



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Department of Physics and Astronomy "A. Righi"

Theoretical Physics

Statistical Mechanics

Lectures held by
Prof. Elisa Ercolessi

Notes by
Gioele Mancino

Academic Year 2025/2026

Contents

1 Thermodynamics Review	3
1.1 Laws of Thermodynamics	5
1.2 Thermodynamic Potentials	7
2 Classical Mechanics Review	9
2.1 Introduction to Classical Systems	9
2.2 Statistical Description	11
2.3 Ergodicity	15
3 Classical Microcanonical Ensemble	17
3.1 Isolated System	18
3.2 Microcanonical Entropy	19
3.3 Perfect Gas	22
4 Classical Canonical Ensemble	25
4.1 System in a Thermal Bath	26
4.2 Partition Function and Thermodynamic Quantities	28
4.3 Generalized Equipartition Theorem	31
5 Classical Grandcanonical Ensemble	33
5.1 System in a Particle and Thermal Bath	34
5.2 Partition Function and Thermodynamic Quantities	38
5.3 Virial Expansion	40
6 Phase Transitions	45
6.1 Phase Diagrams	46
6.2 Classification of Phase Transitions	49
6.3 Discontinuities and Phase Transitions	50
6.4 Heisenberg Model	54

6.5	Ising Model	60
6.6	Critical Exponents	66
7	QM and Statistics	69
7.1	Density Matrix	72
7.2	Qubit	75
7.3	Identical Particles	77
7.4	Quantum Statistics	79
8	Second Quantization	83
8.1	Creation and Annihilation Operators	84
8.2	Fock Space	87
8.3	Field Operators	90
8.4	Observable Operators	94
9	Quantum Ensembles	101
9.1	Microcanonical Ensemble	102
9.2	Canonical Ensemble	104
9.3	Grand Canonical Ensemble	106
10	Quantum Gases	109
10.1	Quantum Statistics	111
10.2	Non Relativistic 3D Quantum Gases	114
10.3	Fermi-Dirac Gas	121
10.4	Bose-Einstein Gas	125
10.5	Fenomenology of BEC and Experimental Realizations	133
11	Exercises	141
11.1	Classical Statistical Mechanics	141
Appendices		155
A	Notation and Conventions	155
B	Computations and further developments	156

Introduction

Statistical Mechanics stands as one of the pillars of modern theoretical physics. It provides the formal bridge between the microscopic world—governed by the dynamical laws of Classical or Quantum Mechanics—and the macroscopic world described by Thermodynamics.

In this course, we move beyond the first phenomenological approach to thermodynamics. Our goal is to derive macroscopic observables (such as temperature, pressure, heat capacity, and magnetization) from the fundamental Hamiltonian describing a system of N interacting particles, where N is of the order of the Avogadro number ($N \sim 10^{23}$).

The Fundamental Problem

Consider a macroscopic sample of gas or a magnetic solid. A complete dynamical description would require solving the equations of motion for every constituent particle. In a classical framework, this means tracking $6N$ variables in phase space $\{q_i(t), p_i(t)\}$. This task is not only computationally impossible but, more importantly, physically redundant.

We do not care about the exact position of every atom in a gas cylinder; we care about the pressure exerted on the walls and the temperature of the gas. The central insight of Statistical Mechanics is that we can trade *exact dynamical knowledge* for *probabilistic information*. By assuming that the system explores its accessible phase space according to specific probability distributions, we can predict macroscopic behavior with overwhelming accuracy.

The Thermodynamic Limit

A recurring theme in these notes is the **Thermodynamic Limit**:

$$N \rightarrow \infty, \quad V \rightarrow \infty, \quad \frac{N}{V} = n = \text{const}$$

Strictly speaking, sharp thermodynamic behavior—such as phase transitions and non-analyticities in the free energy—only emerges in this limit. While finite-size effects are important in mesoscopic physics, the foundational theory relies on N being large enough that relative fluctuations, which typically scale as $1/\sqrt{N}$, become negligible.

Structure of the Notes

These notes are organized to guide you from classical foundations to advanced quantum field theoretical methods applied to statistical systems.

Part I: Classical Foundations (Chapters 1–5)

We begin by reviewing the axiomatic structure of **Thermodynamics** (Ch. 1) and **Classical Mechanics** (Ch. 2), specifically Liouville's theorem and the concept of phase space flow. We then construct the three fundamental ensembles:

- The **Microcanonical Ensemble** (Ch. 3) for isolated systems, defining entropy via the counting of microstates ($S = k_B \ln \Omega$).
- The **Canonical Ensemble** (Ch. 4) for systems in thermal equilibrium with a bath, introducing the partition function Z .
- The **Grand Canonical Ensemble** (Ch. 5) for systems exchanging both energy and particles, crucial for studying phase transitions and quantum gases.

Part II: Interactions and Phase Transitions (Chapter 6)

One of the most fascinating aspects of many-body physics is how simple microscopic interactions give rise to complex collective phenomena. We will study **Phase Transitions**, focusing on the breakdown of mean-field theory, the concept of Spontaneous Symmetry Breaking, and the universality of Critical Exponents. The Ising and Heisenberg models will serve as our primary playgrounds.

Part III: Quantum Statistical Mechanics (Chapters 7–10)

As we approach low temperatures, classical descriptions fail (e.g., the freezing out of degrees of freedom, the Gibbs paradox). We reformulate statistical mechanics using the **Density Matrix** formalism (Ch. 7).

A significant portion of this course is dedicated to **Second Quantization** (Ch. 8). We will move from wavefunctions to the formalism of creation and annihilation operators on Fock Space. This is the natural language for indistinguishable particles and forms the basis for modern condensed matter theory.

Finally, we apply these tools to **Quantum Gases** (Ch. 10). We will derive the Fermi-Dirac and Bose-Einstein distributions and study their low-temperature limits, culminating in the phenomenology of **Bose-Einstein Condensation (BEC)** and superfluidity—macroscopic quantum phenomena that have been experimentally realized in ultracold atomic gases.

Prerequisites: *A solid understanding of Thermodynamics, Hamiltonian Mechanics, and basic Quantum Mechanics is assumed.*

1 | Thermodynamics Review

We begin our exploration of statistical mechanics with a review of the fundamental concepts of thermodynamics. Thermodynamics is the study of energy, heat, work, and the macroscopic properties of systems in equilibrium. It provides the foundation for understanding how microscopic interactions give rise to macroscopic phenomena.

Thermodynamic Systems and States

A thermodynamic system is defined as a macroscopic region in which we can study the behavior of matter and energy. The boundaries of this system can be real or imaginary, and they separate the system from its surroundings. The surroundings are everything outside the system that can interact with it.

Thermodynamic systems can be classified into three main types:

- **Isolated systems:** These systems do not exchange matter or energy with their surroundings. An example is a thermos bottle that keeps its contents insulated from the external environment.
- **Closed systems:** These systems can exchange energy (in the form of heat or work) but not matter with their surroundings. A common example is a sealed container of gas that can expand or contract but does not allow gas to enter or leave.
- **Open systems:** These systems can exchange both energy and matter with their surroundings. An example is a boiling pot of water, where heat is transferred to the water from the stove, and water vapor escapes into the air.

The state of a thermodynamic system is described by its **macroscopic properties**, such as temperature, pressure, volume, and internal energy. These properties can change during thermodynamic processes, which are the transformations that occur within the system, but we will analyze only systems already at equilibrium.

We can divide thermodynamic variables into two categories of *conjugate*¹ quantities:

- **Intensive variables**, which do not depend on the size or extent of the system.
- **Extensive variables**, which depend on the size or extent of the system.

	Extensive		Intensive
E	Energy		—
S	Entropy	\longleftrightarrow	T Temperature
V	Volume	\longleftrightarrow	p Pressure
N	Particles	\longleftrightarrow	μ Chem. potential
\mathbf{P}	Polarization	\longleftrightarrow	\mathbf{E} Electric field
\mathbf{M}	Magnetization	\longleftrightarrow	\mathbf{B} Magnetic field

These variables are related by **state equations**, which express how the properties of a system are interdependent. For example, the ideal gas law relates pressure, volume, and temperature for an ideal gas:

$$pV = nRT,$$

where n is the number of moles of gas, R is the ideal gas constant.

¹Conjugate variables are pairs of variables that are related to each other in a specific way, often through a mathematical relationship, we will soon see examples of this.

1.1 | Laws of Thermodynamics

We now summarize the four fundamental laws of thermodynamics, which govern the behavior of thermodynamic systems.

1.1.1 | Zeroth Law of Thermodynamics

If two systems A and B are described by their set of thermodynamic variables \mathcal{M}_A and \mathcal{M}_B respectively, and they are both at equilibrium, we can use only a subset of $\mathcal{M}_A \times \mathcal{M}_B$: $a \times b$ ($a \in \mathcal{M}_A$ and $b \in \mathcal{M}_B$), which reduces the expression of the equilibrium to a function of these variables: $\mathcal{F}(a, b) = 0$.

Thanks to this law, we can introduce the concept of temperature as a property that determines whether two systems are in thermal equilibrium. To do so we know that:

- (i) $A \sim A$ always holds (reflexivity).
- (ii) If $A \sim B$ then $B \sim A$ (symmetry).
- (iii) If $A \sim B$ and $B \sim C$ then $A \sim C$ (transitivity).

Then we want $\mathcal{F}(a, b)$ to respect these properties, so we can write:

$$\mathcal{F}(a, b) = f(a) - f(b) = 0 \implies f(a) = f(b) = T.$$

1.1.2 | First Law of Thermodynamics

The First Law of Thermodynamics is a statement of the **conservation of energy**. It states that the change in internal energy of a system is given by the heat exchanged with the environment minus the work done by the system on its surroundings, plus the energy associated with the change in the number of particles (if applicable):

$$dE = \delta Q - \delta L + \mu dN. \quad (1.1.1)$$

This law is valid for any generic transformation, and it tells us how one can operate on the energy of a system. The terms δQ and δL are not exact differentials, meaning that they depend on the path taken during the transformation, while E and N are exact differentials, meaning that they depend only on the initial and final states of the system. In general:

$$\oint \delta Q \neq 0, \quad \oint \delta L \neq 0, \quad \oint dE = 0, \quad \oint dN = 0.$$

1.1.3 | Second Law of Thermodynamics

The Second Law of Thermodynamics states that natural processes tend to move towards a state of maximum **entropy**. This law introduces the concept of irreversibility in thermodynamic processes and implies that heat cannot spontaneously flow from a colder body to a hotter body. Mathematically, the Second Law can be expressed as:

$$dS \geq \frac{\delta Q}{T}, \quad (1.1.2)$$

where the equality holds just for reversible processes.

If we combine the First and Second Laws for reversible processes, we can derive an important relation known as the **Gibbs relation** for a system with a variable number of particles:

$$dE = TdS - pdV + \mu dN. \quad (1.1.3)$$

which will be fundamental to derive the thermodynamic potentials.

1.1.4 | Third Law of Thermodynamics

The Third Law of Thermodynamics states that as the temperature of a system approaches absolute zero (0 Kelvin), the entropy of a perfect crystal approaches a constant minimum value, which is typically taken to be zero. This law implies that it is impossible to reach absolute zero through any finite number of processes.

1.2 | Thermodynamic Potentials

Thermodynamic potentials are functions that describe the state of a thermodynamic system and are particularly useful for analyzing systems under different constraints.

If we take the Gibbs relation (1.1.3) it is clear that the natural variables of the internal energy E are S , V , and N . Thus $E = E(S, V, N)$ is a homogeneous function of degree 1 in its extensive variables, meaning that if we scale all extensive variables by a factor $\lambda > 0$, the internal energy scales by the same factor:

$$E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N).$$

The only functions that satisfy this property are linear functions, so we can write:

$$E = TS - pV + \mu N.$$

We obtained the expression of our first thermodynamic potential, the internal energy E . Note that it is a function of conjugate variables.

From this expression is easy to derive dependencies between the variables, for example:

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

We can define other thermodynamic potentials by performing **Legendre transforms** on the internal energy to change its natural variables. The most common thermodynamic potentials (defined here for reversible processes) are:

Internal Energy $E(S, V, N)$

$$E = TS - pV + \mu N \tag{1.2.1}$$

$$dE = T dS - p dV + \mu dN \tag{1.2.2}$$

Helmholtz Free Energy $F(T, V, N)$

$$F = E - TS = -pV + \mu N \tag{1.2.3}$$

$$dF = -S dT - p dV + \mu dN \tag{1.2.4}$$

Gibbs Free Energy $G(T, p, N)$

$$G = E - TS + pV = \mu N \tag{1.2.5}$$

$$dG = -S dT + V dp + \mu dN \tag{1.2.6}$$

Enthalpy $H(S, p, N)$

$$H = E + pV = TS + \mu N \tag{1.2.7}$$

$$dH = T dS + V dp + \mu dN \tag{1.2.8}$$

Grand Potential $\Omega(T, V, \mu)$

$$\Omega = E - TS - \mu N = -pV \tag{1.2.9}$$

$$d\Omega = -S dT - p dV - N d\mu \tag{1.2.10}$$

For general processes (not necessarily reversible) we can write the differentials of the thermodynamic potentials considering the inequalities from the Second Law: every equality becomes an inequality, ($= \rightarrow \leq$).

Thermodynamic limit. In the thermodynamic limit, where the number of particles N and the volume V of the system approach infinity while maintaining a constant density $n = N/V$, the thermodynamic potentials become extensive quantities; this means that they scale linearly with the size of the system. This property is crucial for ensuring that the thermodynamic description remains valid and consistent for large systems, allowing us to apply thermodynamic principles to real-world scenarios where systems contain a vast number of particles. For simplicity, we will often work with a fixed number of particles, and later generalize the results to the thermodynamic limit.

Variational principle. The thermodynamic potentials also satisfy a variational principle: for a system in equilibrium, the appropriate thermodynamic potential is minimized (or maximized) under the given constraints. Fixing a triplet of conjugate variables, the corresponding thermodynamic potential reaches an extremum at equilibrium. We can imagine the hessian matrix of the potential with respect to its natural variables, and we can see that it is positive definite (minimum) or negative definite (maximum).

$$\left[\frac{\partial^2 F}{\partial T^2} \right]_V < 0, \quad \left[\frac{\partial^2 F}{\partial V^2} \right]_T > 0, \quad \frac{\partial \mu}{\partial T} = \frac{\partial^2 F}{\partial T \partial N} \leq 0.$$

2 | Classical Mechanics Review

2.1 | Introduction to Classical Systems

In classical mechanics, the fundamental task is to describe the state and the evolution of a physical system composed of a certain number of particles. Each particle is uniquely characterized by its position and momentum, which together form a set of canonical conjugate variables. For N identical particles moving in a space of dimension d , the state of the i -th particle is represented by the pair:

$$(q_i, p_i) \in \mathcal{M},$$

where \mathcal{M} denotes the **phase space** of a single particle, a $2d$ -dimensional manifold that encompasses all possible positions and momenta. The total state of the system, or its *microscopic state*, is then given by the collection of all the particle states:

$$\{(q_i, p_i)_{i=1}^N\} \in \mathcal{M}^N.$$

An **observable** is any physical quantity that can be assigned a numerical value once the state of the system is specified. Mathematically, an observable is represented by a smooth real function defined on the phase space:

$$f : \mathcal{M}^N \rightarrow \mathbb{R}.$$

The **measurement** of such an observable in a particular state corresponds simply to evaluating the function at the point representing that state:

$$f(\bar{q}_i, \bar{p}_i).$$

Typical examples of observables include the total energy, the angular momentum, or the center-of-mass position of the system.

2.1.1 | Dynamics and Hamilton's Equations

The time evolution of the system is not arbitrary: it is determined by a distinguished observable, the *Hamiltonian function* $H(q_i, p_i; t)$, which typically represents the total energy of the system. According to Hamilton's formulation of mechanics, the equations of motion are given by the system of first-order differential equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

These equations describe how positions and momenta vary in time under the action of the Hamiltonian flow.

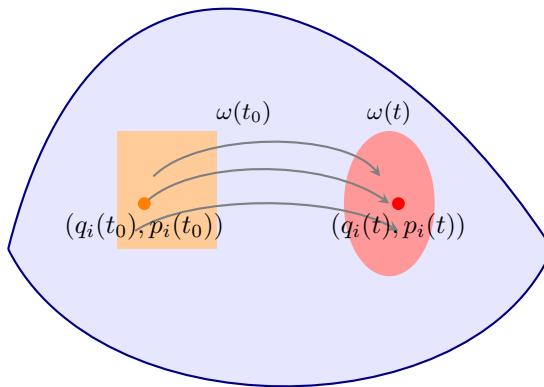
Conservation of energy. If the Hamiltonian does not depend explicitly on time, the system is said to be *autonomous*, and the total energy is conserved along the trajectories:

$$E \equiv H(q_i(t), p_i(t)) = H(q_i(0), p_i(0)).$$

This conservation law is a direct consequence of Hamilton's equations and reflects the time-translation symmetry of the system. In the autonomous case, the motion is confined to a *constant-energy hypersurface* in phase space, defined by

$$S_E = \{(q_i, p_i) \in \mathcal{M}^N \mid H(q_i, p_i) = E\}.$$

Liouville's theorem. A fundamental geometric property of Hamiltonian systems is that their flow in phase space preserves the natural volume element, a result known as *Liouville's theorem*. This means that although the system evolves and individual trajectories may become highly complex, the total volume in phase space occupied by a collection of states remains invariant over time. This property underpins many results in statistical mechanics, where ensembles of states are studied rather than single trajectories.



Initial condition $(q_i(t_0), p_i(t_0))$ evolves to $(q_i(t), p_i(t))$ under Hamiltonian flow. The volume of any region $\omega(t_0)$ in phase space is preserved over time, i.e., $\text{Vol}(\omega(t_0)) = \text{Vol}(\omega(t))$. Consequently, the density of states in phase space remains constant along the trajectories, which implies the conservation of the total number of states.

2.2 | Statistical Description

Hamilton's equations are deterministic: once the initial conditions $q_i(t_0), p_i(t_0)$ are specified, the system of first-order differential equations admits a unique solution $(q_i(t), p_i(t))$ for all times, provided the Hamiltonian is sufficiently smooth. Thus, the evolution in phase space is uniquely determined and reversible, as the flow defines a one-to-one mapping between initial and final states. Each such trajectory corresponds to a **microstate** of the system, fully characterizing its dynamical evolution.

However, in many practical situations the exact specification of the initial conditions is either impossible or irrelevant. What we can measure in experiments are not the detailed positions and momenta of all particles, but rather macroscopic quantities such as energy, volume, pressure, or particle number. Different microscopic configurations $(q_i(t), p_i(t))$ may correspond to the same values of these thermodynamic variables, and are therefore indistinguishable from a physical point of view. This motivates the introduction of a **macrostate**, which is defined not by a single trajectory in phase space, but by a whole set of microstates compatible with the observed macroscopic conditions.

To analyze such systems, statistical mechanics introduces the concept of an **ensemble**: a large collection of virtual copies of the system, each representing a possible microstate consistent with the given macrostate. The ensemble provides a probabilistic description of the system, allowing us to replace detailed knowledge of microscopic trajectories with statistical averages over all accessible states. In this framework, thermodynamic quantities emerge naturally as expectation values of the corresponding observables.

2.2.1 | Probability Distributions in Phase Space

For a large ensemble, the microscopic uncertainty about the state of the system is encoded in a probability density function

$$\rho(q_i, p_i; t)$$

defined on the phase space \mathcal{M}^N . This function satisfies the standard properties of a probability distribution, i.e. *positivity* and *normalization*:

$$\rho(q_i, p_i; t) \geq 0, \quad \int_{\mathcal{M}^N} \rho(q_i, p_i; t) \prod_i dq_i dp_i = 1,$$

so that the total probability of finding the system somewhere in phase space is normalized to unity.

The probability of finding the system in a particular subset $\mathcal{U} \subset \mathcal{M}^N$ is then obtained by integrating the density over that region:

$$\mathbb{P}[(q_i, p_i) \in \mathcal{U}] = \int_{\mathcal{U}} \rho(q_i, p_i; t) \prod_i dq_i dp_i.$$

Since the phase-space measure $\prod_i dq_i dp_i$ carries physical dimensions, it is often convenient to render it dimensionless. This is achieved by introducing the elementary phase-space volume element

$$d\Omega = \prod_i \frac{dq_i dp_i}{h} = \prod_i \frac{dq_i dp_i}{2\pi\hbar},$$

or, in a generic dimensional space,

$$d\Omega = \frac{1}{\xi_N} \prod_{i=1}^N \frac{d^d q_i d^d p_i}{h^d} = \frac{1}{\xi_N} \prod_{i=1}^N \frac{d^d q_i d^d p_i}{(2\pi\hbar)^d}, \quad (2.2.1)$$

where $h = 2\pi\hbar$ is a constant with dimensions of action (commonly chosen as Planck's constant) and ξ_N is a dimensionless factor accounting for particle indistinguishability:

$$\xi_N = \begin{cases} N! & \text{for identical particles,} \\ 1 & \text{for distinguishable particles.} \end{cases}$$

With this convention, probability distributions and state counting become properly normalized (the ξ_N factor is fundamental in this regard, as it corrects for the overcounting of indistinguishable particle configurations, naturally arising from quantum statistics in the phase space) and dimensionless, a step that is essential when passing to statistical mechanics and thermodynamic limits.

From now on we will use the dimensionless measure $d\Omega$ in all phase-space integrals, and denote the total phase-space as \mathcal{M} , letting the number of particles N and the space dimension d be implicit.

2.2.2 | Time Evolution and Stationarity

Theorem 2.1 (Liouville's theorem). *Let $\rho(q_i, p_i; t)$ be a probability density on the phase space \mathcal{M} and let $H(q_i, p_i; t)$ be a (sufficiently smooth) Hamiltonian generating the flow*

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

Then ρ is constant along the Hamiltonian trajectories, i.e.

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \{\rho, H\} = 0, \quad (2.2.2)$$

where $\{\cdot, \cdot\}$ denotes the canonical Poisson bracket.

The physical interpretation of Liouville's theorem is that the probability density is conserved along the flow generated by the Hamiltonian. In other words, as the system evolves in time according to Hamilton's equations, the value of ρ at the point representing the system's state remains unchanged. The Poisson bracket $\{\rho, H\}$ captures the infinitesimal change of ρ due to the Hamiltonian dynamics, and are defined as

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right). \quad (2.2.3)$$

Proof. We give two equivalent derivations.

(I) *Direct computation using the total (material) derivative.* The total derivative of ρ along a trajectory is

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_i \left(\dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i} \right).$$

Substituting Hamilton's equations yields

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_i \left(\frac{\partial H}{\partial p_i} \frac{\partial \rho}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial \rho}{\partial p_i} \right) = \frac{\partial \rho}{\partial t} + \{\rho, H\},$$

which is exactly (2.2.2). Hence $d\rho/dt = 0$ if $\partial_t \rho + \{\rho, H\} = 0$.

(II) *Continuity equation and divergence-free flow.* Consider the phase-space flow vector field $V = (\dot{q}_1, \dots, \dot{q}_N, \dot{p}_1, \dots, \dot{p}_N)$. The probability density satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla_{(q,p)} \cdot (\rho V) = 0.$$

For Hamiltonian vector fields one checks

$$\nabla_{(q,p)} \cdot V = \sum_i \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = \sum_i \left(\frac{\partial^2 H}{\partial p_i \partial q_i} - \frac{\partial^2 H}{\partial q_i \partial p_i} \right) = 0,$$

so the flow is divergence-free. Using $\nabla \cdot (\rho V) = V \cdot \nabla \rho + \rho \nabla \cdot V$ we obtain

$$\frac{\partial \rho}{\partial t} + V \cdot \nabla_{(q,p)} \rho = 0,$$

which is again the same as $\partial_t \rho + \{\rho, H\} = 0$. This shows that the density is conserved along trajectories.

Combining the two viewpoints yields the stated result. \square

Definition 2.1 (Stationary system). A system is said to be **stationary** if its probability density function does not depend explicitly on time, i.e.

$$\frac{\partial \rho}{\partial t} = 0.$$

This condition is necessary for thermodynamic equilibrium, since it ensures that the statistical state of the system remains invariant under time evolution. In this case, Liouville's equation reduces to

$$\{\rho, H\} = 0,$$

meaning that the stationary distribution must be a constant of motion.

The requirement $\{\rho, H\} = 0$ is satisfied in the following physically relevant situations:

- *Microcanonical ensemble:* the distribution is uniform within the constant-energy hypersurface,

$$\rho = \text{constant},$$

corresponding to an isolated system with fixed energy, volume, and particle number.

- *Canonical and grand canonical ensembles:* the distribution depends only on the Hamiltonian,

$$\rho = \rho(H),$$

for instance $\rho \propto e^{-\beta H}$ in the canonical ensemble, where $\beta = (k_B T)^{-1}$. These ensembles describe systems in contact with a reservoir fixing temperature (canonical) or both temperature and chemical potential (grand canonical).

2.2.3 | Time-Independent Hamiltonians and Density of States

When the Hamiltonian is independent of time, $H = H(q_i, p_i)$, the total energy of the system is conserved. Consequently, the system's trajectory is confined to a *constant-energy hypersurface* in phase space,

$$S_E = \{(q_i, p_i) \in \mathcal{M} \mid H(q_i, p_i) = E\}.$$

This geometric structure is central in statistical mechanics, since many physical properties depend only on the energy of the system. A useful quantity is the phase-space volume corresponding to all states with energy less than or equal to E :

$$\Sigma(E) = \int_{0 \leq H(q_i, p_i) \leq E} d\Omega,$$

where $d\Omega$ is the dimensionless phase-space volume element defined in (2.2.1). Conceptually, this integral can be computed in two steps: first, by integrating over the hypersurface of fixed energy $S_{\mathcal{H}}$, and then by integrating over all energies from 0 to E :

$$\Sigma(E) = \int_0^E d\mathcal{H} \int_{S_{\mathcal{H}}} dS_{\mathcal{H}} = \int_0^E d\mathcal{H} \omega(\mathcal{H}),$$

where $dS_{\mathcal{H}}$ denotes the natural measure on the energy surface.

Definition 2.2 (Density of states). The **density of states** at energy E is defined as

$$\omega(E) = \int_{\mathcal{M}} d\Omega \delta(H(q_i, p_i) - E), \quad (2.2.4)$$

where δ is the Dirac delta function, which selects points in phase space with energy exactly equal to E . Equivalently, it can be expressed as the derivative of the cumulative phase-space volume:

$$\frac{\partial \Sigma(E)}{\partial E} = \lim_{\Delta E \rightarrow 0} \frac{\Sigma(E + \Delta E) - \Sigma(E)}{\Delta E} = \lim_{\Delta E \rightarrow 0} \frac{1}{\Delta E} \int_E^{E+\Delta E} d\mathcal{H} \omega(\mathcal{H}) \sim \frac{\omega(E)\Delta E}{\Delta E} = \omega(E),$$

thus enforcing the relation

$$\omega(E) = \frac{\partial \Sigma(E)}{\partial E}.$$

From the fundamental theorem of calculus, $F = \int dx f(x) \implies F' = f(x)$. It is important to note that $\omega(E)$ has dimensions of inverse energy, as it represents the number of microstates per unit energy interval at energy E , and it is recognized as the surface area of the constant-energy hypersurface S_E in phase space.

The density of states counts the number of accessible microstates at a given energy and provides a key link between microscopic Hamiltonian dynamics and macroscopic thermodynamic quantities such as entropy and temperature.

2.3 | Ergodicity

Let us now consider a system described by a stationary probability distribution $\rho(q_i, p_i)$ on the phase space \mathcal{M} . The statistical properties of the system can be extracted by computing ensemble averages of observables.

Definition 2.3 (Ensemble average). Given an observable $f : \mathcal{M} \rightarrow \mathbb{R}$, its **ensemble average** is defined as

$$\langle f \rangle = \int_{\mathcal{M}} \rho(q_i, p_i) f(q_i, p_i) d\Omega. \quad (2.3.1)$$

This represents the expected value of f when the system is sampled according to the probability distribution ρ , providing a statistical description of the observable's behaviour across the ensemble of microstates.

Definition 2.4 (Standard deviation). The fluctuations of an observable f around its mean value are quantified by the **standard deviation**, given by

$$(\Delta f)^2 = \langle f^2 \rangle - \langle f \rangle^2. \quad (2.3.2)$$

In the thermodynamic limit, fluctuations of macroscopic observables become negligible compared to their averages, ensuring the reproducibility of thermodynamic quantities.

The average and fluctuations of observables are central to the connection between microscopic dynamics and macroscopic thermodynamics, and they obviously depend on the probability distribution ρ . A fundamental question in statistical mechanics is whether these ensemble averages can be related to time averages computed along a single trajectory of the system. This leads us to the concept of **ergodicity**, which establishes a bridge between the long-term behaviour of individual systems and the statistical properties of ensembles. Before defining ergodicity, we need to introduce two more types of averages.

The **microcanonical average** of an observable f computed over the constant-energy hypersurface S_E is defined as:

$$\langle f \rangle_E = \frac{1}{\omega(E)} \int_{S_E} f(q_i, p_i) dS_E, \quad (2.3.3)$$

where dS_E is the natural measure on the constant-energy hypersurface and $\omega(E)$ is the density of states. This represents the average of f when all accessible microstates at fixed energy are equally probable, as in the microcanonical ensemble.

The **time average** of f along a trajectory starting from initial condition $(q_i(t_0), p_i(t_0))$ is given by

$$\langle f \rangle_\infty = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(q_i(t), p_i(t)) dt. \quad (2.3.4)$$

This describes the long-term average value of the observable along the dynamical evolution of a single system; it depends on the initial conditions and the specific trajectory followed, thus it may not be defined for some sets of initial conditions.

Definition 2.5 (Ergodic system). A system is said to be **ergodic** iff¹ the time average of any observable equals its microcanonical average for almost all initial conditions:

$$\langle f \rangle_\infty = \langle f \rangle_E. \quad (2.3.5)$$

¹If and only if.

Ergodicity provides the fundamental justification for using ensemble averages to describe macroscopic properties at thermodynamic equilibrium, as it ensures that the long-time behaviour of a single system is representative of the statistical behaviour of the entire ensemble.

Ergodicity is a non-trivial property that depends on the specific Hamiltonian and the nature of the interactions within the system; it's a strong assumption, which involves that the system explores the entire constant-energy hypersurface given enough time: every trajectory is dense in S_E .²

Integrable Systems. A special class of Hamiltonian systems are the *integrable systems*, which possess as many independent constants of motion (integrals) as degrees of freedom. In such systems, the motion is confined to invariant tori in phase space, and trajectories do not explore the entire constant-energy hypersurface. As a result, integrable systems are seen as the opposite of ergodic systems, since time averages depend on the initial conditions and do not coincide with microcanonical averages.

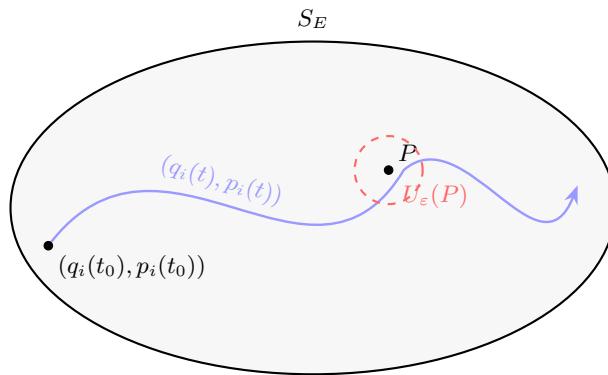


Figure 2.1: In an ergodic system, a single trajectory (blue curve) starting from an initial condition $(q_i(t_0), p_i(t_0))$ (black dot) will eventually come arbitrarily close to any point P on the constant-energy hypersurface S_E (gray ellipse). The red dashed circle represents a small neighborhood $U_\varepsilon(P)$ around the point P , which the trajectory will revisit infinitely often over time. This illustrates the concept of ergodicity, where time averages along a single trajectory coincide with ensemble averages over the entire hypersurface.

