{x}_j(0) = \gamma_{jl}X_l(0)$ and $\dot{x}_i(0) = \gamma_{il}X_l(0)$ (recall that we use the Einstein summation convention) which gives

$$\langle x_i(0)\gamma_{jl}X_l(0) \rangle = \langle \gamma_{il}X_l(0)x_j(0) \rangle \quad (8.102)$$

and then, using Eq. (8.96),

$$-\gamma_{jl}k_B\delta_{il} = -\gamma_{il}k_B\delta_{jl} \quad (8.103)$$

and finally

$$\gamma_{ji} = \gamma_{ij}. \quad (8.104)$$

This is the Onsager reciprocity relation.

Given that the kinetic coefficients are symmetric and that we expect that to lowest order, the currents are linear functions of forces, we may write

$$\dot{x}_i = \frac{\partial f}{\partial X_i}, \quad (8.105)$$

where the generating function f is a quadratic form containing X_i and X_j :

$$f = \frac{1}{2}\gamma_{ij}X_iX_j. \quad (8.106)$$

This ansatz respect the symmetry of the coefficients as $\dot{x}_i = \partial f/\partial X_i = [(\gamma_{ij} + \gamma_{ji})/2]X_j = \gamma_{ij}X_j$ and $\dot{x}_j = \partial f/\partial X_j = [(\gamma_{ji} + \gamma_{ij})/2]X_i = \gamma_{ji}X_i$. The properties of the generating function are determined by the condition that the rate of entropy production \dot{S} must be positive as the system approaches equilibrium:

$$\dot{S} = \frac{\partial S}{\partial x_i} \frac{\partial x_i}{\partial t} = X_i \dot{x}_i = X_i \frac{\partial f}{\partial X_i} = 2f. \quad (8.107)$$

This condition imposes restrictions on the kinetic coefficients.

Nematodynamics

As an illustration of these ideas, let us briefly sketch how they can be employed to construct a theory of dynamics of nematic liquid crystals [11]. Here the variables of interest are collective degrees of freedom (such as the velocity field) rather than microscopic degrees of freedom (such as positions, orientations, and momenta of individual particles); this is generally referred to as *hydrodynamics*. In this materials, motion is characterized by both the flow field and by the reorientation of the nematic director. Like in isotropic fluids, the flow field is described by the symmetric part of the velocity gradient

$$A_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right), \quad (8.108)$$

whereas the reorientation of the nematic director relative to the background fluid is given by

$$\mathbf{N} = \dot{\mathbf{n}} - \boldsymbol{\omega} \times \mathbf{n}. \quad (8.109)$$

The first term is the material derivative of the director with respect to time,

$$\dot{\mathbf{n}} = \frac{d\mathbf{n}}{dt} = \frac{\partial \mathbf{n}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{n}, \quad (8.110)$$

and $\boldsymbol{\omega}$ in the second term is the antisymmetric part of the derivative of \mathbf{v} :

$$\boldsymbol{\omega} = \nabla \times \mathbf{v}. \quad (8.111)$$

The second term in Eq. (8.109) ensures that the rotation of the nematic director together with the fluid as a whole does not contribute to the reorientation; this is merely the rotation of the coordinate system.

We denote the force conjugate to the velocity gradient tensor \mathbf{A} by σ^s and the force conjugate to the rate of change of the director \mathbf{N} by \mathbf{h} . Here σ^s is the stress and \mathbf{h} is the so-called molecular field, which is equal to minus the functional derivative of the nematic elastic free energy F_d with respect to the director \mathbf{n} :

$$h_i = -\frac{\partial f_d}{\partial n_i} + \frac{\partial}{\partial x_j} \frac{\partial f_d}{\partial (\partial n_i / \partial x_j)}. \quad (8.112)$$

Here f_d is the elastic free energy density. (Note that since \mathbf{n} is a unit vector, in equilibrium \mathbf{h} must be parallel to \mathbf{n} rather than 0. In the one-constant approximation, $\mathbf{h} = K \nabla^2 \mathbf{n}$.)

