

Quantum Statistical Mechanics

Matúš Medo
Physics Department

December 4, 2012

Contents

1	Fundamentals	1
1.1	The state operator	1
1.2	Entropy as a lack of information	4
1.3	The fundamental postulate	6
2	Statistical ensembles	8
2.1	Ideal gas of distinguishable particles	8
2.2	Two systems in thermal contact	10
2.3	Canonical ensemble	11
2.4	Einstein's model for the specific heat of a solid	13
2.5	Grand canonical ensemble	15
3	Quantum statistics of the ideal gas	18
3.1	Indistinguishable particles	18
3.2	Occupation numbers	19
3.3	Non-interacting fermions and bosons	20
3.4	Relations between ensembles	23
3.5	Black body radiation	23
3.6	Electrons in metal	25
3.7	The Bose-Einstein condensation	28
3.8	Degenerate atomic gases	30
3.9	Debye's model for the specific heat of a solid	34
	Appendices	37
A	Bose-Einstein condensation in complex networks	37
B	Physics of white dwarfs	43
B.1	Relativistic Fermi gas	43
B.2	White Dwarfs	46
C	Path integrals	50
C.1	Brownian motion	50
C.2	Path integral formulation of quantum mechanics	53
C.3	Path integrals in statistical mechanics	57

Literature

1. W. Greiner, L. Neise, H. Stöcker, *Thermodynamics and statistical mechanics* (also available in French and German)
2. R. K. Pathria, *Statistical Mechanics*

These lecture notes on Quantum Statistical Mechanics are closely based on previous lecture notes developed by Dionys Baeriswyl and Yi-Cheng Zhang.

1 Fundamentals

The theory of heat, one of the cornerstones of physics, can be approached from two sides, from the world of macroscopic phenomena using some general principles or from the microscopic world of molecules, atoms, elementary particles. A macroscopic approach is used in thermodynamics, while the microscopic approach constitutes the framework of statistical mechanics. Depending on whether the microscopic dynamics is treated according to the rules of classical mechanics or those of quantum mechanics, one differentiates between classical and quantum statistical mechanics. This course deals with quantum statistical mechanics. A basic acquaintance with thermodynamics and quantum mechanics is presupposed, but not with classical statistical mechanics.

The main aim of statistical mechanics is to describe macroscopic properties of a system on the basis of the behavior of its constituents (atoms, molecules, ...). The number of particles in a macroscopic volume is usually huge. For instance, 1 mm^3 of air at ambient pressure and temperature contains about 3×10^{16} molecules. It is pointless to even try to follow the evolution of such a big number of particles or to try to describe the state at a certain time, because it is not only impossible to achieve that computationally, but quite generally macroscopic measurements are coarse-grained and thus depend only on the average behavior of the constituents.

Statistical mechanics uses probabilistic methods for linking the microscopic and macroscopic worlds. This approach is justified by the law of large numbers, which states that probabilistic results are essentially exact if the number of variables involved is very large.

Quantum statistical mechanics is based on the quantum-mechanical description of many-particle systems. Therefore, even before introducing statistical distributions (for microscopic states), we have to face the intrinsic probabilistic nature of quantum-mechanical predictions. It is very important to distinguish carefully between the intrinsic statistical nature of quantum mechanics and the statistical aspects introduced by the probability distribution of quantum-mechanical states.

1.1 The state operator

We have already mentioned the impossibility of describing precisely the state vector of the constituents of a macroscopic body, which suggests that we have to resort to probabilistic methods. There is an additional reason for choosing a probabilistic approach, namely the impossibility of preparing a truly stationary state of a large system. The fundamental reason for this fact is the extremely small difference between energy levels (exponentially small in the number of particles). In view of the uncertainty relation $\Delta E \Delta t \geq \hbar$, it would require an astronomically large time Δt to prepare the system in a state where the energy uncertainty ΔE is smaller than the level spacing. But even if the system would be in a stationary state at some time, it would quickly make a transition to a different state because it is impossible to separate completely a system from the environment. We thus have to account for two sources of randomness and uncertainty in quantum systems: intrinsic randomness which is due to quantum mechanics (recall the probabilistic nature of

measurement in QM) and statistical randomness which is due to our lack of information about the system. Mastering the latter is the main goal of statistical physics.

A very useful device for describing the statistical nature due to both the intrinsic quantum-mechanical uncertainties and the incomplete knowledge of the state is the state operator. To find its form, we start with a standard quantum-mechanical case: a system which can be in various states $\Psi_n(\mathbf{x})$ where \mathbf{x} is a position vector.¹ In a particular state Ψ_n , the expected value of an operator \hat{A} reads

$$\langle \hat{A} \rangle_{\Psi_n} = \int \Psi_n^*(\mathbf{x}) \hat{A} \Psi_n(\mathbf{x}) d\mathbf{x} \equiv \langle \Psi_n | \hat{A} | \Psi_n \rangle. \quad (1.1)$$

How to generalize this to the case of our interest where the exact state of the system is unknown? Denote the probability that the system is in fact in state Ψ_n as p_n —we call this a “mixed” state and it is, as we shall soon see, fundamentally different from the usual quantum superposition of states. It is now natural to modify the above expression to the form

$$\langle \hat{A} \rangle_{\text{mixed}} = \sum_n p_n \int \Psi_n^*(\mathbf{x}) \hat{A} \Psi_n(\mathbf{x}) d\mathbf{x} \equiv \sum_n p_n \langle \Psi_n | \hat{A} | \Psi_n \rangle. \quad (1.2)$$

That is, if the system is with probability 25% in state 1 where $\langle \hat{A} \rangle_1 = 1$ and with probability 75% in state 2 where $\langle \hat{A} \rangle_2 = 5$, we expect $\langle \hat{A} \rangle_{\text{mixed}} = 0.25 \times 1 + 0.75 \times 5 = 4$.

If we choose any complete and orthonormal basis $\{\Phi_a\}$, Equation (1.2) can be further developed using the closure relation $\sum_a |\Phi_a\rangle \langle \Phi_a| = \hat{1}$. We find

$$\begin{aligned} \langle \hat{A} \rangle_{\text{mixed}} &= \sum_n p_n \langle \Psi_n | \left(\sum_a |\Phi_a\rangle \langle \Phi_a| \right) \hat{A} | \Psi_n \rangle = \sum_n p_n \sum_a \langle \Psi_n | \Phi_a \rangle \langle \Phi_a | \hat{A} | \Psi_n \rangle = \\ &= \sum_a \langle \Phi_a | \hat{A} \left(\sum_n p_n |\Psi_n\rangle \langle \Psi_n| \right) | \Phi_a \rangle = \sum_a \langle \Phi_a | \hat{A} \hat{\varrho} | \Phi_a \rangle = \text{Tr}(\hat{A} \hat{\varrho}) \end{aligned}$$

where we denoted

$$\hat{\varrho} = \sum_n p_n |\Psi_n\rangle \langle \Psi_n| \quad (1.3)$$

and from where follows the following important formula

$$\langle \hat{A} \rangle_{\text{mixed}} = \text{Tr}(\hat{A} \hat{\varrho}) = \text{Tr}(\hat{\varrho} \hat{A}). \quad (1.4)$$

Here $\hat{\varrho}$ is a so-called state operator (often referred to as the density matrix) of a given mixed state. The possibility to switch the order of operators under the trace comes from the fact that $\text{Tr}(\hat{A}\hat{B}) = \text{Tr}(\hat{B}\hat{A})$ which we prove in the following paragraph. Since $\hat{\varrho}$ doesn't depend on the chosen basis $\{\Phi_a\}$, neither on the operator \hat{A} , we can conclude that it provides a sufficient description of the system at hand in the sense that it can be used to predict the expected value of any quantum-mechanical measurement. It is further useful

¹Equivalently, we could choose the so-called momentum representation and work with $\tilde{\Psi}(\mathbf{p})$ in the momentum space.

to note that $\text{Tr}\hat{\rho} = 1$. To show that, it suffices to set $\hat{A} = \hat{1}$, leading to $\langle \hat{1} \rangle_{\text{mixed}} = \sum_n p_n \langle \Psi_n | \hat{1} | \Psi_n \rangle = \sum_n p_n \langle \Psi_n | \Psi_n \rangle = \sum_n p_n = 1$ (what else, right?) and, at the same time, $\langle \hat{1} \rangle_{\text{mixed}} = \text{Tr}(\hat{1}\hat{\rho}) = \text{Tr}\hat{\rho}$.

Traces of operators which we use now are closely related to traces of matrices from ordinary linear algebra where $\text{Tr}A = \sum_i A_{ii}$. In the case of an operator, we simply sum-up the diagonal elements in a chosen basis: $\text{Tr}\hat{A} = \sum_i \langle \Phi_i | \hat{A} | \Phi_i \rangle$. This interpretation relies on an important fact that trace of a matrix is independent of the basis. To prove that, one first shows that $\text{Tr}(AB) = \text{Tr}(BA)$ and then realizes that since change of a basis is due to a unitary matrix U which transforms A to $A' = U^\dagger A U$, we have $\text{Tr}A' = \text{Tr}(U^\dagger A U) = \text{Tr}(U U^\dagger A) = \text{Tr}A$.²

Time evolution of the state operator can be derived easily by making time derivative of its basic formula which yields

$$\frac{\partial \hat{\rho}}{\partial t} = \sum_n p_n \left(\frac{\partial |\Psi_n\rangle}{\partial t} \langle \Psi_n| + |\Psi_n\rangle \frac{\partial \langle \Psi_n|}{\partial t} \right)$$

where we benefit from the fact that probabilities p_n are simply given at a certain moment and thus do not evolve with time. Time evolution of $|\Psi_n\rangle$ is given by the Schroedinger equation in the form $\partial |\Psi_n\rangle / \partial t = \frac{1}{i\hbar} \hat{H} |\Psi_n\rangle$. Time evolution of $\langle \Psi_n|$ follows by complex conjugation of this in the form $\partial \langle \Psi_n| / \partial t = (\partial |\Psi_n\rangle / \partial t)^* = -\frac{1}{i\hbar} \hat{H}^\dagger (|\Psi_n\rangle)^* = -\frac{1}{i\hbar} \langle \Psi_n| \hat{H}$. Putting everything together, we obtain

$$\frac{\partial \hat{\rho}}{\partial t} = \sum_n p_n \frac{1}{i\hbar} (\hat{H} |\Psi_n\rangle \langle \Psi_n| - |\Psi_n\rangle \langle \Psi_n| \hat{H}) = \frac{1}{i\hbar} (\hat{H} \hat{\rho} - \hat{\rho} \hat{H}) = \frac{1}{i\hbar} [\hat{H}, \hat{\rho}]. \quad (1.5)$$

We finally discuss the difference between mixed states described by state operators $\hat{\rho}$ and usual quantum states Ψ which are often referred to as “pure” states in this context. First of all, we can write state operator for an arbitrary mixture of states—even for a mixture where there is in fact only one state Ψ . State operator of this pure state reads $\hat{\rho} = |\Psi\rangle \langle \Psi|$. Its square is then $\hat{\rho}^2 = |\Psi\rangle \langle \Psi| |\Psi\rangle \langle \Psi| = |\Psi\rangle \langle \Psi| = \hat{\rho}$. The opposite is also true: if $\hat{\rho}^2 = \hat{\rho}$, then the corresponding state is a pure state. To see this, it is enough to take state operator of an arbitrary mixed state, $\hat{\rho} = \frac{1}{2}(|\Psi_1\rangle \langle \Psi_1| + |\Psi_2\rangle \langle \Psi_2|)$ and verify that indeed $\hat{\rho}^2 \neq \hat{\rho}$. Instead of working with states, one can also work with their vector representation. We know that the two basic states of a two-level system can be represented by vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. If the system is with probability p in the up state and with probability $1 - p$ in the down state, the matrix form of the state operator follows in the form

$$\hat{\rho} = p \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}^\dagger + (1 - p) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}^\dagger = \dots = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}$$

which has expected properties (trace one, $\hat{\rho}^2 \neq \hat{\rho}$, $\langle \sigma_z \rangle = \text{Tr}(\sigma_z \hat{\rho}) = \text{Tr} \begin{pmatrix} p & 0 \\ 0 & p-1 \end{pmatrix} = 2p - 1$, etc.).

²To prove $\text{Tr}(AB) = \text{Tr}(BA)$, we write $\text{Tr}(AB) = \sum_i (AB)_{ii} = \sum_i (\sum_j A_{ij} B_{ji}) = \sum_j (\sum_i B_{ji} A_{ij}) = \sum_j (BA)_{jj} = \text{Tr}(BA)$.

A simple example of a pure state is an unpolarized photon beam whose photons can have two helicities, corresponding to two orthogonal quantum states, $|R\rangle$ (right circular polarization) and $|L\rangle$ (left circular polarization). One can wonder if a superposition of these two states, for example $|X\rangle = (|R\rangle + |L\rangle)/\sqrt{2}$, does not represent unpolarized light. That's of course false— $|X\rangle$ corresponds to vertically polarized light which passes through a vertical polarizer without absorption and is absorbed entirely by a horizontal polarizer. By contrast, unpolarized light can be described perfectly by assuming that each photon is either $|R\rangle$ with 50% probability or $|L\rangle$ with 50% probability—a situation which can be easily described by state operator. Regardless of the type and orientation of a polarizer, 50% of such unpolarized light passes through it. For more details see L.E. Ballentine, *Quantum Mechanics*, Prentice Hall 1990 (chapter 2) or R. P. Feynman, *Statistical Mechanics*, W. A. Benjamin, 1972 (chapter 2).

1.2 Entropy as a lack of information

The central concept of the theory of heat is that of entropy. In thermodynamics, this concept is usually introduced via the Carnot cycle, in the kinetic theory of gases via Boltzmann's H -theorem, in statistical mechanics it is related to the number of possible states compatible with given macroscopic constraints. Entropy is used also in theory of information where it measures the lack of information. To illustrate this, let us consider an ensemble of events e_n ($n = 1, \dots, \Gamma$) with event probabilities p_n satisfying the non-negativity and normalization conditions

$$p_n \geq 0, \quad \sum_{n=1}^{\Gamma} p_n = 1. \quad (1.6)$$

In the context of information theory, events e_n represent various letters³ in an incoming message (for example, e_1 represents 'A', e_2 represents 'B', and so on). Imagine that we need to encode texts in a binary digital form where each letter is translated into a series of zeros and ones (for example, 'A' would be '001', and so on). The question now is that if we know all event probabilities in advance, how should we construct the letter codes so that the resulting average length of our message is minimal. It is obviously advantageous to assign short codes to frequent letters and long codes to rare ones. But how exactly?

Denoting the length of code assigned to event e_n as l_n , the expected length required to encode one event is

$$\langle l \rangle = \sum_{n=1}^{\Gamma} p_n l_n.$$

Without any constraints, this quantity is minimized by setting $l_n = 0$ for all events (negative lengths are obviously impossible). However, such empty messages would have no sense because it would be impossible to decode them back to a sequence of letters. Therefore,

³One could also let events to represent higher units in a language, such as words in a given language. The derivation presented here does not depend on what exactly the events are.

there obviously must exist some constraint on l_n values which guarantees that the message can be decoded. This constraint is called the Kraft inequality⁴ and states that for a code to uniquely decipherable, it must hold that

$$\sum_{n=1}^{\Gamma} 2^{-l_n} \leq 1.$$

If the equality holds, the code has no redundancy (letter codes are not longer than they need to be). It is now straightforward to carry out constrained minimization of $\langle l \rangle$. The Lagrange function has the form

$$\mathcal{L}(\{l_n\}, \lambda) = \sum_{n=1}^{\Gamma} p_n l_n - \lambda \left(\sum_{n=1}^{\Gamma} 2^{-l_n} - 1 \right)$$

and leads to $l_n = -\log_2 p_n$.⁵ The expected message length thus has the form $\langle l \rangle = -\sum_{n=1}^{\Gamma} p_n \log_2 p_n$ which, up to a multiplicative term, is identical with the entropy as we know it from statistical physics

$$S(p_1, \dots, p_{\Gamma}) = -k_B \sum_{n=1}^{\Gamma} p_n \ln p_n, \quad (1.7)$$

where $k_B = 1.38 \times 10^{-23} \text{ JK}^{-1}$ is the Boltzmann constant. This is a so-called Gibbs formula for entropy. If one of the events is certain ($p_n = 1$), one can easily see that entropy equals zero because $1 \times \ln 1 = 0$ and $\lim_{x \rightarrow 0} x \ln x = 0$.

One can show that Eq. (1.7) is the only one that fulfills the following conditions which one would expect from a measure of a lack of information:

- (i) $S(0, \dots, 0, 1, 0, \dots, 0) = 0$ (no missing information) and $S > 0$ in all other situations.
- (ii) S is a continuous and symmetric function of its parameters (*i.e.*, the order in which one specifies the events does not matter).
- (iii) An impossible event does not influence S , $S(p_1, \dots, p_{\Gamma}, 0) = S(p_1, \dots, p_{\Gamma})$.
- (iv) For equally probable events, S increases with the number of events.
- (v) S is additive. For independent events e_n ($n = 1, \dots, \Gamma$) and e'_m ($m = 1, \dots, \Gamma'$) where the probability of the composed event $e_n \cap e'_m$ is $p_n p'_m$ this additivity implies that $S(\{p_n \times p'_m\}) = S(p_1, \dots, p_{\Gamma}) + S(p'_1, \dots, p'_{\Gamma'})$.

⁴See http://en.wikipedia.org/wiki/Kraft_inequality for more information about the inequality and its proof.

⁵Constructing an optimal code is an entirely different matter (for example, code lengths prescribed by the formula above might be non-integer which is impossible to achieve in reality) which we do not discuss here. A starting point for further reading is provided at http://en.wikipedia.org/wiki/Huffman_coding.

For a proof that these conditions indeed result in a measure that we found above, see for example B. Diu et al., *Physique Statistique, Complément I.G.* Conversely, it is simple to verify that Eq. (1.7) fulfills all five conditions.

We can now ask, which situation represented by event probabilities $\{p_n\}_{n=1}^{\Gamma}$ yields maximal “disinformation” (maximal lack of information). This corresponds to maximizing $S(p_1, \dots, p_{\Gamma})$ with the constraint $\sum_{n=1}^{\Gamma} p_n = 1$. The Lagrange function is now of the form $\mathcal{L}(\{p_n\}, \lambda) = S(\{p_n\}) - \lambda(\sum_n p_n - 1)$. By deriving it with respect to p_m , one obtains

$$\ln p_m = -\frac{k_B + \lambda}{k_B} \quad (1.8)$$

which is simply the same probability of all events. Due to the normalization condition, $p_m = 1/\Gamma$ and therefore finally

$$\max_{\{p_n\}} \left(-k_B \sum_n p_n \ln p_n \right) = -k_B \sum_n \frac{1}{\Gamma} \ln \frac{1}{\Gamma} = k_B \ln \Gamma. \quad (1.9)$$

This is the entropy of the most disordered state without constraints. If some constraints are added (such as the expected energy of the system, for example), different solutions arise. We shall return to those cases later.

1.3 The fundamental postulate

In an unpolarized light beam all polarization directions are equally probable. In the spirit of information theory, this amounts to a maximum of ignorance, *i.e.*, to maximum entropy. In statistical mechanics, the events correspond to different stationary states of a system. For a thermally isolated system where the total energy is constant (up to small fluctuations mentioned earlier), the principle of maximum entropy tells us that all states having the same total energy are equally probable. Stated otherwise: The occupation probability is uniform for states in the subspace defined by the external constraints (such as values of total energy, volume, number of particles and so on). This is the fundamental postulate of statistical mechanics. With our definition of entropy this principle can be expressed as $S = k_B \ln \Gamma$ where Γ is the number of states compatible with the constraints and the proportionality has been chosen to be Boltzmann’s constant. For a given number of particles N and a given volume V , the number of states is a function of all three variables— E , V , and N . Correspondingly, the entropy is also a function of these three macroscopic variables and we thus get

$$S(E, V, N) = k_B \ln \Gamma(E, V, N). \quad (1.10)$$

The state defined by macroscopic observables (such as pressure, number of particles, etc.) is called *macrostate*, while the state defined by all its microscopical properties (the exact quantum numbers of all its constituents) is called *microstate*. A single macrostate thus can be composed from a large number of microstates—that is the meaning of $\Gamma(E, V, N)$ here.

In order to obtain an explicit expression for the state operator $\hat{\varrho}$ in this situation, we consider an essentially closed system with total energy in a narrow interval between E and $E + \delta E$ ($\delta E \ll E$). The number of the available energy eigenstates $|n\rangle$ is therefore

$$\Gamma(E, V, N) = \sum_{\substack{n \\ E < E_n < E + \delta E}} 1 \quad (1.11)$$

and the probability of finding the system in a specific eigenstate $|n\rangle$ is

$$p_n = \begin{cases} 1/\Gamma & \text{when } E < E_n < E + \delta E, \\ 0 & \text{otherwise.} \end{cases} \quad (1.12)$$

This corresponds to the state operator

$$\hat{\varrho}_{\text{micro}} = \sum_n p_n |n\rangle \langle n| = \frac{1}{\Gamma(E, V, N)} \sum_{\substack{n \\ E < E_n < E + \delta E}} |n\rangle \langle n|. \quad (1.13)$$

In close analogy to an unpolarized photon beam, this state operator can be interpreted as a collection of systems, each of which has a different energy E_n in the narrow interval $E < E_n < E + \delta E$. Such a collection of systems is referred to as ensemble. When we specifically speak about a collection of systems with a given energy E (up to the usual uncertainty), this is referred to as *microcanonical ensemble*.

2 Statistical ensembles

In the preceding chapter we have established the microcanonical ensemble of closed systems with essentially the same energy E (up to an uncertainty δE) represented by the state operator $\hat{\rho}_{\text{micro}}$ given by Eq. (1.13). The expected value of an arbitrary observable \hat{A} then reads $\langle A \rangle = \text{Tr}(\hat{\rho}_{\text{micro}} \hat{A})$ and the entropy is related to Γ by Boltzmann's equation (1.10). In this chapter, we use microcanonical ensemble to treat ideal gas, neglecting its symmetry constraints. These constraints will be elaborated later on in Section 3 where we will see that at low temperatures they lead to dramatic changes in behavior. Microcanonical ensemble is little suitable for study of other, less trivial systems. After using it to solve the ideal gas example, we thus proceed to alternative ensembles—the canonical ensemble representing a system in thermal equilibrium with a heat bath at a certain temperature, and the grand canonical ensemble where in addition to energy, particles too can be exchanged with the environment.