²A set is dense in a space if every point in the space is either in the set or is a limit point of the set, i.e., can be approximated arbitrarily closely by points in the set.

3 | Classical Microcanonical Ensemble

The microcanonical ensemble describes an isolated system with fixed total energy E , volume V , and number of particles N . It is the most fundamental statistical ensemble, appropriate for systems that do not exchange energy or particles with their surroundings. All accessible microstates are assumed to be equally probable, reflecting a state of maximal ignorance beyond the conserved quantities.

In this ensemble, the macroscopic state is specified by the constraints E, V, N , while the microscopic states are distributed uniformly over the hypersurface of constant energy in phase space. The key quantity is the *density of states* $\omega(E)$, which counts the number of microstates compatible with the fixed energy.

The microcanonical ensemble provides a direct link between mechanics and thermodynamics: the entropy becomes a function of the conserved quantities, and thermodynamic relations—such as temperature and pressure—can be derived from its derivatives. In the thermodynamic limit, the predictions of the microcanonical ensemble coincide with those of the canonical and grandcanonical ensembles, provided the system is sufficiently large and ergodic.

This ensemble is particularly useful for understanding the foundations of statistical mechanics, the emergence of equilibrium, and the behavior of isolated systems such as closed atomic clusters or idealized models in theoretical physics.

3.1 | Isolated System

A **macroscopic state** of an isolated system is specified by fixing three thermodynamic parameters: the total energy E , the volume V , and the particle number N . By definition, an isolated system does not exchange either energy or matter with the environment. As a consequence, the microscopic motion of the system is restricted to the constant-energy hypersurface

$$S_E = \{(q_i, p_i) \in \mathcal{M} : \mathcal{H}(q_i, p_i) = E\}.$$

The Hamiltonian $\mathcal{H}(q_i, p_i)$ is time independent, since E is constant and the macroscopic quantities do not vary with time.

Definition 3.1 (Microcanonical distribution). In the microcanonical ensemble we assume, *a priori*, that all accessible microstates compatible with the macroscopic constraints are equally probable. This is expressed by the probability density function

$$\rho_{mc}(q_i, p_i) = \frac{1}{\omega(E)} \delta(\mathcal{H}(q_i, p_i) - E), \quad (3.1.1)$$

where $\omega(E)$ is the density of states on the energy surface, ensuring proper normalization.¹

This expression can be obtained as the limiting case of a finite energy window. Suppose that the energy of the system is known only within an interval $[E, E + \Delta E]$. In this case, the probability density is defined as

$$\rho_{mc}(q_i, p_i) = \begin{cases} \frac{1}{\Gamma(E)} & \text{if } E \leq \mathcal{H}(q_i, p_i) \leq E + \Delta E, \\ 0 & \text{otherwise,} \end{cases}$$

where the normalization factor is given by

$$\Gamma(E) = \int_E^{E+\Delta E} \omega(E') dE' \simeq \omega(E) \Delta E.$$

In the limit $\Delta E \rightarrow 0$, the distribution becomes sharply concentrated on the surface S_E , recovering the delta-function form introduced above. This construction reflects the idea that, in the absence of further information, the best statistical description of an isolated system is to assign equal probability to all microstates consistent with the conservation laws.

¹Since the integral over the phase space is equal to 1 for a properly normalized distribution, the normalization condition reads:

$$\int_{\mathcal{M}} d\Omega \rho_{mc}(q_i, p_i) = 1 \implies \int_{\mathcal{M}} d\Omega \delta(\mathcal{H}(q_i, p_i) - E) = \frac{1}{C} = \omega(E).$$

3.2 | Microcanonical Entropy

The central thermodynamic quantity associated with the microcanonical ensemble is the entropy, defined according to Boltzmann's principle.

Definition 3.2 (Microcanonical entropy). For an isolated system with fixed energy E , volume V , and particle number N , the microcanonical entropy is

$$S_{mc}(E, V, N) \equiv k_B \log \omega(E), \quad (3.2.1)$$

where $\omega(E)$ denotes the density of states at energy E and k_B is the Boltzmann constant: $k_B \simeq 1.38 \times 10^{-23} \text{ J/K}$, the fundamental constant relating temperature to energy.

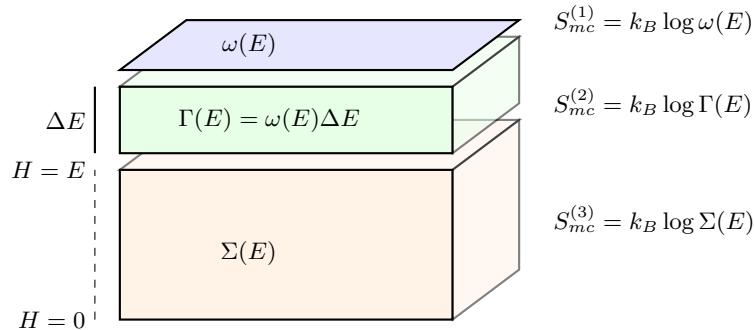
The density of states $\omega(E)$ practically takes into account the number of microstates accessible to the system at the specified energy. The Boltzmann constant k_B serves to convert the logarithmic measure of microstate multiplicity into physical units of entropy.

While all the quantities upon which entropy depends are macroscopic and extensive, meaning that they scale linearly with the size of the system (e.g., with N and V), the density of states $\omega(E)$ scales exponentially with N : $\omega(E) \sim e^{\beta N}$ for some β . That's why the logarithm is taken in the definition of entropy, ensuring that S_{mc} itself is extensive and scales linearly with N .

3.2.1 | Thermodynamic Limit

In the thermodynamic limit (denoted by \lim_{td} , corresponding to $N \rightarrow \infty$, $V \rightarrow \infty$, with $n = N/V$ finite), the **specific entropy** is defined as:

$$s_{mc} = \lim_{td} \frac{S_{mc}(E, V, N)}{N}.$$



By exploiting the relations between $\omega(E)$, $\Gamma(E)$ and $\Sigma(E)$, one can equivalently write:

$$s_{mc} = k_B \lim_{td} \frac{\log \omega(E)}{N} = k_B \lim_{td} \frac{\log \Gamma(E)}{N} = k_B \lim_{td} \frac{\log \Sigma(E)}{N} = s_{th}.$$

This identity is formally always true, although its physical realization can be explicitly verified only in specific models, such as the ideal gas. Thus we found that in the thermodynamic limit for the microcanonical ensemble there is a uniquely defined entropy.

At equilibrium, the properties of entropy satisfy the following key principles:

- **Additivity:** For two independent subsystems, entropy is additive:

$$S_{mc}^{(\text{tot})} = S_{mc}^{(1)} + S_{mc}^{(2)}.$$

Proof. Consider two independent subsystems with energies E_1 and E_2 , such that the total energy is conserved: $E = E_1 + E_2$. The two subsystems are independent:

$$\mathcal{S}_1 : \begin{cases} \mathcal{H}_1(q_i^{(1)}, p_i^{(1)}) = E_1, & (q_i^{(1)}, p_i^{(1)}) \in \mathcal{M}_1, \\ \omega_1(E_1) = \int_{S_{\mathcal{H}_1=E_1}} dS_{\mathcal{H}_1} \implies S_{mc}^{(1)} = k_B \log \omega_1(E_1), \end{cases}$$

and the same applies to the second subsystem \mathcal{S}_2 such that $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$. The total system $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ has a total density of states given by

$$\omega(E) = \int_{\mathcal{M}} d\Omega \delta(\mathcal{H} - E) = \int_{\mathcal{M}_1} \int_{\mathcal{M}_2} d\Omega_1 d\Omega_2 \delta(\mathcal{H}_1 + \mathcal{H}_2 - E).$$

This integral can be decomposed by introducing an integral over the energy partition:

$$\omega(E) = \int d\mathcal{H}_1 \int d\mathcal{H}_2 \int dS_{\mathcal{H}_1} \int dS_{\mathcal{H}_2} \delta(\mathcal{H}_2 - (E - \mathcal{H}_1)) = \int_0^E d\mathcal{H}_1 \omega_1(\mathcal{H}_1) \omega_2(E - \mathcal{H}_1).$$

This quantity can be bounded above by the product of the densities of states of the subsystems, evaluated at the most probable energy partition (i.e. the one that maximizes the product) and multiplied by the total energy E (the upper bound of the integral):

$$\omega(E) \leq E \omega_1(E_1^*) \omega_2(E_2^*),$$

where E_1^* is the value that maximizes the product $\omega_1(E_1)\omega_2(E-E_1)$, and $E_2^* = E - E_1^*$. For a sufficiently small energy resolution ΔE , we can write the following bounds:

$$\Delta E \omega_1(E_1^*) \omega_2(E_2^*) \leq \omega(E) \leq E \omega_1(E_1^*) \omega_2(E_2^*),$$

Multiplying both sides by ΔE , we obtain:

$$(\Delta E)^2 \omega_1(E_1^*) \omega_2(E_2^*) \leq \Delta E \omega(E) \leq E \Delta E \omega_1(E_1^*) \omega_2(E_2^*).$$

Defining $\Gamma(E) = \Delta E \omega(E)$ as the number of microstates in the energy shell $[E, E + \Delta E]$, we get:

$$\Gamma_1(E_1^*) \Gamma_2(E_2^*) \leq \Gamma(E) \leq \frac{E}{\Delta E} \Gamma_1(E_1^*) \Gamma_2(E_2^*).$$

To reconstruct the entropy, we take the logarithm of the inequality and multiply by Boltzmann's constant k_B :²

$$S_1(E_1^*) + S_2(E_2^*) \leq S(E) \leq S_1(E_1^*) + S_2(E_2^*) + k_B \log \left(\frac{E}{\Delta E} \right).$$

Now we consider the thermodynamic limit, dividing by the total number of degrees of freedom N and taking $N \rightarrow \infty$:

$$\frac{S_1(E_1^*) + S_2(E_2^*)}{N} \leq \frac{S(E)}{N} \leq \frac{S_1(E_1^*) + S_2(E_2^*)}{N} + \frac{k_B \log \left(\frac{E}{\Delta E} \right)}{N}.$$

Since the logarithmic correction vanishes in the thermodynamic limit (because \log grows sublinearly with N), we obtain:

$$s_1(E_1^*) + s_2(E_2^*) \leq s(E) \leq s_1(E_1^*) + s_2(E_2^*),$$

²The logarithm is a monotonic function, so the inequality signs are preserved.

which implies:

$$S_{mc}^{(\text{tot})} = S_{mc}^{(1)} + S_{mc}^{(2)}.$$

□

Remark. This property of entropy emerges in the thermodynamic limit at equilibrium, where the energies E_1^* and E_2^* are those that maximize the product of the densities of states of the two subsystems. This reflects the fact that the equilibrium configuration is the most probable one, corresponding to the maximum number of accessible microstates.

- **Equivalence with thermodynamic entropy:** In the thermodynamic limit,

$$S_{mc} = S_{th},$$

where S_{th} denotes the entropy defined by macroscopic thermodynamics.

Proof. Since $\Gamma_1(E_1)\Gamma_2(E_2) = (\Delta E)^2 \omega_1(E_1)\omega_2(E_2)$, we can take its variation at the maximum point (E_1^*, E_2^*) , which must vanish at equilibrium:

$$\delta \{\Gamma_1(E_1)\Gamma_2(E_2)\}_{(E_1^*, E_2^*)} = 0 = \frac{\partial \Gamma_1(E_1)}{\partial E_1} \Bigg|_{E_1^*} \Delta E_1 \Gamma_2(E_2^*) + \Gamma_1(E_1^*) \frac{\partial \Gamma_2(E_2)}{\partial E_2} \Bigg|_{E_2^*} \Delta E_2.$$

Since the total energy is conserved, we have $\Delta E = \Delta E_1 + \Delta E_2 = 0$, which implies $\Delta E_1 = -\Delta E_2$. Substituting this, we obtain:

$$\frac{1}{\Gamma_1(E_1^*)} \frac{\partial \Gamma_1(E_1)}{\partial E_1} \Bigg|_{E_1^*} = \frac{1}{\Gamma_2(E_2^*)} \frac{\partial \Gamma_2(E_2)}{\partial E_2} \Bigg|_{E_2^*},$$

which corresponds to the equality of logarithmic derivatives. Recalling the definition of microcanonical entropy, we can rewrite this as:

$$\frac{\partial S_{mc}^{(1)}}{\partial E_1} \Bigg|_{E_1^*} = \frac{\partial S_{mc}^{(2)}}{\partial E_2} \Bigg|_{E_2^*}.$$

This condition is equivalent to the thermodynamic equilibrium condition:

$$\frac{\partial S_{mc}^{(1)}}{\partial E_1} \Bigg|_{E_1^*} = \frac{\partial S_{mc}^{(2)}}{\partial E_2} \Bigg|_{E_2^*} = \frac{1}{T},$$

which expresses the equality of temperatures at equilibrium. Therefore, we conclude that the microcanonical entropy S_{mc} coincides with the thermodynamic entropy S_{th} in the thermodynamic limit. □

Finally, the microcanonical entropy can also be expressed in the form of Boltzmann's universal formula, which relates entropy to the average of the logarithm of the distribution function:

Proposition 3.1 (Boltzmann's formula). *For the microcanonical ensemble one has*

$$s_{mc} = -k_B \lim_{td} \frac{1}{N} \langle \log \rho_{mc} \rangle_{mc} = -k_B \lim_{td} \frac{1}{N} \int d\Omega \rho_{mc} \log \rho_{mc}. \quad (3.2.2)$$

This shows the direct connection between Boltzmann's statistical definition of entropy and the Gibbs–Shannon entropy of probability distributions, when restricted to the microcanonical setting.

3.3 | Perfect Gas

As an explicit example of the microcanonical formalism, let us consider a gas of N non-relativistic and non-interacting monoatomic particles in three spatial dimensions. This is the paradigmatic system used to test the consistency of statistical mechanics with the laws of thermodynamics.

The phase space of the system is

$$\mathcal{M} = V^N \times \mathbb{R}^{3N} = \left\{ \{(q_i, p_i)\}_{i=1}^N : q_i \in V, p_i \in \mathbb{R}^3 \right\},$$

where each particle can move freely inside the container of volume V , and the momentum space is unbounded. The Hamiltonian is purely kinetic:

$$\mathcal{H}(q_i, p_i) = \sum_{i=1}^N \frac{p_i^2}{2m}, \quad p_i = (p_i^x, p_i^y, p_i^z).$$

Volume and density of states

The accessible phase space volume at fixed energy E can be computed explicitly as the integral over the phase space constrained by the energy shell,³ yielding

$$\Sigma(E) = \frac{2}{3} \left(\frac{V}{h^3} \right)^N \frac{(2\pi m E)^{3N/2}}{N \Gamma(3N/2)}.$$

From this, the density of states follows:

$$\omega(E) = \frac{\partial \Sigma(E)}{\partial E} = \frac{3N}{2E} \Sigma(E), \quad \Gamma(E) = \frac{3N}{2E} \Sigma(E) \Delta E.$$

Hence, in the thermodynamic limit,

$$\lim_{td} \frac{\log \Sigma(E)}{N} = \lim_{td} \frac{\log \omega(E)}{N} = \lim_{td} \frac{\log \Gamma(E)}{N}.$$

Entropy for distinguishable particles

For distinguishable particles, the entropy can be computed in the thermodynamic limit as $S = k_B \log(\Sigma(E))$, remembering that

$$\Gamma(x+1) = x! = x \Gamma(x) \Rightarrow \Gamma(1/2) = \sqrt{\pi}.$$

Thus, using Stirling's approximation for large N , $\log N! \approx N \log N - N$, one obtains

$$S = \frac{3}{2} N k_B + N k_B \log \left[V \left(\frac{4\pi m E}{3N h^2} \right)^{3/2} \right].$$

From this expression one recovers the standard thermodynamic relations:

$$\frac{1}{T} = \frac{\partial S}{\partial E} \Big|_{V,N} \Rightarrow E = \frac{3}{2} N k_B T,$$

$$\frac{p}{T} = \frac{\partial S}{\partial V} \Big|_{E,N} \Rightarrow pV = N k_B T.$$

However, this entropy is not extensive, as it grows faster than linearly with N . Furthermore, we will interpret the result for the energy as the *classical equipartition theorem*, assigning an average energy of $\frac{1}{2}k_B T$ to each quadratic degree of freedom.

³The integral over space will give the volume to the power of N , while for the integral over momenta one can go to spherical coordinates and use the volume of a $3N$ -dimensional sphere: $\frac{\pi^{d/2}}{\Gamma(d/2+1)} R^d$.

Entropy for indistinguishable particles

Correct extensivity is restored by taking into account the indistinguishability of identical particles, which introduces the Gibbs correction factor $1/N!$.⁴ The resulting entropy is

$$S = \frac{5}{2}Nk_B + 3Nk_B \log\left(\frac{d}{\lambda_T}\right),$$

where

$$d^3 \equiv \frac{V}{N}, \quad \lambda_T \equiv \sqrt{\frac{h^2}{2\pi mk_B T}}$$

is the *thermal wavelength*. This expression is extensive and coincides with the classical Sackur–Tetrode formula.

Limits of validity

The microcanonical entropy becomes negative for sufficiently low temperatures,

$$T < T^* \equiv \frac{e^{-5/3}h^2}{2\pi mk_B d^2},$$

corresponding to densities such that $d \lesssim \lambda_T$. This signals the breakdown of the classical approximation, and the necessity to include quantum statistics (Bose–Einstein or Fermi–Dirac) in the description of the gas.

⁴This factor was naturally incorporated inside the measure of the phase space in (2.2.1) in order to reach a rigorous treatment of the indistinguishable states (renormalizing accordingly) in the phase space.

4 | Classical Canonical Ensemble

In many physical situations, a system cannot be regarded as completely isolated, but rather as being in thermal contact with a large reservoir that can exchange energy with it. While the total energy of the combined system (the system plus its environment) is conserved, the energy of the subsystem of interest fluctuates around an equilibrium value determined by the temperature of the reservoir.

The statistical description appropriate for such conditions is the *canonical ensemble*. In this ensemble, the macroscopic state of the system is specified by fixing the temperature T , the volume V , and the number of particles N , while the energy is allowed to vary according to a well-defined probability distribution. The fundamental quantity that characterizes the canonical ensemble is the *partition function*, which encodes the statistical weight of all accessible microstates.

From the partition function, all thermodynamic quantities—such as the internal energy, entropy, and Helmholtz free energy—can be derived. The canonical ensemble thus provides a natural bridge between microscopic mechanics and macroscopic thermodynamics, allowing us to compute equilibrium properties of systems in contact with a thermal bath.

4.1 | System in a Thermal Bath

Let us consider a system \mathcal{S} in thermal contact with a large reservoir \mathcal{E} , which we may refer to as the *environment* or *thermal bath*. From a macroscopic point of view, the system is characterized by fixed values of temperature, volume, and number of particles (T, V, N) . Microscopically, however, its state is determined by a point in its phase space,

$$\mathcal{M}_{\mathcal{S}} = \{(q_i, p_i)\}, \quad i = 1, \dots, N,$$

where each pair (q_i, p_i) specifies the coordinates and momenta of the particles that constitute the system.

At thermal equilibrium, both the system and the reservoir share the same temperature,

$$T_1 = T_2 = T,$$

and are free to exchange energy through microscopic interactions. Nevertheless, the total energy of the composite system, which we may call the *universe* $\mathcal{U} = \mathcal{S} + \mathcal{E}$, remains constant:

$$E = E_1 + E_2 = \text{const.}$$

Therefore, while the subsystem \mathcal{S} and its environment fluctuate in energy, the complete system \mathcal{U} as a whole is described by a **microcanonical ensemble**, since its total energy is fixed.

The goal is to determine the probability distribution describing the microscopic states of the subsystem \mathcal{S} , when it is in contact with the thermal bath. Starting from the microcanonical distribution of the total system, we write

$$\rho_{mc}(q_i^{(1)}, p_i^{(1)}, q_i^{(2)}, p_i^{(2)}) = \frac{1}{\omega(E)} \delta(\mathcal{H}(q_i^{(1)}, p_i^{(1)}, q_i^{(2)}, p_i^{(2)}) - E),$$

where $\omega(E)$ is the density of states of the entire universe at total energy E . To obtain the effective probability density for the subsystem, we integrate out the environmental degrees of freedom:

$$\rho_c^{(1)}(q_i^{(1)}, p_i^{(1)}) = \int_{\mathcal{M}_2} d\Omega_2 \rho_{mc}(q_i^{(1)}, p_i^{(1)}, q_i^{(2)}, p_i^{(2)}).$$

By recalling the expression for the microcanonical probability distribution, we get

$$\rho_c^{(1)} = \frac{1}{\omega(E)} \int_{\mathcal{M}_2} d\Omega_2 \delta(E_1 + E_2 - E) = \frac{1}{\omega(E)} \omega_2(E - E_1),$$

where $\omega_2(E_2)$ denotes the density of states of the reservoir evaluated at energy $E_2 = E - E_1$. The probability of finding the subsystem in a state of energy E_1 is therefore proportional to the number of accessible microstates of the environment compatible with energy $E_2 = E - E_1$.

To proceed, it is convenient to express the density of states of the environment in terms of its entropy:

$$S_2(E_2) = k_B \ln \omega_2(E_2).$$

Since the environment is assumed to be much larger than the subsystem, its energy variations $\Delta E_2 = -E_1$ are extremely small compared with its total energy. This allows us to expand $S_2(E_2)$ in a Taylor series around the point $E_2 = E$:

$$S_2(E - E_1) \approx S_2(E) - E_1 \left(\frac{\partial S_2}{\partial E_2} \right)_{E_2=E} + \frac{E_1^2}{2} \left(\frac{\partial^2 S_2}{\partial E_2^2} \right)_{E_2=E} + \dots$$

For a large reservoir, the higher-order terms can be neglected, since its temperature is essentially unaffected by the small energy exchanges. From thermodynamics we know that

$$\frac{\partial S}{\partial E} = \frac{1}{T}, \text{ from } dE = TdS - pdV + \mu dN,$$

so that

$$S_2(E - E_1) \approx S_2(E) - \frac{E_1}{T}.$$

By exponentiating this relation, we recover the corresponding expression for the density of states:

$$\omega_2(E - E_1) = e^{S_2(E - E_1)/k_B} \approx e^{S_2(E)/k_B} e^{-E_1/(k_B T)}.$$

The first factor $e^{S_2(E)/k_B}$ is constant because the total energy E of the universe is fixed. Consequently, the probability of finding the subsystem in a microstate of energy E_1 is proportional to

$$\rho_c^{(1)} \propto e^{-\beta E_1}, \quad \beta = \frac{1}{k_B T}.$$

This exponential dependence on the subsystem energy defines the canonical ensemble. Physically, it reflects the fact that configurations of the subsystem with higher energy correspond to fewer available microstates for the environment, and are thus less probable. The term with the opposite sign in the exponent, $e^{+\beta E_1}$, would imply an increasing probability for higher energies and would make normalization impossible, so only the decaying exponential form is physically acceptable.

Hence, the canonical probability distribution for a system in thermal equilibrium with a heat bath is given by

$$\rho_c(q_i, p_i) = \frac{1}{Z} e^{-\beta \mathcal{H}(q_i, p_i)}, \quad (4.1.1)$$

where the normalization factor Z is known as the **canonical partition function**. It is obtained by requiring that the total probability over phase space be one:

$$Z = \int_{\mathcal{M}_S} d\Omega e^{-\beta \mathcal{H}(q_i, p_i)}. \quad (4.1.2)$$

For a system of N particles in d dimensions this becomes¹

$$Z = \frac{1}{\xi_N} \int \prod_{i=1}^N \frac{d^d q_i d^d p_i}{h^d} e^{-\beta \mathcal{H}(q_i, p_i)} = \frac{1}{\xi_N} \int_0^\infty dE e^{-\beta E} \omega(E).$$

The partition function plays a central role in statistical mechanics, as it contains all the thermodynamic information about the system. Once Z is known, quantities such as the internal energy, entropy, and free energy can be derived from it through standard thermodynamic relations.

Finally, let us define the notion of an average quantity within this ensemble.

Definition 4.1. Canonical average. If $f(q_i, p_i)$ is any observable function defined on the phase space of the system, its average over the canonical ensemble is given by

$$\langle f \rangle_c = \int_{\mathcal{M}_S} d\Omega \rho_c(q_i, p_i) f(q_i, p_i) = \frac{1}{Z} \int_{\mathcal{M}_S} d\Omega e^{-\beta \mathcal{H}(q_i, p_i)} f(q_i, p_i). \quad (4.1.3)$$

This operation, often called the *canonical average*, provides the bridge between the microscopic description of the system and the macroscopic quantities that can be experimentally measured.

In equilibrium statistical mechanics, macroscopic observables are nothing but ensemble averages of microscopic functions evaluated with the canonical distribution.

¹Note the measure of the phase space: we have incorporated the *indistinguishability factor* ξ_N : $\xi_N = N!$ for indistinguishable particles, and $\xi_N = 1$ for distinguishable ones. Also, the factor h^{dN} ensures the correct dimensionality and connects classical and quantum descriptions. Now we have properly defined a general measure on phase space.

4.2 | Partition Function and Thermodynamic Quantities

Recalling the expression found for the partition function, we can now analyze its physical dependencies:

$$Z = \int_{\mathcal{M}_S} d\Omega e^{-\beta \mathcal{H}} = Z(V, T, N), \quad (4.2.1)$$

that is, the partition function depends on the thermodynamic variables characterizing the canonical ensemble: the volume V , the temperature T , and the number of particles N . In this sense, Z plays the role of a thermodynamic potential from which all equilibrium properties can be derived.

A particularly useful relation expresses the Helmholtz free energy in terms of the partition function, providing a bridge between thermodynamics and statistical mechanics:

$$F(T, V, N) = E - TS = -\frac{1}{\beta} \log Z.$$

The logarithm is essential in this definition, since it restores the correct *extensivity* of the thermodynamic potential. Indeed, while the partition function Z scales exponentially with the system size ($Z \propto V^N$), its logarithm grows linearly with V , N , and T :

$$\log Z \propto N \log V \Rightarrow F \propto V, N, T.$$

Hence, by taking the logarithm we recover a quantity that behaves properly under a rescaling of the system. Let's demonstrate the relation between F and Z .

From the definition of F we can write:

$$Z = e^{-\beta F} = \int_{\mathcal{M}_S} d\Omega e^{-\beta \mathcal{H}}.$$

This identity can be rearranged as:

$$\int_{\mathcal{M}_S} d\Omega e^{\beta(F-\mathcal{H})} = 1.$$

Differentiating both sides with respect to β gives:

$$\frac{\partial}{\partial \beta} \int_{\mathcal{M}_S} d\Omega e^{\beta(F-\mathcal{H})} = 0,$$

from which it follows:

$$\int_{\mathcal{M}_S} d\Omega e^{\beta(F-\mathcal{H})} \left[(F - \mathcal{H}) + \beta \frac{\partial F}{\partial \beta} \right] = 0.$$

Since the integrand is weighted by the canonical probability density

$$\frac{e^{-\beta \mathcal{H}}}{Z},$$

we can rewrite the previous expression in the form:²

$$F \int_{\mathcal{M}_S} d\Omega \frac{e^{-\beta \mathcal{H}}}{Z} = \int_{\mathcal{M}_S} d\Omega \frac{e^{-\beta \mathcal{H}}}{Z} \mathcal{H} + T \frac{\partial F}{\partial T} \int_{\mathcal{M}_S} d\Omega \frac{e^{-\beta \mathcal{H}}}{Z}.$$

Recognizing the expression for the canonical average, we obtain:

$$F = \langle \mathcal{H} \rangle_c + T \frac{\partial F}{\partial T}.$$

Finally, recalling that $\langle \mathcal{H} \rangle_c = E$ and $S = -\frac{\partial F}{\partial T}$, we recover the thermodynamic definition of the Helmholtz free energy:

$$F = E - TS.$$

²We have also used $\beta \frac{\partial}{\partial \beta} = -T \frac{\partial}{\partial T}$.

Remark. The Helmholtz free energy $F(T, V, N)$ is therefore the natural thermodynamic potential of the canonical ensemble. Its differential form,

$$dF = -S dT - p dV + \mu dN,$$

directly provides the equilibrium relations between macroscopic observables and serves as the starting point for the derivation of all other state functions.

Another useful relation connects the internal energy to the partition function:

$$E = \langle \mathcal{H} \rangle_c = \int_{\mathcal{M}_S} d\Omega \frac{e^{-\beta \mathcal{H}}}{Z} \mathcal{H} = \frac{1}{Z} \left(-\frac{\partial Z}{\partial \beta} \right) = -\frac{\partial}{\partial \beta} \log Z. \quad (4.2.2)$$

This expression highlights how the mean energy of the system can be derived directly from the temperature dependence of the partition function. Once the partition function Z is known, all macroscopic thermodynamic quantities follow systematically: the Helmholtz free energy from $F = -k_B T \log Z$, the internal energy from $E = -\partial_\beta \log Z$, and finally the entropy from the thermodynamic identity $S = (E - F)/T$. In summary,

$$Z \Rightarrow F, E, S,$$

showing that the knowledge of the partition function completely determines the equilibrium thermodynamics.

Entropy. Let us recall the result obtained for the microcanonical ensemble, where the statistical and thermodynamic definitions of entropy coincide,

$$S_{mc} = -k_B \langle \log \rho_{mc} \rangle_{mc} = S_{th}.$$

We now aim to demonstrate the corresponding relation in the canonical ensemble:

$$\begin{aligned} S_c &= -k_B \langle \log \rho_c \rangle_c = -k_B \int_{\mathcal{M}_S} d\Omega \rho_c \log \rho_c \\ &= -k_B \int_{\mathcal{M}_S} d\Omega \rho_c (\log(e^{-\beta \mathcal{H}}) - \log(Z)) \\ &= k_B \int_{\mathcal{M}_S} d\Omega \rho_c (\beta \mathcal{H}) + k_B \log(Z) \int_{\mathcal{M}_S} d\Omega \rho_c \\ &= \frac{k_B}{k_B T} \langle \mathcal{H} \rangle_c + k_B \log(Z) \\ &= \frac{E - F}{T} = S_{th} = S_c, \end{aligned}$$

practically proving Boltzmann's universal formula. Hence, in the thermodynamic limit we have

$$s_{mc} = s_c = s_{th}.$$

4.2.1 | Distinguishable and Indistinguishable Systems

Consider a system composed of multiple independent and distinguishable subsystems. If we denote by A and B two such subsystems, the total system can be expressed as:

$$\mathcal{S} = A \cup B, \quad \mathcal{H} = \mathcal{H}_A + \mathcal{H}_B, \quad \mathcal{M} = \mathcal{M}_A \otimes \mathcal{M}_B, \quad d\Omega = d\Omega_A d\Omega_B.$$

The statistical independence of the two parts implies that the total partition function factorizes as:

$$Z = \int_{\mathcal{M}_A \otimes \mathcal{M}_B} d\Omega_A d\Omega_B e^{-\beta(\mathcal{H}_A + \mathcal{H}_B)} = Z_A Z_B.$$

More generally, for a system composed of N independent, distinguishable subsystems, we obtain the product rule:

$$Z = \prod_{i=1}^N Z_i.$$

This property reflects the fact that, for independent subsystems, the total probability measure in phase space factorizes, and so does the corresponding statistical weight.