Instead of writing the currents \mathbf{A} and \mathbf{N} in terms of σ^s and \mathbf{h} as implied by Eq. (8.89), we can also define forces in terms of currents, that is in the

form $X_j = \gamma_{ji}^{-1} \dot{x}_i$:

$$\sigma_{ij}^s = L_{ijkl} A_{kl} + M_{ijk} N_k, \quad (8.113)$$

$$h_i = M_{ijk} A_{jk} + P_{ij} N_j, \quad (8.114)$$

where we have already employed the Onsager reciprocity relation and used the same off-diagonal term in the two forces. The symmetry of the nematic liquid crystal is $D_{\infty h}$ which imposes certain restrictions on the coefficients. For an incompressible material,

$$\begin{aligned} \sigma_{ij}^s &= \alpha_1 n_i n_j n_k n_l A_{kl} + \alpha_4 A_{ij} + \frac{1}{2}(\alpha_5 + \alpha_6)(n_i A_{kj} + n_j A_{ki}) n_k \\ &\quad + \frac{1}{2} \gamma_2 (n_i N_j + n_j N_i), \end{aligned} \quad (8.115)$$

$$h_i = \gamma_2 n_j A_{ji} + \gamma_1 N_i. \quad (8.116)$$

Thus all seven coefficients introduced here have the dimension of viscosity. Notice that the cross-terms have the same coefficients; the factor of $1/2$ in the last term in σ_{ij}^s is there so as to ensure that the rate of change of the director enters the stress in a symmetrized form as it should in order to respect the $\mathbf{n} \leftrightarrow -\mathbf{n}$ equivalence. In particular, if $\mathbf{v} = 0$ then $\dot{\mathbf{N}} = \partial \mathbf{n} / \partial t$ and $(n_i N_j + n_j N_i) / 2 = (1/2) \partial (n_i n_j) / \partial t$.

Finally, the entropy production rate reads

$$T \dot{S} = \int (A_{ij} \sigma_{ji}^s + h_i N_i) \, d\mathbf{r}; \quad (8.117)$$

here the factor of T was also included because this is more convenient from the dimensional point of view.

8.5 Kubo formula

Let us return to the fluctuation-dissipation theorem for a moment, in part so as to illustrate its meaning and power but also to show that it can be applied to any system involving dissipation. This time consider electrons in a resistor. Just like a steady external mechanical driving force gives rise to the drift of a colloidal particle, a steady external electromotive force U would lead to a steady current according to the Ohm law $U = RI$. In case of a random fluctuating electromotive force $U(t)$, the current fluctuates too and according to the Faraday law a time-dependent current $I(t)$ is associated with an electromotive force given by LdI/dt so that

$$L \frac{dI}{dt} + RI = U(t), \quad (8.118)$$

where L is inductance. Note that this equation is completely analogous to Eq. (8.39); we can identify L with the mass of the particle M , the resistance R with inverse mobility $1/B$ or equivalently with the friction coefficient γ , and the voltage $U(t)$ with the external mechanical force $F(t)$. Then Eq. (8.82) can be rewritten as

$$R = \frac{1}{2k_B T} \int_{-\infty}^{\infty} \langle U(0)U(s) \rangle ds. \quad (8.119)$$

On dividing by R^2 , this result can be recast into

$$\frac{1}{R} = \frac{1}{2k_B T} \int_{-\infty}^{\infty} \langle I(0)I(s) \rangle ds \quad (8.120)$$

because the fluctuating voltage can be equivalently represented by the fluctuating current using Ohm's law. We recognize that the left-hand side is the conductance of the resistor and the right-hand side contains the current autocorrelation function.

This form can be generalized to 3D

$$\frac{1}{R} = \frac{1}{6k_B T} \int_{-\infty}^{\infty} \langle \mathbf{I}(0) \cdot \mathbf{I}(s) \rangle ds \quad (8.121)$$

and further to the so-called Kubo formula which relates the current density $\mathbf{j}_e(t)$ to the applied electric field $\mathbf{E}(t)$:

$$j_i(t) = \int_{-\infty}^t E_l(t') \Phi_{li}(t-t') dt'; \quad (8.122)$$

as before, summation over repeated index l is implied. Here

$$\Phi_{li}(s) = \frac{1}{k_B T} \langle j_l(0)j_i(s) \rangle \quad (8.123)$$

so that $k_B T \Phi_{li}$ represents a component of the autocorrelation tensor of the current density $\mathbf{j}(t)$. Now we consider a static external field along the x axis $\mathbf{E} = (E, 0, 0)$, calculate the component of the current along the field, and divide it by E ; the sum on l contains only 1 term. This gives the conductivity of the resistor

$$\sigma_{xx} = \frac{j_x}{E} = \int_{-\infty}^t \Phi_{xx}(t-t') dt' = \int_0^{\infty} \Phi_{xx}(s) ds = \frac{1}{k_B T} \int_0^{\infty} \langle j_x(0)j_x(s) \rangle ds. \quad (8.124)$$