2.1 Ideal gas of distinguishable particles

Ideal gas of N particles with mass m is described by the Hamiltonian

$$\hat{H} = \sum_{j=1}^N \frac{\hat{\mathbf{p}}_j^2}{2m}. \quad (2.1)$$

For distinguishable particles, the eigenstates are simply products of single-particle states

$$|\mathbf{k}_1\rangle \otimes |\mathbf{k}_2\rangle \otimes \cdots \otimes |\mathbf{k}_N\rangle. \quad (2.2)$$

For identical particles, such as electrons or helium atoms, one has to use states with appropriate symmetry which cannot be written as simple product states (see Section 3). As we remember from the basic quantum mechanics course, the energy eigenvalues of this system are

$$E(\mathbf{k}_1, \dots, \mathbf{k}_N) = \sum_{j=1}^N \frac{\hbar^2 \mathbf{k}_j^2}{2m}. \quad (2.3)$$

State of the system is thus described by a $3N$ -dimensional vector $(k_{1x}, k_{1y}, k_{1z}, k_{2x}, \dots, k_{Nz})$. We imagine the system to be enclosed in a cubic box of side length L and volume $V = L^3$ and use periodic boundary conditions for the wave functions, *i.e.*,

$$e^{ik_{j\alpha}L} = 1, \quad \alpha = x, y, z, \quad j = 1, \dots, N. \quad (2.4)$$

Possible wave vectors \mathbf{k}_j therefore form a discrete cubic lattice,

$$\mathbf{k}_j = \frac{2\pi}{L}(\nu_{jx}, \nu_{jy}, \nu_{jz}) = \frac{2\pi}{L} \boldsymbol{\nu}, \quad \nu_{j\alpha} \in \mathcal{Z}. \quad (2.5)$$

Since quantum numbers corresponding to ordinary systems are exceedingly large (when $L = 1$ m and $T = 300$ K, the typical quantum number is of the order of 10^{10}), we can replace

sums over \mathbf{k} by integrals. For instance, the number of different states $(k_{1x}, k_{1y}, \dots, k_{3Nz})$ with total energy below E is the volume of the $3N$ -dimensional sphere

$$\sum_{j=1}^N \frac{\hbar^2 \mathbf{k}_j^2}{2m} = \sum_{j=1}^N \frac{\hbar^2 \nu_j^2}{2mL^2} < E \quad (2.6)$$

divided by the volume per point $(2\pi/L)^{3N}$. We can thus write this number of states as

$$\Omega(E_{\mathbf{k}_1, \dots, \mathbf{k}_N} < E) = \left(\frac{L}{2\pi}\right)^{3N} \int dk_{1x} \dots \int dk_{3Nz} \Theta \left(E - \sum_{j=1}^N \frac{\hbar^2 \mathbf{k}_j^2}{2m} \right), \quad (2.7)$$

where $\Theta(x)$ is the Heaviside step function ($\Theta(x) = 1$ for $x \geq 0$ and zero otherwise). As will be proven in exercises, the volume of a d -dimensional sphere with radius r is given by $\pi^{d/2} r^d / (d/2)!$. In our case, radius of the sphere is given by the maximal possible wave vector element k_{\max} which is achieved when all other wave vectors are zero. From Eq. (2.6) it follows that $\hbar^2 k_{\max}^2 / (2m) = E$ and thus we find

$$\Omega(E_{\mathbf{k}_1, \dots, \mathbf{k}_N} < E) = \frac{1}{(3N/2)!} \left(\frac{mEL^2}{2\pi\hbar^2} \right)^{3N/2}. \quad (2.8)$$

For any large number n we may use Stirling's formula,

$$n! \sim (n/e)^n \text{ for } n \rightarrow \infty, \quad (2.9)$$

and obtain

$$\Omega(E_{\mathbf{k}_1, \dots, \mathbf{k}_N} < E) \approx e^{3N/2} V^N \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3N/2}.$$

This formula allows us to calculate the number of available states between E and $E + \delta E$ (δE has to be much smaller than E but large enough to accommodate many states between E and $E + \delta E$),

$$\Gamma(E, V, N) = \Omega(E + \delta E) - \Omega(E) \approx \frac{\partial \Omega}{\partial E} \delta E \approx e^{3N/2} V^N \frac{3N}{2} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3N/2} \frac{\delta E}{E}. \quad (2.10)$$

By neglecting terms of order 1, $\ln N$ and $\ln(\delta E/E)$, and identifying E with the thermodynamic energy U , we obtain entropy of the system in the form

$$S(U, V, N) = k_B \ln \Gamma(U, V, N) = Nk_B \left[\frac{3}{2} + \ln V + \frac{3}{2} \ln \left(\frac{mU}{3\pi\hbar^2 N} \right) \right]. \quad (2.11)$$

This relation could be readily transformed into the fundamental thermodynamic relation $U = U(S, V, N)$ but instead, we invert the differential $dU = -p dV + T dS + \mu dN$ to obtain

$$dS = \frac{1}{T} dU + \frac{p}{T} dV - \frac{\mu}{T} dN. \quad (2.12)$$

E_1 N_1, V_1	E_2 N_2, V_2
---------------------	---------------------

Figure 2.1: Thermal contact between two systems.

Using Eq. (2.11), we now get

$$\begin{aligned}\frac{1}{T} &:= \left(\frac{\partial S}{\partial U} \right)_{V,N} = \frac{3}{2} N k_B \frac{1}{U}, \\ \frac{p}{T} &:= \left(\frac{\partial S}{\partial V} \right)_{U,N} = N k_B \frac{1}{V}, \\ -\frac{\mu}{T} &:= \left(\frac{\partial S}{\partial N} \right)_{U,V} \approx \frac{S}{N}.\end{aligned}$$

which are the familiar relations that hold for the ideal gas.

2.2 Two systems in thermal contact

We now consider two systems separated by a diathermal wall which allows energy to be exchanged while the volumes and the particle numbers as well as the total energy E remain fixed. We assume that the interaction between the two subsystems is negligible for the calculation of energy eigenvalues. Therefore the number of microstates for given energies E_1 and $E - E_1$ (up to a small uncertainty) can be taken as the product

$$\Gamma_1(E_1, V_1, N_1) \Gamma_2(E - E_1, V_2, N_2), \quad (2.13)$$

where $\Gamma_i(E_i, V_i, N_i)$ is the number of available microstates in the two subsystems ($i = 1, 2$) for the given macroscopic constraints (see Figure 2.1. According to Eq. (1.10), the total entropy is now given by the sum

$$S_1(E_1, V_1, N_1) + S_2(E - E_1, V_2, N_2). \quad (2.14)$$

This needs to be maximized, so its differential with respect to E_1 must be zero. We thus have

$$\frac{\partial S_1}{\partial E_1} + \frac{\partial S_2}{\partial E_1} = 0 \implies \frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2} = 0 \implies T_1 = T_2$$

where we used the fact that since $E_2 = E - E_1$, $\partial S_2 / \partial E_1 = -\partial S_2 / \partial E_2$ and we used then the definition of temperature $1/T_i := \partial S_i / \partial E_i$ in the last step. Thermal equilibrium of two systems in thermal contact thus follows directly from the maximum principle for the entropy.

2.3 Canonical ensemble

The microcanonical ensemble, representing a single closed system, is not very practical for actual calculations. An alternative approach is based on a composite system consisting of two subsystems: a small one, for which we want to find a statistical distribution function (or a state operator), in thermal equilibrium with a large one. It turns out that the resulting *canonical ensemble* is much easier to handle than the microcanonical ensemble. The key step is to associate the probability for a particular state of the small system (with energy E_n) with the number of available microstates of the large system (at energy $E - E_n$).

We consider a small subsystem 1 in thermal equilibrium with system 2 which is much bigger than system 1. If $\Gamma(E)$ is the number of microstates of the composed system 1 + 2 with energy E , then the probability of finding system 1 in a particular state n with energy E_n is

$$P_n = \frac{\Gamma_2(E - E_n)}{\Gamma(E)} \quad (2.15)$$

where $\Gamma_2(E - E_n)$ is the number of states of system 2 with energy $E - E_n$. From Eq. (1.10) follows that $\Gamma = \exp(S/k_B)$ and we thus obtain

$$P_n = \frac{e^{S_2(E - E_n)/k_B}}{e^{S(E)/k_B}} = e^{S_2(E - E_n)/k_B - S(E)/k_B}. \quad (2.16)$$

If U_1 is the average energy of system 1, Taylor expansion of S_2 gives

$$\begin{aligned} S_2(E_2) &= S_2(E - E_n) = S_2(E - U_1 + (U_1 - E_n)) \approx \\ &\approx S_2(E - U_1) + \frac{\partial S_2}{\partial E_2}(U_1 - E_n) = S_2(E - U_1) + \frac{U_1 - E_n}{T} \end{aligned}$$

where T is temperature. Additivity of entropy, $S(E) = S_1(U_1) + S_2(E - U_1)$, now implies

$$P_n = \exp\left(\frac{U_1 - TS_1(U_1) - E_n}{k_B T}\right) = e^{\beta F_1} e^{-\beta E_n} \quad (2.17)$$

where $F_1 = U_1 - TS_1(U_1)$ is the Helmholtz free energy of system 1 and $\beta := 1/(k_B T)$. Since P_n is normalized by definition ($\sum_n P_n = 1$), we find

$$1 = e^{\beta F_1} \sum_n e^{-\beta E_n}. \quad (2.18)$$

This motivates us to introduce a so-called partition function

$$Z(T, V, N) := \sum_n e^{-\beta E_n} \quad (2.19)$$

which then allows us to rewrite the previous equation to the form $e^{\beta F_1} = 1/Z$, yielding

$$F(T, V, N) = -\ln Z/\beta = -k_B T \ln Z(T, V, N). \quad (2.20)$$

which is an analog of the formula $S(E, V, N) = k_B \ln \Gamma(E, V, N)$ from the microcanonical ensemble.

Once Z is known, all other properties follow easily. In addition to the free energy above, the probability of a microstate n is

$$P_n = \frac{e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} = \frac{1}{Z} e^{-\beta E_n}. \quad (2.21)$$

From now on, we omit the index 1 for the investigated system. The internal energy can be written as

$$U := \sum_n E_n \frac{e^{-\beta E_n}}{Z} = \frac{1}{Z} \sum_n E_n e^{-\beta E_n} = \frac{-1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta} \quad (2.22)$$

which is an essential relation between Z and one of the key quantities of interest, U . We can now conclude the derivations by computing the entropy using the Gibbs formula

$$S = -k_B \sum_n P_n \ln P_n = k_B \sum_n P_n (\beta E_n + \ln Z) = \frac{U - F}{T} \quad (2.23)$$

which leads to a formula that we already met above

$$F = U - TS. \quad (2.24)$$

From here it follows that $\partial F / \partial T = -S$. We have seen before that the microcanonical ensemble corresponds to maximal entropy (or maximal lack of information). A similar relation can be found for canonical ensemble which maximizes the entropy under constraints of normalization ($\sum_n P_n = 1$) and fixed average energy ($\sum_n E_n P_n = U$). Furthermore, canonical ensemble follows from minimization of free energy $F = U - TS$ with the only constraint being that of P_n normalization. We can thus say that in the canonical approach, a system attempts to minimize U and maximize S which are two contradictory requirements and their respective weights are set by temperature. When T is very small, minimization of F is equivalent to minimization of U , whereas when T is large, it is mainly the entropy term what plays a role and minimization of F is equivalent to maximization of S .

As a simple illustration of the canonical ensemble at work, one can study a two-level system with energy levels $\pm E$ and only one particle where the partition function turns out in the form $Z_1 = e^{-\beta E} + e^{\beta E} = 2 \cosh(\beta E)$ where we use index 1 to stress that it relates to a one-particle system. Internal energy reads $U_1 = -E \tanh(\beta E)$ which behaves in agreement with expectations: $U_1 \rightarrow -E$ when $T \rightarrow 0$ (only the ground state is populated then) and $U_1 \rightarrow 0$ when $T \rightarrow \infty$ (when $k_B T \gg E$, occupation probabilities of both levels are almost the same).

Assuming that there are N identical, distinguishable, and non-interacting particles, the N -particle partition function is based on summing over all N -particle microstates $\{n_i\}_1^N$ where energy $E(\{n_i\}_1^N) = E_1 + \dots + E_N$ and therefore

$$Z_N(T, V, N) = \sum_{\{n_i\}_1^N} e^{-\beta E_1} \dots e^{-\beta E_N} = \underbrace{\left(\sum_{n_1} e^{-\beta E_1} \right)}_{Z_1(T, V, 1)} \times \sum_{\{n_i\}_2^N} e^{-\beta E_2} \dots e^{-\beta E_N}. \quad (2.25)$$

where we separated the sum over quantum numbers describing particle 1, n_1 , by keeping quantum numbers of all other particles fixed under the summation. By repeating this step for all N particles, we obtain

$$Z_N(T, V, N) = \prod_{i=1}^N Z_1(T, V, 1) = Z_1(T, V, 1)^N$$

which is a general relationship between 1-particle and N -particle partition function for non-interacting and distinguishable particles. Internal energy finally follows from $U := -\partial \ln Z_N / \partial \beta$ in the form $U_N = NU_1$ which is again a very expected result.

We can again write down the corresponding state operator which has the form

$$\hat{\varrho}_{\text{can}} = \frac{1}{Z} \sum_n e^{-\beta E_n} |n\rangle \langle n| = \frac{1}{Z} \sum_n e^{-\beta \hat{H}} |n\rangle \langle n| = \frac{1}{Z} e^{-\beta \hat{H}} \quad (2.26)$$

where

$$Z = \sum_n e^{-\beta E_n} = \text{Tr}(e^{-\beta \hat{H}}). \quad (2.27)$$

The relations

$$\text{Tr} \hat{\varrho}_{\text{can}} = 1, \quad \ln \hat{\varrho}_{\text{can}} = -\ln Z - \beta \hat{H}$$

imply

$$\text{Tr}(\hat{\varrho}_{\text{can}} \ln \hat{\varrho}_{\text{can}}) = -\ln Z - \beta \text{Tr}(\hat{\varrho}_{\text{can}} \hat{H}) = \beta(F - U).$$

We can thus express the entropy in a compact way in terms of the state operator as

$$S = -k_B \text{Tr}(\hat{\varrho} \ln \hat{\varrho}). \quad (2.28)$$

This formula applies also to the entropy in the microcanonical ensemble (see Eqs. (1.10) and (1.13)).

2.4 Einstein's model for the specific heat of a solid

In 1900, Planck derived his famous formula for the black-body radiation by assuming that the energy exchange between the electromagnetic radiation and a body occurs only in discrete lumps. Five years later, Einstein radically modified Planck's picture by postulating that, independent of the interaction with matter, light itself consists of energy quanta (nowadays called photons). Shortly after (in an article published in 1907), Einstein applied his approach to the vibrations of a solid, by assuming each atom to execute harmonic oscillations with a fixed frequency ω_0 and discrete energies $\varepsilon_n = \hbar\omega_0(n + \frac{1}{2})$ where $n = 0, 1, 2, \dots$. In this way, he could resolve a long-standing discrepancy between the specific heat predicted by the classical theory of the ideal gas and the specific heat values observed experimentally, especially for experiments carried out on solids containing elements with a light mass.

Einstein's analysis can be put onto a firm basis by modeling the atoms of a solid as independent harmonic oscillators with the Hamiltonian

$$\hat{H} = \sum_{j=1}^N \left(\frac{1}{2M} \hat{\mathbf{p}}_j^2 + \frac{1}{2} M \omega_0^2 \hat{\mathbf{r}}_j^2 \right). \quad (2.29)$$

Here we have neglected the possible interactions of atoms and therefore we do not have to worry about the symmetry of the resulting many-atom states (see Chapter 3).

We will use the canonical ensemble to find thermodynamical properties of this system. The Hamiltonian eigenvalues are

$$E_{n_1, \dots, n_{3N}} = \sum_{i=1}^{3N} \hbar \omega_0 \left(n_i + \frac{1}{2} \right) = \hbar \omega_0 (n_1 + n_2 + \dots + n_{3N} + \frac{3}{2} N) \quad (2.30)$$

and the partition function reads

$$\begin{aligned} Z &= e^{-\frac{3}{2} N \beta \hbar \omega_0} \sum_{\{n_i\}} e^{-\beta \hbar \omega_0 (n_1 + n_2 + \dots + n_{3N})} = e^{-\frac{3}{2} N \beta \hbar \omega_0} \sum_{\{n_i\}} \left(\prod_{i=1}^{3N} e^{-\beta \hbar \omega_0 n_i} \right) = \\ &= e^{-\frac{3}{2} N \beta \hbar \omega_0} \prod_{i=1}^{3N} \left(\sum_{n_i=0}^{\infty} e^{-\beta \hbar \omega_0 n_i} \right) = e^{-\frac{3}{2} N \beta \hbar \omega_0} \prod_{i=1}^{3N} \frac{1}{1 - e^{-\beta \hbar \omega_0}} = \left(\frac{e^{-\frac{1}{2} \beta \hbar \omega_0}}{1 - e^{-\beta \hbar \omega_0}} \right)^{3N}. \end{aligned} \quad (2.31)$$

Here we exchanged the order of summation and product (note that when doing so, the nature of the sum changes from a sum over all possible microstates $\{n_i\}$ to a sum over all possible values of one quantum number n_i) similarly as we did in Eq. (2.25). This can be justified simply by noticing that while $\sum_{\{n_i\}} \prod_i$ and $\prod_i \sum_{n_i}$ look different, there is a perfect one-to-one correspondence between individual terms in these two expressions.

Eq. (2.22) can be used to find the internal energy in the form (note how a product of terms in Z transforms to a sum of contributions in U)

$$U = \frac{3}{2} N \hbar \omega_0 + \sum_{i=1}^{3N} \frac{\hbar \omega_0}{e^{\beta \hbar \omega_0} - 1} = 3N \hbar \omega_0 \left(\frac{1}{2} + \frac{1}{e^{\beta \hbar \omega_0} - 1} \right). \quad (2.32)$$

The specific heat then reads

$$C := \frac{\partial U}{\partial T} = 3N k_B \left(\frac{\hbar \omega_0}{k_B T} \right)^2 \frac{e^{\beta \hbar \omega_0}}{(e^{\beta \hbar \omega_0} - 1)^2} \quad (2.33)$$

with the asymptotic limits

$$C \sim \begin{cases} 3N k_B & \beta \hbar \omega_0 \ll 1 \ (T \rightarrow \infty), \\ 3N k_B \left(\frac{\hbar \omega_0}{k_B T} \right)^2 e^{-\frac{\hbar \omega_0}{k_B T}} & \beta \hbar \omega_0 \gg 1 \ (T \rightarrow 0). \end{cases} \quad (2.34)$$

In the limit $T \rightarrow \infty$ we recover the Dulong-Petit law known from classical analysis of the system whereas in the limit $T \rightarrow 0$ we find specific heat decaying vanishing exponentially which qualitatively agrees with experiments. A more realistic model for specific heat of crystals was introduced by Debye in 1913 where interactions of atoms are included and the specific heat vanishes with temperature polynomially (for small T , it is proportional to T^3) which better agrees with experiments. We shall study that in Sec. 3.9.

The Helmholtz free energy now follows in the form

$$F = -k_B T \ln Z = 3N \left(\frac{1}{2} \hbar \omega_0 + k_B T \ln (1 - e^{-\beta \hbar \omega_0}) \right). \quad (2.35)$$

and the entropy is

$$S = -\frac{\partial F}{\partial T} = 3N k_B \left(-\ln(1 - e^{-\beta \hbar \omega_0}) + \frac{\beta \hbar \omega_0}{e^{\beta \hbar \omega_0} - 1} \right). \quad (2.36)$$

When $T \rightarrow 0$, entropy goes to zero

$$S \sim 3N k_B \left(1 + \frac{\hbar \omega_0}{k_B T} \right) e^{-\frac{\hbar \omega_0}{k_B T}} \quad (2.37)$$

in agreement with the fourth law of thermodynamics. In general, the zero-temperature entropy is proportional to $\ln \Gamma_0$ where Γ_0 is the degeneracy of the system's ground state. In Einstein's model, the ground state is unique and thus $\Gamma_0 = 1$ and $S = 0$. In general, one often faces some degeneracy of the ground state with $\Gamma_0 \sim N$. In that case, entropy is proportional to $\ln N$ which is negligible with respect to entropy values at non-zero temperature which are proportional to N .

2.5 Grand canonical ensemble

In the canonical ensemble, the number of particles is fixed and only the energy of the system is allowed to fluctuate. If one considers a small system in contact with a bigger one (reservoir), it is natural to consider also the situation where the number of particles is not fixed but fluctuates around a given average value. Here we carry out this analysis and study a small system 1 in thermal (related to exchange of energy) and chemical (related to exchange of particles) equilibrium with a reservoir. Similarly as in Sec. 2.3, the probability of finding system 1 in a particular state n with N particles and energy $E_n^{(N)}$ is

$$P_n^{(N)} = \frac{\Gamma_2(E - E_n^{(N)}, N_{\text{tot}} - N)}{\Gamma(E, N_{\text{tot}})} \quad (2.38)$$

where $\Gamma_2(E - E_n^{(N)}, N_{\text{tot}} - N)$ is the number of states of reservoir with energy $E - E_n^{(N)}$ and number of particles $N_{\text{tot}} - N$ and $\Gamma(E, N_{\text{tot}})$ is the number of states of the whole system with energy E and N_{tot} particles (both these quantities are fixed). Eq. (1.10) allows us to write

$$P_n^{(N)} = \exp \left(\frac{S_2(E - E_n^{(N)}, N_{\text{tot}} - N) - S(E, N_{\text{tot}})}{k_B} \right). \quad (2.39)$$

The average energy of system 1 and its average number of particles are labeled as U_1 and N_1 , respectively. Entropy S_2 can now be expanded in a Taylor series as

$$\begin{aligned} S_2(E - E_n^{(N)}, N_{\text{tot}} - N) &\approx S_2(E - U_1, N_{\text{tot}} - N_1) + \frac{\partial S_2}{\partial E_2} (U_1 - E_n^{(N)}) + \frac{\partial S_2}{\partial N_2} (N_1 - N) = \\ &= S_2(E - U_1, N_{\text{tot}} - N_1) + \frac{1}{T} (U_1 - E_n^{(N)}) - \frac{\mu}{T} (N_1 - N). \end{aligned}$$

We can now introduce the grand canonical potential

$$\Omega(T, V, \mu_1, \dots, \mu_n) := U - S \frac{\partial U}{\partial S} - \frac{\partial U}{\partial N} = U - TS - \mu N \quad (2.40)$$

which will allow us to express Eq. (2.39) in a more compact way. Denoting $\Omega_1(T, V_1, \mu) = U_1 - TS_1 - \mu N_1$ and using the additivity of entropy

$$S(E, N_{\text{tot}}) = S_1(U_1, N_1) + S_2(E - U_1, N_{\text{tot}} - N_1) \quad (2.41)$$

we find

$$P_n^{(N)} = e^{\beta \Omega_1(T, V_1, \mu)} e^{-\beta(E_n^{(N)} - \mu N)}. \quad (2.42)$$

From now on we drop index 1 labeling the system of our interest. System 2 (reservoir) appears only through parameters T and μ . Normalization of P_n implies

$$\sum_{n, N} P_n^{(N)} = e^{\beta \Omega(T, V, \mu)} \mathcal{Z}(T, V, \mu) = 1 \quad (2.43)$$

where we sum over all possible states of N particles and then over all particle numbers N . Here $\mathcal{Z}(T, V, \mu)$ is the grand canonical partition function

$$\mathcal{Z}(T, V, \mu) = \sum_{n, N} e^{-\beta(E_n^{(N)} - \mu N)} \quad (2.44)$$

which is a direct generalization of the canonical partition function $Z(T, V, N)$. The normalization condition above yields the fundamental relation

$$\Omega(T, V, \mu) = -k_B T \ln \mathcal{Z}(T, V, \mu) \quad (2.45)$$

which is a generalization of Eq. (2.20). The average number of particles now reads

$$\langle N \rangle = \frac{1}{\mathcal{Z}} \sum_{n, N} N e^{-\beta(E_n^{(N)} - \mu N)} = \frac{1}{\beta \mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \mu} = k_B T \frac{\partial \ln \mathcal{Z}}{\partial \mu} = -\frac{\partial \Omega}{\partial \mu}. \quad (2.46)$$

It is simple to show that despite a conceptual difference in their basic assumptions, canonical and grand canonical ensemble are equivalent. It basically suffices to show that fluctuations of the number of particles are negligible for macroscopic systems. To achieve that, we evaluate

$$k_B T \frac{\partial \langle N \rangle}{\partial \mu} = (k_B T)^2 \frac{\partial}{\partial \mu} \left(\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \mu} \right) = (k_B T)^2 \left(-\frac{1}{\mathcal{Z}^2} \left(\frac{\partial \mathcal{Z}}{\partial \mu} \right)^2 + \frac{1}{\mathcal{Z}} \frac{\partial^2 \mathcal{Z}}{\partial \mu^2} \right) = \langle N^2 \rangle - \langle N \rangle^2.$$

Where we have $\sigma_N^2 = \langle N^2 \rangle - \langle N \rangle^2$ on the right side and a quantity proportional to N on the left side. We thus see that the relative fluctuations of the number of particles, $\sigma_N / \langle N \rangle$ are proportional to $N^{-1/2}$ which vanishes for macroscopic systems where $N \rightarrow \infty$ (so-called thermodynamic limit). Knowing that the two ensembles produce identical results, we can always choose the one that is easier to handle in the situation at hand. For example, systems of identical particles are easier to treat in grand canonical ensemble than in the canonical ensemble. The microcanonical ensemble is usually the most difficult one to use but it still did its job in providing the basis for developing the other two, more practical ensembles.