Example (Identical particles). Consider a system of N non-interacting, non-relativistic, identical but *distinguishable* free particles in one dimension. The Hamiltonian is additive:

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

and the total partition function follows immediately:

$$Z^{(\text{dist})} = \int e^{-\beta\mathcal{H}} d\Omega = (Z_1)^N,$$

where Z_1 is the single-particle partition function.

However, if the particles are *indistinguishable*, the situation changes. Since any permutation of the N particles represents the same microscopic state, we must divide by the number of equivalent configurations, namely $N!$. The correct expression for the canonical partition function is therefore:

$$Z^{(\text{ind})} = \frac{(Z_1)^N}{N!}.$$

This correction accounts for the overcounting of identical configurations in phase space and ensures the proper extensivity of thermodynamic quantities such as entropy and free energy.

4.3 | Generalized Equipartition Theorem

The equipartition theorem provides a fundamental link between the microscopic degrees of freedom of a system and its macroscopic thermal properties. In its generalized form, it states that each parameter ξ_i , coordinate or its conjugate momentum for $i = 1, \dots, 2dN$, contribute to the mean energy according to their appearance in the Hamiltonian. This result extends the classical formulation beyond the simple quadratic case, offering a deeper understanding of how thermal energy is distributed among different types of degrees of freedom.

Theorem 4.1 (Generalized Equipartition Theorem). *For a canonical ensemble described by the Hamiltonian $\mathcal{H}(\{\xi_i\})$, the following relation holds for each generalized coordinate or momentum $\xi_j \in [a, b]$:*

$$k_B T = \left\langle \xi_j \frac{\partial \mathcal{H}}{\partial \xi_j} \right\rangle_c, \quad (4.3.1)$$

provided that the boundary term vanishes:

$$\xi_j e^{-\beta \mathcal{H}} \Big|_{\xi_j=a}^{\xi_j=b} = 0, \quad (4.3.2)$$

where a and b are typically extended to $\pm\infty$.

Proof. Consider the canonical measure on phase space:³

$$d\Omega e^{-\beta \mathcal{H}} = \left(\prod_{i=1}^{2dN} d\xi_i \right) e^{-\beta \mathcal{H}}.$$

We can isolate the differential with respect to one variable ξ_j and differentiate the product $\xi_j e^{-\beta \mathcal{H}}$ to obtain

$$d\Omega e^{-\beta \mathcal{H}} = \left(\prod_{i \neq j} d\xi_i \right) d\xi_j e^{-\beta \mathcal{H}} = \left(\prod_{i \neq j} d\xi_i \right) \left[d(\xi_j e^{-\beta \mathcal{H}}) - \xi_j e^{-\beta \mathcal{H}} \left(-\beta \frac{\partial \mathcal{H}}{\partial \xi_j} \right) d\xi_j \right],$$

which simplifies to

$$d\Omega e^{-\beta \mathcal{H}} = \left(\prod_{i \neq j} d\xi_i \right) d(\xi_j e^{-\beta \mathcal{H}}) + d\Omega \beta \xi_j \frac{\partial \mathcal{H}}{\partial \xi_j} e^{-\beta \mathcal{H}}.$$

Using the normalization condition of the canonical ensemble,

$$1 = \frac{1}{Z} \int d\Omega e^{-\beta \mathcal{H}},$$

and substituting the expression above, we get:

$$1 = \frac{1}{Z} \left[\int \left(\prod_{i \neq j} d\xi_i \right) \int_a^b d(\xi_j e^{-\beta \mathcal{H}}) + \beta \int d\Omega \xi_j \frac{\partial \mathcal{H}}{\partial \xi_j} e^{-\beta \mathcal{H}} \right].$$

Requiring the boundary term to vanish,

$$\xi_j e^{-\beta \mathcal{H}} \Big|_{\xi_j=a}^{\xi_j=b} = 0,$$

³Since we consider a dimensionless measure and we include the ξ_N factor in $d\Omega$, we will consider these contributions to have been absorbed in the definition of the generalized variables ξ_i (it also helps to avoid confusion between the ind. factor ξ_N with the generalized variables ξ_i in the computation).

with $a, b \rightarrow \pm\infty$, the first term disappears and we obtain:

$$\frac{1}{\beta} = k_B T = \left\langle \xi_j \frac{\partial \mathcal{H}}{\partial \xi_j} \right\rangle_c.$$

□

Corollary 4.2 (Classical Equipartition Theorem). *If the Hamiltonian depends quadratically on one of its canonical variables, namely*

$$\mathcal{H} = A \xi_j^2 + \tilde{\mathcal{H}}(\xi_{i \neq j}),$$

with A constant and $\xi_j \in (-\infty, \infty)$, the boundary term

$$\xi_j e^{-\beta \mathcal{H}} \Big|_{-\infty}^{\infty} \sim \xi_j e^{-\beta A \xi_j^2} \Big|_{-\infty}^{\infty} = 0$$

automatically vanishes,⁴ and the generalized equipartition theorem yields:

$$\left\langle \xi_j \frac{\partial \mathcal{H}}{\partial \xi_j} \right\rangle_c = \langle 2A \xi_j^2 \rangle_c = k_B T.$$

Hence, the mean energy associated with each quadratic degree of freedom is

$$E_{\xi_j} = \langle A \xi_j^2 \rangle_c = \frac{1}{2} k_B T. \quad (4.3.3)$$

Therefore, every quadratic canonical variable, whether a coordinate or a conjugate momentum, contributes an average energy of $\frac{1}{2} k_B T$ to the total internal energy of the system.

The equipartition theorem thus provides a direct and general rule for distributing the thermal energy among the accessible microscopic degrees of freedom. Its consequences become particularly clear when applied to systems with quadratic Hamiltonians, as illustrated in the following examples.

Example (Ideal Gas). For a gas of N non-interacting, non-relativistic particles in d spatial dimensions, the Hamiltonian depends quadratically only on the dN momentum components:

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

Each quadratic term contributes an average energy of $\frac{1}{2} k_B T$, hence:

$$E = \frac{dN}{2} k_B T.$$

This result shows that the internal energy of a classical ideal gas is entirely kinetic and depends only on the number of degrees of freedom, not on the nature of the particles or their interactions.

Example (Harmonic Oscillators). For a system of N independent harmonic oscillators in d dimensions, the Hamiltonian is quadratic in both momenta and coordinates:

$$\mathcal{H} = \sum_{i=1}^N \left(\frac{|p_i|^2}{2m} + \frac{1}{2} m \omega^2 |q_i|^2 \right).$$

Since each oscillator contributes two quadratic terms (one kinetic, one potential), the total internal energy becomes:

$$E = dN k_B T.$$

Thus, the equipartition theorem immediately provides the temperature dependence of the mean energy and, consequently, of the heat capacity for harmonic systems.

⁴The integrand is Gaussian and centered at $\xi_j = 0$, so the product decays exponentially at infinity.

5 | Classical Grandcanonical Ensemble

In many physical situations, the number of particles in a system is not strictly fixed, but can fluctuate due to exchange with a surrounding reservoir. This occurs, for example, in open systems that allow particle transfer—such as chemical reactions, adsorption processes, or systems in contact with a particle bath. While the total energy and total particle number of the combined system (system plus reservoir) remain conserved, the subsystem of interest exhibits fluctuations in both energy and particle number.

The appropriate statistical framework for describing such systems is the *grandcanonical ensemble*. In this ensemble, the macroscopic state of the system is specified by fixing the temperature T , the volume V , and the chemical potential μ , while both energy and particle number are allowed to fluctuate according to a well-defined probability distribution.

The central quantity that characterizes the grandcanonical ensemble is the *grand partition function*, which encodes the statistical weight of all accessible microstates across different particle numbers. It plays a role analogous to the canonical partition function, but extended to include particle exchange. From the grand partition function, one can derive all relevant thermodynamic quantities—such as the average energy, average particle number, entropy, and the grand potential $\Phi = -k_B T \log \mathcal{Z}$. The grandcanonical ensemble thus provides a powerful bridge between microscopic dynamics and macroscopic thermodynamics for open systems, allowing us to compute equilibrium properties in the presence of both thermal and particle exchange.

5.1 | System in a Particle and Thermal Bath

Let us now consider a system \mathcal{S} that is not only in thermal contact with a reservoir \mathcal{E} , but also able to exchange particles with it. This situation arises naturally in open systems, such as chemical reactions, adsorption processes, or quantum gases in contact with a particle bath.

From a macroscopic perspective, the state of the system is specified by fixed values of temperature, volume, and chemical potential (T, V, μ) , while both energy and particle number are allowed to fluctuate. Microscopically, the system is described by its phase space, now extended to include configurations with varying particle number:

$$\mathcal{M}_{\mathcal{S}} = \bigcup_N \{(q_i, p_i)\}_{i=1}^N.$$

At equilibrium, the system and the reservoir share the same temperature and chemical potential:

$$T_1 = T_2 = T, \quad \mu_1 = \mu_2 = \mu.$$

Energy and particles can be exchanged freely, but the total energy and total particle number of the composite system, the *universe* $\mathcal{U} = \mathcal{S} + \mathcal{E}$, remain conserved:

$$E = E_1 + E_2 = \text{const}, \quad N = N_1 + N_2 = \text{const}.$$

Thus, the full system \mathcal{U} is described by a microcanonical ensemble in the extended space of energy and particle number.

Our goal is to derive the probability distribution that governs the microscopic states of the subsystem \mathcal{S} , when it is allowed to exchange both energy and particles with the reservoir. As in the canonical case, we start from the microcanonical distribution of the universe and integrate out the environmental degrees of freedom.

The resulting marginal distribution for the subsystem will depend on both its energy E_1 and particle number N_1 , and will be proportional to the number of microstates available to the reservoir with energy $E_2 = E - E_1$ and particle number $N_2 = N - N_1$:

$$\rho_{gc}^{(1)}(E_1, N_1) \propto \omega_2(E - E_1, N - N_1).$$

To proceed, we express the reservoir's density of states in terms of its entropy:¹

$$S_2(E_2, N_2) = k_B \ln \omega_2(E_2, N_2),$$

and expand it around the equilibrium point (E, N) , assuming the reservoir is much larger than the system. Neglecting higher-order terms, we obtain:

$$S_2(E - E_1, N - N_1) \approx S_2(E, N) - \frac{E_1}{T} + \frac{\mu N_1}{T},$$

where we have used the thermodynamic relations $(\partial S / \partial E)_{N,V} = 1/T$, and $(\partial S / \partial N)_{E,V} = -\mu/T$. Exponentiating this expression yields:

$$\omega_2(E - E_1, N - N_1) \propto e^{-\beta(E_1 - \mu N_1)}.$$

¹This derivation is similar to the one presented for the canonical ensemble. However, we will later introduce an alternative method that is physically instructive for different reasons.

Hence, the grandcanonical probability distribution for the subsystem takes the form:

$$\rho_{gc}(q_i, p_i, N) = \frac{1}{\mathcal{Z}} e^{-\beta(\mathcal{H}(q_i, p_i) - \mu N)},$$

where \mathcal{Z} is the **grand partition function**, defined by the normalization condition:

$$\mathcal{Z} = \sum_{N=0}^{\infty} \int_{\mathcal{M}_S} d\Omega e^{-\beta(\mathcal{H}(q_i, p_i) - \mu N)}. \quad (5.1.1)$$

We can obtain the same results for the grandcanonical ensemble by integrating the *canonical probability distribution of the universe*² with respect to the environment variables:

$$\rho_{gc}^{(1)}(q_i^{(1)}, p_i^{(1)}, N_1) = \int_{\mathcal{M}_2} d\Omega_2 \rho_c(q_i^{(1)}, p_i^{(1)}, q_j^{(2)}, p_j^{(2)}) = \frac{e^{-\beta\mathcal{H}_1}}{Z(T, V, N)} \int_{\mathcal{M}_2} d\Omega_2 e^{-\beta\mathcal{H}_2},$$

where $d\Omega_2 = \frac{1}{h^{dN_2} \xi_{N_2}} \prod_{j=1}^{N_2} d^d q_j^{(2)} d^d p_j^{(2)}$ is the measure on the phase space of the environment.

Because the subsystem can exchange both energy and particles with the environment, the normalization condition for the total probability requires integrating over all subsystem variables and summing over all possible particle numbers N_1 :

$$\sum_{N_1=0}^{\infty} \int_{\mathcal{M}_1} d\Omega_1 \rho_{gc}^{(1)}(q_i^{(1)}, p_i^{(1)}, N_1) = 1.$$

Let us check that this normalization indeed holds. Substituting the explicit form of $\rho_{gc}^{(1)}$, we have:

$$\begin{aligned} \sum_{N_1=0}^{\infty} \int_{\mathcal{M}_1} d\Omega_1 \rho_{gc}^{(1)}(q_i^{(1)}, p_i^{(1)}, N_1) &= \sum_{N_1=0}^{\infty} \int_{\mathcal{M}_1} d\Omega_1 \frac{e^{-\beta\mathcal{H}_1}}{Z(T, V, N)} \int_{\mathcal{M}_2} d\Omega_2 e^{-\beta\mathcal{H}_2} \\ &= \sum_{N_1=0}^{\infty} \frac{\int_{\mathcal{M}_1} d\Omega_1 e^{-\beta\mathcal{H}_1} \int_{\mathcal{M}_2} d\Omega_2 e^{-\beta\mathcal{H}_2}}{\int_{\mathcal{M}} d\Omega e^{-\beta(\mathcal{H}_1 + \mathcal{H}_2)}}. \end{aligned}$$

Introducing the explicit expressions for the phase-space measures, we can rewrite this ratio as:

$$\sum_{N_1=0}^{\infty} \frac{\xi_N}{\xi_{N_1} \xi_{N_2}} \frac{\int_{V_1} \prod_{i=1}^{N_1} d^d q_i^{(1)} \int_{V_2} \prod_{j=1}^{N_2} d^d q_j^{(2)} \Phi_1 \Phi_2}{\int_V \prod_{i=1}^{N_1} d^d q_i^{(1)} \int_V \prod_{j=1}^{N_2} d^d q_j^{(2)} \Phi},$$

where we have defined:

$$\begin{aligned} \Phi_1 &= \int_{\mathbb{R}^{dN_1}} \prod_{i=1}^{N_1} d^d p_i^{(1)} e^{-\beta\mathcal{H}_1}, \\ \Phi_2 &= \int_{\mathbb{R}^{dN_2}} \prod_{j=1}^{N_2} d^d p_j^{(2)} e^{-\beta\mathcal{H}_2}, \\ \Phi &= \int_{\mathbb{R}^{dN}} \prod_{i=1}^{N_1} d^d p_i^{(1)} \prod_{j=1}^{N_2} d^d p_j^{(2)} e^{-\beta(\mathcal{H}_1 + \mathcal{H}_2)} = \Phi_1 \Phi_2, \end{aligned}$$

where the last equality follows from the statistical independence of subsystem and environment.

²Although the universe as a whole is strictly microcanonical, it is common to derive the grandcanonical ensemble by starting from a canonical distribution and integrating out the environmental degrees of freedom. This procedure, though approximate, becomes exact in the thermodynamic limit, where ensemble equivalence guarantees that the subsystem distribution coincides with the correct marginal of the global microcanonical ensemble.

To simplify the spatial integrations, let us introduce the notation for the average of a function over a dN -dimensional configuration space of volume V :

$$\langle \Phi \rangle_V := \frac{1}{V^N} \int_V \prod_{i=1}^N d^d q_i \Phi.$$

In the thermodynamic limit ($V \rightarrow \infty$ at fixed density), we assume that such averages become independent of the specific integration domain, provided its surface-to-volume ratio vanishes.³ Hence,

$$\sum_{N_1=0}^{\infty} \int_{\mathcal{M}_1} d\Omega_1 \rho_{gc}^{(1)}(q_i^{(1)}, p_i^{(1)}, N_1) = \sum_{N_1=0}^{\infty} \frac{\xi_N}{\xi_{N_1} \xi_{N_2}} \frac{\langle \Phi_1 \rangle_{V_1} \langle \Phi_2 \rangle_{V_2}}{\langle \Phi \rangle_V} \frac{V_1^{N_1} V_2^{N_2}}{V^N}.$$

For indistinguishable particles, this becomes:

$$\sum_{N_1=0}^{\infty} \frac{N!}{N_1!(N-N_1)!} \left(\frac{V_1}{V}\right)^{N_1} \left(\frac{V-V_1}{V}\right)^{N-N_1},$$

which is a binomial probability distribution. In the thermodynamic limit, using $(V_1 + V_2)/V \rightarrow 1$, the sum converges to unity, proving the normalization of the derived distribution.

We can thus write the **grandcanonical probability density** for the subsystem as:

$$\rho_{gc}^{(1)}(q_i^{(1)}, p_i^{(1)}, N_1) = e^{-\beta \mathcal{H}_1} \frac{Z^{(2)}(T, V - V_1, N - N_1)}{Z(T, V, N)}.$$

To make this expression explicit, we note that the ratio of the two partition functions can be Taylor-expanded around the equilibrium values of the environment variables. Using $Z = e^{-\beta F}$ and expanding the free energy $F(T, V, N)$ for small variations of V_1 and N_1 , we find:

$$\begin{aligned} \frac{Z^{(2)}(T, V - V_1, N - N_1)}{Z(T, V, N)} &= e^{-\beta[F(T, V - V_1, N - N_1) - F(T, V, N)]} \\ &\simeq e^{-\beta \left[-\left(\frac{\partial F}{\partial N}\right)_{V,T} N_1 - \left(\frac{\partial F}{\partial V}\right)_{N,T} V_1 \right]} \\ &= e^{-\beta(-\mu N_1 + p V_1)}, \end{aligned}$$

where we have identified the thermodynamic relations $\mu = (\partial F / \partial N)_{V,T}$ and $p = -(\partial F / \partial V)_{N,T}$.

Substituting this back, we obtain the final expression for the **grandcanonical probability density function**:⁴

$$\rho_{gc}(q_i, p_i, N) = e^{-\beta \mathcal{H}(q_i, p_i)} e^{-\beta(-\mu N + p V)},$$

with the normalization condition:

$$\sum_{N=0}^{\infty} \int_{\mathcal{M}_N} d\Omega e^{-\beta \mathcal{H}} e^{-\beta(-\mu N + p V)} = 1.$$

We can now identify the corresponding **grandcanonical partition function**. Introducing the *fugacity* $z = e^{\beta \mu}$, the normalization condition reads:

$$e^{-\beta p V} \sum_{N=0}^{\infty} z^N \int_{\mathcal{M}_N} d\Omega e^{-\beta \mathcal{H}} = 1.$$

³A rigorous justification of this statement would require ergodicity assumptions and a detailed measure-theoretic analysis, which are beyond our present scope.

⁴At this point we rename subsystem variables as global ones, since all explicit dependence on the environmental quantities has been eliminated.

Recognizing the canonical partition function in the integral, we find:

$$e^{-\beta pV} \sum_{N=0}^{\infty} z^N Z_N = 1 \quad \Rightarrow \quad \mathcal{Z} = \sum_{N=0}^{\infty} z^N Z_N = e^{\beta pV} = e^{-\beta \Omega},$$

where $\Omega = -pV$ is the *grand potential*. Finally, the normalized probability density becomes:

$$\rho_{gc}(q_i, p_i, N) = e^{-\beta \mathcal{H}(q_i, p_i)} \frac{z^N}{\mathcal{Z}}. \quad (5.1.2)$$

Finally, we define the ensemble average of an observable in the grandcanonical ensemble.

Definition 5.1. Grandcanonical average. Let $f(q_i, p_i, N)$ be an observable defined on the extended phase space of the system. Its average over the grandcanonical ensemble is given by

$$\begin{aligned} \langle f(q_i, p_i, N) \rangle_{gc} &= \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} \frac{1}{\xi_N} \int \prod_{i=1}^N \frac{d^d q_i d^d p_i}{h^d} e^{-\beta(\mathcal{H}(q_i, p_i) - \mu N)} f(q_i, p_i, N) \\ &= \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N Z_N \langle f(q_i, p_i, N) \rangle_c. \end{aligned} \quad (5.1.3)$$

This defines the connection between microscopic dynamics and macroscopic observables in open systems.

5.2 | Partition Function and Thermodynamic Quantities

As in the canonical ensemble, the **grand partition function** contains all the thermodynamic information about the system. Once \mathcal{Z} is known, all macroscopic quantities can be derived from it. The central thermodynamic potential of the grandcanonical ensemble is the **grand potential**:

$$\Omega(T, V, \mu) = -k_B T \log \mathcal{Z}. \quad (5.2.1)$$

It plays a role analogous to the Helmholtz free energy in the canonical ensemble, with natural variables (T, V, μ) instead of (T, V, N) .

From Ω we can obtain all other quantities of interest through the relations:

$$p = -\left(\frac{\partial \Omega}{\partial V}\right)_{T, \mu}, \quad \mathcal{N} = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T, V}, \quad S = -\left(\frac{\partial \Omega}{\partial T}\right)_{V, \mu},$$

which can easily be derived from the thermodynamic identity:

$$d\Omega = -S dT - p dV - \mathcal{N} d\mu.$$

However, it is instructive to re-derive these expressions directly from statistical averages, to make explicit the connection between ensemble averages and thermodynamic observables.

The average energy is the grandcanonical expectation value of the Hamiltonian. Formally, this is analogous to the canonical ensemble, except that the sum runs over all possible particle numbers:

$$E = \langle \mathcal{H}(q_i, p_i, N) \rangle_{gc} = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \int d\Omega e^{-\beta \mathcal{H}} \mathcal{H}.$$

To evaluate this, notice that the Hamiltonian appears only in the exponential. We can therefore express it as a derivative with respect to β :

$$\mathcal{H} e^{-\beta \mathcal{H}} = -\frac{\partial}{\partial \beta} (e^{-\beta \mathcal{H}}),$$

which yields:

$$E = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \left(-\frac{\partial}{\partial \beta} \right) \int d\Omega e^{-\beta \mathcal{H}} = \frac{1}{\mathcal{Z}} \left(-\frac{\partial}{\partial \beta} \right) \sum_{N=0}^{\infty} z^N \int d\Omega e^{-\beta \mathcal{H}} \Big|_z,$$

thus leading to the final expression:

$$E = -\frac{1}{\mathcal{Z}} \left(\frac{\partial \mathcal{Z}}{\partial \beta} \right)_z = -\left(\frac{\partial \ln \mathcal{Z}}{\partial \beta} \right)_z. \quad (5.2.2)$$

The derivative is taken at fixed fugacity $z = e^{\beta \mu}$. This means that when β varies, the chemical potential μ must adjust so as to keep z constant. Physically, this corresponds to varying the temperature while maintaining the same average particle number.

The average number of particles $\mathcal{N} = \langle N \rangle_{gc}$ can be derived in a completely analogous manner:

$$\mathcal{N} = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \int d\Omega e^{-\beta \mathcal{H}} N = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} N z^N Z_N.$$

Recognizing that differentiation with respect to z brings down a factor of N , we obtain:

$$\mathcal{N} = \frac{1}{\mathcal{Z}} \left(z \frac{\partial}{\partial z} \sum_{N=0}^{\infty} z^N Z_N \right)_\beta = z \left(\frac{\partial \ln \mathcal{Z}}{\partial z} \right)_\beta. \quad (5.2.3)$$

This formula explicitly relates the average number of particles to the dependence of the partition function on the fugacity, illustrating that μ controls the probability distribution over different particle numbers.

Entropy. Finally, we show that the statistical entropy in the grandcanonical ensemble coincides with the thermodynamic entropy, as it did for the microcanonical and canonical ensembles in the thermodynamic limit. Using Boltzmann's universal definition of entropy:

$$S_{gc} = -k_B \langle \log \rho_{gc} \rangle_{gc},$$

and substituting the expression for the grandcanonical probability density $\rho_{gc} = e^{-\beta(\mathcal{H}-\mu N)}/\mathcal{Z}$, we find:

$$\begin{aligned} S_{gc} &= -\frac{k_B}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \int d\Omega e^{-\beta\mathcal{H}} \log \left(\frac{e^{-\beta(\mathcal{H}-\mu N)}}{\mathcal{Z}} \right) \\ &= -\frac{k_B}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \int d\Omega e^{-\beta\mathcal{H}} (-\beta\mathcal{H} + \beta\mu N - \log \mathcal{Z}). \end{aligned}$$

Performing the integrations and recognizing the ensemble averages, this becomes:

$$S_{gc} = -k_B [-\beta E + \beta\mu\mathcal{N} - \log \mathcal{Z}].$$

Recalling that $\Omega = -k_B T \log \mathcal{Z}$, we can rewrite this as:

$$S_{gc} = \frac{1}{T} (E - \mu\mathcal{N} - \Omega).$$

This is precisely the **thermodynamic identity** linking entropy, energy, chemical potential, and the grand potential:

$$S_{th} = \frac{E - \mu\mathcal{N} - \Omega}{T}.$$

Hence, we have proven that:

$$S_{gc} = S_{th},$$

that is, in the thermodynamic limit, the entropy computed from the grandcanonical ensemble coincides with the macroscopic thermodynamic entropy.

Remark. This result completes the chain of equivalences among entropy ensembles. It confirms that all three statistical ensembles yield the same thermodynamics when extensive variables are large and fluctuations are negligible.

Now that we have demonstrated the entropy equivalence in all ensembles, we can state the universal Boltzmann relation between entropy and the averages of each ensemble:

Theorem 5.1 (Boltzmann universal relation.). *In the thermodynamic limit, the entropy of a system is the same in all statistical ensembles and is given by:*

$$S = -k_B \langle \log \rho_{ens} \rangle_{ens}, \tag{5.2.4}$$

where ρ_{ens} is the probability density function of the chosen ensemble (microcanonical, canonical, or grandcanonical), and $\langle \cdot \rangle_{ens}$ denotes the corresponding ensemble average. The entropy computed via this relation coincides with the thermodynamic entropy S_{th} :

$$S_{mc} \longrightarrow S_c \longrightarrow S_{gc} = S_{th}.$$

5.3 | Virial Expansion

The virial expansion provides a systematic framework to describe the thermodynamic properties of real gases in terms of either their density or fugacity. It allows us to express quantities such as the pressure or the average particle number as **power series** that account for the effects of interparticle interactions.

In the **dilute limit**, the gas behaves as an ideal gas, while higher-order terms in the virial expansion capture *corrections due to correlations* among pairs, triplets, and larger groups of particles. The basic idea is to expand thermodynamic quantities in powers of the density $n = \mathcal{N}/V$ or the fugacity $z = e^{\beta\mu}$, obtaining a series representation that systematically describes deviations from ideal gas behavior.

In what follows, we will use this expansion to derive an equation of state for weakly interacting classical gases. We begin with the expansion in fugacity z , and later we will recover the equivalent expansion in the density n :

$$\begin{aligned}\Omega = -pV &= -\frac{1}{\beta} \log \mathcal{Z} \implies \frac{p}{k_B T} = \frac{1}{V} \log \mathcal{Z}, \\ \mathcal{N} &\implies n = \frac{\mathcal{N}}{V} = \frac{z}{V} \left(\frac{\partial \log \mathcal{Z}}{\partial z} \right)_\beta.\end{aligned}$$

5.3.1 | Fugacity Expansion of the Grand Partition Function

We start from the grand canonical partition function:

$$\mathcal{Z}(T, V, \mu) = \sum_{N=0}^{\infty} z^N Z_N(T, V),$$

where $Z_N(T, V)$ is the canonical partition function for N particles, and $z = e^{\beta\mu}$ is the *fugacity*.

We consider a classical gas of interacting particles described by the Hamiltonian

$$\mathcal{H}(q_i, p_i) = \sum_i^N \frac{p_i^2}{2m} + \sum_{i < j}^N U_{ij}, \quad U_{ij} = U(|q_i - q_j|),$$

where the second term accounts for pairwise interactions through the potential U_{ij} . We assume that the interaction depends only on the distance between particles, i.e., it is isotropic, and that the gas is dilute and weakly interacting. In this regime, interactions can be treated perturbatively, allowing us to expand the grand canonical partition function in powers of the fugacity z .

We can then compute the first few canonical partition functions:

$$\begin{aligned}Z_0 &= 1, \\ Z_1 &= \int_V d\Omega_1 e^{-\beta\mathcal{H}_1} = \frac{V}{\lambda^3}, \quad \lambda = \frac{h}{\sqrt{2\pi m k_B T}}, \\ Z_2 &= \int_V d\Omega_2 e^{-\beta\mathcal{H}_2} = \frac{1}{2!\lambda^6} \int_V d^3r_1 d^3r_2 e^{-\beta U(r)} \\ &= \frac{1}{2\lambda^6} \int_V d^3r_{\text{cm}} \int_V d^3r e^{-\beta U(r)} = \frac{V}{2\lambda^6} \int_V d^3r e^{-\beta U(r)}.\end{aligned}$$

Here, $\mathcal{H}_1 = \frac{p_1^2}{2m}$ contains no interaction term, while $\mathcal{H}_2 = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + U_{12}$, and recall the phase space measure for N particles, as defined in (2.2.1), is given by:

$$d\Omega_N = \frac{1}{h^{3N} N!} \prod_{i=1}^N d^3 q_i d^3 p_i.$$

In the last step of the Z_2 calculation, we switched to center-of-mass and relative coordinates, $r_{\text{cm}} = \frac{r_1+r_2}{2}$ and $r = r_1 - r_2$, so that the integral over r_{cm} simply gives a factor V .

Thus, the grand partition function up to second order in z becomes:

$$\mathcal{Z} = 1 + z \frac{V}{\lambda^3} + z^2 \frac{V}{2\lambda^6} \int_V d^3 r e^{-\beta U(r)} + O(z^3).$$

It is often convenient to express the logarithm of the grand partition function as a power series in z , since all thermodynamic quantities can be derived from it. Recalling that for a small parameter x ,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots,$$

we apply this expansion to \mathcal{Z} and neglect terms beyond second order in z :

$$\begin{aligned} \log \mathcal{Z} &= \frac{zV}{\lambda^3} + \frac{z^2 V}{2\lambda^6} \int_V d^3 r e^{-\beta U(r)} - \frac{1}{2} \frac{z^2 V^2}{\lambda^6} + O(z^3) \\ &= \frac{zV}{\lambda^3} + \frac{z^2 V}{2\lambda^6} \left(\int_V d^3 r e^{-\beta U(r)} - V \right) + O(z^3) \\ &= \frac{zV}{\lambda^3} + \frac{z^2 V}{2\lambda^6} \left(\int_V d^3 r e^{-\beta U(r)} - \int_V d^3 r \right) + O(z^3) \\ &= \frac{zV}{\lambda^3} + \frac{z^2 V}{2\lambda^6} J_2(\beta) + O(z^3), \end{aligned}$$

where we have defined

$$J_2 = \int d^3 r (e^{-\beta U(r)} - 1) = 4\pi \int_0^\infty dr r^2 (e^{-\beta U(r)} - 1).$$

In general, the logarithm of the grand partition function can be written as a power series in z :

$$\log \mathcal{Z} = V \sum_{\ell=1}^{\infty} b_\ell(T) z^\ell, \quad (5.3.1)$$

where the coefficients $b_\ell(T)$ are known as the **cluster integrals** or **virial coefficients in the fugacity expansion**. They depend only on the temperature and on the interparticle potential.