Given that the autocorrelation function is invariant to time reversal [Eq. (8.97)], this result can be recast as

$$\sigma_{xx} = \frac{1}{2k_B T} \int_{-\infty}^{\infty} \langle j_x(0)j_x(s) \rangle ds \quad (8.125)$$

which is equivalent to the fluctuation-dissipation theorem [Eq. (8.120)]. This equality provides a reverse-engineering proof of the Kubo formula.

Alternatively, one may consider the response to an oscillating field $\mathbf{E} = (E \cos \omega t, 0, 0)$. Then, again resorting to the fact that $\langle j_x(0)j_x(s) \rangle$ is an even function of s , that is $\langle j_x(0)j_x(s) \rangle = \langle j_x(0)j_x(-s) \rangle$, we find that

$$\sigma_{xx}(\omega) = \frac{1}{2k_B T} \int_{-\infty}^{\infty} \langle j_x(0)j_x(s) \rangle \exp(-i\omega s) ds. \quad (8.126)$$

In turn, this means that

$$\langle j_x(0)j_x(s) \rangle = \frac{k_B T}{\pi} \int_{-\infty}^{\infty} \sigma_{xx}(\omega) \exp(i\omega s) d\omega, \quad (8.127)$$

which means that the current-current autocorrelation function is a Fourier transform of the conductivity. In particular, if $\sigma_{xx} = \sigma$ is independent of frequency, then the autocorrelation function is a delta function

$$\langle j_x(0)j_x(s) \rangle = 2k_B T \sigma \delta(s) \quad (8.128)$$

characteristic of the white noise.

The Kubo formula Eq. (8.122) can also be used to show that the Onsager reciprocity relations follow from the fluctuation-dissipation theorem. To this end, let us use a more abstract notation where \dot{x}_i is a generalized current and X_j is the generalized force:

$$\dot{x}_i(t) = \frac{1}{k_B T} \int_{-\infty}^t X_j(t') \langle \dot{x}_j(t') \dot{x}_i(t) \rangle dt', \quad (8.129)$$

which has the form $\dot{x}_i = \gamma_{ij} X_j$ except that it includes a time dependence not addressed in the Onsager formulation. Now we introduce $s = t - t'$, which goes from 0 (present) to ∞ (distant past), and thus

$$\dot{x}_i(t) = \frac{1}{k_B T} \int_0^{\infty} X_j(t-s) \langle \dot{x}_j(t-s) \dot{x}_i(t) \rangle ds. \quad (8.130)$$

To show that the Onsager relations hold, we would like to compare this formula to its analog where i and j are swapped:

$$\dot{x}_j(t) = \frac{1}{k_B T} \int_0^{\infty} X_i(t-s) \langle \dot{x}_i(t-s) \dot{x}_j(t) \rangle ds. \quad (8.131)$$

Since the correlation function is invariant to time shift and to time reversal as a consequence of the reversibility of microscopic processes,

$$\begin{aligned}\langle \dot{x}_j(t-s)\dot{x}_i(t) \rangle &= \langle \dot{x}_j(0-s)\dot{x}_i(0) \rangle = \langle \dot{x}_j(s)\dot{x}_i(0) \rangle = \langle \dot{x}_i(0)\dot{x}_j(s) \rangle \\ &= \langle \dot{x}_i(t-s)\dot{x}_j(t) \rangle.\end{aligned}\quad (8.132)$$

Thus the correlation functions in the integrands in Eqs. (8.130) and (8.131) are the same. Moreover, if they are delta functions like in Eq. (8.128) then the current at a given time t only depends on the force at the same time and then the Kubo formula completely reduces to the form $\dot{x}_i = \gamma_{ij}X_j$. In this case, the content of the Onsager reciprocity relations $\gamma_{ij} = \gamma_{ji}$ stems directly from the equality $\langle \dot{x}_j(t-s)\dot{x}_i(t) \rangle = \langle \dot{x}_i(t-s)\dot{x}_j(t) \rangle$.