To find a corresponding state operator, we introduce an orthonormal basis $|n, N\rangle$ for which it holds that

$$\hat{N}|n, N\rangle = N|n, N\rangle, \quad \hat{H}|n, N\rangle = E_n^{(N)}|n, N\rangle \quad (2.47)$$

where \hat{N} is the number of particles operator. We therefore get

$$\hat{\mathcal{Q}}_{\text{grand}} = \frac{1}{\mathcal{Z}} \sum_{n, N} e^{-\beta(E_n^{(N)} - \mu N)} |n, N\rangle \langle n, N| = \frac{1}{\mathcal{Z}} e^{-\beta(\hat{H} - \mu \hat{N})} \quad (2.48)$$

where

$$\mathcal{Z} = \text{Tr}(e^{-\beta(\hat{H} - \mu \hat{N})}). \quad (2.49)$$

Entropy $S = -\partial\Omega/\partial T$ can be again expressed by Eq. (2.28).

3 Quantum statistics of the ideal gas

In classical mechanics particles can be distinguished, at least in principle, by attaching some label to each of them and following their individual trajectories. Due to Heisenberg's uncertainty principle this is no longer possible for quantum particles, at least if the average distance between them is smaller than the spread of the single-particle wave functions. This indistinguishability has important consequences for the many-particle states and accordingly for the statistical description of a quantum many-particle system.

3.1 Indistinguishable particles

We present first a simple argument why indistinguishable particles have either symmetric or antisymmetric wave functions. Consider two quantum particles described by the wave function $\psi(\mathbf{r}_1, \mathbf{r}_2)$, representing an eigenstate of some two-particle Hamiltonian. If the two particles are indistinguishable, the Hamiltonian is invariant with respect to a permutation of spatial coordinates and momenta. Therefore $\psi(\mathbf{r}_2, \mathbf{r}_1)$ represents also an eigenstate and has the same eigenvalue as $\psi(\mathbf{r}_1, \mathbf{r}_2)$. We cannot distinguish the two states and therefore the two wave functions have to be the same up to a constant complex phase factor λ

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \lambda \psi(\mathbf{r}_2, \mathbf{r}_1). \quad (3.1)$$

Applying the same reasoning one more time, we arrive at

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \lambda^2 \psi(\mathbf{r}_1, \mathbf{r}_2) \quad (3.2)$$

from where follows that $\lambda^2 = 1$ and thus $\lambda = \pm 1$. Introducing the permutation operator \hat{P} through the relation $\hat{P}\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_2, \mathbf{r}_1)$, we see that Eq. (3.1) can be rewritten as

$$\hat{P}\psi(\mathbf{r}_1, \mathbf{r}_2) = \lambda \psi(\mathbf{r}_1, \mathbf{r}_2) \quad (3.3)$$

and two-particle states of two indistinguishable particles fulfill

$$\hat{P}\psi(\mathbf{r}_1, \mathbf{r}_2) = \pm \psi(\mathbf{r}_1, \mathbf{r}_2). \quad (3.4)$$

States with $\lambda = 1$ are called symmetric whereas states with $\lambda = -1$ are called antisymmetric.

It is an empirical fact that all known elementary particles are described either by symmetric or antisymmetric wave functions. The former are called *bosons*, the latter *fermions*. There are two known classes of fermions—twelve leptons (the electron, the muon, and the τ -particle with their corresponding neutrini as well as one anti-particle for each particle), and six quarks (up, down, strange, charm, bottom, top) together with their antiparticles, the anti-quarks. The interactions between fermions are mediated by bosons—the electromagnetic interaction by photons, the weak interaction (responsible for the β decay) by the (massive) W^+ , W^- and Z bosons. Quarks (and anti-quarks) are also coupled by electromagnetic and weak forces, but in addition they experience the strong

interaction mediated by gluons. Gravitation is imagined to be mediated by gravitons, but these particles have not yet been detected. According to a fundamental result of relativistic quantum field theory there is an intimate relationship between spin and statistics. Under some rather weak assumptions one finds that bosons have integer spins and fermions have half-integer spins. In fact, all the known elementary particles agree with the spin-statistics connection. Leptons and quarks have spin $1/2$, the photon and the W^+ , W^- and Z bosons have spin 1, while the graviton is a spin 2 particle.

Baryons, such as the neutron and the proton, are bound states of three quarks. Mesons, such as π^+ , π^0 and π^- , consist of two quarks. At not too large energies, these compound particles behave as elementary particles, with spin $1/2$ in the case of neutrons and protons and spin zero in the case of pions. Ambient temperatures ($k_B T \approx 25$ meV) are of course many orders of magnitude below the energies required for uncovering the internal structure of hadrons or pions ($\gtrsim 500$ MeV). At typical thermal energies even atoms can be treated as elementary particles with bosonic or fermionic character. Thus, ^3He , with its two protons, one neutron and two electrons, behaves as a fermion, while ^4He , which has one neutron more than ^3He , can be treated as a boson. Clearly the interaction potential between the Helium atoms reflects the atomic structure, with the long-range Van der Waals attraction induced by virtual electronic excitations and the hard core repulsion reflecting the effects of the Pauli principle on electrons of different atoms.

3.2 Occupation numbers

In order to understand how the state selection works for Fermi and Bose statistics, we consider a very simple example. We choose a single-particle space with only three basis vectors, $|\phi_1\rangle$, $|\phi_2\rangle$, $|\phi_3\rangle$. We construct all the possible three-particle states, using the tensor product

$$|i, j, k\rangle := |\phi_i\rangle \otimes |\phi_j\rangle \otimes |\phi_k\rangle.$$

For distinguishable particles, there are 3^3 independent 3-particle states. The situation is of course different for indistinguishable particles, where the vector $|i, j, k\rangle$ represents the same state as, for example, $|j, i, k\rangle$. For fermions there is only one antisymmetric state (up to a phase factor),

$$|\psi_f\rangle = \frac{1}{\sqrt{6}} (|1, 2, 3\rangle - |1, 3, 2\rangle - |2, 1, 3\rangle + |2, 3, 1\rangle + |3, 1, 2\rangle - |3, 2, 1\rangle) \quad (3.5)$$

where the sign of each term is given by the number of particle permutations that are necessary to move from $|1, 2, 3\rangle$ to this particular state vector. For bosons there are ten states fulfilling the requirement of symmetry, $|\psi_{b1}\rangle = |1, 1, 1\rangle$, $|\psi_{b2}\rangle = |2, 2, 2\rangle$, $|\psi_{b3}\rangle = |3, 3, 3\rangle$, $|\psi_{b4}\rangle = \frac{1}{\sqrt{3}}(|1, 1, 2\rangle + |1, 2, 1\rangle + |2, 1, 1\rangle)$, $|\psi_{b5}\rangle = \frac{1}{\sqrt{3}}(|1, 1, 3\rangle + |1, 3, 1\rangle + |3, 1, 1\rangle)$, $|\psi_{b6}\rangle = \frac{1}{\sqrt{3}}(|1, 2, 2\rangle + |2, 2, 2\rangle + |2, 2, 1\rangle)$, $|\psi_{b7}\rangle = \frac{1}{\sqrt{3}}(|1, 3, 3\rangle + |3, 1, 3\rangle + |3, 3, 1\rangle)$, $|\psi_{b8}\rangle = \frac{1}{\sqrt{3}}(|2, 2, 3\rangle + |2, 3, 2\rangle + |3, 2, 2\rangle)$, $|\psi_{b9}\rangle = \frac{1}{\sqrt{3}}(|2, 3, 3\rangle + |3, 2, 3\rangle + |3, 3, 2\rangle)$, $|\psi_{b10}\rangle = \frac{1}{\sqrt{6}}(|1, 2, 3\rangle + |1, 3, 2\rangle + |2, 1, 3\rangle + |2, 3, 1\rangle + |3, 1, 2\rangle + |3, 2, 1\rangle)$.

In addition to this rather tedious representation (which, however, is often very useful in computation), for both fermions and for bosons there exists an alternative simple representation of the many-particle states. This representation is based on so-called occupation numbers n_i which indicate how many particles are in a given single-particle state $|\phi_i\rangle$. Because of the Pauli exclusion principle, these numbers are restricted to 0 and 1 for fermions, whereas they can take arbitrary values for bosons. In both cases, the constraint $\sum_i n_i = N$ applies where N is the total number of particles. Our bosonic state $|\psi_{b9}\rangle$ can be thus represented by $n_1 = 0, n_2 = 1, n_3 = 2$ (one can easily check that all ten bosonic states above are represented by mutually different sets of occupation numbers).

In Section 2.1 we have calculated the entropy of N independent particles in a volume V with a given total energy E . This was possible for the case of distinguishable particles for which the product states $|\mathbf{k}_1\rangle \otimes \dots \otimes |\mathbf{k}_N\rangle$ are distinct. Clearly the number of states of given energy will be strongly reduced by restricting the ensemble to states with the appropriate symmetry. Thus the entropy for bosons or fermions will be smaller than that of distinguishable particles for the same macroscopic conditions. Unfortunately, the enumeration of available many-body states with bosonic or fermionic symmetry is very difficult, and also the evaluation of the canonical partition function poses problems due to the constraint of a fixed total particle number. We will see in the next section that for the grand canonical ensemble it is straightforward to calculate the partition function, both for independent bosons and for independent fermions.

3.3 Non-interacting fermions and bosons

Let's consider the Hamiltonian of N free particles

$$\hat{H} = \sum_{j=1}^N \frac{\hat{\mathbf{p}}_j^2}{2m} \quad (3.6)$$

where $\hat{\mathbf{p}}_j = -i\hbar \nabla_j$ is the momentum operator of particle j . We allow symmetric (for bosons) and antisymmetric (for fermions) wave functions which obey

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N) = \pm \psi(\mathbf{r}_1, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N). \quad (3.7)$$

For a single particle in a cubic box with volume $V = L^3$ and periodic boundary conditions, the eigenstate fulfills

$$\phi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (3.8)$$

where $\mathbf{k} = \frac{2\pi}{L} (\nu_1, \nu_2, \nu_3)$ where ν_1, ν_2, ν_3 are integers. The eigenstates of the N -particle Hamiltonian are

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \sum_P \sigma_P \phi_{\mathbf{k}_1}(\mathbf{r}_{p_1}) \dots \phi_{\mathbf{k}_N}(\mathbf{r}_{p_N}) \quad (3.9)$$

where $P = (p_1, \dots, p_N)$ enumerates permutations of $(1, 2, \dots, N)$ and $\sigma_P = 1$ for bosons while $\sigma_P = \pm 1$ for fermions (the sign is determined by the sign of permutation, *i.e.*, by the

number of two-particle exchanges necessary to move from $(1, 2, \dots, N)$ to P . (Note that this is simply a generalization the above example with three-particles states of indistinguishable particles.)

Now we switch to the occupation number representation to determine behavior of this system. We introduce $n_{\mathbf{k}}$ where $\sum_{\mathbf{k}} n_{\mathbf{k}}$ is free to vary (we use the grand canonical ensemble) and $n_{\mathbf{k}} = 0, 1$ for fermions while $n_{\mathbf{k}}$ are arbitrary natural numbers for bosons. (One can note that Eq. (3.9) yields zero for fermions if any two \mathbf{k}_i and \mathbf{k}_j are equal. This equation thus implicitly contains the Pauli exclusion principle.) The total energy and the total number of particles of a given system configuration $\{n_{\mathbf{k}}\}$ is given by the sum of contributions from individual levels

$$E = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}}, \quad N = \sum_{\mathbf{k}} n_{\mathbf{k}} \quad (3.10)$$

where $\varepsilon_{\mathbf{k}} = \hbar^2 \mathbf{k}^2 / (2m)$. The grand canonical partition function reads

$$\begin{aligned} \mathcal{Z}(T, V, \mu) &= \sum_{\{n_{\mathbf{k}}\}} e^{-\beta(E - \mu N)} = \sum_{\{n_{\mathbf{k}}\}} e^{-\beta \sum_{\mathbf{k}} (\varepsilon_{\mathbf{k}} - \mu) n_{\mathbf{k}}} = \\ &= \sum_{\{n_{\mathbf{k}}\}} \prod_{\mathbf{k}} e^{-\beta (\varepsilon_{\mathbf{k}} - \mu) n_{\mathbf{k}}} = \prod_{\mathbf{k}} \left(\sum_{n_{\mathbf{k}}} e^{-\beta (\varepsilon_{\mathbf{k}} - \mu) n_{\mathbf{k}}} \right) \end{aligned} \quad (3.11)$$

where $n_{\mathbf{k}} = 0, 1, 2, \dots$ for bosons and $n_{\mathbf{k}} = 0, 1$ for fermions. To see that the exchange of summation and product above is really correct, one needs to realize that for each term in the former expression, there is an identical term in the latter expression and vice versa.

The summation over one occupation number $n_{\mathbf{k}}$ is easy to carry out and we obtain

$$\mathcal{Z}(T, V, \mu) = \begin{cases} \prod_{\mathbf{k}} (1 - e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)})^{-1} & \text{for bosons,} \\ \prod_{\mathbf{k}} (1 + e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}) & \text{for fermions.} \end{cases}$$

The grand canonical potential $\Omega = -k_B T \ln \mathcal{Z}$ is

$$\Omega(T, V, \mu) = \begin{cases} k_B T \sum_{\mathbf{k}} \ln [1 - e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}] & \text{for bosons,} \\ -k_B T \sum_{\mathbf{k}} \ln [1 + e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}] & \text{for fermions.} \end{cases} \quad (3.12)$$

The average number of particles follows from Eq. (2.46) in the form

$$N = -\frac{\partial \Omega}{\partial \mu} = \begin{cases} \sum_{\mathbf{k}} \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} - 1} & \text{for bosons,} \\ \sum_{\mathbf{k}} \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} + 1} & \text{for fermions.} \end{cases} \quad (3.13)$$

With the help of the Euler equation $U = TS - pV + \mu N$ and the definition of the grand canonical potential $\Omega = U - TS - \mu N$ one obtains $\Omega = -pV$. Eq. (3.12) can thus be interpreted as an equation of state

$$\frac{pV}{k_B T} = \begin{cases} -\sum_{\mathbf{k}} \ln [1 - e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}] & \text{for bosons,} \\ \sum_{\mathbf{k}} \ln [1 + e^{-\beta(\varepsilon_{\mathbf{k}} - \mu)}] & \text{for fermions.} \end{cases} \quad (3.14)$$

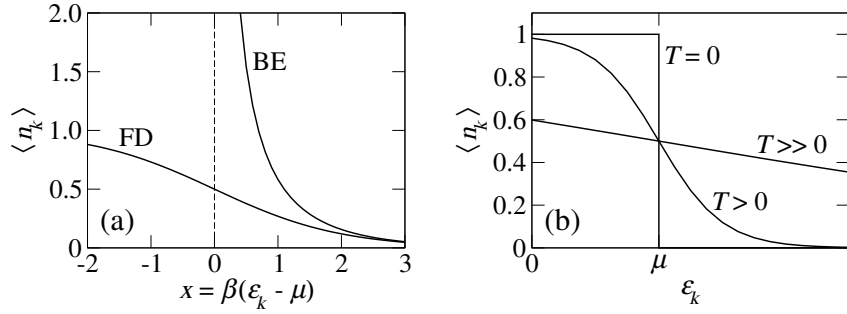


Figure 3.1: (a) Mean occupation number $\langle n_{\mathbf{k}} \rangle$ versus $x = \beta(\varepsilon_{\mathbf{k}} - \mu)$ for bosons and fermions. (b) Mean occupation number $\langle n_{\mathbf{k}} \rangle$ versus $\varepsilon_{\mathbf{k}}$ for fermions at various temperatures.

If instead of the of the chemical potential μ , the expected number of particles N is known, one has to first obtain μ using Eq. (3.13) and then plug the resulting value in Eq. (3.14).

The average occupation number of the state corresponding to \mathbf{k} is by definition

$$\langle n_{\mathbf{k}} \rangle = \frac{1}{\mathcal{Z}} \sum_{\{n_{\mathbf{k}'}\}} n_{\mathbf{k}} e^{-\beta \sum_{\mathbf{k}'} (\varepsilon_{\mathbf{k}'} - \mu) n_{\mathbf{k}'}}$$

which immediately motivates us to express it as $\langle n_{\mathbf{k}} \rangle = -k_B T \partial \ln \mathcal{Z} / \partial \varepsilon_{\mathbf{k}} = \partial \Omega / \partial \varepsilon_{\mathbf{k}}$ and therefore

$$\langle n_{\mathbf{k}} \rangle = \begin{cases} \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} - 1} & \text{for bosons,} \\ \frac{1}{e^{\beta(\varepsilon_{\mathbf{k}} - \mu)} + 1} & \text{for fermions.} \end{cases} \quad (3.15)$$

These are two famous distributions: that of Bose-Einstein applies to bosons and that of Fermi-Dirac applies to fermions.

Now it is time to study the differences between the two distributions. As one can see in Fig. 3.1a, occupation numbers for bosons and fermions become identical when $\beta(\varepsilon_{\mathbf{k}} - \mu) \gg 1$. That's because the expected occupation number $\langle n_{\mathbf{k}} \rangle$ is then small which implies that there are little interactions between particles and it thus does not matter whether the particles are bosons or fermions. In the case of bosons, when $\varepsilon_{\mathbf{k}} \rightarrow \mu^+$, $\langle n_{\mathbf{k}} \rangle$ grows without any bounds. This is of course unphysical (there true expected number of particles of the studied system is fixed) which simply means that the chemical potential of a bosonic system is always smaller than the ground energy and thus $\varepsilon_{\mathbf{k}} - \mu > 0$. In particular, if the ground energy is zero, the chemical potential has to be negative. Figure 3.1b then shows specifically fermions at various temperatures. Fermions at zero temperature occupy the lowest available energy levels with exactly one particle and all other levels remain empty. Energy of the highest occupied level is called the Fermi energy, ε_F , and the chemical potential of a fermionic system at zero temperature equals this Fermi energy. As temperature grows, more and more particles are excited to energy levels above ε_F .

3.4 Relations between ensembles

Equation (2.19) introduces the canonical partition function by summing over all possible microstates n . We rewrite it now to a sum over all possible energetic levels

$$Z(T, V, N) = \sum_n e^{-\beta E_n} = \sum_E g(E) e^{-\beta E} = \sum_E \Gamma(E, V, N) e^{-\beta E}. \quad (3.16)$$

where $g(E)$ is the degeneracy of the energy level E which is the same quantity as the number of microstates with energy E , $\Gamma(E, V, N)$. Equation (3.16) presents the canonical partition function as a weighted sum of the number of states corresponding to different energies where weights decay exponentially with energy E . Formally it corresponds to a discrete Laplace transformation (which is often labeled as the Z-transformation)⁶ of $\Gamma(E, V, N)$ into a new function $Z(T, V, N)$ where T (or, equivalently, $\beta = 1/(k_B T)$) is a newly introduced variable.

Similarly in Equation (3.11), one can split the summation over all sets of occupation numbers into a summation over various numbers of particles and then summation over all corresponding sets of occupation numbers. This yields

$$\mathcal{Z}(T, V, \mu) = \sum_{N=0}^{\infty} \sum_{\substack{\{n_{\mathbf{k}}\} \\ \sum n_{\mathbf{k}} = N}} e^{-\beta(E - \mu N)} = \sum_{N=0}^{\infty} e^{\beta \mu N} \sum_{\substack{\{n_{\mathbf{k}}\} \\ \sum n_{\mathbf{k}} = N}} e^{-\beta E} = \sum_{N=0}^{\infty} z^N Z(T, V, N)$$

where we introduced a so-called fugacity $z := \exp(\beta \mu)$. This tells us that the grand canonical partition function is simply a weighted sum of canonical partition functions with weights changing exponentially with N and depending on the system's fugacity z . Again, formally this corresponds to a discrete Laplace transformation of the canonical partition function into a grand canonical partition function where a new variable μ is introduced instead of N in the process. In summary, we uncovered a matryoshka doll of ensembles where $Z(T, V, N)$ is hiding in $\mathcal{Z}(T, V, \mu)$ and $\Gamma(E, V, N)$ is hiding in $Z(T, V, N)$.

3.5 Black body radiation

We now consider gas of photons for which it is customary to use angular frequency ω instead of wave vectors \mathbf{k} . The conversion relation is simple, $\omega_{\mathbf{k}} = c|\mathbf{k}|$, and the energy is $E_{\mathbf{k}} = \hbar \omega = \hbar c|\mathbf{k}|$. Until now we worked with density $V/(2\pi)^3$ in the three-dimensional \mathbf{k} -space. We now introduce the number of states in a small range of frequencies $g(\omega) d\omega$. Starting from the familiar \mathbf{k} -space density, we can write

$$g(\omega) d\omega = \frac{V}{(2\pi)^3} \times 4\pi k^2 dk \times 2 \quad (3.17)$$

⁶Both integral and discrete Laplace transformation are extensions of the standard Fourier transformation. For a continuous function $f(t)$, the Fourier transformation is $F(s) = \int_{-\infty}^{\infty} f(t) e^{-ist} dt$ while the Laplace transformation is $L(s) = \int_0^{\infty} f(t) e^{-st} dt$.

where the second term corresponds to the volume of the \mathbf{k} -space corresponding to the angular frequency ω and the third term captures the fact that there are two independent photon polarizations. By dividing with $d\omega$, we create $dk/d\omega$ on the right side which equals $1/c$. Then after substituting $k^2 = \omega^2/c^2$, we obtain

$$g(\omega) = \frac{V\omega^2}{\pi^2 c^3}.$$

Gas of photons always has zero chemical potential because the number of photons is not conserved (a photon can be absorbed when exciting an electron in an atom to a higher energetic level and later emitted) and the corresponding Lagrange multiplier is thus zero. The expected occupation number of state with energy $\hbar\omega$ is thus $n_{\text{occup}} = 1/(e^{\hbar\omega/k_B T} - 1)$ and the spectral density of photons is simply

$$n(\omega) = N(\omega)/V = g(\omega)n_{\text{occup}}/V = \frac{1}{\pi^2 c^3} \frac{\omega^2}{e^{\hbar\omega/k_B T} - 1}.$$

The spectral energy density $u(\omega)$ is then obtained by multiplying $n(\omega)$ with energy of the corresponding photons, yielding

$$u(\omega) = \hbar\omega n(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\hbar\omega/k_B T} - 1}. \quad (3.18)$$

which is the famous Planck's law.