5.3.2 | Thermodynamic Quantities in the Dilute Limit

From the grand potential $\Omega = -k_B T \log \mathcal{Z}$, the pressure is given by

$$p = -\frac{\Omega}{V} = k_B T \sum_{\ell=1}^{\infty} b_\ell(T) z^\ell = k_B T \left[\frac{zV}{\lambda^3} + \frac{z^2 V}{2\lambda^6} J_2(\beta) + O(z^3) \right].$$

The average number of particles follows from

$$\mathcal{N} = z \frac{\partial \log \mathcal{Z}}{\partial z} \Big|_{\beta} = V \sum_{\ell=1}^{\infty} \ell b_\ell(T) z^\ell = V \left[\frac{zV}{\lambda^3} + z^2 \frac{V}{\lambda^6} J_2(\beta) + O(z^3) \right].$$

Density expansion. We define the number density as

$$n = \frac{\langle N \rangle}{V} = \sum_{\ell=1}^{\infty} \ell b_{\ell}(T) z^{\ell}. \quad (5.3.2)$$

Both pressure and density are therefore expressed as power series in the fugacity z . Working in the limit of small z corresponds to a dilute gas regime, characterized by low density and low pressure. Let us recover the expression for n in the non-interacting case:

$$J_2(\beta) \Big|_{U=0} = \int d^3r (1 - 1) = 0 \implies n = \frac{z}{\lambda^3} + O(z^3),$$

which is the expected result for an ideal classical gas.

Pressure expansion. By inverting the series to express z as a function of n and substituting it into the expression for the pressure, we obtain an expansion of p in powers of the density:

$$\frac{p}{k_B T} = n + B_2(T) n^2 + B_3(T) n^3 + \dots,$$

where the coefficients $B_n(T)$ are the *virial coefficients*.

- $B_2(T)$ accounts for pairwise interactions and is directly related to the intermolecular potential $u(r)$ through

$$B_2(T) = \frac{1}{2} \int (e^{-\beta u(r)} - 1) d^3r = \frac{1}{2} J_2(\beta).$$

Thus, $B_2(T)$ captures the leading-order correction to ideal gas behavior due to two-body correlations.

- $B_3(T)$ describes the effect of three-body correlations, involving integrals over triplets of particles and their mutual interactions, we are not deriving it here.
- Higher-order coefficients $B_n(T)$ for $n \geq 4$ involve increasingly complex integrals over n -particle configurations and account for many-body effects.

Higher-order virial coefficients describe increasingly subtle correlation effects and become significant at higher densities or for stronger interactions. In summary, the virial expansion establishes a direct connection between the microscopic interaction potential and the macroscopic equation of state:

$$\frac{p}{k_B T n} = 1 + B_2(T)n + B_3(T)n^2 + \dots$$

In the case of non-interacting particles, all virial coefficients beyond the first vanish, and we recover the ideal gas law:

$$\begin{aligned} \frac{p}{k_B T} &= \left[\frac{zV}{\lambda^3} + \frac{z^2 V}{2\lambda^6} J_2(\beta) + O(z^3) \right], \\ J_2(\beta) \Big|_{U=0} &= \int d^3r (1 - 1) = 0, \\ \implies \frac{p}{k_B T} &= \frac{zV}{\lambda^3} = n, \end{aligned}$$

thus recovering the ideal gas equation of state.

Remark. The virial expansion is particularly useful for gases at low to moderate densities, where truncating the series after a few terms yields an accurate equation of state. At high densities or near phase transitions, however, the series may fail to converge, and more sophisticated approaches are required.

Fugacity expansion in the density. Since we have expressions for both pressure and density as power series in fugacity z , we can combine them to express pressure directly as a function of density: the idea is to invert the density expansion to express z in terms of n , and then substitute this into the pressure expansion. To express the fugacity z as a function of the density n , we can invert the series as follows:

$$\begin{aligned} z &= An + Bn^2 + O(n^3), \\ n &= \frac{z}{\lambda^3} + \frac{z^2}{\lambda^6} J_2(\beta) + O(z^3) \\ &= \frac{An + Bn^2}{\lambda^3} + \frac{(An + Bn^2)^2}{\lambda^6} J_2 + O(n^3) \\ &= \frac{An + Bn^2}{\lambda^3} + \frac{A^2 n^2}{\lambda^6} J_2 + O(n^3), \end{aligned}$$

whose only consistent solution is $A = \lambda^3$ and $B = -J_2 \lambda^3$. Hence,

$$z = \lambda^3 n - J_2 \lambda^3 n^2 + O(n^3).$$

Substituting this into the pressure expansion, we recover the virial equation of state for a weakly interacting classical gas:

$$\frac{p}{k_B T} = \frac{z}{\lambda^3} + \frac{z^2}{2\lambda^6} J_2(\beta) + O(z^3) = n - \frac{J_2}{2} n^2 + O(n^3),$$

which again reduces to the ideal gas law when $U = 0$.

5.3.3 | Van der Waals Equation of State

To connect the virial expansion with a more familiar and physically intuitive equation of state, we will now consider a simple yet effective model for the intermolecular potential $U(r)$. A convenient choice is the **hard-sphere potential** with an additional attractive tail. In this model, molecules are treated as rigid spheres of diameter σ , meaning they cannot overlap; this represents short-range repulsion force, due to the finite size of particles.

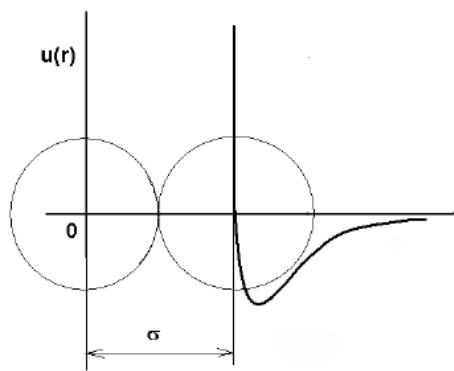


Figure 5.1: Hard-sphere potential with attractive tail, where particles cannot approach closer than a distance σ .

At larger separations, an attractive potential $f(r) < 0$ is introduced to mimic cohesive forces such as van der Waals or dispersion interactions:

$$U(r) = \begin{cases} \infty & \text{if } r < \sigma, \\ f(r) & \text{if } r \geq \sigma, \quad \text{with } f(r) < 0. \end{cases}$$

This potential effectively separates the microscopic interaction into two competing contributions:

- a **repulsive core**, enforcing an excluded volume around each particle;
- an **attractive tail**, lowering the energy when particles are moderately close.

Using this potential, we can compute the second cluster integral, $J_2(\beta)$, which directly encodes the

two-body correlations induced by the interaction:

$$\begin{aligned} J_2(\beta) &= 4\pi \int_0^\infty dr r^2 (e^{-\beta U(r)} - 1) \\ &= 4\pi \left[\int_0^\sigma dr r^2 (0 - 1) + \int_\sigma^\infty dr r^2 (e^{-\beta f(r)} - 1) \right] \\ &= -\frac{4\pi\sigma^3}{3} + 4\pi \int_\sigma^\infty dr r^2 (e^{-\beta f(r)} - 1) \\ &\doteq -2b + 2\beta a, \end{aligned}$$

where $a, b > 0$ are constants that characterize the potential: after expanding the exponential for weak interactions ($\beta|f(r)| \ll 1$), they are given by

$$b = \frac{2\pi}{3}\sigma^3, \quad a = 2\pi \int_\sigma^\infty dr r^2 |f(r)|.$$

The first term represents the *excluded volume* due to hard-sphere repulsion being negative, while the second captures the *attractive tail*, which is positive since $f(r) < 0$. At high temperatures, thermal motion dominates and the attractive contribution becomes negligible; at low temperatures, attractive interactions prevail. This competition explains why, at intermediate temperatures, the effective interaction can change sign.

Substituting this result into the virial expansion of the pressure yields:

$$\begin{aligned} \frac{p}{k_B T} &= n - (-b + \beta a)n^2 + O(n^3) \\ &= \frac{1}{v} + \left(b - \frac{a}{k_B T} \right) \frac{1}{v^2} + O\left(\frac{1}{v^3}\right), \end{aligned}$$

from which we can rearrange to obtain the celebrated **Van der Waals equation of state**:

$$\left(p + \frac{a}{v^2} \right) (v - b) = k_B T. \quad (5.3.3)$$

This equation provides a phenomenological correction to the ideal gas law, capturing finite particle size and intermolecular attractions within a simple yet remarkably accurate model for real gases.

6 | Phase Transitions

Phase transitions are among the most striking manifestations of collective behavior in statistical physics. They represent abrupt qualitative changes in the macroscopic state of matter that arise from the smooth variation of a few external control parameters, such as temperature, pressure, or magnetic field. At the microscopic level, these transitions emerge from the cooperative dynamics of a very large number of interacting degrees of freedom, leading to singularities in thermodynamic quantities and to the spontaneous breaking of symmetries.

In this chapter, we develop the thermodynamic and statistical foundations needed to describe and classify phase transitions. We begin by introducing the concept of a *phase diagram*, which organizes the possible equilibrium states of a system and identifies coexistence regions where two or more phases can be in equilibrium. We then derive the *Clausius–Clapeyron equation*, which provides a quantitative description of coexistence curves and the relation between intensive thermodynamic variables along phase boundaries.

Next, we discuss the *classification of phase transitions* according to the nature of the non-analyticity of the free energy, distinguishing between first-order transitions, characterized by discontinuities in first derivatives (such as latent heat or density jumps), and continuous or higher-order transitions, where thermodynamic quantities remain continuous but their derivatives diverge. This classification naturally leads to the study of critical phenomena and universality.

Finally, to illustrate how these ideas emerge from microscopic models, we consider the paradigmatic example of the *Ising model*. Despite its simplicity, this lattice model captures the essential features of cooperative ordering and provides a concrete realization of a continuous phase transition, allowing us to connect the abstract thermodynamic description with a statistical-mechanical one.

Throughout this chapter, our aim is to bridge the phenomenological thermodynamics of phase transitions with their microscopic statistical interpretation, highlighting how non-analytic behavior in macroscopic observables originates from the collective organization of microscopic states.

6.1 | Phase Diagrams

A classical fluid can exist in different phases—solid, liquid, and gas—depending on its temperature T and pressure p . In the $p - T$ diagram, these phases occupy distinct regions separated by phase boundaries, along which two phases can coexist in equilibrium.

Coexistence lines and triple point. In a phase diagram, the *coexistence lines* (or phase boundaries) represent the set of thermodynamic states where two phases of the same substance can coexist in equilibrium. Along each line, the temperature, pressure, and chemical potential of the two phases are equal,

$$T_1 = T_2, \quad p_1 = p_2, \quad \mu_1 = \mu_2,$$

ensuring that the transformation can occur reversibly under constant thermodynamic conditions. These three coexistence lines meet at a unique point in the $p-T$ plane called the *triple point*, where solid, liquid, and vapor coexist simultaneously in thermodynamic equilibrium. At this point, the equality of temperature, pressure, and chemical potential extends to all three phases,

$$T_s = T_l = T_v, \quad p_s = p_l = p_v, \quad \mu_s = \mu_l = \mu_v,$$

so that mass can be exchanged among them without altering the macroscopic state of the system. The triple point thus defines a fundamental reference in thermodynamics, often used for the calibration of temperature and pressure scales.

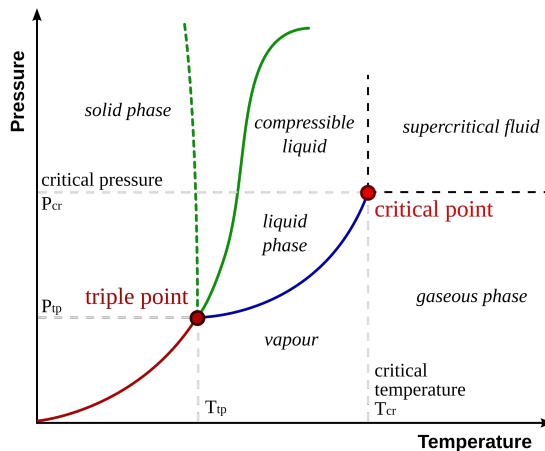


Figure 6.1: A schematic phase diagram showing different phases of a classical fluid, with regions of solid, liquid, and gas separated by phase boundaries. Critical points and triple points are also indicated, along with the coexistence curves.

The solid–liquid line marks the conditions of melting and freezing, while the liquid–vapor line corresponds to vaporization and condensation. The solid–vapor line instead represents sublimation and deposition processes. All three boundaries meet at the triple point, where the three phases coexist simultaneously in thermodynamic equilibrium. The liquid–vapor coexistence curve terminates at the critical point, beyond which the distinction between liquid and gas disappears and the substance forms a supercritical fluid. Each phase transition involves a discontinuous change in certain thermodynamic properties, such as density or enthalpy, and can be associated with a latent heat exchange between the phases.

Critical point and superfluidity. The *critical point* marks the upper end of the liquid–vapor coexistence curve in the $p-T$ diagram. At this point, the distinction between liquid and gas disappears: the densities of the two phases become equal, and the surface tension vanishes. Beyond the critical point, the substance forms a **supercritical fluid**, a single homogeneous phase that exhibits properties intermediate between those of a liquid and a gas, such as high diffusivity combined

with a relatively large density. The system does not dissipate energy from friction as a normal fluid would, allowing it to flow without viscosity. The critical point is characterized by continuous (second-order) phase transition behavior, where response functions such as compressibility and heat capacity diverge according to universal critical exponents.

In certain quantum fluids, such as liquid helium-4, a different kind of continuous phase transition occurs at very low temperatures, leading to the onset of **superfluidity**.

Below the so called λ -point (at $T_\lambda \approx 2.17$ K under saturated vapor pressure), helium-4 undergoes a transition from the normal liquid phase (He I) to the superfluid phase (He II). The superfluid phase exhibits remarkable macroscopic quantum phenomena, including the absence of viscosity, the ability to flow through extremely narrow capillaries, and the quantization of vorticity. Although both the critical point and the superfluid transition correspond to continuous changes in the state of the system, they originate from different microscopic mechanisms: the former arises from the disappearance of density fluctuations at large scales, while the latter emerges from Bose-Einstein condensation of helium atoms.

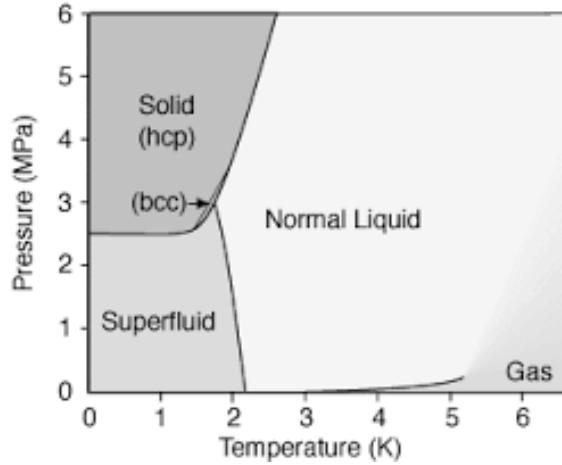


Figure 6.2: Schematic phase diagram of liquid helium-4, showing solid, liquid, gas, and superfluid regions separated by phase boundaries. Coexistence curves are indicated where appropriate, while the boundary between the liquid and superfluid phases is shown for reference, even though this transition is continuous.

6.1.1 | Clausius–Clapeyron Equation

Consider a system at equilibrium along a coexistence line between two phases, denoted α and β . The Gibbs free energy of the system is

$$G(T, p) = \mu N, \quad (6.1.1)$$

where μ is the chemical potential and N the number of particles. Per particle, the Gibbs free energy is

$$g = \frac{G}{N} = \mu.$$

At equilibrium along the coexistence line, the chemical potentials of the two phases are equal:

$$\mu_\alpha(T, p) = \mu_\beta(T, p) \implies g_\alpha(T, p) = g_\beta(T, p).$$

Differentiating both sides along the coexistence line gives

$$d\mu_\alpha = d\mu_\beta,$$

which, using the thermodynamic identity for the Gibbs free energy per particle¹

$$d\mu = -s dT + v dp,$$

¹ $dg = d\mu = -s dT + v dp$ since $dG = \mu dN + N d\mu$, leading to the deletions which yield $dg = d\mu$.

where s and v are the molar entropy and molar volume, respectively, we obtain

$$-s_\alpha dT + v_\alpha dp = -s_\beta dT + v_\beta dp.$$

Now we can differentiate with respect to temperature:

$$\frac{d\mu}{dT} = \left(\frac{\partial\mu}{\partial T}\right)_p + \left(\frac{\partial\mu}{\partial p}\right)_T \frac{dp}{dT} = -s + v \frac{dp}{dT},$$

and rewriting this, the slope of the coexistence curve in the p - T plane is

$$\frac{dp}{dT} = \frac{s_\beta - s_\alpha}{v_\beta - v_\alpha} = \frac{\Delta s}{\Delta v}.$$

Introducing the **latent heat** q per mole, which is the heat absorbed or released during the phase transition at constant temperature and pressure, we have

$$q = T \Delta s. \quad (6.1.2)$$

Substituting this relation into the previous expression, we obtain the **Clausius–Clapeyron equation**:

$$\frac{dp}{dT} = \frac{q}{T \Delta v}. \quad (6.1.3)$$

This equation provides a fundamental link between measurable thermodynamic quantities along the coexistence line: the slope of the phase boundary is determined by the latent heat of the transition and the change in molar volume between the two phases.

6.2 | Classification of Phase Transitions

Phase transitions can be broadly divided into two main categories: *first-order* and *continuous* (or *second-order*) transitions, depending on the nature of the discontinuities that appear in the thermodynamic description of the system.

Historically, phase transitions were classified according to the *Ehrenfest scheme*, which identified the order of the transition with the lowest derivative of the Gibbs free energy that exhibits a discontinuity. Although this classification remains conceptually useful, it has been largely superseded by a more general framework developed in modern statistical mechanics, which highlights critical behavior and universality near continuous transitions.

6.2.1 | First-Order Phase Transitions

First-order phase transitions are characterized by a discontinuous change in the first derivatives of the free energy with respect to the thermodynamic variables. In particular, quantities such as volume or entropy experience finite jumps at the transition point. These discontinuities imply the exchange of a finite amount of heat at constant temperature and pressure, known as the *latent heat* of the transition:

$$q = T \Delta s,$$

where Δs is the entropy difference between the two coexisting phases, as defined in (6.1.2). First-order transitions are accompanied by the coexistence of phases and occur along well-defined lines in the p - T plane. Typical examples include the melting of a solid, the vaporization of a liquid, and the sublimation of a solid into vapor.

6.2.2 | Continuous Phase Transitions

Continuous (or second-order) phase transitions are characterized by the continuity of the first derivatives of the free energy, while higher-order derivatives become discontinuous or diverge. No latent heat is involved, and the transition occurs smoothly, often associated with the emergence of long-range correlations and critical fluctuations. Near the critical point, the system exhibits *scale invariance* and *universal behavior*, where macroscopic properties depend only on general features such as symmetry and dimensionality, rather than on microscopic details. Examples include the liquid-gas critical point and the superfluid transition in helium-4.

Remark. *The presence or absence of latent heat provides a clear thermodynamic signature distinguishing first-order from continuous phase transitions. In first-order transitions, the system absorbs or releases a finite amount of energy to reorganize its microscopic structure, while in continuous transitions, this reorganization occurs gradually, with thermodynamic quantities varying smoothly but response functions (such as heat capacity or compressibility) diverging near the critical point.*

6.3 | Discontinuities and Phase Transitions

Phase transitions are intimately connected with the analytic structure of the partition function. For finite systems, $\mathcal{Z}_N(\lambda)$ is an analytic function of its parameters, so true singularities cannot occur. However, the distribution of its zeros in the complex plane reveals where non-analytic behavior may emerge in the thermodynamic limit.

This analytic mechanism was first formalized by Lee and Yang in a pair of theorems that establish both the existence and the analyticity properties of the thermodynamic potential in the complex plane of the control parameter.

Theorem 6.1 (Lee–Yang I | Existence of the thermodynamic limit). *Consider a system of N interacting particles with potential energy U_N satisfying the stability bound*

$$U_N \geq -KN, \quad K > 0,$$

and contained in a volume V whose boundary increases no faster than $V^{2/3}$. Then, for any complex fugacity $z = e^{\beta\mu}$, the limit

$$\Psi(z, v, T) = \lim_{V \rightarrow \infty} \frac{1}{V} \log \mathcal{Z}(z, v, T)$$

exists and defines a continuous, monotonically increasing function of $z \in \mathbb{R}^+$.

Theorem 6.2 (Lee–Yang II | Analyticity and distribution of zeros). *Let R be an open subset of the complex z -plane containing a portion of the positive real axis. If, for all finite V and T , the grand partition function $\mathcal{Z}(z, v, T)$ has no zeros in R , then the sequence $(\log \mathcal{Z}(z, v, T))/V$ converges uniformly as $V \rightarrow \infty$ in any closed subset of R , and therefore the thermodynamic potential $\Psi(z, v, T)$ is analytic in R .*

The two results together imply that a phase transition can only occur if, in the thermodynamic limit, a locus of zeros of the finite-volume partition function $\mathcal{Z}(z, V, T)$ approaches the positive real axis. The point z_c where this accumulation occurs divides the real axis into distinct analytic regions, corresponding to different thermodynamic phases. At $z = z_c$, the potential $\Psi(z, v, T)$ remains continuous but becomes non-analytic: a cusp singularity corresponds to a first-order phase transition, while higher-order singularities describe continuous phase transitions. If we have regions where the zeroes of \mathcal{Z} do not accumulate near the real axis then there will be no PT in that region.

Remark. *The celebrated Lee–Yang circle theorem is a special case of this general framework, proved for the ferromagnetic Ising model, where all zeros of the partition function in the complex fugacity plane $z = e^{-2\beta h}$ lie on the unit circle $|z| = 1$. In more general systems, such as lattice gases or fluids with realistic interactions, the zeros no longer lie on a perfect circle, yet their accumulation and contact with the real axis remain the universal analytic mechanism underlying phase transitions.*

To illustrate the qualitative content of the **first Lee–Yang theorem**, let us consider a system described by a mean-field Hamiltonian, for which the grand partition function can be written as a polynomial of finite degree N in the complex fugacity z . Starting from the canonical partition function of a system with Hamiltonian

$$\mathcal{H} = \sum_i \frac{p_i^2}{2m} + U_N(q_1, \dots, q_N),$$

we can integrate over the momenta and separate the purely configurational contribution:

$$Z_N(V, T) = \int_V d\Omega e^{-\beta \mathcal{H}} = \frac{Q_N}{N! \lambda_{\text{th}}^{3N}}, \quad Q_N(V, T) = \int \prod_i d^3 q_i e^{-\beta U_N(q_1, \dots, q_N)}.$$

Here, Q_N encodes the interaction between particles, while the prefactor $1/(N! \lambda_{\text{th}}^{3N})$ represents the contribution of indistinguishability and kinetic energy.

Hence, the grand partition function of the system reads

$$\mathcal{Z}_N(z) = \sum_{n=0}^N \frac{z^n Q_n}{n! \lambda_{\text{th}}^{3n}} = \sum_{n=0}^N a_n z^n, \quad a_n > 0 \ \forall n.$$

At finite N , this is a polynomial in z with positive coefficients. Even before taking the thermodynamic limit, this already implies that the zeros of $\mathcal{Z}_N(z)$ cannot lie on the positive real axis, since all the coefficients of the polynomial are strictly positive. Therefore, for any finite system, the grand partition function is analytic and positive for $z \in \mathbb{R}^+$.

To examine whether the thermodynamic limit exists, we must ensure that the series remains convergent as $N \rightarrow \infty$. We assume, as in the standard proof, that the interaction potential is bounded from below by a function linear in N :

$$U_N \geq -KN, \quad K > 0.$$

This lower bound guarantees the system's stability: the energy cannot decrease faster than linearly with the number of particles. From this condition, it follows that

$$e^{-\beta U_N} \leq e^{\beta KN}.$$

Using this inequality, we can bound the configurational integral:

$$Q_N \leq V^N e^{\beta KN}.$$

Substituting this estimate into the series for $\mathcal{Z}_N(z)$, we find

$$\mathcal{Z}_N(z) \leq \sum_{n=0}^N \frac{1}{n!} \left(\frac{zV e^{\beta K}}{\lambda_{\text{th}}^3} \right)^n \xrightarrow{N \rightarrow \infty} \exp \left(\frac{zV e^{\beta K}}{\lambda_{\text{th}}^3} \right).$$

The thermodynamic limit is therefore well defined: the grand partition function approaches a continuous, positive, monotonically increasing function for all $z \in \mathbb{R}^+$.

In summary, the essential point of the first Lee–Yang theorem is that, under the assumption of thermodynamic stability (i.e., a potential bounded from below), the grand partition function $\mathcal{Z}(z)$ is analytic for positive real fugacity and its thermodynamic limit is well defined.

The **second Lee–Young theorem** help us understand how singularities can appear in the thermodynamic limit: we have to focus on the logarithm of the grand partition function. All thermodynamic quantities, such as pressure, density, or free energy, are obtained from derivatives of

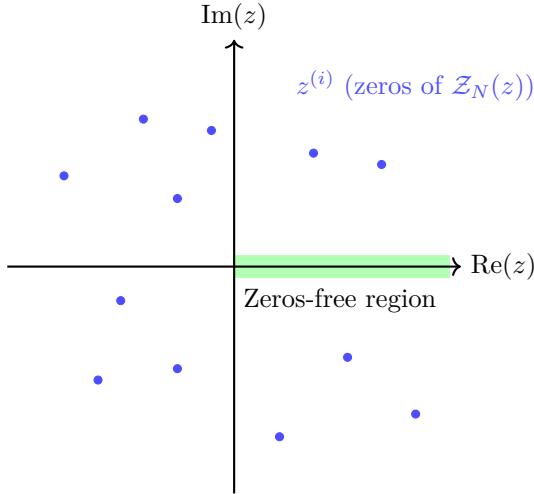
$$\Phi = \log \mathcal{Z} \implies \begin{cases} p = k_B T \Psi(z, v, T), \\ n = \frac{1}{v} = z \frac{\partial}{\partial z} \Psi(z, v, T) \Big|_\beta. \end{cases}$$

Therefore, the zeros of \mathcal{Z} correspond to singularities of Φ , and thus to possible singularities of thermodynamic observables. In other words, any non-analytic behaviour of macroscopic quantities must originate from points in the complex fugacity plane where $\mathcal{Z} = 0$.

From the previous discussion, we know that at finite N the grand partition function

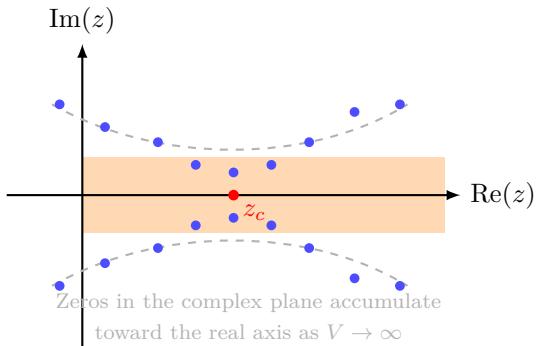
$$\mathcal{Z}_N(z) = \sum_{n=0}^N a_n z^n, \quad a_n > 0,$$

is a polynomial with strictly positive coefficients.



By the fundamental theorem of algebra, it possesses N complex zeros, but since all coefficients are positive, none of these zeros can lie on the positive real axis. As a result, for any finite N , $\mathcal{Z}_N(z)$ is analytic and strictly positive for real positive z , which means that no true phase transition can occur in a system of finite size. The partition function, and hence all thermodynamic functions derived from it, remain smooth and differentiable.

However, the situation changes qualitatively in the thermodynamic limit $N, V \rightarrow \infty$ with fixed density. As N increases, the zeros of $\mathcal{Z}_N(z)$ in the complex z -plane can become denser and may start to accumulate along certain curves. In this limit, the set of zeros can form continuous lines or regions in the complex plane.



If such a line of zeros $z^{(i)}$ approaches the positive real axis, then the logarithm of the partition function, $\Phi(z) = \log \mathcal{Z}(z)$, becomes non-analytic at the point z_c where the accumulation occurs:

$$z^{(i)} \in \mathbb{C} - \mathbb{R}^+, \quad z^{(i)} \xrightarrow{N \rightarrow \infty} z_c \in \mathbb{R}^+.$$

This non-analyticity is precisely what we interpret as a phase transition.

Hence, the second Lee–Yang theorem provides a qualitative explanation for the origin of thermodynamic singularities: the partition function itself is always analytic for finite systems, but as the system size grows, its zeros in the complex fugacity plane can condense and “pinch” the real axis. The point where this happens marks the emergence of a macroscopic phase transition. If we can demonstrate that for an infinite volume the zeroes remain out of the considered region, then there will be no PT in this region.

Remark. *In practice, even though the Lee–Yang theorems show that true non-analyticities can only appear strictly in the thermodynamic limit, real systems with a large but finite number of particles can still exhibit apparent phase transitions. For instance, a small bottle of water can freeze or boil sharply, even though it contains a finite (and thus analytic) number of degrees of freedom.*

freedom. This happens because, for macroscopic N , the zeros of the partition function $\mathcal{Z}_N(z)$ already lie extremely close to the real axis. As a consequence, the derivatives of $\log \mathcal{Z}_N$ can vary very rapidly near those quasi-singular points, producing an abrupt but smooth crossover that is experimentally indistinguishable from a genuine discontinuity.

6.4 | Heisemberg Model

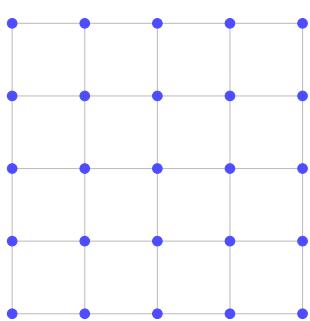
The Ising model represents one of the simplest yet most profound models in statistical mechanics. It captures the essential mechanism behind magnetic ordering, spontaneous symmetry breaking, and the appearance of phase transitions in systems composed of many interacting constituents. Although originally conceived to describe ferromagnetism, its conceptual framework extends to a wide class of collective phenomena, from lattice gases to binary mixtures and neural networks.

In this chapter, we will start by analyzing the simplest possible case: a system of magnetic moments embedded in a crystalline solid lattice, each associated with a magnetic moment, but *without* any interaction between the spins. This will be done with the Heisemberg model, which accounts for $3D$ -spins. This non-interacting model describes a collection of distinguishable atoms with magnetic moments oscillating in a thermal bath, where the total energy is determined solely by their coupling with an external magnetic field. Such a system corresponds to a **paramagnet**, and its behavior can be completely understood within the framework of classical equilibrium statistical mechanics.

We will then introduce interactions between neighboring magnetic moments, which tend to align or anti-align depending on the sign of the coupling constant. It is precisely this introduction of *spin–spin interaction* that gives rise to the emergence of cooperative behavior and, eventually, to a phase transition between the paramagnetic and **ferromagnetic** phases. Thus, the Ising model clearly illustrates how the collective properties of matter — such as magnetization and critical phenomena — arise from the interplay between microscopic interactions and thermal fluctuations.

6.4.1 | Non-interacting Magnetic Spins

We consider a crystalline solid consisting of N atoms arranged on a regular lattice of fixed volume V and temperature T . Each lattice site hosts a single atom, whose center of mass is bound to oscillate around its equilibrium position due to thermal vibrations.



For the purpose of studying magnetic properties, we assume that these translational vibrations are small and can be effectively decoupled from the magnetic degrees of freedom. Each atom carries a *magnetic moment* (or spin) represented by a three-dimensional vector μ_i of fixed magnitude $\mu = |\mu_i|$, free to orient in any direction in space. This is the characteristic of the **Heisemberg model**. Hence, the configuration space associated with the magnetic degrees of freedom of each atom is the surface of a sphere of radius μ .