8.6 Time-dependent Ginzburg-Landau theory

In the above discussion of nematodynamics, we have already used the concepts of fluxes and forces in the hydrodynamic context. The same can be done with the theory developed for the description of Brownian motion, which can be applied to study the dynamics of a suitable order parameter m . We already hinted at this when mentioning the Landau-Khalatnikov theory [Eq. (8.57)]. Let us now extend these ideas by incorporating stochastic forces acting on the collective degrees of freedom; this is known as the time-dependent Ginzburg-Landau model [12].

We begin by the zero-field Landau free energy²

$$F[m(\mathbf{r}, t)] = \int d\mathbf{r} \left\{ \frac{1}{2} a \tilde{t} m^2(\mathbf{r}, t) + \frac{1}{4} a_4 m^4(\mathbf{r}, t) + \frac{1}{2} b [\nabla m(\mathbf{r}, t)]^2 \right\}. \quad (8.133)$$

We know from Eq. (7.21) that the correlation length diverges with a critical exponent of 1/2, and it is natural to expect that then the fluctuations must become ever slower on approaching the critical point. We know that the thermodynamic force can be defined by $\delta F / \delta m(\mathbf{r}) = a \tilde{t} m(\mathbf{r}) + a_4 m^3(\mathbf{r}) - b \nabla^2 m(\mathbf{r})$, which vanished in equilibrium, and then the simplest stochastic equation of motion for the order parameter is of the form

$$\frac{\partial m(\mathbf{r}, t)}{\partial t} = -\Gamma \frac{\delta F[m(\mathbf{r}, t)]}{\delta m(\mathbf{r}, t)} + \mathcal{F}(\mathbf{r}, t). \quad (8.134)$$

Here $\mathcal{F}(\mathbf{r}, t)$ is the stochastic force arising from all other degrees of freedom, analogous to the force exerted by water molecules on the Brownian particle. As before $\langle \mathcal{F}(\mathbf{r}, t) \rangle = 0$. We furthermore assume that the spatial and

²In this section, \tilde{t} rather than t stands for the reduced temperature $T/T_c - 1$; t is time.

temporal correlations of the stochastic force are very short-ranged so that

$$\langle \mathcal{F}(\mathbf{r}, t) \mathcal{F}(\mathbf{r}', t') \rangle = 2\Gamma k_B T \delta(\mathbf{r} - \mathbf{r}') \delta(t - t') \quad (8.135)$$

which is in agreement with Eq. (8.82).

Now we revert to the Fourier space:

$$m(\mathbf{k}) = \int m(\mathbf{r}) \exp(-i\mathbf{k} \cdot \mathbf{r}) d\mathbf{r} \quad (8.136)$$

where the equation of motion reads

$$\frac{\partial m(\mathbf{k}, t)}{\partial t} = -\Gamma \frac{\partial F[m(\mathbf{k}, t)]}{\partial m(-\mathbf{k}, t)} + \mathcal{F}(\mathbf{k}, t). \quad (8.137)$$

Here we have replaced the functional derivative by a partial derivative because in Fourier space, the elastic term reads $bk^2 m(\mathbf{k})m(-\mathbf{k})/2$. In addition,

$$\langle \mathcal{F}(\mathbf{k}, t) \mathcal{F}(\mathbf{k}', t') \rangle = 2\Gamma k_B T \delta(\mathbf{k} - \mathbf{k}') \delta(t - t'). \quad (8.138)$$

Now we can deal with Eq. (8.137), restricting the discussion to the high-temperature regime where the thermodynamic force reads

$$\frac{\partial F[m(\mathbf{k}, t)]}{\partial m(-\mathbf{k}, t)} = (a\tilde{t} + bk^2)m(\mathbf{k}, t). \quad (8.139)$$

We first introduce the wavevector-dependent relaxation rate

$$\gamma(\mathbf{k}) = \Gamma(a\tilde{t} + bk^2) \quad (8.140)$$

and then formally solve Eq. (8.137) just like we did in Sec. 8.3 when discussing Brownian motion:

$$m(\mathbf{k}, t) = m(\mathbf{k}, 0) \exp(-\gamma(\mathbf{k})t) + \int_0^t \exp(-\gamma(\mathbf{k})(t-u)) \mathcal{F}(\mathbf{k}, u) du. \quad (8.141)$$