The low frequency behavior of $u(\omega)$ can be found by approximating $e^{\hbar\omega/k_B T} \approx 1 + \hbar\omega/k_B T$ which leads to

$$u(\omega) = \frac{\omega^2}{\pi^2 c^3} k_B T = \frac{g(\omega)}{V} k_B T$$

which is usually referred to as the Rayleigh-Jeans law. It agrees with the equipartition theorem known from classical statistical mechanics (or as high-temperature limit in quantum statistical mechanics): both degrees of freedom (represented by two light polarizations) bring $\frac{1}{2}k_B T$ which is further multiplied with the density of states in the ω -space. This form of $u(\omega)$ is obviously flawed when applied to the whole frequency spectrum because it implies that energy density monotonically grows with ω , thus resulting in everything being “fried” by gamma radiation (a so-called ultraviolet catastrophe). This problem was only solved after seeds of quantum mechanics began to emerge. The central point of the correct derivation (which was presented by Max Planck in 1900) was the postulate that energy of the electromagnetic field is not infinitely divisible but it instead comes in the elementary units (which were recognized to represent individual photons only later in Einstein's seminal work on the photoelectric effect) of size $\hbar\omega$. If such quanta exist, they imply that there exist an energy gap between the ground state (with no photons of frequency ω) and the first excited state (with one photon of frequency ω). The exponentially decaying Boltzmann's weight $e^{-\beta E}$ then automatically suppresses high-frequency states exactly as we need. Without the energy gap, the equipartition theorem would enter the stage and let the ultraviolet catastrophe happen.

The next step is to compute the total energy density of the electromagnetic field. This is straightforward to write down as

$$u = \int_0^\infty u(\omega) d\omega = \frac{\hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3 d\omega}{e^{\hbar\omega/k_B T} - 1} = \frac{k_B^4 T^4}{\pi^2 c^3 \hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} \quad (3.19)$$

where we used the substitution $x := \hbar\omega/k_B T$. It is important to note that the integral here is just a numerical factor which does not depend on properties of the system (such as temperature T) or some physical constants. We thus see that the energy density is proportional to T^4 which agrees with the Stefan-Boltzmann law (this law describes the total power P emitted per surface area of a black body which one can show to be closely related to the energy density u via $P = uc/4$). To evaluate the integral, we write $1/(e^x - 1) = e^{-x}/(1 - e^{-x})$ where $e^{-x} < 1$ for $x > 0$. The fraction can now be expanded through the geometric series as $e^{-x}/(1 - e^{-x}) = e^{-x} \sum_{n=0}^\infty (e^{-x})^n = \sum_{n=1}^\infty e^{-nx}$. After exchanging the order of integration and summation, the integral of our interest becomes

$$\sum_{n=1}^\infty \int_0^\infty x^3 e^{-nx} dx = \sum_{n=1}^\infty \frac{6}{n^4} = 6\zeta(4) = \frac{\pi^4}{15}.$$

Here we benefited from the fact that integrals $\int x^3 e^{-nx} dx$ are easily evaluated by, for example, repeated application of per-parts integration, leading to the result $6/n^4$. Finally, $\zeta(m)$ is the well-known Riemann zeta function and $\zeta(4) = \pi^4/90 \approx 1.08232$.

3.6 Electrons in metal

The Hamiltonian of free particles given by Eq. (3.6) is a good approximation for conduction electrons in simple metals such as potassium (K) and sodium (Na). At zero temperature, electrons occupy the lowest available energy levels. Energy of the highest occupied level is referred to as the Fermi energy, ε_F , to which corresponds the wave vector with magnitude k_F . The number of all states with $|\mathbf{k}| \leq k_F$ must be enough to incorporate N electrons which leads to

$$2 \sum_{\substack{\mathbf{k} \\ |\mathbf{k}| \leq k_F}} 1 = N \quad (3.20)$$

where the additional factor of two stands for two different spin directions which allow two electrons to share the same wave vector \mathbf{k} . This sum can be changed into integration over the sphere with radius k_F which has volume $\frac{4}{3}\pi k_F^3$ and thus

$$N = 2 \times \frac{V}{(2\pi)^3} \times \frac{4}{3}\pi k_F^3.$$

This allows us to express k_F as a function of electron density $n := N/V$ and then finally express $\varepsilon_F = \hbar^2 k_F^2 / (2m)$ in the form

$$k_F = (3n\pi^2)^{1/3}, \quad \varepsilon_F = \frac{\hbar^2}{2m} (3n\pi^2)^{2/3}. \quad (3.21)$$

As we saw before from the Fermi-Dirac distribution, ε_F is equal to the chemical potential at zero temperature. For alkali metals, the Fermi energy is of the order of electron volts and ε_F/k_B is of the order of $3 \cdot 10^4$ K. Therefore at room temperature, $k_B T \ll \varepsilon_F$.

The density of states with energy ε can be computed similarly as before for the gas of bosons. We write

$$g(\varepsilon) d\varepsilon = 2 \times \frac{V}{(2\pi)^3} \times 4\pi k^2 dk \implies g(\varepsilon) = \frac{3nV}{2} \frac{\sqrt{\varepsilon}}{\varepsilon_F^{3/2}} \quad (3.22)$$

The Fermi-Dirac distribution of occupation numbers now leads to

$$N = \int_0^\infty g(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} d\varepsilon, \quad U = \int_0^\infty \varepsilon g(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} d\varepsilon. \quad (3.23)$$

The usual volume-normalized quantities n and u are then defined as $n := N/V$ and $u := U/V$. Note that these are in fact the final results which allows us to determine μ for any given N and T (from the first equation)⁷ and then compute the resulting U (from the second equation). The only problem is that these integral equations cannot be solved analytically. One has to invest quite some work to find the system's behavior in the limit of small temperature (that is, $k_B T \ll \varepsilon_F$ which, as said above, holds up to thousands of Kelvins). That is exactly what we are going to do now.

Introducing $f(\varepsilon) := 1/(e^{\beta(\varepsilon-\mu)} + 1)$, we can write

$$\frac{df}{d\varepsilon} = -\frac{\beta}{4} \frac{1}{\cosh^2 \frac{\beta(\varepsilon-\mu)}{2}} = -\frac{\beta}{4} \operatorname{sech}^2 \frac{\beta(\varepsilon-\mu)}{2}. \quad (3.24)$$

Now we are ready for per-parts integration

$$\int_0^\infty \varepsilon^\nu f(\varepsilon) d\varepsilon = -\frac{1}{\nu+1} [\varepsilon^{\nu+1} f(\varepsilon)]_0^\infty - \frac{1}{\nu+1} \int_0^\infty \varepsilon^{\nu+1} \frac{df}{d\varepsilon} d\varepsilon$$

where the first term equals zero in both 0 and ∞ . The second term can be developed further by substituting $x := \varepsilon - \mu$ and extending the lower integration bound to $-\infty$ which we can do because $\beta\mu \gg 1$ and the value of the integrand is therefore very small for $x < -\mu$. We obtain

$$\frac{\beta}{4(\nu+1)} \int_{-\mu}^\infty (\mu+x)^{\nu+1} \operatorname{sech}^2 \frac{\beta x}{2} dx.$$

Since $\operatorname{sech}^2 \frac{\beta x}{2}$ is negligible for $\beta x \gg 1$, the relevant contribution to the integral comes from $|x| \ll \mu$ and we can thus extend the integration region to $(-\infty, \infty)$ and then expand the integrand in powers of x/μ , neglecting higher-order terms. The above expression then leads to

$$\frac{\beta\mu^{\nu+1}}{4(\nu+1)} \left\{ \int_{-\infty}^\infty \operatorname{sech}^2 \frac{\beta x}{2} dx + \frac{(\nu+1)\nu}{2\mu^2} \int_{-\infty}^\infty x^2 \operatorname{sech}^2 \frac{\beta x}{2} dx + O[(k_B T)^4] \right\}. \quad (3.25)$$

⁷This agrees what we said about computation with the grand canonical ensemble in general: the requirement of achieving a certain expected number of particles fixes the chemical potential.

The terms proportional to odd powers of x do not appear here because they are odd functions and their integral over the symmetric range $(-\infty, \infty)$ is therefore zero. Using the simple substitution $\beta x = y$ and integrals

$$\int_{-\infty}^{\infty} \text{sech}^2 \frac{y}{2} dy = 4, \quad \int_{-\infty}^{\infty} y^2 \text{sech}^2 \frac{y}{2} dy = \frac{4\pi^2}{3}$$

we finally obtain

$$\int_0^{\infty} \varepsilon^\nu f(\varepsilon) d\varepsilon \approx \frac{\mu^{\nu+1}}{\nu+1} \left\{ 1 + \frac{\pi^2}{6} \nu(\nu+1) \left(\frac{k_B T}{\mu} \right)^2 + O[(k_B T/\mu)^4] \right\}.$$

Equations (3.23) now yield

$$n = n \left(\frac{\mu}{\varepsilon_F} \right)^{3/2} \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 \right], \quad (3.26)$$

$$u = \frac{3}{5} n \varepsilon_F \left(\frac{\mu}{\varepsilon_F} \right)^{5/2} \left[1 + \frac{5\pi^2}{8} \left(\frac{k_B T}{\mu} \right)^2 \right]. \quad (3.27)$$

Note that particle densities on both sides of Eq. (3.26) cancel and only an equation of one variable (μ) remains. For $T = 0$ K, this further simplifies to $\mu = \varepsilon_F$ (as always, the zero-temperature chemical potential equals ε_F) and the second to $u = \frac{3}{5} n \varepsilon_F$ (this simple result can be derived directly from averaging the electron energy inside the Fermi sphere).

For small temperatures where $k_B T \ll \varepsilon_F$, we can assume that the chemical potential changes only little from its zero-temperature limit. If we write $\mu = \varepsilon_F + (\mu - \varepsilon_F)$, $\mu - \varepsilon_F$ thus becomes a small parameter and we can develop equations for n and u in powers of $\mu - \varepsilon_F$. Eq. (3.26) gives

$$1 = \left(1 + \frac{3}{2} \frac{\mu - \varepsilon_F}{\varepsilon_F} \right) \left[1 + \frac{\pi^2}{8} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right] \implies \frac{\mu - \varepsilon_F}{\varepsilon_F} = -\frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2.$$

where we neglected the second-order term proportional to $k_B T(\mu - \varepsilon_F)$ and replaced $k_B T/\mu$ with $k_B T/\varepsilon_F$ because $k_B T/\mu$ is already a small number and thus we don't need its second-order correction which could be found by writing $\mu = \varepsilon_F + (\mu - \varepsilon_F)$ also here.

Eq. (3.27) can be also approximated using $(\mu/\varepsilon_F)^{5/2} \approx 1 + \frac{5}{2} (\mu - \varepsilon_F)/\varepsilon_F$. In combination with the previous result for $(\mu - \varepsilon_F)/\varepsilon_F$, we get

$$u = \frac{3}{5} n \varepsilon_F \left[1 - \frac{5\pi^2}{24} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right] \left[1 + \frac{5\pi^2}{8} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right] \approx \frac{3}{5} n \varepsilon_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right]. \quad (3.28)$$

The specific heat per electron follows in the form

$$\frac{C_V}{N} = \frac{\partial u}{\partial T} = \frac{\pi^2}{2} k_B \left(\frac{k_B T}{\varepsilon_F} \right). \quad (3.29)$$

Finally, the equation of state (3.14) can be integrated per parts, leading to

$$p = \frac{k_B T}{V} \int_0^\infty g(\varepsilon) \ln [1 + e^{-\beta(\varepsilon-\mu)}] d\varepsilon = \frac{2}{3V} \int_0^\infty g(\varepsilon) \frac{\varepsilon}{e^{\beta(\varepsilon-\mu)} + 1} d\varepsilon = \frac{2}{3} u. \quad (3.30)$$

where we used that $g(\varepsilon) \sim \sqrt{\varepsilon}$ and thus $\int g(\varepsilon) d\varepsilon = \frac{2}{3} g(\varepsilon) \varepsilon$. Because of the Pauli exclusion principle, the pressure does not vanish at $T = 0$ K but remains at some base level proportional to u .

3.7 The Bose-Einstein condensation

As in the black body radiation section, we shall study again non-interacting free bosons with one important difference: this time it will be bosonic particles with non-zero mass m . This has two implications: particle energy now reads $\varepsilon = \mathbf{p}^2/2m$ where $\mathbf{p} = \hbar \mathbf{k}$ and the chemical potential μ does not need to be zero because these particles cannot be destroyed or created (at least not at low energies) and their number is thus conserved. What consequences follow from these seemingly small differences?

We start again with the density of states in the energy space

$$g(\varepsilon) d\varepsilon = \frac{V}{(2\pi)^3} \times 4\pi k^2 dk \implies g(\varepsilon) = \frac{Vm^{3/2}}{\sqrt{2\pi^2\hbar^3}} \sqrt{\varepsilon}$$

where we used that $k^2 = 2m\varepsilon/\hbar^2$ and therefore $dk/d\varepsilon = \sqrt{m}/(\hbar\sqrt{2\varepsilon})$. The expected number of particles is obtained by combining $g(\varepsilon)$ with the average occupation number for the Bose-Einstein statistics

$$N(\mu) = \int_0^\infty \frac{g(\varepsilon)}{e^{(\varepsilon-\mu)/k_B T} - 1} d\varepsilon. \quad (3.31)$$

This equation again allows us to achieve a desired number of particles by varying μ . One can see that as μ increases, $N(\mu)$ grows. However, we know that chemical potential of a bosonic system cannot be greater than the ground level energy ε_0 (if this condition is violated, the exponent value in the Bose-Einstein distribution is less than zero which results in negative occupation probability). Since $\varepsilon_0 = 0$ here, μ must be negative or, in the limit case, zero. The highest possible number of particles is thus achieved when μ .

If it would hold that $\lim_{\mu \rightarrow 0^-} N(\mu) = \infty$, we could stop here with the statement that nothing interesting happens.⁸ However, the integral corresponding to $\mu = 0$ can be evaluated and leads to the maximal possible number of particles

$$N_{\max} = \int_0^\infty \frac{g(\varepsilon)}{e^{\varepsilon/k_B T} - 1} d\varepsilon = \frac{Vm^{3/2}}{\sqrt{2\pi^2\hbar^3}} \int_0^\infty \frac{\sqrt{\varepsilon}}{e^{\varepsilon/k_B T} - 1} d\varepsilon = \frac{V(mk_B T)^{3/2}}{\sqrt{2\pi^2\hbar^3}} \int_0^\infty \frac{\sqrt{z}}{e^z - 1} dz \quad (3.32)$$

⁸Such a divergence of $N(\mu)$ is in principle imaginable because the occupation number of the ground state diverges when $\mu = 0$. One can check that when the same computation is done in two dimensions, the resulting $N(\mu)$ diverges when $\mu \rightarrow 0^-$.

where we substituted $z := \varepsilon/k_B T$. After introducing $h = 2\pi\hbar$ and the thermal de Broglie wavelength of particles $\lambda = h/\sqrt{2\pi m k_B T}$, we obtain

$$N_{\max} = \frac{V}{\lambda^3} \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{\sqrt{z}}{e^z - 1} dz.$$

The integral can be evaluated using expansion in a geometric series⁹

$$\begin{aligned} \int_0^\infty \frac{\sqrt{z}}{e^z - 1} dz &= \int_0^\infty e^{-z} \sqrt{z} \frac{1}{1 - e^{-z}} dz = \int_0^\infty dz e^{-z} \sqrt{z} \sum_{n=0}^\infty e^{-zn} = \sum_{n=1}^\infty \int_0^\infty e^{-zn} \sqrt{z} dz = \\ &= \left| \begin{array}{l} \sqrt{z} = t \\ dz = 2t dt \end{array} \right| = 2 \int_0^\infty e^{-nt^2} t^2 dt = -2 \frac{\partial}{\partial n} \int_0^\infty e^{-nt^2} dt = -\frac{\partial}{\partial n} \sqrt{\frac{\pi}{n}} = \frac{\sqrt{\pi}}{2n^{3/2}}. \end{aligned}$$

The resulting N_{\max} is

$$N_{\max} = \frac{V}{\lambda^3} \sum_{n=1}^\infty \frac{1}{n^{3/2}} = \frac{V}{\lambda^3} \zeta(3/2).$$

where $\zeta(3/2) \approx 2.61238$.

Imagine that we start at a situation where $N < N_{\max}$ and thus no problems with the above computation arise. When N and V are fixed and temperature T decreases, λ grows without any bounds and N_{\max} decreases until it becomes equal to N . This occurs at a so-called critical temperature T_c which is easy to compute as

$$T_c = \frac{h^2}{2\pi m k_B} \left(\frac{n}{\zeta(3/2)} \right)^{2/3}.$$

When the temperature decreases further and the number of particles is thus higher than N_{\max} , the extra particles are forced into the ground state which, as we saw, has an infinite “storage capacity” at $\mu = 0$. Denoting the number of these particles as N_0 , it follows that $N_0 = N - N_{\max}$ and therefore

$$\frac{N_0}{N} = 1 - \frac{\zeta(3/2)}{n\lambda^3} = 1 - \left(\frac{T}{T_c} \right)^{3/2}. \quad (3.33)$$

where we denoted $\lambda = x/\sqrt{T}$ and used that $\zeta(3/2)/(nx^3) = T_c^{-3/2}$. When $T < T_c$, a macroscopic fraction of all particles occupy the ground state—this phenomenon is known as the Bose-Einstein condensate. The ratio N_0/N is zero above T_c and non-zero under T_c and thus acts as an order parameter for the phase transition between the normal and condensed state. Fig. 3.2 shows its dependency on temperature.

We can further compute the corresponding energy of particles in the condensation phase ($T \leq T_c$ and thus $\mu = 0$). Similarly to Eq. (3.32), we write

$$U = \int_0^\infty \frac{\varepsilon g(\varepsilon)}{e^{\varepsilon/k_B T} - 1} d\varepsilon = \dots = \frac{3}{2} k_B T \frac{V}{\lambda^3} \zeta(5/2) \quad (3.34)$$

⁹The general formula is $\int_0^\infty x^a/(e^x - 1) dx = \zeta(a+1)a!$. Factorial of half-integer a can be computed from the usual recursive equation $a! = a \times (a-1)!$ using $(1/2)! = \sqrt{\pi}/2$.

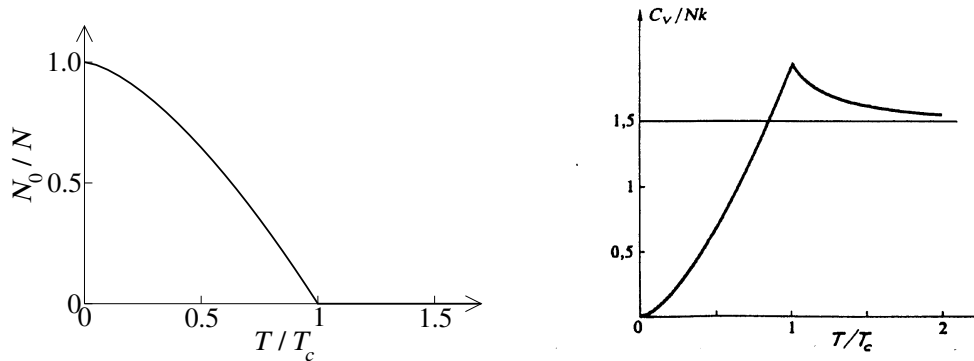


Figure 3.2: The Bose-Einstein condensate: the dependency of the order parameter N_0/N (left) and heat capacity (right) on temperature.

where exactly the same steps as before for N_{\max} are necessary and $\zeta(5/2) \approx 1.34149$. Particles in the ground state require no special treatment in this case as their energy is zero. Again denoting $\lambda = x/\sqrt{T}$, heat capacity follows as

$$C_V = \frac{\partial U}{\partial T} = \frac{15}{4} \frac{V k_B T^{3/2}}{x^3} \zeta(5/2) = \frac{15\zeta(5/2)}{4\zeta(3/2)} N k_B \left(\frac{T}{T_c} \right)^{3/2} \approx 1.926 N k_B \left(\frac{T}{T_c} \right)^{3/2} \quad (3.35)$$

and vanishes polynomially as $T \rightarrow 0$. It is also possible to derive C_V for $T > T_c$, that derivation is more complicated because μ is then not zero but depends on T instead. We do not pursue this here and only show the result in Fig. 3.2. Note that C_V is discontinuous at T_c but it does not diverge (thus there is no latent heat present and we therefore speak about a second-order transition¹⁰ here). The high temperature limit of C_V is $\frac{3}{2} N k_B$ as one would expect.

The Bose-Einstein condensation was first predicted by Satyendra Nath Bose and Albert Einstein in 1924–25. However, it was not until 1995 when the first condensate was produced at the NIST-JILA lab at the Colorado University a gas of rubidium atoms cooled to 170 nK. Experiment leaders Eric Cornell and Carl Wieman, together with Wolfgang Ketterle from MIT who produced the Bose-Einstein condensate four months later using a gas of sodium atoms, received the 2001 Nobel Prize in Physics. Fig. 3.3 illustrates these experimental results.

3.8 Degenerate atomic gases

During the last decade a whole new field has been created thanks to experiments on ultracold trapped gases of alkali atoms. These experiments were made possible thanks to advances in the techniques of trapping (magnetically or optically) and of cooling (down to

¹⁰Familiar examples of first-order phase transitions are the melting of ice or the boiling of water where at the critical temperature (0°C and 100°C, respectively) a large amount of heat needs to be added to the system at fixed temperature to allow for the phase to change.

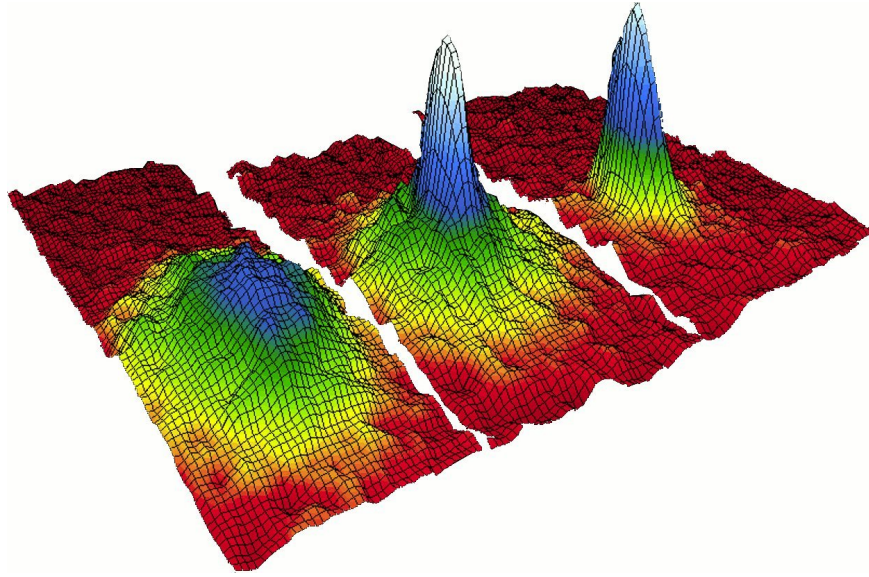


Figure 3.3: Velocity distribution confirming the discovery of the Bose-Einstein condensate in a gas of rubidium atoms: just before the appearance of the condensate ($T > T_c$, left), just after the appearance of the condensate ($T < T_c$, middle), and after further evaporation which left a sample of nearly pure condensate (right). Credits: NIST-JILA lab.

fractions of μK). In this way, atomic clouds of up to 10^{10} atoms have been formed, with densities between 10^{11} cm^{-3} and $5 \cdot 10^{15} \text{ cm}^{-3}$. We can estimate the critical temperature, *i.e.*, the temperature below which quantum-statistical effects become important, using the relation $\lambda^3 n \approx 1$ between the atomic density n and the thermal de Broglie wavelength λ . The masses of the bosonic alkali atoms ^{87}Rb , ^{23}Na , ^7Li and of the fermionic atoms ^{40}K , ^6Li are of the order of 10^{-22} g , leading to critical temperatures between $0.15 \mu\text{K}$ and $3 \mu\text{K}$.

