STATISTICAL MECHANICS

Lecture Notes L3 Physique Fondamentale
Université Paris-Sud

H.J. Hilhorst
2016-2017

Laboratoire de Physique Théorique, bâtiment 210
CNRS and Université Paris-Sud, 91405 Orsay Cedex, France
PREFACE

These Notes were written for the benefit of the students of the Statistical Mechanics I course at level L3 of the “Fundamental Physics” track. They can in no way replace attending the course, participating in the problem sessions, and reading books.

The choice of subjects has been chiefly determined by the requirement of coherence with the French version of the course. Several important subjects left out will be dealt with in the Statistical Mechanics II course or in advanced courses at the M1 and M2 level.

These Notes should be considered as an incomplete draft. Unavoidably, there must be errors and misprints that have escaped my attention. Criticism from colleagues and students is welcome.

H.J. Hilhorst
November 2016
# TABLE OF CONTENTS

I. FROM MECHANICS TO STATISTICAL MECHANICS  
II. MICROCANONICAL ENSEMBLE  
III. CANONICAL ENSEMBLE  
IV. HARMONIC OSCILLATOR  
V. GRAND-CANONICAL ENSEMBLE  
VI. STATISTICAL MECHANICS OF QUANTUM SYSTEMS  
VII. IDEAL QUANTUM GASES  
VIII. IDEAL FERMI GAS  
IX. IDEAL BOSE GAS  
X. PHOTON GAS  
XI. CORRELATION FUNCTIONS  
XII. SYSTEMS OF CHARGED PARTICLES  
XIII. CHEMICAL EQUILIBRIUM  

Appendix A. PROBABILITY DISTRIBUTIONS  
Appendix B. MISCELLANEOUS MATHEMATICS
I. FROM MECHANICS TO STATISTICAL MECHANICS

I.1. Reminders about classical mechanics

I.1.1. Single particle

The motion of a classical particle of mass \( m \) is given by Newton’s law

\[ F = ma. \]  

Here \( a = \dot{v} = \ddot{r} \) is the particle’s acceleration and \( F \) is the sum of all forces acting on it. These include

- forces due to other particles;
- forces due to external electric, magnetic, and/or gravitational fields;
- forces due to the walls, if any, between which the particle is enclosed.

We will set \( p = mv \). In general \( F \) will be a function of \( r, p, \) and \( t \). The equation of motion is then equivalent to the pair of equations

\[ \dot{r}(t) = \frac{\dot{p}}{m}, \]  
\[ \dot{p}(t) = F(r(t), p(t), t), \]  

which are of first order in the derivatives with respect to time. If at some arbitrary initial instant one knows

\[ r(t_0) = r_0, \quad p(t_0) = p_0, \]  

then one can calculate, in principle, \( r(t) \) and \( p(t) \) at every later instant \( t > t_0 \). In order to convince yourself of this, discretize time into small intervals (“time steps”) \( \Delta t \) and write, neglecting terms of order \( \Delta t^2 \),

\[ r(t_0 + \Delta t) = r(t_0) + \Delta t \frac{p(t_0)}{m}, \]  
\[ p(t_0 + \Delta t) = p(t_0) + \Delta t F(r(t_0), p(t_0), t_0). \]  

These equations express the position and momentum at time \( t_0 + \Delta t \) in terms of those at \( t_0 \). By iterating them one obtains the solution \((r(t); p(t))\) for all \( t > t_0 \). This solution may be visualized as a curve (a “trajectory”) parametrized by \( t \) in a space of dimension six. This space is called the particle’s phase space.

**Exercise.** Show that you may also find the trajectory at all earlier times \( t < t_0 \).

I.1.2. \( N \) particles

Let \( N \) particles, all of mass \( m \), be numbered by \( i = 1, 2, \ldots, N \). Newton’s equations read

\[ F_i = m a_i, \quad i = 1, 2, \ldots, N, \]  

where \( a_i = \dot{v}_i = \ddot{r}_i \) is the acceleration of the \( i \)th particle. Since it is not our purpose to treat the most general case, let us suppose that the forces at work are due only to
(i) An external potential $V_{\text{ext}}(r_i)$ acting on each of the particles $i = 1, 2, \ldots, N$. This force may represent, for example, a gravitational field

$$V_{\text{ext}}(r_i) = m g r_i^3$$

where $r_i = (r_{i1}, r_{i2}, r_{i3})$ and $g$ is the acceleration of gravity, or the effects of, for example, impenetrable walls,

$$V_{\text{ext}} = \begin{cases} 0 & \text{if } 0 \leq r_{i1}, r_{i2}, r_{i3} \leq L, \\ \infty & \text{elsewhere}, \end{cases}$$

(corresponding to particles enclosed in a cubic container of edge length $L$).

(ii) An interaction potential between the particles. More specifically we will suppose that there is only a spherically symmetric pair potential $V_{\text{int}}(|r_i - r_j|)$ between them.

Remark. This potential applies to spherically symmetric particles. For example, the potential $V_{\text{int}}(r)$ between two atoms of a noble gas at distance $r = |r|$ has in good approximation the shape indicated in Fig. 1. For argon (Ar) we have $\sigma = 3.4 \times 10^{-10}$ m and $\epsilon = 1.65 \times 10^{-21}$ J.

Suppose we know the $r_i$ and $p_i$ for all $i = 1, 2, \ldots, N$. The kinetic and potential energy of this system are then given by

$$E_{\text{kin}} = \sum_{i=1}^{N} \frac{p_i^2}{2m},$$

$$E_{\text{pot}} = \sum_{1 \leq i < j \leq N} V_{\text{int}}(|r_i - r_j|) + \sum_{i=1}^{N} V_{\text{ext}}(r_i),$$

so that its total energy $H$ becomes

$$H = \mathcal{H}(r_1, r_2, \ldots, r_N; p_1, p_2, \ldots, p_N)$$

$$= E_{\text{kin}} + E_{\text{pot}}.$$
The force on the $i$th particle is equal to minus the gradient of the total potential energy with respect to $r_i$ which gives

$$ F_i = -\frac{\partial E_{\text{pot}}}{\partial r_i}, \quad \frac{\partial}{\partial r_i} \equiv \left( \frac{\partial}{\partial r_{i1}}, \frac{\partial}{\partial r_{i2}}, \frac{\partial}{\partial r_{i3}} \right). \quad (11) $$

**Exercise.** Show that

$$ \frac{\partial V_{\text{int}}(|r_i - r_j|)}{\partial r_i} = \hat{r}_{ij} V'_{\text{int}}(|r_i - r_j|), \quad (12) $$

in which $V'_{\text{int}}(r) \equiv dV_{\text{int}}/dr$ and $\hat{r}_{ij} = (r_i - r_j)/|r_i - r_j|$. Define the external force $F_{\text{ext}}(r_i)$ and show that

$$ F_i = \sum_{j=1, j \neq i}^{N} \hat{r}_{ij} V_{\text{int}}(r_{ij}) + F_{\text{ext}}(r_i). \quad (13) $$

The equations of motion for the $N$-particle system can now be written as

$$ \dot{r}_i(t) = \frac{p_i(t)}{m}, \quad (14) $$

$$ \dot{p}_i(t) = -\frac{\partial}{\partial r_i} E_{\text{pot}}(r_1(t), r_2(t), \ldots, r_N(t)) $$

for $i = 1, 2, \ldots, N$. This is a closed system of $6N$ coupled first order ordinary differential equations. For a given set of initial values $\{r_i(0), p_i(0)\}$ they have a unique solution $\{r_i(t), p_i(t)\}$, which represents a trajectory in the $6N$-dimensional phase space of the $N$-particle system.

One easily verifies that Eq. (14) is equivalent to

$$ \dot{r}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial r_i}, \quad (15) $$

which are called the Hamilton or Hamilton-Jacobi equations of the particle system. In this context $H$ is called the Hamiltonian.

The advantage of this slightly more abstract form of the equations of motion is that it remains valid in more general cases, in particular when $r_i$ and $p_i$ are replaced with an arbitrary pair of “canonically conjugate” coordinates.\(^1\) Furthermore, Eqs. (15) are easy to use in formal proofs, some of which will follow.

**Exercise.** Argue why two trajectories in phase space cannot intersect. Can there be closed loop trajectories?

**Exercise.** The Hamiltonian of a classical one-dimensional oscillator of mass $m$ and angular frequency $\omega$ is

$$ H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \quad (16) $$

\(^1\)The definition of “canonically conjugate” is given in your Classical Mechanics course; here it suffices to know that $r_i$ and $p_i$ satisfy the condition of being canonically conjugate.
Figure 2: A trajectory in phase space, confined to a surface of constant energy.

Check that the Hamilton equations for this oscillator are the usual equations of motion that you know. Find the trajectories in phase space.

In order to discuss the consequences of Eqs. (15) more easily we will set

\[ r^N(t) = (r_1(t), \ldots, r_N(t)), \]
\[ p^N(t) = (p_1(t), \ldots, p_N(t)), \] (17)

and

\[ \Gamma(t) = (r^N(t); p^N(t)) \] (18)

Now \( \Gamma(t) \) is a trajectory in a 6N-dimensional phase space. For a given \( \Gamma(0) \) the Hamilton-Jacobi equations fully determine \( \Gamma(t) \) at all later (and earlier) times. This is a purely mechanical problem.

Remark. In classical mechanics the notation \((q_i, p_i)\) instead of \((r_i, p_i)\) is usual.

I.2. Ensembles: general

I.2.1 Probability distributions in phase space

The number of gas molecules in a cm\(^3\) is of the order of 10\(^{23}\), and it is of course in practice possible neither to determine the initial \(r_i(0)\) and \(p_i(0)\) nor, even if we knew them, to calculate all the \(r_i(t)\) and \(p_i(t)\). And moreover, we would not be interested in so much detail.

We will use the terminology

**microstate** = point \( \Gamma \) in phase space

Since we can never know the microstate of a system, we now pass from simple mechanics to statistical mechanics by introducing the concept of a probability density \( P(\Gamma) = P(r^N; p^N) \) for the system to be in microstate \( \Gamma = (r^N; p^N) \). We will call

**macrostate** = probability distribution \( P(\Gamma) \) in phase space
The probability distributions that we are going to use will always depend on a few parameters, such as the volume, the temperature, etc. of the system; so the macrostate will depend on these parameters. We should evidently have the normalization

\[ \int d\Gamma P(\Gamma) \equiv \int d\mathbf{r}^N d\mathbf{p}^N P(\mathbf{r}^N; \mathbf{p}^N) = 1, \]  

(19)

where \( d\mathbf{r}^N \equiv d\mathbf{r}_1 d\mathbf{r}_2 \ldots d\mathbf{r}_N \) and \( d\mathbf{p}^N = d\mathbf{p}_1 d\mathbf{p}_2 \ldots d\mathbf{p}_N \). The actual choice of \( P(\Gamma) \) will be motivated by the experimental conditions: the way the system has been prepared, the absence or presence of contacts with its environment, etc. In general an initial \( P(\Gamma) \) will evolve in time and we will write it as \( P(\Gamma, t) \). For a physical system in equilibrium \( P(\Gamma) \) is independent of time. We will soon see examples.

I.2.2. Link with experiment

Let \( A \) be a measurable physical quantity (such as the pressure or the temperature of a gas) whose value in microstate \( \Gamma \) is equal to \( A(\Gamma) \). Suppose we measure \( A \) at some instant of time \( t \). Such a measurement necessarily takes a short but finite time \( \tau \). The true experimental result obtained will be the value of \( A(\Gamma) \) averaged along the segment of the system’s trajectory between two times \( t \pm \tau/2 \),

\[ A_{\text{exper}} = \frac{1}{\tau} \int_{t-\tau/2}^{t+\tau/2} \! dt' \, A(\Gamma(t')). \]  

(20)

Although \( \tau \) is short at the macroscopic (human) scale, it will be a very long time at the microscopic scale. The essential hypothesis now is that this segment of a trajectory passes through a sequence of points of phase space that are sufficiently representative for the full region of phase space in which \( P(\Gamma) \) is sensibly different from zero. The average along the trajectory may therefore be replaced with the phase space average

\[ \langle A \rangle = \int d\Gamma P(\Gamma) A(\Gamma). \]  

(21)

We consider this average as the theoretical prediction for the result of the experimental measurement.

We will henceforth consider the equivalence between (21) and (20) as established. This is called the ergodic hypothesis. Systems that can be proven to satisfy this hypothesis are called ergodic; however, for the vast majority of Hamiltonians of interest such proofs do not exist.

I.2.3. Energy conservation

Can a trajectory go everywhere in phase space? The answer is “no”, for the simple reason that – if the system is isolated from its environment (as we will suppose), and if the Hamiltonian does not depend explicitly on the time \( t \) – the total energy \( E \) of the system remains constant. Now the equation

\[ \mathcal{H}(\Gamma) = E \]  

(22)

represents a hypersurface of dimension \( 6N - 1 \) in phase space. Hence every trajectory is confined to one of these hypersurfaces. This is also expressed by saying that \( \mathcal{H} \) is an invariant.

One should wonder if there are maybe other invariants. These would then divide every hypersurface \( \mathcal{H} = \text{cst} \) into smaller sectors to which the trajectories would be confined. In general, except for special cases that one recognizes as such, it is supposed that \( \mathcal{H} \) is the only invariant.
The proof that $\mathcal{H}$ is an invariant takes a single line,

$$\frac{d\mathcal{H}}{dt} = \sum_{i=1}^{N} \left[ \frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i + \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i \right] = 0,$$

(23)

where the second equality results from the equations of motion (15).

I.2.4. Ensembles

We may consider a plain density $\rho(\Gamma)$ in phase space such that $\rho(\Gamma) \geq 0$, but which is not normalized. One passes easily from $\rho(\Gamma)$ to a probability density $P(\Gamma)$ by means of the relation

$$P(\Gamma) = \frac{\rho(\Gamma)}{\int d\Gamma \rho(\Gamma)}.$$

(24)

One may represent any positive function $\rho(\Gamma)$ on phase space as a collection of points such that the larger $\rho(\Gamma)$ is, the denser the points are locally. At this stage of our discussion the total number of points has no significance. Then, abandoning the original interpretation of a probability law for a single physical system, we imagine that each of these points corresponds to a different system. Whence the name ensemble for $\rho(\Gamma)$. This alternative interpretation is often useful. The phase space average (21) will also be referred to as the ensemble average.

Remark. The ensemble formalism that we are about to develop does not exclude the study of a precise microstate. If you know the microstate $\Gamma_0$ of the system of interest to you, then you can express this knowledge by setting

$$\rho(\Gamma) = \delta(\Gamma - \Gamma_0).$$

(25)

Remark. The phase space of a classical physical system is a continuum: there is an uncountable infinity of microstates $\Gamma$. However, quantum mechanically the energy levels and eigenstates of a particle (harmonic oscillator, particle in a box) are discrete and countable. As a consequence, so is the phase space of a quantum system: it is discrete. This observation does not play a role right away – for the moment we will study only classical systems – but we have to keep it in mind.

Concerning ensembles, a more limited question is the following one. Suppose we have let the system relax to equilibrium. We then expect it to be described by a stationary (= time-independent) density, that we will denote by $\rho(\mathbf{r}^N; \mathbf{p}^N)$. Can we find such a density? See section I.4.

I.3. The flow in phase space

Let there be given at a certain time $t_0$ an ensemble of points (microstates) in phase space, described by a density $\rho(\Gamma, t_0)$. When the time evolution is “turned on”, each of these microstates starts moving along a trajectory determined by the Hamilton equations. At every later instant of time $t > t_0$ we therefore have a new set of microstates, described by a density $\rho(\Gamma, t)$. The relation between $\rho(\Gamma, t)$ and $\rho(\Gamma, t_0)$ is, evidently, completely determined by the Hamilton equations. Question: how can we obtain it explicitly?

However, even before invoking these equations, we may observe that every point in the ensemble present at $t_0$ will still be present at time $t$, and no new point will have appeared. It
follows that the density $\rho$ is conserved, that is, satisfies the continuity equation
\[ \frac{\partial \rho}{\partial t} + \text{div } \rho V = 0, \]  
where $V$ is the $6N$-dimensional velocity in phase space (note: only 3-dimensional vectors will be printed in boldface). For later use we write Eq. (26) as
\[ \frac{\partial \rho}{\partial t} + V \cdot \text{grad } \rho + \rho \text{div } V = 0. \]  
These equations require the following explanation. If at some instant $t$ a microstate of the ensemble is at $\Gamma = (r_1, \ldots, r_N; p_1, \ldots, p_N)$, then it has at that moment along its trajectory the velocity
\[ V(\Gamma) = \frac{d\Gamma}{dt} = (\dot{r}_1, \ldots, \dot{r}_N; \dot{p}_1, \ldots, \dot{p}_N). \]

The current density in phase space is therefore equal to $\rho V$. Furthermore, in the continuity equation the operators “div” and “grad” are the $6N$-dimensional generalizations of those that you know,
\[ \text{grad } \rho = \left( \frac{\partial \rho}{\partial r_1}, \ldots, \frac{\partial \rho}{\partial r_N}; \frac{\partial \rho}{\partial p_1}, \ldots, \frac{\partial \rho}{\partial p_N} \right), \]
\[ \text{div } V = \sum_{i=1}^N \sum_{\alpha=1}^3 \left[ \frac{\partial \dot{r}_{i\alpha}}{\partial r_{i\alpha}} + \frac{\partial \dot{p}_{i\alpha}}{\partial p_{i\alpha}} \right]. \]
If now we employ the Hamilton-Jacobi equations (15) we may write (28) as
\[ V = \left( \frac{\partial \mathcal{H}}{\partial p_1}, \ldots, \frac{\partial \mathcal{H}}{\partial p_N}; -\frac{\partial \mathcal{H}}{\partial r_1}, \ldots, -\frac{\partial \mathcal{H}}{\partial r_N} \right). \]

With the help of definition (30) one sees from (31) that
\[ \text{div } V = 0, \]
which means that the flow in phase space is “incompressible”, this term being borrowed from the flow of a real three-dimensional fluid. From (32) and (27) we obtain another equation that expresses this incompressibility, namely
\[ \frac{\partial \rho}{\partial t} + V \cdot \text{grad } \rho \equiv \frac{D\rho}{Dt} = 0. \]

The symbol $D/Dt$ defined here is the hydrodynamic derivative: the equation says that the density stays constant for an observer traveling along with the moving points. Equations (32) and (33) are equivalent formulations of what is called the Liouville theorem.

\[^2\text{It has many other names: material derivative, convective derivative, . . .}\]
If in Eq. (33) we substitute for \( V \) the explicit expression (31), we find that the density \( \rho(\Gamma, t) \) obeys
\[
\frac{\partial \rho}{\partial t} = \sum_{i=1}^{N} \left[ \frac{\partial H}{\partial r_i} \cdot \frac{\partial \rho}{\partial p_i} - \frac{\partial H}{\partial p_i} \cdot \frac{\partial \rho}{\partial r_i} \right],
\]
which is called the Liouville equation. This answers the question asked at the beginning of this section.

**Summary.** Let the Hamiltonian of an \( N \)-particle system be given. Eqs. (15) then show how to calculate the time evolution of an individual system; and similarly, Eq. (34) shows how to calculate the time evolution of an ensemble of systems.

Eq. (34) is linear in \( \rho \). Therefore the Liouville equation may be written succinctly as
\[
\frac{\partial \rho}{\partial t} = \mathcal{L} \rho, \tag{35}
\]
where \( \mathcal{L} \), called the Liouville operator, is a very complicated but linear operator acting in the space of all functions \( \rho(\Gamma) \). Suppose that at some initial instant of time \( t = 0 \) we know \( \rho(\Gamma, 0) \). The linearity of Eq. (35) then allows us to write the solution \( \rho(\Gamma, t) \) at any later time \( t > 0 \) as
\[
\rho(\Gamma, t) = e^{\mathcal{L} t} \rho(\Gamma, 0). \tag{36}
\]
This solution, however, is only formal: it is impossible, in general, to calculate the operator \( e^{\mathcal{L} t} \) and no practical conclusion can be drawn directly from (36). Nevertheless, Eq. (35) is the starting point for the statistical mechanics of (classical) time dependent systems. In this introductory course, however, we will be concerned almost exclusively with systems at equilibrium: their properties do not change with time.

**Exercise.** Show that the invariance of the energy, expressed by (23), may also be written as
\[
V \cdot \nabla H = 0. \tag{37}
\]
Show furthermore that
\[
|V| = |\nabla H|. \tag{38}
\]
Interpret both relations.

**I.4. Stationary ensembles**

*Stationarity*

In general an ensemble \( \rho(\Gamma, t) \) depends on time and therefore any average \( \langle A \rangle \) calculated in that ensemble will also depend on time. We know experimentally that an isolated system will tend to an equilibrium state in which its macroscopic properties have stationary (= time independent) values. This is why we are interested in stationary ensembles \( \rho(\Gamma) \).

It is not hard to find such ensembles. Whenever we can write \( \rho \) as some function \( \phi \) of \( H(\Gamma) \),
\[
\rho(\Gamma) = \phi \left( H(\Gamma) \right), \tag{39}
\]
then upon substituting it in the Liouville equation (34) we find that \( \partial \rho / \partial t = 0 \) whatever \( \phi \); or more rapidly,
\[
\frac{\partial \rho}{\partial t} = \frac{d\phi}{dt} \frac{dH}{dt} = 0, \tag{40}
\]
where the second equality sign is due to (23). If the system were to have still other invariants than $\mathcal{H}$, we might of course let $\phi$ depend on those, too.

It might seem unsatisfactory that there is an arbitrary function $\phi$ here. Now, as we will see, statistical mechanics is at the basis of thermodynamics, and it will appear – however surprising this may seem at first – that any choice of $\phi$ leads to the same thermodynamics. This is why we say that all stationary ensembles are equivalent. Different choices of $\phi$, however, do have consequences when we ask questions that go beyond simple thermodynamics (and which may, for example, concern fluctuations). The choice of the appropriate ensemble will be dictated by experimental considerations. All this will become clear step by step in the following chapters.

In the remainder of this section we address various points that concern all ensembles.

* **Partition function**
  In stationary ensembles the key role is played by a quantity $Z$, called the partition function, and defined as
  \[ Z = \int d\Gamma \rho(\Gamma). \]  
  (41)
  You might think that this quantity, already encountered in Eq. (24), is nothing but a simple normalization constant without great interest. A normalization constant indeed it is, but its interest goes far beyond what is clear at this stage.

* **Three important ensembles**
  We discuss in the following chapters of these Notes the three most common statistical ensembles, namely
  – the microcanonical ensemble (index “m”);
  – the canonical ensemble (index “c”);
  – the grand-canonical ensemble (index “g”).

* **Walls**
  We call system the part of nature that we select for our investigation. We call wall the separation between a system and its surroundings. The nature of the wall plays a role in choosing the ensemble.
  A wall is said to be (completely) isolating if it does not allow for the exchange of either heat of particles. We call it adiabatic if we want to emphasize that it does not allow for the exchange of heat. A wall is said to be diathermal if it allows for the exchange of heat. It is called permeable if it allows for the exchange of one or more species of particles (and necessarily for the energy that comes along with them). Finally, we may consider a wall able to slide without friction, which is then called a piston.
  The same terminology applies to walls that separate two subsystems of a total system.

* **Intensive and extensive variables**
  When you double a thermodynamic system by putting a copy of it next to it, the volume, number of particles, energy, … of the total system will also double; they are called extensive variables. However, the pressure, density, … will remain the same; they are called intensive variables.
  The ratio of two extensive variables is intensive. In particular, if $N$ is the number of particles in a system and $X_k$ is another extensive variable, then $x_k \equiv X_k/N$ is intensive: it represents
the amount of $X_k$ per particle; similarly $\tilde{x}_k \equiv X_k/V$ represents the amount of $X_k$ per unit of volume.

A consequence is the following. Let $N$ and $X_1, X_2, \ldots$ be extensive variables. Let $Y$ be another extensive variable that may be expressed in terms of them, and set $y = Y/N$.

$$\frac{Y(N, X_1, X_2, \ldots)}{N} = y(x_1, x_2, \ldots),$$

(42)

that is, an intensive variable can depend only on ratios of extensive variables.

Thermodynamics is valid only in the limit of a system that is very large with respect to the microscopic scales of length, energy, etc. This is referred to as the thermodynamic limit. Theoretically it may be taken, for example, by letting $N \to \infty$ while keeping all intensive variables fixed. The above relations between intensive and extensive variables hold, strictly speaking, only in the thermodynamic limit. In statistical mechanics it will often be tacitly understood that the limit $N \to \infty$ has been, or is to be taken.

[*** To be elaborated in class.]

**Exercise.** What contradictions would arise if, for example, the energy were superextensive? Subextensive? This shows that extensivity is a necessary condition for thermodynamic stability.
II. MICROCANONICAL ENSEMBLE

II.1. The ensemble

II.1.1. Boltzmann’s postulate

The microcanonical ensemble is based on Boltzmann’s postulate:

If an isolated macroscopic system has an energy in the interval \((E, E + \delta E)\) and is in equilibrium, then, in the absence of any further information, all its microstates having an energy in that interval have the same probability.

The mathematical expression of this postulate leads us directly to the definition of the microcanonical ensemble, namely

\[
\rho_{\text{m}}(\Gamma) = \begin{cases} 
C & \text{if } E < H(\Gamma) < E + \delta E, \\
0 & \text{otherwise,}
\end{cases}
\]  

(1)

in which \(C\) is a constant.

The simplest case occurs when the system is composed of \(N\) identical particles enclosed in a volume \(V\). One then calls the triplet \((E, V, N)\) the independent variables of the microcanonical ensemble. These are the control parameters that an experimentalist can vary. The other ensembles that we will consider are characterized by other triplets of independent variables.

The constant \(C\)

Its dimension \([C]\) is the inverse of \([d\Gamma]\) (why?) and hence of \([d\mathbf{r}_i d\mathbf{p}_i]^N\). One easily finds, therefore, that \([C] = \text{(Js)}^{-3N} = \text{[action]}^{-3N}\). Its definition: \(C\) is the number of microstates in the ensemble per unit of volume in phase space. This value might seem without importance, since the constant \(C\) drops out of the probability density \(P(\Gamma)\) defined by (24). By convention, and for good reasons, one chooses \(C\) such that \(\rho_{\text{m}}(\Gamma)d\Gamma\) is equal to the number of quantum mechanical microstates that correspond\(^3\) to the classical phase space volume \(d\Gamma\). Hence

\[
\int_{E < H(\Gamma) < E + \delta E} d\Gamma \rho_{\text{m}}(\Gamma) = \text{number of microstates of the corresponding quantum system having an energy between } E \text{ and } E + \delta E.
\]  

(3)

In this classical discussion we therefore fix \(C\) by appealing to the underlying quantum mechanics. We state here without proof that this leads to

\[
C = \frac{1}{N! \hbar^{3N}},
\]  

(4)

where \(\hbar\) is Planck’s constant (remember that \([\hbar]=\text{action}\),

\[
\hbar = 6.63 \times 10^{-34} \text{Js.}
\]  

(5)

\(^3\)In a sense that we will not elaborate upon here.
The presence of the factor $1/\hbar^{3N}$ may be made plausible for the case of free particles (done in class). The appearance of the factor $1/N!$ in Eq. (4) is

- on the one hand, necessary to guarantee the extensivity of the thermodynamic quantities, as we will soon see;
- on the other hand, a consequence of quantum mechanics, which says that the exchange of two identical particles (in a many-particle system) does not lead to a new microstate, the two particles being indistinguishable (or: indiscernible). However, our classical equations (to be precise, the integral $\int d\Gamma$) count two such configurations as distinct microstates. The factor $1/N!$ corrects this redundancy.

Remarks.

(1) For a system with $s$ different species of particles with numbers $N_1, \ldots, N_s$ the factor $1/N!$ should obviously be replaced with $1/(N_1!\ldots N_s!)$.

(2) We will encounter physical systems in which it is impossible for two particles to exchange positions. In such cases the factor $1/N!$ is absent. The best known example is a crystal lattice, where each particle is bound by elastic forces to a well-defined lattice site.

II.1.2. Partition function and entropy

In what follows we denote by an index “m” all quantities defined in the microcanonical ensemble (for certain quantities this index will be suppressed later on). There is, first of all, the microcanonical partition function,

$$Z_m(E, V, N) = \int d\Gamma \rho_m(\Gamma)$$

$$= \frac{1}{N!\hbar^{3N}} \int_{(E, \delta E)} d\Gamma,$$

where $(E, \delta E)$ is shorthand notation for the shell $E < H(\Gamma) < E + \delta E$. Let us now introduce the auxiliary quantity

$$\Phi(E) = \frac{1}{N!\hbar^{3N}} \int_{H(\Gamma)<E} d\Gamma$$

where for convenience we have suppressed the variables $V$ and $N$), which represents the (dimensionless) volume of phase space corresponding to a system energy at most equal to $E$. Then obviously

$$Z_m(E, V, N) = \delta E \Phi'(E),$$

where $\Phi' \equiv (\partial \Phi/\partial E)_{N,V}$ is the density of states,\textsuperscript{4} that is, the number of energy levels per unit interval on the energy axis.

With each partition function one associates a thermodynamic potential, called that way because one deduces all other thermodynamic quantities from it by differentiations. The thermodynamic potential associated with the microcanonical partition function is the (microcanonical) entropy defined by

$$Z_m = e^{k_B^{-1}S_m} \quad \text{or} \quad S_m(E, V, N) = k_B \log Z_m(E, V, N),$$

\textsuperscript{4}Often the symbol $\rho(E)$ is used for this quantity; if so, it should not be confused with the ensembles $\rho(\Gamma)$. Furthermore, $\Omega(E, V, N)$ is an alternative notation for $Z_m(E, V, N)$.
Figure 3: Microcanonical ensemble.

where $k_B$ is Boltzmann’s constant,

$$k_B = 1.38 \times 10^{-23} \text{ JK}^{-1}.$$  \hspace{1cm} (10)

This entropy is a function on the three-dimensional space of coordinates $(E, V, N)$. It is not clear at this point – but will soon become so – that this $S_m$ is the same thing as the thermodynamic entropy $S$.

II.2. Examples: ideal gas and harmonic oscillator

II.2.1. Partition function and entropy of the ideal gas

Is it possible to calculate a microcanonical partition function? Yes, but only for certain model systems.

Let us consider an ideal (= noninteracting, $V_{int}(r) = 0$) gas\(^5\) of $N$ particles enclosed in a volume $V$ and having a total energy $E$. The microcanonical partition function of this system may now be calculated as follows. It is convenient to start from Eq. (8) and to begin by calculating $\Phi(E)$,

$$\Phi(E) = \frac{1}{N! h^{3N}} \int_{r_i \in V} \int \sum_i p_i^2 < 2mE \, dp_1 \ldots dp_N$$

$$= \frac{V^N}{N! h^{3N}} \left(2mE\right)^{\frac{3N}{2}}\frac{\pi^{\frac{3N}{2}}}{\Gamma\left(\frac{3N}{2} + 1\right)},$$  \hspace{1cm} (11)

where $\Gamma(x)$ is the Gamma function and where we have used the formula for the volume of a $d$-dimensional sphere. We see that

$$\Phi'(E) = \frac{3N}{2E} \Phi(E).$$  \hspace{1cm} (12)

\(^5\)More rarely called a perfect gas.
Upon applying Stirling’s formula,

\[ N! = N^N e^{-N} \sqrt{2\pi N} \left[ 1 + O(N^{-1}) \right], \quad N \to \infty, \quad (13) \]

to the factorial and the Gamma function in (11) we obtain

\[ \Phi(E) = \frac{e^{\frac{3}{2}N}}{\pi \sqrt{6N}} \left( \frac{V}{N} \right)^N \left( \frac{mE}{3\pi \hbar^2 N} \right)^{\frac{3}{2}N} \left[ 1 + O(N^{-1}) \right]. \quad (14) \]

When substituting (14) in (8) and (8) in (9) we obtain an explicit expression for the micro-canonical entropy of the ideal gas,

\[ S_m(E,V,N) = Nk_B \left[ \frac{5}{2} + \log \left( \frac{V}{N} \left( \frac{mE}{3\pi \hbar^2 N} \right)^{3/2} \right) \right] + O(\log N) + O \left( \log \frac{N \delta E}{E} \right). \quad (15) \]

In the thermodynamic limit the \( O \) terms may be ignored; Eq. (15) is then called the Sackur-Tetrode formula (1912). Note that the entropy is obviously extensive and that the RHS of Eq. (15) has the structure of Eq. (I.42),

\[ S_m(E,V,N) = Ns_m(\epsilon,v), \quad (16) \]

where we have set \( \epsilon = E/N \) and \( v = V/N \).

\textit{Remark.} In Eq. (11) the logarithm of neither \( V^N \) nor \( N! \) is extensive, but the \( \log V^N/N! \) is. This shows once more the need for the factor \( 1/N! \) when all particles are free to move through the whole volume. Furthermore, the logarithm of neither \( (2mE)^{\frac{3}{2}N} \) nor \( \Gamma(\frac{3}{2}N + 1) \) is extensive, but the logarithm of their ratio is.