Hence

$$\begin{aligned} \langle m(\mathbf{k}, t) m(\mathbf{k}', t') \rangle &= \langle m(\mathbf{k}, 0) m(\mathbf{k}', 0) \rangle \exp(-[\gamma(\mathbf{k})t + \gamma(\mathbf{k}')t']) \\ &+ \int_0^t du \int_0^{t'} du' \exp(-\gamma(\mathbf{k})(t-u) - \gamma(\mathbf{k}')(t'-u')) \\ &\quad \times \langle \mathcal{F}(\mathbf{k}, u) \mathcal{F}(\mathbf{k}', u') \rangle \end{aligned} \quad (8.142)$$

The cross terms containing $\langle \mathcal{F}(\mathbf{k}, u) \rangle$ and $\langle \mathcal{F}(\mathbf{k}', u') \rangle$ vanish. In writing the exponential time dependence, we anticipated that the correlation function contains $\delta(\mathbf{k} - \mathbf{k}')$ and thus $\gamma(\mathbf{k})$ and $\gamma(\mathbf{k}')$ were identified ahead of time. — If t and t' are both large compared to $1/\gamma(\mathbf{k})$ and $1/\gamma(\mathbf{k}')$, respectively, the first term is also negligible and as before, the correlation function $\langle \mathcal{F}(\mathbf{k}, u)\mathcal{F}(\mathbf{k}', u') \rangle$ only depends on $u' - u$ so that it can be written as $\langle \mathcal{F}(\mathbf{k}, 0)\mathcal{F}(\mathbf{k}', u' - u) \rangle$. Then

$$\begin{aligned} \langle m(\mathbf{k}, t)m(\mathbf{k}', t') \rangle &= \int_0^t du \int_0^{t'} du' \exp(-\gamma(\mathbf{k})(t - u + t' - u')) \\ &\quad \times \langle \mathcal{F}(\mathbf{k}, 0)\mathcal{F}(\mathbf{k}', u' - u) \rangle. \end{aligned} \quad (8.143)$$

We introduce $s' = u' - u$ and $s = t + t' - 2u$; s' ranges between $-\infty$ and ∞ . Without loss of generality we can assume that $t' > t$: Then s ranges between $t' - t$ and ∞ (corresponding to $u \ll t, t'$). Also note that $du' = ds'$ but $du = -ds/2$. Thus

$$\begin{aligned} \langle m(\mathbf{k}, t)m(\mathbf{k}', t') \rangle &= \frac{1}{2} \int_{t'-t}^{\infty} ds \exp(-\gamma(\mathbf{k})s) \\ &\quad \times \int_{-\infty}^{\infty} ds' \exp(\gamma(\mathbf{k})s') \langle \mathcal{F}(\mathbf{k}, 0)\mathcal{F}(\mathbf{k}', s') \rangle. \end{aligned} \quad (8.144)$$

We insert Eq. (8.138) and integrate over s' to find

$$\langle m(\mathbf{k}, t)m(\mathbf{k}', t') \rangle = \Gamma k_B T \delta(\mathbf{k} - \mathbf{k}') \int_{t'-t}^{\infty} ds \exp(-\gamma(\mathbf{k})s). \quad (8.145)$$

Finally,

$$\begin{aligned} \langle m(\mathbf{k}, t)m(\mathbf{k}', t') \rangle &= \frac{\Gamma k_B T}{\gamma(\mathbf{k})} \delta(\mathbf{k} - \mathbf{k}') \exp(-\gamma(\mathbf{k})(t' - t)) \\ &= \frac{k_B T}{a\tilde{t} + bk^2} \delta(\mathbf{k} - \mathbf{k}') \exp(-\gamma(\mathbf{k})(t' - t)). \end{aligned} \quad (8.146)$$

By spelling out $\gamma(\mathbf{k})$ in the prefactor, we emphasize that the magnitude of the long-wavelength fluctuations with $k = 0$ diverges at the critical point $t = 0$; in particular,

$$\gamma(\mathbf{k} = 0) \sim \tilde{t} \sim \xi^{-2}. \quad (8.147)$$

At the same time, their relaxation time $\tau = 1/\gamma(\mathbf{k} = 0) = 1/\gamma a\tilde{t}$ also diverges, which is referred to as the critical slowing down. Note that these results were obtained within the mean-field theory.