The Hamiltonian of the system naturally separates into two distinct contributions:

$$\mathcal{H} = \mathcal{H}_{\text{kin}}(\{q_i, p_i\}) + \mathcal{H}_{\text{mag}}(\{\mu_i\}),$$

where \mathcal{H}_{kin} accounts for the kinetic and potential energy of the lattice vibrations, and \mathcal{H}_{mag} describes the interaction of the magnetic moments with external fields and, possibly, with one another. Since our interest lies in the thermodynamic and collective magnetic properties, and the kinetic part merely contributes a multiplicative factor to the partition function, we will focus on the magnetic Hamiltonian \mathcal{H}_{mag} .

Remark (Origin of the atomic magnetic moment). The magnetic moment of an atom originates from the microscopic motion of charged particles within it. In classical terms, a circulating electric charge produces a magnetic dipole moment; in quantum mechanics, this effect arises from two distinct contributions:

1. **Orbital motion of electrons:** each electron moving around the nucleus carries an orbital angular momentum \mathbf{L} , associated with a magnetic moment $\boldsymbol{\mu}_{\text{orb}} = -\frac{e}{2m_e} \mathbf{L}$.
2. **Intrinsic spin:** the electron possesses an intrinsic angular momentum \mathbf{S} , giving rise to a spin magnetic moment $\boldsymbol{\mu}_{\text{spin}} = -g_s \frac{e}{2m_e} \mathbf{S}$, where $g_s \simeq 2.0023$ is the spin g-factor.

The total magnetic moment of an atom is the vector sum of these contributions, weighted by the number and configuration of electrons. In many materials, the magnetic moments of different atoms are randomly oriented, leading to no macroscopic magnetization. However, when an external magnetic field is applied, these moments tend to align with it, giving rise to the paramagnetic response described by the Curie law.

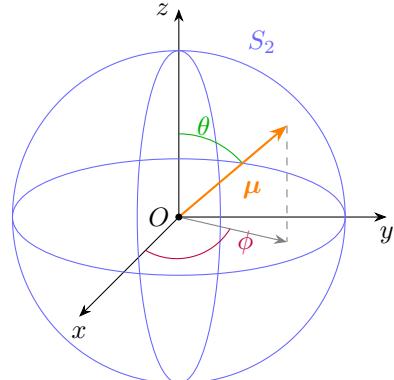
To describe the orientation of a magnetic moment in space, it is convenient to adopt spherical coordinates. We can parametrize each moment as

$$\boldsymbol{\mu}_i = \mu(\sin \theta_i \cos \phi_i, \sin \theta_i \sin \phi_i, \cos \theta_i),$$

where $\phi_i \in [0, 2\pi)$ is the azimuthal angle and $\theta_i \in [0, \pi]$ the polar angle.

Within this framework, the pair $(\phi_i, \cos \theta_i)$ can be regarded as a set of generalized canonical variables (q_i, p_i) for the rotational degrees of freedom of each magnetic moment, since the conjugate momentum to the coordinate ϕ_i is proportional to $\cos \theta_i$. This representation allows for a compact and geometrically intuitive formulation of the magnetic phase space as a product of N spheres:

$$\mathcal{M}_{\text{mag}} = \bigotimes^N (S_2).$$



For a lattice of non-interacting magnetic moments, if we consider an external magnetic field $\mathbf{H} = H \hat{\mathbf{z}}$ applied to the system, the magnetic part of the Hamiltonian reads:

$$\mathcal{H}_{\text{mag}}(\{\boldsymbol{\mu}_i\}) = - \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{H}. \quad (6.4.1)$$

The negative sign ensures that the lowest energy configuration corresponds to spins aligned with the field. Since the spins do not interact, the partition function factorizes:

$$Z_N = (Z_1)^N,$$

where the single-spin partition function is

$$\begin{aligned} Z_1 &= \int_{S_2} d\Omega e^{-\beta \mathcal{H}_{\text{mag}}(\boldsymbol{\mu})} = \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi e^{\beta H \mu \cos \theta} \\ &= 2\pi \int_0^\pi d\theta \sin \theta e^{\beta H \mu \cos \theta} = 2\pi \int_{-1}^1 d(\cos \theta) e^{\beta H \mu \cos \theta} \\ &= \frac{2\pi}{\beta H \mu} (e^{\beta H \mu} - e^{-\beta H \mu}) = 4\pi \frac{\sinh(\beta H \mu)}{\beta H \mu}. \end{aligned}$$

Thus, the total partition function is

$$Z_N = \left(4\pi \frac{\sinh(\beta H\mu)}{\beta H\mu} \right)^N, \quad \Rightarrow \quad \log Z_N = N \log \left(4\pi \frac{\sinh(\beta H\mu)}{\beta H\mu} \right). \quad (6.4.2)$$

From this, we can compute the thermodynamic quantities:

- Free energy:

$$F = -\frac{1}{\beta} \log Z_N = -Nk_B T \log \left(4\pi \frac{\sinh(\beta H\mu)}{\beta H\mu} \right),$$

- Internal energy:

$$\begin{aligned} E &= -\frac{\partial}{\partial \beta} \log Z_N = -N \frac{\partial}{\partial \beta} \log \left(\frac{\sinh(\beta H\mu)}{\beta H\mu} \right) \\ &= -N \left[\frac{\mu H \cosh(\beta \mu H)}{\sinh(\beta \mu H)} - \frac{1}{\beta} \right] = -N \mu H \left[\coth(\beta \mu H) - \frac{1}{\beta \mu H} \right] \end{aligned}$$

- Entropy:

$$S = \frac{E - F}{T} = k_B \beta (E - F) = N k_B \left[\log \left(4\pi \frac{\sinh(\beta \mu H)}{\beta \mu H} \right) - \beta \mu H \coth(\beta \mu H) + 1 \right].$$

Zero temperature limit ($T \rightarrow 0$, $\beta \rightarrow \infty$): At very low temperatures, $\beta \mu H \gg 1$, we have $\coth(\beta \mu H) \rightarrow 1$, so

$$E(T \rightarrow 0) \simeq -N \mu H + 0 = -N \mu H.$$

This corresponds to all spins perfectly aligned with the external field, i.e., the system is in its **ground state** with minimal energy.

High temperature limit ($T \rightarrow \infty$, $\beta \rightarrow 0$): For $\beta \mu H \ll 1$, we expand $\coth(\beta \mu H) \simeq \frac{1}{\beta \mu H} + \frac{\beta \mu H}{3} + \dots$ ², giving

$$E(T \rightarrow \infty) \simeq -N \mu H \left(\frac{1}{\beta \mu H} - \frac{\beta \mu H}{3} - \frac{1}{\beta \mu H} \right) = 0.$$

In this limit, thermal fluctuations dominate and the spins are **randomly oriented**, leading to a vanishing average magnetic energy.

In summary at $T = 0$ the system is fully ordered along the field (maximal magnetization), while at $T \rightarrow \infty$ the system is completely disordered due to thermal agitation, so the magnetic energy averages to zero.

As it is easy to notice, our system will never undergo a phase transition, since all these quantities are analytical and do not present any discontinuities or singularities; we have to introduce interactions in order to get our system to be able to undergo a "change of state". But before doing that it is very useful to define a quantity that can help us understand the system dynamics better: the **Magnetization**, which will behave as the *order parameter* of the system.

²The Taylor expansion of the hyperbolic cotangent around $x = 0$ is $\coth(x) = \frac{1}{x} \left(1 + \frac{x^2}{3} \right) + \mathcal{O}(x^2)$.

6.4.2 | Magnetization

The total magnetization of the system is defined as the sum of the magnetic moments of all spins:

$$M = \left\langle \sum_{i=1}^N \boldsymbol{\mu}_i \right\rangle_c. \quad (6.4.3)$$

In the canonical ensemble, the thermal average is expressed as

$$\begin{aligned} & \frac{1}{Z_N} \int d\Omega_N \left(\sum_{i=1}^N \boldsymbol{\mu}_i \right) e^{\beta \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{H}} = \frac{1}{\beta Z_N} \left(\frac{\partial}{\partial H} \right) \int d\Omega_N e^{\beta \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{H}} \\ &= \frac{1}{\beta Z_N} \left(\frac{\partial}{\partial H} Z_N \right) = \frac{1}{\beta} \frac{\partial}{\partial H} \log(Z_N) = -\frac{\partial F}{\partial H} = N\mu \left[\coth(\beta\mu H) - \frac{1}{\beta\mu H} \right]. \end{aligned}$$

Zero temperature limit ($T \rightarrow 0, \beta \rightarrow \infty$): At very low temperatures, thermal fluctuations are negligible compared to the interaction with the external field. In this limit, $\beta\mu H \gg 1$ and $\coth(\beta\mu H) \rightarrow 1$, so

$$M \simeq N\mu(1 - 0) = N\mu.$$

All spins align with the external field, reaching the **saturation magnetization**. This represents the *ground state* of the system, in which the magnetic energy is minimized and thermal disorder is absent.

High temperature limit ($T \rightarrow \infty, \beta \rightarrow 0$): At high temperatures, thermal energy dominates over the interaction with the field. For $\beta\mu H \ll 1$, we expand

$$\coth(\beta\mu H) \simeq \frac{1}{\beta\mu H} + \frac{\beta\mu H}{3} + \dots,$$

giving

$$M \simeq N\mu \frac{\beta\mu H}{3} = \frac{N\mu^2}{3k_B T} H,$$

but zero in the first order. Thermal fluctuations randomize the orientations of the spins, leading to a small net magnetization. The proportionality constant

$$C = \frac{N\mu^2}{3k_B}, \quad (6.4.4)$$

is called the **Curie coefficient**, and the relation

$$M \simeq \frac{C}{T} H, \quad (6.4.5)$$

is the classical **Curie law** for paramagnets. It highlights that in the high-temperature regime, the magnetization decreases inversely with temperature, reflecting the competition between thermal disorder and alignment with the field. This constant is connected to the *magnetic susceptibility*:

$$\chi = \frac{\partial M}{\partial H}, \quad (6.4.6)$$

because in this regime we find that

$$\chi \xrightarrow{T \rightarrow \infty} \frac{M}{H} = \frac{C}{T}.$$

6.4.3 | Interacting Magnetic Spins

To understand the emergence of phase transitions, it is essential to include *interactions* among the magnetic moments of the lattice. In real materials, atomic magnetic moments do not behave independently: the orientation of each spin tends to influence its nearest neighbours through short-range exchange interactions.

If we consider such local couplings, the magnetic Hamiltonian can be written as

$$\mathcal{H}_{\text{mag}} = - \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{H} - J \sum_{\langle ij \rangle} \boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j,$$

where the first term describes the interaction with the external magnetic field \mathbf{H} , while the second term accounts for spin–spin interactions between nearest neighbours, denoted by the notation $\langle ij \rangle$. The number of nearest neighbours of each lattice site defines the **coordination number** z , which depends on the lattice geometry (for example, $z = 2$ in one dimension, $z = 4$ for the square lattice, and $z = 6$ for the triangular lattice).

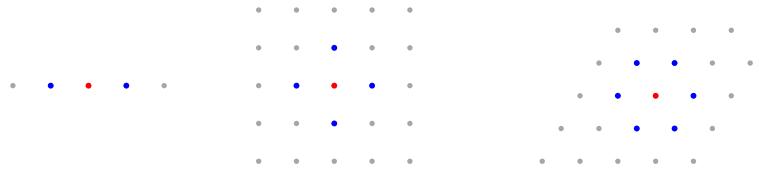


Figure 6.3: Comparison of coordination numbers z for lattices of increasing dimensionality: (a) one-dimensional chain ($z = 2$), (b) two-dimensional square lattice ($z = 4$), (c) two-dimensional triangular lattice ($z = 6$), which has the same coordination number as a simple cubic lattice in 3D.

Ferromagnetic and Paramagnetic Couplings. The sign of the coupling constant J determines the nature of the interaction:

- For $J > 0$, the energy is minimized when neighbouring spins are *aligned*, leading to a **ferromagnetic** interaction. At sufficiently low temperatures, this cooperative tendency gives rise to *spontaneous magnetization*, even in the absence of an external field.
- For $J < 0$, the energy is minimized when neighbouring spins are *anti-aligned*, leading to an **antiferromagnetic** interaction. The resulting long-range order alternates between up and down spins, producing a vanishing net magnetization.

In the absence of spin–spin interactions ($J = 0$), the system reduces to an ideal **paramagnet**, where spins respond independently to the external field and no cooperative ordering occurs.

The Ising Simplification. The full vector model above, known as the **Heisenberg model**, is generally difficult to solve analytically because spins can point in any direction on the unit sphere, and the interaction term involves continuous degrees of freedom. A powerful simplification is obtained by restricting each magnetic moment to take only two possible orientations along a fixed axis (conventionally the z -axis):

$$\boldsymbol{\mu}_i = \pm \mu \hat{\mathbf{z}}.$$

This leads to the **Ising model**, where each spin is represented by a discrete variable $\sigma_i = \pm 1$, and the Hamiltonian becomes

$$\mathcal{H}_{\text{Ising}} = -\mu H \sum_i \sigma_i - \mu^2 J \sum_{\langle ij \rangle} \sigma_i \sigma_j.$$

Despite its simplicity, this model captures the essential physics of cooperative phenomena and phase transitions in magnetic systems.

Remark. *The Ising model, though based on a binary spin variable, provides a universal paradigm for studying critical phenomena. Its conceptual importance lies in showing that sharp macroscopic transitions — such as the onset of magnetization — can emerge from the collective behaviour of many microscopic degrees of freedom interacting through local rules. Moreover, the Ising model's simplicity makes it a cornerstone not only in statistical mechanics, but also in diverse fields such as neural networks, lattice gases, and social dynamics, wherever binary interactions play a central role.*

Notice that, in this discrete formulation, the system no longer possesses a continuous phase space. Instead, its microstates form a finite ensemble of configurations given by all possible combinations of up and down spins on the lattice.

6.5 | Ising Model

Consider a hypercubic lattice whose sites carry discrete variables $\sigma_i = \pm 1$. The Ising Hamiltonian

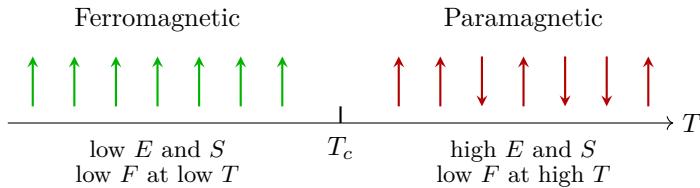
$$\mathcal{H}_{\text{Ising}} = -H \sum_i \sigma_i - J \sum_{\langle ij \rangle} \sigma_i \sigma_j, \quad (6.5.1)$$

where $J > 0$ favors configurations in which nearest-neighbor spins are aligned, while the external field H tends to align the spins along its direction; we are considering for simplicity $\mu = 1$.

To determine the equilibrium state, we minimize the free energy

$$F = E - TS = \langle \mathcal{H} \rangle - T(k_B \log \Gamma(E)),$$

which makes explicit the competition between the internal energy E and the entropy S . Energetically favorable states have aligned spins (low E), whereas entropy favors disordered configurations (high S). These two contributions typically push the system in opposite directions, and the temperature controls the relative weight of the entropic term. Thus, T emerges as the key parameter that determines which term dominates in the minimization of the free energy, and consequently which phase—ordered or disordered—the system will adopt.



If we inspect the Ising phase diagram (see figure 6.4), we immediately recognize the typical signature of a phase transition: a thermodynamic quantity develops a discontinuity at a specific value of the control parameters. Whenever such a non-analytic behaviour appears, a phase transition takes place. To distinguish the different phases, we introduce a quantity whose expectation value can indicate the degree of order in the system: the **order parameter**. For the Ising model—as well as for the Heisenberg model—the natural choice is the *magnetization*. Its behaviour clearly displays a discontinuity near the critical temperature, separating the ferromagnetic and paramagnetic regimes. Above T_c the magnetization vanishes, while below T_c it acquires two opposite non-zero values (corresponding to the two possible orientations of the ferromagnetic phase):

$$M = \begin{cases} 0, & \text{if } T \geq T_c, \\ M_0(T) \neq 0, & \text{if } T < T_c. \end{cases}$$

Thus the magnetization not only signals the phase transition but also encodes the ordering properties of the system. If we look at the M - H diagram, we can also observe the appearance of the *hysteresis loop* in the ferromagnetic phase. When a ferromagnet is magnetized and the external field is subsequently removed, the magnetization does not relax back to zero but to a finite residual value—an effect absent in the paramagnetic phase.

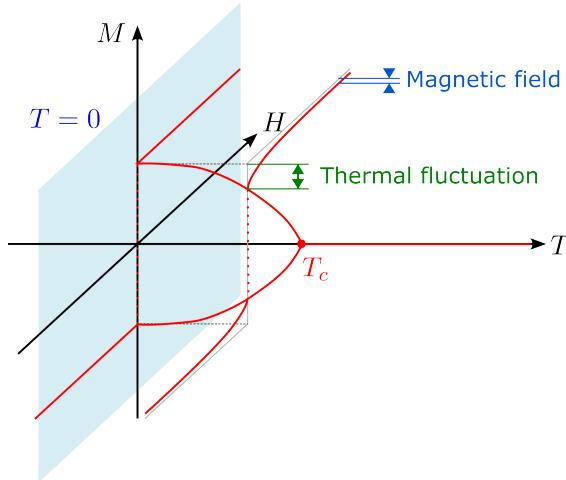


Figure 6.4: Schematic phase diagram of the Ising model, showing the relationship between magnetization M , external field H , and temperature T . The critical temperature T_c marks the boundary between the ferromagnetic (ordered) and paramagnetic (disordered) phases.

Below T_c , however, the system develops a non-zero magnetization $M \neq 0$, and the symmetry is no longer reflected in the equilibrium state: it is *spontaneously broken*, without the need for any explicit symmetry-breaking term.

6.5.1 | Mean-Field Treatment

The mean-field approach is a powerful approximation scheme that allows us to decouple interacting degrees of freedom by replacing the effect of all neighbours with an average, or **mean field**. Since the Ising model is exactly solvable only in 1D (and in 2D only for $H = 0$), while for 3D and higher dimensions no closed-form solution exists, we need an approximation scheme to obtain a tractable and physically meaningful solution.

We start from the partition function:

$$Z_N = \sum_{\{\sigma_i\}} e^{-\beta \mathcal{H}}, \quad \mathcal{H} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - H \sum_i \sigma_i,$$

and we immediately observe that the Hamiltonian cannot be written as a sum of independent single-spin contributions because of the nearest-neighbour interaction term. The key idea of the mean-field approximation is to replace the interaction $\sigma_i \sigma_j$ with an approximate expression depending only on the average magnetization m . With this replacement, the Hamiltonian becomes effectively decoupled and takes the form

$$Z_N^{(\text{mf})} \simeq \sum_{\{\sigma_i\}} e^{-\beta \mathcal{H}^{(\text{mf})}}, \quad \mathcal{H}^{(\text{mf})} = N \frac{J z m^2}{2} - (J z m + H) \sum_i \sigma_i,$$

where z is the coordination number of the lattice. In this form, the Hamiltonian is fully factorized into single-spin contributions.

Another key feature of phase transitions is the rate at which the order parameter changes as the system approaches the critical point. Close to T_c , the magnetization exhibits a power-law behaviour:

$$M(T \rightarrow T_c) \sim |T - T_c|^\beta,$$

where β is the **critical exponent** of the magnetization, characterizing how sharply the transition occurs.

Finally, it is useful to anticipate how symmetries enter the description of phase transitions. When the external field is turned off ($H = 0$) and the temperature is above the critical value ($T > T_c$), the Hamiltonian is invariant under a global \mathbb{Z}_2 transformation (under flipping of all spins).

Before carrying out the explicit computation, it is useful to summarize the logic of the procedure. We will first approximate the interaction term $\sigma_i \sigma_j$ up to a suitable order in fluctuations around the average magnetization. We then reorganize the Hamiltonian until it becomes factorized. Finally, we compute the single-spin partition function, from which the full partition function follows as its N -th power.

Magnetization per spin m . We introduce here the quantity that will play a central role in the mean-field construction: the *average magnetization per spin*. It measures the degree of alignment of the system and will serve as the parameter around which we expand the interaction term:

$$\begin{aligned} m &= \frac{M}{N} = \frac{\langle \sum_i \sigma_i \rangle}{N} = \frac{1}{N} \sum_{\{\sigma_i\}} \left(\sum_i \sigma_i \right) \frac{e^{-\beta \mathcal{H}(\{\sigma_i\})}}{Z_N} \\ &= \frac{1}{Z_N \beta N} \sum_{\{\sigma_i\}} \frac{\partial}{\partial H} \left[e^{-\beta \left(-H \sum_i \sigma_i - J \sum_{\langle ij \rangle} \sigma_i \sigma_j \right)} \right] \\ &= \frac{1}{\beta N} \frac{\partial}{\partial H} \log Z_N = -\frac{1}{N} \frac{\partial F}{\partial H}. \end{aligned}$$

This expression shows that the magnetization is the thermodynamic conjugate variable to the external field H , and it will naturally emerge as the self-consistent order parameter in the mean-field approximation.

Mean field and partition function. We are now ready to apply the mean-field approximation. The idea is the following: instead of treating the interaction of each spin σ_i with its nearest neighbours exactly, we replace the surrounding spins by an average field, effectively producing a new **mean magnetization field**. In practice, we rewrite each spin as

$$\sigma_i \rightarrow \sigma_i + m - m = m + (\sigma_i - m),$$

so that fluctuations appear only through the deviation $\sigma_i - m$. A preliminary approximation follows from the definition of the magnetization:

$$m = \frac{\langle \sum_i \sigma_i \rangle}{N} \sim \frac{\sum_i \langle \sigma_i \rangle}{N} = \langle \sigma_i \rangle,$$

where we have exploited translational invariance. We now expand the interaction term and neglect contributions of higher order in the fluctuations:

$$\begin{aligned} \sigma_i \sigma_j &= [m + (\sigma_i - m)][m + (\sigma_j - m)] \\ &= m^2 + m(\sigma_i - m) + m(\sigma_j - m) + (\sigma_i - m)(\sigma_j - m) \\ &\sim m^2 + m(\sigma_i + \sigma_j) - 2m^2 = m(\sigma_i + \sigma_j) - m^2, \end{aligned}$$

which corresponds to assuming that the fluctuations $\sigma_i - m$ are small with respect to m . With this approximation the Hamiltonian becomes decoupled:

$$\begin{aligned} \mathcal{H}(\{\sigma_i\}) &= -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - H \sum_i \sigma_i \\ &\sim -J \sum_{\langle ij \rangle} (m(\sigma_i + \sigma_j) - m^2) - H \sum_i \sigma_i \\ &= J \frac{Nz}{2} m^2 - J \sum_{\langle ij \rangle} 2m\sigma_i - H \sum_i \sigma_i, \end{aligned}$$

where we used that the number of nearest-neighbour pairs is $\frac{Nz}{2}$. We also used that each σ_i appears twice in the sum over pairs, hence $\sum_{\langle ij \rangle} (\sigma_i + \sigma_j) = \sum_{\langle ij \rangle} 2\sigma_i$. Continuing the computation:

$$\begin{aligned}\mathcal{H}(\{\sigma_i\}) &= J \frac{Nz}{2} m^2 - J \sum_{\langle ij \rangle} 2m\sigma_i - H \sum_i \sigma_i \\ &= J \frac{Nz}{2} m^2 + \sum_i \left(-J \frac{z}{2} 2m\sigma_i - H\sigma_i \right) \\ &= \sum_i \left[J \frac{z}{2} m^2 - (Jzm + H)\sigma_i \right],\end{aligned}$$

since each σ_i is counted $\frac{z}{2}$ times among the nearest-neighbour pairs. At this stage the Hamiltonian is explicitly factorized into single-spin contributions, and we may therefore introduce the single-spin mean-field Hamiltonian

$$\mathcal{H}^{(mf)}(\sigma_i) = J \frac{z}{2} m^2 - (Jzm + H)\sigma_i. \quad (6.5.2)$$

The canonical partition function now follows straightforwardly:

$$\begin{aligned}Z_N = (Z_1)^N &= (e^{-\beta \mathcal{H}^{(mf)}(\sigma_i)})^N = e^{-\beta J \frac{Nz}{2} m^2} \left[e^{\beta \sum_{\sigma_i=\pm 1} (Jzm + H)\sigma_i} \right]^N \\ &= e^{-\beta J \frac{Nz}{2} m^2} \prod_{\sigma_i=\pm 1} \left[e^{\beta (Jzm + H)\sigma_i} \right]^N = e^{-\beta J \frac{Nz}{2} m^2} (2 \cosh(\beta(H + Jzm)))^N.\end{aligned}$$

(?) how do we get the cosh term? its not a sum of exponentials, its a product.

We have thus obtained the mean-field expression for the partition function. It is worth noting that, although the approximation fails in $d = 1$ and is only qualitatively accurate in $d = 2$, it becomes exact for $d \geq 4$. This behaviour is consistent with the Mermin–Wagner theorem, which forbids spontaneous symmetry breaking in low spatial dimensions.

Magnetization and self-consistency. We now aim to obtain an explicit expression for the order parameter of the system. Using the relation between the magnetization m and the Helmholtz free energy, we begin by computing the latter:

$$\begin{aligned}F &= -\frac{1}{\beta} \log Z_N = -\frac{1}{\beta} \left[-\beta J \frac{Nz}{2} m^2 + N \log(2 \cosh(\beta(H + Jzm))) \right] \\ &= J \frac{Nz}{2} m^2 - \frac{N}{\beta} \log(2 \cosh(\beta(H + Jzm))).\end{aligned}$$

Taking the derivative with respect to the magnetic field, we obtain the magnetization:

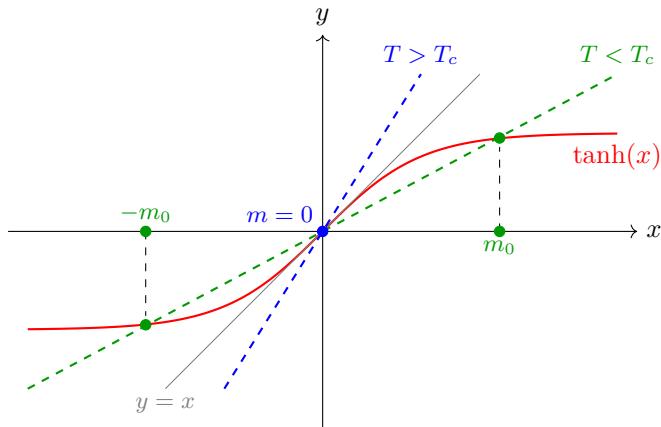
$$\begin{aligned}m &= -\frac{1}{N} \frac{\partial F}{\partial H} = -\frac{1}{\beta} \frac{\partial}{\partial H} \log(2 \cosh(\beta(H + Jzm))) \\ &= -\frac{2 \sinh(\beta(H + Jzm)) \beta}{2 \beta \cosh(\beta(H + Jzm))} = \tanh(\beta(H + Jzm)).\end{aligned}$$

This equation contains m on both sides and therefore constitutes a *self-consistency relation*. Such relations are typical of mean-field treatments and express the fact that the order parameter must be compatible with the field it generates.

Although the equation cannot be solved analytically, it already reveals important physical information. Since the right-hand side is monotonic, for $H > 0$ we find $m > 0$, and for $H < 0$ we find $m < 0$. Let us therefore focus on the case $H = 0$:

$$\begin{aligned}m &= \tanh(\beta Jzm) = \tanh\left(\frac{T_c}{T} m\right) = \tanh(x), \\ y &= \frac{T}{T_c} x = m = \tanh(x), \quad T_c = \frac{Jz}{k_B}.\end{aligned}$$

This form makes the role of temperature explicit: it controls the slope of the straight line that must intersect the curve $y = \tanh(x)$.



The graphical solution shows that for $T > T_c$ the only intersection is at the origin, yielding $m = 0$: this corresponds to the **paramagnetic phase**. For $T < T_c$, instead, the line intersects the curve at two symmetric non-zero points, corresponding to two stable ordered states with opposite magnetization: this is the **ferromagnetic phase**.

6.5.2 | Spontaneous Symmetry Breaking

As we have established, the Ising Hamiltonian is symmetric under a global $\mathbb{Z}_2 = \{\mathbb{I}, g\}$ transformation. However, this invariance of the Hamiltonian does not imply that the thermodynamic state of the system remains symmetric. We have observed that in the paramagnetic phase, the magnetization transforms as:

$$M \xrightarrow{\sigma_i \rightarrow -\sigma_i} -M, \quad \text{with } M = -M = 0.$$

Here, the symmetry is respected. In contrast, in the ferromagnetic state, the net magnetization is non-zero and thus is not invariant under \mathbb{Z}_2 . This phenomenon is known as *spontaneous symmetry breaking* (SSB).

At lower temperatures, the system enters an ordered phase, and the effective symmetries are reduced to a subgroup $G_0 \subset G$ of the original symmetry group. Let the order parameter depend on an observable ϕ that is not invariant under the transformations of G . If the thermal average behaves as follows:

$$\Phi = \langle \phi \rangle = \begin{cases} 0, & \text{if } T \geq T_c, \\ \Phi_0(T) \neq 0, & \text{if } T < T_c, \end{cases}$$

then we say *the symmetry is spontaneously broken*. The magnetization serves as a perfect example of such an order parameter. It is important to note that while every instance of SSB implies the existence of a Phase Transition (PT), not all Phase Transitions are accompanied by SSB.

Correlation. Another quantity worthy of analysis is the *correlation* (or covariance), which describes how spins influence each other over large distances. Intuitively, we expect the correlation to decrease as the distance increases, but the specific mode of this decay offers deep physical insights. The correlation function between two spins is defined as:

$$G_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle = \langle \sigma_i \sigma_j \rangle - m^2,$$

assuming translational invariance where $\langle \sigma_i \rangle = \langle \sigma_j \rangle = m$. Letting $r = |i - j|$ be the distance, the asymptotic behavior for large r is given by:

TODO: plot the correlation

$$G(r) = \begin{cases} e^{-\frac{r}{\xi}}, & \text{if } T \neq T_c; \\ (r^{d-2-\eta})^{-1}, & \text{if } T = T_c. \end{cases} \quad (6.5.3)$$

For $T \neq T_c$, the correlation $G(r)$ remains significantly large only for $r < \xi$ and becomes negligible for $r > \xi$. The parameter ξ is called the **correlation length**, representing the typical scale of fluctuations or the size of the regions within which spins are strongly correlated.

Crucially, we observe that $\xi \rightarrow \infty$ as $T \rightarrow T_c$. Physically, this means that as the critical temperature is approached, distant spins become increasingly correlated. Exactly at $T = T_c$, the correlation length diverges, and the system exhibits long-range order, resulting in a power-law decay for $G(r)$ rather than an exponential one.