II.2.2. \textit{Harmonic oscillator.} Will be done in the problem session.

II.3. \textbf{Identification of thermodynamics and statistical mechanics}

After these examples we continue to develop the general theory.

Thermodynamics deals with the behavior of macroscopic bodies (gases, liquids, magnets, . . . ) (and of radiation) when subjected to temperature changes and external forces. Thermodynamics is a self-contained theory. It relates directly observable quantities to each other and to new more abstract quantities (such as entropy, internal energy,. . . ). Thermodynamics describes matter without referring to its microscopic structure, and without making any assumptions about it. A certain thermodynamic phenomenon (for example, melting) may result from molecular rearrangements, but it will never be possible to identify such microscopic causes from thermodynamics alone. Thermodynamics was developed principally on the basis of the experimental and theoretical study of gases (motivated by their application to heat engines). It was then gradually realized that thermodynamic theory applies in the same way\(^6\) to a whole range of other systems, whence its great value as a general macroscopic theory. In these Notes we will develop a large collection of statistical mechanical expressions. We will then see that thermodynamics results from statistical mechanics if one is willing to identify certain of these expressions with well-known thermodynamic quantities.

\(^6\)In many cases by just making the proper substitution of variables.
Recall now the following from your thermodynamics course. Let there be a gas of \( N \) particles enclosed in a volume \( V \) and having a total ("internal") energy \( E \). These are three parameters that we may control in an experiment. Thermodynamics says that the triplet \((E, V, N)\) completely characterizes the state of this system. Well then, how about this system’s temperature, pressure, and so on? Thermodynamics asserts that there is a state function \( S(E, V, N) \) called entropy from which all these may be derived by simple differentiations. Expressed differently, if its variables undergo changes \( dN, dV, \) and \( dE \), then the entropy change \( dS \) is given by

\[
dS = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN,
\]

in which \( T, p, \) and \( \mu \) are the temperature, the pressure, and the chemical potential of the system, respectively. It follows that

\[
\frac{1}{T} = \left( \frac{\partial S}{\partial E} \right)_{V,N}, \quad \frac{p}{T} = \left( \frac{\partial S}{\partial V} \right)_{E,N}, \quad -\frac{\mu}{T} = \left( \frac{\partial S}{\partial N} \right)_{E,V}.
\] (18)

Returning to statistical mechanics we consider the derivatives of \( S_m(E, V, N) \) along the three main axes in \( EVN \) space. By analogy to Eqs. (18) we now define the (microcanonical) temperature \( T_m \), pressure \( p_m \), and chemical potential \( \mu_m \) according to the three formulas

\[
\frac{1}{T_m} = \frac{\partial S_m(E, V, N)}{\partial E}, \quad \frac{p_m}{T_m} = \frac{\partial S_m(E, V, N)}{\partial V}, \quad -\frac{\mu_m}{T_m} = \frac{\partial S_m(E, V, N)}{\partial N}.
\] (19-21)

If the three independent variables undergo simultaneously infinitesimal changes \( dN, dV, \) and \( dE \), then the resulting entropy change \( dS_m \) will be given by

\[
dS_m = \frac{1}{T_m}dE + \frac{p_m}{T_m}dV - \frac{\mu_m}{T_m}dN,
\] (22)

which is the exact microcanonical analog of the thermodynamical relation (17).

We may now continue and find the microcanonical equivalents of all other thermodynamic quantities of interest. Let us take the heat capacity as an example.

In thermodynamics one defines the heat capacity\(^{10} \) at constant volume \( C_V \) by \( C_V = (\partial E/\partial T)_{N,V} \). The microcanonical analog is to define

\[
C_{V,m} \equiv \frac{\partial E}{\partial T_m} = \left( \frac{\partial T_m}{\partial E} \right)^{-1},
\] (23)

\(^{7}\)In thermodynamics the quantity of matter may be expressed in terms of its mass rather than in the “number of particles”, if one wishes to avoid all reference to microscopic structure.

\(^{8}\)In thermodynamics a very common symbol for this quantity is \( U \).

\(^{9}\)This is why, on the basis of a superficial analogy to electrostatics, the entropy is called a thermodynamic potential.

\(^{10}\)The specific heat \( c_V \) is the intensive counterpart of the heat capacity, \( c_V = C_V/N \).
the derivations being carried out at constant \(N\) and \(V\). Using that \(T_m = 1/(\partial S_m/\partial E)\) we get from (23) after some rewriting the expression for \(C_{V,m}\) entirely in terms of derivatives of the entropy \(S_m(E,V,N)\),

\[
C_{V,m} = -\frac{\left(\frac{\partial S_m}{\partial E}\right)^2}{\frac{\partial^2 S_m}{\partial E^2}} = -\frac{1}{T_m^2} \left(\frac{\partial^2 S_m}{\partial E^2}\right)^{-1}.
\]

(24)

We note that as a consequence of Eq. (23) and the fact that \(C_{V,m} > 0\) we also have that \(\partial T_m/\partial E > 0\) and that \(S_m\) is a concave function of \(E\).

II.4. Example: the ideal gas, again

II.4.1. Thermodynamics of the ideal gas

In section II.2 we have calculated the partition function \(Z_m(E,V,N)\) of the ideal gas and found its entropy \(S_m(E,V,N)\). We will now proceed to the thermodynamics of the ideal gas, based on the relation established in section II.3.

By deriving Eq. (15) it successively with respect to each of its three variables one finds straightforwardly

\[
\frac{\partial S_m}{\partial E} = \frac{3Nk_B}{2E}, \quad \frac{\partial S_m}{\partial V} = \frac{Nk_B}{V}, \quad \frac{\partial S_m}{\partial N} = k_B \log \left\{ \frac{V}{N \left( \frac{mE}{3\pi\hbar^2N} \right)^{3/2}} \right\}.
\]

(25)\hspace{1cm}(26)\hspace{1cm}(27)

By comparing Eqs. (19) and (25) we find an explicit expression for the microcanonical temperature,

\[
E = \frac{3}{2} Nk_B T_m;
\]

(28)

and by comparing Eqs. (20) and (26) we obtain the ideal gas law

\[
p_m V = Nk_B T_m.
\]

(29)

Both relations are familiar from thermodynamics, and this shows the consistency of our identification of \(k_B \log Z_m\) with the thermodynamic entropy.

Exercise. Check that Eqs. (25)-(29) all have both members either extensive or intensive.

Exercise. Check that for an ideal gas Eq. (23) leads to \(C_{V,m} = \frac{3}{2}Nk_B\), another formula well-known from thermodynamics.

Exercise. Show that

\[
s_m(\epsilon,v) = k_B \left[ \frac{5}{2} + \log \frac{v}{\Lambda^3} \right] = k_B \left[ \frac{5}{2} - \log n\lambda^3 \right],
\]

(30)
where $\lambda \equiv h/(2\pi mk_BT)^{1/2}$ is the de Broglie wavelength\textsuperscript{11} of a thermal particle of mass $m$, we have as before $v = V/N$, and in the second line we have introduce the particle density $n = v^{-1} = N/V$. The dimensionless combination $n\lambda^3$ will appear again in other problems.

II.4.2. Limits on the validity

There are two physical effects that set a limit to the range of validity of Eqs. (15) and (30).

1. Interactions.

In real (= nonideal) gases there are interactions between the atoms or molecules. When these interactions are short-ranged, as we assume throughout, they will play a role only when particles are close to one another. We may therefore rightfully expect that the expressions that we found above for the ideal gas are also valid for sufficiently dilute real gases. When the typical distance between particles becomes of the order of the range of the interaction, corrections to the ideal gas expressions should be expected. There exist several methods of calculating such corrections perturbatively, most of them developed in the first half of the 20th century. We will see an example in Sec. V.5.

Remark. It is clear that for systems with long-range interactions we should reason differently. Examples are (i) plasmas (= fluids of which all or an important fraction of the particles are electrically charged) and (ii) systems with only gravitational interactions. We will leave such systems aside for the moment.

2. Quantum effects.

The real world is quantum mechanical and this is a good reason for considering a quantum mechanical ideal gas. When we will do this, in later chapters, it will turn out to be essential to distinguish between bosons and fermions. We restrict ourselves here to the following remarks.

We have argued that in the quantum limit $Z_m$ should reduce to the count of the number of microstates in a narrow energy interval. The result of such a count can only be a positive integer (1,2,3,...). Therefore we have the inequality $S_m = k_B \log Z_m \geq 0$. Obviously Eqs. (15) and (30) violate this inequality: when $T_m$ becomes small, or $n$ large, these expressions for the entropy become negative. The reason is that they have been obtained as integrals on a phase space approximated as a continuum, instead of the actual space of discrete microstates. We may expect the classical approximation to be a good one as long as it gives a positive entropy. This leads to

$$n\lambda^3 \ll 1$$

as the condition of validity for the classical expressions. Again, they appear to certainly hold in the dilute limit. [We may write $n\lambda^3 = \left(\frac{h}{\Delta p \Delta x}\right)^3$, although with an interpretation of $\Delta p$ and $\Delta x$ slightly different from the one in quantum mechanics. This suggests that quantum mechanics comes in when the uncertainty relations are in danger of being violated. Discussed in class.]

II.5. Microcanonical systems in contact

Study here Appendix B.

\textsuperscript{11}Recall that $\lambda = 2\pi/k$ and that in quantum mechanics $p = h\kappa$, whence $\lambda = h/p$. 

21
Exercise. The entropy of two independent (= noninteracting) systems is the sum of their individual entropies. Show this for \( S_m \) (easy). Why is it plausible that this additivity remains true for two systems in contact only via short-range interactions?

II.5.1. Thermal contact

Equilibrium condition. We consider an isolated system composed of two subsystems 1 and 2 that are separated by an adiabatic wall. The systems are therefore microcanonical. We will denote their parameters by \( E_1, V_1, N_1 \) and \( E_2, V_2, N_2 \) and their partition functions by \( Z_{m,1} \) and \( Z_{m,2} \), respectively.

Let us now replace the adiabatic wall by a diathermal one. The two systems will start exchanging energy until again an equilibrium is reached. Our purpose here is to study this new equilibrium.

We will assume that the phase spaces of the individual systems have not changed,\(^{12}\) that is, we neglect any interactions between these systems. This will certainly be a good approximation for interaction forces that are short-ranged. We now ask what the properties are of this new equilibrium.

Let us set \( E = E_1 + E_2, \ V = V_1 + V_2 \), and \( N = N_1 + N_2 \). The partition function \( Z_{m,12}(N, V, E) \) of the total system involves a sum on all ways of distributing the total energy between its two components and is given by

\[
Z_{m,12}(E, V, N) = \delta E \int dE_1 \Phi_1(E_1, N_1, V_1) \int dE_2 \Phi_2(E_2, N_2, V_2) \delta(E - E_1 - E_2)
\]

\[
= \delta E \int dE_1 \Phi_1(E_1, N_1, V_1) \Phi_2(E - E_1, N_2, V_2)
\]

\[
= \int \frac{dE_1}{\delta E} Z_{m,1}(E_1, V_1, N_1) Z_{m,2}(E - E_1, V_2, N_2)
\]

\[
= \int \frac{dE_1}{\delta E} \exp \left[ k_B^{-1} S_{m,1}(E_1, V_1, N_1) + k_B^{-1} S_{m,2}(E - E_1, V_2, N_2) \right],
\]

where \( S_{m,12}(E, V, N; E_1) \) is the entropy\(^ {13} \) for a given arbitrary partitioning of the energy into portions \( E_1 \) and \( E - E_1 \) for the subsystems 1 and 2, respectively. The integral on \( E_1 \) is strongly dominated by the value of the integrand in its maximum, that is, by values \( E_1 \approx E_1^{\text{max}} \). It follows – as we will show in detail in section II.5.2. – that

\[
Z_{m,12}(E, V, N) = \exp \left[ k_B^{-1} S_{m,1}(E_1^{\text{max}}, V_1, N_1) + k_B^{-1} S_{m,2}(E - E_1^{\text{max}}, V_2, N_2) + \mathcal{O}(\log N) \right],
\]

in which the maximum value \( E_1^{\text{max}} \) is the solution of

\[
\frac{\partial S_{m,12}(E, V, N; E_1)}{\partial E_1} = 0,
\]

whence

\[
\frac{\partial S_{m,1}(E_1^{\text{max}}, V_1, N_1)}{\partial E_1} = \frac{\partial S_{m,2}(E - E_1^{\text{max}}, V_2, N_2)}{\partial E_2}.
\]

\(^{12}\)In quantum mechanical language: that the energy levels of the two systems have not changed.

\(^{13}\)C. Texier calls this quantity the entropie réeudite. I am not aware of a specific English term for it.
Invoking now definition (19) of the microcanonical temperature $T_m$ we see that (35) is equivalent to

$$T_{m,1} = T_{m,2},$$

(36)

which is the equilibrium condition that we know from thermodynamics (and daily practice!). This confirms the consistency of interpreting in the microcanonical ensemble the quantity $k_B \log Z_m$ as the entropy. It shows, moreover, that the equilibrium state of the total system is determined by the condition that its entropy be maximized as a function of $E_1$, which after the contact is established becomes a free parameter.

II.5.2. Proof of Eq. (35)

The exact reasoning goes as follows. Write $\nu_1 = N_1/N$ and $\nu_2 = N_2/N$. By the “thermodynamic limit” we will mean here the limit $N \to \infty$ with all the following ratios fixed: $\nu_1, \nu_2, v_1 = V_1/N_1, v_2 = V_2/N_2,$ and $\epsilon = E/N$. We let $\epsilon_1 = E_1/N$. Using (32) we may then write

$$e^{k_B^{-1}S_{m,12}(E,V,N;E_1)} = e^{Ng(\epsilon_1)}$$

(37)

where

$$g(\epsilon_1) = k_B^{-1} \nu_1 s_{m,1}(\nu_1 \epsilon_1, \nu_1 v_1) + k_B^{-1} \nu_2 s_{m,2}(\nu_2(\epsilon - \epsilon_1), \nu_2 v_2).$$

(38)

The quantity $\tilde{S}_{m,12}$ defined here is the entropy of the total system constrained to have its energy partitioned into two quantities $E_1$ and $E - E_1$; after establishment of the thermal contact $E_1$ becomes a free parameter. Eq. (32) may then be rewritten as

$$Z_{m,12}(E,V,N) = \frac{N}{\delta E} \int d\epsilon_1 e^{Ng(\epsilon_1)},$$

(39)

which is an integral exactly of the form discussed in Appendix B. We therefore know without any further calculation that

$$Z_{m,12}(E,V,N) = e^{Ng(\epsilon_1^\text{max})} + O(\log N),$$

(40)

where $\epsilon_1^\text{max} = E_1^\text{max}/N$ is the value of its argument for which $g(\epsilon_1)$ is maximal. The equation $\frac{dg}{d\epsilon_1}(\epsilon_1^\text{max}) = 0$, when the original variables are restored, is identical to Eq. (35). Furthermore $S_{m,12}(E,V,N) = \tilde{S}_{m,12}(E,V,N;E_1^\text{max})$.

II.5.3. Fluctuations

The quantity $\tilde{S}_{m,12}(E,V,N;E_1)$ is the entropy of the total system subject to the condition that the energies of the two subsystems be $E_1$ and $E - E_1$. In order that there be an $E_1^\text{max}$ where this entropy is maximal (rather than just stationary) it is needed that

$$\left. \frac{\partial^2 \tilde{S}_{m,12}}{\partial E_1^2} \right|_{E_1 = E_1^\text{max}} < 0.$$ 

(41)

This condition translates into $C_V > 0$, that is, the heat capacity should be positive.

From the properties of the steepest descent integral it follows that the range of energies around $E_1^\text{max}$ that contribute effectively to the result is determined by $\langle (\epsilon_1 - \epsilon_1^\text{max})^2 \rangle^{\frac{1}{2}} = \epsilon_1^\text{max}$. 

23
The function \( T_{m,1}(E_1) \) and \( T_{m,2}(E - E_1) \) are increasing and decreasing, respectively, with \( E_1 \). Since their point of intersection determines the equilibrium of the total system, we see that the system with the initially higher temperature furnishes energy to the one having the initially lower temperature: heat flows from high to low temperature (draw a figure!).

II.5.4. The piston: a sliding wall

Suppose now that the two systems 1 and 2 are separated by a piston: an adiabatic wall that may slide without friction (which is certainly an idealization). This piston is initially blocked and the two microcanonical subsystems are characterized by the triplets \((E_1, V_1, N_1)\) and \((E_2, V_2, N_2)\) as before. At some point we will
- release the blocking;
- make the wall diathermal.\(^{14}\)

The piston will then relax to a new equilibrium position: this may be considered as an “exchange of volume” between the two subsystems; and furthermore there will be exchange of energy. Therefore \( E_1 \) and \( V_1 \) become free parameters. Our purpose is again to study the new equilibrium state. Let

\[
\hat{S}_{m,12}(E, V, N; E_1, V_1) = S_{m,1}(E_1, V_1, N_1) + S_{m,2}(E - E_1, V - V_1, N - N_1)
\]

be its total entropy under the constraint of a fixed partitioning of the energy and the volume between the two subsystems determined by the pair \((E_1, V_1)\) of free variables. The partition function \( Z_{m,12} \) of the total system may again be expressed in terms of those of its subsystems, now as a convolution with respect to \( E_1 \) and to \( V_1 \). We will not write these down explicitly. The equilibrium conditions may be derived as for the previous case and we find

\[
\frac{\partial \hat{S}_{m,12}(E, V, N; E_1, V_1)}{\partial E_1} = 0, \quad \frac{\partial \hat{S}_{m,12}(E, V, N; E_1, V_1)}{\partial V_1} = 0.
\]

Its solution is \((E_1, V_1) = (E_1^{\text{max}}, V_1^{\text{max}})\). The equilibrium condition of therefore finally takes the form

\[
T_{m,1} = T_{m,2}, \quad p_{m,1} = p_{m,2}.
\]

Convexity in the variable \( V_1 \) at the point \((E_1^{\text{max}}, V_1^{\text{max}})\) means

\[
\frac{\partial^2 S_{m,2}}{\partial V^2} < 0 \iff \frac{\partial p_{m,2}}{\partial V} < 0.
\]

The second one of these relations says that when the volume increases (at constant energy and particle numbers), the pressure goes down. Again we see that the equilibrium state is the one that maximizes the total entropy.

**Exercise.** Verify by a method analogous to the one employed above that the system with the initially higher pressure will increase its volume at the cost of the volume of the system with the initially lower pressure.

\(^{14}\)A different problem arises if we keep the wall adiabatic. That problem is more difficult and we leave it aside.
II.5.5. Permeable wall

Consider again the same two subsystems, initially described by the triplets \((E_1, V_1, N_1)\) and \((E_2, V_2, N_2)\). Let us now make the wall (now immobile, again) permeable for particles, which we take to be of the same kind in the two subsystems. Let

\[
\tilde{S}_{m,12}(E, V, N; E_1, N_1) = S_{m,1}(E_1, V_1, N_1) + S_{m,2}(E - E_1, V - V_1, N - N_1)
\]

be its total entropy under the constraint of a fixed partitioning of the energy and the number of particles between the two subsystems determined by the pair \((E_1, N_1)\) of free variables. The equilibrium conditions may be derived as for the previous case and we find

\[
\frac{\partial \tilde{S}_{m,12}(E, V, N; E_1, N_1)}{\partial E_1} = 0, \quad \frac{\partial \tilde{S}_{m,12}(E, V, N; E_1, N_1)}{\partial N_1} = 0.
\]

(47)

Its solutions are \((E_1, N_1) = (E_1^{\text{max}}, N_1^{\text{max}})\). The equilibrium condition therefore finally takes the form

\[
T_{m,1} = T_{m,2}, \quad \mu_{m,1} = \mu_{m,2}.
\]

(48)

**Exercise.** Verify by methods analogous to the ones used above that the system with the lower chemical potential gains particles at the cost of the system with the higher one.

**Example.** All particles initially in one half of a box.

II.5.6. Arbitrary free variable \(X\)

We consider the general case of an arbitrary extensive state variable \(X\) in a microcanonical system of energy \(E\). Let \(\tilde{S}_m(E; X_1)\) be its the entropy subject to the condition of \(X\) taking the value \(X_1\); that is, \(\tilde{S}_m(E; X_1) = k_B \log Z_m(E; X_1)\), where \(Z_m(E; X_1)\) is the volume of phase space (the number of microstates) having \(X = X_1\). If \(X_1\) is unconstrained, the probability distribution of \(X_1\) then is

\[
P(X_1) = \frac{Z_m(E; X_1)}{Z_m(E)}.
\]

(49)

Since \(X\) is extensive, the probability distribution \(P(X_1)\) will be strongly peaked about its maximum, which occurs for the value \(X = X_1^{\text{max}}\) that solves

\[
\frac{\partial \tilde{S}_m(E; X_1)}{\partial X_1} \bigg|_{X_1^{\text{max}}} = 0.
\]

(50)

Let \(\Delta X_1 = X_1 - X_1^{\text{max}}\). Then the rms deviation of \(X_1\) from its equilibrium value is \((\Delta X_1^2)^{\frac{1}{2}}\), and this quantity is \(\propto \sqrt{N}\) for the same reasons as above. That is, for large \(N\) the rms deviation is negligible on the scale of the variable itself, or put otherwise, in the thermodynamic limit the random variable \(X_1\) takes the sure (= nonrandom) value \(X_1^{\text{max}}\). This shows how determinism at the macroscopic scale arises out of randomness in the microscopic world.

Finally the entropy of the system is given by \(S_m(E) = \tilde{S}_m(E; X_1^{\text{max}})\) up to a subextensive correction that is \(O(\log N)\).

**Remark.** We call \(X\) and \(\partial S/\partial X\) a pair of conjugate variables. The three pairs that we have encountered are \(E\) and \(1/T_m\), \(V\) and \(p_m/T_m\), and \(N\) and \(-\mu_m/T_m\). Loosely speaking one says
that $T$ is conjugate to $E$, $p$ to $V$, and $\mu$ to $N$. The first one of each pair may be considered as a “generalized force” capable of changing the second one.

II.5.7. *Irreversible increase of entropy*

We have found here a general rule: when in an isolated system a constraint $X = X_1$ is released, then the system will tend to a macrostate where $X$ has the value $X = X_1^{\text{max}}$ that maximizes the constraint-dependent entropy $\tilde{S}_m(X)$. Put more concisely,

> When in an isolated system in equilibrium a constraint is released, the system evolves towards a new equilibrium state that maximizes the entropy (subject to the remaining constraints).

The fact that the entropy does not decrease implies that at the macroscopic (thermodynamic) level the laws of nature are not time reversal invariant: they tell us that certain phenomena occur irreversibly. This is in contrast to the microscopic (Newton, Hamilton-Jacobi) equations from which we started. In our above argument we have concluded to this irreversibility only by comparing an initial equilibrium state (with a constraint) to a final equilibrium state (obtained when waiting long enough after release of the constraint). We have not derived the equations of motion that would describe how the initial state evolves into the final one. Such equations should be irreversible, and much literature has been devoted to the problem of how it is possible, if at all, to derive irreversible macroscopic equations from reversible microscopic ones: in the process a symmetry gets lost.

The macroscopic irreversibility is sometimes referred to as the *arrow of time.*

II.5.8. *Gibbs’ paradox*

Let the same two subsystems as before now contain gases at the same temperature and pressure. What happens if the wall is removed? They mix! Or do they? Two cases must be distinguished:

(a) 1 and 2 contained particles of the same kind;
(b) 1 and 2 contained particles of different kinds.

In case (b) the gases start mixing. This is an irreversible process and is accompanied by an increase $\Delta S$ of entropy (the increase is called the “entropy of mixing”). It may be calculated with the aid of the Sackur-Tetrode formula applied to the initial and to the final state (in the latter case, consider the volume as filled by two separate gases and add their entropies).

Gibbs’ paradox arises when we argue that in case (a) the gases also start mixing. False! When we put the wall back in place, then because of the indistinguishability of the particles we will again be in the initial state; and since the entropy can only increase, the state with the wall removed has $\Delta S = 0$ compared to the initial state. In case (a), too, the Sackur-Tetrode formula gives the correct answer for $\Delta S$.

This formula takes the quantum indiscernability of identical particles correctly into account. The term “paradox” dates from the time (end of the 19th century, before the advent of quantum theory) when in a purely classical context this was a serious problem.
II.6. More about entropy

II.6.1. Gibbs entropy

We will reason now on the basis of a discrete state space of microstates $i = 1, 2, \ldots, M$. Suppose a system may be in a microstate $i$ with probability $p_i$. Around 1875 Gibbs defined the entropy $S_G$ of this system as

$$S_G = -k_B \sum_i p_i \log p_i.$$  \hspace{1cm} (51)

For $p_i = 0$ it is understood that $p_i \log p_i = \lim_{\epsilon \to 0} \epsilon \log \epsilon = 0$. It follows that $S_G \geq 0$.

Relation to the microcanonical entropy. The microcanonical partition function $Z_m$ is a sum on all microstates in a given small energy interval $(E, E + \delta E)$, all with unit weight. If there are $M$ such microstates $i$, and we assign equal probability $1/M - 1$ to each of them, then (51) becomes

$$S_G = -k_B \sum_{i=1}^M \frac{1}{M-1} \log \frac{1}{M-1} = k_B \log M.$$  \hspace{1cm} (52)

This is exactly equal to what we defined as $S_m$. We will return to Gibbs’ expression for the entropy when studying the canonical ensemble.

II.6.2. Shannon entropy

If one sets $k_B = 1$, the entropy of a bit (= two-state variable) as used in computers is $-\log 1/2 = \log 2$. Let $p_i$ be the probability that a sequence of $M$ bits be equal to the specific sequence $i$. Introducing a prefactor $1/\log 2$ we then define the entropy of this random sequence as

$$S_S = -\frac{1}{\log 2} \sum_i p_i \log p_i = -\sum_i p_i \log_2 p_i,$$  \hspace{1cm} (53)

which is called the Shannon entropy. So normalized, the entropy of a single bit that is in either of its two states with probability one half is equal to unity. Shannon (1948) used expression (52) to develop a theory of the information content of a stream of symbols. Since that time the Shannon entropy plays a role in many interdisciplinary applications.

In information theory this entropy is interpreted as a quantification of the “lack of information.” To see this, consider a variable that may be in any one of a set of states $i$. If we know its exact state, say $i = i_0$, then $p_i = \delta_{i_0}$ and (52) gives $S_S = 0$: no information is lacking. In the opposite case where all information is lacking, the best we can do is to consider all states as having the same probability $p_i = 1/M$, which gives $S_S = (\log M)/(\log 2) = \log_2 M$. It may be shown that this is the maximum value that $S_S$ can take: all information is lacking.

Exercise. Show that the Gibbs or Shannon entropy is nonnegative; and that for two independent systems it is additive.

II.6.3. Origin of $S_G$

Let $\nu = 1, 2, \ldots, n$ be single-particle states (as opposed to microstates) for structureless particles. Hence $\nu$ is a point $(r, \nu)$ in the six-dimensional single-particle phase space. When $N$ independent particles are distributed over these states with the constraints that there be exactly $N_\nu$ in state $\nu$, then, in Boltzmann’s notation, the number of ways $W$ to realize this distribution is

$$W = \frac{N!}{\prod \nu N_\nu !}.$$  \hspace{1cm} (53)
The Boltzmann entropy of this distribution is $S_B(N_1, N_2, \ldots) = -k_B \log W$. Upon using Stirling’s formula for the factorials and setting $p_\nu = N_\nu / N$ we get

$$S_B = -Nk_B \sum_\nu p_\nu \log p_\nu,$$

where subextensive terms have been neglected. Eq. (51) may be seen as the generalization of (54) to microstates $i$ and a probability distribution $p_i$ in an $N$-particle phase space. Since each of the $W$ realizations is a microstate, we may also consider this $S_B$ as an entropy of the same type as $S_m$ (in the sense that it is based on counting microstates), but with constraints $\nu = 1, 2, \ldots, n$ instead of the $(E, V, N)$.

Note. Eq. (51) may be used also when there are correlations between the particles due to interaction terms in their Hamiltonian.

II.6.4. The Maxwell-Boltzmann distribution

Suppose we consider $N$ independent particles in an external potential $V_{\text{ext}}(r)$. Let the system be isolated, hence microcanonical, with a total energy $E = N\epsilon$.

We imagine the six-dimensional single-particle phase space $(r, \mathbf{v})$ partitioned into volume elements labeled by an index $\nu$, every element large enough to contain many particles but small enough so that the potential energy of a particle in that volume element is practically constant and given by $\epsilon_\nu = \frac{1}{2}m\mathbf{v}_\nu^2 + V_{\text{ext}}(r_\nu)$.

In Eq. (54) the quantity $p_\nu$ is the fraction of particles, out of a set of $N$ independent particles, that are in the single-particle state $\nu$. Boltzmann maximized the entropy (54) by allowing the constraints $\{p_\nu\}$ to vary subject only to the conditions

$$\sum_\nu p_\nu = 1, \quad \sum_\nu p_\nu \epsilon_\nu = \epsilon. \quad (55)$$

Introducing Lagrange multipliers $a$ and $b$ (see Appendix C) we are led to maximize

$$F(\{p_\nu\}) = -\sum_\nu p_\nu \log p_\nu - a(\sum_\nu p_\nu - 1) - b(\sum_\nu \epsilon_\nu p_\nu - \epsilon). \quad (56)$$

From $\partial F / \partial p_\nu = 0$ we get $\log p_\nu + 1 + a + b\epsilon_\nu = 0$, whence

$$p_\nu = e^{-1-a-b\epsilon_\nu}. \quad (57)$$

The constants $a$ and $b$ follow from the two conditions (55). Returning to the original coordinates and setting $C = e^{-1-a}$ we get the Maxwell-Boltzmann distribution

$$p(r, \mathbf{v}) = Ce^{-\frac{1}{2} bm\mathbf{v}^2 - bV_{\text{ext}}(r)}, \quad (58)$$

which is the joint distribution of the velocity and the position coordinates of a single particle in a gas of $N$ noninteracting ones. For $V_{\text{ext}} = 0$ it is the well-known Maxwell distribution of the particle velocities. It appears that $b$ has the interpretation of $1/k_B T$ (see next chapter).

II.6.5. Continuum limit

Let $p(x)$ be a probability distribution of a continuous variable $x$, taken here one-dimensional. Can we associate an entropy with this distribution just as we did with the distribution $p_i$?
Suppose we partition the \( x \) axis into intervals \( \Delta x \) that we label by an index \( i \). Then \( p_i = p(x_i)\Delta x \) is the probability content of the \( i \)th interval, supposed to be located around \( x_i \). This leads to an entropy

\[
S = -k_B \sum_i [p(x_i)\Delta x] \log[p(x_i)\Delta x]
\]

\[
= -k_B \int dx p(x) \log p(x) - k_B \log \Delta x,
\]

where in the second line we have taken the limit \( \Delta x \to 0 \) and now see that this leads to a diverging additive constant \( -\log \Delta x \). The first term in (59) is often retained as the definition of the entropy associated with the distribution of a continuous variable. As long as we are interested in differences of such entropies, the infinite constant cancels out and plays no role. Of course, if \( x \) is a physical variable and we know how to treat it quantum mechanically, this will provide a physical value for \( \Delta x \).

II.6.6. Time dependent phenomena and \( H \) theorem

It is not evident, whether in thermodynamics or in statistical mechanics, how the entropy should be defined for a time dependent system. However, if one accepts the “\( \sum p \log p \)” formula, then time dependent probabilities may be substituted:

\[
S(t) = -k_B \sum_i p_i(t) \log p_i(t).
\]

The time evolution of \( S(t) \) depends on the time evolution equations that one has derived or postulated for the \( p_i(t) \). For several types of equations\(^{15}\) one may show that in isolated systems

\[
\frac{dS(t)}{dt} \geq 0,
\]

with the equality sign holding only in equilibrium. Such a relation is called a Boltzmann \( H \) theorem. However, time dependent phenomena are outside the scope of this course.

\(^{15}\)Among which first of all the Boltzmann equation (see elsewhere).
III. CANONICAL ENSEMBLE

III.1. The ensemble

The microcanonical ensemble applies to systems whose energy $E$ is strictly constant, that is, to systems that are perfectly isolated from their environment.

The canonical ensemble applies to systems that are in thermal contact with their environment. That is to say, they may release or absorb energy (heat). We will consider the idealized case of an environment having an infinite heat capacity, so that its temperature $T$ cannot change. Such an idealized system is called a thermostat.

The energy $E$ of a system in contact with a thermostat will fluctuate, even at equilibrium, around some average value $\langle E \rangle$ determined by the properties of the system itself and the temperature $T$ of the thermostat; the latter will also be the temperature of the system. This is why the canonical ensemble is governed by a new independent parameter, the temperature $T$, which replaces the energy $E$ of the microcanonical ensemble.