Atomic gases can be considered to be systems of weakly interacting particles, at least in some ranges of parameters. We neglect in the following the interaction, but take into account the confinement in terms of a harmonic trapping potential

$$U_{\text{tr}}(\mathbf{r}) = \frac{1}{2} m \omega_0^2 \mathbf{r}^2. \quad (3.36)$$

In this case the single-particle basis is given by eigenfunctions of the three-dimensional harmonic oscillator $\varphi_{n_1, n_2, n_3}(\mathbf{r})$ with energy eigenvalues

$$\varepsilon_{n_1, n_2, n_3} = \hbar \omega_0 \left(\frac{3}{2} + n_1 + n_2 + n_3 \right), \quad n_i = 0, 1, 2, \dots \quad (3.37)$$

The lowest level ($n_1 = n_2 = n_3 = 0$) has the wave function

$$\varphi_{0,0,0}(\mathbf{r}) = \left(\frac{m \omega_0}{\pi \hbar} \right)^{3/4} e^{-\frac{m \omega_0 \mathbf{r}^2}{2 \hbar}}. \quad (3.38)$$

Alkali atoms have both nuclear spin \mathbf{I} and electron spin \mathbf{S} , $S = 1/2$, leading to the total spin $\mathbf{F} = \mathbf{I} + \mathbf{S}$. The *bosonic* atoms ^{87}Rb , ^{23}Na , ^7Li have nuclear spin $I = 3/2$ and therefore

the total spin is either $F = 1$ or $F = 2$. In most experiments on bosonic atoms the total spin is polarized, for instance $F = 2$, $m_F = 2$ where m_F is the magnetic quantum number, and the spin degrees of freedom have not to be taken into account, at least at temperatures for which other hyperfine levels are not thermally excited. *Fermionic* gases are preferentially prepared in two different and approximately equally populated spin states, for instance $F = 9/2$, $m_F = 9/2$ together with $F = 9/2$, $m_F = 7/2$ for ^{40}K .

In the following we limit ourselves to the bosonic case with spin zero at $T = 0$. For N bosons the ground state wave function is given by

$$\Psi_0(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_{i=1}^N \varphi_{0,0,0}(\mathbf{r}_i). \quad (3.39)$$

The Bose distribution function is

$$\langle n_{n_1, n_2, n_3} \rangle = \frac{1}{e^{\beta(\varepsilon_{n_1, n_2, n_3} - \mu)} - 1}, \quad (3.40)$$

where the energy eigenvalues $\varepsilon_{n_1, n_2, n_3}$ are given by Eq. (3.37). The analysis proceeds now very much like in the case of free particles. For $T < T_c$, *i.e.*, for temperatures where $\mu = 0$, we treat the number N_0 of particles occupying the lowest level separately and write

$$N = N_0 + \sum_{n_1, n_2, n_3} \frac{1}{e^{\beta \varepsilon_{n_1, n_2, n_3}} - 1} \quad (3.41)$$

where the sum is only over positive n_1, n_2, n_3 . To evaluate the sum for large enough N we assume $k_B T \gg \hbar \omega_0$. This allows us to neglect the zero-point energy and to replace the sum by an integral,

$$N \approx N_0 + \int_0^\infty dn_1 \int_0^\infty dn_2 \int_0^\infty dn_3 \frac{1}{e^{\beta \hbar \omega_0 (n_1 + n_2 + n_3)} - 1}. \quad (3.42)$$

Introducing the integration variable $n = n_1 + n_2 + n_3$, we get

$$\begin{aligned} \int_0^\infty dn_1 \int_0^\infty dn_2 \int_0^\infty dn_3 f(n_1 + n_2 + n_3) &= \int_0^\infty dn_1 \int_0^\infty dn_2 \int_{n_1 + n_2}^\infty dn f(n) = \\ &= \int_0^\infty dn f(n) \int_0^n dn_1 \int_0^{n - n_1} dn_2 = \frac{1}{2} \int_0^\infty dn n^2 f(n). \end{aligned}$$

With the substitution $x = \beta \hbar \omega_0 n$ we end up with

$$N \approx N_0 + \frac{1}{2} \left(\frac{k_B T}{\hbar \omega_0} \right)^3 \int_0^\infty dx \frac{x^2}{e^x - 1} = N_0 + \zeta(3) \left(\frac{k_B T}{\hbar \omega_0} \right)^3, \quad (3.43)$$

where $\zeta(3) \approx 1.202$. The resulting fraction of atoms in the condensate is

$$\frac{N_0}{N} \approx 1 - \left(\frac{T}{T_c} \right)^3 \quad (3.44)$$

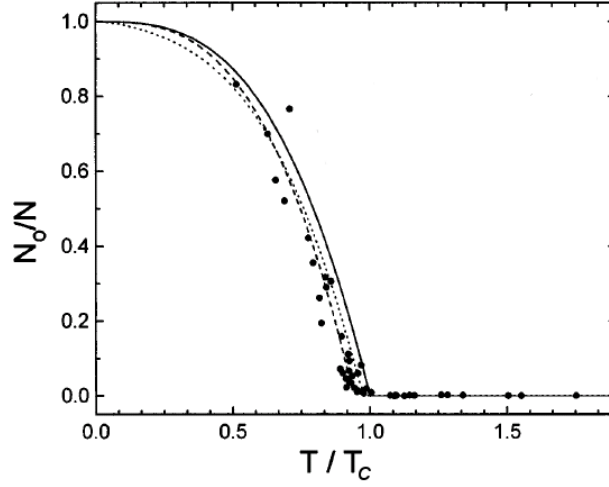


Figure 3.4: Condensed fraction N_0/N as a function of T/T_c in a dilute gas of Rubidium atoms. Symbols are experimental results (J. R. Ensher *et al.*, Physical Review Letters **77**, 4984 (1996)), the solid line represents Eq. (3.43). The dashed line is the best-performing least-squares fit of the form $N_0/N = 1 - (T/T_x)^3$ which yields $T_x = 0.94T_c$. The dotted line is the analytical result including corrections for finite N .

which is zero at the critical temperature

$$k_B T_c = \hbar \omega_0 \left(\frac{N}{\zeta(3)} \right)^{1/3}. \quad (3.45)$$

Clearly $k_B T_c \gg \hbar \omega_0$, our analysis is therefore justified below T_c .

The Fourier transform of the particle density $N|\varphi_{0,0,0}(\mathbf{r})|^2$ gives a momentum distribution of width $\sqrt{m\hbar\omega_0}$. This is much narrower than the width $\sqrt{mk_B T}$ of a classical Maxwell distribution at typical temperatures (or also the momentum distribution of excited particles in the case we have studied, *i.e.*, for $k_B T \gg \hbar \omega_0$). Experimentally, the Bose-Einstein condensation is evidenced by the appearance of a narrow peak over a broader thermal distribution, both in the spatial coordinates and in the velocity (see Fig. 3.3). The spatial distribution is determined directly by light scattering. In order to determine the momentum distribution, the (optical or magnetic) trap potential is switched off and after a (tiny) lapse of time the shadow of the expanding cloud is measured by optical means. For ideal gas there are no collisions and therefore one can determine the velocity distribution from data measured at different times. Fig. 3.4 shows a comparison with our analytical result for N_0/N with experimental results. Our prediction based on the ideal Bose gas works very well. For further information about the Bose-Einstein condensation, see W. Ketterle, Physics Today, December 1999 (a brief review of experiments), K. Burnett *et al.*, Physics Today, December 1999 (a brief review of the theoretical basis), and A. J. Leggett, Review of Modern Physics **73**, 307, 2001 (fundamental theoretical concepts).

3.9 Debye's model for the specific heat of a solid

As we already mentioned in Sec. 2.4, Einstein's model for the specific heat of solids fails in the low-temperature regime where it predicts $C_V \sim e^{-x/T}/T^2$ for $T \rightarrow 0$ whereas from experiments follows $C_V \sim T^3$. The obvious limitation of the Einstein's model is that it assumes only one oscillation frequency ω_0 for all atoms which then results in an energy gap $\hbar\omega_0$ between the ground state and the first excited state. As we have already seen it many times, the existence of such an energy gap implies exponentially suppressed occupation of the first excited state when $T \rightarrow 0$ and, in consequence, a specific heat that vanishes exponentially fast. We can improve on that by including interactions of atoms in our analysis in a simplified, yet sufficient way. These interactions allow the atoms to oscillate jointly, thus allowing waves with arbitrary amplitudes and energies to be propagate in the solid. Due to the existence of large-amplitude (low-energy) waves in a macroscopic crystal where $V \rightarrow \infty$, $N \rightarrow \infty$, and $N/V = \text{const.}$, the energy gap does not exist anymore and our prospects of obtaining non-exponentially vanishing C_V are good.

Collective excitations of periodic solids are usually referred to as *phonons*. The name phonon comes from the Greek word for sound or voice because long-wavelength phonons give rise to sound. The language similarity between photons and phonons is also no coincidence: while photons are excitations of the electromagnetic field, phonons are vibrational excitations of the lattice. There are two possible transverse (perpendicular to the propagation) displacement directions and one possible longitudinal (parallel to the propagation) displacement direction. Propagation velocities for longitudinal and transverse waves—usually labeled as c_l and c_t , respectively—are different in general. While the classical interpretation of both electromagnetic field and vibrational excitations is based on waves, they are both treated as particles in the scope of quantum mechanics. Similarly as photons, phonons have spin one. Projections $s_z = \pm 1$ of this spin correspond to the two transverse modes, projection $s_z = 0$ corresponds to the longitudinal mode. (Recall that such a longitudinal mode does not exist for the electromagnetic field which shows in Eq. (3.17) where we multiply with two for two different polarizations.) We now proceed with the Debye model and leave a more solid introduction of the phonon formalism to later courses on condensed matter and many-body problems.

Phonons have the same momentum-energy-frequency relationships as photons: $\varepsilon = c_s p$, $\varepsilon = \hbar\omega$, $p = \hbar k$ where c_s is the corresponding (transverse or longitudinal) sound velocity. State density in the ω -space for one mode is the same as for photons in Sec. 3.5, $g(\omega) = V\omega^2/(2\pi^2 c^3)$ (an additional factor of 1/2 is due to considering only one mode here). The total density from the two transverse and one longitudinal mode is

$$g(\omega) = V \left(\frac{\omega^2}{2\pi^2 c_l^3} + \frac{\omega^2}{\pi^2 c_t^3} \right) = \frac{V}{2\pi^2} \left(\frac{1}{c_l^3} + \frac{2}{c_t^3} \right) \omega^2.$$

While the total number of modes is infinite in the case of the electromagnetic field, here we are constrained by the total number of degrees of freedom in the solid which is $3N$. (We cannot create an arbitrary number of independent oscillation modes with a limited number of particles.) This implies that the density of states must be truncated at a

maximum frequency ω_c (so-called cut-off frequency). We easily determine ω_c from

$$\int_0^{\omega_c} g(\omega) d\omega = 3N \implies \omega_c^3 = 18\pi^2 n \left(\frac{1}{c_l^3} + \frac{2}{c_t^3} \right)^{-1}$$

where $n := N/V$. We can now rewrite $g(\omega)$ as

$$g(\omega) = \begin{cases} 9N \frac{\omega^2}{\omega_c^3} & \omega \leq \omega_c, \\ 0 & \omega > \omega_c. \end{cases}$$

The energy density $u(\omega)$ is now obtained by multiplying this $g(\omega)$ with the expected energy of phonons with frequency ω which is

$$\langle \varepsilon_\omega \rangle = \hbar\omega \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega} - 1} \right)$$

where $1/2$ stands for the ground state energy and $1/(e^{\beta\hbar\omega} - 1)$ is the expected number of phonons (which have chemical potential $\mu = 0$ as photons) with frequency ω . The total energy is obtained by integration of $Uu(\omega)$ over the whole frequency range. There is zero contribution from $\omega > \omega_c$ and we get

$$U(T, V, N) = \frac{9N\hbar}{\omega_c^3} \int_0^{\omega_c} \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega} - 1} \right) \omega^3 d\omega.$$

The equation for specific heat now follows as

$$C_V := \frac{\partial U}{\partial T}_{V,N} = \frac{9N\hbar^2}{\omega_c^3 k_B T^2} \int_0^{\omega_c} d\omega \frac{\omega^4 e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2} = 9Nk_B \left(\frac{k_B T}{\hbar\omega_c} \right)^3 \int_0^{x_0} dx \frac{x^4 e^x}{(e^x - 1)^2} \quad (3.46)$$

where we used substitution $x := \beta\hbar\omega$ and $x_0 = \beta\hbar\omega_c$.

The obtained specific heat is in a good agreement with experiments. To study its limit properties, we introduce the Debye function

$$D(x_0) = \frac{3}{x_0^3} \int_0^{x_0} dx \frac{x^4 e^x}{(e^x - 1)^2}$$

which allows us to write $C_V = 3Nk_B D(x_0)$. Finally, the cut-off frequency ω_c has the corresponding temperature $\Theta_D = \hbar\omega_c/k_B$ which is known as Debye temperature. The two limits of interest are now $x_0 \ll 1$ and $x_0 \gg 1$. In the former case, we use $e^x \approx 1 + x$ to get

$$D(x_0) \approx \frac{3}{x_0^3} \int_0^{x_0} dx \frac{x^4(1+x)}{x^2} = \frac{3}{x_0^3} \int_0^{x_0} dx x^2 = 1.$$

We thus see that $C_V = 3Nk_B$ in the high-temperature limit. To find the next-order correction, it is best to note that $d_x 1/(e^x - 1) = -e^x/(e^x - 1)^2$ and use per-parts integration to obtain

$$D(x_0) = \frac{3}{x_0^3} \left\{ -\frac{x^4}{e^x - 1} \Big|_0^{x_0} + \int_0^{x_0} dx \frac{4x^3}{e^x - 1} \right\} = -\frac{3x_0}{e^{x_0} - 1} + \frac{12}{x_0^3} \int_0^{x_0} dx \frac{x^3}{e^x - 1}. \quad (3.47)$$

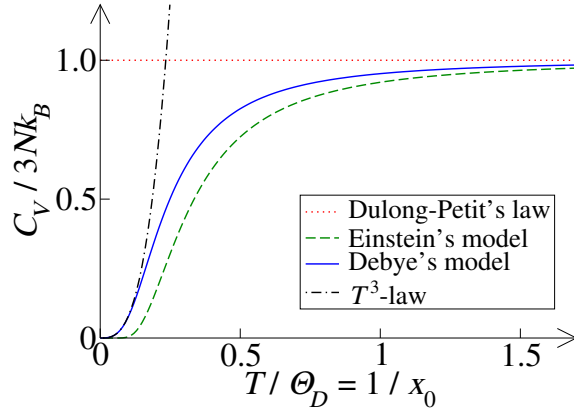


Figure 3.5: Specific heat of solids in various models.

Here $x/(e^x - 1)$ can be expanded as $1 - x/2 + x^2/12$, leading to $D(x_0) \approx 1 - x_0^2/20$. Eq. (3.47) can be used also to study the low-temperature limit $x_0 \rightarrow \infty$. The first term vanishes in this limit and the integral appearing in the second term we already know to be equal to $\pi^4/15$. We thus have

$$D(x_0) \approx \frac{4}{5} \frac{\pi^4}{x_0^3} = \frac{4}{5} \pi^4 \left(\frac{k_B T}{\hbar \omega_c} \right)^3.$$

We find that at low temperatures, C_V vanishes as T^3 —the goal has been fulfilled. In Fig. 3.5, a comparison of various results for the specific heat is shown. A direct comparison of the Einstein and Debye model is not straightforward because their results depend on different material properties: ω_0 and ω_c , respectively. Values of Θ_D for different materials are shown in Table 1.

material	Pb	Ag	Zn	Cu	Al	diamond	NaCl	KCl	MgO
Θ_D [K]	88	215	308	345	398	1850	308	233	850

Table 1: Debye temperatures of different materials obtained by fitting the experimental specific heat curves to $D(x_0)$ (taken from Greiner's book). High values of Θ_D for diamond and, partially also for MgO, imply that their specific heat is not constant at room temperature and also that they are much lower than that of metals.

A Bose-Einstein condensation in complex networks

We will discuss here a very particular case of Bose-Einstein condensation in a system where it is hardly expected: in a complex network. Complex networks build on graph theory which is a long-studied branch of mathematics. The field's foundations were laid by the Swiss mathematician Leonhard Euler to tackle the famous problem of seven bridges of Königsberg (see http://en.wikipedia.org/wiki/Seven_Bridges_of_Koenigsberg). A *graph* (which is a synonym for a network) consists of *nodes* and *edges* (links). Each link connects a pair of nodes (there are also directed graphs where each link has its start and end, *i.e.*, it points from one node to another). The *node degree* k_i is the number of edges connected to node i (see Fig. A.1 for an illustration). For any given graph, one can assess a number of characteristics. Here we will use only two of them: the *average degree* $z := \frac{1}{N} \sum_{i=1}^N k_i$ and the *degree distribution* $P(k)$ which is simply the probability that a randomly selected node has degree k . The difference between a mathematical graph and a complex network is that complex networks usually represent real systems such as a social network (where nodes represent people and links represent social relations between them) or a transportation network (where nodes represent, for example, airports and links represent regular flights between them). The study of complex networks consists of mapping and measuring real networks, proposing simple models which reproduce some of the features seen in real networks, and finally studying processes which take place on networks (for example, how a contagious disease spreads in a social network). The review article by Mark Newman (available at arxiv.org/abs/cond-mat/0303516) provides an extensive and very readable introduction to this field, covering all its various aspects. Another point of reference is provided by en.wikipedia.org/wiki/Network_science.

One of the most typical features of real complex networks is that their degree distribution is very broad. A vast majority of nodes is of a small degree but there are a few nodes whose degree exceeds the average degree by several orders of magnitude. From the mathematical point of view, the degree distribution is of the power-law form $P(k) \sim k^{-\beta}$ (at least for large degree values k ; for this reason we often speak about a power-law tail of the degree distribution). The power-law exponent β is typically between two and four.

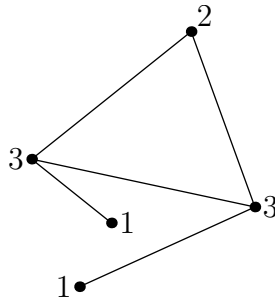


Figure A.1: A simple undirected graph where each node is labeled with its degree. This graph's degree distribution is non-zero only for $k = 1, 2, 3$ where $P(1) = 0.4$, $P(2) = 0.2$, and $P(3) = 0.4$.

The lower the exponent, the broader the distribution. When $\beta \leq 3$, standard deviation of degree diverges with the network size. When $\beta \leq 2$, even the average degree itself diverges. Broad degree distributions are not only interesting by themselves (imagine what would happen if a broad distribution would govern the heights of humans) but they often have important implications on the network's behavior. For example, a hacking attack on one highly connected server can put the whole communication network at risk, and spreading of a contagious disease can get a disastrous boost when a very social person gets infected. An early model of a complex network by Barabási and Albert (see en.wikipedia.org/wiki/Barabasi-Albert_model or Newman's paper) is famous for producing a power law degree distribution $P(k) \sim k^{-3}$. The Bose-Einstein condensation does not appear directly in this model but in its generalized version. We first briefly see the original model and its solution.

The Barabási-Albert (BA) model assumes that the network grows from an initial state (usually represented by two mutually connected nodes) in steps which we label with t . In each step, we introduce one new node and connect it with m links to the nodes that are already present in the system. The probability of selecting a node to attach to is assumed proportional to this node's degree (for this reason is this mechanism often referred to as preferential attachment). Node i with degree $k_i(t)$ is selected with the probability

$$P(i, t) = \frac{k_i(t)}{\sum_j k_j(t)} \quad (\text{A.1})$$

which is properly normalized to one at any time t . This growth mechanism causes that nodes that already have high degree attract more links than low-degree nodes and their degree thus becomes even higher, leading to a broad degree distribution (effects of this kind are often referred to as positive feedback loops). Evidence for this kind of attachment probability has been found in a wide range of systems, including the World Wide Web, citations of scientific papers, and some social networks.

The simplest way to find the degree distribution of the BA model is based on studying the average degree $\langle k_i(t) \rangle$. Assuming that the network is initiated with two connected nodes, $\sum_{j=1}^2 k_j(2) = 2$. In one time step, m new links are added, each of which increases degree of two nodes by one. The sum of degrees therefore grows by $2m$ in each step and we can write $\sum_j k_j(t) = 2m(t - 2) + 2$. The probability that node i with degree $k_i(t)$ attracts a newly forming link is therefore $k_i(t)/[2m(t - 2) + 2]$. Since there are m new links created in each time steps, degree of node i grows on average by

$$\langle \Delta k_i(t) \rangle = m \frac{k_i(t)}{2m(t - 2) + 2} \approx \frac{k_i(t)}{2t} = \frac{\langle k_i(t) \rangle}{2t} \Delta t$$

where in the last step we formally introduced $\Delta t = 1$ and replaced $k_i(t)$ with its average value. We can now pass to the differential equation

$$\begin{aligned} \frac{\langle \Delta k_i(t) \rangle}{\Delta t} &\approx \frac{d\langle k_i(t) \rangle}{dt} = \frac{\langle k_i(t) \rangle}{2t} \implies \frac{d\langle k_i(t) \rangle}{\langle k_i(t) \rangle} = \frac{dt}{2t} \implies \\ \implies \ln \langle k_i(t) \rangle &= C + \frac{1}{2} \ln t = C + \ln t^{1/2} \implies \langle k_i(t) \rangle = e^C \sqrt{t}. \end{aligned}$$

Node i appears at time τ_i with degree m , which leads us to

$$\langle k_i(t) \rangle = m\sqrt{t/\tau_i}. \quad (\text{A.2})$$

We see that the expected node degree strongly depends on the node's arrival time τ_i .