6.6 | Critical Exponents

We now investigate the behavior of the system as $T \rightarrow T_c$ to characterize the properties of the Phase Transition. The strategy is to perform a Taylor expansion of $\tanh(x)$. First, let us define a reduced temperature variable t (representing an infinitesimal deviation):

$$t = \frac{T_c - T}{T_c} = 1 - \frac{T}{T_c} \ll 1. \quad (6.6.1)$$

From this, we can write the inverse temperature ratio as:

$$\frac{T_c}{T} = (1 - t)^{-1} \sim 1 + t.$$

Recall the mean-field equation involving the magnetization m . We can make m explicit inside the argument of the hyperbolic tangent, or equivalently work with the inverse. Using the relation $\tanh^{-1}(m) = \frac{T_c}{T}m$, and substituting our expression for T_c/T :

$$\tanh^{-1}(m) = (1 + t)m.$$

Next, we Taylor expand the inverse hyperbolic tangent in powers of m (since m is small near T_c):

$$\tanh^{-1}(m) \sim m + \frac{m^3}{3} = (1 + t)m \implies \frac{m^3}{3} \approx tm \implies m^2 = 3t.$$

We have thus found how the magnetization scales in proximity to the Phase Transition:

$$m(T \rightarrow T_c) \sim t^{\frac{1}{2}}, \implies \beta = \frac{1}{2}.$$

The reduced temperature t allows us to quantify the rapidity of the transition. In terms of the absolute temperature difference, we found:

$$m(T \rightarrow T_c) \sim |T - T_c|^{\beta=\frac{1}{2}}.$$

This exponent is one of several universal parameters defined to monitor the rate of change (scaling) of thermodynamic variables near a phase transition. These are called **critical exponents**. Formally, a critical exponent λ is associated with a variable $f(t)$ as follows:

$$\lambda = \lim_{t \rightarrow 0} \frac{\log(f(t))}{\log(t)} \longrightarrow f(t \rightarrow 0) = g(t)|t|^\lambda, \quad \text{where } g(0) \neq 0. \quad (6.6.2)$$

These exponents characterize the singular behaviour of various thermodynamic quantities as the system approaches the critical point. We typically consider six of these standard exponents:

- α for **specific heat**: $C_{H=0} \sim |t|^{-\alpha}$;
- β for **order parameter** (magnetization): $\phi \sim |t|^\beta$;
- γ for **susceptibility**: $\chi_{H=0} \sim |t|^{-\gamma}$;
- δ for **external field** (at the critical isotherm $T = T_c$): $H \sim \text{sgn}(\phi)|\phi|^\delta$;
- ν for **correlation length**: $\xi \sim |t|^{-\nu}$;
- η for **anomalous dimension** (at $T = T_c$): $G(r) \sim \frac{1}{r^{d-2+\eta}}$.

Scaling Hypothesis. When the system undergoes a Phase Transition (PT), correlations spread throughout the nodes. Once the PT is completed, the system settles into a single thermodynamic phase. However, approaching the critical point, one can recognize fluctuations where different local states effectively coexist. During a continuous PT, the **correlation length** (or *scale length*) diverges to infinity, until the entire system becomes correlated (in the Ising model, this corresponds to macroscopic magnetic ordering). At this precise moment, the system is said to be **scale invariant**. Conversely, as long as $T \neq T_c$, the system is said to be *holding a scale* (characterized by finite ξ), allowing us to distinguish the thermodynamic phases.

Theorem 6.3. *At a critical point during a PT, the system becomes scale invariant (this forms the basis of Widom's hypothesis). Consequently, the singular part of the free energy must be a homogeneous function of the reduced temperature t and the external field H , respecting the scaling relation:*

$$F(\lambda^a t, \lambda^b H) = \lambda F(t, H), \quad (6.6.3)$$

for some scaling parameters a and b .

From this hypothesis, we can compute the relations connecting the scaling parameters a and b with our critical exponents:

$$\begin{aligned} \alpha &= 2 - \frac{1}{a}, & \beta &= \frac{1-b}{a}, \\ \gamma &= \frac{2b-1}{a}, & \delta &= \frac{b}{1-b}, \\ \nu &= \frac{1}{da}, & \eta &= 2 - d(2b-1). \end{aligned}$$

Finally, it is useful to derive the *scaling laws* (identities connecting the critical exponents), which act as a consistency check:

$$\begin{cases} (2-\eta)\nu = \gamma, \\ \alpha + \beta(1+\delta) = 2, \\ d\nu = 2 - \alpha, \\ \alpha + 2\beta + \gamma = 2. \end{cases} \quad (6.6.4)$$

Universality Classes. Due to scale invariance at the critical point, systems with different Hamiltonians and distinct microscopic descriptions might display the exact same type of PT. Microscopically, these systems may appear completely unrelated; however, as they approach the PT, their macroscopic behavior coincides. This implies that the numerical values of the six exponents $\alpha, \beta, \gamma, \delta, \nu, \eta$ are identical across these systems.

We can thus classify PTs into **universality classes**. Membership in a class does not depend on the explicit form of the Hamiltonian or the specific details of the interactions, but only upon:

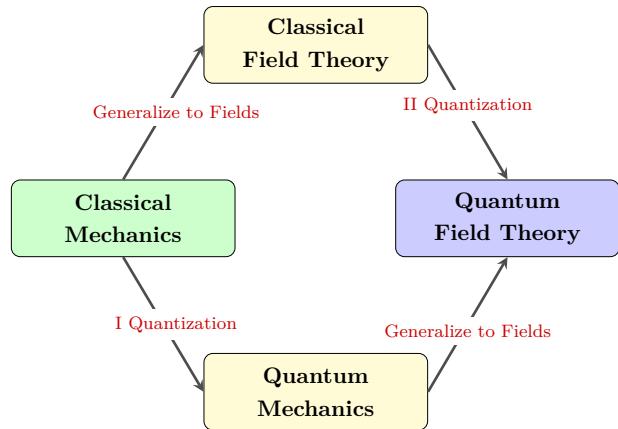
- the *dimension* of the physical space $d = 1, 2, 3 \dots$;
- the *symmetry group* G of the Hamiltonian;
- the *residual symmetry subgroup* G_0 in the *broken phase* after the Spontaneous Symmetry Breaking (SSB).

If these three factors coincide, the Hamiltonians exhibit the same phase transition behavior and belong to the same universality class, characterized by a unique set of critical exponents.

Remark. In $d = 1$ there is no finite-temperature PT because $T_c \rightarrow 0$. Approaching T_c from above implies $T \rightarrow 0$, and "crossing" it would physically imply negative temperatures (which are not accessible in equilibrium). Thus, the standard critical exponents are not well-defined for the 1D case.

7 | QM and Statistics

The starting point is a classical system with a finite number of degrees of freedom, from which we can either move to a classical field theory with infinite degrees of freedom (with the idea to quantize after it), or to a first quantization of our finite classical system (with the idea to extend observables to field later). It's not trivial that the result is the same taking both procedures, but it's true.



Let's briefly review quantum mechanics, in comparison with a classical system as a start.

Classical Mechanics (CM)		Quantum Mechanics (QM)
A precise point in the phase space (\mathbf{q}, \mathbf{p}) , where position and momentum are simultaneously determined.	State	A normalized state vector $ \psi\rangle \in \mathcal{H}$ (Hilbert Space), defined up to a global phase: equivalence classes called rays .
Functions defined on the phase space, e.g. $f(\mathbf{q}, \mathbf{p}) \in \mathbb{R}$.	Observables	Linear, self-adjoint (Hermitian) operators $\hat{A} = \hat{A}^\dagger$ acting on the Hilbert space.
Direct and simultaneous. The act of measurement does not alter the system's state. We can measure a value of the observable $f(\mathbf{q}^o, \mathbf{p}^o)$ at a specific point $(\mathbf{q}^o, \mathbf{p}^o)$.	Measurements	The result is a spectrum of eigenvalues λ_n of \hat{A} . The state instantly collapses to the corresponding eigenstate $ \lambda_n\rangle$. The probability of λ_n as an outcome is $p_n = \langle\psi \lambda_n\rangle ^2 = \langle\psi \mathbb{P}_n \psi\rangle$, if $ \psi\rangle$ is the generic state and $ \lambda_n\rangle$ is the eigenstate $\hat{A} \lambda_n\rangle = \lambda_n \lambda_n\rangle$.
Determined by Hamilton's equations $\dot{q}_i = -\frac{\partial H}{\partial p_i}$ and $\dot{p}_i = \frac{\partial H}{\partial q_i}$. Evolving the system, the observables can be measured in the new points on the phase space.	Time Evolution	Determined by the Schrödinger equation for the states $i\hbar\frac{\partial}{\partial t} \psi\rangle = \hat{H} \psi\rangle$. The time evolution operator is unitary and defined by the Hamiltonian.

In more detail, the degrees of freedom are described in terms of a vector in the Hilbert space $|\psi\rangle, |\phi\rangle, \dots \in \mathcal{H}$, such that

$$\alpha |\psi\rangle + \beta |\phi\rangle \in \mathcal{H}, \text{ and } \langle \psi | \phi \rangle \in \mathbb{C}.$$

A **pure state** is a *ray* belonging to an **equivalence class**, defined up to a global phase

$$|\psi\rangle \sim e^{i\theta} |\psi\rangle, \quad \langle \psi | \psi \rangle = 1.$$

An **observable** is given by a *self-adjoint* operator $\hat{A} = \hat{A}^\dagger$, for which the *spectral theorem* holds

$$\hat{A} |\psi_j\rangle = \lambda_j |\psi_j\rangle, \text{ where } \lambda_j \in \mathbb{R},$$

since the outcome of a measure, the eigenvalues of the interesting observable, must be real.

Projectors. It is very useful to introduce operators which can *project* a given state, an arbitrary vector, in the direction spanned by a chosen vector:

$$\hat{\mathbb{P}}_\psi = |\psi\rangle \langle \psi| = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} a_1^* & a_2^* & \dots & a_n^* \end{pmatrix} = (n \times n) \text{ matrix.}$$

These **projectors** have to be *idempotent* matrices, such that

$$\hat{\mathbb{P}}^\dagger = \hat{\mathbb{P}}, \quad \hat{\mathbb{P}}^2 = \hat{\mathbb{P}}.$$

They project vectors onto the 1D Hilbert space spanned by their defining vector. Considering an orthonormal basis $\{|\psi_n\rangle\}$, we can exploit some interesting properties:

$$\hat{\mathbb{P}}_n \hat{\mathbb{P}}_m = |\psi_n\rangle \langle \psi_n| \psi_m \rangle \langle \psi_n| = |\psi_n\rangle \langle \psi_n| \delta_{nm},$$

hence if $m \neq n$ then the two defining vectors are orthonormal and $\hat{\mathbb{P}}_n \hat{\mathbb{P}}_m = 0$. If we sum the projectors, i.d. we apply them consequently, we find interesting compositions:

- $\hat{\mathbb{P}}_1 + \hat{\mathbb{P}}_2$ 2D: projector on the plane spanned by $\{|\psi_1\rangle, |\psi_2\rangle\}$;
- $\hat{\mathbb{P}}_1 + \hat{\mathbb{P}}_2 + \hat{\mathbb{P}}_3$ 3D: projector on the volume spanned by $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$;
- $\hat{\mathbb{P}}_1 + \dots + \hat{\mathbb{P}}_n$ nD: projector that coincides with the identity:

$$\sum_n \hat{\mathbb{P}}_n = \sum_n |\psi_n\rangle \langle \psi_n| = \mathbb{I}_n.$$

According to the *spectral theorem*, any hermitian operator can be decomposed in:

$$\hat{A} = \sum_n \lambda_n \hat{\mathbb{P}}_n = \sum_n \lambda_n |\psi_n\rangle \langle \psi_n|.$$

Measure. A measure of an observable \hat{A} on a state $|\psi\rangle$ yields a set of possible outcomes λ_n weighted by a probability p_n given by

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle \implies p_n = |c_n|^2.$$

The average value of a measure is given by

$$\langle \hat{A} \rangle = \sum_n \lambda_n p_n = \langle \psi | \hat{A} | \psi \rangle,$$

with associated standard deviation

$$(\Delta \hat{A})^2 = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2.$$

One important feature of the measure is that it represents a *destructive process*: once measured, the state $|\psi\rangle$ collapses into a state in the eigenspace associated to the eigenvalue which is the outcome of the measure.

Time evolution. The time evolution of the system is governed by a special observable, the Hamiltonian; from Schrödinger equation we deduce the time evolution of each state:

$$i\hbar \frac{\partial}{\partial t} |\psi\rangle = \hat{H} |\psi\rangle.$$

We can also create an operator responsible for the active implementation of this equation, assuming the time independence of the Hamiltonian, which results to be a unitary operator:

$$|\psi(t)\rangle = \hat{U}(t) |\psi(0)\rangle, \quad \hat{U}(t) = e^{\frac{i}{\hbar} \hat{H} t}.$$

7.1 | Density Matrix

7.1.1 | Pure State

We are going to introduce one of the most powerful tools in the quantum description of statistical mechanics, the **density matrix**.

Definition 7.1 (Density Matrix of a Pure State). A pure density matrix $\hat{\rho}$ has the following properties:

- it is a bounded operator : $\|\hat{\rho}\| \leq 1$;
- it is self-adjoint: $\hat{\rho}^\dagger = \hat{\rho}$;
- it is a positive operator: $\langle \alpha | \hat{\rho} | \alpha \rangle \geq 0 \forall |\alpha\rangle \in \mathcal{H}$;
- it has unitary trace: $\text{Tr}(\hat{\rho}) = 1$;
- it is idempotent: $\hat{\rho}^2 = \hat{\rho}$.

For a pure state, unambiguously defined by a ray $|\psi\rangle$, the density matrix acquire the meaning of projector onto the subspace spanned by the state:

$$\hat{\rho}_\psi = |\psi\rangle \langle \psi|. \quad (7.1.1)$$

We can use it to define **expection values** for operators, which can be written in terms of the density matrix as

$$\langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle = \text{Tr}(\hat{A} \hat{\rho}_\psi), \quad (7.1.2)$$

where the trace of an operator with elements $A_{mn} = \langle \psi_m | \hat{A} | \psi_n \rangle$ is given by

$$\text{Tr}(\hat{A}) = \sum_n \langle \psi_n | \hat{A} | \psi_n \rangle = \sum_n A_{nn}.$$

A **trace-class operator** \hat{B} is defined as, given any ON basis $\{|e_j\rangle\}$

$$\text{Tr}(\hat{B}) = \sum_n \langle e_n | \hat{B} | e_n \rangle < \infty, \quad (7.1.3)$$

i.e. the class of operators with finite trace (for an infinite dimensional operators is absolutely non trivial to have finite trace, even the simplest one, the identity matrix, has a diverging trace).

Let's show the unitarity of the density operator trace: given any ON basis $\{|e_j\rangle\}$

$$\begin{aligned} \text{Tr}(\hat{\rho}) &= \sum_n \langle e_n | \hat{\rho} | e_n \rangle = \sum_n \langle e_n | \psi \rangle \langle \psi | e_n \rangle \\ &= \langle \psi | \left(\sum_n |e_n\rangle \langle e_n| \right) | \psi \rangle = \langle \psi | \mathbb{I} | \psi \rangle = 1. \end{aligned}$$

7.1.2 | Mixed State

The pure state, as we said, is unambiguously defined by a single ray $|\psi\rangle$, but it might happen to encounter a system prepared into a ensamble of states $\{|\psi_\alpha\rangle, p_\alpha\}_\alpha$, where p_α is the probability of finding the system in the state $|\psi_\alpha\rangle$ and obviously they respect: $0 \leq p_\alpha \leq 1, \sum_\alpha p_\alpha = 1$.

Definition 7.2 (Density Matrix of a Mixed State). The density matrix of a classical mixture $\{|\psi_\alpha\rangle, p_\alpha\}_\alpha$ is given by

$$\hat{\rho} = \sum_{\alpha} p_{\alpha} \hat{\rho}_{\alpha} = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}|, \quad (7.1.4)$$

where $\hat{\rho}_{\alpha}$ is the density matrix of the pure state $|\psi_{\alpha}\rangle$. It has the following properties:

- it is a bounded operator : $\|\hat{\rho}\| \leq 1$;
- it is self-adjoint: $\hat{\rho}^\dagger = \hat{\rho}$;
- it is a positive operator: $\langle \alpha | \hat{\rho} | \alpha \rangle \geq 0 \forall |\alpha\rangle \in \mathcal{H}$;
- it has unitary trace: $\text{Tr}(\hat{\rho}) = 1$.

Example. Consider an experiment using thousands of electrons, where each electron can be in one of two orthogonal spin states along the z -axis:

1. **Pure State 1:** Spin up, $|\psi_1\rangle = |\uparrow\rangle$.
2. **Pure State 2:** Spin down, $|\psi_2\rangle = |\downarrow\rangle$.

If the electron beam is unpolarized (not specially prepared), the overall system is a statistical mixture described by the ensemble:

$$\{|\uparrow\rangle, p_1 = 1/2\} \quad \text{and} \quad \{|\downarrow\rangle, p_2 = 1/2\}$$

This collection of electrons is thus described not by a single $|\psi\rangle$ (pure state), but by the density operator (mixed state):

$$\hat{\rho} = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}| = \frac{1}{2} |\uparrow\rangle \langle \uparrow| + \frac{1}{2} |\downarrow\rangle \langle \downarrow|$$

Notice how the density matrix of a mixed state is **not idempotent** anymore. Let us analyze further its properties:

- Self adjoint operator:

$$\hat{\rho}^\dagger = \left(\sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}| \right)^\dagger = \sum_{\alpha} p_{\alpha} \langle \psi_{\alpha}^\dagger | \psi_{\alpha}^\dagger \rangle = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}| = \hat{\rho}.$$

- Positive operator:

$$\langle \psi | \hat{\rho} | \psi \rangle = \langle \psi | \left(\sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}| \right) | \psi \rangle = \sum_{\alpha} p_{\alpha} \| \langle \psi_{\alpha} | \psi \rangle \|^2 \geq 0.$$

- Unitary trace:

$$\begin{aligned} \text{Tr}(\hat{\rho}) &= \sum_n \langle e_n | \hat{\rho} | e_n \rangle = \sum_n \langle e_n | \left(\sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}| \right) | e_n \rangle \\ &= \sum_{\alpha n} p_{\alpha} \langle \psi_{\alpha} | e_n \rangle \langle e_n | \psi_{\alpha} \rangle = \sum_{\alpha} p_{\alpha} \langle \psi_{\alpha} | \psi_{\alpha} \rangle = \sum_{\alpha} p_{\alpha} = 1. \end{aligned}$$

Let us show and demonstrate an important theorem:

Theorem 7.1. If the density matrix is idempotent, then it refers to a pure state:

$$\hat{\rho}^2 = \hat{\rho} \iff \hat{\rho} = |\psi\rangle\langle\psi| \text{ with } |\psi\rangle \text{ pure.}$$

Proof. (\Leftarrow) trivial, by definition of density matrix of a pure state. (\Rightarrow) $\hat{\rho}$ has to satisfy at least the four properties reported above for a density matrix. Then if we apply the spectral theorem we can write

$$\hat{\rho} = \sum_n \lambda_n |e_n\rangle\langle e_n|, \quad \{|e_n\rangle\} \text{ any ON basis,}$$

and this implies

$$\hat{\rho}|e_n\rangle = \lambda_n |e_n\rangle,$$

as any hermitian operator acting on an ON basis. Since $\hat{\rho}$ is positive, we know $\lambda_n \geq 0$, and since it has unitary trace, we can write

$$\text{Tr}(\hat{\rho}) = 1 \iff 0 \geq \lambda_n \leq 1.$$

If we now compute $\hat{\rho}^2$, we get

$$\begin{aligned} \hat{\rho}^2 &= \left(\sum_n \lambda_n |e_n\rangle\langle e_n| \right) \left(\sum_m \lambda_m |e_m\rangle\langle e_m| \right) \\ &= \sum_{nm} \lambda_n \lambda_m |e_n\rangle\langle e_n| |e_m\rangle\langle e_m| = \sum_{nm} \lambda_n \lambda_m |e_n\rangle \delta_{nm} \langle e_m| = \sum_n \lambda_n^2 |e_n\rangle\langle e_n|, \end{aligned}$$

and by imposing $\hat{\rho}^2 = \hat{\rho}$ we find out that

$$\hat{\rho}^2 = \hat{\rho} \implies \sum_n \lambda_n^2 |e_n\rangle\langle e_n| = \sum_n \lambda_n |e_n\rangle\langle e_n|$$

hence $\lambda_n^2 = \lambda_n$, which admits $\lambda_n = 0, 1$. But since $\sum_j \lambda_j = 1$, then the solution imply one $\lambda_j = 1$ and $\lambda_i = 0$ for all other indices $i \neq j$. But if only one eigenvalue survives, then we can only deduce it is a pure density matrix:

$$\hat{\rho} = \sum_n \lambda_n |e_n\rangle\langle e_n| = \lambda_j |e_j\rangle\langle e_j| = |e_j\rangle\langle e_j|.$$

□

7.2 | Qubit

The qubit is a quantum system with only two available states: $|0\rangle$ and $|1\rangle$. If we consider the following Hilbert space

$$\mathcal{H} = \mathbb{C}^2 = \left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix}; \quad \alpha, \beta \in \mathbb{C} \right\}$$

and the two states $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, they respect

$$\langle 0|0\rangle = \langle 1|1\rangle = 1, \text{ and } \langle 1|0\rangle = \langle 0|1\rangle = 0.$$

Then any vector in the Hilbert space can be written as a linear combination of these two states:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad |\alpha|^2 + |\beta|^2 = 1,$$

for all $|\psi\rangle \in \mathcal{H}$. Let's notice that $p_0 = |\alpha|^2$ and $p_1 = |\beta|^2$, since they are the weight of the state with respect to $|0\rangle$ and $|1\rangle$: they are the probabilities to get an outcome of $|0\rangle$ or $|1\rangle$ from a measure on $|\psi\rangle$. Thus we call $|\psi\rangle$ a **quantum mixture**, which is a pure state. Its density matrix is defined as:

$$\hat{\rho} = |\psi\rangle\langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix},$$

where the off-diagonal elements are proportional to a phase $\alpha\beta^*$, making them responsible for **interference** effect between the two states.

If we instead prepare from the beginning a state $|\psi\rangle$ with the same probabilities as before ($p_0 = |\alpha|^2$ and $p_1 = |\beta|^2$) but not related to expectation values of observables when measured on $|\psi\rangle$, direct probabilities that the system will be either in $|0\rangle$ or $|1\rangle$, then we have a **classical mixture**. Its density matrix is defined by

$$\hat{\rho} = |\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1| = |\alpha|^2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + |\beta|^2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}.$$

If we compare the two expression we can find interesting that the diagonal is the same, while the off diagonal elements are not trivial: while the classical mixture has null off diagonal elements (respecting a classical probability distribution without interactions among the two states), the quantum mixture presents terms correlated to **interference** in the measures among the two states.

Observables and expectation values. Lets consider a quantum system in a pure state $|\psi\rangle$; an observable \hat{A} (which we remember to be hermitian) will have an expectation value given by

$$\langle \hat{A} \rangle_\psi = \langle \psi | \hat{A} | \psi \rangle = \text{Tr}(\hat{A}\hat{\rho}_\psi).$$

Let us show that the last equation holds:

$$\begin{aligned} \langle \hat{A} \rangle_\psi &= \text{Tr}(\hat{A}\hat{\rho}_\psi) = \sum_n \langle e_n | \hat{A}\hat{\rho}_\psi | e_n \rangle \\ &= \sum_n \langle e_n | \psi \rangle \langle \psi | \hat{A} | e_n \rangle = \sum_n \langle \psi | \hat{A} | e_n \rangle \langle e_n | \psi \rangle \\ &= \langle \psi | \lambda_n \sum_n |e_n\rangle \langle e_n| \psi \rangle = \langle \psi | \hat{A} | \psi \rangle = \langle \hat{A} \rangle_\psi. \end{aligned}$$

This expression is very useful because it can be generalized to treat expectation values of observables acting on mixed states. The density matrix for a mixed state $|\psi\rangle$ is

$$\hat{\rho}_\psi = \sum_{\alpha} p_{\alpha} \hat{\rho}_{\alpha} = \sum_{\alpha} p_{\alpha} |\psi_{\alpha}\rangle \langle \psi_{\alpha}|,$$

then if we compute the expectation value of \hat{A} :

$$\langle \hat{A} \rangle_{\psi} = \text{Tr}(\hat{A} \hat{\rho}_{\psi}) = \text{Tr}\left(\sum_{\alpha} p_{\alpha} \hat{\rho}_{\alpha} \hat{A}\right) = \sum_{\alpha} p_{\alpha} \text{Tr}(\hat{\rho}_{\alpha} \hat{A}) = \sum_{\alpha} p_{\alpha} \langle \hat{A} \rangle_{\alpha},$$

which is the sum of the expectations value of \hat{A} in each of the pure states, weighted by the probability of the mixed state to be observed in that state.¹

¹A clearer but heavier notation would have called $\hat{\rho}_{\alpha} \rightarrow \hat{\rho}_{\psi_{\alpha}}$ and $\langle \hat{A} \rangle_{\alpha} \rightarrow \langle \hat{A} \rangle_{\psi_{\alpha}}$.

7.3 | Identical Particles

A system of N particles, each of which are described by a subsystem and its Hilbert space \mathcal{H}_j , lives in the bigger Hilbert space defined by

$$\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_N.$$

We can find a base of our system by composing the basis of each subsystem:

$$\mathcal{H}_j \longrightarrow \{|e_{k_j}\rangle\}_{k_j} \quad \forall k_j = 1, \dots, \dim(\mathcal{H}_j),$$

so that the general base will take form

$$\mathcal{H} \longrightarrow \{|e_{k_1}\rangle, \dots, |e_{k_N}\rangle\}_{k_1, \dots, k_N} \quad \forall k_j = 1, \dots, \dim(\mathcal{H}_j).$$

Now we can write a generic state (vector) in coordinates of this base as

$$|\psi\rangle = \sum_{k_1, \dots, k_N} \alpha_{k_1, \dots, k_N} |e_{k_1}\rangle \dots |e_{k_N}\rangle.$$

The dimensionality of this space will be given by the product of the dimensionalities of the subsystems

$$\dim(\mathcal{H}) = \dim(\mathcal{H}_1) \dots \dim(\mathcal{H}_N),$$

thus the dimensionality of the system grows really fast with the number of particles considered.

Note that if the particles are identical, then

$$\mathcal{H} = (\mathcal{H}_1)^{\otimes N}, \quad \dim(\mathcal{H}) = \dim(\mathcal{H}_1)^N.$$

The system of **indistinguishable** particles acquire also an important property: it should be invariant under transformations of the **permutation group**, up to a phase. If we take as an example a system of two indistinguishable classical particles, then the phase space $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2)$ should be invariant under the exchange of the particles:

$$(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) = (\mathbf{q}_2, \mathbf{q}_1, \mathbf{p}_2, \mathbf{p}_1).$$

Thus a system of N identicle particles should be invariant under a set of $N!$ permutations, with the states spanning only a subspace of the bigger Hilbert space just described.

To clarify further, let's take two particles e and f , both of which can be observed in the state $|\cdot\rangle_1$ or $|\cdot\rangle_2$: we have two possible global states

$$|e\rangle_1 |f\rangle_2 \quad \text{or} \quad |f\rangle_1 |e\rangle_2.$$

But if we consider the case in which the two particles are identical $e \equiv f$ or the two states are the same $|e\rangle_1 \equiv |e\rangle_2$, then the system can be written in thw previous two ways, but now they are completely undistinguishable, i.e. a single state:

$$|e\rangle_1 |f\rangle_2 \equiv |f\rangle_1 |e\rangle_2.$$

So if a state of N particles living in $|\psi\rangle \in \mathcal{H}$ and it is left invariant by any of the $N!$ permutations, then the particles are indistinguishable and $|\psi\rangle \in \mathcal{H}_1^{\otimes N}$:

$$|\psi\rangle \xrightarrow{P} |\psi'\rangle = P|\psi\rangle = e^{i\alpha}|\psi\rangle.$$

Let's then analyze further the *permutation group* of these transformations.

7.3.1 | Permutation Group

The set of all possible permutations P_j of N elements defines the permutation group, which respects the usual properties of a transformations group:

- Closed under a *composition operation*;
- It has an *inverse element* with respect to the composition law;
- It has a *neutral element* with respect to the composition law;
- It respects *associativity* under the composition law.

So if we consider two permutations P_i and P_j , then their consequent application (their composition), is still a permutation. Since there are $N!$ ways of permuting N elements, we conclude that the permutation group is made of $N!$ elements.

We can study the simplest transformation of this group, the **elementary permutation** (or transposition) exchanging the i -th element with the $(i+1)$ -th one, in order to understand better this group and its properties. The elementary permutation is so important because it is the **generator** of this transformation group: we can write any other permutation by composing different elementary permutations. Thus σ_i with $i = 1, \dots, N$ are the generators of the group, but decomposing a permutation in such way is not uniquely defined:

$$\sigma = \sigma_{\alpha_1} \sigma_{\alpha_2} \dots \sigma_{\alpha_N},$$

there may be different combinations of $(\alpha_1, \alpha_2, \dots, \alpha_N)$ which give the same result. But it can be shown that a permutation made by an even (or odd) number of elementary permutations, will always be made of an even (or odd) number of elementary permutations. Furthermore, the number of transpositions defines the sign of the decomposition:

$$\text{sgn}(\sigma) = \begin{cases} +1, & \text{for an even number of transpositions,} \\ -1, & \text{for an odd number of transpositions.} \end{cases}$$

It can be shown that the generators of the permutation group satisfy the following properties:

1. $\sigma_i \sigma_j = \sigma_j \sigma_i, \forall |i - j| \geq 2$ ²
2. $\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}$;
3. $\sigma_i^2 = \mathbb{I}$.

²The condition on the indexes is due to the fact that if $|i - j| = 1$ then we are consequently applying the transposition on the same element: $\sigma_i \sigma_{i+1}$ puts the i -th element in $i+2$ slot and the $i+2$ -th in the $i+1$, while $\sigma_{i+1} \sigma_i$ does the opposite $i \rightarrow i+1$ and $i+2 \rightarrow i$, clearly different transpositions.

7.4 | Quantum Statistics

If we consider a generic state of N indistinguishable particles $\psi(1, 2, \dots, N) \in \mathcal{H}^{\otimes N}$, invariant under permutation up to a phase (which we will not assume permutation independent to be general):

$$\psi(1, 2, \dots, N) \rightarrow \psi(P(1, 2, \dots, N)) = e^{i\phi_P} \psi(1, 2, \dots, N).$$

We know that any permutation can be decomposed into a composition of transpositions

$$P = \sigma_{\alpha_1} \sigma_{\alpha_2} \dots \sigma_{\alpha_k},$$

then the action of P on ψ is the same as the consequent application of the corresponding transpositions on ψ :

$$\psi \rightarrow P\psi = \sigma_{\alpha_1} \sigma_{\alpha_2} \dots \sigma_{\alpha_k} \psi = e^{i\phi_1} e^{i\phi_2} \dots e^{i\phi_k} \psi = e^{i\phi_P} \psi$$

since we can write the action of the j -th transposition on ψ as

$$\sigma_{\alpha_j} \psi = e^{i\phi_j} \psi.$$

Thus if a permutation is decomposed into its transpositions, the phase of the permutation is given by the sum of the phases of the transpositions, i.e.

$$\phi_P = \phi_1 + \phi_2 + \dots + \phi_k.$$

We can now use the properties of the generators to exploit some properties of the phases of the transpositions:

1. $\sigma_i \sigma_j = \sigma_j \sigma_i \iff \phi_P = \phi_i + \phi_j = \phi_j + \phi_i$ clearly;
2. $\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1} \iff \phi_P = \phi_i + \phi_{i+1} + \phi_i = \phi_{i+1} + \phi_i + \phi_{i+1}, \implies \phi_i = \phi_{i+1}$, so that every transposition has the same effect on the phase of the permutation invariant vector: $\sigma_i \psi = e^{i\phi} \psi \forall i$;
3. $\sigma_i^2 = \mathbb{I} \iff (e^{i\phi})^2 = 1$, which is true for two values of ϕ : $\phi = 0, \phi = \pi$.