The canonical ensemble is defined by the following choice of the density $\rho(\Gamma)$,

$$\rho_c(\Gamma) = \frac{1}{N!h^{3N}} e^{-\beta H},$$

in which the exponential is called the Boltzmann factor. Here $\beta \equiv 1/k_B T$ is a parameter about which we make no a priori assumptions. The canonical distribution may be viewed as a sum of juxtaposed microcanonical distributions, the one on $(E,E+\delta E)$ being weighted by the factor $e^{-\beta E}$. The canonical partition function $Z_c$ (for a classical system) is therefore

$$Z_c(T,V,N) = \frac{1}{N!h^{3N}} \int d\Gamma e^{-\beta H},$$

in which the integration is over all of phase space. It depends on the triplet of independent variables

$$(T,V,N).$$

We may relate $Z_c$ to $Z_m$ by rewriting (2) as

$$Z_c(T,V,N) = \int \frac{dE}{\delta E} e^{-\beta E} Z_m(E,V,N) = \int dE e^{-\beta E} \Phi'(E),$$

which is the Laplace transform (see Appendix D) of the density of states $\Phi'(E)$.

For a quantum mechanical system the phase space integral in (4) is of course replaced with a sum on all eigenstates of the system. If $\ell$ is such a state and if we denote its energy by $E_\ell$, then

$$Z_c = \sum_\ell e^{-\beta E_\ell} = \text{tr } e^{-\beta H}.$$

Calculating this sum requires knowing the energy levels, but does not require knowing the corresponding quantum mechanical states. Note that in the sum (5) the same energy level will occur as many times as it is degenerate. It follows that the probability for the system to be in eigenstate $\ell$ is

$$P_c(\ell) = \frac{1}{Z_c} e^{-\beta E_\ell}.$$
III.2. Helmholtz free energy

III.2.1. Legendre transformation between $S_m$ and $F_c$

The thermodynamic potential associated with $Z_c$ is the (canonical) free energy\footnote{It is also called the Helmholtz free energy. Another symbol in use is $A$ instead of $F$.} $F_c$ defined by

$$Z_c = e^{-\beta F_c} \iff F_c(T, V, N) = -k_B T \log Z_c(T, V, N).$$

If $F_c$ is extensive as we expect it to be, then we must have

$$F_c(T, V, N) = N f(T, v), \quad N \to \infty,$$

where as before $v = V/N$.

Can we establish a relation between this canonical thermodynamic potential and the micro-canonical one? We will do this now. For similar reasons as before, the integral on $E$ occurring in (7) is again strongly peaked when $N$ is large. Let the maximum occur at $E = E_c$; this will then also be the average value of $E$ in the canonical ensemble. Using that $Z_m = \exp(k_B^{-1} S_m)$ we then have from (4) (for simplicity not indicating the dependence on $V$ and $N$)

$$Z_c(T) = \int \frac{dE}{\delta E} e^{-\beta E} Z_m(E)$$

$$= \max_E e^{-\beta E + k_B^{-1} S_m(E)}$$

$$= e^{-\beta E_c + k_B^{-1} S_m(E_c)},$$

(9)
where in the second step we have neglected all subextensive terms in the exponential. It follows that

\[ F_c = E_c - TS_m. \]  

This relation between \( F_c \) and \( S_m \) is called a Legendre transformation. It is a kind of transformation that makes sense especially for convex functions, as we know \( S_m(E) \) is. Indeed, the relation \( F = E - TS \) is well-known in thermodynamics.\(^{17}\)

The value \( E_c(E,V,N) \) is the solution of

\[
\left[ -\beta E + S_m(E) \right] \bigg|_{E=E_c} = 0,
\]

from which we see directly that the parameter \( \beta \equiv 1/k_B T \) of the canonical ensemble is given by

\[ T = T_m. \]

When its natural variables \( T, V, N \) undergo variations \( dT, dV, \) and \( dN \), the resulting infinitesimal change in \( F_c \) is

\[
dF_c = \left( \frac{\partial F_c}{\partial T} \right)_{V,N} dT + \left( \frac{\partial F_c}{\partial V} \right)_{T,N} dV + \left( \frac{\partial F_c}{\partial N} \right)_{T,V} dN.
\]

In this equation we have called the three partial derivatives that appear by the names \(-S_c, -p_c, \) and \( \mu_c \). To show that they actually do correspond to the thermodynamic entropy, pressure, and chemical potential we do not any more this time have to appeal to thermodynamics. In fact, the three derivatives may be calculated directly in terms of the microcanonical quantities that we know from chapter II, and we find that when (12) is satisfied

\[ S_c = S_m, \quad p_c = p_m, \quad \mu_c = \mu_m. \]

This shows that the canonical and microcanonical ensemble lead to the same thermodynamics.

**Proof.** (Done in class.)

III.2.2. *Other thermodynamic quantities*

It is easy to show (done in class) that the average energy of a system described by the canonical ensemble may be obtained as

\[
\langle E \rangle_c = -\frac{\partial \log Z_c}{\partial \beta}.
\]

Because of the peaked nature of the energy distribution we must have \( \langle E \rangle_c = E_c \). One shows furthermore that, with \( \Delta E \equiv E - E_c \),

\[
\langle \Delta E^2 \rangle_c \equiv \langle E^2 \rangle_c - \langle E \rangle_c^2 = \frac{\partial^2 \log Z_c}{\partial \beta^2}.
\]

\(^{17}\)Another frequently used notation is \( F = U - TS \).
Since log $Z_c$ is extensive, this implies that the typical deviation $\Delta E$ from its average $\langle E \rangle_c$ is proportional to $N^{\frac{1}{2}}$. Hence the relative fluctuations $\Delta E / \langle E \rangle_c$ of the total energy are vanishingly small in the thermodynamic limit.

**Remark.** Relation (16) is not surprising from the following perspective. If in the last line of Eq. (4) we consider $\Phi(E)$ as an (unnormalized) probability distribution on the $E$ axis, then $Z_c(T,V,N)$ is its generating function; here $-\beta$ plays the role of $ik$ in (A.26). It follows that log $Z_c(T,V,N)$ is the moment generating function of the energy variable, whence (16).

But we also have

$$C_{V,c} = \frac{\partial \langle E \rangle_c}{\partial T} = -k_B \beta^2 \frac{\partial \langle E \rangle_c}{\partial \beta} = k_B \beta^2 \frac{\partial^2 \log Z_c}{\partial \beta^2}. \quad (17)$$

Comparison of (17) and (16) establishes a proportionality between the heat capacity $C_{V,c}$ and the fluctuations of the energy,

$$C_{V,c} = k_B \beta^2 \langle \Delta E^2 \rangle_c. \quad (18)$$

By the way, notice the conversion between $T$ and $\beta$ derivatives,

$$\frac{\partial}{\partial T} = -k_B \beta^2 \frac{\partial}{\partial \beta}, \quad \frac{\partial}{\partial \beta} = -k_B T^2 \frac{\partial}{\partial T}. \quad (19)$$

### III.3. Derivation of the canonical ensemble

We consider a microcanonical system $\mathcal{T}$ of $N$ particles and volume $V$, and a subsystem $\mathcal{S}$ of $N_1$ particles and volume $V_1$ that may exchange heat with $\mathcal{T}$. We will now do something different from what we did in section II.5.1, where we let $N$ and $N_1$ tend to infinity in a fixed ratio. Here we let $N \to \infty$ while keeping $N_1$ fixed (which does not exclude that at a later stage we may let $N_1$ become very large). We will show that $\mathcal{T}$ then acts as a thermostat for $\mathcal{S}$ and that $\mathcal{S}$ is described by the canonical ensemble. To do so we write the microcanonical partition function of the total system as

$$Z_m(E,V,N) = \int \frac{dE}{\delta E} e^{\mathcal{O}(N_1)} Z_m(E_1,V_1,N_1) Z_m(E - E_1,N - N_1,V - V_1) e^{\mathcal{O}(N)}$$

$$= e^{k_B^{-1} S_m(E,V-V_1,N-N_1)} \int \frac{dE}{\delta E} Z_m(E_1,V_1,N_1) e^{-k_B^{-1} E_1 \frac{\partial S_m}{\partial E}} + \mathcal{O}(N^2/N)$$

$$= e^{k_B^{-1} S_m(E,V-V_1,N-N_1)} \int \frac{dE_1}{\delta E} Z_m(E_1,V_1,N_1) e^{-\beta_m E_1} Z_{c}(T_m,V_1,N_1) \quad (20)$$

where in the last line contributions of $\mathcal{O}(N^2/N)$ in the exponent have been neglected; and where we see that the parameter $\beta = 1/k_B T$ of the canonical ensemble appears to have the value $\beta_m = 1/k_B T_m = k_B^{-1} \frac{\partial S_m}{\partial E}$ of the thermostat.

**Remarks.**

The interaction Hamiltonian between the system $\mathcal{S}$ and the thermostat $\mathcal{T}$ has been neglected. We may now let $N_1 \to \infty$ if we wish (but need not).
III.4. Example: the ideal gas

We consider $N$ noninteracting particles of mass $m$ in a volume $V$. The Hamiltonian is

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m}. \tag{21}$$

The canonical partition function becomes

$$Z_c(T, V, N) = \frac{1}{N! h^{3N}} \int d\Gamma e^{-\beta\mathcal{H}(\Gamma)}$$

$$= \frac{V^N}{N! h^{3N}} \int dp_1 \cdots \int dp_N \; e^{-\beta(p_1^2 + \cdots + p_N^2)/2m}$$

$$= \frac{V^N}{N! h^{3N}} \left( \frac{2\pi m}{\beta h^2} \right)^{\frac{3N}{2}}$$

$$= \frac{1}{N!} \left( \frac{V}{\lambda^3} \right)^N \tag{22}$$

with as before $\lambda = h/\sqrt{2\pi m k_B T}$. Note that $V/\lambda^3$ is just the single-particle canonical partition function. Use Stirling’s formula and neglect nonextensive terms. Then (22) leads to

$$F_c(T, V, N) = N k_B T \left[ \log n \lambda^3 - 1 \right]. \tag{23}$$

Of course the same reservations concerning its validity hold as in the microcanonical case (section II.4.2). We derive from (23) that for the ideal gas

$$E_c = -\frac{\partial \log Z_c(T, V, N)}{\partial \beta} = \frac{3}{2} N k_B T,$$

$$S_c = -\left( \frac{\partial F_c(T, V, N)}{\partial T} \right)_{V, N} = -N k_B \left[ \log n \lambda^3 - \frac{5}{2} \right],$$

$$p_c = -\left( \frac{\partial F_c(T, V, N)}{\partial V} \right)_{T, N} = \frac{N k_B T}{V} = n k_B T. \tag{24}$$

All these relations agree with the microcanonical ones and with thermodynamics. All further thermodynamic quantities may be obtained. We present again the example of the heat capacity $C_V = \partial E_c / \partial T$, which is related to the thermodynamic potential $F_c(T, V, N)$ by

$$C_V = k_B \beta^2 \frac{\partial^2}{\partial \beta^2} \beta F_c, \tag{25}$$

and which you may compare to the microcanonical expression (II.24). The indices “m” and “c” indicate only how the expression for a quantity has been derived, but do not alter its expression; to keep things simple we have dropped the index “c” on $C_V$. 

34
III.5. Independent particles

Suppose we have a system of \( N \) classical identical noninteracting particles. Let their Hamiltonian be

\[
\mathcal{H} = \sum_{j=1}^{N} \mathcal{H}_j, \quad \mathcal{H}_j = h(r_j, p_j, \ldots),
\]

(26)

where the dots in \( h \) stand for any further internal degrees of freedom that particle \( j \) may have.

We may distinguish two cases.

Case (a): each particle is confined to the neighborhood of a specific position in space (a “site”). The particles are distinguishable. This happens, for example, for an atom at a fixed position in a crystal. The partition function should then be defined without the factor \( 1/N! \) and we get,

\[
Z_c(T, V, N) = \frac{1}{h^{3N}} \int \text{d}r_1 \ldots \text{d}r_N \text{d}p_1 \ldots \text{d}p_N e^{-\beta \mathcal{H}}
= \left[ \frac{1}{h^3} \int \text{d}r \text{d}p \ldots e^{-\beta h(r, p, \ldots)} \right]^N
= z^N
\]

(27)

in which \( z \) is called the single-particle (canonical) partition function. The free energy becomes

\[
F_c(T, V, N) = -Nk_B T \log z.
\]

(28)

Clearly \( Z_c \) is the product of \( N \) single-particle partition functions and \( F_c \) is the sum of \( N \) single-particle free energies. Clearly, \( F_c \) is extensive if \( z \) is intensive.

Case (b): each particle may move through the whole volume. The particles are indistinguishable. The canonical partition function then is, with the same expression for \( z \) as before,

\[
Z_c(T, V, N) = \frac{z^N}{N!}.
\]

(29)

The free energy becomes

\[
F_c(T, V, N) = -Nk_B T \log \frac{z}{N} - Nk_B T
\]

(30)

and in order that it be properly extensive we must have that \( z \sim N \) in the thermodynamic limit; which will be true exactly because the integration on the volume will cause a factor \( V \) to appear in \( z \). The ideal gas of the previous section is the simplest example of this case; we will see another one in the next section.

Quantum mechanics.

Case (a) follows straightforwardly from the underlying quantum mechanics. If \( \lambda \) labels the eigenstates of the single-particle Hamiltonian \( h \) in (26) and \( \epsilon_\lambda \) stands for their energies, then \( z = \sum_\lambda \exp(-\beta \epsilon_\lambda) \) and \( Z_c \) is the product of \( N \) factors \( z \).

Case (b) arises from the underlying quantum mechanics in a more complicated way. The fundamental reason is that identical particles that occupy the same volume are indistinguishable and have to be treated according to quantum statistics. The quantum case will be dealt with in detail in chapters VII and VIII.
III.6. Example: independent particles in a gravitational field

This is an example of case (b) of the previous section. Let \( N \) noninteracting particles of coordinates \( \mathbf{r}_j = (x_j, y_j, z_j) \), momenta \( \mathbf{p}_j \), and mass \( m \) be placed in a gravitational field of acceleration \( g \), in equilibrium at a temperature \( T \). This is an approximate model for the earth’s atmosphere.

We study a system in which we confine the particles to a half-way infinite vertical column of cross section \( A = L^2 \). To be specific, we let \( 0 < x_j, y_j < L \) and \( z_j > 0 \). The system Hamiltonian then is

\[
\mathcal{H} = \sum_{j=1}^{N} \frac{p_{j}^2}{2m} + \sum_{j=1}^{N} mgz_j , \tag{31}
\]

which is of the form (26). We easily calculate

\[
z = \frac{1}{h^{3N}} \int dp_1 e^{-\frac{\beta p_1^2}{2m}} \int_0^L dx_1 \int_0^L dy_1 \int_0^\infty dz_1 e^{-\beta mgz_1} = A \times \left( \frac{\sqrt{2\pi mk_BT}}{h} \right)^3 \times \frac{1}{\beta mg} , \tag{32}
\]

whence, see (29),

\[
Z_c(T, A, N) = \frac{1}{N!} \left( \frac{\sqrt{2\pi mk_BT}}{h} \right)^{3N} \left( \frac{A}{\beta mg} \right) . \tag{33}
\]

In the thermodynamic limit \( N \to \infty \) the free energy \( F_c = -k_BT \log Z_c \) becomes

\[
F_c(T, A, N) = Nk_BT \left[ \log (n_2\lambda^2 \cdot \lambda \beta mg) - 1 \right] , \tag{34}
\]

where we have used Stirling’s formula and taken the limit \( N \to \infty \) at fixed \( n_2 \equiv N/L^2 = N/A \). Interpretation: \( n_2 \) is the average number of particles on the earth’s surface contained in a column of unit cross section.

**Exercise.**

a. Derive the probability distribution \( p(z_1) \) of the height \( z_1 \) of particle 1 as a marginal distribution of the single-particle Boltzmann distribution. Calculate the average height \( \langle z_1 \rangle \).

b. Interpret the factors \( n_2\lambda^2 \) and \( \lambda \beta mg \) that occur inside the logarithm in (34).

**Exercise.**

a. Calculate the average particle density \( \langle n(z) \rangle \) at height \( z \) above the earth’s surface.

b. Interpret the quantity \(- (\partial F_c(T, A, N)/\partial A)_{T,N} \). Show that it may be cast in the form

\[
- \left( \frac{\partial F_c}{\partial A} \right)_{T,N} = \int_0^\infty dz \ n(z)k_BT . \tag{35}
\]

Let \( p(z) \) stand for the atmospheric pressure at altitude \( z \). Argue that (if the hypotheses of these exercises are right) the ideal gas law holds locally in the atmosphere. Find an explicit expression for \( p(z) \) (called the barometric height formula).
Exercise. Estimate the thickness of the earth’s atmosphere. Some data: \( m_{N_2} = 4.65 \times 10^{-26} \text{kg} \), \( g = 9.8 \text{ms}^{-2} \), \( T = -50^\circ \text{C} \), \( k_B T = 1 \text{eV} \Leftrightarrow T = 11600 \text{K} \).

III.7. Example: one-dimensional gas of hard rods

(Done in class)

III.8. Magnetism

Introduction. Analogy of \((p,V)\) to \((M,B)\). Elaborate.

III.8.1. General

Let \( \vec{S}_1, \vec{S}_2, \ldots, \vec{S}_N \), be a collection of \( N \) spins, such that \( \vec{S}_j \equiv (S^x_j, S^y_j, S^z_j) \) is attached to site \( j \) of a crystal lattice. Let \( \mathcal{H}_0(\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_N) \) be the system Hamiltonian; it describes the interaction between the spins. Typically, there will be an interaction between spin pairs on nearest neighbor sites and weaker interactions between spins further apart. When the crystal is placed in an external magnetic field \( \vec{B} \), the total Hamiltonian becomes

\[
\mathcal{H} = \mathcal{H}_0 - \gamma \mu_B \sum_{j=1}^{N} \vec{B} \cdot \vec{S}_j
\]

in which the gyromagnetic factor \( \gamma \) and the Bohr magneton \( \mu_B \) are the quantities you know from quantum mechanics, and \( \vec{M} \) is the total magnetization. It is always possible to rotate the coordinate system such that we get \( \vec{B} = (0,0,B) \) and the magnetic part of the Hamiltonian becomes \( -\gamma \mu_B B \sum_j S^z_j \).

Since spin is a quantum phenomenon, the partition function associated with \( \mathcal{H} \) involves a sum on its discrete eigenstates, which in the general case are impossible to find (this would require diagonalizing \( \mathcal{H} \), i.e., solving the Schrödinger equation \( \mathcal{H}|\psi\rangle = \epsilon|\psi\rangle \)). The crystal has, nevertheless, a canonical partition function \( Z_c(T,B,N) \) and a free energy \( F_c(T,B,N) = -k_B T \log Z_c(T,B,N) \) that derives from it.

Replacement of \((V,-p)\) by \((B,M)\).

We interrupt here the general development in order to first study now a simple special case.

III.8.2. Example 1: Paramagnetic crystal

Let us take for the \( \vec{S}_j \) spins \( 1/2 \) that are independent i.e., \( \mathcal{H}_0 = 0 \). The \( z \) component \( S^z_i \) of each spin may take the values \( \pm \frac{1}{2} \hbar \). We will choose new dimensionless units such that these \( z \) components are represented by the variable \( s_j = \pm 1 \). and we will denote the associated magnetic moment by \( \mu s_j \) (we have \( \mu = \frac{1}{2} \gamma \mu_B \hbar \)). When placed in a magnetic field \( B \) the energy of spin...

---

\(^{18}\)We are discussing here the magnetic partition function, tacitly supposing that the spins are not coupled to the other degrees of freedom of the crystal.

\(^{19}\)Do not confuse this \( \mu \) with the chemical potential, denoted elsewhere by the same symbol.
\( i \) will be \(-\mu B s_i\). The Hamiltonian of this system of independent spins therefore is
\[
\mathcal{H} = -\mu B \sum_{i=1}^{N} s_i.
\] (37)

A microstate of the system may be specified by \( s \equiv (s_1, s_2, \ldots, s_N) \) (how many microstates are there?). It is easy to show that each microstate is an eigenstate of \( \mathcal{H} \), so the diagonalization problem that renders the general case difficult, has disappeared. Hence the Hamiltonian (37) is to magnetism what the ideal gas Hamiltonian is to fluids.

A. Canonical calculation.

The canonical partition function at inverse temperature \( \beta \) is
\[
Z_c(T, B, N) = \sum_{s} e^{-\beta \mathcal{H}}.
\] (38)

Its calculation is very easy and runs as follows.

\[
Z_c(T, B, N) = \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1} e^{\beta \mu B \sum_{i=1}^{N} s_i} = \left[ \sum_{s_1=\pm 1} e^{\beta \mu B s_1} \right]^{N} = \left( \frac{2 \cosh(\beta \mu B)}{z} \right)^{N},
\] (39)
in which now \( z = 2 \cosh(\beta \mu B) \) is the single-particle partition function. The free energy is
\[
F_c(T, B, N) = -N k_B T \log z = -N k_B T \log \cosh \beta \mu B.
\] (40)

The average energy in the canonical ensemble is now obtained as
\[
\langle E \rangle_c = -\frac{\partial \log Z_c}{\partial \beta} = \ldots = -N \mu B \tanh \beta \mu B.
\] (41)

Another quantity of great interest is the total magnetization \( M = \mu \sum_{i=1}^{N} s_i \), or equivalently the magnetization per spin \( m = M/N \). It so happens that in this system the total energy and magnetization are the same variable, up to the constant \( \mu B \); this is a peculiarity of this too simple model and not the general situation. Using the result (41) we therefore get
\[
\langle m \rangle_c = \mu \tanh \beta \mu B
\] (42)
whence the magnetic susceptibility
\[
\chi(B) \equiv \frac{\partial \langle m \rangle_c}{\partial B} = \beta \mu^2 (1 - \tanh^2 \beta \mu B).
\] (43)

The zero-field susceptibility is given by \( \chi(0) = \beta \mu^2 \) and diverges for \( T \to 0 \).
Exercise. Take the limit $T \to 0$ at fixed $B \neq 0$ and explain your finding intuitively.

B. Microcanonical calculation.

It is instructive to see now how we would we deal with this system in the microcanonical ensemble. The microcanonical partition function is $Z_m(E, B, N)$ in which the energy has a prescribed value, say $E$. The microcanonical temperature $T_m$ is then calculated according to (see the problem session)

$$\frac{1}{T_m} = \frac{\partial S_m}{\partial E} = \ldots = -\frac{k_B}{2\mu B} \log \frac{1 + \frac{e}{\mu B}}{1 - \frac{e}{\mu B}}$$

$$= -\frac{k_B}{\mu B} \text{artanh} \frac{e}{\mu B}. \quad (44)$$

You now see that if we identify $E = \langle E \rangle_c$ and $T_m = T$, equations (41) and (44) are identical.

Exercise. Calculate the entropies $S_c$ and $S_{rmm}$ of this system and show that they are identical.

III.8.3. General (continued)

Magnetization and magnetic susceptibility

Having obtained the canonical partition function (or equivalently, the free energy) we easily find the average value of the total magnetization $M = \mu \sum_{i=1}^{N} s_i$ by carrying out a derivation with respect to $B$,

$$\langle M \rangle = k_B T \frac{\partial \log Z_c}{\partial B}$$

$$= -\frac{\partial F_c}{\partial B}. \quad (45)$$

Deriving $\log Z_c$ once more in the same way gives

$$(k_B T)^2 \frac{\partial^2 \log Z_c}{\partial B^2} = \ldots$$

$$= \langle M^2 \rangle - \langle M \rangle^2$$

$$\equiv \langle \Delta M^2 \rangle, \quad (46)$$

a relation to which we will return in a moment.

The magnetic susceptibility

$$\chi \equiv \frac{1}{N} \frac{\partial \langle M \rangle}{\partial B} \quad (47)$$

measures how easily the system responds to a change in the magnetic field. Note that $\chi$ has been defined such as to be intensive. From the two preceding equations we have

$$\chi = \frac{k_B T}{N} \frac{\partial^2 \log Z_c}{\partial B^2}. \quad (48)$$

Upon comparing relations (47) and (48) we find that

$$\langle \Delta M^2 \rangle = N k_B T \chi. \quad (49)$$
It is worth reflecting a moment on the interpretation of this equation. It establishes a relation between the fluctuations of a physical quantity (in casu the magnetization) in a state of equilibrium and the way this equilibrium is modified when one of the system parameters (in casu $B$) is changed. Note that we have encountered a similar relation earlier, viz. Eq. (18).

**Remark.** Nowadays some people refer to equations of the type (18) and (49) as fluctuation-dissipation relations. This name, however, is traditionally reserved for the time-dependent extensions of these relations, not discussed in this course, and whose proofs are highly nontrivial.

### III.8.4. Example 2: Ising model

Let us consider a system of electrons $j$, each one carrying a spin $\vec{S}_j = (S^x_j, S^y_j, S^z_j)$ and attached to an atom occupying a site of a crystal. Heisenberg (1928) supposed that in a real crystal the interaction between the electron spins is in good approximation limited to nearest-neighbor electrons. A general expression for the energy of such a spin system then is $i$ and $j$ is

$$H_{\text{Heis}} = -\sum_{\langle i,j \rangle} \sum_{\alpha} J_{\alpha} S^\alpha_i S^\alpha_j,$$

(50)

known today as the (anisotropic) Heisenberg Hamiltonian, and in which $\langle i,j \rangle$ denotes a pair of nearest-neighbor sites. Here the $J_{\alpha}$ are so-called exchange energies that depend on the specifics of the crystal; they must be calculated quantum mechanically. In statistical physics these energies are usually supposed to be given constants.

In case all $J_{\alpha}$ are equal, we speak of the isotropic Heisenberg interaction; in case $J_z = 0$, we speak of an “XY” interaction; and in case $J_x = J_y = 0$ we speak of an Ising interaction.

Consider now a crystal of $N$ spins with an Ising interaction $J_z$ between nearest neighbors. We introduce the spin variables $s_i = \pm 1$ with $i = 1, 2, \ldots, N$ as in section III.8.2 and set $J = (\hbar/2)^2 J_z$. Then the Hamiltonian of the electron spin system may be written as

$$H_{\text{Ising}} = -J \sum_{\langle i,j \rangle} s_i s_j.$$

(51)

The Ising Hamiltonian has the special property that all “operators” commute and its spins may therefore be treated as classical (although discrete) variables.

The Ising model has been studied for almost a century now, which is why it is mentioned here. It will not be discussed any further in this course, however.

### III.9. Energy distribution in the canonical ensemble

What is the probability density $p(E)$ that a system described by the canonical ensemble at a temperature $T$ have an energy $E$? We have seen above [Eqs. (15) and (16)] that the mean $\langle E \rangle_c$ and variance $\langle E^2 \rangle_c - \langle E \rangle_c^2$ of $E$ follow form the partition function $Z_c$. If we knew that $p(E)$ was a Gaussian, then it would be fixed completely by its mean and variance. We will now show that indeed $p(E)$ is a Gaussian.

---

20This spin-spin interaction originates from the electrostatic Coulomb interaction between the electron charges, combined with Pauli’s exclusion principle. The magnetic interaction between two electron magnetic moments $\gamma \mu_B \vec{S}_i$ and $\gamma \mu_B \vec{S}_j$ is very much weaker.
We start from Eq. (4) and refine the analysis that led us to Eq. (9). Upon expanding the microcanonical entropy $S_m(E)$ about $E = E_c$ to second order we get

$$Z_c(T) = \text{e}^{-\beta E_c + k_B^{-1} S_m(E_c)} \int \frac{dE}{\delta E} \exp \left( \frac{1}{2} k_B^{-1} (E - E_c)^2 \frac{\partial^2 S_m(E)}{\partial E^2} \bigg|_{E_c} + \frac{1}{6} k_B^{-1} (E - E_c)^3 \frac{\partial^3 S_m(E)}{\partial E^3} \bigg|_{E_c} + \ldots \right)$$

(52)

We will now do this properly. We are interested in $p(E)$ in the limit $N \to \infty$ and will therefore make all $N$ dependence under the integral sign in (52) explicit. The $k$th $E$-derivative of $k_B^{-1} S_m$ is proportional to $N^{1-k}$ (why?) and we will write it for this occasion as $N^{1-k}s(k)$. Let us scale the energy difference with respect to $E_c$ as $\Delta E \equiv E - E_c = N^{1/2}/\epsilon$. Then we may rewrite (52) as

$$Z_c(T) = \text{cst} \times \frac{1}{N^{1/2} \delta E} \int d\epsilon \exp \left( \frac{1}{2} \epsilon^2 s^{(2)} + \frac{1}{6} N^{-1/2} \epsilon^3 s^{(3)} + O(N^{-1}) \right)$$

(53)

in which “cst” stands for an expression independent of $E$. It is now clear that for $N \to \infty$ the third and higher order terms in the exponent vanish and that integrand draws its contribution from $\epsilon$ values that are typically of order $N^0$. The expression under the integral tends towards

$$\exp \left( \frac{1}{2} \epsilon^2 s^{(2)} \right) = \exp \left( \frac{1}{2} k_B^{-1} (E - E_c)^2 \frac{\partial^2 S_m(E)}{\partial E^2} \bigg|_{E_c} \right)$$

(54)

where we have restored the original variables. Eq. (54) represents, up to a normalization factor, the distribution of the energy $E$ in the canonical ensemble.

The method applied here is a special case of our general discussion in Appendix B.

**Exercise.** Find $p(E)$ including its normalization. Using (54) and your knowledge of the microcanonical ensemble, show that the fluctuations of the energy are given by

$$\langle \Delta E^2 \rangle_c = k_B T_m^2 C_{V,m}. \quad (55)$$

Compare to (18) and comment.

### III.10. Marginal laws of the canonical distribution

Maxwell distribution for the velocity of a particle.
Barometric height formula.
Joint distribution for $N$ or $k$ particle positions.

### III.11. General pair of conjugate variables

Let $H(\Gamma) = H_0(\Gamma) - \phi X(\Gamma)$ in which $H_0$ is an “unperturbed” (but otherwise arbitrary) Hamiltonian, and $\phi$ a control parameter that couples to a physical quantity $X(\Gamma)$, also arbitrary in the present discussion. In practice $X$ will be the sum of some intensive quantity $x$ on the volume of the system, and will therefore be extensive.

Let $Z_c(T, V, \phi, N)$ be the canonical partition function of this system. It is easy to show that the average of $X$ is given by

$$\langle X \rangle_c = k_B T \frac{\partial \log Z_c}{\partial \phi}. \quad (56)$$
The “susceptibility” \( \chi_\phi \) is the per particle increase of this average for a unit increase of the control parameter, that is
\[
\chi_\phi \equiv \frac{1}{N} \frac{\partial \langle X \rangle_c}{\partial \phi}. \tag{57}
\]
One derives almost directly
\[
\langle \Delta X^2 \rangle_c = N k_B T \chi_\phi. \tag{58}
\]
This generalizes earlier relations of such type found for \( X \) equal to the energy and the magnetization.

III.12. Canonical systems in contact

III.12.1. Equilibrium condition

Consider two canonical systems, 1 and 2, at equilibrium at a temperature \( T \), and characterized by variables \( V_1, N_1, \ldots \) and \( V_2, N_2, \ldots \), respectively. We may see these variables as “constraints.” When the two systems are brought into contact such as to be able to exchange volume (by means of a piston), the constraints \( V_1 \) and \( V_2 \) are released and replaced with the single constraint \( V = V_1 + V_2 \). When they are allowed to exchange particles (through a permeable wall) the constants \( N_1 \) and \( N_2 \) are released and replaced by the single constraint \( N = N_1 + N_2 \). And so on.

In the general discussion below we will consider extensive constraints \( X_1 \) and \( X_2 \) of arbitrary type and let \( X = X_1 + X_2 \). Before the contact is established the systems have free energies \( F_1(T, X_1) \) and \( F_2(T, X_2) \) and the partition function of the combined system is\(^{21}\)
\[
Z_{c,1}(T, X_1)Z_{c,2}(T, X_2) = e^{-\beta F_1(T, X_1) - \beta F_2(T, X_2)}. \tag{59}
\]
After release of the constraint (and once the new equilibrium has gotten established) the partition

\(^{21}\)Any possible interaction terms between the two systems are assumed to be negligible.
function is
\[
Z_{c,12}(T, X) = \sum_{X} e^{-\beta F_1(T, X) - \beta F_2(T, X - \bar{X})}
= \sum_{\bar{X}} e^{-\beta \tilde{F}_{12}(T, \bar{X}; \bar{X})}
\]  
(60)

(if \( \bar{X} \) may vary continuously, we will have an integral instead of a sum). Here \( \tilde{F}_{12} \) is the free energy when the two systems are constrained to have values \( X_1 = \bar{X} \) and \( X_2 = X - \bar{X} \). The total partition function is a sum on all possible values of the constraint \( \bar{X} \).