We now interpret Eq. (A.2) as transformation from variable τ to variable k . Since τ_i is simply equal to i , the distribution of τ is uniform over the range from 1 to t and thus $\varrho(\tau) \approx 1/t$. The degree distribution $P(k)$ follows immediately as

$$\varrho(\tau) d\tau = P(k) dk \implies P(k) = \varrho(\tau) \left| \frac{d\tau}{dk} \right| = \frac{1}{t} \times \frac{2tm^2}{k^3} = \frac{2m^2}{k^3}. \quad (\text{A.3})$$

We finally obtain the power-law decay $P(k) \sim k^{-3}$. It is interesting to note that despite all our approximations, $P(k)$ is roughly normalized

$$\sum_{k=m}^{\infty} P(k) \approx 2m^2 \int_m^{\infty} \frac{dk}{k^3} = 2m^2 \left[\frac{-1}{2k^2} \right]_m^{\infty} = 1.$$

A similar computation of the average degree $z = \sum_{k=m}^{\infty} kP(k)$ gives $z = 2m$ which agrees with the sum of all node degrees divided by the number of nodes, $[\sum_k k_j(t)]/t = 2m$.

The Barabási-Albert model has been generalized in many different ways. We will study the generalization where the attachment probability is driven not only by the node degree k_i but also by a so-called node fitness η_i . The idea behind this is that some nodes are simply more fit in competition with the other nodes and attract links at higher pace than the others. Eq. (A.1) generalizes to

$$P(i, t) = \frac{\eta_i k_i(t)}{\sum_j \eta_j k_j(t)} \quad (\text{A.4})$$

The rest of the model remains the same: the network grows by gradual addition of new nodes which create m new links each.

Preferential attachment with fitness can be mapped onto a Bose gas as follows. We first assign energy level i to each node where the level's energy ε_i is given by the node's fitness η_i as $\varepsilon_i = -\frac{1}{\beta} \ln \eta_i$ and $\beta = 1/T$ (since we are not going to measure “temperature” of the system in any physical units, we can omit the usual k_B here; the exact meaning of this temperature will be discussed later). Every link attached to a node is then represented as one particle at the corresponding energy level. A link between nodes i and j is therefore represented by one particle with energy ε_i and another with energy ε_j . When a new node i is introduced and creates m links to the existing nodes, this is represented by creating a new energy level ε_i populated with m particles (links from node i) and adding other m particles to the existing energy levels $\varepsilon_1, \dots, \varepsilon_m$ (selection of those levels is driven by preferential attachment). Particles stay in the originally assigned energy levels (a particle hopping from one level to another would correspond to rewiring a link). Figure A.2 illustrates this mapping on a small network. In this new representation, the probability distribution

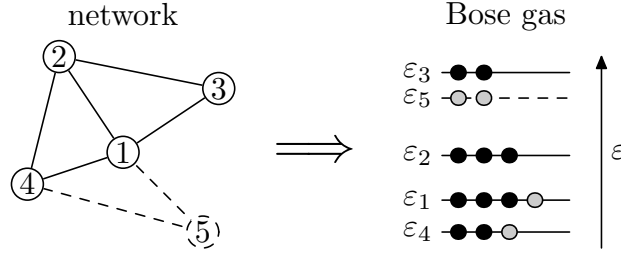


Figure A.2: Illustration of the mapping between the network model with $m = 2$ and the Bose gas. When node 5 is added in the system and links to nodes 1 and 4, energy level 5 with energy ε_5 is created and $2m$ new particles are added to levels 1, 4, and 5 (two particles for the last one).

of fitness values $\varrho(\eta)$ translates into a probability distribution of energy levels $g(\varepsilon)$ as $g(\varepsilon) = \varrho(e^{-\beta\varepsilon}) d\eta/d\varepsilon = \beta\varrho(e^{-\beta\varepsilon})e^{-\beta\varepsilon}$.

Each energy level (node) i added to the system at time τ_i with energy ε_i has its occupation number (degree) $k_i(t; \varepsilon_i, \tau_i)$ which changes with time. As before, we build on the average growth rate

$$\left\langle \frac{\partial k_i(t; \varepsilon_i, \tau_i)}{\partial t} \right\rangle = m \frac{\eta_i k_i(t; \varepsilon_i, \tau_i)}{\sum_j \eta_j k_j(t; \varepsilon_j, t_j)} = m \frac{e^{-\beta\varepsilon_i} k_i(t; \varepsilon_i, \tau_i)}{Z_t} \quad (\text{A.5})$$

where

$$Z_t = \sum_j e^{-\beta\varepsilon_j} k_j(t; \varepsilon_j, t_j)$$

is the system's partition function at time t . Its average value can be obtained by averaging over energy values ε_j under the summation. We obtain

$$\langle Z_t \rangle = \int_1^t dt_0 \int d\varepsilon g(\varepsilon) k(t; \varepsilon, t_0) e^{-\beta\varepsilon} \quad (\text{A.6})$$

where we replaced the summation over nodes with integration over their appearance times.

Motivated by the degree growth relation $\langle k_i(t) \rangle = m\sqrt{t/\tau_i}$ which holds for the original model without fitness, we now assume that the generalized growth has the form

$$\langle k_i(t; \varepsilon_i, \tau_i) \rangle = m (t/\tau_i)^{f(\varepsilon_i)}$$

where the exponent $f(\varepsilon)$ somehow depends on the level's energy (occupation of low-energy states must grow faster than that of high-energy states). Obviously, it must hold that $f(\varepsilon_i) \leq 1$ because otherwise $\langle k_i(t) \rangle$ grows with time faster than linearly and its increasing share of all $2mt$ particles present in the system eventually exceeds 1 which is physically impossible. Based on the assumed form of $\langle k_i(t) \rangle$, Eq. (A.6) can be solved and yields

$$\langle Z_t \rangle = m \int d\varepsilon g(\varepsilon) e^{-\beta\varepsilon} t^{f(\varepsilon)} \underbrace{\int_1^t dt_0 t_0^{-f(\varepsilon)}}_{(t^{1-f(\varepsilon)} - 1)/(1-f(\varepsilon))} = mt \int d\varepsilon \frac{g(\varepsilon) e^{-\beta\varepsilon}}{1-f(\varepsilon)} (1 - t^{f(\varepsilon)-1}).$$

Since $f(\varepsilon) - 1 < 0$, term $t^{f(\varepsilon)-1}$ is less and less important as the network grows and we can thus neglect it in the thermodynamic limit. The previous formula can be now written as $\langle Z_t \rangle = mt/z$ where

$$\frac{1}{z} = \int d\varepsilon \frac{g(\varepsilon)e^{-\beta\varepsilon}}{1 - f(\varepsilon)} \quad (\text{A.7})$$

is the inverse fugacity. Since $z > 0$ for any β , we can finally introduce the chemical potential μ as $z = e^{\beta\mu}$ as is usually the case in the grand-canonical ensemble. It follows that $\langle Z_t \rangle = mte^{-\beta\mu}$. Note that μ is independent of t (at least when t is large enough so that we can neglect the term $t^{f(\varepsilon)-1}$ in the previous formula for $\langle Z_t \rangle$). Eq. (A.5) can thus be solved by separation of variables and gives

$$k_i(t) = m(t/\tau_i)e^{-\beta(\varepsilon-\mu)}$$

which implies that we found that the form of $f(\varepsilon)$ (which was unknown to us until now) is $f(\varepsilon) = e^{-\beta(\varepsilon-\mu)}$. By introducing this in Eq. (A.7), we obtain

$$e^{-\beta\mu} = \int d\varepsilon g(\varepsilon) \frac{e^{-\beta\varepsilon}}{1 - e^{-\beta(\varepsilon-\mu)}} \implies 1 = \int d\varepsilon \frac{g(\varepsilon)}{e^{\beta(\varepsilon-\mu)} - 1} \quad (\text{A.8})$$

which allows us to determine the chemical potential for any β and $g(\varepsilon)$.

Eq. (A.8) is very particular because it exactly matches the equation for the average number of bosonic particles where the mean occupation number is $n(\varepsilon) = 1/(e^{\beta(\varepsilon-\mu)} - 1)$. This holds despite two important differences between our system and a usual gas of bosons. First, our “particles” choose an energy level and stay there forever as opposed to usual quantum systems where jumps to other levels can occur due to thermal fluctuations. Second, the number of particles grows in time as opposed to usual quantum systems where the expected number of particles is fixed. Third, the number of energy levels grows with time as opposed to usual quantum systems where the energy spectrum is given by geometry of the system and it is therefore fixed. Fourth, indistinguishability of particles and the implied symmetries of their quantum states play no role here. Despite all this, we effectively see the Bose-Einstein statistics in action.

In a direct analogy with the standard Bose-Einstein condensation, we now examine whether there is always μ satisfying Eq. (A.8). The condensation appears when the maximum possible result of the integration is smaller than one. Since as always for the Bose-Einstein distribution, μ cannot be greater than the ground state energy ε_0 . Since the right hand side of Eq. (A.8) grows with μ , the maximum possible value of the integral is achieved for $\mu = \varepsilon_0$. The critical temperature $T_c = 1/\beta_c$ therefore satisfies

$$\int d\varepsilon \frac{g(\varepsilon)}{e^{\beta_c(\varepsilon-\varepsilon_0)} - 1} = 1. \quad (\text{A.9})$$

As an example we take $g(\varepsilon) = 2\varepsilon$ where $\varepsilon \in [0, 1]$ and $\varepsilon_0 = 0$. Integral in the condition for β_c cannot be solved using elementary functions. A rough analytical approximation can be obtained by replacing the upper integration bound with ∞ , yielding $\beta_c^A \approx \sqrt{\zeta(2)\Gamma(3)} =$

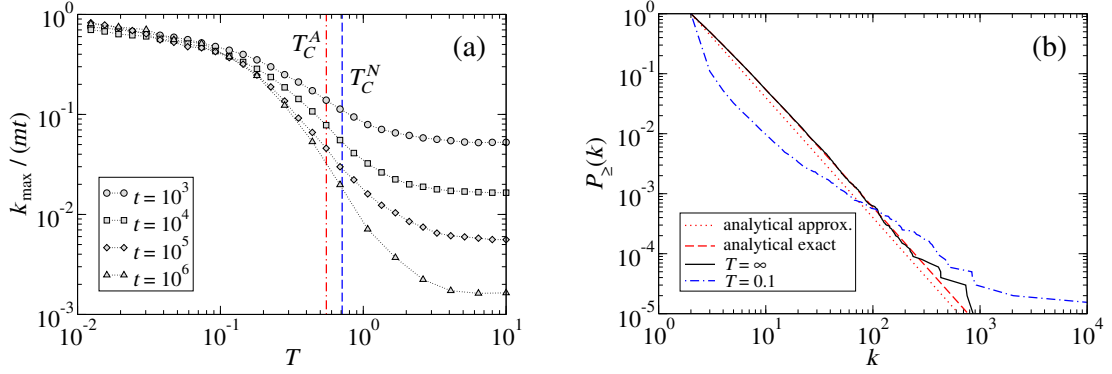


Figure A.3: The fitness model for $m = 2$ and $g(\varepsilon) = 2\varepsilon$ for $\varepsilon \in [0, 1]$. (a) Relative degree $k_{\max}/(mt)$ of the node with the highest degree as a function of temperature for various network sizes t . Theoretically predicted critical temperature values are shown with vertical lines. Results are averaged over 100 (for $t = 10^3$ and $t = 10^4$) and 10 network realizations (for $t = 10^5$ and $t = 10^6$). (b) Cumulative degree distributions in log-log scale for $t = 10^5$. When $T = \infty$ (which is equivalent to the original model without fitness), every other node has degree two (it attracted no links) and $P(k) \sim 1/k^3$ (that is, $P_{\geq}(k) \sim 1/k^2$). When $T < T_c$, a node with a macroscopic number of links appears (here is $k_{\max} \approx 150,000$) and the proportion of nodes with degree two increases to 90%.

$\sqrt{\pi^2/3} \approx 1.814$ and $T_c^A = 1/\beta_c^A \approx \sqrt{3/\pi^2} \approx 0.551$. It is also possible to find β_c satisfying Eq. (A.9) numerically, yielding $\beta_c^N \approx 1.405$ and $T_c^N = 1/\beta_c^N \approx 0.712$.

Network simulations shown in Fig. A.3 confirm that around T_c there really is a fundamental change in the system's behavior. When $T \gg T_c$, the largest occupation number (node degree) k_{\max} is small in comparison with the maximal possible value mt which is achieved when *all* new links go to one single node. When $T \ll T_c$, k_{\max} is large and, even more, it grows fast with t so that $k_{\max}/(mt)$ is roughly constant. This is the sign of condensation: even in the thermodynamic limit (which is obtained by $t \rightarrow \infty$ in our case), a macroscopic fraction of all particles (links) end in the same state (node). We close by a discussion of what is the meaning of temperature in our system. As we saw at the very beginning, temperature here maps node fitness values η into state energy values ε by $\eta = e^{-\beta\varepsilon}$. At the end we assumed a fixed density of the energy values $g(\varepsilon)$ and temperature therefore determines how are the fitness values distributed. When T is large, β and $\beta\varepsilon$ are small, leading to all fitness values η being close to one. In such a case, as we already know, no condensation occurs and it is not the fit nodes but the early introduced nodes that receive the largest number of links. In the original Barabási-Albert model (which is recovered when $T \rightarrow \infty$) is $k_{\max} \sim \sqrt{t}$ and therefore $k_{\max}/(mt) \sim 1/\sqrt{t} \rightarrow 0$. When T is small, β is large and the fitness values are therefore widely distributed. One of the early knows that at the same time has a particularly high fitness value then becomes a “condensation” core. There can be nodes of higher fitness introduced later in the network's evolution but they still cannot overcome the time advantage of that early-and-fit node which in consequence attracts a substantial fraction of all future links.

B Physics of white dwarfs

This section is based on examples 14.4 and 14.5 in the book “Thermodynamics and statistical mechanics” by Greiner, Neise, and Stöcker. We will study and use here a relativistic Fermi gas at zero temperature. While this may sound as a contradiction, we will examine a situation where this limit is relevant.

B.1 Relativistic Fermi gas

We want to derive the thermodynamic property of a relativistic Fermi gas at temperature $T = 0$. As usual, we start with the logarithm of the grand partition function $\ln Z = \sum_k \ln(1 + z \exp[-\beta \varepsilon_k])$. The one-particle energies ε_k in the momentum state k are now taken to be relativistic, and we subtract the rest mass of the particles to yield an easier check of the non-relativistic limit,

$$\varepsilon = mc^2 \left(\sqrt{1 + \left(\frac{p}{mc} \right)^2} - 1 \right). \quad (\text{B.1})$$

Next we write the sum over all one-particle states as an integral over all momenta

$$\ln Z = g \frac{V}{h^3} \int_0^\infty dp 4\pi p^2 \ln(1 + z \exp[-\beta \varepsilon]). \quad (\text{B.2})$$

The logarithm can be removed by integration by parts,

$$\ln Z = \frac{4\pi g V}{h^3} \frac{\beta}{3} \int_0^\infty dp p^3 \frac{d\varepsilon}{dp} \frac{1}{z^{-1} \exp[\beta \varepsilon] + 1}. \quad (\text{B.3})$$

The fugacity $z = \exp[\beta \mu]$ is to be determined for a given particle number from

$$N(T, V, z) = \sum_k \langle n_k \rangle = \frac{4\pi g V}{h^3} \int_0^\infty dp p^2 \frac{1}{z^{-1} \exp[\beta \varepsilon] + 1}. \quad (\text{B.4})$$

We now want to consider the case $T = 0$ K. Then the mean occupation number has a simple step-function form, $\langle n_k \rangle = \Theta(\varepsilon_f - \varepsilon)$ (as we already discussed in Sec. 3.6, only the Fermi sphere states are occupied at zero temperature), and thus it holds that

$$\ln Z = \frac{4\pi \beta g V}{3h^3} \int_0^{p_f} dp p^3 \frac{d\varepsilon}{dp}, \quad (\text{B.5})$$

$$N(T, V, z) = \frac{4\pi g V}{h^3} \int_0^{p_f} p^2 dp = \frac{4\pi g V}{3h^3} p_f^3. \quad (\text{B.6})$$

Here p_f is the momentum corresponding to the Fermi energy ε_f according to Eq. (B.1), which can be determined from the particle density with the help of Eq. (B.6) as

$$p_f = \left(\frac{3}{4\pi} \frac{N h^3}{V g} \right)^{1/3}. \quad (\text{B.7})$$

Thus p_f is proportional to $n^{1/3}$ ($n = N/V$). As before, the chemical potential is equal to the Fermi energy and thus $\mu = \varepsilon_f = \varepsilon(p_f)$. We now compute $d\varepsilon/dp$ from Eq. (B.1)

$$\frac{d\varepsilon}{dp} = c \frac{\frac{p}{mc}}{\left[1 + \left(\frac{p}{mc}\right)^2\right]^{1/2}}. \quad (\text{B.8})$$

If we insert this into Equation (B.5) and consider $\ln Z = PV/(k_B T)$, it follows that (we write here P for the pressure to avoid confusion with the momentum)

$$P = \frac{4\pi g}{3h^3} \int_0^{p_f} mc^2 \frac{\frac{p}{mc}}{\left[1 + \left(\frac{p}{mc}\right)^2\right]^{1/2}} p^2 dp. \quad (\text{B.9})$$

Before evaluating this integral, we also want to present an explicit expression for the energy. This is simpler than differentiating the function afterwards

$$U = \sum_k \varepsilon_k \langle n_k \rangle = \frac{4\pi g V}{h^3} \int_0^{p_f} dp p^2 mc^2 \left(\left[1 + \left(\frac{p}{mc}\right)^2\right]^{1/2} - 1 \right) \quad \text{for } T = 0 \text{ K}. \quad (\text{B.10})$$

To calculate the integrals (B.9) and (B.10), we make a substitution which is often useful in the relativistic case: $p = mc \sinh x$. Then we have $\varepsilon = mc^2(\cosh x - 1)$ and $d\varepsilon/dp = c \tanh x$. If we finally set $p_f = mc \sinh x_f$, the following integrals have to be performed

$$P = \frac{4\pi g m^4 c^5}{3h^3} \int_0^{x_f} \sinh^4 x \, dx, \quad (\text{B.11})$$

$$U = \frac{4\pi g V m^4 c^5}{h^3} \int_0^{x_f} (\cosh x - 1) \sinh^2 x \cosh x \, dx. \quad (\text{B.12})$$

These integrals are quite simple if we use the properties of hyperbolic functions:

$$\begin{aligned} \cosh^2 x - \sinh^2 x &= 1, & \cosh 2x &= \cosh^2 x + \sinh^2 x, & \sinh 2x &= 2 \cosh x \sinh x, \\ \frac{d}{dx} \sinh x &= \cosh x, & \frac{d}{dx} \cosh x &= \sinh x. \end{aligned}$$

For example, we have

$$\begin{aligned} \sinh^4 x &= \sinh^2 x (\cosh^2 x - 1) = \frac{1}{4} \sinh^2 2x - \sinh^2 x = \\ &= \frac{1}{8} \cosh 4x - \frac{1}{2} \cosh 2x + \frac{3}{8} = \frac{d}{dx} \left(\frac{1}{32} \sinh 4x - \frac{1}{4} \sinh 2x + \frac{3}{8} x \right), \end{aligned}$$

which allows us to directly integrate Eq. (B.11). Analogously, one finds

$$\begin{aligned} (\cosh x - 1) \sinh^2 x \cosh x &= \frac{1}{4} \sinh^2 2x - \sinh^2 x \cosh x = \\ &= \frac{1}{8} \cosh 4x - \frac{1}{8} - \sinh^2 x \frac{d}{dx} \sinh x = \\ &= \frac{d}{dx} \left(\frac{1}{32} \sinh 4x - \frac{x}{8} - \frac{1}{3} \sinh^3 x \right), \end{aligned}$$

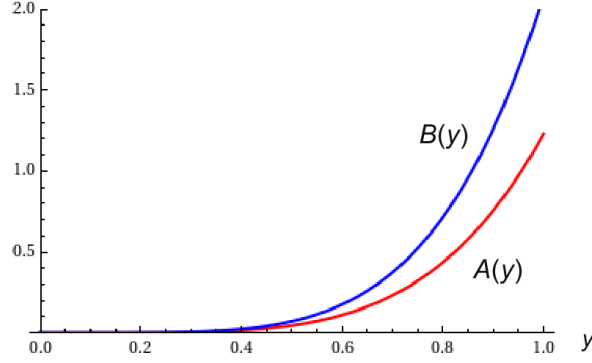


Figure B.1: Plot of $A(y)$ and $B(y)$.

which solves (B.12). One can again transform the results of the integration (hyperbolic functions of multiple of the argument) into hyperbolic functions of the argument alone

$$\int_0^{x_f} \sinh^4 x \, dx = \frac{1}{8} (3x_f - 3 \sinh x_f \cosh x_f + 2 \sinh^3 x_f \cosh x_f), \quad (\text{B.13})$$

$$\int_0^{x_f} (\cosh x - 1) \sinh^2 x \cosh x \, dx = \frac{1}{8} (-x_f + \sinh x_f \cosh x_f + 2 \sinh^3 x_f \cosh x_f - \frac{8}{3} \sinh^3 x_f) \quad (\text{B.14})$$

For sake of abbreviation, we introduce $A(y)$ and $B(y)$ (see Fig. B.1)

$$A(y) = \sqrt{1+y^2} (2y^3 - 3y) + 3 \operatorname{arcsinh} y, \quad (\text{B.15})$$

$$B(y) = 8y^3 (\sqrt{1+y^2} - 1) - A(y). \quad (\text{B.16})$$

where $y = \sinh x = p/(mc)$.

Equations (B.13) and (B.14) can be conveniently expressed in the form

$$\int_0^{x_f} \sinh^4 x \, dx = \frac{1}{8} A(y_f),$$

$$\int_0^{x_f} (\cosh x - 1) \sinh^2 x \cosh x \, dx = \frac{1}{24} B(y_f)$$

We thus obtain

$$P = \frac{g\pi m^4 c^5}{6h^3} A(y_f) \quad (\text{B.17})$$

$$U = \frac{g\pi V m^4 c^5}{6h^3} B(y_f) \quad (\text{B.18})$$

with the dimensionless quantity $y_f = \sinh x_f = p_f/(mc)$.

It is useful to investigate the non-relativistic limit $y_f \ll 1$ and the ultra-relativistic limit $y_f \gg 1$. To this end, we need simply the respective expansions of $A(y)$ and $B(y)$. With $\operatorname{arcsinh} y = \ln(y + \sqrt{1 + y^2})$ one obtains

$$A(y) \approx \frac{8}{5}y^5 - \frac{4}{7}y^7 + \dots \quad y \ll 1, \quad (\text{B.19})$$

$$B(y) \approx \frac{12}{5}y^5 - \frac{3}{7}y^7 + \dots \quad y \ll 1. \quad (\text{B.20})$$

Also, for large arguments one may write down an expansion if one writes

$$A(y) = 2y^4\sqrt{1 + y^{-2}} - 3y^2\sqrt{1 + y^{-2}} + 3\ln\left\{y\left(1 + \sqrt{1 + y^{-2}}\right)\right\}$$

and takes y^{-2} to be a small quantity,

$$A(y) \approx 2y^4 - 2y^2 + 3\ln 2y - \frac{7}{4} + \frac{5}{4}y^{-2} + \dots \quad y \gg 1, \quad (\text{B.21})$$

$$B(y) \approx 6y^4 - 8y^3 + 6y^2 - 3\ln 2y + \frac{3}{4} - \frac{3}{4}y^{-2} + \dots \quad y \gg 1. \quad (\text{B.22})$$

In the non-relativistic case $y_f \ll 1$, if we restrict ourselves to the lowest order approximation in Equations (B.19) and (B.20), we obtain from Equations (B.17) and (B.18) just the results for the ideal Fermi gas:

$$P = \frac{g\pi m^4 c^5}{6h^3} \frac{8}{5} \left(\frac{p_f}{mc}\right)^5 = \frac{2}{3} \left(\frac{2g\pi}{5h^3 m} p_f^5\right) = \frac{2}{3} \frac{U}{V} \quad (\text{B.23})$$

since

$$U = \frac{g\pi V m^4 c^5}{6h^3} \frac{12}{5} \left(\frac{p_f}{mc}\right)^5 = \frac{2g\pi V}{5h^3 m} p_f^5 \quad (\text{B.24})$$

In the ultra-relativistic case $y_f \gg 1$, we get

$$P = \frac{g\pi m^4 c^5}{6h^3} 2 \left(\frac{p_f}{mc}\right)^4 = \frac{1}{3} \left(\frac{g\pi c}{h^3} p_f^4\right) = \frac{1}{3} \frac{U}{V} \quad (\text{B.25})$$

since

$$U = \frac{g\pi V m^4 c^5}{6h^3} 6 \left(\frac{p_f}{mc}\right)^4 = \frac{g\pi V c}{h^3} p_f^4 \quad (\text{B.26})$$

which will be used later.