Chapter 9

Stochastic thermodynamics

The domain of classical thermodynamics is more or less restricted to systems in equilibrium, exchanging heat, work, and matter with the environment. The same applies to the equilibrium statistical mechanics based on the probability for the system to be in a given microstate that is proportional to the Boltzmann factor. Also incorporated within this framework is the description of small fluctuations around equilibrium and the response of the system to small external forces as discussed in the previous Chapter; these states are by definition not in equilibrium but they are still related to the behavior of system in equilibrium. For example, the fluctuation-dissipation theorem can be used to calculate the transport properties of the system based on the equilibrium correlation functions, which is quite remarkable.

Many of the reassuring statements concerning fluctuations around equilibrium rely on the thermodynamic limit, which is certainly not applicable when studying small systems. The second class of systems outside the equilibrium thermodynamics consist of strongly driven systems. Since the beginning of the 1990s, theoretical developments including the steady-state fluctuation theorem, the Jarzynski relation, the Crooks fluctuation theorem, and many other results began to emerge, eventually forming a comprehensive apparatus which allows one to introduce work, heat, and entropy at the level of trajectories in non-equilibrium systems [13]. Known as *stochastic thermodynamics*, this apparatus has thus far been applied to several experimental systems such as colloidal particles in laser traps, enzymes, molecular motors etc.

These systems can be displaced from equilibrium in three different ways: (i) they can be prepared in a non-equilibrium state and left to relax towards equilibrium; (ii) they can be driven by time-dependent forces; and (iii) they

can be driven by steady forces giving rise to non-equilibrium steady states. Given that the systems of interest are typically in contact with a heat reservoir (usually the surrounding water bath), they are still characterized by a well-defined temperature.

9.1 Dissipated work

To appreciate the difference between the classical and the stochastic thermodynamics, consider a macroscopic ensemble such as a gas exchanging work and heat with a heat reservoir, which maintains a constant temperature T so that all changes in the reservoir are quasistatic [14]. The first law for the system states that

$$\Delta U = W + Q. \quad (9.1)$$

The second law of thermodynamics states that the total entropy change of both the reservoir and the system must be nonnegative:

$$\Delta S_{tot} = \Delta S_{res} + \Delta S \geq 0. \quad (9.2)$$

Here the subscript "res" refers to the reservoir and no subscript refers to the system. Given that heat can flow to the system solely from the reservoir, $Q_{res} = -Q$ and thus $\Delta S_{res} = -Q/T$ because all changes in the reservoir are quasistatic. On the other hand, the change of entropy of the system is

$$\Delta S \geq \frac{Q}{T}. \quad (9.3)$$

It is reasonable to divide ΔS into a reversible part $\Delta_e S = Q/T$, which is associated with a trajectory from the initial to the final state via equilibrium intermediate states, and an irreversible part $\Delta_i S$. Then

$$\Delta S_{tot} = \underbrace{-\frac{Q}{T}}_{=\Delta S_{res}} + \underbrace{\frac{Q}{T} + \Delta_i S}_{=\Delta S} \geq 0 \quad (9.4)$$

so that

$$\Delta S_{tot} = \Delta_i S \geq 0. \quad (9.5)$$

As expected, the irreversible part of the entropy change in the system represents the entropy production.

Now consider the free energy difference between the initial and the final states. Given that in both states, the temperature is the same as the system is in contact with the heat bath,

$$\Delta F = \Delta U - \Delta(TS) = \Delta U - T\Delta S = W + Q - T\Delta S = W - T\Delta_i S \quad (9.6)$$

and since $\Delta_i S \geq 0$ we have

$$T\Delta_i S = W - \Delta F \geq 0. \quad (9.7)$$

This result relates work expended in a process, which is not a function of state, to the change of the free energy, which is a function of state; $T\Delta_i S$ may also be viewed as dissipated work. We stress that it only provides a lower bound for work rather than an exact value: $W \geq \Delta F$ so that if the free energy is decreased in a process ($\Delta F < 0$), the work W may be negative (so that work is expended by the system) but $|W|$ must be smaller than $|\Delta F|$. For example, work that can be expended by an expanding gas with an initial volume V_i and a final volume V_f at a constant temperature is

$$-W \leq -\Delta F_{\text{eq}} = Nk_B T \ln \frac{V_f}{V_i}; \quad (9.8)$$

here we took into account that the internal energy of an ideal gas depends only on temperature. In a quasistatic (that is, slow) expansion, $-W = -\Delta F_{\text{eq}}$ but otherwise the work expended by the gas is smaller than $-\Delta F_{\text{eq}}$.