We apply once more the usual argument: \( \tilde{F} \) is extensive, that is, it is of the form \( \tilde{F}_{12} = N \tilde{f}(T, \frac{\bar{X}}{N}, \frac{\bar{X}}{N}) \). Hence the sum on \( \bar{X} \) will have a sharp (and effectively Gaussian) peak occurring at a maximum \( \bar{X} = X_1^{\text{max}} \) which is the solution of
\[
\frac{\partial \tilde{F}_{12}(T, X; \bar{X})}{\partial \bar{X}} = 0.
\]  
(61)

This equation is the condition that determines the new equilibrium. The free energy \( F_{12}(T, X) \) of the combined system in this equilibrium will be given by
\[
Z_{c,12}(T, X) = e^{-\beta \tilde{F}_{12}(T, X; X_1^{\text{max}})} + o(N)
\]  
(62)

The equilibrium condition (61) becomes a relation between the derivatives of \( F_1 \) and \( F_2 \), namely
\[
\frac{\partial F_1(T, X_1^{\text{max}})}{\partial X_1} = \frac{\partial F_2(T, X - X_1^{\text{max}})}{\partial X_2}
\]  
(63)

This condition may easily be worked out more explicitly for specific cases. We thus have
\[
X = V : \quad p_{c,1} = p_{c,2}
\]
\[
X = N : \quad \mu_{c,1} = \mu_{c,2}
\]
\[
\text{etc.}
\]  
(64)

We draw the following important conclusion:

When in a system in equilibrium in contact with a heat bath a constraint is released, the system evolves towards a new equilibrium state that minimizes the free energy (subject to the remaining constraints).

Note the similarity and the differences with what you know about the microcanonical ensemble (section II.5.6-7).

III.12.2. Fluctuations (once more)

In order that (61) correspond to a maximum in the integrand we need the “stability” condition
\[
\left. \frac{\partial^2 \tilde{F}_{12}}{\partial \bar{X}^2} \right|_{X_1^{\text{max}}} > 0 \implies \left. \frac{\partial^2 F_{12}(T, X)}{\partial X^2} \right| > 0.
\]  
(65)
In case $X = V$ this leads to
\[
\frac{\partial^2 F_{12}}{\partial V^2} = -\left(\frac{\partial p}{\partial V}\right)_{N,T} = V\kappa_T > 0,
\] (66)
in which $\kappa_T$ is the isothermal compressibility.

As before, the (effectively Gaussian) sum on $\tilde{X}$ has a width determined by the second order term in the expansion
\[
\tilde{F}(T, X; X_1) = F_{12}(T, X) + \frac{1}{2}(\tilde{X} - X_1^{\text{max}})^2 \frac{\partial^2 F_{12}(T, X)}{\partial X^2} + \ldots,
\] (67)
which shows that the width is
\[
\langle \Delta X^2 \rangle_c = \frac{1}{\frac{\partial^2 F_{12}}{\partial X^2}}.
\] (68)

**Exercise** Consider the two systems in contact through a permeable wall and let $X = V$. Argue why small (large) fluctuations of the two subvolumes $V_1$ and $V_2$ go hand in hand with a low (high) compressibility.

Write down the probability distribution $p(X)$ of the variable $X$ when the combined system is in equilibrium.

### III.13. The Gibbs or Shannon entropy

In section II.6 we exhibited a definition of the “Gibbs” or “Shannon” entropy ($S_G$ or $S_S$, respectively) not based on thermodynamics but on the probabilities $p_i$ of a set of microstates $i$. Working with a discrete set of microstates, we then showed the relation of $S_G$ to the microcanonical entropy.

Let $E_\ell$ be the energy of microstate $\ell$. In the canonical ensemble we have that $p_\ell = e^{-\beta E_\ell}/Z_c$ and we may try to see what we get when we substitute this in the expression for the Gibbs or Shannon entropy:
\[
S_G = -k_B \sum_\ell p_\ell \log p_\ell
\] (69)
\[
= -k_B \sum_\ell \frac{e^{-\beta E_\ell}}{Z_c} \log \frac{e^{-\beta E_\ell}}{Z_c}
\]
\[
= \ldots
\]
\[
= \frac{1}{T} \sum_\ell \frac{E_\ell e^{-\beta E_\ell}}{Z_c} + k_B \log Z_c \sum_\ell \frac{e^{-\beta E_\ell}}{Z_c}
\] (70)

so that
\[
F_c = \langle E \rangle_c - TS_G.
\] (71)
It follows that the Gibbs-Shannon expression (69) for the entropy, when combined with the canonical microstate probabilities, yields the canonical entropy: $S_G = S_c = S$.

Expression (69), although not directly necessary for statistical thermodynamics, provides an important complementary view of the concept of entropy.
IV. HARMONIC OSCILLATOR

Applications. Discussed in class.

IV.1. The classical harmonic oscillator

The Hamiltonian of a classical one-dimensional oscillator of mass $m$ and angular frequency $\omega$ is

$$\mathcal{H}(x,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2. \quad (1)$$

The calculation of its canonical partition function is straightforward. We have $N = 1$ particles and the volume $V$ plays no role (why?). Hence

$$Z_c(T) = \frac{1}{\hbar} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dp \ e^{-\beta \mathcal{H}(x,p)} = \frac{1}{\hbar} \sqrt{\frac{2\pi m}{\beta}} \sqrt{\frac{2\pi}{\beta m \omega^2}} = \frac{2\pi k_B T}{\hbar \omega}. \quad (2)$$

We deduce from it the free energy,

$$F_c(T) = -k_B T \log Z_c(T) = -k_B T \log \frac{2\pi k_B T}{\hbar \omega}, \quad (3)$$

and the energy

$$\langle E \rangle_c = -\frac{\partial \log Z_c(T)}{\partial \beta} = -\frac{\partial}{\partial \beta} \log \frac{2\pi}{\beta \hbar \omega} = k_B T. \quad (4)$$

From the above and the thermodynamic relation $F_c = \langle E \rangle_c - T S_c$ one finds for the entropy of the oscillator

$$S_c = k_B \left( \log \frac{2\pi k_B T}{\hbar \omega} + 1 \right). \quad (5)$$

We observe that, according to (5), at sufficiently low temperature this entropy will become negative. As discussed in class, this is a sign that the classical calculation is no longer valid.

The heat capacity of this single oscillator, that we will continue to denote by $C_V$, is given by

$$C_V = \frac{\partial \langle E \rangle_c}{\partial T} = k_B. \quad (6)$$

One may also calculate the average kinetic and potential energy, which are given by

$$\langle \frac{1}{2} m v^2 \rangle_c = \frac{1}{2} k_B T, \quad \langle \frac{1}{2} m \omega^2 x^2 \rangle_c = \frac{1}{2} k_B T. \quad (7)$$

The quantity $\langle E \rangle_c$ of Eq. (4) is their sum.
Remark. Eqs. (7) are an example of what is called equipartition (which is short for: equipartition of the energy between the degrees of freedom of the system). The equipartition theorem says that an energy equal to $\frac{1}{2}k_B T$ is stored in every degree of freedom of a classical system that occurs quadratically in the Hamiltonian. One consequence is the well-known expression
\begin{equation}
\langle \frac{1}{2} m v^2 \rangle = \frac{3}{2} k_B T
\end{equation}
for the average kinetic energy of a real-world (three-dimensional) particle.

More generally, as an important but approximate rule of thumb, an energy of the order of the thermal energy $k_B T$ is stored in every classical degree of freedom.

Exercise. Calculate the average kinetic and potential energy of a particle in a gravitational field, that is, described by the Hamiltonian
\begin{equation}
\mathcal{H} = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + mgx_3, \quad x_3 > 0.
\end{equation}
Three-dimensional harmonic oscillator

Its Hamiltonian is

\[ H(r, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 r^2. \tag{10} \]

It is equivalent to three decoupled one-dimensional oscillators and therefore its canonical partition function is the cube of the partition function found above. As a consequence the free energy and its derivatives acquire a factor 3 with respect to those found above.

IV.2. The quantum harmonic oscillator

The quantum harmonic oscillator is described by Hamiltonian (1), but in which \( p \) and \( x \) are operators. This Hamiltonian may be rewritten [see your Quantum Mechanics course] as

\[
H(x, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 \\
= -\frac{\hbar}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 \\
= \ldots \\
= \hbar\omega (a^\dagger a + \frac{1}{2}),
\]

in which \( a^\dagger \) and \( a \) are the usual creation and destruction operators, respectively. The last line shows that \( H \) has the energy levels

\[
\epsilon_n = (n + \frac{1}{2})\hbar\omega, \quad n = 0, 1, 2, \ldots
\]  

The canonical partition function therefore is

\[
Z_c(T) = \sum_{n=0}^{\infty} e^{-\beta\epsilon_n} \\
= e^{-\frac{1}{2}\beta\hbar\omega} \frac{1}{1 - e^{-\beta\hbar\omega}}
\]

[which is also equal to \( 1/(2\sinh \frac{1}{\beta}\hbar\omega) \)]. It follows that

\[
\langle E \rangle_c = -\frac{\partial \log Z_c(T)}{\partial \beta} \\
= \ldots \\
= \frac{\hbar\omega}{2} \frac{1 + e^{-\beta\hbar\omega}}{1 - e^{-\beta\hbar\omega}}
\]

(which is also equal to \( \frac{\hbar\omega}{2} \coth \frac{\hbar\omega}{2} \)). At high temperature the quantum curve approaches the classical curve asymptotically. At low temperature \( \langle E \rangle_c \) tends to the ground state energy, which is nonzero in quantum mechanics. See Fig. 6.

For the specific heat one finds

\[
C_V = \frac{\partial \langle E \rangle_c}{\partial T} \\
= k_B\beta^2 \hbar^2 \omega^2 \frac{e^{\beta\hbar\omega}}{(e^{\beta\hbar\omega} - 1)^2}.
\]

47
We may now consider the two limits

\[ T \to \infty \quad \text{or} \quad \beta \to 0 : \quad C_V \simeq k_B, \]
\[ T \to 0 \quad \text{or} \quad \beta \to \infty : \quad C_V \simeq k_B \left( \frac{\hbar \omega}{k_B T} \right)^2 e^{-\hbar \omega / k_B T}. \]  

(16)

The exponential suppression of \( C_V \) for \( T \to 0 \) is characteristic of an energy spectrum in which a gap separates the ground state from the excited states. One says that for \( k_B T \lesssim \hbar \omega \) the degrees of freedom of the oscillator are “frozen.” This is a quantum effect. See Fig. 7.

### IV.3. The harmonic crystal

#### IV.3.1. The classical crystal

We consider a crystal of \( N \) atoms described by the Hamiltonian

\[ H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + V(r_1, \ldots, r_N). \]  

(17)

Let \( V \) be such that it is minimal when the atoms are located at the lattice positions \( r_i = r_i^0 \) (positions of mechanical equilibrium). We know experimentally that this is what happens. For small deviations of the \( r_i = (r_{ix}, r_{iy}, r_{iz}) \) from their equilibrium values the potential energy becomes, with \( V_0 = V(r_1^0, \ldots, r_N^0) \),

\[ V(r_1, \ldots, r_N) = V_0 + \left. \frac{1}{2} \sum_{i,i'=1}^{N} \frac{\partial^2 V}{\partial r_i \partial r_{i'}} \right|_0 (r_i - r_i^0)(r_i' - r_i^0) + \ldots. \]  

(18)

In order to simplify notation we set \( \iota = (i, \alpha) \) and introduce coordinates \( q_\iota = r_{i \alpha} - r_{i0 \alpha} \) and \( p_\iota = m \dot{q}_\iota \). Upon neglecting the higher order terms in Eq. (18) we obtain the crystal Hamiltonian in the harmonic approximation,

\[ H = \sum_{\iota,\iota'}^{3N} \frac{p_\iota^2}{2m} + \frac{1}{2} m \sum_{\iota,\iota'}^{3N} \Omega^2_{\iota\iota'} q_\iota q_{\iota'}, \]  

(19)

in which the \( 3N \times 3N \) matrix \( \Omega^2 \) is given by \( m \Omega^2_{\iota\iota'} = \left[ \partial^2 V / \partial r_{i \alpha} \partial r_{i' \alpha'} \right]_0 \). Eq. (19) is a Hamiltonian of 3N coupled harmonic oscillators. The matrix may be diagonalized by linearly transforming from the \( p_\iota, q_\iota \) to new variables \( \tilde{p}_\kappa, \tilde{q}_\kappa \), where the index \( \kappa = (k, \lambda) \) consists of a wavevector \( k \) and a polarization index \( \lambda \).\(^{22}\) We then get

\[ H = \sum_{k,\lambda} \left[ \frac{\tilde{p}_{k\lambda}^2}{2m} + \frac{1}{2} m \omega_{k\lambda}^2 \tilde{q}_{k\lambda}^2 \right]. \]  

(20)

The wavevector \( k \) may take any of \( N \) discrete values in a finite region of the \( k \) plane called the first Brillouin zone (see your course on solid state physics); the polarization index \( \lambda \) may take three values (for one longitudinal and two transverse polarizations). Eq. (20) shows therefore that there are 3N independent vibrational modes.

#### IV.3.2. The quantum crystal

\(^{22}\)Do not confuse this \( \lambda \) with a wavelength, often indicated by the same symbol.
The Hamiltonian has the same form as in Eq. (17), but in which \( p_i = \frac{\hbar}{i} \partial/\partial r_i \) and \( r_i \) are operators. The Hamiltonian is diagonalized by the same transformation as in the classical case. In the quantum case this leads to

\[
\mathcal{H} = \sum_{k, \lambda} \hbar \omega_{k\lambda} (a_{k\lambda}^\dagger a_{k\lambda} + \frac{1}{2}),
\]

which is a sum of 3N independent quantum oscillators. The energy levels of the crystal are

\[
\sum_{k, \lambda} (n_{k\lambda} + \frac{1}{2}) \hbar \omega_{k\lambda} \quad \text{with} \quad n_{k\lambda} = 0, 1, 2, \ldots
\]

One says that the system contains \( n_{k\lambda} \) phonons (= quanta of vibrational energy) of type \((k, \lambda)\).

We return to the simplified notation \( \kappa = (k, \lambda) \), where \( \kappa \) runs through 3N distinct values.

The canonical partition function is

\[
Z_c(T) = \prod_{\kappa} \frac{e^{-\frac{1}{2} \beta \hbar \omega_{\kappa}}}{1 - e^{-\beta \hbar \omega_{\kappa}}}. \tag{23}
\]

The average energy is given by

\[
\langle E \rangle_c = \sum_{\kappa} \frac{\hbar \omega_{\kappa} (1 + e^{-\beta \hbar \omega_{\kappa}})}{2} \tag{24}
\]

and the heat capacity by

\[
C_V = k_B \sum_{\kappa} (\beta \hbar \omega_{\kappa})^2 \frac{e^{\beta \hbar \omega_{\kappa}}}{(e^{\beta \hbar \omega_{\kappa}} - 1)^2}. \tag{25}
\]

For \( T \to \infty \) one sees that \( C_V \simeq 3Nk_B \), as for the classical crystal. For \( T \to 0 \) things are slightly more complicated.

One expects that when \( N \) is large, the frequencies \( \omega_{\kappa} \) constitute a quasi-continuum on the frequency axis. Let \( \mathcal{N}(\omega) d\omega \) be the number of frequencies in the interval \([\omega, \omega + d\omega]\); we may write this as

\[
\mathcal{N}(\omega) = \sum_{\kappa} \delta(\omega - \omega_{\kappa}) \tag{26}
\]

[to see this, integrate both sides between \( \omega_0 \) and \( \omega_0 + d\omega \); by the way, what is the physical dimension of \( \mathcal{N} \)?].

Different types of crystals will have different spectra, hence different functions \( \mathcal{N}(\omega) \). We may then write (25) as

\[
C_V = k_B \int_0^\infty d\omega \mathcal{N}(\omega) (\beta \hbar \omega)^2 \frac{e^{\beta \hbar \omega}}{(e^{\beta \hbar \omega} - 1)^2}. \tag{27}
\]

We must of course have

\[
\int_0^\infty d\omega \mathcal{N}(\omega) = 3N. \tag{28}
\]

The true problem is how to find the spectrum and hence \( \mathcal{N}(\omega) \). The details of the heat capacity depend on the details of \( \mathcal{N}(\omega) \), but there are some general features. For \( T \to 0 \) the integral (27)
Figure 8: Debye density of frequencies.

draws its main contribution from the low frequencies. We may then use the fact that for \( \omega \to 0 \) we have

\[
\mathcal{N}(\omega) \simeq C_1 \omega^2 + \ldots,
\]  

(no proof given here) in which the dots stand for higher order terms in \( \omega \). Substitution of (29) in (27) gives

\[
C_V = k_B \int_0^\infty d\omega \left( C_1 \omega^2 + \ldots \right) \frac{(\beta \hbar \omega)^2}{(e^{\beta \hbar \omega} - 1)^2}
\]

\[
\simeq k_B C_1 \frac{2^5}{4} (\beta \hbar)^{-3} \int_0^\infty dx \frac{x^4}{\sinh^2 x}
\]

\[
= k_B \frac{4\pi^4 C_1}{15} \left( \frac{k_B T}{\hbar} \right)^3, \quad (T \to 0).
\]  

In going from the first to the second line we set \( x = \beta \hbar \omega/2 \) and in passing to the third line we used that the integral on \( x \) equals \( \pi^3/30 \). The important conclusion that appears here is that the phonon specific heat goes to zero as \( \sim T^3 \) when \( T \to 0 \), whatever the crystal may be.

We note that \( C_1 \) must be a time-cube in order that everything be dimensionally consistent; however, a more elegant rewriting of (30) is possible, as we will see now.

The Debye model. Debye proposed in 1912 to model a crystal by introducing a cutoff frequency \( \omega_D \) and taking

\[
\mathcal{N}(\omega) = \begin{cases} 
C_1 \omega^2 & \omega < \omega_D, \\
0 & \omega > \omega_D.
\end{cases}
\]  

Substituting this in the normalization condition (28) yields \( C_1 = 9N/\omega_D^3 \) and hence (30) becomes

\[
C_V \simeq Nk_B \frac{12\pi^4}{5} \left( \frac{k_B T}{\hbar \omega_D} \right)^3
\]

\[
= Nk_B \frac{12\pi^4}{5} \left( \frac{T}{T_D} \right)^3, \quad T \to 0,
\]  

(32)
where in the last line we have introduced the \textit{Debye temperature} defined by $\hbar \omega_D = k_B T_D$.

The $T^3$ behavior of the phonon specific heat is indeed observed in experiments. By fitting the data to expression (32) one may characterize each crystal by its Debye temperature. We have, for example, $T_D \approx 170$ K for gold, $T_D \approx 215$ K for silver, and $T_D \approx 400$ K for lead. Put simply, for each crystal this is the temperature below which there are almost no lattice vibrations.

\textbf{Exercise.} Show that the dot terms in (30) lead to corrections having higher powers in $T/T_D$ as $T \to 0$.

\textit{Note.} We will see that the specific heat of the \textit{electrons} is linear in $T$ as $T \to 0$. Hence it dominates the phonon specific heat at low enough temperature.

We will return to the subject of harmonic oscillators in the chapter about photons (= quanta of oscillation of the electromagnetic field).

\textbf{Exercise.} The quantization of (20), which leads to (21), is possible only if the pairs $(p_\iota, q_\iota)$, when viewed as quantum mechanical operators, satisfy the appropriate commutation relations. State these relations and explain why they are satisfied.
V. GRAND-CANONICAL ENSEMBLE

V.1. The ensemble

V.1.1. Definition

The grand-canonical ensemble is not, strictly speaking, an ensemble of the type (39), but a slight generalization of it. We will now denote an $N$ particle microstate by $\Gamma^N = (r^N, p^N)$. The grand-canonical ensemble is a density $\rho_g(N, \Gamma^N)$ on both $N$ and $\Gamma^N$. It depends on a parameter $\mu$ and is defined by

$$\rho_g(N, \Gamma^N) = \frac{1}{N!h^{3N}} e^{\beta \mu N - \beta H}. \quad (1)$$

The expression on the RHS is of course equal to $e^{\beta \mu N} \rho_c(\Gamma^N)$. The grand-canonical distribution may be viewed as a weighted sum of juxtaposed weighted canonical distributions. The associated grand-canonical partition function, $Z_g$ is therefore

$$Z_g(T, V, \mu) = \sum_{N=0}^{\infty} \int d\Gamma \frac{1}{N!h^{3N}} e^{\beta \mu N - \beta H}, \quad (2)$$

which depends on the triplet of independent variables

$$(T, V, \mu). \quad (3)$$

Eq. (2) may be rewritten in various ways, in particular as

$$Z_g(T, V, \mu) = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_c(T, V, N) \quad (4)$$

$$= \sum_{N=0}^{\infty} e^{\beta \mu N - F_c(T, V, N)}. \quad (5)$$

The first line of this equation shows that $Z_g$ is the generating function of the $Z_c(T, V, N)$ for $N = 0, 1, 2, \ldots$. Passing from $Z_c$ to its generating function is not very different from doing a Laplace transform, albeit a discrete one (sum on $N$ instead of an integral).

The quantity $e^{\beta \mu} \equiv z$ is called the fugacity of the system\(^\text{23}\). A high (low) fugacity or chemical potential corresponds to a high (low) tendency for the presence of particles. The chemical potential $\mu$ can be positive, negative, or zero; it is measurable only indirectly.

V.1.2. Grand-canonical averages

The probability distribution $P_g(N, \Gamma^N)$ associated with the grand-canonical ensemble is

$$P_g(N, \Gamma^N) = \frac{\rho_g(N, \Gamma^N)}{Z_g} \quad (6)$$

It is normalized to unity,

$$\sum_{N=0}^{\infty} \int d\Gamma^N P_g(N, \Gamma^N) = 1. \quad (7)$$

\(^{23}\)In chemical thermodynamics it is rather called the activity or activity coefficient.
Let \( A(N, \Gamma^N) \) be an observable. Its grand-canonical average is
\[
\langle A \rangle_g = \sum_{N=0}^{\infty} \int d\Gamma^N A(N, \Gamma^N) P_g(N, \Gamma^N)
\]
\[
= \frac{1}{Z_g} \sum_{N=0}^{\infty} \frac{1}{N! h^{3N}} e^{\beta \mu N} \int d\Gamma^N A(N, \Gamma^N) e^{-\beta H(N, \Gamma^N)},
\]
(8)
in which \( H(N, \Gamma^N) \) is the \( N \) particle Hamiltonian. Hence \( \langle A \rangle_g \) is the result of an average that also runs over a varying number of particles.

There may be several reasons for wanting to work in the grand-canonical ensemble. Among these:

(i) it allows for the description of the physical phenomenon of a fluctuating particle number, for example (but not only) in the presence of permeable walls;

(ii) for certain types of problems, usually easy to recognize, the thermodynamics is derived most easily in the grand canonical ensemble;

(iii) the grand-canonical ensemble provides a suitable framework for setting up a theory for nonideal gases – which is one of the final purposes of statistical mechanics.

V.2. The grand-canonical potential

V.2.1. Legendre transformation between \( F_c \) and \( J_g \)

The thermodynamic potential associated with \( Z_g \) is the grand-canonical potential (or just grand potential) \( J_g \) defined by \(^{24}\)
\[
J_g(T, V, \mu) = -k_B T \log Z_g(T, V, \mu).
\]
(9)
If \( J_g \) is extensive, as we expect it is, then we must have
\[
J_g(T, V, \mu) = V j(T, \mu), \quad V \to \infty;
\]
(10)
notice that since \( N \) is no longer an independent variable, we have used here proportionality to \( V \) as the criterion for extensivity.

Eq. (5) shows that \( Z_g \) is obtained as the sum over the exponential of an extensive expression. The sum is therefore strongly peaked and, for reasons discussed several times before, will be equal to the value of the summand taken in the peak. Therefore, if this peak occurs for \( N = N_g(T, V, \mu) \), we have
\[
Z_g(\mu) = \max_N e^{\beta \mu N - \beta F_c(N)}
\]
\[
= e^{\beta \mu N_g - \beta F_c(N_g)},
\]
(11)
where subextensive terms in the exponential have been neglected and where, for simplicity, we have not indicated the dependence on \( V \) and \( T \). It follows that
\[
J_g = F_c - \mu N_g
\]
(12)

\(^{24}\)There is no general agreement on the symbols. Christophe Texier uses \( \Xi \) instead of \( Z_g \). It is common to use \( \Omega \) instead of \( J_g \). Many authors use \( z = e^{\beta \mu} \) for the fugacity.
This relation between \( J_g \) and \( F_c \) is again a Legendre transformation; it is the consequence of the “Laplace” transformation (4) and the fact that the thermodynamic limit is being taken. Indeed, the relation \( J = F - \mu N \) is well-known in thermodynamics.

The value \( N_g(T,V,\mu) \) is the solution of \( \frac{\partial}{\partial N} [\mu N - F_c(N)] \bigg|_{N_g} = 0 \), from which we see directly that the parameter \( \mu \) of the grand-canonical ensemble has the interpretation

\[
\mu = \mu_c. \tag{13}
\]

When its natural variables \( T, V, \mu \) undergo variations \( dT, dV, \) and \( d\mu \), the resulting infinitesimal change in \( J_g \) is

\[
dJ_g = \left( \frac{\partial J_g}{\partial T} \right)_{V,\mu} dT + \left( \frac{\partial J_g}{\partial V} \right)_{T,\mu} dV + \left( \frac{\partial J_g}{\partial \mu} \right)_{T,V} dN. \tag{14}
\]

The three new partial derivatives that appear here may be calculated directly in terms of known canonical quantities and we find (reasoning in a similar way as in section III.2.1) that, when (13) is satisfied,

\[
S_g = S_c, \quad p_g = p_c. \tag{15}
\]

Hence the grand-canonical ensemble leads to the same thermodynamics as the canonical (and the microcanonical) ensemble; we may suppress the indices \( m, c, \) and \( g \) on the thermodynamic quantities: these refer only to how they have been calculated.

**Remark.** If we compare the definition of the pressure \( p_g \) in (14) to (10), we see that the derivation of \( J_g \) with respect to \( V \) actually amounts to a simple division. It follows that the grand-canonical potential satisfies

\[
J_g = -pV. \tag{16}
\]

**V.2.2. Other thermodynamic quantities**

It is easy to show that the average number of particles of a system described by the grand-canonical ensemble may be obtained as

\[
\langle N \rangle_g = k_B T \frac{\partial \log Z_g}{\partial \mu}. \tag{17}
\]

Because of the peaked nature of the distribution of \( N \) we must have \( \langle N \rangle_g = N_g \) and in fact (17) is just another way of expressing the third partial derivative in (14).

**Note.** In spite of \( N, V, T \) being the independent variables, one often sets \( \beta = 1/k_B T \) (as always) and \( \alpha = -\mu/k_B T \), and considers \( \alpha \) and \( \beta \) as independent. Eq. (17) then takes the form

\[
\langle N \rangle_g = - \frac{\partial \log Z_g}{\partial \alpha}. \tag{18}
\]

One shows furthermore that

\[
\frac{\partial^2 \log Z_g}{\partial \alpha^2} = \frac{\langle N^2 \rangle_g - \langle N \rangle_g^2}{\sigma_N^2}. \tag{19}
\]
Since \( \log Z_g \) is extensive (proportional to \( V \)), this shows that the typical deviation \( \Delta N = N - \langle N \rangle_g \) of the particle number from its average scales as \( V^{1/2} \). The relative fluctuation \( \Delta N/\langle N \rangle_g \sim V^{-1/2} \) of the particle number is therefore very small and negligible in the thermodynamic limit.

**Exercise.** Relate the fluctuations (19) in the number of particles to the isothermal compressibility \( \kappa_T \equiv -\frac{1}{V} \frac{\partial V}{\partial p} \). Discuss beforehand in what sense you expect \( \langle \Delta N^2 \rangle_g \) to vary with \( \kappa_T \).

You may check for yourself that the average energy in the grand-canonical ensemble is obtained as

\[
\langle E \rangle_g = -\left. \frac{\partial \log Z_g}{\partial \beta} \right|_{\alpha \text{ fixed}}.
\]

[Remark on mixed derivatives with respect to \( \alpha \) and \( \beta \).]

**V.3. Derivation of the grand-canonical ensemble**

The grand-canonical ensemble applies to any subsystem (whether finite or thermodynamically large) of a canonical or microcanonical system with which it can exchange both energy and particles. The subsystem may be separated from the total system by permeable walls or simply by an imaginary closed surface. Formal derivations of the grand-canonical ensemble go along the same lines as the derivation of the canonical from the microcanonical one; we will not do that here.

**V.4. The ideal gas**

Recall that the canonical partition function of the ideal gas, given by (III.22), reads

\[
Z_c(T, V, N) = \frac{1}{N!} \left( \frac{V}{\lambda^3} \right)^N
\]

Substituting this in (4) and doing the sum we find

\[
Z_g(T, V, \mu) = \exp \left( e^{\beta \mu} \frac{V}{\lambda^3} \right)
\]

whence

\[
J_g(T, V, \mu) = -k_B T e^{\beta \mu} \frac{V}{\lambda^3}.
\]

It follows immediately that the grand-canonical pressure and average particle number of the ideal gas are given by

\[
p_g = -\left. \frac{\partial J_g}{\partial V} \right|_{\mu} = k_B T e^{\beta \mu} \frac{V}{\lambda^3},
\]

\[
N_g = -\left. \frac{\partial J_g}{\partial \mu} \right|_{V} = e^{\beta \mu} \frac{V}{\lambda^3}.
\]

For the particle density this gives \( n_g = N_g/V = e^{\beta \mu}/\lambda^3 \), which goes down with the chemical potential as expected. When \( \mu \) is eliminated from equations (24) and (25) we find that

\[
p_g = n_g k_B T,
\]
which is again the equation of state of the ideal gas, expressed in terms of the usual physical quantities, but now derived within the framework of the grand-canonical ensemble.

To obtain the average grand-canonical energy we may proceed by calculating

\[
\langle E \rangle_g = -\frac{\partial \log Z_g}{\partial \beta} \bigg|_{\alpha \text{ fixed}} = -\frac{\partial}{\partial \beta} e^{-\alpha V} \frac{V}{\lambda^3} = -e^{-\alpha V} \frac{\partial}{\partial \beta} \left( \frac{2\pi m}{\beta h^2} \right)^{\frac{3}{2}}
\]

\[
= \frac{3}{2} N_g k_B T.
\]

(27)
in agreement with what the other ensembles gave.

**Exercise.** Rederive the Sackur-Tetrode formula (I.15) for the entropy within the framework of the grand-canonical ensemble.

### V.5. Nonideal gases: the virial expansion

We have studied ideal gases in three ensembles. Real gases may be close to ideal, but differences will necessarily become more apparent as the density goes up. Suppose that for a nonideal gas we may write the equation of state as

\[
\frac{p}{k_B T} = n + B_2 n^2 + B_3 n^3 + \ldots,
\]

(28)
in which the first term is the ideal gas expression and the higher order terms account for deviations from the ideal gas law as the density goes up. The series (28) is called a *virial expansion*; the coefficients \(B_\ell\), which depend on temperature, are the *virial coefficients*. They may of course be determined from experimental data and will depend on the kind of gas under consideration.

**Question:** For a nonideal gas with a given interparticle potential \(U(|\mathbf{r}_i - \mathbf{r}_j|)\), can we determine the coefficients of the correction terms? If so, the measured equation of state will tell us something about the interparticle potential and *vice versa*.

**Answer:** Yes, this can be done, and the grand-canonical ensemble is a very suitable framework for the theory.

Let us consider the \(N\) particle Hamiltonian

\[
\mathcal{H}(N, \Gamma^N) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{1 \leq i < j \leq N}^{U_N} U(r_{ij}'),
\]

(29)
where \(r_{ij'} \equiv |\mathbf{r}_i - \mathbf{r}_{j'}|\). Let in this section \(Q_N\) denote the configurational partition function of the \(N\) particle system,

\[
Q_N(T,V) = \int d\mathbf{r}_1 \ldots d\mathbf{r}_N \ e^{-\beta U_N},
\]

(30)
with \(Q_0 \equiv 1\). [Briefly discuss here \(Q_1\) and \(q_2\).] Since \(Q_N\) involves interacting particles, there is no hope of ever being able to calculate it exactly.

We will set \(\zeta = e^{\beta \mu}/\lambda^3\). The grand-canonical partition function for the system described by Hamiltonian (29) then takes the form

\[
Z_g(T,V,\mu) = \sum_{N=0}^{\infty} \frac{\zeta^N}{N!} Q_N(T,V).
\]

(31)
Let us now expand its logarithm as a series in powers of $\zeta$. This gives
\[
\log Z_g(T,V,\mu) = Q_1 \zeta + \frac{1}{2} (Q_2 - Q_1^2) \zeta^2 + \frac{1}{6} (Q_3 - 3Q_1Q_2 + 2Q_1^3) \zeta^3 + \ldots
\]
(32)

where the second line defines the $b_\ell(T,V)$. You may check easily that $Q_1 = V$ whence $b_1 = 1$. Notice that $Q_N$ involves an $N$-fold volume integral: we expect it to be proportional to $V^N$ in the thermodynamic limit. However, since we know that $\log Z_g \sim V$, there must occur cancelations of all nonlinear powers of $V$ in the expressions for the $b_\ell$. Or put differently, we expect that
\[
\lim_{V \to \infty} b_\ell(V,T) \equiv b_\ell(T)
\]
(34)

exists.