B.2 White Dwarfs

The main application of the cold relativistic Fermi gas arises in astrophysics. In the 1930 Chandrasekhar found that stars which are called *white dwarfs* are very well described in the framework of this model system for electrons. In the following example we shall to investigate this in more detail.

Consider the following model of a white dwarf: a gas sphere, consisting of helium, of mass $M \approx 10^{30}$ kg, at a density $\rho = 10^{10}$ kg/m³ and (central) temperature $T = 10^7$ K (at these temperatures the helium atoms are nearly completely ionized). Our work plan is

1. Show that despite the large temperature the electron gas can be considered to be very cold ($k_B T \ll \varepsilon_f$), but that on the other hand relativistic effects become important.
2. Show that the main contribution to the total pressure originates from free electrons, while the contribution of the helium nuclei is very small. Show in particular that the helium nuclei may be considered as a classical ideal gas.
3. Calculate the equilibrium radius of a white dwarf under the assumption that in equilibrium the pressure of the electrons just balances the gravitational pressure. Neglect here the radial change of density and pressure. Show that there is a relationship between the mass of the star and its radius, and investigate the limits of this relation.
4. What happens, if the mass of the star is 10–20 times larger?

We first estimate the Fermi energy and the Fermi momentum of the electrons in the gas with the help of the mass density ρ . Each ionized helium atom contributes two electrons and four nucleons to the total mass. The helium nuclei may be treated non-relativistically, since their mean kinetic energy, due to the thermal energy $k_B T \approx 1 \text{ keV}$, is very small compared to their rest mass $m_{\text{He}} c^2 \approx 4 \text{ GeV}$. Also, for the electrons, the contribution of the kinetic energy to the total mass is still rather small ($m_e c^2 \approx 511 \text{ keV}$), so that we may write (if N denotes the number of electrons in the state),

$$M \approx N(m_e + 2m_n) \approx 2m_n N \quad (\text{B.27})$$

since $m_e \ll m_n$ and two nucleon masses belong to one electron. Therewith, the particle density of the electrons in the star may be estimated as

$$n = \frac{N}{V} \approx \frac{M/2m_n}{M/\rho} \approx \frac{\rho}{2m_n} \approx 3 \cdot 10^{-9} \frac{\text{electrons}}{\text{fm}^3} \quad (\text{B.28})$$

($1 \text{ fm} = 10^{-15} \text{ m}$). From this density we calculate for the Fermi momentum of the electrons, according to Equation (B.7)

$$p_f = h \left(\frac{3n}{4\pi g} \right)^{1/3} \approx 5 \cdot 10^{-22} \frac{\text{kg m}}{\text{s}} \approx 0.9 \frac{\text{MeV}}{c} \quad (\text{B.29})$$

If we insert this into the relativistic energy-momentum relation, we obtain for the (kinetic) Fermi energy, without the rest mass of the electrons, $\varepsilon_f \approx 0.5 \text{ MeV}$. Thus, relativistic effects become important, but because $k_B T \approx 1 \text{ keV} \ll \varepsilon_f$, the electron gas may be considered as cold.

Second, for the helium nuclei we calculate the parameter $n\lambda^3$, which tells us whether quantum effects are important:

$$\lambda_{\text{He}} = \left(\frac{h^2}{2\pi m k_B T} \right)^{1/2} \approx 247 \text{ fm}. \quad (\text{B.30})$$

Since the particle number density of the helium nuclei is half that of the electrons, it follows that $n\lambda_{\text{He}}^3 \approx 2.27 \cdot 10^{-2}$, which is small compared to 1. Thus, we may apply the Boltzmann limit. The helium nuclei therefore contribute a pressure

$$P_{\text{He}} = n_{\text{He}} k_B T \approx 1.5 \cdot 10^{-12} \frac{\text{MeV}}{\text{fm}^3}. \quad (\text{B.31})$$

On the other hand, according to Equation (B.17), we find for the Fermi pressure of the electrons with $y_f = p_f/(mc) \approx 2$ and with the function $A(y)$ introduced in the example (which we here crudely approximate by its expansion $A(y_f) \approx A(2) \approx 26.7$ for $y_f \gg 1$):

$$P_e = \frac{1}{24\pi^2} \frac{(mc^2)^4}{(\hbar c)^3} A(y_f) \approx 10^{-9} \frac{\text{MeV}}{\text{fm}^3} \quad (\text{B.32})$$

which is 1000-times larger than the pressure of the helium nuclei.

Up to now, we have assumed the gas is enclosed by a box. This is of course not the case, but gravity prevents the gas from leaking out. If the gas expands by a volume dV while the gas sphere is enlarged by dR , the energy

$$dE_P = -P dV = -P(R) 4\pi R^2 dR \quad (\text{B.33})$$

will be gained. However, the pressure is a function of the Fermi momentum (see Equation (B.17)), and the latter depends in turn on the volume or radius (at given particle number). On the other hand, when the sphere is enlarged the potential energy increases by the amount

$$dE_g = \frac{dE_g(R)}{dR} dR = \alpha \frac{GM^2}{R^2} dR \quad \text{where } E_g(R) = -\alpha GM^2/R. \quad (\text{B.34})$$

The additional factor α stands for eventual correlations which result from an inhomogeneous density distribution. However, in the case considered here it is on the order of 1. In thermodynamic equilibrium, the free energy has to have a minimum; *i.e.*, $dF = 0$. Since we consider the system at $T = 0$ K, it holds that $F = E - TS = E$ and thus

$$dF = dE_g + dE_P = 0 = \alpha \frac{GM^2}{R^2} - P(R) 4\pi R^2 \quad (\text{B.35})$$

or

$$P(R) = \frac{\alpha}{4\pi} \frac{GM^2}{R^4} \quad (\text{B.36})$$

A relationship between the mass and the radius of the star follows from this equation. If we insert Equation (B.17) for the pressure and Equation (B.7) for the Fermi momentum, Equation (B.36) becomes

$$A \left[\left(\frac{9\pi M}{8m_n} \right)^{1/3} \frac{\hbar c}{R m_e c^2} \right] = 6\pi\alpha \left(\frac{\hbar c}{R m_e c^2} \right)^3 \frac{1}{m_e c^2} \frac{GM^2}{R}. \quad (\text{B.37})$$

The units in this equation for $R(M)$ are remarkable: the mass M of the star is measured in units of the nucleon mass m_n , and the radius R of the star in units of the Compton

wavelength of the electrons ($\hbar c/m_e c^2$). Finally, the gravitational energy GM^2/R appears on the right hand side, measured in units of the electron mass $m_e c^2$. Equation (B.37) thus relates quantum mechanics, special relativity and classical gravitational theory.

Unfortunately, Eq. (B.37) can be solved neither for $R(M)$ nor for $M(R)$. However, the limiting cases, where the argument of the function becomes very large or very small, are analytically solvable. First we note that with $M \approx 10^{30}$ kg, $m_n \approx 1.6 \cdot 10^{-27}$ kg, $\hbar c = 179$ MeV fm, and $m_e c^2 \approx 0.5$ MeV, the argument of the function A is 1 if $R \approx 5 \cdot 10^6$ m. Therefore, for small arguments $y_f = p_f/mc \ll 1$ we have $R \gg 10^6$ m, and Equation (B.37) becomes, with $A(y) \approx \frac{8}{5}y^5$,

$$R \approx \frac{3(9\pi)^{2/3}}{40\alpha} \frac{\hbar^2}{Gm_n^{5/3}m_e} M^{-1/3}. \quad (\text{B.38})$$

In the cases of a large argument $y \gg 1$ (or $R \ll 10^6$ m), $A(y) \approx 2y^4 - 2y^2$, and one obtains

$$R \approx \frac{(9\pi)^{1/3}}{2} \frac{\hbar c}{m_e c^2} \left(\frac{M}{m_n}\right)^{1/3} \left[1 - \left(\frac{M}{M_0}\right)^{2/3}\right]^{1/2} \quad (\text{B.39})$$

where we used the abbreviation

$$M_0 = \frac{9}{64} \left(\frac{3\pi}{\alpha^3}\right)^{1/2} \left(\frac{\hbar c}{Gm_n^2}\right)^{3/2} m_n \quad (\text{B.40})$$

and also took into account the second order terms in $A(y)$. For $M_0 < M$ there are no real solutions for R . As one observes, the radius of the star approaches zero if M converges towards the (finite) mass M_0 . Consequently, there are no white dwarfs with a mass larger than M_0 . One calls M_0 the *Chandrasekhar limit*. Obviously, the Fermi pressure can no longer compensate the gravitational pressure of the star for $M_0 > M$, and the collapse of the star is the consequence. However, for mass M in the vicinity of M_0 , which yields a very small white dwarf, strong effects originating from general relativity occur. These can no longer be neglected if the radius of the star is of the order of the Schwarzschild radius $R_s = 2GM/c^2$.

The considerations presented here were worked out in detail by Chandrasekhar in the years 1931–1935. The first investigations concerning this subject were done by Fowler (1926), who realized that the Fermi gas in a white dwarf is completely degenerate (in analogy to the electron gas in metals), while the perception that a relativistic treatment is in order originates from Anderson (1929) and Stoner (1929–1930). The numerical value of the Chandrasekhar mass above is $M_0 \approx 1.6 \cdot 10^{30}$ kg/ $\alpha^{3/2}$ where, as already said before, $\alpha \approx 1$. Detailed studies of Chandrasekhar yielded

$$M_0 = \frac{5.75}{\mu_e^2} M_{\text{sun}} \quad (\text{B.41})$$

where μ_e^2 takes care of the ionization degree of the He atom (the number of free electrons per He nucleus). This is, to a good approximation, $\mu_e \approx 2$ from which $M_0 \approx 1.44 M_{\text{sun}}$ follows. We now see that our sun will end its life as a white dwarf.

C Path integrals

Quantum mechanics, as formulated by Schrödinger, Heisenberg, Dirac and others between 1923 and 1927, is based on the concept of observables, represented by Hermitean operators, and states, represented by vectors of a Hilbert space. An alternative formulation was developed by Feynman in his doctoral thesis around 1945, following ideas of Dirac. His approach is based on amplitudes associated with possible motions (or paths), which have to be summed up (integrated) to obtain the probability of a certain outcome. This is, as Feynman himself points out, not a new theory, but rather a new formulation of quantum mechanics, which is useful for improving our conceptual understanding and, sometimes, for actual calculations. More recently, this formalism has become very popular, as it allows to pass readily from quantum field theory to statistical mechanics. Therefore methods developed for the former can be applied to the latter and *vice versa*. Before embarking upon the formalism of path integrals, it is useful to touch briefly the closely related problem of Brownian motion.

C.1 Brownian motion

We consider a random walk on a square lattice with lattice constant a and primitive vectors $\mathbf{a}_1, \mathbf{a}_2$. A particle jumps at regular time steps ($t = n \Delta t$, $n = 1, 2, 3, \dots$) from the site it occupies at that moment to one of the four neighboring sites without preference (*i.e.*, each of the possible neighboring sites has equal probability $1/4$ to be chosen). Successive moves are uncorrelated. Fig. C.1 shows the result of such a random walk.

The key quantity is the probability $P(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0)$ of finding the particle at site \mathbf{r}_1 at time t_1 if at time t_0 it was at the site \mathbf{r}_0 . The initial condition is straightforward

$$P(\mathbf{r}_1, t_0; \mathbf{r}_0, t_0) = \delta_{\mathbf{r}_1, \mathbf{r}_0}. \quad (\text{C.1})$$

The particle has to be somewhere on the lattice at time t_1 , therefore the normalization condition

$$\sum_{\mathbf{r}_1} P(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0) = 1 \quad (\text{C.2})$$

has to be fulfilled at any t_1 . Assuming that $P(\mathbf{r}, t; \mathbf{r}_0, t_0)$ is known for all sites \mathbf{r} at time t , we can obtain the function for the next time step by the recursive relation

$$P(\mathbf{r}, t + \Delta t; \mathbf{r}_0, t_0) = \frac{1}{4} \sum_{\substack{\text{neighbors} \\ \mathbf{r}' \text{ of } \mathbf{r}}} P(\mathbf{r}', t; \mathbf{r}_0, t_0) \quad (\text{C.3})$$

where the four possible neighboring sites are $\mathbf{r}' = \mathbf{r} \pm \mathbf{a}_i$ with $i = 1, 2$. This equation can be rewritten as

$$\begin{aligned} & P(\mathbf{r}, t + \Delta t; \mathbf{r}_0, t_0) - P(\mathbf{r}, t; \mathbf{r}_0, t_0) = \\ &= \frac{1}{4} \sum_{\alpha=1}^2 [P(\mathbf{r} + \mathbf{a}_\alpha, t; \mathbf{r}_0, t_0) - 2P(\mathbf{r}, t; \mathbf{r}_0, t_0) + P(\mathbf{r} - \mathbf{a}_\alpha, t; \mathbf{r}_0, t_0)] . \end{aligned} \quad (\text{C.4})$$

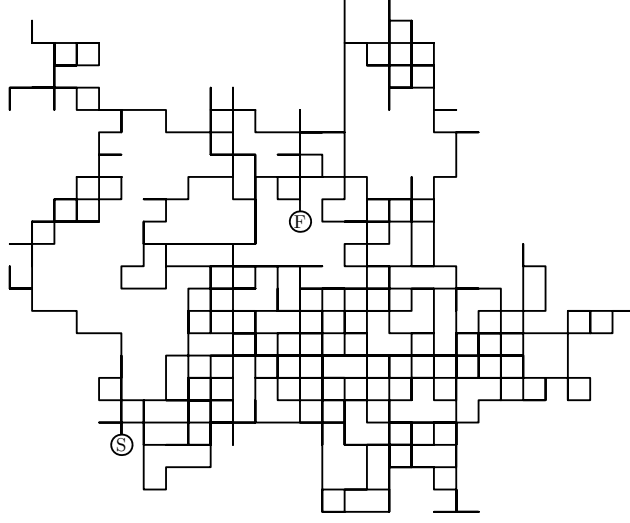


Figure C.1: Random walk on a square lattice for 1000 time steps. S and F mark the start and finish site, respectively.

where we simply wrote the summation over neighbors explicitly and added $-P(\mathbf{r}, t; \mathbf{r}_0, t_0)$ to both sides of the equation.

The probability does not change by applying a translation in space or in time, therefore we can assume

$$P(\mathbf{r}, t; \mathbf{r}_0, t_0) = P(\mathbf{r} - \mathbf{r}_0, t - t_0). \quad (\text{C.5})$$

In the following we will replace $\mathbf{r} - \mathbf{r}_0$ by \mathbf{r} and $t - t_0$ by t . We now let both the lattice constant a and the time interval Δt tend to zero. It is then convenient to introduce the probability density

$$p(\mathbf{r}, t) := \frac{P(\mathbf{r}, t)}{a^2}, \quad (\text{C.6})$$

which is normalized because of Eq. (C.2) it holds that

$$\int p(\mathbf{r}, t) d^2\mathbf{r} = \lim_{a \rightarrow 0} \sum_{\mathbf{r}} a^2 p(\mathbf{r}, t) = 1. \quad (\text{C.7})$$

Dividing Eq. (C.1) by Δt , we get

$$\frac{p(\mathbf{r}, t + \Delta t) - p(\mathbf{r}, t)}{\Delta t} = \frac{a^2}{4\Delta t} \sum_{\alpha=1}^2 \frac{p(\mathbf{r} + \mathbf{a}_\alpha, t) - 2p(\mathbf{r}, t) + p(\mathbf{r} - \mathbf{a}_\alpha, t)}{a^2}. \quad (\text{C.8})$$

We now take the limits $a \rightarrow 0$, $\Delta t \rightarrow 0$ in such a way that the prefactor remains a finite constant

$$\frac{a^2}{4\Delta t} = D. \quad (\text{C.9})$$

The left hand side of Eq. (C.8) tends to the first derivative with respect to time while the two terms of the right hand side tend to second spatial derivatives. In this way we arrive at the diffusion equation

$$\frac{\partial}{\partial t} p(\mathbf{r}, t) = D \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) p(\mathbf{r}, t). \quad (\text{C.10})$$

In this limit, the initial condition (C.1) becomes $p(\mathbf{r}, 0) = \delta(\mathbf{r})$. To solve Eq. (C.10) in terms of the probability density $p(\mathbf{r}, t)$, we go over to the Fourier transform

$$p(\mathbf{r}, t) = \int \frac{d^2 \mathbf{q}}{(2\pi)^2} \tilde{p}(\mathbf{q}, t) e^{i\mathbf{q} \cdot \mathbf{r}}. \quad (\text{C.11})$$

It is easy to verify that the Fourier coefficients $\tilde{p}(\mathbf{q}, t)$ satisfy

$$\frac{\partial}{\partial t} \tilde{p}(\mathbf{q}, t) = -D(q_x^2 + q_y^2) \tilde{p}(\mathbf{q}, t) \quad (\text{C.12})$$

with the initial condition $\tilde{p}(\mathbf{q}, 0) = 1$. The solution is

$$\tilde{p}(\mathbf{q}, t) = e^{-Dq^2 t}. \quad (\text{C.13})$$

The Fourier integral (C.11) is then readily evaluated,

$$p(\mathbf{r}, t) = \int_{-\infty}^{\infty} \frac{dq_x}{2\pi} e^{i(q_x x - Dq_x^2 t)} \int_{-\infty}^{\infty} \frac{dq_y}{2\pi} e^{i(q_y y - Dq_y^2 t)} = \frac{1}{4\pi Dt} e^{-\frac{r^2}{4Dt}}. \quad (\text{C.14})$$

The procedure outlined above for the special case of two dimensions can be easily extended to other dimensions (and also to other lattices). Using the original set-up with an initial position \mathbf{r}_0 at time t_0 and a final position \mathbf{r}_1 at time t_1 , one obtains the probability density in the continuum limit ($t_1 > t_0$) of a d -dimensional lattice

$$P(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0) = \frac{1}{[4\pi D(t_1 - t_0)]^{d/2}} e^{-\frac{|\mathbf{r}_1 - \mathbf{r}_0|^2}{4D(t_1 - t_0)}}. \quad (\text{C.15})$$

This solution satisfies the equation

$$P(\mathbf{r}_2, t_2; \mathbf{r}_0, t_0) = \int d^d \mathbf{r} P(\mathbf{r}_2, t_2; \mathbf{r}, t) P(\mathbf{r}, t; \mathbf{r}_0, t_0) \quad (\text{C.16})$$

for $t_0 < t < t_2$. Eq. (C.16) expresses the fact that at the intermediate time t the particle must be somewhere in space. It can be explicitly checked for the solution (C.15). Next, we divide the time interval $t_n - t_0$ into n time slices. Generalizing Eq. (C.16) to this case we obtain

$$P(\mathbf{r}_n, t_n; \mathbf{r}_0, t_0) = \int d^d \mathbf{r}_1 \cdots \int d^d \mathbf{r}_{n-1} P(\mathbf{r}_n, t_n; \mathbf{r}_{n-1}, t_{n-1}) \cdots P(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0) \quad (\text{C.17})$$

or, using the explicit form (C.15),

$$P(\mathbf{r}_n, t_n; \mathbf{r}_0, t_0) = \frac{1}{[4\pi D(t_1 - t_0)]^{d/2}} \left(\prod_{j=1}^{n-1} \frac{\int d^d \mathbf{r}_j}{[4\pi D(t_{j+1} - t_j)]^{d/2}} \right) \exp \left[- \sum_{j=1}^n \frac{|\mathbf{r}_j - \mathbf{r}_{j-1}|^2}{4D(t_j - t_{j-1})} \right].$$

In the limit $n \rightarrow \infty$, this result is usually presented by the symbolic formula

$$P(\mathbf{r}_f, t_f; \mathbf{r}_i, t_i) = \int_{\substack{\mathbf{r}(t_i) = \mathbf{r}_i \\ \mathbf{r}(t_f) = \mathbf{r}_f}} \mathcal{D}\mathbf{r}(t) \exp \left[- \frac{1}{4D} \int_{t_i}^{t_f} dt \dot{\mathbf{r}}^2 \right]. \quad (\text{C.18})$$

This is a sum over all possible paths from the initial position \mathbf{r}_i at time t_i to the final position \mathbf{r}_f at time t_f , with a certain weight associated with each path (a so-called Wiener measure). We see that the probability function $P(\mathbf{r}, t; \mathbf{r}_0, t_0)$ can be considered either as a solution of the diffusion equation (C.10), corresponding to the original recursion relation (C.3), or as Wiener's path integral (C.18). We will discover a similar relation between the Schrödinger equation and Feynman's path integral. Here we merely remark that the diffusion equation

$$\frac{\partial}{\partial t} p(\mathbf{r}, t) = D \nabla^2 p(\mathbf{r}, t) \quad (\text{C.19})$$

is transformed into the Schrödinger equation by replacing: dt by $i dt$, $p(\mathbf{r}, t)$ by $\psi(\mathbf{r}, t)$ and D by $\hbar/(2m)$. The exponent in Eq. (C.18) then becomes $-i S[\mathbf{r}(t)]/\hbar$ where

$$S[\mathbf{r}(t)] = \int_{t_i}^{t_f} dt \frac{m \dot{\mathbf{r}}^2}{2} \quad (\text{C.20})$$

is the action integral for the path $\mathbf{r}(t)$ with $\mathbf{r}(t_i) = \mathbf{r}_i$ and $\mathbf{r}(t_f) = \mathbf{r}_f$ (recall that action S is defined as integral of $T - U$ over a path $\mathbf{r}(t)$ and $U(\mathbf{r}) = 0$ in our case).

C.2 Path integral formulation of quantum mechanics

The theory of Brownian motion allows us to calculate the probability density $p(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ of finding a particle at a position \mathbf{r}_2 at time t_2 if it has started its random walk at a position \mathbf{r}_1 at time t_1 . In quantum mechanics the dynamics is deterministic, *i.e.*, the state vector $|\psi(t_2)\rangle$ is well defined if we know the state vector $|\psi(t_1)\rangle$,

$$|\psi(t_2)\rangle = \hat{U}(t_2, t_1) |\psi(t_1)\rangle. \quad (\text{C.21})$$

The time evolution operator $\hat{U}(t_2, t_1)$ satisfies the equation of motion

$$i\hbar \frac{\partial}{\partial t_2} \hat{U}(t_2, t_1) = \hat{H} \hat{U}(t_2, t_1) \quad (\text{C.22})$$

where \hat{H} is the Hamiltonian of the system. The probabilistic aspect of quantum mechanics appears if we want to extract the value of a specific observable from a given quantum state.