The last observation is the most important, giving powerful insights into the system's symmetry: $(e^{i\phi})^2 = 1 \implies e^{i\phi} = \pm 1$, so that each transposition gives a factor of 1 ($\phi = 0$) or -1 ($\phi = \pi$). But since the phase is the same for all transposition, then we have two fixed case:

- $\phi = 0$ in which every transposition leaves the state totally invariant: it is a completely symmetric system made of **Bosons**:

$$\begin{cases} \sigma_i : \psi \rightarrow \psi; \\ P : \psi \rightarrow \psi. \end{cases}$$

- $\phi = \pi$ in which every transposition changes the sign of the state: it is a completely antisymmetric system made of **Fermions**:

$$\begin{cases} \sigma_i : \psi \rightarrow -\psi; \\ P : \psi \rightarrow \text{sgn}(P)\psi. \end{cases}$$

7.4.1 | Two Particle State

If we consider N identical particles living in their generical d -dimensional Hilbert space, the total space is given by

$$\mathcal{H} = \mathcal{H}_1^{\otimes N} = \mathcal{L}^2(\mathbb{R}^{dN}).$$

Let us restrict the case to $N = 2$, then $\mathcal{H} = \mathcal{H}_1^{\otimes 2} = \mathcal{L}^2(\mathbb{R}^{2d})$. So we can consider states composed by the singular substates of the two particles such that, if they are **distinguishable**

$$\psi_1(x_1), \psi_2(x_2) \implies \psi = \psi_1(x_1)\psi_2(x_2) \neq \psi_1(x_2)\psi_2(x_1) \in \mathcal{H}^{\otimes 2}.$$

In general the last equation does not hold, they are different states, but if we consider **undistinguishable** particles then we have *total symmetry* or *total antisymmetry* depending on the particles' nature.

We can introduce two operators, \hat{S} and \hat{A} which project our state onto the totally antisymmetric or symmetric subspaces as follows:

$$\begin{aligned}\hat{S}\psi &= \psi_S(x_1, x_2) = \frac{\psi(x_1, x_2) + \psi(x_2, x_1)}{2} = \frac{\psi_1(x_1)\psi_2(x_2) + \psi_1(x_2)\psi_2(x_1)}{2} = \psi_S(x_2, \psi_1); \\ \hat{A}\psi &= \psi_A(x_1, x_2) = \frac{\psi(x_1, x_2) - \psi(x_2, x_1)}{2} = \frac{\psi_1(x_1)\psi_2(x_2) - \psi_1(x_2)\psi_2(x_1)}{2} = -\psi_A(x_2, \psi_1).\end{aligned}$$

Thus any state can be symmetrized or antisymmetrized with the action of these operators. Notice how, being projectors, these operators respect

$$\begin{aligned}\hat{S} &= \hat{S}^\dagger, \quad \hat{S}^2 = \hat{S}; \\ \hat{A} &= \hat{A}^\dagger, \quad \hat{A}^2 = \hat{A}.\end{aligned}$$

Another very important property is the orthogonality of these operators: they commute and any consequent application of both of them on a state is projected in the null space:

$$\hat{S}\hat{A} = \hat{A}\hat{S} = 0, \quad \langle \psi_S | \psi_A \rangle = \langle \psi_A | \psi_S \rangle = 0.$$

Thus we can decompose our general Hilbert space into two direct sum subspaces:

$$\mathcal{H} = \mathcal{H}_S \oplus_{\perp} \mathcal{H}_A, \quad \implies \quad \psi = \psi_S + \psi_A.$$

Thus the two subspaces are orthogonal and they share only the null vector.

Pauli exclusion principle. It's interesting to notice how the important **Pauli exclusion principle** arise from here, only by applying statistical prescription: if we consider two fermions in the same state

$$\psi(x_1, x_2) = \psi_1(x_1)\psi_1(x_2),$$

we know this state to be antisymmetric, so if we apply the \hat{A} operators we are not changing the state nor projecting it into a subspace, it's the same; then by applying it we notice how the wavefunction vanishes:

$$\hat{A}\psi(x_1, x_2) = \frac{\psi_1(x_1)\psi_1(x_2) - \psi_1(x_2)\psi_1(x_1)}{2} = 0.$$

Thus the principle: *two fermions cannot occupy the same physical state.*

7.4.2 | N Particle State

For a generic number of indistinguishable particle, we had $\mathcal{H} = \mathcal{H}_1^{\otimes N}$, with states of the form $\psi(x_1, \dots, x_N) = \psi_1(x_1)\psi_2(x_2) \dots \psi_N(x_N)$, which can be written in a more convenient form

$$|\psi\rangle = |\psi_1\rangle_1 |\psi_2\rangle_2 \dots |\psi_N\rangle_N \in \mathcal{H}.$$

The index ψ_j refers to a particle state, a set of quantum numbers defining it, while the index on $|\rangle_j$ refers to the particle itself. Thus we can express the action of the symmetry projectors as:

$$\begin{aligned}\hat{S}: |\psi\rangle &= \frac{1}{N!} \sum_{\sigma} |\psi_{\sigma(1)}\rangle_1 |\psi_{\sigma(2)}\rangle_2 \dots |\psi_{\sigma(N)}\rangle_N \in \mathcal{H}_S; \\ \hat{A}: |\psi\rangle &= \frac{1}{N!} \sum_{\sigma} \text{sgn}(\sigma) |\psi_{\sigma(1)}\rangle_1 |\psi_{\sigma(2)}\rangle_2 \dots |\psi_{\sigma(N)}\rangle_N \in \mathcal{H}_A.\end{aligned}$$

These operators still respects the properties of projectors and they still commute: we can decompose the Hilbert space in:

$$\mathcal{H} = \mathcal{H}_S \oplus_{\perp} \mathcal{H}_A \oplus_{\perp} \mathcal{H}',$$

where the last subspace \mathcal{H}' contains non physical states and solutions: we can think and write of a neither symmetric nor antisymmetric state, but it is non physical as all particles are fermions or bosons.

We will work in a space where we consider a variable number of particles, as we will describe in detail in the next chapter, direct sum of the spaces for every possible number of particle:

$$\mathcal{H}_{S/A} = \bigoplus_{N=0}^{\infty} \mathcal{H}_{S/A}^{(N)},$$

aiming now to the intrinsic description of each Hilbert space for a fixed N : $\mathcal{H}_{S/A}^{(N)} = \mathcal{H}_1^{\otimes N}$. We need an arbitrary ON basis, which we will construct from the basis of the single particle space:

$$\mathcal{H}_1: \{|\psi_n\rangle\}_{n=1}^{\infty}.$$

Thus if we combine N of this spaces (and basis), we find a basis which spans the whole N particle space $\mathcal{H}^{\otimes N} = \mathcal{H}_{S/A}^{(N)}$:

$$\mathcal{H}_{S/A}^{(N)}: \{|\psi_{\alpha_1}\rangle |\psi_{\alpha_2}\rangle \dots |\psi_{\alpha_N}\rangle\}_{\alpha_1, \alpha_2, \dots, \alpha_N=1}^{\infty}.$$

Remembering that we are speaking of undistinguishable particles, we can now expose a new description of our states: since we are not interested in the position and state of each particle, we consider just the states and their population, using the **occupation number** n_k : we have an infinite amount of occupation numbers (infinite dimensionality of Hilbert space for one particle), each of whose can assume a value between 0 and N :

$$0 < n_1, \dots, n_k, \dots < N, \quad n_1 + \dots + n_k + \dots = N,$$

their sum has to be equal to the number of particles.

We can express a state

$$\psi_{n_1, \dots, n_k, \dots}(x_1, x_2, \dots, x_N)$$

in the base $\{|\psi_{\alpha_1}\rangle |\psi_{\alpha_2}\rangle \dots |\psi_{\alpha_N}\rangle\}_{\alpha_1, \alpha_2, \dots, \alpha_N=1}^{\infty}$ by setting accordingly the pedices, such that n_1 pedices will be equal to one ($\alpha_j = 1$), n_k pedices will be equal to k ($\alpha_j = k$) eccetera.

This occupation numbers behave differently with respect to the particles' nature:

Bosons: $n_k = 0, 1, 2, \dots, \infty;$

Fermions: $n_k = 0, 1;$

this is due to the Pauli exclusion principle: there cannot exist two fermions in the same state, while bosons do not have such kind of restrictions over occupation numbers.

8 | Second Quantization

In this chapter, we introduce the formalism of second quantization, which provides a powerful framework for describing many-particle quantum systems. This approach is particularly useful for systems of identical particles, such as bosons and fermions, where the symmetrization or antisymmetrization of the wavefunction plays a crucial role.

We will begin by defining creation and annihilation operators, which allow us to add or remove particles from specific quantum states. We will then construct the Fock space, which is the Hilbert space that accommodates varying numbers of particles. Following this, we will introduce field operators that enable us to describe particle fields in a continuous manner.

Finally, we will discuss observable operators in the context of second quantization and provide several examples, including the density operator, number operator, and the free Hamiltonian for non-interacting particles.

8.1 | Creation and Annihilation Operators

When dealing with many-particle systems, it is often convenient to use creation and annihilation operators. These operators facilitate the addition and removal of particles from specific quantum states. They obey different commutation or anticommutation relations depending on the nature of particles. We will discuss both bosonic and fermionic cases.

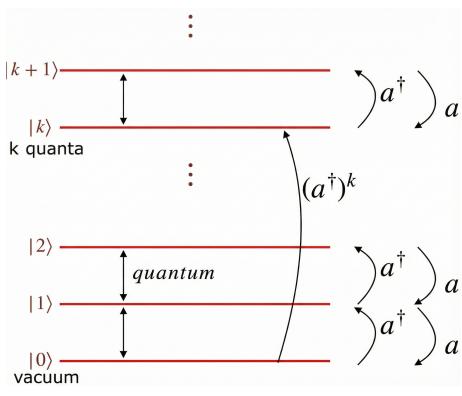
8.1.1 | Bosonic Case

We will first consider bosons, which are particles that obey Bose-Einstein statistics. The **creation operator** \hat{a}^\dagger adds a particle to the quantum state, while the **annihilation operator** \hat{a} removes a particle from that state: they add or remove a quantum of energy when applied to the wavefunction. These operators satisfy the following commutation relations:

$$[\hat{a}, \hat{a}^\dagger] = \mathbb{I}, \quad [\hat{a}, \hat{a}] = [\hat{a}^\dagger, \hat{a}^\dagger] = 0, \quad (8.1.1)$$

where \mathbb{I} is the identity operator. These relations reflect the indistinguishable nature of bosons and their ability to occupy the same quantum state. Note that these operators are not Hermitian, as they do not correspond to observable quantities directly.

Before showing how these operators act on states, we use them to define the **number operator** \hat{N} :



$$\hat{N} = \hat{a}^\dagger \hat{a}, \quad (8.1.2)$$

which counts the number of particles in a given quantum state. The action of the creation and annihilation operators on the number states $|n\rangle$ is given by:

$$\begin{aligned} \hat{N} |n\rangle &= n |n\rangle, \\ \hat{a} |n\rangle &= \sqrt{n} |n-1\rangle, \\ \hat{a}^\dagger |n\rangle &= \sqrt{n+1} |n+1\rangle. \end{aligned}$$

We also have the following useful commutation relations involving the number operator:

$$[\hat{N}, \hat{a}] = -\hat{a}, \quad [\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger.$$

Example. Consider a simple harmonic oscillator, which can be described using bosonic creation and annihilation operators since the excitations of the oscillator correspond to quanta of energy, thus scalar particles. When considering the Hilbert space $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$ of a quantum harmonic oscillator, the creation and annihilation operators can be expressed in terms of the position \hat{x} and momentum \hat{p} operators as:

$$\hat{a} = \frac{1}{\sqrt{2}} (\hat{x} + i\hat{p}), \quad \hat{a}^\dagger = \frac{1}{\sqrt{2}} (\hat{x} - i\hat{p}).$$

from $[\hat{x}, \hat{p}] = i\hbar$. Then we could also write the hamiltonian operator as:

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) = \hbar\omega \left(\hat{N} + \frac{1}{2} \right).$$

The energy eigenvalues of the harmonic oscillator are then given by:

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots$$

since the number operator \hat{N} has eigenvalues n .

This example starts to illustrate how creation and annihilation operators can be used to describe quantum systems in terms of quantized energy levels. This formalism is particularly powerful when extended to many-particle systems, where the number of particles can vary. We will soon use the ladder operators to build the Fock space representation of many-particle states from an orthonormal basis.

8.1.2 | Fermionic Case

Next, we consider fermions, which are particles that obey Fermi-Dirac statistics. The creation operator \hat{a}^\dagger adds a fermion to a quantum state, while the annihilation operator \hat{a} removes a fermion from that state. These operators satisfy the following anticommutation relations:

$$\{\hat{a}, \hat{a}^\dagger\} = \mathbb{I}, \quad \{\hat{a}, \hat{a}\} = \{\hat{a}^\dagger, \hat{a}^\dagger\} = 0, \quad (8.1.3)$$

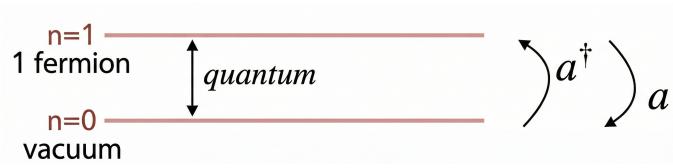
where \mathbb{I} is the identity operator. These relations reflect the Pauli exclusion principle, which states that no two fermions can occupy the same quantum state simultaneously.

Similar to the bosonic case, we define the number operator \hat{N} for fermions:

$$\hat{N} = \hat{a}^\dagger \hat{a}, \quad (8.1.4)$$

which counts the number of fermions in a given quantum state.

The action of the creation and annihilation operators on the number states $|n\rangle$ (where $n = 0$ or 1 due to the Pauli exclusion principle, eigenstates of the number operator) is given by:



$$\begin{aligned} \hat{N} |n\rangle &= n |n\rangle, \\ \hat{a} |1\rangle &= |0\rangle, \quad \hat{a} |0\rangle = 0, \\ \hat{a}^\dagger |0\rangle &= |1\rangle, \quad \hat{a}^\dagger |1\rangle = 0, \end{aligned}$$

since a fermionic state can either be unoccupied ($n = 0$) or occupied by a single fermion ($n = 1$): from the null anticommutation relations among creation operators we have

$$\{\hat{a}^\dagger, \hat{a}^\dagger\} = 0 \implies (\hat{a}^\dagger)^2 = 0,$$

We also have the same commutation relations involving the number operator:

$$[\hat{N}, \hat{a}] = -\hat{a}, \quad [\hat{N}, \hat{a}^\dagger] = \hat{a}^\dagger.$$

Example. If we consider a system of spin- $\frac{1}{2}$ fermions in $\mathcal{H} = \mathbb{C}^2$, such as electrons, the creation and annihilation operators can be defined with Pauli spin matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

respecting $[\sigma^i, \sigma^j] = \epsilon_{ijk}\sigma^k$.

For example, the ladder operators for an electron with spin up can be expressed as $\sigma^{(\pm)} = \frac{1}{2}(\sigma_x \pm i\sigma_y)$, such that

$$\begin{aligned} \sigma^{(+)} &= (\sigma^{(-)})^\dagger, \\ (\sigma^{(+)})^2 &= (\sigma^{(-)})^2 = 0, \\ \{\sigma^{(+)}, \sigma^{(-)}\} &= \mathbb{I}, \\ \{\sigma^{(+)}, \sigma^{(+)}\} &= \{\sigma^{(-)}, \sigma^{(-)}\} = 0. \end{aligned}$$

Thus the creation operator $\sigma^{(+)}$ adds an electron with spin up to the state, while the annihilation operator $\sigma^{(-)}$ removes an electron with spin up from the state. The number operator $\hat{N} = \sigma^{(+)}\sigma^{(-)}$ counts the number of electrons with spin up in the state, which can be either 0 or 1 due to the Pauli exclusion principle.

Thus, we have established the fundamental properties of creation and annihilation operators for both bosons and fermions. These operators form the basis for constructing the Fock space, which we will explore in the next section.

8.2 | Fock Space

The Fock space is a Hilbert space that accommodates states with varying numbers of particles. It is constructed as the direct sum of tensor products of single-particle Hilbert spaces. For a single-particle Hilbert space $\mathcal{H}^{(1)}$, the Fock space $\mathcal{F}(\mathcal{H})$ is defined as:

$$\mathcal{F}(\mathcal{H}) = \bigoplus_{N=0}^{\infty} \mathcal{H}^{(N)}, \quad (8.2.1)$$

If we describe a system of identical particles, the N -particle Hilbert space $\mathcal{H}^{(N)}$ is constructed as the symmetric (for bosons) or antisymmetric (for fermions) tensor product of N copies of the single-particle Hilbert space:

$$\mathcal{H}_{S/A}^{(N)} = |n_1, n_2, \dots, n_N\rangle, \text{ where } \begin{cases} S: & \text{Symmetric for bosons, } n_k = 0, 1, 2, \dots \\ A: & \text{Antisymmetric for fermions, } n_k = 0, 1, \end{cases}$$

where n_k represents the occupation number of the k -th quantum state and $\mathcal{H} = \mathcal{H}_A \oplus_{\perp} \mathcal{H}_S$. These results are a direct consequence of the **canonical commutation and anticommutation relations** satisfied by the creation and annihilation operators, and they are encoded in the structure of the Fock space itself.

If we denote the occupation number basis states in Fock space as $\{|n_1, n_2, \dots, n_k, \dots\rangle\}_{n_1, n_2, \dots}$, where n_k is the number of particles in the k -th single-particle state, we can define a couple of creation and annihilation operators \hat{a}_n^\dagger and \hat{a}_n for each single-particle state n (which in the single particle Hilbert space corresponds to the n -th vector of the basis $\{|e_n\rangle\}_n$). As we said, these operators induce the statistical nature of the particles, so they satisfy either the commutation relations (8.1.1) for bosons or the anticommutation relations (8.1.3) for fermions:

$$\begin{cases} \text{Bosons:} & [\hat{a}_m, \hat{a}_n^\dagger] = \delta_{mn} \mathbb{I}, \quad [\hat{a}_m, \hat{a}_n] = [\hat{a}_m^\dagger, \hat{a}_n^\dagger] = 0, \\ \text{Fermions:} & \{\hat{a}_m, \hat{a}_n^\dagger\} = \delta_{mn} \mathbb{I}, \quad \{\hat{a}_m, \hat{a}_n\} = \{\hat{a}_m^\dagger, \hat{a}_n^\dagger\} = 0, \end{cases}$$

which can be compacted into the single relation

$$[\hat{a}_n, \hat{a}_m^\dagger]_{\pm} = \delta_{mn} \mathbb{I}, \quad [\hat{a}_n, \hat{a}_m]_{\pm} = [\hat{a}_n^\dagger, \hat{a}_m^\dagger]_{\pm} = 0, \quad (8.2.2)$$

where the $+$ sign corresponds to fermions (anticommutator) and the $-$ sign corresponds to bosons (commutator).

The algebra generated by the fermionic CCR encodes the Pauli exclusion principle, while the bosonic CCR allows for multiple occupancy of the same state. The vacuum state $|0\rangle$ in Fock space is defined as the state with no particles, satisfying:

$$\hat{a}_n |0\rangle = 0, \quad \forall n.$$

The action of the creation and annihilation operators on the occupation number basis states can be used to build up or reduce the number of particles in each state: we can create from the vacuum state any occupation number state by applying the creation operators the appropriate number of

times

$$\begin{aligned}\mathcal{H}_{S/A}^{(0)} &= \{\lambda |0\rangle \quad \lambda \in \mathbb{C}\}, \\ \mathcal{H}_{S/A}^{(1)} &= \left\{ |0, \dots, n_k = 1, 0, \dots\rangle = \hat{a}_k^\dagger |0\rangle \right\} \sim |e_k\rangle, \\ \mathcal{H}_{S/A}^{(2)} &= \left\{ |0, \dots, n_i = 1, \dots, n_j = 1, 0, \dots\rangle = \hat{a}_i^\dagger \hat{a}_j^\dagger |0\rangle \right\} \sim |e_i\rangle \otimes_{S/A} |e_j\rangle, \\ &\vdots \\ \mathcal{H}_{S/A}^{(N)} &= \left\{ |n_1, n_2, \dots, n_k, \dots\rangle = \prod_k \frac{(\hat{a}_k^\dagger)^{n_k}}{\sqrt{n_k!}} |0\rangle \right\} \sim \Psi_{n_1, n_2, \dots, n_k, \dots},\end{aligned}$$

where the last line represents a general N -particle state with occupation numbers n_k for each single-particle state k , thus has to respect $\sum_k n_k = N$ and the normalization factor $\frac{1}{\sqrt{n_k!}}$ is included for bosons to account for multiple occupancy (it is omitted for fermions since n_k can only be 0 or 1) since we are treating indistinguishable particles.¹

Two particle state. If we focus for a moment on the two particle case, we can see how the symmetrization or antisymmetrization of the states arises naturally from the commutation or anticommutation relations of the creation operators. For two particles in states i and j , we have:

$$\mathcal{H}_{S/A}^{(2)} = \left\{ \hat{a}_i^\dagger \hat{a}_j^\dagger |0\rangle, \quad \forall i, j \right\} \sim \mathcal{H} \otimes_{S/A} \mathcal{H}.$$

We are considering explicitly:

- one particle in state $|e_i\rangle$, so that $n_i = 1$;
- one particle in state $|e_j\rangle$, so that $n_j = 1$;
- all other states unoccupied, so that $n_k = 0$ for $k \neq i, j$.

Thus, we have to write the two-particle states in a symmetrized or antisymmetrized form, depending on whether we are dealing with bosons or fermions:

$$\begin{aligned}|1_i, 1_j\rangle_S &= \hat{a}_i^\dagger \hat{a}_j^\dagger |0\rangle = \hat{a}_j^\dagger \hat{a}_i^\dagger |0\rangle = |1_j, 1_i\rangle_S, \\ |1_i, 1_j\rangle_A &= \hat{a}_i^\dagger \hat{a}_j^\dagger |0\rangle = -\hat{a}_j^\dagger \hat{a}_i^\dagger |0\rangle = -|1_j, 1_i\rangle_A,\end{aligned}$$

which shows that the bosonic state is symmetric under particle exchange, while the fermionic state is antisymmetric. This construction can be extended to any number of particles, since the symmetrization or antisymmetrization arises naturally from the commutation or anticommutation relations of the creation operators, leading to the full Fock space representation of many-particle quantum systems.

8.2.1 | Properties of Fock Space

We can summarize some important properties of Fock space, which were already hinted at in the previous discussion:

1. **Occupation Number Basis:** The occupation number basis states $|n_1, n_2, \dots\rangle$ form a complete orthonormal basis for Fock space:

$$\{|n_1, n_2, \dots\rangle\}_{n_1, n_2, \dots} \implies \langle n'_1, n'_2, \dots | n_1, n_2, \dots \rangle = \delta_{n'_1 n_1} \delta_{n'_2 n_2} \dots$$

¹But this factor comes out algebraically from the action of the creation operator, in order to simplify its eigenvalues which would be multiplying the state.

2. **Annihilation:** The annihilation operator \hat{a}_k acting on an occupation number state $|n_1, n_2, \dots, n_k, \dots\rangle$ decreases the occupation number of the k -th state by one, such that $\hat{a}_k : \mathcal{H}_{S/A}^{(N)} \rightarrow \mathcal{H}_{S/A}^{(N-1)}$

$$\begin{aligned} B) \quad & \hat{a}_k |n_1, n_2, \dots, n_k, \dots\rangle = \sqrt{n_k} |n_1, n_2, \dots, n_k - 1, \dots\rangle, \\ F) \quad & \hat{a}_k |n_1, n_2, \dots, n_k, \dots\rangle = \eta_k \sqrt{n_k} |n_1, n_2, \dots, n_k - 1, \dots\rangle, \end{aligned}$$

where $\eta_k = (-1)^{\sum_{j < k} n_j}$ accounts for the sign change due to the antisymmetry of fermionic states.

3. **Creation:** The creation operator \hat{a}_k^\dagger acting on an occupation number state $|n_1, n_2, \dots, n_k, \dots\rangle$ increases the occupation number of the k -th state by one, such that $\hat{a}_k^\dagger : \mathcal{H}_{S/A}^{(N)} \rightarrow \mathcal{H}_{S/A}^{(N+1)}$

$$\begin{aligned} B) \quad & \hat{a}_k^\dagger |n_1, n_2, \dots, n_k, \dots\rangle = \sqrt{n_k + 1} |n_1, n_2, \dots, n_k + 1, \dots\rangle, \\ F) \quad & \hat{a}_k^\dagger |n_1, n_2, \dots, n_k, \dots\rangle = \eta_k \sqrt{n_k + 1} |n_1, n_2, \dots, n_k + 1, \dots\rangle. \end{aligned}$$

4. **Particle Number Operator:** The total particle number operator $\hat{N} = \sum_k \hat{n}_k = \sum_k \hat{a}_k^\dagger \hat{a}_k$ counts the total number of particles in the system:

$$\hat{N} |n_1, n_2, \dots\rangle = \left(\sum_k n_k \right) |n_1, n_2, \dots\rangle = N |n_1, n_2, \dots\rangle.$$

To build a generic state $|f\rangle$ in Fock space, we can take linear combinations of the occupation number basis states:

$$|f\rangle = \sum_{n_1, n_2, \dots} f(n_1, n_2, \dots) |n_1, n_2, \dots\rangle,$$

where $f(n_1, n_2, \dots)$ are complex coefficients that determine the contribution of each occupation number state to the overall state $|f\rangle$. But since any state in Fock space can be constructed by applying creation operators to the vacuum state, we can also express $|f\rangle$ as:

$$|f\rangle = \sum_{n_1, n_2, \dots} f(n_1, n_2, \dots) \left[(\hat{a}_1^\dagger)^{n_1} (\hat{a}_2^\dagger)^{n_2} \dots \right] |0\rangle.$$

8.3 | Field Operators

In the context of second quantization, field operators provide a way to describe particle fields in a continuous manner, rather than focusing on individual particles. These operators are constructed from the creation and annihilation operators defined in the Fock space framework.

Let us consider a Fock state built from a single-particle Hilbert space $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^d)$ with an orthonormal basis $\{|e_n\rangle\}_n$. Each of these basis states corresponds to a specific quantum state of a single particle, such as a momentum eigenstate, which can be constructed as:

$$|e_n(\mathbf{k}_n)\rangle = |e_n\rangle = \hat{a}_n^\dagger |0\rangle.$$

Thus we could think of a generic Fock state as:

$$|f\rangle = \sum_n f_n |e_n\rangle = \sum_n f_n \hat{a}_n^\dagger |0\rangle = \hat{\psi}^\dagger(f) |0\rangle,$$

where we have defined the **field creation operator** $\hat{\psi}^\dagger(f)$ as:

$$\hat{\psi}^\dagger(f) = \sum_n f_n \hat{a}_n^\dagger. \quad (8.3.1)$$

and similarly the **field annihilation operator** $\hat{\psi}(f)$ as:

$$\hat{\psi}(f) = \sum_n f_n^* \hat{a}_n. \quad (8.3.2)$$

As we will see, these field operators play a crucial role in describing many-particle systems in a continuous manner, allowing us to express observables and interactions in terms of fields rather than individual particles. They also satisfy commutation or anticommutation relations depending on whether we are dealing with bosons or fermions:

$$[\hat{\psi}(f), \hat{\psi}^\dagger(g)]_\mp = \left[\sum_n f_n^* \hat{a}_n, \sum_m g_m \hat{a}_m^\dagger \right]_\mp = \sum_{n,m} f_n^* g_m [\hat{a}_n, \hat{a}_m^\dagger]_\mp = \sum_n f_n^* g_n = \langle f | g \rangle,$$

where the $-$ sign corresponds to bosons (commutator) and the $+$ sign corresponds to fermions (anticommutator). The other commutators or anticommutators vanish:

$$[\hat{\psi}(f), \hat{\psi}(g)]_\mp = 0, \quad [\hat{\psi}^\dagger(f), \hat{\psi}^\dagger(g)]_\mp = 0.$$

If we choose the basis states $\{|e_n\rangle\}_n$ to be position eigenstates, we can express the field operators in terms of plane waves or localized states:

$$\{|e_n\rangle = u_n(\mathbf{x})\}_{n=1}^\infty,$$

such that the **plane waves basis vectors** are given by:

$$u_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}},$$

where \mathbf{k} is the wavevector associated with the momentum of the particle. Note that these functions form a complete orthonormal set in $\mathcal{L}^2(\mathbb{R}^d)$, but for now we are ignoring normalization factors for simplicity.

The **field operators** $\hat{\psi}(\mathbf{x})$ and its conjugate $\hat{\psi}^\dagger(\mathbf{x})$ ² can be thus expressed in terms of these basis functions as a Fourier expansion, where the Fourier coefficients are the annihilation and creation operators for the momentum eigenstates:

$$\hat{\psi}(\mathbf{x}) = \sum_n u_n(\mathbf{x}) \hat{a}_n = \sum_{\mathbf{k}_n} e^{i\mathbf{k}_n \cdot \mathbf{x}} \hat{a}_{\mathbf{k}_n}, \quad \hat{\psi}^\dagger(\mathbf{x}) = \sum_n u_n^*(\mathbf{x}) \hat{a}_n^\dagger = \sum_{\mathbf{k}_n} e^{-i\mathbf{k}_n \cdot \mathbf{x}} \hat{a}_{\mathbf{k}_n}^\dagger, \quad (8.3.3)$$

giving us a way to describe the field operators in real space: the localized excitations of the particle field at position \mathbf{x} are given by a superposition of plane waves with different momenta, weighted by the corresponding creation and annihilation operators.

We can now express the previously defined field creation and annihilation operators $\hat{\psi}^\dagger(f)$ and $\hat{\psi}(f)$ in terms of the position-space field operators:

$$\begin{aligned} \hat{\psi}^\dagger(f) &= \sum_n f_n \hat{a}_n^\dagger = \int d^d \mathbf{x} \hat{\psi}^\dagger(\mathbf{x}) f(\mathbf{x}), \\ \hat{\psi}(f) &= \sum_n f_n^* \hat{a}_n = \int d^d \mathbf{x} \hat{\psi}(\mathbf{x}) f^*(\mathbf{x}), \end{aligned}$$

where $f(\mathbf{x}) = \sum_n f_n u_n(\mathbf{x})$ is a test function that describes the spatial profile of the particle field. We can obtain this results by substituting the expressions for $\hat{\psi}(\mathbf{x})$ and $\hat{\psi}^\dagger(\mathbf{x})$ from equation (8.3.3) into the integrals and using the orthonormality of the basis functions $u_n(\mathbf{x})$:

$$\begin{aligned} \hat{\psi}^\dagger(f) &= \int d^d \mathbf{x} \hat{\psi}^\dagger(\mathbf{x}) f(\mathbf{x}) \\ &= \int d^d \mathbf{x} \left(\sum_n u_n^*(\mathbf{x}) \hat{a}_n^\dagger \right) \left(\sum_m f_m u_m(\mathbf{x}) \right) \\ &= \sum_{n,m} f_m \hat{a}_n^\dagger \int d^d \mathbf{x} u_n^*(\mathbf{x}) u_m(\mathbf{x}) = \sum_n f_n \hat{a}_n^\dagger, \end{aligned}$$

and similarly for $\hat{\psi}(f)$. We can now see that the field operators $\hat{\psi}(\mathbf{x})$ and $\hat{\psi}^\dagger(\mathbf{x})$ serve as the building blocks for constructing the field creation and annihilation operators $\hat{\psi}(f)$ and $\hat{\psi}^\dagger(f)$ through integration over space.