From (33) we have, if we use that $\log Z_g = pV/k_BT$ and divide both members by $V$,
\[
\frac{p}{k_BT} = \sum_{\ell=1}^{\infty} b_\ell(T,V) \zeta^\ell.
\]
(35)

Consider now the series for the grand-canonical particle density $n = \langle N \rangle_g / V$ (we set again $-\beta\mu \equiv \alpha$) that comes along with (33),
\[
n = -\frac{1}{V} \frac{\partial \log Z_g}{\partial \alpha} = \frac{\zeta}{V} \frac{\partial \log Z_g}{\partial \zeta} = \sum_{\ell=1}^{\infty} \ell b_\ell(T,V) \zeta^\ell.
\]
(36)

Upon eliminating $\zeta$ from Eqs. (35) and (36) we find for $p/k_BT$ a series of the exact type (28), in which, in the thermodynamic limit, the coefficients are given by
\[
B_2 = -b_2(T), \quad B_3 = 4b_2^2(T) - 2b_3(T), \ldots
\]
(37)

We now investigate the second virial coefficient $B_2$. From its definition we have
\[
b_2(T) = \lim_{V \to \infty} \frac{1}{2V} \left[ \int_{V} \int_{V} \, dr_1 \, dr_2 \, e^{-\beta U(r_{12})} - V^2 \right]
\]
\[
= \lim_{V \to \infty} \frac{1}{2V} \int_{V} \int_{V} \, dr_1 \, dr_2 \left[ e^{-\beta U(r_{12})} - 1 \right]
\]
(38)

whence
\[
B_2(T) = -\frac{1}{2} \int_{0}^{\infty} 4\pi r^2 \, dr \left[ e^{-\beta U(r)} - 1 \right].
\]
(39)

If needed, this integral is easily evaluated numerically. Virial coefficients have been measured and tabulated by experimentalists for many real gases. Knowing the precise equation of state of a gas is of immense importance for a multitude of industrial purposes.

**Exercise.** Discuss the sign of $B_2$ for attractive and repulsive potentials.
**Exercise.** Derive the expression for $B_3$ given in (37).

**Exercise.** Calculate the second virial coefficient for a $d$-dimensional gas of hard spheres of diameter $a$, for $d = 3, 2, 1$. Compare with the exercise about the one-dimensional hard rods that we did in chapter III.

*Epilog to this section.* Calculating the consequences of interactions between particles is one of the fundamental and very difficult questions of statistical physics. It has no all-encompassing answer.

In many approaches to this problem there is a “zeroth order” Hamiltonian that one knows how to treat exactly, and an interaction term that leads to a perturbation series. The virial expansion is but one implementation of this general scheme; here the zeroth order system is the ideal gas. In other cases the zeroth order system might be, for example, the ground state, or the harmonic crystal.

Most gases are known to have critical points (recall the typical phase diagram and what is meant by a critical point). It so happens that the critical point is not close to any zeroth order Hamiltonian. The study of critical phenomena therefore presents specific theoretical difficulties and has developed into a scientific domain of its own. We must refer you to other courses to learn more about this.
VI. STATISTICAL MECHANICS OF QUANTUM SYSTEMS

In chapter I we showed how for a classical system you can go from the basic equations of motion – the Newton or Hamilton-Jacobi equations – to the Liouville equation for an ensemble $\rho$ in classical phase space.

In this chapter we will start all over again and do the same thing for a quantum system, starting from the Schrödinger equation. It will appear that the classical ensemble $\rho$ is replaced with a density matrix or density operator also called $\rho$, and the classical Liouville equation of motion for $\rho$ by its quantum mechanical counterpart, called the von Neumann equation. The stationary solutions of this equation are the quantum analogs of the classical ensembles.

Our dealings with quantum systems in earlier chapters (in particular in chapter IV: “Harmonic oscillators”) anticipated upon the results of the present chapter.

The material of this chapter may be found in part in C. Texier, Chapitre 3, sections II.A and II.B. You may also consult B. Diu et al., Physique Statistique (ed. Herrmann, Paris, 1989), Compléments I.H and V.E.

VI.1. Reminders about quantum mechanics

Let us consider the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi$$

where $\Psi(r_1, r_2, \ldots, r_N; t)$ is the wave function\textsuperscript{25} of a many-body system of $N$ particles. The number $|\Psi(r_N, t)|^2 dr^N$ denotes the probability of finding particle $i$ between $r_i$ and $r_i + dr_i$ for all $i = 1, 2, \ldots, N$. The corresponding ket will be denoted\textsuperscript{26} $|\Psi\rangle$. It is a vector in the Hilbert space of the $N$ particle system. In the present context we call $|\Psi\rangle$ a pure state of the system. All states that you have encountered in an introductory course to quantum mechanics were pure states. It is not needed for $\Psi$ to be an eigenstate of what operator soever. The number $N$ may be very large, for example of the order of $10^{23}$. A macroscopic system is described by a wave function in the same way as a microscopic one.

The Hamiltonian is an operator in the space of pure states. In one of the simplest cases it might be of the form

$$H = \frac{1}{2m} \sum_{i=1}^{N} \mathbf{p}_i^2 + \sum_{1 \leq i < i' \leq N} U(|\mathbf{r}_i - \mathbf{r}_{i'}|),$$

$$p_i = \frac{\hbar}{i} \frac{\partial}{\partial r_i}.$$\textsuperscript{22}

The quantum mechanical scalar product is

$$\langle \Psi_1 | \Psi_2 \rangle = \int d\mathbf{r}^N \Psi_1^*(\mathbf{r}^N) \Psi_2(\mathbf{r}^N)$$

\textsuperscript{25}If needed, the spin is easily incorporated in the same formalism.

\textsuperscript{26}We avoid the notation $|\Psi\rangle$ since we are using angular brackets to define ensemble averages.
and the normalization condition
\[
(\Psi|\Psi) = \int dr^N |\Psi(r^N, t)|^2 = 1.
\] (4)

Let \( n \) be an arbitrary index (no need to specify whether it is discrete or continuous, or partly both; it may take an infinity of values) and let \( \{\phi_n(r^N)\} \) be a time independent orthonormal basis, that is,
\[
(\phi_m|\phi_n) = \delta_{mn}.
\] (5)

An arbitrary pure state is a superposition of the \( \phi_n \) with time dependent coefficients \( c_n(t) \),
\[
|\Psi(t)\rangle = \sum_n c_n(t)|\phi_n\rangle,
\] (6)

where \( c_n(t) = (\phi_n|\Psi(t)) \) and where for convenience we have suppressed the spatial coordinates \( r^N \). The normalization (4) of \( \Psi(t) \) implies that
\[
\sum_n |c_n(t)|^2 = 1.
\] (7)

The time evolution of the \( c_n(t) \) is determined by the Schrödinger equation (1). Upon substituting (6) in (1) and then left-multiplying by \( (\phi_n| \) one obtains
\[
i\hbar \frac{dc_n(t)}{dt} = \sum_m H_{nm} c_m(t)
\] (8)
in which
\[
H_{nm} = (\phi_n|H|\phi_m)
\] (9)
is the matrix representation of the Hamiltonian in the basis \( \{\phi_n(r^N)\} \).

Let \( A \) be an operator corresponding to a physical observable. Quantum mechanics tells us that \( \overline{A} \equiv (\Psi|A|\Psi) \) is the expectation value of \( A \) that will be observed when the system is in pure state \( \Psi \).

At this point we will choose the \( \phi_n \) to be eigenstates of \( A \),
\[
A\phi_n = A_n \phi_n,
\] (10)
where \( A_n \) is one of the eigenvalues of \( A \). It then follows that in the pure state \( \Psi(t) \) the expectation value of \( A \) may be expressed as
\[
\overline{A}(t) = (\Psi(t)|A|\Psi(t))
= \sum_n \sum_m A_m c_n(t)^* c_m(t) (\phi_n|\phi_m)
= \sum_n A_n |c_n(t)|^2.
\] (11)
Here the $A_n$ are the possible outcomes of a measurement of $A$ performed on the system and $|c_n(t)|^2$ is the quantum mechanical probability for the system to be in eigenstate $n$.

**VI.2. The density operator / density matrix**

Suppose now we do not know the exact pure state of the system. Suppose all we know is that it is in one of a set of pure states $\Psi^k$, where $k$ is an index. In analogy to what precedes we set

$$\Psi^k(t) = \sum_n c_n^k(t)\phi_n$$  \hspace{1cm} (12)

and will write the expectation value of a physical observable $A$ in pure state $\Psi^k$ as

$$\overline{A}^{(k)} = (\Psi^k | A | \Psi^k) = \sum_n |c_n^k(t)|^2 A_n.$$  \hspace{1cm} (13)

Let now $P_k$ be the (classical $=$ nonquantum) probability that the system is in $\Psi^k$. It will be convenient to work with unnormalized weights $p_k$ from which the $P_k$ may be derived according to

$$P_k = \frac{p_k}{Z}, \quad Z = \sum_k p_k.$$  \hspace{1cm} (14)

The collection of weights $\{p_k\}$ constitutes again an ensemble; it is an ensemble of pure states. The expectation value (the average) of observable $A$ in this ensemble will be denoted by $\langle A \rangle$. It is obviously given by

$$\langle A \rangle = \sum_k P_k \overline{A}^{(k)}.$$  \hspace{1cm} (15)

This formula expresses the essence of what we want to convey in this chapter. However, we are now going to rewrite it in several other useful ways. We have

$$\langle A \rangle = Z^{-1} \sum_k p_k \overline{A}^{(k)} = Z^{-1} \sum_k p_k (\Psi^k | A | \Psi^k).$$  \hspace{1cm} (16)

Upon inserting the complete basis of states $\sum_n |\phi_n\rangle\langle\phi_n|$ before and after the operator $A$ we get

$$\langle A \rangle = Z^{-1} \sum_n \sum_m \sum_k (\phi_m | \Psi^k \rangle p_k (\Psi^k | \phi_n \rangle \langle \phi_n | A | \phi_m \rangle) \rho_{mn} A_{nm}$$  \hspace{1cm} (17)

in which $\rho_{mn}$ is the $mn$ matrix element of the density operator $\rho$ defined by

$$\rho = \sum_k |\Psi^k\rangle p_k (\Psi^k |.$$  \hspace{1cm} (18)
Hence

\[ \langle A \rangle = \frac{1}{Z} \sum_n \sum_m \rho_{nm} A_{mn} \]

\[ = \frac{1}{Z} \text{tr} \rho A \]

\[ = \frac{\text{tr} \rho A}{\text{tr} \rho} . \]  \hspace{1cm} (19)

Remarks.
1. The factor $1/Z$ is sometimes incorporated in the definition of $\rho$. In that case one has $\text{tr} \rho = 1$ and $\langle A \rangle = \text{tr} \rho A$.
2. The density operator $\rho$ is an operator in Hilbert space in the same way as $\mathcal{H}, A, B, \ldots$. It is time dependent (why?). As any operator, it may be expressed as a matrix once a basis has been chosen.
3. The “collection of pure states” with which we started is also called a statistical mixture of pure states or a mixed state. One might be tempted to speak of an “ensemble of pure states”; however, it is the density operator $\rho$, rather than this collection of pure states, that is the counterpart of the classical ensembles. In particular, $\rho$ determines all the expectation values of the physical observables.
4. Different statistical mixtures may lead to the same density operator.

VI.3. The von Neumann equation

The time dependence of the density operator $\rho$ is determined by the Schrödinger equation and we will now establish how.

We have

\[ \rho_{mn}(t) = \sum_k p_k c_{mk}^*(t)c_{nk}(t) . \]  \hspace{1cm} (20)

Employing Eq. (8) for the time evolution of $c_n(t)$ and its complex conjugate and obtain

\[ i\hbar \frac{d}{dt} \rho_{mn}(t) = \ldots \]

\[ = \sum_\ell \left( \mathcal{H}_{m\ell} \rho_{\ell n}(t) - \rho_{m\ell}(t) \mathcal{H}_{\ell n} \right) \]  \hspace{1cm} (21)

or, in operator notation,

\[ i\hbar \frac{d}{dt} \rho = [\mathcal{H}, \rho] . \]  \hspace{1cm} (22)

Eq. (22) is called the von Neumann or Liouville–von Neumann equation; it is the quantum counterpart of the classical Liouville equation.

One advantage of the formulas where the density operator appears is that it is not necessary to specify a basis. For example, $\text{tr} \rho$ is a scalar whose value is independent of the basis selected to evaluate it.
VI.4. Stationary density operators

A density operator is stationary if \( d\rho/dt = 0 \), that is, if

\[
[\mathcal{H}, \rho] = 0. \tag{23}
\]

A sufficient condition for (23) to hold is that we have \( \rho = f(\mathcal{H}) \), where \( f \) is an arbitrary function. In that case \( \rho \) is diagonal in the basis of eigenfunctions of \( \mathcal{H} \). Let us denote these by \( \psi_n \) and their eigenvalues by \( E_n \),

\[
\mathcal{H}\psi_n = E_n\psi_n. \tag{24}
\]

Using these as a basis we find

\[
Z = \text{tr} \rho = \sum_n (\psi_n|\rho|\psi_n) = \sum_n \rho_{nn} \tag{25}
\]

and

\[
\langle A \rangle = \frac{\text{tr} \rho A}{\text{tr} \rho} = \frac{\sum_n (\psi_n|\rho A|\psi_n)}{\sum_n (\psi_n|\rho|\psi_n)} = \frac{\sum_n \rho_{nn} (\psi_n|A|\psi_n)}{\sum_n \rho_{nn}}. \tag{26}
\]

**Interpretation**

In thermal equilibrium the expectation value of an observable \( A \) of a quantum system is obtained by

1. calculating it in an eigenstate \( \psi_n \) of the system Hamiltonian;
2. averaging the result with weight \( \rho_{nn} \) on all eigenstates.

Different choices of \( f \) lead to different weights \( \rho_{nn} \). These \( \rho_{nn} \) are the analogs of the classical ensembles. In several exercises we have already carried out steps (1) and (2). They appear here as a logical consequence within the framework of the density operator formalism.

The stationary ensembles usually considered in quantum mechanics are the same as those used in classical mechanics. In particular

\[
\rho_{mn} = \begin{cases} 
\delta_{mn} & \text{if } E < E_n < E + \delta E, \\
0 & \text{otherwise,}
\end{cases} \quad \text{microcanonical ensemble,}
\]

\[
\rho_{nn} = \delta_{nn} e^{-\beta E_n}, \quad \text{canonical ensemble.} \tag{27}
\]

In the latter case one may also write

\[
\rho = e^{-\beta \mathcal{H}}. \tag{28}
\]

In this formalism the canonical partition function and the free energy become

\[
Z_c = \text{tr} e^{-\beta \mathcal{H}} = \sum_n e^{-\beta E_n} \tag{29}
\]
and
\[ F_c = -k_B T \log Z_c = -k_B T \log \text{tr} e^{-\beta H}, \tag{30} \]
respectively, which are relations already encountered before.

Let now \( \mathcal{H}_N \) be the Hamiltonian of an \( N \) particle system and let \( \rho_N = \exp(-\beta \mathcal{H}_N) \).
Then the grand-canonical partition function is
\[
Z_g(\mu, V, T) = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_c(T, V, N) \\
= \sum_{N=0}^{\infty} e^{\beta \mu N} \text{tr} \rho_N \\
= \sum_{N=0}^{\infty} \text{tr} e^{\beta (\mu N - \mathcal{H}_N)} \tag{31}
\]

In earlier chapters we have calculated quantum mechanical partition functions for a number of simple systems (spins, harmonic oscillators). The present formalism adds nothing new to that; it simply confirms that we did it the correct way.

Remark. From the above it is clear that in equilibrium statistical mechanics the off-diagonal elements (\( \rho_{mn} \) with \( m \neq n \)) of the density matrix play no role. They do play a role, however, in time-dependent phenomena, and in such cases the present formalism is the indispensable starting point.

Example
We consider a spin \( \frac{1}{2} \) whose Hilbert space is spanned by the two kets \( |+\rangle \) and \( |−\rangle \). An arbitrary state may be written as \( \alpha |+\rangle + \beta |−\rangle \), where \( \alpha \) and \( \beta \) are complex coefficients. The probability distribution \( p_k \) discussed above becomes, in this example, a distribution \( p(\alpha, \beta) \). [To be completed.]
VII. IDEAL QUANTUM GASES

By an ideal gas we mean a system of particles without interaction. The following observations, made earlier for a classical system, remain true in quantum mechanics.
1. An ideal gas which is not at equilibrium will never be able to reach equilibrium. This is because equilibration is due to the exchange of energy and momentum between the particles.
2. Ideal gases do not exist: there are always interactions, even if they may be weak, that guarantee that a real gas will tend to equilibrium. Nevertheless, for several problems of interest in physics an ideal gas may be a good approximation to a real gas.

VII.1. The ideal gas Hamiltonian and its eigenfunctions

Let us consider $N$ identical noninteracting particles of spin $S$ enclosed in a cubic volume $V = L^3$. We will write $r_j = (x_j, y_j, z_j) \in V$ for the position of the $j$th particle and $s_j = -S, -S + 1, \ldots, S$ for its spin state. The $N$ particle Hamiltonian is spin independent and given by

$$H = \frac{1}{2m} \sum_{j=1}^{N} p_j^2$$

$$= -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \left( \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right)$$

$$= \sum_{j=1}^{N} H_j,$$  \hspace{1cm} (1)

in which the $H_j$ are single-particle Hamiltonians. We adopt periodic boundary conditions, that is, we identify the positions $r$ and $r + (m_1, m_2, m_3)L$ where $m_1, m_2,$ and $m_3$ are arbitrary integers.

The single-particle eigenkets and eigenvalues of $H_j$ depend on a wavevector $k$ and a spin index $\sigma$ and are given by

$$|k\sigma\rangle$$  \hspace{1cm} (2)

and

$$\epsilon_{k\sigma} = \frac{\hbar^2 k^2}{2m},$$  \hspace{1cm} (3)

respectively, in which

$$k = \frac{2\pi}{L} n = \frac{2\pi}{L} (n_x, n_y, n_z)$$  \hspace{1cm} (4)

where $n_x, n_y,$ and $n_z$ are arbitrary integers, and

$$\sigma = -S, -S + 1, \ldots, S.$$  \hspace{1cm} (5)

The $|k\sigma\rangle$ are orthonormal; they are a special choice for the single-particle basis states that were called $|\lambda\rangle$ in your Quantum Mechanics course. As a consequence, the labels
correspond to single-particle energies. Note that the energy \( \epsilon_{k\sigma} \) is spin independent here. In position-spin representation we may write

\[
\psi_{k\sigma}(r_j, s_j) = (r_j s_j | k\sigma) = \frac{1}{\sqrt{V}} e^{ik \cdot r_j} \delta_{\sigma,s_j}
\]  

(6)

Suppose now that from the infinite set of possible labels \( \{(k\sigma)\} \) we select \( N \) specific ones, namely \( (k_1, \sigma_1), (k_2, \sigma_2), \ldots, (k_N, \sigma_N) \). The \( N \) particle product states that may be constructed from the single-particle states with these labels are

\[
\Psi_{k_1\sigma_1 \ldots k_N\sigma_N}(r_1, s_1 \ldots r_N, s_N) = \psi_{k_1\sigma_1}(r_1, s_1)\psi_{k_2\sigma_2}(r_2, s_2) \ldots \psi_{k_N\sigma_N}(r_N, s_N)
\]  

(7)

in which \( k_j \) is the wavevector of particle \( j \). They are also eigenfunctions of \( \mathcal{H} \) with eigenvalues

\[
E_{k_1\sigma_1 \ldots k_N\sigma_N} = \epsilon_{k_1\sigma_1} + \epsilon_{k_2\sigma_2} + \ldots + \epsilon_{k_N\sigma_N}.
\]  

(8)

Even though every function of type (7) is mathematically an eigenfunction of \( \mathcal{H} \), not every such function is allowed to describe a physical system. Two cases have to be distinguished.

I. **Bosons.** The particles are bosons (have an integer spin). In that case, out of all mathematically possible product states (7) with the selected labels the only allowable one is the symmetric linear combination

\[
\Psi_{\text{symm}}^{k_1 \sigma_1 \ldots k_N \sigma_N} = \sum_P \psi_{k_1 \sigma_1}(r_{P1}, s_{P1})\psi_{k_2 \sigma_2}(r_{P2}, s_{P2}) \ldots \psi_{k_N \sigma_N}(r_{PN}, s_{PN}),
\]  

(9)

where \( P \) is a permutation of the indices \( \{1, 2, \ldots, N\} \) and where we ignore a normalization factor which in the present discussion is of no importance. The lower index of \( \Psi_{\text{symm}} \) is a collection of \( N \) labels \( (k_j, \sigma_j) \) that need not be all different. Note that a permutation of two labels does not give a new state. The particles are said to obey Bose-Einstein (BE) statistics.

II. **Fermions.** The particles are fermions (have half-integer spin). In that case, out of all mathematically possible product states (7) with the selected labels the only allowable one is the antisymmetric linear combination

\[
\Psi_{\text{antisymm}}^{k_1 \sigma_1 \ldots k_N \sigma_N} = \sum_P (-1)^P \psi_{k_1 \sigma_1}(r_{P1}, s_{P1})\psi_{k_2 \sigma_2}(r_{P2}, s_{P2}) \ldots \psi_{k_N \sigma_N}(r_{PN}, s_{PN})
\]  

(10)

in which \( (-1)^P \) stand for the signature of the permutation; and where the second line is a rewriting of the first one as a *Slater determinant*. In order for this wavefunction \( \Psi_{\text{antisymm}}^{k_1 \sigma_1 \ldots k_N \sigma_N} \)
to be nonzero its lower index must be a collection of labels \((k_j, \sigma_j)\) that are all different. This is called the Pauli exclusion principle: no two particles can be in the same state. The particles are said to obey Fermi-Dirac (FD) statistics.

VII.2. Characterization with the aid of occupation numbers

The wavevectors \(k\) constitute an infinite three-dimensional lattice each of whose sites represents \(2S + 1\) single-particle states. A microstate of the quantum \(N\) particle system may be characterized

- either by listing the \(N\) pairs \((k_j, \sigma_j)\);
- or by specifying for each single-particle state \((k, \sigma)\) the number \(n_{k\sigma}\) of particles in that state; we call \(n_{k\sigma}\) the occupation number.

In the latter case the list of occupation numbers is infinite and such that

\[
\sum_{k, \sigma} n_{k\sigma} = N \tag{11}
\]

in which we have the possible values

\[
\begin{cases}
\text{for BE: } n_{k\sigma} = 0, 1, 2, \ldots, \\
\text{for FD: } n_{k\sigma} = 0, 1.
\end{cases} \tag{12}
\]

This alternative but equivalent representation will greatly facilitate the calculations below.

VII.3. Important remark

One often encounters quantum systems of \(N\) independent particles but which are not described by the ideal gas Hamiltonian (1). Examples are

- \(N\) independent particles in an external potential.
- \(N\) interacting particles whose Hamiltonian, after diagonalization, takes the form of a sum of independent single-particle Hamiltonians. These are then called quasi-particles. What changes in such cases? The only changes are
  
  (i) the label \((k, \sigma)\) is replaced with a more general label \(\lambda\) that contains all relevant quantum numbers. The plane waves \(\psi_{k\sigma}\) are replaced with more general wave functions \(\psi_{\lambda}\), but as we have seen the nature of the eigenfunctions of the Hamiltonian does not have any consequences for the thermodynamics.

  (ii) the expression \(\epsilon_{k\sigma} = \frac{\hbar^2 k^2}{2m}\), which is characteristic of free particles, is replaced with a more general expression \(\epsilon_{\lambda}\).

VII.4. Partition function of a system of independent identical quantum particles

VII.4.1. Arbitrary single-particle energies

We are looking for the quantum analog of the classical (= nonquantum) equation of state \(pV = Nk_B T\). We summarize the relevant formulas. The canonical partition function
of a system of \( N \) independent quantum particles is
\[
Z_c(T, V, N) = \sum_{k_1\sigma_1, \ldots, k_N\sigma_N}^{\text{BE,FD}} e^{-\beta E_{k_1\sigma_1, \ldots, k_N\sigma_N}},
\]
(13)
in which the summation must take into account the particle statistics, either BE or FD. The total energy of the \( N \) particle system is given by
\[
E_{k_1\sigma_1, \ldots, k_N\sigma_N} = \sum_{j=1}^{N} \epsilon_{k_j\sigma_j} \quad [\text{which is (8)}]
\]
\[
= \sum_{k,\sigma} n_{k\sigma} \epsilon_{k\sigma}.
\]
(14)

In (13) we may alternatively sum on the \( n_{k\sigma} \), which gives
\[
Z_c(T, V, N) = \sum_{\{n_{k\sigma}\}}^{\text{BE,FD}} \delta_{N,\sum_{k',\sigma'} n_{k'\sigma'}} e^{-\beta \sum_{k',\sigma'} n_{k'\sigma'} \epsilon_{k'\sigma'}}.
\]
(15)

The delta function constraint in this sum is inconvenient. A way to lift it is to pass to the grand-canonical partition function,
\[
Z_g(T, V, \mu) = \sum_{N=0}^{\infty} e^{\beta \mu N} Z_c(N, V, T)
\]
\[
= \sum_{\{n_{k\sigma}\}}^{\text{BE,FD}} e^{\beta \sum_{k',\sigma'} (\mu - \epsilon_{k'\sigma'}) n_{k'\sigma'}}
\]
\[
= \prod_{k,\sigma} \left[ \sum_{n_{k\sigma}}^{\text{BE,FD}} e^{\beta (\mu - \epsilon_{k\sigma}) n_{k\sigma}} \right] z_{k\sigma}.
\]
(16)

The third line of this equation makes \( Z_g \) appear as the product of grand-canonical partition functions \( z_{k\sigma} \) of a single energy level \( \epsilon_{k\sigma} \) that may be populated by a variable number \( n_{k\sigma} \) of particles.

*Remark.* A simpler example of the mathematical operation that has taken place above is the following,
\[
\sum_{n_{1,2}} e^{\sum_{j=1,2} a_j n_j} = \prod_{i=1,2} \left[ \sum_{n_i} e^{a_i n_i} \right].
\]
Check this.

We now need to specify whether the particles are bosons (BE) or fermions (FD). This leads to the two distinct explicit expressions
\[
Z_g^{\text{BE}}(T, V, \mu) = \prod_{k,\sigma} \left[ 1 - e^{\beta (\mu - \epsilon_{k\sigma})} \right]^{-1},
\]
\[
Z_g^{\text{FD}}(T, V, \mu) = \prod_{k,\sigma} \left[ 1 + e^{\beta (\mu - \epsilon_{k\sigma})} \right].
\]
(17)
We will suppose for simplicity that the single-particle energies do not depend on the spin variable and write $G = 2S + 1$ for the spin degeneracy. We then have from (17)

$$\log Z_{BE,FD}^g(T, V, \mu) = \mp G \sum_{\mathbf{k}} \log[1 \mp e^{\beta(\mu - \epsilon_{\mathbf{k} \sigma})}].$$  \hspace{2cm} (18)

In the limit $V \to \infty$ the wavevectors in $k$ space get closer and closer and we may replace the sum in (18) by an integral according to

$$\sum_{\mathbf{k}} \to \frac{V}{(2\pi)^3} \int_{\mathbb{R}^3} d\mathbf{k}.$$  \hspace{2cm} (19)

Recall that $(2\pi)^3/V$ is the volume of the cubic unit cell of the lattice in $k$ space.

VII.4.2. Density expansion for an ideal gas of free particles

So far everything is valid for arbitrary single-particle energies $\epsilon_{\mathbf{k} \sigma}$. We will now restrict ourselves to expression (3), valid for an ideal gas of free particles. Upon substituting (3) in (18) and using (19) we get

$$\log Z_{BE,FD}^g(T, V, \mu) = \mp G \frac{V}{8\pi^3} \int d\mathbf{k} \log[1 \mp e^{\beta(\mu - \beta^2 k^2/2m)}].$$  \hspace{2cm} (20)

We now use in (20) the identity

$$\mp \log(1 \mp e^{x}) = \sum_{\ell=1}^{\infty} (\pm)^{\ell-1} \frac{e^{x \ell}}{\ell}, \quad x < 0,$$  \hspace{2cm} (21)

and obtain

$$\log Z_{BE,FD}^g(T, V, \mu) = G \frac{V}{8\pi^3} \sum_{\ell=1}^{\infty} (\pm)^{\ell-1} \frac{e^{\beta \mu \ell}}{\ell} \int d\mathbf{k} e^{-\beta \ell^2 k^2/2m}$$

$$= G \frac{V}{\lambda^3} \sum_{\ell=1}^{\infty} (\pm)^{\ell-1} \frac{e^{\beta \mu \ell}}{\ell^{5/2}}, \quad \mu < 0,$$  \hspace{2cm} (22)

in which

$$\beta = \frac{1}{k_B T}, \quad \lambda = \frac{h}{\sqrt{2\pi mk_B T}}.$$  \hspace{2cm} (23)

Recall that $\lambda$ is the de Broglie wavelength of a typical particle of mass $m$ in a gas of temperature $T$. Thermodynamics tells us that the pressure $p$ of a gas is related to the grand-canonical partition function by $\log Z_g = \beta pV$. Hence (22) leads us to

$$p_{BE,FD}(\mu, T) = G \frac{k_B T}{\lambda^3} \sum_{\ell=1}^{\infty} (\pm)^{\ell-1} \frac{e^{\beta \mu \ell}}{\ell^{5/2}}.$$  \hspace{2cm} (24)
The chemical potential cannot be measured directly and for that reason we rather express $p$ in terms of the density,

$$n = \frac{\langle N \rangle}{V} = \left( \frac{\partial p}{\partial \mu} \right)_{V,T}. \quad (25)$$

From (24) and (25) we get

$$n = \frac{G}{\lambda^3} \sum_{\ell=1}^{\infty} (\pm)^{\ell-1} \frac{e^{\beta \mu \ell}}{\ell^{3/2}}. \quad (26)$$

We solve (26) for $e^{\beta \mu}$ and substitute the result in (24). From here on we make an expansion for low density $n$, useful only (as we will see in a moment) when the quantum effects are weak. To this end we set

$$e^{\beta \mu} = a_1 n + a_2 n^2 + \ldots \quad (27)$$

and find the coefficients $a_i$ successively,

$$a_1 = \frac{\lambda^3}{G},$$

$$a_2 = \mp a_1^2 \frac{3}{2^{3/2}},$$

$$a_3 = \left( \frac{1}{4} - \frac{1}{3^{3/2}} \right) a_1^3, \quad (28)$$

and so on.

Substitution of (27)-(28) in (24) leads to

$$p^{BE,FD} = nk_B T \left[ 1 \mp 2^{-5/2} \frac{n \lambda^3}{G} + \left( \frac{1}{8} - 2.3^{-5/2} \right) \left( \frac{n \lambda^3}{G} \right)^2 + \ldots \right]$$

$$= nk_B T \left[ 1 \mp 0.1768 \frac{n \lambda^3}{G} - 0.0033 \left( \frac{n \lambda^3}{G} \right)^2 + \ldots \right], \quad (29)$$

which is the desired equation of state, albeit in the form of a series expansion. Several comments are in place now.

**Exercise.** Check the expressions for $a_1, a_2, \text{ and } a_3$ given in Eq. (28). Check that you find (29).

**Exercise.** Show that $pV = \frac{2}{3}E$ for both bosons and fermions.

**Exercise.** Show that the density of single-particle energy levels $\rho(\epsilon)$ is given by $\rho(\epsilon) = VA \times \epsilon^{1/2}$ and find the expression for the constant $A$.

*Answer:* $A = \frac{G}{4 \pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2}$.

**VII.5. Comments**
(1) The above expansion is analogous to the virial expansion.\textsuperscript{27} Even though the particles have no interaction, the effect of quantum statistics are those of an interaction.

(2) The quantum effects are important when \( n\lambda^3 \gg 1 \) or equivalently

\[
\frac{\lambda^3}{v} \gg 1,
\]

where \( v \equiv 1/n \) is the volume per particle. This condition may be interpreted as follows. A single particle has quantum behavior on length scales less than its thermal de Broglie wavelength \( \lambda \). The other particles that notice this quantum character are those within the volume \( \lambda^3 \), and this number is equal to \( \lambda^3/v = n\lambda^3 \).

(3) The sign \( \mp \) is such that the ‘−’ is for bosons and the ‘+’ for fermions. The Pauli principle says (we simplify) that fermions ‘repel’ each other. The pressure of a gas of fermions is therefore larger than that of a classical gas. Bosons, on the contrary, may ‘pile up’ in the same single-particle state without limit. The pressure of a gas of bosons is therefore less than that of a classical gas.

(4) One sometimes says that classical particles obey Maxwell-Boltzmann (MB) statistics. This is the statistics that prevails when \( \langle n_k \sigma \rangle \ll 1 \).

(5) From what precedes it is clear that quantum effects may be enhanced in the following ways.

(i) By lowering the temperature \( T \). Usually you have to be as low as a few degrees Kelvin or less.

(ii) By increasing the particle density \( n \). This means considering solids or liquids rather than gases.

(iii) By working with particles of light mass \( m \), such as electrons or He atoms.