An often-cited microscopic version of such an experiment is stretching of an RNA strand attached to two colloidal beads held by micropipets [15] (Fig. 9.1a). If stretched slowly (Fig. 9.1b, blue curve), the force extension curve is essentially reversible but when the beads move faster (red curve) the unfolding and refolding are not reversible as witnessed by the hysteresis. Figure 9.1c shows several superposed measurements at low speed (blue), medium speed (green) and fast speed (red). The superposed curves in the same run show how important are the fluctuations, and the ever larger hysteresis on increasing the speed of beads shows that the process is generally irreversible.

Figure 9.2 shows histograms of dissipated work at several extensions and several pulling speeds (blue, green, and red corresponding to slow, medium, and fast speeds, respectively). Note that the dissipated work can be negative and that the mean increases with the pulling speed, that is with the degree of irreversibility.

Had this experiment been done with many strands, the dissipated work per strand would tend to the average value $\langle W \rangle$ and Eq. (9.7) would read $\langle W \rangle - \Delta F_{\text{eq}} > 0$. Thus all that the second law of thermodynamics says about the distribution of the dissipated work concerns the lower bound of its mean, which does not look like much.

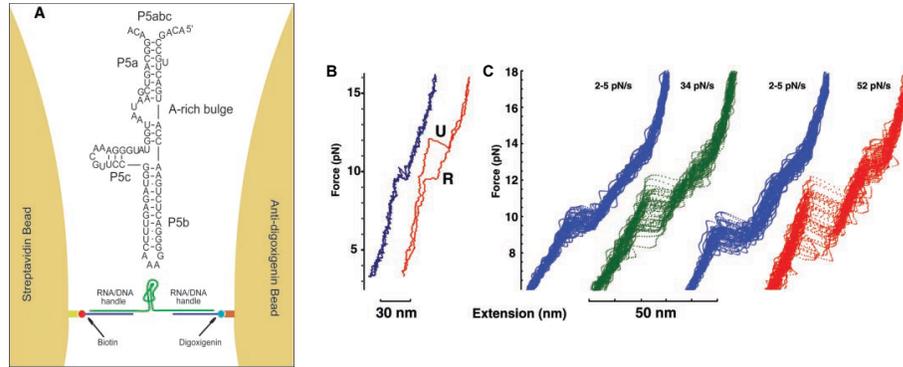


Figure 9.1: Schematic of the RNA stretching experiment (a) and a typical force-extension curve (b) at slow and fast bead speed (blue and red, respectively). Also shown are force-extension curves for dozens of runs, indicating that fluctuations are very prominent (c) [15].

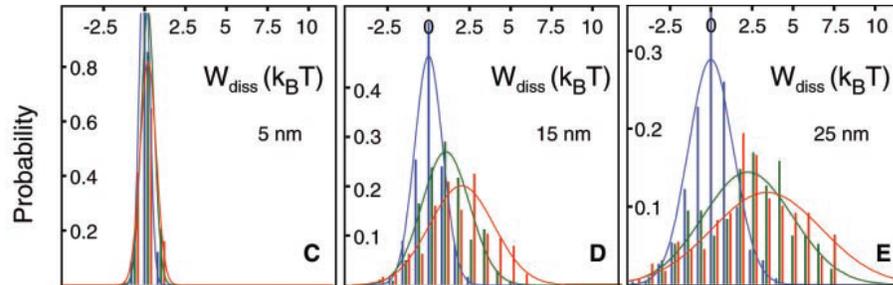


Figure 9.2: Distribution of dissipated work for several RNA extensions (left to right, 5, 15, and 25 nm) at several bead speeds, blue, green, and red corresponding to slow, medium, and fast speeds [15].