Thus for a particle described by the state vector $|\psi(t_2)\rangle$ at time t_2 the probability density of finding it at \mathbf{r}_2 is given by $|\langle\mathbf{r}_2|\psi(t_2)\rangle|^2$. In the particular case where the particle was (for sure) at the point \mathbf{r}_1 (*i.e.*, in the state $|\mathbf{r}_1\rangle$ at time t_1), we can use Eq. (C.21) to get

$$|\langle\mathbf{r}_2|\psi(t_2)\rangle|^2 = |\langle\mathbf{r}_2|\hat{U}(t_2, t_1)|\mathbf{r}_1\rangle|^2. \quad (\text{C.23})$$

Therefore the important quantity is the amplitude (or the propagator)

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) := \langle\mathbf{r}_2|\hat{U}(t_2, t_1)|\mathbf{r}_1\rangle. \quad (\text{C.24})$$

It can be interpreted as the wave function of a particle that was at a position \mathbf{r}_1 at time t_1 . Therefore in coordinate representation the evolution equation (C.22) of a free particle with Hamiltonian $\hat{H} = \hat{\mathbf{p}}^2/(2m)$ is

$$i\hbar \frac{\partial}{\partial t_2} K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = -\frac{\hbar^2}{2m} \nabla_2^2 K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) \quad (\text{C.25})$$

where the subscript 2 at ∇ reminds that the spatial derivative are with respect to \mathbf{r}_2 . The initial condition is given by

$$K(\mathbf{r}_2, t_1; \mathbf{r}_1, t_1) = \langle\mathbf{r}_2|\mathbf{r}_1\rangle = \delta(\mathbf{r}_2 - \mathbf{r}_1). \quad (\text{C.26})$$

Moreover, the group property

$$\hat{U}(t_2, t_0) = \hat{U}(t_2, t_1)\hat{U}(t_1, t_0) \quad (\text{C.27})$$

together with the closure relation

$$\int d^3\mathbf{r}_1 |\mathbf{r}_1\rangle\langle\mathbf{r}_1| = \hat{1} \quad (\text{C.28})$$

yields the recursion relation

$$\begin{aligned} K(\mathbf{r}_2, t_2; \mathbf{r}_0, t_0) &= \langle\mathbf{r}_2|\hat{U}(t_2, t_0)|\mathbf{r}_0\rangle = \langle\mathbf{r}_2|\hat{U}(t_2, t_1)\hat{1}\hat{U}(t_1, t_0)|\mathbf{r}_0\rangle = \\ &= \int d^3\mathbf{r}_1 \langle\mathbf{r}_2|\hat{U}(t_2, t_1)|\mathbf{r}_1\rangle\langle\mathbf{r}_1|\hat{U}(t_1, t_0)|\mathbf{r}_0\rangle = \\ &= \int d^3\mathbf{r}_1 K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)K(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0). \end{aligned} \quad (\text{C.29})$$

As already anticipated at the end of the previous section, the equations for the quantum amplitude, Eqs. (C.25), (C.26) and (C.29), are indeed obtained from those for the probability density of Brownian motion by replacing $p(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ by $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$, dt_2 by $i dt_2$ and D by $\hbar/2m$.

To calculate explicitly the propagator of a free particle we use the momentum representation

$$\langle\mathbf{p}|\hat{U}(t_2, t_1)|\mathbf{p}'\rangle = \langle\mathbf{p}|e^{-\frac{i(t_2-t_1)}{2\hbar m}\hat{\mathbf{p}}^2}|\mathbf{p}'\rangle = e^{-\frac{i(t_2-t_1)}{2\hbar m}\mathbf{p}^2}\delta(\mathbf{p} - \mathbf{p}') \quad (\text{C.30})$$

together with the closure relation

$$\int d^3\mathbf{p} |\mathbf{p}\rangle \langle \mathbf{p}| = \hat{1}. \quad (\text{C.31})$$

We obtain

$$\begin{aligned} K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) &= \int d^3\mathbf{p} \langle \mathbf{r}_2 | \mathbf{p} \rangle e^{-\frac{i(t_2-t_1)}{2\hbar m} \mathbf{p}^2} \langle \mathbf{p} | \mathbf{r}_1 \rangle = \\ &= \frac{1}{(2\pi\hbar)^3} \int d^3\mathbf{p} e^{\frac{i}{\hbar} [\mathbf{p} \cdot (\mathbf{r}_2 - \mathbf{r}_1) - \mathbf{p}^2 (t_2 - t_1)/(2m)]} = \\ &= I(x_2 - x_1, t_2 - t_1) I(y_2 - y_1, t_2 - t_1) I(z_2 - z_1, t_2 - t_1) \end{aligned} \quad (\text{C.32})$$

where

$$I(x, t) := \int_{-\infty}^{\infty} \frac{dp}{2\pi\hbar} e^{\frac{i}{\hbar} [px - p^2 t/(2m)]}. \quad (\text{C.33})$$

The remaining integral can be reduced to a Gaussian integral by analytical continuation in the complex plane. Completing the square in the exponent and introducing the integration variable $u = \sqrt{t/(2\hbar m)}[p - (mx/t)]$ (for $t > 0$), we find

$$I(x, t) := \sqrt{\frac{m}{2\pi\hbar t}} e^{i\frac{mx^2}{2\hbar t}} \int_{-\infty}^{\infty} \frac{du}{\sqrt{\pi}} e^{-iu^2}. \quad (\text{C.34})$$

The Gaussian integral

$$\int_{-\infty}^{\infty} \frac{du}{\sqrt{\pi}} e^{-\lambda u^2} = \frac{1}{\sqrt{\lambda}} \quad (\text{C.35})$$

can be continued analytically to complex values of the parameter λ . With $\lambda = i$ we get

$$I(x, t) = \sqrt{\frac{m}{2\pi i \hbar t}} e^{i\frac{mx^2}{2\hbar t}}. \quad (\text{C.36})$$

The final result for the free propagator is

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = \left(\frac{m}{2\pi i \hbar (t_2 - t_1)} \right)^{3/2} e^{i\frac{m|\mathbf{r}_2 - \mathbf{r}_1|^2}{2\hbar(t_2 - t_1)}}. \quad (\text{C.37})$$

The probability density

$$|K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)|^2 = \left(\frac{m}{2\pi\hbar(t_2 - t_1)} \right)^3 \quad (\text{C.38})$$

does not depend on the spatial coordinates and is therefore not normalizable. This peculiarity is due to the choice of an unphysical initial state $|\mathbf{r}_1\rangle$, which is a superposition of state vectors $|\mathbf{p}\rangle$ with equal weights for all momenta and therefore it cannot be normalized. For a physical initial state the weight of states $|\mathbf{p}\rangle$ would tend to zero for $p \rightarrow \infty$. In

this case one would recover a normalized probability density. Let us then go back to the evolution of a physical state $|\psi(t)\rangle$ with $\langle\psi(t)|\psi(t)\rangle = 1$. Using Eqs. (C.21) and (C.28) we find

$$\begin{aligned}\psi(\mathbf{r}_2, t_2) &:= \langle\mathbf{r}_2|\psi(t_2)\rangle = \int d^3\mathbf{r}_1 \langle\mathbf{r}_2|\hat{U}(t_2, t_1)|\mathbf{r}_1\rangle \langle\mathbf{r}_1|\psi(t_1)\rangle = \\ &= \int d^3\mathbf{r}_1 K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) \psi(\mathbf{r}_1, t_1).\end{aligned}\quad (\text{C.39})$$

We see that the propagator provides a link between wave functions at two different times.

We consider now a particle in an external potential, which may depend on time. The Hamiltonian is given by $\hat{H} = \hat{H}_0 + \hat{V}(t)$ where \hat{H}_0 is the kinetic part and the potential is a function $V(\mathbf{r}, t)$ in the coordinate representation. The recursion relation (C.29) is also valid in this case and can be generalized to

$$K(\mathbf{r}_n, t_n; \mathbf{r}_0, t_0) = \int d^3r_1 \cdots \int d^3r_{n-1} K(\mathbf{r}_n, t_n; \mathbf{r}_{n-1}, t_{n-1}) \cdots K(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0) \quad (\text{C.40})$$

as in the case of Brownian motion. We choose equidistant time intervals $t_j = t_{j-1} + \varepsilon$ with $n\varepsilon = t_n - t_0$ and consider the limit $\varepsilon \rightarrow 0, n \rightarrow \infty$, keeping $n\varepsilon$ fixed. The evolution operator

$$\hat{U}(t + \varepsilon, t) = \hat{1} - \frac{i}{\hbar} \hat{H}(t) \varepsilon + \mathcal{O}(\varepsilon^2) \quad (\text{C.41})$$

can be rewritten as

$$\hat{U}(t + \varepsilon, t) = e^{-\frac{i}{\hbar} \hat{H}_0 \varepsilon} e^{-\frac{i}{\hbar} \hat{V}(t) \varepsilon} + \mathcal{O}(\varepsilon^2). \quad (\text{C.42})$$

To linear order in ε we obtain

$$\begin{aligned}K(\mathbf{r}_j, t_j; \mathbf{r}_{j-1}, t_{j-1}) &= \langle\mathbf{r}_j, t_j|\hat{U}(t_j - \varepsilon, t_{j-1})|\mathbf{r}_{j-1}, t_{j-1}\rangle = \\ &= \langle\mathbf{r}_j, t_j|e^{-\frac{i\varepsilon}{\hbar} \hat{H}_0}|\mathbf{r}_{j-1}, t_{j-1}\rangle e^{-\frac{i\varepsilon}{\hbar} V(\mathbf{r}_{j-1}, t_{j-1})} = \\ &= \left(\frac{m}{2\pi i \hbar \varepsilon}\right)^{3/2} e^{\frac{i\varepsilon}{\hbar} [m|\mathbf{r}_j - \mathbf{r}_{j-1}|^2/(2\varepsilon^2) - V(\mathbf{r}_{j-1}, t_{j-1})]}\end{aligned}\quad (\text{C.43})$$

where the expression (C.37) for the free electron propagator has been used. Inserting this result into Eq. (C.40) we arrive at

$$K(\mathbf{r}_n, t_n; \mathbf{r}_0, t_0) = \left(\frac{m}{2\pi i \hbar \varepsilon}\right)^{3n/2} \int d^3r_1 \cdots \int d^3r_{n-1} e^{\frac{i\varepsilon}{\hbar} \sum_{j=1}^n \left[\frac{m}{2} \frac{|\mathbf{r}_j - \mathbf{r}_{j-1}|^2}{\varepsilon^2} - V(\mathbf{r}_{j-1}, t_{j-1}) \right]} \quad (\text{C.44})$$

which we write formally as a path integral in the limit $\varepsilon \rightarrow 0$,

$$K(\mathbf{r}_f, t_f; \mathbf{r}_i, t_i) = \int_{\mathbf{r}_i, t_i}^{\mathbf{r}_f, t_f} \mathcal{D}(\mathbf{r}, t) e^{\frac{i}{\hbar} S(t_f, t_i)} \quad (\text{C.45})$$

where

$$S(t_f, t_i) := \int_{t_i}^{t_f} dt \left(\frac{m}{2} |\dot{\mathbf{r}}|^2 - V(\mathbf{r}, t) \right) \quad (\text{C.46})$$

is the classical action for the Lagrange function

$$L(\mathbf{r}, \dot{\mathbf{r}}, t) = \frac{m}{2} |\dot{\mathbf{r}}|^2 - V(\mathbf{r}, t) \quad (\text{C.47})$$

and the integration in (C.45) is over all paths $\mathbf{r}(t)$ with $\mathbf{r}(t_i) = \mathbf{r}_i$ and $\mathbf{r}(t_f) = \mathbf{r}_f$.

Path integrals can be worked out explicitly for some simple cases, such as the harmonic oscillator, but the physical insight provided by this formalism is equally important. A quantum particle which starts at a point \mathbf{r}_1 at t_1 explores all possible paths on its way to the point \mathbf{r}_2 at time t_2 . Its wave function at \mathbf{r}_2, t_2 is a coherent superposition of all the complex amplitudes associated with the different trajectories. The important parameter is of course \hbar . In the limit $\hbar \rightarrow 0$ the integrand of Eq. (C.45) oscillates rapidly from one path to a neighboring path, except for a stationary point where the variation of the action vanishes, $\delta S = 0$. This is just Hamilton's principle of classical mechanics, which determines the classical path from an initial point \mathbf{r}_i, t_i to a final point \mathbf{r}_f, t_f . In cases where $\hbar \ll S(t_f, t_i)$ a “semi-classical” approximation can be used where the action is expanded around its extremal value. Mathematically this corresponds to the stationary phase approximation. The path integral formalism is also useful for perturbative calculations, where $e^{iS/\hbar}$ is expanded in powers of S/\hbar . This is possible if the potential is weak enough.

In contrast to the Wiener integral of Brownian motion, Feynman's path integral remains rather formal and mathematically ill defined. To circumvent this difficulty, one can apply a “Wick rotation” and replace t by $-it$, *i.e.*, one starts with the evolution operator $e^{-\hat{H}(t_2-t_1)/\hbar}$. In this “Euclidean” formulation Feynman's path integral becomes a Wiener integral and is well defined. The transition to the quantum-mechanical propagator is made at the end of the calculation by applying the inverse rotation (t replaced by it). Euclidean path integrals appear in a natural way in quantum statistical mechanics.

C.3 Path integrals in statistical mechanics

After this detour to Brownian motion and time evolution in quantum mechanics we return to equilibrium statistical mechanics, where the path integral formulation is also useful and, in particular, yields a simple limiting procedure for recovering classical statistical mechanics. In order to keep the presentation simple, we restrict ourselves to

- (i) non-interacting particles in an external potential,
- (ii) one dimension,
- (iii) Boltzmann (and not Fermi-Dirac or Bose-Einstein) statistics.

The many-particle Hamiltonian is given by

$$\hat{H}_N = \sum_{i=1}^N \left(\frac{1}{2m} \hat{\mathbf{p}}_i^2 + V(\hat{\mathbf{x}}_i) \right). \quad (\text{C.48})$$

Its eigenstates are tensor products of single-particle states

$$|\nu_1, \dots, \nu_N\rangle = |\nu_1\rangle \otimes \dots \otimes |\nu_N\rangle \quad (\text{C.49})$$

where

$$\hat{H}|\nu\rangle = \varepsilon_\nu |\nu\rangle, \quad \hat{H} = \frac{1}{2m} \hat{\mathbf{p}}^2 + V(\hat{\mathbf{x}}). \quad (\text{C.50})$$

Boltzmann statistics does not impose any restrictions on the symmetry of the many-particle states and the (canonical) partition function is simply given by

$$Z_N = \prod_{i=1}^N \left(\sum_{\nu_i} e^{-\beta \varepsilon_{\nu_i}} \right) = \left(\sum_{\nu} e^{-\beta \varepsilon_{\nu}} \right)^N \quad (\text{C.51})$$

as already discussed in the special case of Einstein's model of the specific heat (Section 2.4). We have therefore just to calculate the partition function for a single particle

$$Z = \sum_{\nu} e^{-\beta \varepsilon_{\nu}} = \text{Tr} \left(e^{-\beta \hat{H}} \right). \quad (\text{C.52})$$

We can also formulate this problem in terms of the state operator $\hat{\varrho}$. In contrast to Section 2.3, we use here the unnormalized state operator

$$\hat{\varrho} := e^{-\beta \hat{H}}. \quad (\text{C.53})$$

It allows us to calculate, in principle, the thermal average of any observable A ,

$$\langle A \rangle = \frac{1}{Z} \text{Tr}(\hat{\varrho} \hat{A}), \quad Z = \text{Tr}(\hat{\varrho}). \quad (\text{C.54})$$

In coordinate representation we have the density matrix

$$\varrho(x, x') := \langle x | \hat{\varrho} | x' \rangle = \langle x | e^{-\beta \hat{H}} | x' \rangle. \quad (\text{C.55})$$

The closure relations

$$\sum_{\nu} |\nu\rangle \langle \nu| = \hat{1}, \quad \int_{-\infty}^{\infty} dx |x\rangle \langle x| = \hat{1} \quad (\text{C.56})$$

or, in coordinate representation with $\varphi_{\nu}(x) := \langle x | \nu \rangle$,

$$\sum_{\nu} \varphi_{\nu}(x) \varphi_{\nu}^*(x') = \delta(x - x'), \quad \int dx \varphi_{\nu}^*(x) \varphi_{\nu'}(x) = \delta_{\nu, \nu'} \quad (\text{C.57})$$

allow us to express the partition function as an integral over the density matrix,

$$\begin{aligned} Z &= \sum_{\nu} \langle \nu | e^{-\beta \hat{H}} | \nu \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \sum_{\nu} \varphi_{\nu}^*(x) \langle x | e^{-\beta \hat{H}} | x' \rangle \varphi_{\nu}(x') = \\ &= \int_{-\infty}^{\infty} dx \langle x | e^{-\beta \hat{H}} | x \rangle = \int_{-\infty}^{\infty} dx \varrho(x, x). \end{aligned} \quad (\text{C.58})$$

We divide the “time interval” from 0 to $\hbar\beta$ into n slices $j\varepsilon$, $j = 0, \dots, n$ with $\varepsilon = \hbar\beta/n$ and factorize

$$e^{-\beta\hat{H}} = e^{-\frac{\varepsilon}{\hbar}\hat{H}} \dots e^{-\frac{\varepsilon}{\hbar}\hat{H}} \quad (n \text{ factors}). \quad (\text{C.59})$$

Inserting $n - 1$ integrations over x_j , $j = 1, \dots, n - 1$ through closure relations, we obtain

$$\varrho(x_n, x_0) = \langle x_n | e^{-\beta\hat{H}} | x_0 \rangle = \int_{-\infty}^{\infty} dx_1 \dots \int_{-\infty}^{\infty} dx_{n-1} \langle x_n | e^{-\frac{\varepsilon}{\hbar}\hat{H}} | x_{n-1} \rangle \dots \langle x_1 | e^{-\frac{\varepsilon}{\hbar}\hat{H}} | x_0 \rangle. \quad (\text{C.60})$$

In the limit $\varepsilon \rightarrow 0$ we can separate kinetic and potential energies as

$$\langle x_j | e^{-\frac{\varepsilon}{\hbar}\hat{H}} | x_{j-1} \rangle \approx \langle x_j | e^{-\frac{\varepsilon\hat{p}^2}{2\hbar m}} e^{-\frac{\varepsilon V(\hat{x})}{\hbar}} | x_{j-1} \rangle = \langle x_j | e^{-\frac{\varepsilon\hat{p}^2}{2\hbar m}} | x_{j-1} \rangle e^{-\frac{\varepsilon V(x_{j-1})}{\hbar}}. \quad (\text{C.61})$$

The free “propagator” is evaluated in a similar way as in Eq. (C.32) (here we use the substitution $z = \sqrt{\varepsilon/(2\hbar m)} [p - i m(x_j - x_{j-1})/\varepsilon]$). One finds

$$\langle x_j | e^{-\frac{\varepsilon\hat{p}^2}{2\hbar m}} | x_{j-1} \rangle = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{\frac{ip}{\hbar}(x_j - x_{j-1})} e^{-\frac{\varepsilon p^2}{2\hbar m}} = \sqrt{\frac{m}{2\pi\hbar\varepsilon}} e^{-\frac{m}{2\hbar\varepsilon}(x_j - x_{j-1})^2}. \quad (\text{C.62})$$

Inserting this result into Eq. (C.60), we obtain the density matrix

$$\varrho(x_n, x_0) = \sqrt{\frac{m}{2\pi\hbar\varepsilon}} \prod_{j=1}^{n-1} \left(\sqrt{\frac{m}{2\pi\hbar\varepsilon}} \int_{-\infty}^{\infty} dx_j \right) e^{-\frac{\varepsilon}{\hbar} \sum_{j=1}^n \left[\frac{m(x_j - x_{j-1})^2}{2\varepsilon^2} + V(x_{j-1}) \right]}. \quad (\text{C.63})$$

We can interpret this result as a path integral over random walks from x_0 at $t = 0$ over the intermediate positions x_j at time $t_j = j\varepsilon$, $j = 1, \dots, n - 1$ to the final position x_n at time $\beta\hbar$. Therefore, in exact correspondence to the case of Brownian motion, Eq. (C.18), we can express the density matrix as a (Wiener) path integral

$$\varrho(x_f, x_i) = \int_{\substack{x(0)=x_i \\ x(\hbar\beta)=x_f}} \mathcal{D}x(t) e^{-\frac{1}{\hbar} \int_0^{\hbar\beta} dt \left[\frac{m}{2} \dot{x}^2 + V(x(t)) \right]}. \quad (\text{C.64})$$

The partition function (C.58) thus is expressed as

$$Z = \int_{-\infty}^{\infty} dx \int_{\substack{x(0)=x \\ x(\hbar\beta)=x}} \mathcal{D}x(t) e^{-\frac{1}{\hbar} \int_0^{\hbar\beta} dt \left[\frac{m}{2} \dot{x}^2 + V(x(t)) \right]}. \quad (\text{C.65})$$

Here there are only closed paths, starting at x at time 0 and coming back to x at time $\hbar\beta$. It is straightforward to calculate the path integral (C.64) for the case of the harmonic oscillator. The partition function obtained in this way reproduces what we have found in Section 2.4.

Instead of analyzing a particular case, we discuss more generally what happens in the high-temperature limit, $\hbar\beta \rightarrow 0$. If $\hbar\beta$ is very small, relevant trajectories (those contributing a negligible weight to the path integral) remain very close to the starting

point x (otherwise the term \dot{x}^2 would become very large and suppress the weight). We can therefore replace $V(x(t))$ by $V(x)$ and write

$$Z \approx \int_{-\infty}^{\infty} dx e^{-\beta V(x)} \int_{\substack{x(0)=x \\ x(\hbar\beta)=x}} \mathcal{D}x(t) e^{-\frac{1}{\hbar} \int_0^{\beta\hbar} dt \frac{m}{2} \dot{x}^2}. \quad (\text{C.66})$$

Using Eq. (C.62) for $x_j = x_{j-1} = x$ and replacing ε by $\hbar\beta$ we obtain

$$\langle x | e^{-\frac{\beta \hat{p}^2}{2m}} | x \rangle = \int_{\substack{x(0)=x \\ x(\hbar\beta)=x}} \mathcal{D}x(t) e^{-\frac{1}{\hbar} \int_0^{\beta\hbar} dt \frac{m}{2} \dot{x}^2} = \frac{1}{\hbar} \sqrt{\frac{mk_B T}{2\pi}}. \quad (\text{C.67})$$

This expression may be written as an integral

$$\int_{-\infty}^{\infty} \frac{dp}{2\pi} e^{-\frac{\beta p^2}{2m}} = \sqrt{\frac{mk_B T}{2\pi}}. \quad (\text{C.68})$$

Thus we obtain the simple result

$$Z = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dx dp}{h} e^{-H(x,p)} \quad (\text{C.69})$$

where $H(x, p)$ is the classical Hamilton function. This is the well-known classical partition function.

Literature on path integrals

1. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill 1965 (still the standard reference).
2. R. P. Feynman, *Statistical Mechanics*, W. A. Benjamin 1972 (lecture notes including several applications of path integrals).
3. L. S. Schulman, *Techniques and Applications of Path Integration*, Dover Publications 2005.
4. H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets*, World Scientific 2004.