Exercise. The quantum effects disappear also in the limit \( G \to \infty \). Interpret this result.

Terminology. When the quantum effects are dominant, one speaks of a strongly degenerate quantum system. The above expansion in powers of \( n\lambda^3 \) is useful for weakly degenerate quantum systems.

VII.6. Averages of occupation numbers

We return here to Eqs. (18) and (19), which is the point where we were before doing the low density expansion. In the grand-canonical ensemble the average total number of particles is given by

\[
\langle N \rangle_g = \frac{1}{\beta} \frac{\partial \log Z_{BE,FD}^g}{\partial \mu}.
\]

More in detail one may ask about the average number \( \langle n_{k\sigma} \rangle \) of particles in a specific single-particle state \((k,\sigma)\). Since (16) tells us that \( \log Z_g = \sum_{k,\sigma} \log z_{k\sigma} \), this average is

\textsuperscript{27}Compare Eqs. (24) and (26) above to Eqs. (35) and (36) of Chapter V.5. In both cases the latter equation has an extra factor \( \ell \) in the summand.
Figure 9: Average occupation number of a single-particle level of energy $\epsilon_{k\sigma}$ in Bose-Einstein, Fermi-Dirac, and Maxwell-Boltzmann statistics.

obtained as

$$\langle n_{k\sigma} \rangle = \frac{1}{\beta} \frac{\partial \log z_{k\sigma}}{\partial \mu},$$  \hspace{1cm} (32)

whence

$$\langle n_{k\sigma} \rangle_{BE, FD} = \frac{e^{\beta(\mu - \epsilon_{k\sigma})}}{1 \mp e^{\beta(\mu - \epsilon_{k\sigma})}}.$$  \hspace{1cm} (33)

Notice that in (32) we are dealing with the grand-canonical partition function $z_{k\sigma}$ of one specific single-particle energy level.

Expressions (33) are the Bose-Einstein and Fermi-Dirac distributions for independent particles. We recall that $\mu < 0$ so that also $\beta(\mu - \epsilon_{k\sigma}) < 0$ for all $(k, \sigma)$. The total particle density in the system is obtained as $1/V$ times the sum of the particle numbers in each state, that is,

$$n = \frac{1}{V} \sum_{k, \sigma} \langle n_{k\sigma} \rangle_{BE, FD}$$

$$= \frac{G}{8\pi^3} \int dk \frac{e^{\beta(\mu - \epsilon_k)}}{1 \mp e^{\beta(\mu - \epsilon_k)}},$$  \hspace{1cm} (34)

which of course follows more directly if you substitute (20) in (31) and carry out the differentiation with respect to $\mu$. 72
**Exercise.** Say how the classical limit of Eq. (33) may be taken and show that this expression then becomes the Maxwell-Boltzmann distribution.

**Exercise.** Calculate the first order correction term to the equation of state as given by (29) for a gas of argon at normal temperature and pressure. Compare your answer with the first order correction term for the same gas given by (V.28) and discuss the difference between the two expansions.
VIII. IDEAL FERMI GAS

VIII.1. Introduction

We will now be interested in the ideal Fermi gas (that is, composed of noninteracting particles) at temperatures $T$ low enough so that it is strongly degenerate. Two questions, at least, arise:

1. When may we consider that the temperature is sufficiently low?
2. What is the ground state (the only state that counts when $T \to 0$)?

The best example is the “gas” of conduction/valence electrons in a metal. Another example is a system of liquid $^3\text{He}$. The electron cloud around an atom of large atomic number may also be considered as a Fermi gas, but the Coulomb interactions are certainly not negligible, so this gas is not “ideal.”

Exercise. As we have seen, the inequality $n\lambda^3 \ll 1$ is the condition for quantum effects to be weak. Calculate $n\lambda^3$ for the electron gas in a metal at room temperature, assuming that there is one conduction electron per unit of volume equal to $(3\,\text{Å})^3)$. Does this gas behave quantum mechanically?

The exercise above shows that room temperature is a very low temperature indeed for the electron gas in a metal. As a consequence, a canonical ensemble at room temperature for this system will be dominated by states close to the ground state.

VIII.2. Ground state

We now consider a system (“gas”) of free and independent electrons in a cubic volume $V = L^3$ with periodic boundary conditions. Hence $\epsilon_k^\sigma = \epsilon_k = \hbar^2 k^2 / 2m$ with $k = 2\pi n / L$ and $n = (n_x, n_y, n_z)$ a three-dimensional vector of integers. Furthermore $\sigma = \pm 1/2$ whence $G = 2$.

I. Direct argument

The single particle wavevectors $k$ constitute a cubic lattice in $k$ space (“reciprocal space”) whose lattice distance is $2\pi / L$; the reciprocal-space volume per lattice point is therefore $8\pi^2 / V$. Each lattice point $k$ may store one electron of spin $\sigma = 1/2$ and one electron of spin $\sigma = -1/2$. The relation $\epsilon_k = \text{cst}$ defines a spherical surface in $k$ space. If you add successively electrons to the system while keeping its energy as low as possible, you therefore fill successive spherical shells. We let $k_F$ stand for the radius of the shell when all $N$ electrons have been added. A sphere of radius $k$ contains

$$\frac{4\pi}{3} k^3 \cdot \frac{V}{8\pi^3}$$

lattice points. In order to fill these with $N$ electrons this number should be equal to $N/2$.

---

28 The ions are considered as providing an immobile positive background that screens the Coulomb interactions between the electrons.
Figure 10: Single-particle states of an ideal gas. At high temperature only very few will be actually occupied by a particle.

Figure 11: The Fermi surface: here a circle.
(a factor 1/2 because of the spin) and hence

\[ n = \frac{N}{V} = \frac{k_F^3}{3\pi^2}, \]  

which when solved for \( k \) leads to

\[ k_F = (3\pi^2 n)^{1/3}, \quad \text{the Fermi wavenumber.} \]  

We therefore conclude that the ground state is given by the set of occupation numbers

\[ n_{k\sigma} = \begin{cases} 1 & |k| \leq k_F, \\ 0 & |k| > k_F. \end{cases} \]  

The spherical surface \( k = k_F \) is called the Fermi surface. The last electron added has a single-particle energy

\[ \epsilon_F \equiv \epsilon_{k_F} = \frac{\hbar^2 k_F^2}{2m}, \quad \text{the Fermi energy,} \]  

and a velocity

\[ v_F = \frac{\hbar k_F}{m}, \quad \text{the Fermi velocity.} \]  

The Fermi energy and velocity are characteristic of a temperature

\[ T_F = \frac{\epsilon_F}{k_B}, \quad \text{the Fermi temperature.} \]  

**Example.** Copper has \( \epsilon_F = 7.00 \text{eV} \), whence \( T_F = 8.16 \times 10^4 \text{K} \). Furthermore it has \( v_F = 1.57 \times 10^6 \text{ms}^{-1} \).
**Exercise.** Show that the ground state energy $E_0$ of the ideal Fermi gas is given by

$$E_0 = \frac{3}{5} N \epsilon_F.$$  

(8)

**Exercise.** Show that the condition $n \lambda^3 \approx 1$ corresponds to $T \approx T_F$.

II. *Derivation within the formalism*

We return to Eq. (33). In the limit of low temperature we have

$$\langle n_{k\sigma} \rangle = \frac{1}{e^{-\beta(\mu - \epsilon_{k\sigma})} + 1} \to \begin{cases} 1 & \epsilon_{k\sigma} < \mu, \\ 0 & \epsilon_{k\sigma} > \mu. \end{cases}$$  

(9)

Using now that $\epsilon_{k\sigma} = \hbar^2 k^2 / 2m$ we conclude that

all single-particle levels are \( \{ \text{occupied} \} \) for \(|k| \lesssim k_F \equiv \sqrt{2m \mu / \hbar} \),

in which we still have to determine $\mu$. That may be done by imposing the total density of electrons, $n = N/V$, which in the limit $T \to 0$ leads to

$$n = \frac{2}{8\pi^3} \cdot 4\pi \int_0^{k_F} \, dk \, k^2 = \frac{k_F^3}{3\pi^2}.$$  

(11)

From this equation we obtain the same expression for $k_F$ and hence all further results of the "direct argument," but moreover the extra relation

$$\mu(T = 0, n) = \epsilon_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}.$$  

(12)

We may now use the relation $pV = k_B T \log Z_g$ and Eq. (VII.20) to find the pressure in the limit $T \to 0$,

$$p(T = 0, n) = \lim_{T \to 0} \frac{k_B T}{V} \frac{2V}{8\pi^3} \int_{k < k_F} \, dk \, \log \left[ 1 + e^{\beta(\epsilon_F - \hbar^2 k^2 / 2m)} \right]$$

$$= \frac{1}{4\pi^3} \int_{k < k_F} \, dk \left( \epsilon_F - \frac{\hbar^2 k^2}{2m} \right)$$

$$= \frac{1}{4\pi^3} \left[ \frac{\hbar^2 k_F^2}{2m} \cdot \frac{4\pi}{3} k_F^3 - \frac{\hbar^2}{2m} \cdot \frac{4\pi}{5} k_F^5 \right]$$

$$= \frac{\hbar^2 k_F^5}{15\pi^2 m}$$

$$= \frac{2}{5} n \epsilon_F$$

$$= \frac{(3\pi^2)^{2/3}}{5} \frac{\hbar^2}{m} n^{5/3}.$$  

(13)
We see that even at zero temperature the ideal Fermi gas exerts a pressure, which may be seen as due to the nonzero velocities of the particles.

**VIII.3. Thermodynamics at very low $T$**

**Application:** electron contribution to the specific heat of a metal.

**Method:** at room temperature the electron system is still close to its ground state, so we will set up a perturbative calculation around the latter.

**Question:** Find $C_V(T,V,N) = \frac{dE}{dT}$ at low temperature. Write $E$ (rather than $E_g$ or $\langle E \rangle_g$) for the average total energy of the system and the task becomes finding $E(T,V,N)$.

To answer this question we will start from

$$E(T,V,\mu) = \sum_{k,\sigma} \langle n_{k\sigma} \rangle \epsilon_{k\sigma} = \cdots = \frac{V}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty d\eta \frac{\eta^{3/2}}{e^{\beta(\eta-\mu)} + 1},$$

where we introduced $\eta = \epsilon_{k\sigma} = h^2 k^2 / 2m$ as the variable of integration. Similarly

$$N(T,V,\mu) = \sum_{k,\sigma} \langle n_{k\sigma} \rangle = \frac{V}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty d\eta \frac{\eta^{1/2}}{e^{\beta(\eta-\mu)} + 1}.$$ 

Our strategy to obtain $E(T,V,N)$ will be as follows: we will invert $N(T,V,\mu)$ given by (15) so as to obtain $\mu(T,V,N)$ and substitute the latter in (14). It will appear that this is possible only if we expand all relevant quantities as power series in the ratio $k_B T / \epsilon_F$.

**Remark.** Since $(\epsilon_F / k_B T)^{3/2} = \frac{3}{8} \sqrt{\pi n} \lambda^3$ we see that $k_B T / \epsilon_F$ is small when $n \lambda^3$ is large.

We begin by considering the integral

$$I = \int_0^\infty \frac{f(\eta)d\eta}{e^{\beta(\eta-\mu)} + 1}$$

$$= \int_0^\mu d\eta f(\eta) - \int_0^\mu \frac{f(\eta)d\eta}{e^{-\beta(\eta-\mu)} + 1} + \int_\mu^\infty \frac{f(\eta)d\eta}{e^{\beta(\eta-\mu)} + 1},$$

in which $f(\eta) = \eta^{1/2}$ [see figure 13] or $= \eta^{3/2}$. To begin with, we will expand in inverse powers of $\beta \mu$, supposed to be $\gg 1$. This hypothesis will have to be checked *a posteriori* since, even though $\mu_0 \equiv \mu(T = 0) = \epsilon_F$, and therefore finite, $\mu$ is a still unknown function of $T$. [We impose that $n$ does not vary.]

In the second term in (16) we may replace the lower limit of integration by $-\infty$, which only adds a contribution that is exponentially small in $\beta \mu$. That done, we set $x = \mu - \eta$
in the second term and \( x = \eta - \mu \) in the third one. This yields

\[
I = \int_0^\mu d\eta f(\eta) + \int_0^\infty dx \frac{2x f'(\mu) + \ldots}{e^{3x} + 1} + \ldots
\]

(17)

[Asymptotic expansion discussed in class.] The result is

\[
I = \int_0^\mu d\eta f(\eta) + \frac{\pi^2}{6\beta^2} f'(\mu) + \ldots,
\]

(18)

where we used the identity

\[
\int_0^\infty \frac{z}{e^z + 1} = \frac{\pi^2}{12}.
\]

(19)

and where now the dots stand for terms of higher powers in \( 1/\beta \). This procedure is called a Sommerfeld expansion. Setting \( f(\eta) = \eta^{1/2} \) and \( = \eta^{3/2} \) successively we find

\[
E(T,V,\mu) = \frac{V(2m\mu)^{5/2}}{10\pi^2 m\hbar^3} \left[ 1 + \frac{5\pi^2}{8} \left( \frac{k_B T}{\mu} \right)^2 + \ldots \right],
\]

(20)

\[
N(T,V,\mu) = \frac{V(2m\mu)^{3/2}}{3\pi^2 \hbar^3} \left[ 1 + \frac{\pi^2}{8} \left( \frac{k_B T}{\mu} \right)^2 + \ldots \right].
\]

(21)

Relation (21) gives \( N \) (or \( n = N/V \)) as a function of \( \mu \) and \( T \). In order to find \( \mu \) as a function of \( n \) and \( T \) we divide it by \( V \) and raise it to the power \( \frac{2}{3} \), which gives

\[
\mu = \frac{\hbar^2}{2m} \left( 3\pi^2 n \right)^{2/3} \left[ 1 + \frac{\pi^2}{8} \left( \frac{k_B T}{\mu} \right)^2 + \ldots \right]^{-\frac{2}{3}},
\]

(22)

whence

\[
\mu(T,V,N) = \mu_0 \left[ 1 - \frac{\pi^2}{12} \left( \frac{k_B T}{\mu_0} \right)^2 + \ldots \right]
\]

(23)
in which $\mu_0 = \mu_0(n)$. Substitution of (23) in (20) gives

$$E(T,V,N) = \frac{3}{5} N \epsilon_F \left[ 1 + \frac{5\pi^2}{12} \left( \frac{k_B T}{\mu_0} \right)^2 + \ldots \right].$$

(24)

Hence

$$C_V(T,V,N) = \frac{dE}{dT} = \frac{\pi^2}{2} \frac{k_B T}{\epsilon_F} N k_B + \ldots \quad (T \to 0)$$

(25)

where the dots indicate terms with higher powers of $\frac{k_B T}{\epsilon_F} = \frac{T}{T_F}$. The last two equations constitute the answer to the question that we asked at the beginning of this section. Eq. (25) shows that the electron heat capacity vanishes linearly with temperature in the $T \to 0$ limit. This contrasts with the classical expression $C_V = \frac{3}{2} N k_B$.

In contrast to expansion (VII.29), equation (25) is an expansion in negative powers of $n \lambda^3$. It is not an equation of state, but see the exercise below.

**Exercise.** Use result (24) and the relation $pV = \frac{2}{3} E$ [see Chapter VII] to obtain the equation of state near zero temperature. Write it down explicitly.

**Answer:**

$$p(n,T) = \left( \frac{3\pi^2}{5m} \right) \left[ \frac{5m}{p_0} \right] \left[ 1 + \frac{5\pi^2}{12} \left( \frac{k_B T}{\epsilon_F} \right)^2 + \ldots \right].$$

(26)

**Exercise.** Let $\Delta \epsilon$ be the width of the Fermi surface at temperature $T$. Determine the order of magnitude of $\Delta \epsilon/\epsilon_F$. Same question (in obvious notation) for $\Delta k/k_F$.

**Remarks and comments**

(1) **The final result.** We recall that expression (25) represents the heat capacity of the electrons in a metal; as we have seen, the lattice vibrations (phonons) contribute an amount $\propto T^3$. In the low temperature limit the total heat capacity of a solid is therefore given by

$$C_V^\text{total} = C_V^\text{el} + C_V^\text{ph} = AT + BT^3, \quad T \to 0.$$  

(27)

The electron contribution becomes visible only at very low $T$, typically of the order of a few degrees K. Observe that $A$ and $B$ are fully determined by $T_F$ and $T_D$ (the Fermi and the Debye temperature), respectively.

(2) **Excitations.** For $T > 0$ excitations from the ground state are possible: an electron close to but just inside the Fermi surface occupies a position just outside that surface and
leaves a “hole.” The Fermi surface becomes slightly diffuse. This is best represented by the occupation number $\left< n_{k\sigma} \right>^{FD}$ shown as a function of the energy $\epsilon_k$ [see figure 13].

(3) For independent electrons that are not free but subject to a periodic potential – such as the one created by the atoms of a lattice – there is still a Fermi surface, but it is no longer spherical.
IX. IDEAL BOSE GAS

IX.1. Introduction

An important application: liquid $^4$He (see at the end of this chapter). Below we will take bosons of spin $S = 0$, hence $G = 1$.

Ground state: all particles occupy the “point” (lattice site, single-particle state) $\mathbf{k} = 0$. Therefore the ground state energy is zero. We recall [Eq. VII.(33)] that for bosons without spin

$$\langle n_k \rangle = \frac{1}{e^{-\beta(\mu - \epsilon_k)} - 1}. \quad (1)$$

The lowest single-particle energy is $\epsilon_{k=0} = 0$ and therefore we need $\mu < 0$ if all the $\langle n_k \rangle$ are to be nonnegative. If $-\beta(\mu - \epsilon_k) \ll 1$ we may write (1) as

$$\langle n_k \rangle \approx \frac{k_B T}{\frac{\hbar^2 k^2}{2m} - \mu} \quad (2)$$

with the special case

$$\langle n_0 \rangle \approx -\frac{k_B T}{\mu}. \quad (3)$$

This shows that the average occupation number $\langle n_0 \rangle$ may increase without bounds as $\mu \to 0^-$, while those with $k \neq 0$ stay finite. But in that case we have to redo the mathematics of Chapter VII.

IX.2. Particle density and condensate

In order to derive Eq. (VII.34) for the particle density $n$ we replaced $\sum_k$ with $V(2\pi)^{-3} \int d\mathbf{k}$, which required the summand to be a smooth function of $\mathbf{k}$. This operation is certainly not allowed when the term with $\mathbf{k} = 0$ dominates all the others, in which case it has to be treated separately. Instead of (34) we therefore have for the density

$$n = \frac{1}{V} \sum_k \langle n_k \rangle$$

$$= \frac{1}{V} \langle n_0 \rangle + \frac{1}{(2\pi)^3} \int d\mathbf{k} \frac{e^{\beta(\mu - \epsilon_k)}}{1 - e^{\beta(\mu - \epsilon_k)}}, \quad (4)$$

In this last expression we now expand the denominator in a power series in $e^{\beta(\mu - \epsilon_k)}$, use that $\epsilon_k = \frac{\hbar^2 k^2}{2m}$, pass to spherical coordinates in $\mathbf{k}$ space, and integrate term by term with the aid of the identity

$$\int_0^\infty dx \ x^2 e^{-x^2} = \frac{\sqrt{\pi}}{4}, \quad (5)$$

and multiply by $\lambda^3$. This leads to (check this as an exercise)

$$n\lambda^3 = \frac{\lambda^3}{V} \frac{1}{e^{-\beta \mu} - 1} + \sum_{\ell=1}^{\infty} \frac{e^{\beta \mu \ell}}{\ell^{3/2}}. \quad (6)$$
Figure 14: Reduced density $n\lambda^3$ as a function of the fugacity $e^{\beta\mu}$ for the ideal Bose gas. For comparison the straight line corresponding to the classical gas is also shown.

The second term on the RHS above is identical to what one finds from Eq. (VII.26). For $\mu \to 0^-$ this term tends to

$$\sum_{\ell=1}^{\infty} \frac{1}{\ell^{3/2}} = \zeta\left(\frac{3}{2}\right) = 2.612 \ldots,$$

where $\zeta$ is the Riemann zeta function; the finiteness of this limit value is proof that the excited levels need no special treatment as $\mu \to 0^-$. For small $\beta\mu$ the first term on the RHS may be written as

$$\frac{\lambda^3}{V} \frac{1}{e^{-\beta\mu} - 1} \approx -\frac{\lambda^3 k_B T}{V \mu},$$

which for $\mu \to 0^-$ tends to infinity.

Remark. In the second term the thermodynamic limit $V \to \infty$ has been taken (where?). In the first term it is not allowed to let $V \to \infty$, which would reduce it to zero. The correct procedure is to obtain first all final results for a finite system and only then to let $V \to \infty$.

The first term cannot be neglected compared to the second one only if

$$-\frac{k_B T \lambda^3}{V \mu} \gtrsim 2.612 \quad \text{or} \quad -\frac{k_B T \lambda^3}{2.612 V} \lesssim \mu < 0,$$

that is, for $\mu$ in an interval that tends to zero in the thermodynamic limit. See the figure [not yet included], which represents the number of particles in a volume $\lambda^3$ as a function of $e^{\beta\mu}$ [or: as a function of $\mu$ when $T$ is kept fixed].
Distribution of the particles in $k$ space. When the chemical potential $\mu$ is increased, the density increases; but when $\mu$ gets very close to zero, all new particles are going to occupy the level $k = 0$. In the thermodynamic limit we therefore have a peak at $k = 0$ plus the curve that represents the second term. See figure.

As before, we are interested in the dependence of the observable physical quantities on $T$ and $n$, rather than $\mu$ and $n$. Our procedure will be as follows. Let $n$ be fixed. The figure then shows $\lambda^3$ [which is $\propto T^{-3/2}$] as a function of $\mu/k_B T$. Inversely we may view this curve as representing $\mu/k_B T$ as a function of $\lambda^3$ and observe that $\mu/k_B T$ continues to increase when $T$ goes down so long as $T$ has not reached a value $T_c$ (corresponding to a thermal wavelength $\lambda_c$) which is the solution of

$$n\lambda_c^3 = \frac{nh^3}{(2\pi mk_B)^{3/2}} = 2.612,$$

that is,

$$T_c = \frac{h^2}{2\pi mk_B} \left( \frac{n}{2.612} \right)^{2/3}.$$  

We call $T_c$ a critical temperature. Let us now consider expression (6) divided by $\lambda^3$. For $T > T_c$ it suffices to take the second term into account. For $T < T_c$ one may set $\mu = 0$ in the second term which may then be rewritten as

$$\frac{2.612}{\lambda^3} = \left( \frac{T}{T_c} \right)^{3/2} n.$$  

[This term is in fact $\propto T^{3/2}$ and independent of $n$: the $n$ cancels against the one in the definition of $T_c$.] Hence

$$n = \begin{cases} 
\frac{1}{\lambda^3} \sum_{\ell=1}^{\infty} e^{\beta \mu \ell} \ell^{3/2} & T > T_c, \\
-\frac{1}{V\beta \mu} + \left( \frac{T}{T_c} \right)^{3/2} n & T < T_c.
\end{cases}$$  

Without the need of knowing $\mu$ for $T < T_c$ we now see that the particle density in the state $k = 0$ is given by

$$n_0 = n - \frac{2.612}{\lambda^3} = \left[ 1 - \left( \frac{T}{T_c} \right)^{3/2} \right] n, \quad T < T_c,$$

where now the limit $V \to \infty$ has been taken. This phenomenon is called Bose-Einstein condensation. The particles in the single-particle state $k = 0$ are called the condensate.

---

29We introduce the density $n_0 \equiv \langle n_\vec{0} \rangle/V$; note that $n_\vec{0}$ is a dimensionless number. CT writes $N_0$ instead of our $n_0$. 

84
It is a “condensation” (in the sense of an accumulation or piling up) not in ordinary space but in k space. In the figure [not available yet] we show the number of particles in a shell between $k$ and $k + dk$, namely $4\pi k^2 \langle n_k \rangle dk$.

**IX.3. Thermodynamics**

We investigate several more thermodynamic relations. The particles with $k = 0$ do not contribute to the energy $E$ of the system, which is given by

$$\frac{E}{V} = \frac{1}{V} \sum_{k \neq 0} \langle n_k \rangle \epsilon_k$$

$$= \frac{1}{(2\pi)^3} \int dk \epsilon_k \frac{e^{\beta(\mu - \epsilon_k)}}{1 - e^{\beta(\mu - \epsilon_k)}}$$

$$= \ldots$$

$$= \frac{3}{2} \frac{k_B T}{\lambda^3} \sum_{\ell=1}^{\infty} \frac{e^{\beta \mu \ell}}{\ell^{5/2}}.$$

The calculation is analogous to a preceding one; we used the identity

$$\int_0^\infty dx x^4 e^{-x^2} = \frac{3}{8} \sqrt{\pi}.$$

The formulas (VI.22) and (VI.24) derived earlier for $\log Z_g^{BE}$ and $p^{BE}$, respectively, continue to hold. The replacement of the sum on $k$ by an integral carried out there to obtain them was innocent. The particles in the condensate therefore do not contribute to the pressure (the point $k = 0$ has zero weight in an integration).

If we compare (VI.24) for $p^{BE}$ with the one above for $E/V$ we see that

$$p = \frac{2E}{3V}.$$
For \( T < T_c \) the two formulas take a special form since \( \mu = 0 \),

\[
p = \frac{2E}{3V} = \frac{k_B T}{\lambda^3} \sum_{\ell=1}^{\infty} \frac{1}{(\ell^{5/2})} = \zeta\left(\frac{5}{2}\right) \frac{k_B T}{\lambda^3} = \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} n k_B \frac{T^{5/2}}{T_c^{3/2}}; \quad T < T_c.
\]

Now \( T_c \) is proportional to \( n^{2/3} \); the last line therefore implies that below \( T_c \) the pressure does not depend on the density \( n \). If at fixed \( T \) one increases \( n \)

(i) \( p \) at first increases as it would in an ordinary gas;

(ii) at a well-defined critical density \( n = n_c \), which when \( T \) is fixed follows from the condition \( n_c \lambda^3 = 2.612 \), the system enters its low temperature phase: a condensate forms;

(iii) when the density is made to further increase, the pressure stays at the value it had at the critical density and which is proportional to \( T^{5/2} \);

(iv) explanation: all particles that one continues to add to the system will occupy the single-particle state \( k = 0 \).

Since \( C_V = \frac{\partial E}{\partial T} \) we have for \( T < T_c \) the proportionality \( C_V \propto T^{3/2} \).

**Exercise.** Derive from Eq. (18) the expression for \( C_V \) in the temperature regime \( 0 \leq T < T_0 \). Show that at the critical temperature \( T = T_c^- \) we have the relation

\[
\frac{C_V}{k_B N} = 1.93.
\]

**IX.4. Liquid \(^4\text{He}\)**

Application: liquid \(^4\text{He}\). Recall Eq. (11) for the critical temperature. Upon substituting the values of \( m \) and \( n \) for liquid \(^4\text{He}\) one finds

\[
T_c = 3.14 K.
\]

In reality \(^4\text{He}\) undergoes a phase transition at \( T_\lambda = 2.17 K \), called the *lambda transition* because of the shape of the specific heat shown in the figure. It is, in fact, a Bose-Einstein condensation *modified* by the effects of the interatomic interactions. Below its critical temperature liquid \(^4\text{He}\) is *superfluid*, that is, it flows without resistance.

See Diu, complément VI.D (in particular pp. 892, 893), or J. Bardeen in *Physics Today* Vol. 43 no. 12 (December 1990), page 25 (this article also mentions superconductivity).
X. PHOTON GAS

X.1. Another view of the harmonic oscillator

Recall that a quantum harmonic oscillator of frequency \( \omega \) has energy levels \((n + \frac{1}{2})\hbar \omega\), where \(n = 0, 1, 2, \ldots\). The canonical partition function therefore is

\[
Z_c = e^{-\beta \hbar \omega / 2} \sum_{n=0}^{\infty} e^{-n\beta \hbar \omega}.
\]  

(1)

We may look at this formula differently and say that there exists one energy level, \( \hbar \omega \), which may be occupied by an arbitrary number, \(n\), of fictitious particles (actually quanta of oscillation, also called quasi-particles). These particles are therefore bosons and \(Z_c\) is their grand-canonical partition function at chemical potential \(\mu = 0\).

More generally, if a physical system is composed of \(N_0\) oscillators of frequencies \(\omega_1, \omega_2, \ldots, \omega_{N_0}\), then its state is specified by the set of “occupation numbers” of these oscillators: \(n_1, n_2, \ldots, n_{N_0}\). In that case, as we have seen in Chapter V, the expression for the partition function, here now denoted by \(Z\), is

\[
Z = \frac{e^{-\frac{1}{2} \beta \sum_i \omega_i}}{Z_0} \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \ldots \sum_{n_{N_0}=0}^{\infty} e^{-n_1\beta \omega_1 + n_2\beta \omega_2 + \ldots + n_{N_0}\beta \omega_{N_0}} \frac{1}{Z'}
\]

(2)

in which \(Z'\) is again a grand-canonical partition function at \(\mu = 0\), analogous to the ones encountered in Chapter IX. For the free energy one has

\[
F = -k_B T \log Z = -k_B T \log Z_0 - k_B T \log Z' = \frac{1}{2} \sum_i \hbar \omega_i + k_B T \sum_i \log(1 - e^{-\hbar \omega_i / k_B T})
\]

(3)

in which the zero point energy \(F_0\) is the contribution to the free energy stemming from the ground state energy of the oscillators. Since \(F_0\) is a constant and since all physical observables may be expressed as derivatives of \(F\), we may work with \(F'\) instead of \(F\), and will do so below.\(^{30}\)

A difference with the bosons studied in Chapter IX is that here the total boson number, \(N = \sum_i n_i\), is not conserved and therefore does not play the role of an independent

\(^{30}\)We are in fact eliminating a bit too easily here a nontrivial problem: in the case of electromagnetic radiation one has \(F_0 = \infty\). See section X.5.
parameter. At a given temperature $T$ the system will automatically settle in an equilibrium state with an average $\langle N \rangle$ that is some function of $T$. This equilibrium is not attained as the result of the exchange of a conserved number of bosons with a reservoir, but by their creation and destruction. There is therefore here one fewer independent parameter than in the case of nondestructible bosons: $(T, V)$ replaces $(T, V, N)$ or $(T, V, \mu)$.

### X.2. Maxwell’s equations

The Maxwell equations are classical equations: they are self-contained and without any reference to quantum mechanics. We recall here a few things that you know from your course on Electromagnetism. Let $\mathbf{E}(r, t)$ denote a space and time dependent electric field. Inside a vacuum cavity (that we will take cubic and of volume $V = L^3$) the electric field must satisfy

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \Delta \mathbf{E}, \quad \text{div} \mathbf{E} = 0, \quad (4)$$

in which $c$ is the speed of light in vacuum. Let us look for eigensolutions of (4) by supposing that they have the form of running waves

$$\mathbf{E}(r, t) = \mathbf{E}_0 e^{i \mathbf{k} \cdot \mathbf{r} - i \omega t} \quad (5)$$

in which for the moment $\mathbf{k}$ and $\omega$ are arbitrary. By substitution of (5) in (4) we are led to the following conclusions.

1. It is necessary for $\omega$ to be related to $\mathbf{k}$ by

$$\omega = \omega_k = ck, \quad k \equiv |\mathbf{k}|. \quad (6)$$

2. For periodic boundary conditions $\mathbf{k}$ can take only the values

$$\mathbf{k} = \frac{2\pi}{L} \mathbf{n} \quad (7)$$

where $\mathbf{n}$ is a vector of arbitrary integer components.

3. The vector $\mathbf{E}_0$ must be perpendicular to $\mathbf{k}$, that is, $\mathbf{E}_0 \cdot \mathbf{k} = 0$. Hence $\mathbf{E}_0$ is a linear combination of two unit vectors that are orthogonal to each other and to $\mathbf{k}$, and that we will denote by

$$\mathbf{e}_{k\lambda} \quad \text{with } \lambda = 1, 2 \text{ the polarization index}. \quad (8)$$

We have therefore found eigensolutions [or: vibrational modes] of the electric field (4) that are running waves characterized by a wavevector $\mathbf{k}$ and a polarization $\lambda$,

$$\mathbf{E}(r, t) = A^0_{k\lambda} \mathbf{e}_{k\lambda} e^{i \mathbf{k} \cdot \mathbf{r} - i \omega_k t}, \quad (9)$$

where $A^0_{k\lambda}$ is an amplitude. The general solution $\mathbf{E}(r, t)$ is a sum of (9) over $\mathbf{k}$ and $\lambda$ with arbitrary\footnote{Do not confuse with a wavelength, often also indicated by the same letter.} amplitudes.