9.2 Jarzynski relation

A far stronger statement relating work and free energy change during an irreversible process is provided by the Jarzynski relation (1997; also known as the Jarzynski equality). This relation involves the work expended in driving the system at a temperature T pertaining to the initial equilibrium state to a final equilibrium state, and it states that

$$\langle \exp(-W/k_B T) \rangle = \exp(-\Delta F/k_B T). \quad (9.9)$$

The average is over all trajectories connecting the initial and the final state. Jarzynski relation implies the second law in the form of Eq. (9.7). To show

this, we use Jensen's inequality

$$f\left(\sum_{i=1}^n p_i x_i\right) \leq \sum_{i=1}^n p_i f(x_i), \quad (9.10)$$

where $p_i \geq 0$ are the probabilities that $x = x_i$ and f is a convex function. In other words, $f(\text{mean}) \leq \text{mean}(f)$. If f is an exponential function, then we find that $\exp(-\langle W \rangle / k_B T) \leq \langle \exp(-W / k_B T) \rangle$ so that

$$\exp(-\langle W \rangle / k_B T) \leq \exp(-\Delta F / k_B T) \quad (9.11)$$

or

$$\langle W \rangle \geq \Delta F, \quad (9.12)$$

i.e., $\langle W \rangle - \Delta F \geq 0$.

Jarzynski relation has been verified experimentally, its main strength being in allowing one to determine the free energy difference between two states based on arbitrarily irreversible trajectories connecting them.

9.3 Crooks (fluctuation) theorem

The Jarzynski theorem itself follows from a more general theorem discovered by Crooks in 1998 [14]. The Crooks (fluctuation) theorem gives the ratio of probabilities of the forward and the reverse trajectories between two equilibrium states, $p(W)$ and $p(-W)$, respectively:

$$\frac{p(W)}{p(-W)} = \exp((W - \Delta F_{eq}) / k_B T). \quad (9.13)$$

In the special case of an infinitesimally slow, quasistatic transformation $W = \Delta F_{eq}$ so that $p(W) = p(-W)$.

The above excerpts briefly illustrate the scope of stochastic thermodynamics, which has by now evolved into a comprehensive and coherent theoretical framework. Many of the results were verified experimentally in colloidal systems, electric circuits, etc. where they provided a means of interpreting the measurements or simulations. More details on the topic can be found in the reviews by Seifert[13] and Van den Broeck [14].

Bibliography

- [1] R. K. Pathria and P. D. Beale, *Statistical Mechanics* 3rd Ed. (Elsevier, Amsterdam, 2011).
- [2] C. N. Likos, *Colloidal Interactions: From Effective Potentials to Structure*, in C. Bechinger, F. Sciortino, and P. Zihlerl (eds.), *Physics of Complex Colloids* (Società Italiana di Fisica, Bologna, 2013).
- [3] C. Van Vliet, *Equilibrium and Non-Equilibrium Statistical Mechanics* (World Scientific, Singapore, 2008).
- [4] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (World Scientific, Singapore, 2005).
- [5] M. J. Bowick and A. Travesset, *Phys. Rep.* **344**, 255 (2001).
- [6] H. H. von Grünberg, P. Keim, and G. Maret, *Phase Transitions in Two-Dimensional Colloidal Systems*, in G. Gompper and M. Schick (Eds.), *Soft Matter Volume 3* (Wiley VCH, Weinheim, 2007).
- [7] W. Qi, A. P. Gantapara, and M. Dijkstra, *Soft Matter* **10**, 5449 (2014).
- [8] Z. Rácz, *From Langevin to Fokker-Planck Equation*, <http://cgl.elte.hu/~racz/Stoch-diff-eq.pdf>.
- [9] P. Eastman, *Statistical Mechanics*, <https://web.stanford.edu/~peastman/statmech/>.
- [10] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Butterworth Heinemann, Amsterdam, 1980).
- [11] P.-G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Oxford Science Publications, Oxford, 1993).
- [12] F. Schwabl, *Statistical Mechanics* (Springer, Berlin, 2002).

- [13] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
- [14] C. Van den Broeck, *Stochastic Thermodynamics: A Brief Overview*, in C. Bechinger, F. Sciortino, and P. Zihlerl (eds.), *Physics of Complex Colloids* (Società Italiana di Fisica, Bologna, 2013).
- [15] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Science **296**, 1832 (2002).