\footnote{The electric field is an observable taking real values and therefore we must have $A^0_{-k\lambda} = A^0_{k\lambda}$.}
X.3. Quantizing the electromagnetic field

The eigensolution (9) may be expressed alternatively as

\[ \mathcal{E}(r,t) = A_{k\lambda}(t) e^{ik \cdot r}, \]  

(10)

which is the product of a time dependent amplitude and a fixed function of space. The amplitude \( A_{k\lambda}(t) \) satisfies

\[ \frac{d^2 A_{k\lambda}}{dt^2} + c^2 k^2 A_{k\lambda} = 0, \]  

(11)

as may be verified directly or by substituting (10) in (4). Eq. (11) may be compared to the classical harmonic oscillator with Hamiltonian

\[ H = \frac{p^2}{2m} + m\omega^2 x^2 / 2 \]  

(12)

The vibrational mode \((k, \lambda)\) of the electromagnetic field is therefore a harmonic oscillator of frequency \( \omega = ck \). With this established we may now quantize the electromagnetic field according to the rules of quantum mechanics. It follows that the energy stored in the vibrational mode \((k, \lambda)\) can take only the discrete values

\[ \epsilon_{k\lambda} = (n_{k\lambda} + \frac{1}{2}) \hbar \omega_k, \quad n_{k\lambda} = 0, 1, 2, \ldots. \]  

(13)

We call \( n_{k\lambda} \) the number of photons of type \((k, \lambda)\). Photons, therefore, are bosons. What is more, they are strictly noninteracting among themselves and therefore constitute a “true” ideal gas.

Remarks.

(1) The polarization index \( \lambda \) is analogous to the spin \( \sigma \). A photon has a spin \( S = 1 \), yet \( \lambda \) takes only two values and not \( 2S + 1 = 3 \). This is due to the fact that the photon mass is zero (see a book on quantum mechanics; there is no longitudinal polarization).

(2) There is no Bose-Einstein condensation since always \( \mu = 0 \). There is no control parameter allowing us to fix the total number of photons.

(3) The number \( N_0 \) of modes \((k, \lambda)\) is infinite and since classically in thermal equilibrium each of these modes would contain an energy \( k_B T \), this would lead to a problem: the ultraviolet catastrophe (see below).

X.4. The black body: a photon gas at equilibrium

[Or: Thermodynamics of the radiation field.] By a black body\(^{33}\) we mean a vacuum cavity (or “oven”) whose walls, having a fixed temperature \( T \), are capable of absorbing and emitting radiation of all wavelengths (this latter feature earns it the qualification “black”). The balance between absorption and emission will lead to an equilibrium in which the cavity is filled with a photon gas at temperature \( T \), independently of the details of the emission and absorption processes; it will be entirely determined by

(i) the temperature \( T \) of the walls;

\(^{33}\)Do not confuse with a black hole, which is an altogether different thing!
(ii) the energy levels (modes) of the electromagnetic field inside the cavity. The radiation field thus established is called black-body radiation or sometimes also thermal radiation.

X.4.1. Spectral energy density

One may study the black-body radiation by making a small hole in the cavity (of negligible size) and measuring the radiation that comes out. The quantity easiest to measure is the intensity of the radiation in a small frequency interval between \( \omega \) and \( \omega + \, d \omega \).

According to Chapter IX we have at equilibrium

\[
\langle n_{k\lambda} \rangle = \frac{1}{e^{\beta \hbar \omega_k} - 1}.
\]  

(14)

The number of energy levels corresponding to a wavevector of length between \( k \) and \( k + \, dk \) is equal to

\[
2 \frac{V}{(2\pi)^3} 4\pi k^2 \, dk = \frac{V \omega^2}{\pi^2 c^3} \, d\omega \equiv \rho(\omega) \, d\omega.
\]  

(15)

The number \( n(\omega, T) \, d\omega \) of photons having an energy between \( \hbar \omega \) and \( \hbar \omega + \hbar \, d\omega \) is this same number multiplied by (14),

\[
n(\omega, T) \, d\omega = \frac{V}{\pi^2 c^3} \frac{\omega^2}{e^{\beta \hbar \omega} - 1} \, d\omega.
\]  

(16)

The energy \( dE \) of the photons in this frequency interval is

\[
dE = \hbar \omega n(\omega, T) \, d\omega = V \, \hbar \frac{\omega^3}{\pi^2 c^3} \frac{1}{e^{\beta \hbar \omega} - 1} \, d\omega,
\]

(17)

where \( u(\omega, T) \) is the spectral energy per unit of volume in the cavity, which is a measurable quantity. This expression for \( u(\omega, T) \) is called Planck’s law (1900) for black-body radiation. It is valid for all \( \omega \) and all \( T \), without any restriction, at the sole condition that the system be at equilibrium.

It is of interest to consider two limits of Planck’s law.

(i) For \( \hbar \omega \ll k_B T \) we obtain the classical limit,

\[
u(\omega, T) \simeq \frac{k_B T}{\pi^2 c^3} \omega^2,
\]

(18)

which is called the Rayleigh-Jeans formula (19th century) and does not contain \( \hbar \). When integrated over \( \omega \) it diverges at high frequencies (the “ultraviolet catastrophe”), which

---

\(^{34}\)Here: density = energy per unit interval on the frequency axis.
brings out the necessity for a modification of the classical theory. (ii) For $\hbar \omega \gg k_B T$ we obtain
\begin{equation}
    u(\omega, T) \simeq \frac{\hbar}{\pi^2 c^3} \omega^3 e^{-\hbar \omega / k_B T},
\end{equation}
which is called Wien’s law and was discovered empirically in 1893.

**Exercise.** Show that for given $T$ the spectral energy density $u(\omega, T)$ has its maximum for $\omega = \omega_m$ where
\begin{equation}
    \omega_m = 2.821 \frac{k_B T}{\hbar}.
\end{equation}
The proportionality of this maximum to $T$ is called Wien’s displacement law.

All matter radiates and in fact many ordinary objects emit thermal radiation approximately as black bodies. Thermal radiation at room temperature peaks at infrared wavelengths. The sun emits thermal radiation at a temperature of about 5250 K (the temperature of the surface of the sun), which peaks in the visible light.

**X.4.2. Thermodynamics**

**Free energy.** We start from (3). The free energy $F'$ is given by
\begin{equation}
    F' = 2 \frac{V}{(2\pi)^3} k_B T \int dk \log(1 - e^{-\hbar \omega / k_B T})).
\end{equation}
Use now (15) to write
\begin{equation}
    2 \frac{V}{(2\pi)^3} \int dk = \int_0^\infty d\omega \rho(\omega) = \frac{V}{\pi^2 c^3} \int_0^\infty d\omega \omega^2.
\end{equation}
Then, setting $x = \hbar \omega / k_B T$, we get
\begin{equation}
    F'(T, V) = \frac{V k_B T}{\pi^2} \left( \frac{k_B T}{\hbar c} \right)^3 \int_0^\infty dx x^2 \log(1 - e^{-x})
    \left[ -\frac{1}{3} \int_0^\infty dx \frac{x^3}{e^x - 1} \right]_{n^4/15}
    = -\frac{\pi^2}{45} k_B T \left( \frac{k_B T}{\hbar c} \right)^3 V.
\end{equation}

**Energy.** The total energy $E$ contained in the cavity may be obtained either from (23) according to $E = -T^2(\partial/\partial T)(F'/T)$ or by integrating the spectral energy density (17).
The latter method leads to

\[ E(T, V) = V \int_0^\infty d\omega u(\omega, T) \]

\[ = \frac{V h}{\pi^2 c^3} \left( \frac{k_B T}{\hbar} \right)^4 \int_0^\infty dx \frac{x^3}{e^x - 1} \]

\[ = \frac{\pi^2}{15} V h c \left( \frac{k_B T}{\hbar c} \right)^4 \]

\[ = \frac{\pi^2}{15} k_B T \left( \frac{k_B T}{\hbar c} \right)^3 V, \]  \hspace{1cm} (24)

that is, \( E(T, V) \) is proportional to \( T^4 \). This is called the Stefan-Boltzmann law.

To interpret this formula, recall that a photon has a wavelength \( \lambda \) that may be written

\[ \lambda = \frac{2 \pi}{k} = \frac{2 \pi c}{\omega} = \frac{2 \pi \hbar c}{\omega \hbar c}. \]

Therefore \( 2 \pi \hbar c/k_B T \) is the wavelength of a photon having a thermal energy and is analogous to the de Broglie wavelength of a massive particle.

Radiation pressure. We obtain directly from (23)

\[ p = -\frac{\partial F'}{\partial V} = \frac{\pi^2}{45} \left( \frac{k_B T}{\hbar c} \right)^3 k_B T \]  \hspace{1cm} (25)

and hence, because of (24),

\[ pV = \frac{1}{3} E, \]  \hspace{1cm} (26)

which may be compared to the analogous relations for a classical ideal gas, for which \( pV = \frac{2}{3} E \). Microscopically this pressure is due to the collisions of the photons against the walls. (Can you derive it that way?)

X.5. Casimir force

Exercise. Strictly speaking, the free energy contains a term \( F_0 \), the zero-point (or “vacuum”) energy of the electromagnetic field. We have suppressed it right from the beginning by arguing that it is a constant and that therefore its derivatives vanish. Do you agree with this for the calculation of \( E \) ? For the calculation of \( p \)?

The origin of the Casimir force (1948) between two parallel metal plates in vacuum at distance \( L \) was discussed in class. This attractive force is equal to

\[ F_{\text{Cas}} = -\frac{\pi^2 \hbar c}{240 L^4} \text{ per unit area}. \]

The proof of this formula is beyond the scope of our course. Show, however, that the only combination of \( \hbar, c, \) and \( L \) that makes a pressure is \( \hbar c/L^4 \).
X.6. Cosmic microwave background (CMB)

[Sometimes referred to as relic radiation.
In French: Fond diffus cosmologique or rayonnement fossile.]

During the first \( \approx 0.3 \times 10^6 \) years of its history, the temperature of the universe was above 3000 K and there were many free electric charges around, not combined into neutral atoms. Since the interaction between photons and charged particles is very strong, during that period the photons were being scattered, emitted, and absorbed all the time, which led to an equilibrium between photons and matter. The photons were distributed according to Planck’s law at the prevailing temperature. The universe is expanding and therefore getting colder. When the temperature went down below 3000 K, neutral atoms were formed and the scattering cross section for photons off matter decreased sharply: the universe became “transparent” for photons. Said differently, “radiation decoupled from matter.” Put more quantitatively, the probability for a photon that was free at that moment to have been scattered or absorbed by a particle of matter, between then and now, is negligible. These photons now fill the large empty intergalactic spaces.

The only change is that, since the universe has expanded, all distances have increased since the moment of decoupling by a factor that may be calculated and that is close to 1000. In particular, the photon wavelength has increased by this factor. Planck’s law at the moment of decoupling is thereby transformed into a new law of Planck for a radiation in equilibrium at a temperature a thousand times lower, that is, at \( T \approx 3 \) K. Its spectral density peaks in the range of microwave frequencies.

This radiation is isotropic (arrives on earth with the same intensity from all directions). Its existence was predicted in 1948 and it was discovered accidentally in 1965 by A. Penzias and K. Wilson (Nobel prize 1978). It has been studied since the 1990’s with the aid of satellites (COBE, WMAP, Planck). Its exact temperature has been measured with great accuracy and was determined in 2009 as \( 2.72548 \pm 0.00057 \) K. Small fluctuations of this temperature according to the direction in the sky are of great interest to cosmologists since they contain information about what happened in the early universe.
XI. CORRELATION FUNCTIONS

XI.1. The single-particle density

Let us consider the microstate \( \Gamma = (p_1, \ldots, p_N; r_1, \ldots, r_N) \) that was at the basis of the classical ensembles in Chapters I, II, III, and V. Let \( R \) be some point in space. Remember that we suppose our particles to be point-like. When the system is in microstate \( \Gamma \), then the microscopic particle density \( n(R; \Gamma) \) in \( R \) is zero if there is no particle and infinite if there happens to be one of the \( N \) particles. The function \( A(\Gamma) \) [Chapter I, Eq. (21)] that expresses this is therefore

\[
n(R; \Gamma) = \sum_{i=1}^{N} \delta(r_i - R). \tag{1}\]

Attention: The \( r_i \) are dynamical variables (degrees of freedom of the system), whereas \( R \) is a parameter representing a fixed position.

**Exercise.** What quantity is represented by \( \int_{V \subset V'} dR n(R; \Gamma) \)?

The average density in \( R \) will be denoted by \( n(R) \) and is defined as the ensemble average of \( n(R; \Gamma) \) according to

\[
n(R) = \frac{\int d\Gamma \rho(\Gamma) n(R; \Gamma)}{\int d\Gamma \rho(\Gamma)} \tag{2}\]

in which \( \rho = \rho_m \) or \( \rho = \rho_c \). This is the density that is usually measured in experiments. Note that it may depend on \( R \) even in equilibrium, for example when the system is subject to an external potential.

**Exercise.** How should expression (2) for \( n(R) \) be adapted to the grand-canonical ensemble? Check that

\[
\int_V dR n(R) = \left\{ \begin{array}{ll}
N & \text{in the ensembles ‘m’ and ‘c’}, \\
\langle N \rangle_g & \text{in ensemble ‘g’}. 
\end{array} \right. \tag{3}
\]

From here on we will work in the canonical ensemble, so that \( \rho(\Gamma) = \exp[-\beta \mathcal{H}(\Gamma)] \).

Let as before

\[
\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{1\leq i<j\leq N} U(r_{ij}) + \sum_{i=1}^{N} U_{\text{ext}}(r_i). \tag{4}
\]

We will let \( Z_N \) denote the configurational partition function\(^35\)

\[
Z_N = \int_V \text{d}r_1 \ldots \int_V \text{d}r_N e^{-\beta U_N(r_1, \ldots, r_N)}. \tag{5}
\]

\(^35\)The same quantity was called \( Q_N \) in Chapter V.5.
Substitution of (1) and (4) in (2) gives

\[
n(R) = Z_N^{-1} \int_V dr_1 \ldots \int_V dr_N \sum_{i=1}^N \delta(r_i - R) e^{-\beta U_N(r_1, \ldots, r_N)}
\]

whence

\[
n(r_1) = \frac{N}{Z_N} \int_V dr_2 \ldots \int_V dr_N e^{-\beta U_N(r_1, \ldots, r_N)},
\]

where \( r_1 \) has become a parameter. Eqs. (6) and (7) show that \( n(R) \) is proportional to the partition function restricted to the microstates that have a particle in \( R \).

**Exercise.** Show that for an ideal gas in an external potential \( U_{\text{ext}}(R) \) one has

\[
n(R) = C e^{-\beta U_{\text{ext}}(R)},
\]

and find the expression for \( C \).

The dimensionless quantity \( n(R) \Delta R \) represents the average number of particles in a volume element of size \( \Delta R \) at \( R \). When \( n(R) \Delta R \ll 1 \), it is also the probability that the volume element of size \( \Delta R \) around \( R \) contain a particle. (Question: why is the condition \( \ll 1 \) needed in this last statement?) For a spatially homogeneous system \( n(R) = N/V \equiv n \).

**XI.2. The \( p \)-particle density**

**XI.2.1. \( p \)-particle correlations**

We now ask about the probability \( n^{(p)}(R_1, \ldots, R_p) dr_1 \ldots dr_p \) that \( p \) infinitesimal volume elements around given points be occupied by particles. This requires averaging a generalization of (1),

\[
n^{(p)}(R_1, \ldots, R_p; \Gamma) = \left[ \sum_{i=1}^N \delta(r_i - R_1) \right] \ldots \left[ \sum_{i=1}^N \delta(r_i - R_p) \right]
\]

and one finds\(^{36}\)

\[
n^{(p)}(r_1, \ldots, r_p) = \frac{N!}{(N-p)!} Z_N \int_V dr_{p+1} \ldots \int_V dr_N e^{-\beta U_N(r_1, \ldots, r_N)},
\]

which is called the \( p \)-particle density.

\(^{36}\)You have to fill \( p \) positions with \( p \) particles selected among \( N \); there are \( \binom{N}{p} \) possible selections and then \( p! \) ways to do the filling.
XI.2.2. Pair correlation

An important case is \( p = 2 \). One calls \( n^{(2)}(\mathbf{R}_1, \mathbf{R}_2) \) the pair density. For a system which after ensemble averaging is translation and rotation invariant we have

\[
n^{(2)}(\mathbf{R}_1, \mathbf{R}_2) = n^{(2)}(\mathbf{R}_1 - \mathbf{R}_2) = n^{(2)}(|\mathbf{R}_1 - \mathbf{R}_2|) = n^{(2)}(R_{12}).
\]

(11)

Recall that the thermodynamic limit consists in letting \( N, V \to \infty \) in a fixed ratio \( n = N/V \). After this limit has been taken one may consider the limit \( |\mathbf{R}_1 - \mathbf{R}_2| \to \infty \). It should be expected that

\[
n^{(2)}(R_{12}) \to n^2 \quad \text{for } R_{12} \to \infty
\]

(12)

(why?) and therefore we set

\[
n^{(2)}(R_{12}) = n^2 g(R_{12}).
\]

(13)

Here the dimensionless quantity \( g(R) \) is called the pair distribution function and is such that \( g(\infty) = 1 \). Finally we introduce

\[
h(R) = g(R) - 1,
\]

(14)

where \( h(R) \) is called the (pair) correlation function and is such that \( h(\infty) = 0 \).

**Remark.** Eq. (9) leads to (10) only if \( \mathbf{R}_1, \ldots, \mathbf{R}_p \) are all different. One may also consider the RHS of (9) for arbitrary \( \mathbf{R}_1, \ldots, \mathbf{R}_p \), not necessarily all different. For \( p = 2 \) we then get

\[
\left[ \sum_{i=1}^{N} \delta(\mathbf{r}_i - \mathbf{R}_1) \right] \left[ \sum_{i=1}^{N} \delta(\mathbf{r}_i - \mathbf{R}_2) \right] = \sum_{i=1}^{N} \frac{\delta(\mathbf{r}_i - \mathbf{R}_1)}{\delta(\mathbf{r}_1 - \mathbf{R}_2)} \delta(\mathbf{r}_i - \mathbf{R}_2)
\]

\[
+ \sum_{i \neq j} \delta(\mathbf{r}_i - \mathbf{R}_1) \delta(\mathbf{r}_j - \mathbf{R}_2)
\]

\[
= \delta(\mathbf{R}_1 - \mathbf{R}_2)n(\mathbf{R}_1; \Gamma) + n^{(2)}(\mathbf{R}_1, \mathbf{R}_2; \Gamma).
\]

(15)

This expression will be useful in the next section.

In order to discuss the behavior of the correlation function, let us consider for definiteness a fluid composed of particles interacting by a Lennard-Jones potential,

\[
V_{\text{LJ}}(r) = \epsilon_0 \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right].
\]

(16)

This potential reproduces in a good approximation the force acting between two noble gas atoms (Ne, Ar, Xe, . . .). The parameters \( \epsilon_0 \) and \( \sigma \) (see figure [not included]) are adjustable to the atom species under consideration.

---

37 Except in case (c) below.
38 This same name is often also applied to \( g(R) \) or to \( n^{(2)}(R) \).
39 The behavior \( \propto r^{-6} \) as \( r \to \infty \) follows from a quantum mechanical calculation that takes into account that atoms, although neutral, have fluctuating dipole moments.
(a) At high temperature and relatively low density the system is in its gas phase and
the correlation function \(g(r)\) behaves as in the figure [not included].

(b) At lower temperature and higher density it is in its liquid phase and \(g(r)\) has a
behavior as in the figure [not included]. The same behavior prevails in an amorphous
solid (as may be obtained, e.g., by rapidly quenching a liquid), which however is not an
equilibrium state.

(c) At low enough temperature the system will become a crystalline solid and \(g(r)\)
will behave as in the figure [not included]. In this case it will no longer tend to unity as
\(r \to \infty\); we say that the system has long range order.

XI.2.3. Correlation length

The oscillations of \(g(R)\), if any, occur at a scale of typically less than a few tens of Å.
They are representative for the degree of local order in the fluid. For \(R \to \infty\) (in practice
a few tens or at most a few hundreds of Å) one may ask how \(g(R)\) tends to its limit value
\(g(\infty) = 1\). Most often this approach will be exponential and one sets
\[
\Delta (R) = g(R) - 1 \sim e^{-R/\xi}
\]
where \(\xi(n, T)\) is the correlation length and depends on the two intensive variables \(n\) and
\(T\). Typically, \(\xi\) increases with decreasing temperature. Under special circumstances not
to be discussed here the correlation length may tend to infinity: the system is then said
to become ‘critical.’

XI.3. Structure factor

The (static) structure factor \(S(q)\) is defined as
\[
S(q) = \frac{1}{N} \left\langle \left| \sum_{i=1}^{N} e^{-iq \cdot r_i} \right|^2 \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\langle e^{-iq \cdot (r_i - r_j)} \right\rangle.
\]

We may relate this quantity to the pair density by writing
\[
S(q) = \frac{1}{N} \int_V dR_1 \int_V dR_2 \ e^{-iq \cdot (R_1 - R_2)} \left\langle \sum_{i=1}^{N} \delta(r_i - R_1) \sum_{j=1}^{N} \delta(r_j - R_2) \right\rangle \nonumber
\]
\[
= 1 + \frac{1}{N} \int_V dR_1 \int d(R_2 - R_1) e^{iq \cdot (R_2 - R_1)} n^{(2)}(R_2 - R_1) \nonumber
\]
\[
= 1 + n \int dR e^{iq \cdot R} g(R) \nonumber
\]
\[
= 1 + (2\pi)^3 n \delta(q) + n \int_V dR e^{iq \cdot R} [g(R) - 1]
\]
so that (for $q \neq 0$)

$$S(q) = 1 + n \int_V dR e^{i q \cdot R} h(R).$$ \hspace{1cm} (20)

The structure factor $S(q)$ derives its importance from the fact that it may be measured directly in an experiment of elastic scattering of slow neutrons or of X rays. See figure [not yet available].

In a point $r$ in space the amplitude of the scattered wave is equal to a sum of scattering contributions from all particle positions $r_i$, that is,

$$\psi_{\text{scatt}}(r) \propto \sum_{i=1}^{N} e^{i k_0 \cdot r_i} e^{i k \cdot (r - r_i)} = e^{i k \cdot r} \sum_{i=1}^{N} e^{-i q \cdot r_i}. \hspace{1cm} (21)$$

where $q = k - k_0$ and $|k| = |k_0|$. The scattered intensity measured in $r$ therefore becomes, after ensemble averaging,

$$I(q) \propto \langle |\psi_{\text{scatt}}(r)|^2 \rangle \propto S(q). \hspace{1cm} (22)$$

By using (22) and inverting the Fourier transform in (20) one recovers $h(R)$ and hence $g(R)$ from a measurement of $I(q)$.

**XI.4. Pair correlation and thermodynamics**

Let us consider again Hamiltonian (4), but in the absence of an external field,

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{1 \leq i < j \leq N} U(|r_i - r_j|),$$ \hspace{1cm} (23)

where we will suppose that the particles are enclosed in a volume $V = L^3$. We now write
the average potential energy $E_{\text{pot}}$ as

$$
\langle E_{\text{pot}} \rangle = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle U(|\mathbf{r}_i - \mathbf{r}_j|) \rangle
$$

$$
= \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \int d\mathbf{R}_1 \int d\mathbf{R}_2 U(|\mathbf{R}_1 - \mathbf{R}_2|) \langle \delta(\mathbf{r}_i - \mathbf{R}_1) \delta(\mathbf{r}_j - \mathbf{R}_2) \rangle
$$

$$
= \frac{1}{2} N \int d\mathbf{r} g(\mathbf{r}) U(\mathbf{r})
$$

$$
= 2\pi N \int_{0}^{\infty} d\mathbf{r} r^2 g(r) U(r).
$$

(24)

We now look for a similar expression for the pressure $p$. To that end we start from the definition

$$
\beta p = \frac{\partial \log Z_c}{\partial V} = \frac{1}{Z_c} \frac{\partial}{\partial V} \frac{1}{N! h^{3N}} \int d\mathbf{p}^N \int d\mathbf{r}^N e^{-\beta \sum_{i=1}^{N} \frac{p_{i}^2}{2m} - \beta U_N(r^N)}.
$$

(25)

How to calculate the derivative with respect to $V$? We should take into account that the limits of integration of the $\mathbf{r}_i$ depend on $V$. Here comes the trick to do that: in order to eliminate this dependence we set

$$
\mathbf{r}_i = L \mathbf{r}_i', \quad i = 1, 2, \ldots, N,
$$

(26)

so that $0 \leq x_{i}', y_{i}', z_{i}' \leq 1$ and $d\mathbf{r}^N = L^{3N} d\mathbf{r}'^N = V^N d\mathbf{r}'^N$. In terms of these new variables of integration Eq. (25) becomes

$$
\beta p = \frac{1}{Z_c} \frac{\partial}{\partial V} \frac{1}{N! h^{3N}} V^N \int d\mathbf{p}^N \int d\mathbf{r}'^N e^{-\beta \sum_{i=1}^{N} \frac{p_{i}^2}{2m} - \beta U_N(L \mathbf{r}'^N)},
$$

(27)

in which now all dependence on $V$ [or: $L$] is visible. The result of the differentiation is

$$
\beta p = \frac{N}{V} - \frac{\beta}{3V} \left\langle \sum_{i=1}^{N} \frac{\partial U_N}{\partial \mathbf{r}_i} \cdot \mathbf{r}_i \right\rangle.
$$

(28)

The negative of the expression in angular brackets is called the (internal) virial of the system. This equation is actually a reformulation of the virial theorem (see next section for a derivation from first principles; not available yet). The virial theorem relates the average kinetic energy of a many-particle system to another average that involves the forces (and hence the potential energy) between the particles.

So far we have not used for $U_N$ its specific expression (23) as a sum of pair potentials. Now note the identity

$$
\frac{\partial r_{ij}}{\partial \mathbf{r}_i} = \frac{\mathbf{r}_i - \mathbf{r}_j}{r_{ij}}.
$$

(29)
Upon using it together with the expression for $U_N$ in Eq. (27) we get

\[ p = nk_B T - \frac{1}{3V} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \frac{\partial U(r_{ij})}{\partial r_{ij}} \frac{r_i - r_j}{r_{ij}} \cdot r_j \rangle \]

\[ = nk_B T - \frac{1}{6V} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle U'(r_{ij}) \frac{r_{ij}^2}{r_{ij}} \rangle \]

\[ = nk_B T - \frac{2}{3} \pi n^2 \int_0^\infty dr \ r^3 g(r) U'(r), \tag{30} \]

the last line being obtained in the same way as the derivation of (24). When you compare (30) to (28),

\[ \frac{p}{k_B T} = n + B_2 n^2 + B_3 n^3 + \ldots, \tag{31} \]

you understand why that equation is called the “virial expansion.” The terms of order $n^3$ and higher in (31) appear when $g(r)$ in (30) is expanded as a power series in $n$.

**Exercise.** Check Eq. (30); verify in particular that it is dimensionally correct.

We have therefore succeeded in expressing the pressure with the help of the two-particle correlation function. The second term in (30) is added to the ideal gas result.

**XI.6. Virial and virial theorem**

Done in class. [See Diu, p. 704 sqq. or C. Kittel, *Elementary statistical physics*, p. 222.]

**XI.7. Application: pair correlation in the one-dimensional Ising model**
II. SYSTEMS OF CHARGED PARTICLES

A plasma is an ionized gas. An electrolyte is a system of charged particles dissolved in a liquid. The theoretical study of systems with charged particles requires special methods because of the long range of the Coulomb interaction. Many of the methods developed for systems with short range interactions do not apply any longer. As a general rule, the mathematical problems stem from the fact that the space integrals encountered do not converge, either at infinity or in the origin, or both.

II.1. The problem

We consider $N = N_{\text{ion}} + N_{\text{el}}$ particles, ions and electrons, in a volume $V$. We let $Z$ stand for the atomic number. Out of the $N$ particles $N_{\text{ion}}$ have an electric charge $q_i = Ze$ and a mass $m_i = M$ (where $i = 1, 2, \ldots, N_{\text{ion}}$), and $N_{\text{el}}$ have a charge $q_i = -e$ and a mass $m_i = m$ (where $i = N_{\text{ion}} + 1, \ldots, N_{\text{ion}} + N_{\text{el}}$). We let $N$ be even and impose global neutrality, $\sum_{i=1}^{N} q_i = 0$. This leads to $N_{\text{ion}} = (Z + 1)^{-1}N$ and $N_{\text{el}} = Z(Z + 1)^{-1}N$. The Hamiltonian of this system is

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j, j \neq i}^{N} \frac{q_i q_j}{4\pi \epsilon |r_i - r_j|}, \quad (1)$$

Here $U_i(r_i)$ is the potential at the position of the $i$th particle due to all other particles. The canonical partition function is

$$Z = \frac{1}{N_{\text{el}}!N_{\text{ion}}!} \left( \frac{2\pi m}{\beta \hbar^2} \right)^{3/2} N_{\text{el}}^2 \left( \frac{2\pi M}{\beta \hbar^2} \right)^{3/2} N_{\text{ion}}^2 \frac{Z_{\text{conf}}}{Z}, \quad (2)$$

$$Z_{\text{conf}} = \int \cdots \int e^{-\frac{1}{\beta} \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} \frac{q_i q_j}{4\pi \epsilon |r_i - r_j|}}, \quad (3)$$

where $Z_{\text{conf}}$ is the configurational partition function. There is no known way of calculating $Z_{\text{conf}}$ exactly and the question that arises is: how to find the best possible approximation? A physical idea is the following. A positive ion will attract round itself negative electrons rather than other positive ions. We may therefore imagine that around a given ion there will form a “cloud” of negative charges that “shield” its potential. As a consequence, the ion will be the center of a potential that decays at large distances not as $\sim 1/r$ but faster\footnote{The idea of an effective potential also occurs in van der Waal’s theory of phase transitions (not discussed in these lecture notes); it is a fundamental concept in statistical physics.}.

II.2. Debye-Hückel theory

II.2.1 Equations.
The first particle \((i = 1)\) is an ion of charge \(q_1 = Ze\). We choose its position as the origin of our coordinate system and ask about the spatial distribution of the other charges. [*** We might alternatively have fixed a “test” charge in the origin.]

XII.2.2 Solution.
XIII. CHEMICAL EQUILIBRIUM

This chapter is an application of the canonical and the grand-canonical ensemble.

XIII.1. Introduction

Suppose a volume $V$ contains $N_A$, $N_B$, and $N_C$ particles of three different gases A, B, and C, respectively. We will denote the free energy of this system, when it is at equilibrium at temperature $T$, by $F(T, V, N_A, N_B, N_C)$. It may be calculated, in principle, from the system Hamiltonian $\mathcal{H}(N_A, N_B, N_C)$ via the canonical partition function $Z_c$ according to

$$Z_c = e^{-\beta F} = \frac{1}{h^{3N_A}N_A!} \frac{1}{h^{3N_B}N_B!} \frac{1}{h^{3N_C}N_C!} \int d\Gamma^A \int d\Gamma^B \int d\Gamma^C e^{-\beta \mathcal{H}(N_A, N_B, N_C)}. \quad (1)$$

Each species has its own chemical potential,

$$\mu_A = \frac{\partial F}{\partial N_A}, \quad \mu_B = \frac{\partial F}{\partial N_B}, \quad \mu_C = \frac{\partial F}{\partial N_C}. \quad (2)$$

A variation $(dF_A, dF_B, dF_C)$ of the three particle numbers will change the free energy by

$$dF = \mu_A dN_A + \mu_B dN_B + \mu_C dN_C. \quad (3)$$

XIII.2. Chemical reaction

XIII.2.1. Condition for chemical equilibrium

Suppose now that we “turn on” the possibility of the chemical reaction

$$A + B \rightleftharpoons C. \quad (4)$$

Examples are $e^- + H^+ \rightleftharpoons H$ and $2H \rightleftharpoons H_2$. That is, we have released a constraint. You know from chapter III (the Canonical Ensemble) that in such a case the general rule is that the system will tend to a new equilibrium in which the free energy takes its minimum value compatible with the remaining constraints.

In the present case this means that $N_A$, $N_B$, and $N_C$ will take values $N_A^*$, $N_B^*$, and $N_C^*$ such that $dF = 0$ for all allowed variations $(dF_A, dF_B, dF_C)$. Why “allowed”? That is because the reaction implies that necessarily $dN_C = -dN_A = -dN_B$. Using this in $(3)$ we see that the resulting free energy change is given by

$$dF = (-\mu_A - \mu_B + \mu_C) dN_C. \quad (5)$$

It follows that in the presence of the reaction chemical equilibrium occurs for

$$\mu_A + \mu_B = \mu_C. \quad (6)$$

This relation is easily generalized to more complicated reactions (formula given in class).
XIII.2.2. Finding $N_A^*, N_B^*$, and $N_C^*$

To find the particle numbers at chemical equilibrium a slightly more elaborate version of the above argument is needed. Done in class.

XIII.3. A special case: three ideal gases

Let A, B, and C be three ideal gases. Let C be a bound state of an A and a B whose binding energy is $-u_0$ (we will imagine $u_0 > 0$). We will begin by considering this system in the absence of any reaction, that is, with fixed $N_A$, $N_B$, and $N_C$.

The system Hamiltonian is, in self-evident notation,

$$\mathcal{H}(N_A, N_B, N_C) = \sum_{i=1}^{N_A} \frac{p_i^2}{2m_A} + \sum_{j=1}^{N_B} \frac{p_j^2}{2m_B} + \sum_{k=1}^{N_C} \frac{p_k^2}{2m_C} - N_C u_0. \quad (7)$$

You may easily calculate the canonical partition function using Eq. (1), and then the grand-canonical partition function $Z_g$. The result is that

$$Z_g(T, V, \mu_A, \mu_B, \mu_C) = \sum_{N_A=0}^{\infty} \sum_{N_B=0}^{\infty} \sum_{N_C=0}^{\infty} e^{\beta \mu_A N_A + \beta \mu_B N_B + \beta \mu_C N_C} Z_c(T, V, N_A, N_B, N_C)$$

$$= \exp \left[ \frac{\beta \mu_A V}{\lambda_A^3} + \frac{\beta \mu_B V}{\lambda_B^3} + \frac{\beta (\mu_C + u_0) V}{\lambda_C^3} \right]$$

$$= \exp \left[ \frac{(\zeta_A + \zeta_B + \zeta_C e^{\beta u_0}) V}{\lambda_C^3} \right], \quad (8)$$

in which the last line defines the $\zeta_\alpha$ and where furthermore $\lambda_\alpha = h/\sqrt{2\pi m_\alpha k_B T}$, for $\alpha = A, B, C$. For the densities of the three particle species we have

$$n_A = \frac{\langle N_A \rangle}{V} = \frac{1}{V} \frac{\partial \log Z_g}{\partial \mu_A} = \ldots \zeta_A \quad (9)$$

and similarly

$$n_B = \zeta_B, \quad n_C = e^{\beta u_0} \zeta_C. \quad (10)$$

Now take the reaction into account: this means imposing Eq. (6), $\mu_A + \mu_B = \mu_C$. It then follows from Eqs. (9)-(10) that

$$\frac{\zeta_A \zeta_B}{\zeta_C} = \frac{\lambda_A^3}{\lambda_A^3 \lambda_B^3} = \left( \frac{2\pi k_B T m_A m_B}{h^2 m_C} \right)^{3/2}, \quad (11)$$

which may be rewritten as

$$\frac{n_A n_B}{n_C} = e^{-\beta u_0} \left( \frac{2\pi k_B T m_A m_B}{h^2 m_C} \right)^{3/2}. \quad (12)$$
This is called the Saha equation or alternatively the law of mass action. It tells us where the reaction equilibrium is. It is remarkable that the ratio in Eq. (12) depends only on the particle masses, the temperature, and the binding energy $u_0$.

**Exercise** (done in class). In each of the following cases, determine with the aid of Saha’s equation in which direction the equilibrium shifts:

(a) when the temperature is raised;
(b) when the binding energy $u_0$ is increased;
(c) when the pressure is increased (i.e., the volume decreased).

Could you comment on the case $u_0 < 0$?

### XIII.4. Ionization equilibrium of atomic hydrogen

As an application of the preceding section we consider a gas of atomic hydrogen at high temperature, such as occurs, for example, in stellar atmospheres. Notation:

$$\text{H}^+ + e^- \rightleftharpoons \text{H}$$

masses $m_+$ $m_-$ $m_0$

densities $n_+$ $n_-$ $n_0$ (13)

Since $m_+ \approx m_0$ and $n_+ = n_-$ the Saha equation becomes (check this)

$$\frac{n_+^2}{n_0} = \frac{e^{-\beta u_0}}{\lambda_-^3},$$

where $\lambda_- = h/\sqrt{2\pi m_- k_BT}$. That is, we have one equation for two unknowns. However, there is one more equation. Let us denote by $n$ the total number of H atoms that we would have in the absence of ionization; equivalently, $n = \text{total proton density} = \text{total electron density}$. We may consider $n$ as a given variable. Let furthermore

$$x = \frac{n_+}{n}$$

(15)

stand for the fraction of ionized atoms. Then $n = n_0 + n_+$ whence $n_0 = n(1 - x)$. Using this in the Saha equation (14) we find

$$\frac{x^2}{1 - x} = \frac{e^{-\beta u_0}}{n\lambda_-^3}$$

$$= \frac{1}{n} \left( \frac{h^2}{2\pi m_- k_BT} \right)^{3/2} e^{-u_0/k_BT}.$$ (16)

This quadratic equation for the ionized fraction $x(n, T)$ at given density and temperature is easily solved. The temperature dependence is dominated by the exponential. A value typical to a stellar atmosphere is $n = 10^{20} \text{ m}^{-3}$. With this value one finds from Eq. (16) that the ionized fraction goes up from close to zero below 3000 K to $\frac{1}{2}$ around 10 000 K and almost 1 above 15 000 K.
**Exercise.** Let an atomic hydrogen gas be in equilibrium at temperature $T$. The ratio of the number $N_2$ of atoms in the excited state of quantum number $n = 2$ and energy $-E_I/4$ to the number $N_1$ in the ground state (quantum number $n = 1$, energy $-E_I = -13.6\text{ eV}$) is (why?)

$$\frac{N_2}{N_1} = 4 \exp \left( -\frac{3}{4} \beta E_I \right). \quad (17)$$

Calculate this fraction numerically for $T = 10,000\text{ K}$. Explain why it is in some way “easier” to ionize an H atom than to raise it to an excited state.

**XIII.5. Reactants with internal states**

Discussed in class.

**XIII.6. Reaction kinetics**

Consider again reaction (4),

$$A + B \xrightarrow{k_1} \xleftarrow{k_2} C. \quad (18)$$

Here the *dissociation rate* $k_2$ is the probability per unit of time for a C to dissociate; and the *recombination rate* $k_1$ is the probability per unit of time for an A to recombine with a B, per unit of number density of the B’s.

Let $n_A$, $n_B$, and $n_C$ be the time dependent number densities (or “concentrations”) of the three species. In your chemistry course you have learned to associate with a reaction of this type the time evolution equations (often called *chemical rate equations*)

\[
\begin{align*}
\dot{n}_C & = k_1 n_A n_B - k_2 n_C , \\
\dot{n}_A & = -k_1 n_A n_B + k_2 n_C , \\
\dot{n}_B & = -k_1 n_A n_B + k_2 n_C .
\end{align*}
\] (19)

Note that (19) implies mass conservation, as expressed by $(d/dt)(n_A + n_B + 2n_C) = 0$. The rate equations say that the system is in equilibrium when the right hand members vanish, that is, when

$$\frac{n_A n_B}{n_C} = \frac{k_2}{k_1}. \quad (20)$$

Upon comparing (20) to the Saha equation (12) we see that the reaction rates $k_1$ and $k_2$ cannot be arbitrary but must have a ratio fixed by fundamental constants (the particle masses, the temperature, and the binding energy).
Appendix A. PROBABILITY DISTRIBUTIONS

A.1. Random variables

Example 1. When a die is thrown, the result may be 1 or 2 or \ldots or 6 pips, but the outcome cannot be predicted: it is a random variable. Let’s denote by $P(n)$ the probability that the result be $n$. For a fair die we have $P(n) = 1/6$ for all $n = 1, 2, \ldots, 6$. If the die is biased, the $P(n)$ may be different from $1/6$, but in all cases we should have

\begin{align}
0 \leq P(n) & \leq 1 \quad \text{for } n = 1, 2, \ldots, 6, \quad (1) \\
P(1) + P(2) + \ldots + P(6) &= 1. \quad (2)
\end{align}

This example leads us to the following definition.

Definition. A discrete random variable $n$ is specified by

(i) the set $X$ of all values it may take;
(ii) a function $P(n)$ such that

\begin{align}
0 \leq P(n) & \leq 1 \quad \text{for all } n \in X, \quad (3) \\
\sum_{n \in X} P(n) &= 1 \quad \text{(normalization).} \quad (4)
\end{align}

We call $P(n)$ a probability law or probability distribution. Eq. (4) is the normalization condition of this law: it expresses that we are sure that $n$ will take one of the values in $X$.

Remark. In many applications $X$ will be a subset of the integers, as in this example. However, $X$ may be any set, e.g., $X = \{\text{red, yellow, green}\}$.

Example 2. A needle is dropped onto a two-dimensional surface equipped with a coordinate system. What is the probability that the needle make an angle $\phi$ of, say, $\pi/4$ with the $x$ axis? Obviously, the probability that the angle be exactly $\pi/4 = 0.785398163\ldots$ is zero, and we should rephrase the question as follows: what is the probability that $\phi$ be in the interval $[\pi/4 - \epsilon, \pi/4 + \epsilon]$, for some small given $\epsilon$?

More generally, given a prescribed angle $\phi_0$, let $P(\phi_0)\Delta\phi_0$ be the probability that the angle be in the interval $[\phi_0, \phi_0 + \Delta\phi_0]$. Instead of (3) and (4) we now have

\begin{align}
P(\phi) & \geq 0 \quad \text{for } 0 \leq \phi \leq \pi, \quad (5) \\
\int_0^\pi d\phi P(\phi) &= 1. \quad (6)
\end{align}

Note that whereas $P(\phi_0)\Delta\phi$ is a probability (a number in $[0, 1]$), the function $P(\phi)$ is a probability density (a probability per unit of angle). This example leads us to the following definition.

Definition. A continuous random variable $x$ is specified by
(i) the set $X$ of all values it may take;  
(ii) a function $P(x)$ such that 
\[ 0 \leq P(x) \quad \text{for all } x \in X, \]  
\[ \int_X dx P(x) = 1 \quad \text{(normalization)}. \]  
Again, we refer to $P(x)$ as a probability law or probability distribution. It is actually a probability density. In the case of a continuous variable a frequent misuse of language is an expression like “the probability $P(x)$ that the variable be equal to $x$”. We will tolerate this as long as the speaker is aware of what he actually means.

**Remarks.**
1. A random variable is also called a stochastic variable.
2. An outcome ($n$ or $x$ in the above examples) is also called an event.
3. Mathematicians often use a separate symbol (for example $\underline{n}$ or $\underline{x}$) to distinguish the random variable itself from its possible outcomes. This allows such statements as “the probability that $\underline{n}$ be equal to $n$”. However, we will not complicate our notation; it will always be clear what is meant.
4. The discrete and continuous cases discussed above are actually not very different from one another. In our general discussion below we will denote the random variable by $x$ without specifying whether it is discrete or continuous.
5. Obviously, one may conceive of random variables whose range $X$ is partly discrete and partly continuous. Can you think of an example from physics?

**A.2. Random variables in physics**

A random variable as defined above may be considered as a purely mathematical object. In that case one may choose the probability distributions, $P(n)$ or $P(x)$, without any constraint. However, in statistical mechanics we will be going to consider certain physical observables as random variables, for example

- the number of atoms in a (mentally delineated) subvolume of 1 cm$^3$ inside a container filled with an atomic gas;
- the speed of a Brownian particle suspended in a gas;

in both cases the variable changes rapidly and unpredictably in time and by performing repeated measurement we may determine its probability distribution. One therefore introduces for such a variable a probability law $P(x)$ supposed to summarize our knowledge about it. Evidently, such physical probability laws cannot be arbitrary: they must be compatible with the fundamental laws of physics. The probability laws that we will encounter in this course are all derived from a few basic postulates combined with the laws of physics. We will consider these postulates as correct when the results of our statistical theory are confirmed by the experimental measurements.

In many applications $X$ will be equal to the real axis $\mathbb{R}$ or a subset thereof. However, we will often also consider multicomponent variables $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ taking their
values in the \( m \)-dimensional space \( \mathbb{R}^m \).

**Example 3.** Suppose two particles with positions \( r_1 \) and \( r_2 \) are contained in an external potential \( V_{\text{ext}}(r) \) and interact through \( V(|r_1 - r_2|) \). Their total potential energy \( E_{\text{pot}} \) is then \( E_{\text{pot}} = V(|r_1 - r_2|) + V_{\text{ext}}(r_1) + V_{\text{ext}}(r_2) \). Statistical mechanics asserts that in thermal equilibrium at a temperature \( T \), the probability density \( P(r_1, r_2) \) for finding the two particles at positions \( r_1 \) and \( r_2 \), respectively, is proportional to \( \exp(-E_{\text{pot}}/k_B T) \), that is, to the Boltzmann factor\(^{41} \).

\[
P(r_1, r_2) = C \exp[-(V(|r_1 - r_2|) + V_{\text{ext}}(r_1) + V_{\text{ext}}(r_2))/k_B T],
\]

where \( C \) is the appropriate normalization constant. This \( P \) is a probability distribution in the six-dimensional space \( \mathbb{R}^6 \).

**Example 4.** Statistical mechanics asserts that the probability distribution \( P(v) = P(v_1, v_2, v_3) \) of the velocity \( v = (v_1, v_2, v_3) \) of a particle of mass \( m \) in a gas at equilibrium at a temperature \( T \) is given by

\[
P(v) = \left( \frac{m}{2\pi k_B T} \right)^{3/2} \exp\left(-\frac{mv^2}{2k_B T}\right),
\]

which is called the **Maxwell distribution** and is a probability distribution in \( \mathbb{R}^3 \). It is proportional to the Boltzmann factor of the kinetic energy \( E_{\text{kin}} = \frac{1}{2}mv^2 \).

### A.3. Examples of probability distributions

**A.3.1. Distributions of a continuous variable \( x \)**

\( a. \) The **exponential law**

\[
P(x) = ae^{-ax}, \quad x \geq 0.
\]

\( b. \) The **Lorentz or Cauchy law**

\[
P(x) = \frac{a}{\pi(a^2 + x^2)}, \quad -\infty < x < \infty.
\]

\( c. \) The **uniform distribution** on an interval

\[
P(x) = \frac{1}{b-a}, \quad a \leq x \leq b.
\]

\( d. \) The **Gaussian or normal law**

\[
P(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \quad -\infty < x < \infty.
\]

### A.3.2. Distributions of a discrete variable \( n \)

\(^{41}\text{Here } k_B \text{ is Boltzmann’s constant (more about it later).} \)
e. Poisson’s law

\[ P(n) = e^{-a} \frac{a^n}{n!}, \quad n = 0, 1, 2, \ldots \]  

(15)

f. The binomial law

\[ P(n) = \binom{n_{\text{max}}}{n} p^n (1 - p)^{n_{\text{max}} - n}, \quad n = 0, 1, 2, \ldots, n_{\text{max}}. \]  

(16)

Exercise. Verify for each case that the distribution is correctly normalized.

Exercise. How does Poisson’s law arise? Let \( N \) point-like dots be placed randomly on a line segment of length \( L \), independently and with a uniform distribution. We take the limit \( N, L \to \infty \) with \( N/L = \rho \) fixed. What is the probability \( p_n \) that a subsegment of length \( y \) contain exactly \( n \) dots?

A.4. Averages and moments

Given a probability distribution \( P(x) \), one calls

\[ \langle x \rangle = \int dx x P(x) \]  

(17)

the average or expectation value of \( x \). If \( g(x) \) is a function of \( x \), we call

\[ \langle g(x) \rangle = \int dx g(x) P(x) \]  

(18)

the average of \( g(x) \). The special choice \( g(x) = x^m \) leads to

\[ \langle x^m \rangle = \int dx x^m P(x), \]  

(19)

called the \( m \)th moment of \( x \).

The quantity \( x - \langle x \rangle \) is the deviation of a particular value \( x \) of the random variable from its expectation value \( \langle x \rangle \). It is frequently written as

\[ \Delta x = x - \langle x \rangle. \]  

(20)

Evidently \( \langle \Delta x \rangle = 0 \). However, if we first square and then average we get

\[ \sigma_x^2 \equiv \langle \Delta x^2 \rangle = \langle (x - \langle x \rangle)^2 \rangle, \]  

(21)

called the mean square deviation, or the variance, or in certain contexts the dispersion,\(^{42}\) of \( x \). It is necessarily nonnegative: \( \sigma_x^2 \geq 0 \). The quantity \( \sigma_x \) itself is called the root-mean-square (rms) or standard deviation of \( x \).

\(^{42}\)In your Quantum Mechanics course we used \( \delta x \) and \( \Delta x \) in lieu and stead of our present \( \Delta x \) and \( \sigma_x \), respectively.
Exercise. Show that $\sigma^2_x = \langle x^2 \rangle - \langle x \rangle^2$.

Exercise. Calculate $\langle x \rangle$ and $\sigma_x$ for the probability laws in section A.3. Which new phenomenon appears in case b? Adapt the definitions of $\langle x \rangle$ and $\sigma_x$ to the discrete case for the examples e and f.

A.5. Multivariate probability laws

By a multivariate law we mean one that depends on two or more variables, such as in Examples 3 and 4 above. We return to Example 4 above. The Maxwell distribution (10) may be rewritten as

$$P(v) = p(v_1)p(v_2)p(v_3),$$

(22)

where

$$p(v_i) = \left(\frac{m}{2\pi k_B T}\right)^{1/2} e^{-\frac{mv_i^2}{2k_B T}}, \quad i = 1, 2, 3.$$  

(23)

The important point is that here $P$ factorizes into three separate probability distributions (that in this case happen to be all equal) for $v_1$, $v_2$, and $v_3$. We say that $v_1$, $v_2$, and $v_3$ are independent random variables. This leads us to the following more general definition.

Definition. Whenever a probability distribution factorizes into two or more distributions implying disjoint subsets of its variables, we say that these subsets are independent.

In the general case, whether or not its variables are independent, when a multivariate distribution $P$ of $r$ variables is integrated on one of them, one obtains the probability distribution – let us call it $P'$ – of the remaining $r - 1$ variables:

$$P'(x_1, \ldots, x_{r-1}) = \int dx_r P(x_1, \ldots, x_{r-1}, x_r).$$

(24)

Here $P'$ is called a marginal distribution of $P$. By repeated integrations one obtains marginal distributions of still fewer variables.

For a multivariate distribution one defines the correlation between two variables $x_i$ and $x_j$ as

$$C_{ij} \equiv \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle = \langle \Delta x_i \Delta x_j \rangle.$$

(25)

For $i = j$ we have that $C_{ii} = \sigma_{x_i}$; for $i \neq j$ we call $C_{ij}$ a covariance. The symmetric matrix $(C_{ij})$ is called the covariance matrix. In the frequently occurring case where for all $i$ we have $\langle x_i \rangle = 0$, the (co)variances simplify to $C_{ij} = \langle x_i x_j \rangle$.

Two variables $x_i$ and $x_j$ for which $C_{ij} = 0$ are said to be uncorrelated. Independent variables are uncorrelated; inversely, however, uncorrelated variables are not necessarily independent. Random variables that are independent and drawn from the same probability law are said to be independent and identically distributed, or “i.i.d.”.

Remark. In order to calculate an average involving only the subset of variables $\{x_1, \ldots, x_s\}$, but not its complement $\{x_{s+1}, \ldots, x_r\}$, it suffices to know the marginal law restricted to that subset.
Exercise. Illustrate this remark explicitly by calculating $\langle v_3 \rangle$ with respect to the Maxwell distribution.

Exercise. Consider the one-dimensional analog of Example 3. Calculate the averages and the covariance matrix of the probability distribution $P(x_1, x_2)$. Are the signs (positive/negative) of your answers compatible with what should be expected?

A.6. Generating function

A.6.1. Moments

Let $k$ be any real parameter. The particular average

$$G(k) \equiv \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} dx \, e^{ikx} P(x)$$

is the Fourier transform of $P(x)$ normalized such that $G(0) = 1$. It is also the moment generating function associated with $P(x)$. This becomes clear when we Taylor expand $G(k)$ in $k$,

$$G(k) = \sum_{\ell=0}^{\infty} \frac{(ik)^\ell}{\ell!} \langle x^\ell \rangle.$$  \tag{27}$$

In certain cases the easiest way to calculate the moments $\langle x^\ell \rangle$ may be to first calculate $G(k)$ and then expand it in powers of $k$. The full distribution $P(x)$ may be obtained from $G(k)$ by an inverse Fourier transformation.

Note: Equation (26) acquires a simpler look if you set $e^{ik} = s$ and consider $G$ as a function of $s$ instead of $k$; you may also just set $ik = p$ and consider $G$ as a function of $p$.

If instead of $x$ we have a discrete variable $n$, the generating function (26) obviously becomes (now in terms of the variable $s$)

$$G(s) \equiv \langle s^n \rangle = \sum_n s^n p_n.$$  \tag{28}$$

The inverse of this relation is

$$p_n = \frac{1}{2\pi i} \oint \frac{ds}{s^{n+1}} G(s),$$

where the integration is counterclockwise around the origin of the complex $s$ plane and the residue theorem is used.

A.6.2. Cumulants

We may ask what series is generated not by $G(k)$ but by $\log G(k)$. This yields, instead of (27), the series

$$\log G(k) = \sum_{\ell=1}^{\infty} \frac{(ik)^\ell}{\ell!} \kappa_{\ell}.$$  \tag{30}$$
whose coefficients \( \kappa_\ell \) are called the *cumulants* of the distribution \( P(x) \). The cumulants are nonlinear combinations of the moments. One easily finds

\[
\begin{align*}
\kappa_1 &= \mu_1, \\
\kappa_2 &= \mu_2 - \mu_1^2, \\
\kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3, \\
\kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4.
\end{align*}
\]

(31)

The general formula is not simple. Note that the second cumulant is also the variance of the distribution. Cumulants appear in all sorts of expansions encountered in theoretical physics.

**Exercise.** Find the generating function of the Gaussian distribution (14). Show that all its cumulants beyond the second one vanish, that is, \( \kappa_3 = \kappa_4 = \ldots = 0 \).

**A.7. Transformation of variables**

Let \( P(x) \) be a given probability distribution and let \( y = f(x) \).

Question: What is the probability distribution \( \tilde{P}(y) \) of \( y \)?

Answer: If \( y = f(x) \), the “probability content” of the interval \([x, x + dx]\) must be equal to that of \([y, y + dy]\) that is,

\[
P(x)dx = \tilde{P}(y)dy
\]

(32)

where \( dy = |df(x)/dx|dx \). It follows that

\[
\tilde{P}(y) = P(x) \left| \frac{dx}{dy} \right|.
\]

(33)

For multivariate distributions the absolute value of the Jacobian appears.

**Exercise.** Derive from (10) the marginal distribution \( P_1(v) \) for the velocity \( v \equiv |v| \) and then the distribution \( P_2(E) \) of the kinetic energy \( E = \frac{1}{2}mv^2 \).

**A.8. Central limit theorem**

Let \( x_1, x_2, \ldots, x_r \) be independent variables drawn from distributions \( P_{x_1}(x_1), P_{x_2}(x_2), \ldots, P_{x_r}(x_r) \) whose corresponding generating functions are \( G_{x_1}(k_1), G_{x_2}(k_2), \ldots, G_{x_r}(k_r) \). We will take these variables to have zero average.

We define the sum \( y = x_1 + x_2 + \ldots + x_r \) and are interested in its probability distribution \( P_y(y) \); we will call the corresponding generating function \( G_y(y) \).

By Fourier transforming

\[
P_y(y) = \int dx_1 \ldots \int dx_r P_{x_1}(x_1) \ldots P_{x_r}(x_r) \delta(y - x_1 - x_2 - \ldots - x_r)
\]

(34)

we find

\[
G_y(k) = G_{x_1}(k)G_{x_2}(k) \ldots G_{x_r}(k).
\]

(35)
Suppose now all $x_i$ identically distributed, denote their law by $P_x(x)$, and the corresponding generating function by $G_x(k)$. Then (35) becomes
\[ G_y(k) = [G_x(k)]^r. \] (36)

1. $P_x(x)$ a Gaussian. For the special case of a Gaussian $P_x(x)$ we have the correspondence
\[ P_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad \iff \quad G_x(k) = e^{-\frac{1}{2}\sigma^2 k^2}. \] (37)

In this case we see by combining (36) and (37) that $G_y(k) = \exp(-\frac{1}{2}r\sigma^2 k^2)$ and hence $P_y(y)$ is a Gaussian of variance $r\sigma^2$.

This suggest to define a differently scaled sum variable,
\[ z = \frac{x_1 + x_2 + \ldots + x_r}{\sqrt{r}}. \] (38)

By relating its generating function to that of $y$ we obtain
\[ G_z(k) = \langle e^{ikz} \rangle = \langle e^{\frac{i}{\sqrt{r}} y} \rangle = G_y \left( \frac{k}{\sqrt{r}} \right) = \left[ G_x \left( \frac{k}{\sqrt{r}} \right) \right]^r = e^{-\frac{1}{2}\sigma^2 k^2}, \] (39)

whence it follows that the distribution $P_z(z)$ is a Gaussian of variance $\sigma^2$.

2. $P_x(x)$ arbitrary. Now things become interesting. For an arbitrary distribution $P_x(x)$ we have that
\[ G_x(k) = 1 - \frac{1}{2}\sigma^2 k^2 + \ldots, \] (40)

where the dots stand for higher order terms in $k$. Again we ask about the distribution $P_z(z)$ of the scaled sum defined by (38). Upon substituting (40) in (36) we find
\[ G_z(k) = \left[ 1 - \frac{\sigma k^2}{2r} + \ldots \right]^r. \] (41)

We observe that the dot terms, which are of higher order in $k$, are also of higher order in $1/\sqrt{r}$. Hence, upon taking the limit $r \to \infty$ we get from (41) the final result
\[ G_z(k) = e^{-\frac{1}{2}\sigma^2 k^2}, \quad r \to \infty. \] (42)

The correspondence (37) then leads to the Central Limit Theorem (CLT), that we may now formulate as follows.

Let i.i.d. random variables $x_1, x_2, \ldots, x_r$ be drawn from an arbitrary common probability distribution $P_x(x)$ that has a variance $\sigma^2$. Then the scaled sum $z = (x_1 + x_2 + \ldots + x_r)/\sqrt{r}$ is distributed, in the limit $r \to \infty$, according to the Gaussian
\[ P_z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2\sigma^2}}. \] (43)

Remarks.
(1) It is sometimes said that under addition of i.i.d. random variables the Gaussian distribution is an attractor in the space of probability laws. Mathematically this happens in Eq. (41), where the dot terms, that differentiate between different distributions $P_x(x)$, are seen to be negligible when $r \to \infty$.

(2) The CLT applies to a very wide class of laws $P_x(x)$. Nevertheless, excluded are those that do not satisfy $\sigma^2 < \infty$, that is, do not have a finite variance. For such laws there exist other attractors, called Lévy distributions.
Appendix B. METHOD OF STEEPEST DESCENT

B.1. Maximum inside an interval

Suppose you are given the integral

\[ I_N = \int_{-\infty}^{\infty} dx \ e^{Ng(x)} \]  

in which \( g(x) \) is an arbitrary (sufficiently differentiable) function having a maximum on the \( x \) axis. Of course you cannot calculate \( I_N \) explicitly if you know nothing else about \( g(x) \).

**Question:** In the absence of any further knowledge, what can you say about the behavior of \( I_N \) in the limit of large \( N \)?

**Answer:** Let the maximum of \( g(x) \) occur at \( x = x_0 \). Expand

\[ g(x) = g(x_0) + \frac{1}{2} g''(x_0)(x-x_0)^2 + \frac{1}{6} g'''(x_0)(x-x_0)^3 + \frac{1}{24} g^{(4)}(x_0)(x-x_0)^4 + \mathcal{O}((x-x_0)^5), \]  

substitute this expansion in (1), and change to the variable of integration \( \tilde{y} = x - x_0 \). This leads to

\[ I_N = e^{Ng(x_0)} \int_{-\infty}^{\infty} d\tilde{y} \ e^{N \left[ \frac{1}{2} g''(x_0)\tilde{y}^2 + \frac{1}{6} g'''(x_0)\tilde{y}^3 + \frac{1}{24} g^{(4)}(x_0)\tilde{y}^4 + \mathcal{O}(\tilde{y}^5) \right]} \].

(3)

Now scale \( \tilde{y} \) by a power of \( N \) such that the leading (that is, the quadratic) term in the exponential becomes \( N \)-independent. This fixes the choice \( y = N^{\frac{1}{2}} \tilde{y} \). Hence

\[ I_N = N^{-\frac{1}{2}} e^{Ng(x_0)} \int_{-\infty}^{\infty} dy \ e^{N \frac{1}{2} g''(x_0)y^2 + \frac{1}{6} N^{-\frac{1}{2}} g'''(x_0)y^3 + \frac{1}{24} N^{-1} g^{(4)}(x_0)y^4 + \mathcal{O}(N^{-\frac{3}{2}}y^5)} \].

(4)

Expand now all terms in the integrand that are accompanied by negative powers of \( N \). This gives

\[ I_N = N^{-\frac{1}{2}} e^{Ng(x_0)} \int_{-\infty}^{\infty} dy \ e^{\frac{1}{2} g''(x_0)y^2} \left[ 1 + \frac{1}{6} N^{-\frac{1}{2}} g'''(x_0)y^3 + \frac{1}{24} N^{-1} g^{(4)}(x_0)y^4 + \mathcal{O}(N^{-\frac{3}{2}}y^5) \right]. \]

(5)

Each term on the RHS now contains a Gaussian integral with a power of \( y \) inserted. Using that odd powers yield zero by symmetry and that

\[ \int_{-\infty}^{\infty} dy \ e^{-\frac{1}{2} \alpha y^2} = \frac{(2\pi)^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}}, \quad \int_{-\infty}^{\infty} dy \ y^4 e^{-\frac{1}{2} \alpha y^2} = \frac{3\pi^{\frac{3}{2}}}{\alpha^{\frac{3}{2}}}, \]

we obtain, up to the first correction term,

\[ I_N = N^{-\frac{1}{2}} e^{Ng(x_0)} \left[ 1 + \frac{3\sqrt{2} g^{(4)}(x_0)}{12 g''(x_0)} + \ldots \right]. \]

(7)
This series may be determined to any order desired in $N^{-1}$.

**Exercise.** Slightly generalize this to do the same integral with the insertion of an arbitrary (sufficiently differentiable) function $h(x)$, that is,

$$ I'_N = \int_{-\infty}^{\infty} dx \ h(x) e^{Ng(x)} \quad (8) $$

*Indication:* The first correction term receives an extra contribution proportional to $h'(x_0)g'''(x_0)$.

**Exercise.** Show that the series (7) does not change if the integration in (1) is only over an interval $(a,b)$ containing $x_0$, the corrections being of order $O(e^{-CN})$ for some $c > 0$.

**Remarks**

(1) The method exposed here is called the *method of steepest descent*. It has several alternative names: Laplace’s method, the *saddle point* method, or the *stationary phase* method. The last two names refer to the fact that when viewed in the complex plane the integrand has a saddle point.

(2) In statistical mechanics $N$ usually represents the system size. In many instances where integrals of the above type occur one is interested first of all in the terms in the exponential that are proportional to $N$. In that case we may write (7) as

$$ I_N = e^{Ng(x_0)+O(\log N)} \quad (9) $$

and obtaining this result requires nothing more than the value of $g(x_0)$.

(3) The range of $x$ values that effectively contribute to the integral is determined by the rms deviation $\langle (x - x_0)^2 \rangle^{\frac{1}{2}} = 1/\sqrt{Ng''(x_0)}$.

(4) Notice that when $h(x)$ is inserted, the integration acts on this insertion as an unnormalized Dirac delta at $x = x_0$. This may also be expressed by $I'_N/I_N = h(x_0) + \ldots$, where the dot terms vanish as $N \to \infty$.

**B.2. Boundary maximum**

We consider again the integral (1) but now suppose that $x$ is integrated over a finite interval $[a,b]$ and that $g(x)$ has a boundary maximum in $x = a$. The calculation then proceeds as follows. [*** To be completed.]
Appendix C. LAGRANGE MULTIPLIER METHOD

Suppose you wish to find the stationary points of a function \( f(x_1, x_2, \ldots, x_s) \) under the conditions

\[
g_j(x_1, x_2, \ldots, x_s) = 0, \quad (1)
\]

where \( j = 1, 2, \ldots, r \) and \( r \leq s \). To solve this problem, introduce parameters \( \lambda_1, \lambda_2, \ldots, \lambda_r \), called \((Lagrange)\) multipliers, and consider the more general problem of finding the stationary points of

\[
F(x_1, x_2, \ldots, x_s) - \sum_{j=1}^{r} \lambda_j g_j(x_1, x_2, \ldots, x_s). \quad (2)
\]

This leads to the \( s \) equations

\[
\frac{\partial F}{\partial x_i} = 0, \quad i = 1, 2, \ldots, s, \quad (3)
\]

which may have one or several solutions \((x_1^*(\lambda), x_2^*(\lambda), \ldots, x_s^*(\lambda))\), where \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r) \).

Now substitute this solution in (1). This leads to

\[
g_j(x_1^*(\lambda), x_2^*(\lambda), \ldots, x_s^*(\lambda)) = 0, \quad j = 1, 2, \ldots, r, \quad (4)
\]

which are now \( r \) equations for the unknown \( r \)-component vector \( \lambda \). Let the solution be \( \lambda^* \). Then \((x_1^*(\lambda^*), x_2^*(\lambda^*), \ldots, x_s^*(\lambda^*))\) is a stationary point of \( f(x_1, x_2, \ldots, x_s) \) that satisfies the conditions.