Now that we have defined the field operators in position space, we can also express their commutation or anticommutation relations:

$$[\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})]_\mp = \delta(\mathbf{x} - \mathbf{y}), \quad [\hat{\psi}(\mathbf{x}), \hat{\psi}(\mathbf{y})]_\mp = 0, \quad [\hat{\psi}^\dagger(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})]_\mp = 0,$$

where the $-$ sign corresponds to bosons (commutator) and the $+$ sign corresponds to fermions (anticommutator). We can derive the only non-vanishing relation as follows:

$$\begin{aligned} [\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})]_\mp &= \sum_{n,m} u_n(\mathbf{x}) u_m^*(\mathbf{y}) [\hat{a}_n, \hat{a}_m^\dagger]_\mp \\ &= \sum_{n,m} u_n(\mathbf{x}) u_m^*(\mathbf{y}) \delta_{nm} = \sum_n u_n(\mathbf{x}) u_n^*(\mathbf{y}) \\ &= \sum_n e^{i\mathbf{k}_n \cdot (\mathbf{x} - \mathbf{y})} = \delta(\mathbf{x} - \mathbf{y}), \end{aligned}$$

²They are different from the field creation and annihilation operators defined earlier, as they depend on the continuous position variable \mathbf{x} and conceptually represent the annihilation and creation of a particle at a specific point in space.

but we can obtain the delta function even from another request:

$$\begin{aligned} [\hat{\psi}(f), \hat{\psi}^\dagger(g)]_\mp &= \int d^d\mathbf{x} \int d^d\mathbf{y} f^*(\mathbf{x})g(\mathbf{y}) [\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})]_\mp \\ &= \langle f|g\rangle = \int d^d\mathbf{x} f^*(\mathbf{x})g(\mathbf{x}), \end{aligned}$$

thus to satisfy both we need $[\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{y})]_\mp = \delta(\mathbf{x} - \mathbf{y})$. These relations reflect the local nature of the field operators, indicating that the creation and annihilation of particles at different points in space are independent processes.

Remark. Note that the field operators $\hat{\psi}(\mathbf{x})$ and $\hat{\psi}^\dagger(\mathbf{x})$ can be interpreted as operator-valued distributions rather than ordinary operators, due to their dependence on continuous position variables. This means that they are not well-defined at individual points in space, but rather when integrated against suitable test functions $f(\mathbf{x})$ and $g(\mathbf{x})$.

Furthermore they do not depend upon the choice of basis $\{u_n(\mathbf{x})\}_n$ used to express them, as long as the basis is complete and orthonormal in $L^2(\mathbb{R}^d)$:

$$\hat{\psi}^\dagger(\mathbf{x}) = \sum_n u_n^*(\mathbf{x}) \hat{a}_n^\dagger = \sum_m v_m^*(\mathbf{x}) \hat{b}_m^\dagger.$$

Canonical Commutation Relation Algebra

The field operators $\hat{\psi}(\mathbf{x})$ and $\hat{\psi}^\dagger(\mathbf{x})$ satisfy the canonical commutation or anticommutation relations, which form the basis of the algebraic structure of quantum fields. These relations are the same as those satisfied by the creation and annihilation operators in the single-particle Hilbert space, but they are now expressed in terms of continuous position variables.

For the single-particle Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d)$, with orthonormal complete basis

$$\{|e_n\rangle\}_n \iff \{\hat{a}_n, \hat{a}_n^\dagger\}_n,$$

the canonical commutation or anticommutation relations for the ladder operators are given by:

$$\begin{cases} [\hat{a}_m, \hat{a}_n^\dagger]_\mp = \mathbb{I}\delta_{mn}, \\ [\hat{a}_m, \hat{a}_n]_{mp} = 0, \\ [\hat{a}_m^\dagger, \hat{a}_n^\dagger]_{mp} = 0. \end{cases}$$

We can now build the Fock space $\mathcal{F}(\mathcal{H})$ from the single-particle Hilbert space \mathcal{H} as:

$$\mathcal{F}(\mathcal{H}) = \bigoplus_{N=0}^{\infty} \mathcal{H}_{S/A}^{(N)},$$

where $\mathcal{H}_{S/A}^{(N)}$ is the symmetric (for bosons) or antisymmetric (for fermions) tensor product of N copies of the single-particle Hilbert space. Thus we can have any particle number state $|n_1, n_2, \dots\rangle$ in Fock space, built from the vacuum state $|0\rangle$ as:

$$|n_1, n_2, \dots\rangle = \mathcal{N} \prod_k (\hat{a}_k^\dagger)^{n_k} |0\rangle,$$

where \mathcal{N} is a normalization factor and n_k is the occupation number of the k -th single-particle state (how many particles occupy that same state, how many single particle Hilbert spaces we have to tensor multiply to obtain a state with that occupation).

We know that the occupation numbers n_k can take any non-negative integer value for bosons, while for fermions they can only be 0 or 1 due to the Pauli exclusion principle:

$$\begin{cases} \text{Bosons: } & n_k = 0, 1, 2, \dots \\ \text{Fermions: } & n_k = 0, 1, \end{cases}$$

where the total number of particles N in the system is given by the sum of the occupation numbers. If we are in a system with a fixed number of particles N , we can restrict ourselves to the N -particle subspace $\mathcal{H}_{S/A}^{(N)}$ of Fock space: we would be working in the **canonical ensemble**. However, Fock space allows us to consider states with varying numbers of particles, which is particularly useful in quantum field theory and many-body physics, where particle number is not necessarily conserved, thus we work in the **grand canonical ensemble**.

Let us now study how to represent observables in Fock space using the field operators defined earlier.

8.4 | Observable Operators

Observables in quantum mechanics are represented by Hermitian operators acting on the Hilbert space of the system. In the context of Fock space, we can express these observables in terms of the field operators $\hat{\psi}(\mathbf{x})$ and $\hat{\psi}^\dagger(\mathbf{x})$. These operators allow us to describe physical quantities such as particle number, energy, and momentum in a many-particle system.

Let us start from the single-particle operator. Given a **single-particle observable** represented by a Hermitian operator $\hat{O}^{(1)}$ acting on the single-particle Hilbert space $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^d)$, we can express a **operator on the Fock space** as:

$$\hat{O} = \sum_{j=1}^N \hat{O}^{(1)}(\mathbf{x}_j, \mathbf{p}_j), \quad (8.4.1)$$

where $\mathbf{x}_j, \mathbf{p}_j \in \mathbb{R}^d$ are the position and momentum operators for the j -th particle, and N is the total number of particles in the system. An handy example of such operator is the free Hamiltonian of a system of non-interacting particles:

$$\hat{H} = \sum_{j=1}^N H^{(1)}(\mathbf{x}_j, \mathbf{p}_j) = \sum_{j=1}^N \left(\frac{\mathbf{p}_j^2}{2m} + V(\mathbf{x}_j) \right),$$

where $H^{(1)}(\mathbf{x}_j, \mathbf{p}_j)$ is the single-particle Hamiltonian for the j -th particle, consisting of the kinetic energy term $\frac{\mathbf{p}_j^2}{2m}$ and the potential energy term $V(\mathbf{x}_j)$ (depending solely on position of the current particle). To move onto the Fock space representation, we have to keep in mind that the generic state in Fock space can be expressed as a superposition of occupation number states and has to be properly symmetrized or antisymmetrized depending on whether we are dealing with bosons or fermions.

Let us consider the single-particle observable $\hat{O}^{(1)}$ in the Fock space with plane wave basis states $\{u_n(\mathbf{k}_n)\}_n$. We can decompose the operator $\hat{O}^{(1)}$ in this basis as:

$$\hat{O}^{(1)}(\mathbf{x}_j, \mathbf{p}_j) u_{\mathbf{k}}(\mathbf{x}) = \epsilon(\mathbf{k}) u_{\mathbf{k}}(\mathbf{x}),$$

where $\epsilon(\mathbf{k}) \in \mathbb{R}$ is the real eigenvalue associated with the eigenstate $u_{\mathbf{k}}(\mathbf{x})$: the single-particle operator has to be Hermitian. Thus, we can express a generic state in Fock space as:

$$\psi_{n_1, n_2, \dots, n_N}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \mathcal{N} \left(\hat{S}/\hat{A} \right) u_{\alpha_1}(\mathbf{x}_1) u_{\alpha_2}(\mathbf{x}_2) \dots u_{\alpha_N}(\mathbf{x}_N) \in \mathcal{H}_{S/A}^{(N)},$$

where \hat{S}/\hat{A} indicates the symmetrization or antisymmetrization operator for bosons or fermions respectively. Let us now see how an operator on the Fock space \hat{O} acts on this state as a sum of single-particle operators:

$$\begin{aligned} \hat{O} \psi_{n_1, n_2, \dots, n_N}(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N) &= \sum_{j=1}^N \hat{O}^{(1)}(\mathbf{x}_j, \mathbf{p}_j) \psi_{n_1, n_2, \dots, n_N}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \\ &= \mathcal{N} \sum_{j=1}^N \left(\hat{S}/\hat{A} \right) u_{\alpha_1}(\mathbf{x}_1) \dots \epsilon(\mathbf{k}_{\alpha_j}) u_{\alpha_j}(\mathbf{x}_j) \dots u_{\alpha_N}(\mathbf{x}_N) \\ &= \left(\sum_{j=1}^N \epsilon(\mathbf{k}_{\alpha_j}) \right) \psi_{n_1, n_2, \dots, n_N}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N). \end{aligned}$$

This shows that the action of an operator \hat{O} on the many-particle state $\psi_{n_1, n_2, \dots, n_N}$ results in a sum of the eigenvalues $\epsilon(\mathbf{k}_{\alpha_j})$ of each single-particle operator, corresponding to each occupied single-particle state:

$$\sum_{j=1}^N \epsilon(\mathbf{k}_{\alpha_j}) = \sum_k n_k \epsilon(\mathbf{k}_k),$$

where n_k is the occupation number of the k -th single-particle state with eigenvalue $\epsilon(\mathbf{k}_k)$. Thus, we can express the operator \hat{O} in Fock space as:

$$\hat{O} = \sum_k \epsilon(\mathbf{k}_k) \hat{a}_k^\dagger \hat{a}_k = \sum_k \epsilon(\mathbf{k}_k) \hat{n}_k, \quad (8.4.2)$$

where $\hat{a}_k^\dagger \hat{a}_k$ is the number operator \hat{n}_k for the k -th single-particle state, counting the number of particles occupying that state:

$$\hat{n}_k |n_1, n_2, \dots\rangle = \hat{a}_k^\dagger \hat{a}_k |n_1, n_2, \dots\rangle = n_k |n_1, n_2, \dots\rangle.$$

If we work in a basis in which the single-particle operator $\hat{O}^{(1)}$ is not diagonal, for instance

$$\{v_\beta(\mathbf{x})\}_\beta \iff \hat{b}_\beta, \hat{b}_\beta^\dagger,$$

the Fock space representation of the operator \hat{O} would be constructed using the **field operators**, which can be expressed in terms of the creation and annihilation operators in the new basis:

$$\hat{\psi}'(\mathbf{x}) = \sum_\beta v_\beta(\mathbf{x}) \hat{b}_\beta, \quad \hat{\psi}'^\dagger(\mathbf{x}) = \sum_\beta v_\beta^*(\mathbf{x}) \hat{b}_\beta^\dagger.$$

We can show that any operator \hat{O} can be expressed in terms of the field operators in a general basis independent form as:

$$\begin{aligned} \int d^d \mathbf{x} \hat{\psi}'^\dagger(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) \hat{\psi}'(\mathbf{x}) &= \int d^d \mathbf{x} \left(\sum_\beta v_\beta^*(\mathbf{x}) \hat{b}_\beta^\dagger \right) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) \left(\sum_\gamma v_\gamma(\mathbf{x}) \hat{b}_\gamma \right) \\ &= \sum_{\beta, \gamma} \hat{b}_\beta^\dagger \hat{b}_\gamma \int d^d \mathbf{x} v_\beta^*(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) v_\gamma(\mathbf{x}) \\ &= \sum_{\beta, \gamma} \xi_{\beta\gamma} \hat{b}_\beta^\dagger \hat{b}_\gamma, \end{aligned}$$

while, if we choose the basis in which the single-particle observable $\hat{O}^{(1)}$ is diagonal (\hat{a} , \hat{a}^\dagger), we would recover equation (8.4.2):

$$\begin{aligned} \int d^d \mathbf{x} \hat{\psi}'^\dagger(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) \hat{\psi}'(\mathbf{x}) &= \int d^d \mathbf{x} \left(\sum_n u_n^*(\mathbf{x}) \hat{a}_n^\dagger \right) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) \left(\sum_m u_m(\mathbf{x}) \hat{a}_m \right) \\ &= \sum_{n, m} \hat{a}_n^\dagger \hat{a}_m \int d^d \mathbf{x} u_n^*(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) u_m(\mathbf{x}) \\ &= \sum_n \epsilon_n \hat{a}_n^\dagger \hat{a}_n, \end{aligned}$$

thus confirming the basis independence of the field operator representation of the single-particle operator

$$\sum_{\alpha, \beta} \xi_{\alpha\beta} \hat{b}_\alpha^\dagger \hat{b}_\beta \iff \sum_n \epsilon_n \hat{a}_n^\dagger \hat{a}_n,$$

with coefficients given by the matrix elements of the single-particle operator in the respective bases:

$$\xi_{\alpha\beta} = \int d^d \mathbf{x} v_{\alpha}^{*}(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) v_{\beta}(\mathbf{x}), \quad \epsilon_n = \int d^d \mathbf{x} u_n^{*}(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) u_n(\mathbf{x}).$$

Thus, we can express the single-particle operator \hat{O} in Fock space in a basis-independent form using the field operators as:

$$\hat{O} = \int d^d \mathbf{x} \hat{\psi}^{\dagger}(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) \hat{\psi}(\mathbf{x}). \quad (8.4.3)$$

Resuming these results, we have shown that an operator acting on single-particle states can be lifted to an operator acting on Fock space by summing over all particles, and that this operator can be expressed in terms of creation and annihilation operators or field operators:

- $\hat{O} = \sum_n \epsilon_n \hat{a}_n^{\dagger} \hat{a}_n$ in the diagonal basis $u_n(\mathbf{x}) \iff \hat{a}_n, \hat{a}_n^{\dagger}$, with $\epsilon_n = \langle u_n | \hat{O}^{(1)} | u_n \rangle$ the eigenvalues of the single-particle operator;
- $\hat{O} = \sum_{\alpha,\beta} \xi_{\alpha\beta} \hat{b}_{\alpha}^{\dagger} \hat{b}_{\beta}$ in a generic basis $v_{\beta}(\mathbf{x}) \iff \hat{b}_{\beta}, \hat{b}_{\beta}^{\dagger}$, with $\xi_{\alpha\beta} = \langle v_{\alpha} | \hat{O}^{(1)} | v_{\beta} \rangle$ the matrix elements of the single-particle operator;
- $\hat{O} = \int d^d \mathbf{x} \hat{\psi}^{\dagger}(\mathbf{x}) \hat{O}^{(1)}(\mathbf{x}, \mathbf{p}) \hat{\psi}(\mathbf{x})$ in terms of the field operators, independent from basis choice.

8.4.1 | Density and Number Operators

Let us now consider the **density operator** $\hat{\rho}(\mathbf{x})$, which represents the particle density at a given position in space. In Fock space, the density operator can be expressed in terms of a sum of single-particle delta functions:

$$\hat{\rho}(\mathbf{x}) = \sum_{j=1}^N \delta(\mathbf{x} - \mathbf{x}_j), \quad (8.4.4)$$

where \mathbf{x}_j is the eigenvalue of the position operator for the j -th particle. Looking at the action of the single-particle operators $\hat{\rho}_j = \delta(\mathbf{x} - \mathbf{x}_j)$ (the delta functions) on a field operator, we can see that it picks out its contribution at the position \mathbf{x}_j :

$$\psi(\mathbf{x}) \rightarrow \hat{\rho}_j \psi(\mathbf{x}) = \int d^d \mathbf{y} \delta(\mathbf{y} - \mathbf{x}_j) \psi(\mathbf{y}) = \psi(\mathbf{x}_j).$$

Thus, we can express the density operator $\hat{\rho}(\mathbf{x})$ in terms of the field operators as:

$$\hat{\rho}(\mathbf{x}) = \hat{\psi}^{\dagger}(\mathbf{x}) \hat{\psi}(\mathbf{x}) = \sum_{m,n} u_m^{*}(\mathbf{x}) u_n(\mathbf{x}) \hat{a}_m^{\dagger} \hat{a}_n, \quad (8.4.5)$$

which counts the number of particles at position \mathbf{x} by creating and annihilating a particle at that point. This expression is consistent with the previous definition of a Fock operator in (8.4.3), as one can easily compute:

$$\begin{aligned} \hat{\rho}(\mathbf{x}) &= \int d^d \mathbf{y} \hat{\psi}^{\dagger}(\mathbf{y}) \hat{\rho}(\mathbf{x}) \hat{\psi}(\mathbf{y}) = \int d^d \mathbf{y} \hat{\psi}^{\dagger}(\mathbf{y}) \delta(\mathbf{y} - \mathbf{x}) \hat{\psi}(\mathbf{y}) \\ &= \hat{\psi}^{\dagger}(\mathbf{x}) \hat{\psi}(\mathbf{x}) = \sum_{m,n} u_m^{*}(\mathbf{x}) u_n(\mathbf{x}) \hat{a}_m^{\dagger} \hat{a}_n. \end{aligned}$$

which matches the definition in (8.4.5) (the last equality follows from the definition of the field operators as expansion in ladder operators). If we recognize that the term $u_m^{*}(\mathbf{x}) u_n(\mathbf{x})$ is an outer product of the basis functions evaluated at position \mathbf{x} , we can see that the density operator $\hat{\rho}(\mathbf{x})$

effectively sums over all possible single-particle states, weighted by their contributions at that position.

For what concerns the **number operator** \hat{N} , which counts the total number of particles in the system, we can express it as the integral of the density operator over all space:

$$\hat{N} = \int d^d\mathbf{x} \hat{\rho}(\mathbf{x}) = \int d^d\mathbf{x} \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x}) = \sum_k \hat{a}_k^\dagger \hat{a}_k = \sum_k \hat{n}_k, \quad (8.4.6)$$

where we have used the expression for the density operator in terms of the field operators. The number operator \hat{N} thus counts the total number of particles by summing over the occupation numbers of all single-particle states.

The last equality in equation (8.4.6) can be derived by substituting the expansion of the field operators in terms of the ladder operators into the integral:

$$\begin{aligned} \hat{N} &= \int d^d\mathbf{x} \left(\sum_m u_m^*(\mathbf{x}) \hat{a}_m^\dagger \right) \left(\sum_n u_n(\mathbf{x}) \hat{a}_n \right) \\ &= \sum_{m,n} \hat{a}_m^\dagger \hat{a}_n \int d^d\mathbf{x} u_m^*(\mathbf{x}) u_n(\mathbf{x}) = \sum_n \hat{a}_n^\dagger \hat{a}_n = \sum_n \hat{n}_n, \end{aligned}$$

where we have used the orthonormality of the basis functions $u_n(\mathbf{x})$ to evaluate the integral.

8.4.2 | Free Hamiltonian

Let us now consider the **free Hamiltonian** \hat{H} of a system of non-interacting particles. The free Hamiltonian describes the kinetic energy of the particles and can be expressed in terms of the momentum operator $\hat{\mathbf{p}}$ as:

$$\hat{H} = \sum_{j=1}^N \frac{\hat{\mathbf{p}}_j^2}{2m}, \quad (8.4.7)$$

where m is the mass of the particles and $\hat{\mathbf{p}}_j$ is the momentum operator for the j -th particle.

If we work in the basis of unnormalized plane wave states $u_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}$, which are eigenstates of the momentum operator with eigenvalues $\hbar\mathbf{k}$, we can express the diagonalized free Hamiltonian in Fock space as:

$$-\frac{\hbar^2 \nabla^2}{2m} u_{\mathbf{k}_n}(\mathbf{x}) = \epsilon_{\mathbf{k}_n} u_{\mathbf{k}_n}(\mathbf{x}),$$

where we can compute the energy eigenvalue associated with the plane wave state $u_{\mathbf{k}_n}(\mathbf{x})$ as

$$-\frac{\hbar^2 \nabla^2}{2m} e^{i\mathbf{k}_n \cdot \mathbf{x}} = \frac{\hbar^2 \mathbf{k}_n^2}{2m} e^{i\mathbf{k}_n \cdot \mathbf{x}} \implies \epsilon_{\mathbf{k}_n} = \frac{\hbar^2 \mathbf{k}_n^2}{2m}.$$

Periodic Boundary Conditions. As we have anticipated earlier, the plane wave states $u_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}$ are not normalizable in infinite space, as they extend infinitely in all directions. To address this issue, we can impose **periodic boundary conditions** on a finite volume $V = L^d$ of space, effectively confining the particles to a box of side length L in d dimensions. This allows us to discretize the allowed wavevectors \mathbf{k} and normalize the plane wave states. We can impose in each spatial dimension (for example in three dimensions):

$$\begin{cases} u_{\mathbf{k}}(x, y, z) = u_{\mathbf{k}}(x + L, y, z), \\ u_{\mathbf{k}}(x, y, z) = u_{\mathbf{k}}(x, y + L, z), \\ u_{\mathbf{k}}(x, y, z) = u_{\mathbf{k}}(x, y, z + L), \end{cases}$$

which leads to the quantization of the wavevectors as:

$$e^{k_x x} e^{ik_x L} = e^{ik_x x} \implies k_x = \frac{2\pi n_x}{L}, \quad n_x \in \mathbb{Z},$$

and similarly for the other dimensions. Thus, the allowed wavevectors in three dimensions become:

$$\mathbf{k} = \frac{2\pi}{L} (n_x, n_y, n_z), \quad n_x, n_y, n_z \in \mathbb{Z}.$$

With these periodic boundary conditions, we can normalize the plane wave states over the finite volume V as:

$$\|u_{\mathbf{k}}(\mathbf{x})\|_2^2 = \int_V d^3\mathbf{x} |u_{\mathbf{k}}(\mathbf{x})|^2 = |\mathcal{C}|^2 \int_V d^3\mathbf{x} 1 \implies \mathcal{C} = \frac{1}{\sqrt{V}},$$

ensuring that they satisfy the orthonormality condition:

$$\int_V d^d\mathbf{x} u_{\mathbf{k}}^*(\mathbf{x}) u_{\mathbf{k}'}(\mathbf{x}) = \delta_{\mathbf{k}, \mathbf{k}'},$$

Thus the normalized plane wave states in a finite volume with periodic boundary conditions are given by:

$$u_{\mathbf{k}}(\mathbf{x}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \mathbf{k} = \frac{2\pi}{L} (n_x, n_y, n_z), \quad n_x, n_y, n_z \in \mathbb{Z}.$$

Using these normalized plane wave states as our basis functions $u_n(\mathbf{x})$, we can express the free Hamiltonian in Fock space using equation (8.4.2) as:

$$\hat{H} = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} = \frac{\hbar^2}{2m} \sum_{\mathbf{k}} \mathbf{k}^2 \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}}, \quad (8.4.8)$$

where the sum runs over all allowed wavevectors \mathbf{k} in the finite volume with periodic boundary conditions. This is the standard form of the free Hamiltonian in second quantization, describing a quantum non relativistic perfect gas, where the bosonic or fermionic nature of the particles is encoded in the commutation or anticommutation relations of the ladder operators $\hat{a}_{\mathbf{k}}$ and $\hat{a}_{\mathbf{k}}^\dagger$.

Other hamiltonian systems. If we were to consider a system of photons, this expression would still hold, but we would have to take into account that **photons** are massless particles, thus the energy eigenvalues would be given by

$$\epsilon_{\mathbf{k}} = c|\mathbf{p}| = \hbar c|\mathbf{k}|,$$

where c is the speed of light. This relation is known as the **dispersion relation** for photons, reflecting the fact that their energy is directly proportional to their momentum magnitude. There exist other dispersion relations for different types of particles or quasi-particles, depending on their mass and the nature of their interactions: the **phonons**, for example, have the following dispersion relation

$$\epsilon_{\mathbf{k}} = v|\sin(\mathbf{k})|.$$

Phonons are quantized modes of vibrations in a crystal lattice (quantum particles of sound or lattice vibrations), and their dispersion relation reflects the collective excitations of the lattice structure. Here, v represents the speed of sound in the material, and the sine function captures the periodic nature of the lattice vibrations.

Hints on energy transitions. We are considering in general a system where matter is conserved: it can be transformed into energy and vice versa, as the relativistic equation states, but particles cannot be created or destroyed. However in transitions between different energy states, particles absorbe or emit energy quanta (photons) to move from one energy level to another. In these processes, the number of particles remains constant, but their energy changes due to the absorption or emission of photons, thus the number of photons changes. We could describe this transition as

$$|1, 0, 0, \dots\rangle \rightarrow |0, 1, 0, \dots\rangle \implies \hat{a}_2^\dagger \hat{a}_1 |1, 0, 0, \dots\rangle = |0, 1, 0, \dots\rangle,$$

where we are not considering the photon states explicitly, but we could do it by adding another set of ladder operators for the photons, for example $\hat{A}_q, \hat{A}_q^\dagger$ for the photon mode with wavevector q :

$$\begin{aligned} & |1, 0, 0, \dots\rangle_{\text{mat}} \otimes |n_1, \dots\rangle_{\text{phot}} \rightarrow |0, 1, 0, \dots\rangle_{\text{mat}} \otimes |n_1 - 1, \dots\rangle_{\text{phot}} \\ & \implies \hat{a}_2^\dagger \hat{a}_1 \hat{A}_q |1, 0, 0, \dots\rangle_{\text{mat}} \otimes |n_1, \dots\rangle_{\text{phot}} = |0, 1, 0, \dots\rangle_{\text{mat}} \otimes |n_1 - 1, \dots\rangle_{\text{phot}}. \end{aligned}$$

The notation here is a bit sloppy, but the idea is that we can describe the matter and photon states separately, and use the appropriate ladder operators to describe the transitions between different energy levels and photon states.

9 | Quantum Ensembles

In this chapter, we extend the concepts of statistical mechanics to quantum systems. We will explore the different quantum ensembles and their applications to various physical systems. We will consider only systems in thermal equilibrium, where the principles of quantum mechanics and statistical mechanics intersect; we will discuss the microcanonical, canonical, and grand canonical ensembles in the quantum context.

The systems we are going to address are characterized by discrete energy levels, and the state of the system $|\psi\rangle \in \mathcal{H}_N$ is described by a density matrix rather than a classical probability distribution

$$\hat{\rho}_\psi = |\psi\rangle\langle\psi| \text{ or } \hat{\rho}_\psi = \sum_\alpha p_\alpha |\alpha\rangle\langle\alpha|,$$

Operators are going to be assumed time independent and they will act on the proper Hilbert space of the system: if we are treating a system of N particles, the Hilbert space will be the tensor product of the single-particle Hilbert spaces:

$$\mathcal{H}_N = \mathcal{H}_1^{\otimes N} \text{ for distinguishable particles,}$$

$$\mathcal{H}_N = \mathcal{H}_S^{(N)} \text{ for bosons,}$$

$$\mathcal{H}_N = \mathcal{H}_A^{(N)} \text{ for fermions,}$$

and when we will need N to vary, we will work in the Fock space $\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{H}_{S/A}^{(N)}$.

The Hamiltonian operator \hat{H}_N will have discrete eigenvalues E_n with eigenstates $|n\rangle$, and it will commute with other observables of the system, such as the number operator \hat{N} ¹, to ensure conservation laws. To clarify the procedure, we will replace

- the phase space with the appropriate Hilbert space,
- the Hamiltonian function with the Hamiltonian operator, self-adjoint and assumed time independent,
- the classical probability distribution with the density operator $\hat{\rho}$ to describe the statistical state of the system,
- the partition function with the appropriate quantum renormalization factor, ensuring the trace of the density operator is equal to one,
- the classical expectation values with the quantum expectation values, computed as the trace of the product of the density operator and the observable operator.

¹Even in the GC ensamble the Hamiltonian will commute with the number operator, if we are using the Hamiltonian defined for a fixed number of particles \hat{H}_N ; when we will use the Hamiltonian acting on the Fock space for a variable number of particles \hat{H} things will be more complicated.

9.1 | Microcanonical Ensemble

As in the classical case, the microcanonical ensemble describes an isolated quantum system with fixed energy E (configurations on a constant energy surface), volume V , and number of particles N ; most operators will have a discrete spectrum and a finite degeneracy g_n associated to each energy level E_n , working at fixed volume and number of particles.

Given an ON basis $\{|\psi_{j,\alpha}\rangle\}_{j,\alpha}$, the generic time independent Hamiltonian operator \hat{H} can be diagonalized as

$$\hat{H} |\psi_{j,\alpha}\rangle = E_j |\psi_{j,\alpha}\rangle, \quad \alpha = 1, \dots, g_j,$$

where α encodes the degeneracy g_j of each energy level E_j and we can express the Hamiltonian in terms of its spectral decomposition:

$$\hat{H} = \sum_j \sum_{\alpha=1}^{g_j} E_j |\psi_{j,\alpha}\rangle \langle \psi_{j,\alpha}| = \sum_j E_j \hat{\mathbb{P}}_j,$$

where $\hat{\mathbb{P}}_j = \sum_{\alpha=1}^{g_j} |\psi_{j,\alpha}\rangle \langle \psi_{j,\alpha}|$ is the projector onto the eigenspace associated to the energy level E_j . Since the system is isolated, the energy is constrained to lie within a small interval $[E, E + \delta E]$, and all accessible microstates within this energy range are equally probable: in the microcanonical ensemble we have $E = E_{j_0}$, so that

- $|\psi_{j,\alpha}\rangle$ has zero probability for all $j \neq j_0$,
- $|\psi_{j_0,\alpha}\rangle$ has equal probability $\frac{1}{g_{j_0}}$ for all $\alpha = 1, \dots, g_{j_0}$.

Thus the density operator for a mixed state of the microcanonical ensemble is given by

$$\hat{\rho}_{mc} = \sum_{j \neq j_0, \alpha} (p_{j \neq j_0, \alpha} = 0) |\psi_{j,\alpha}\rangle \langle \psi_{j,\alpha}| + \sum_{\alpha=1}^{g_{j_0}} p_{j_0} |\psi_{j_0,\alpha}\rangle \langle \psi_{j_0,\alpha}|,$$

with $p_{j_0} = \frac{1}{g_{j_0}}$ since we have equal probabilities for all microstates at energy E_{j_0} and they must sum to one; therefore we can rewrite the density operator as

$$\hat{\rho}_{mc} = \frac{1}{g_{j_0}} \hat{\mathbb{P}}_{j_0} = \frac{1}{g_{j_0}} \sum_{\alpha=1}^{g_{j_0}} |\psi_{j_0,\alpha}\rangle \langle \psi_{j_0,\alpha}|, \quad (9.1.1)$$

Since the density operator is normalized, we have $\text{Tr}(\hat{\rho}_{mc}) = 1$. Let us drop the subscript j_0 when there is no ambiguity, since we are working at fixed energy, after selecting a specific eigenspace of the Hamiltonian. We can look at the matrix elements of the density operator in the energy eigenbasis:

$$\hat{\rho}_{mc} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{g} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \vdots & \frac{1}{g} & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

(?) explained badly or badly written matrix? it is a block diagonal matrix with a block of size $g \times g$ with all entries equal to $\frac{1}{g}$ corresponding to the degenerate energy level E , and zeros elsewhere.

The expectation value of an observable \hat{A} in the microcanonical ensemble is given by

$$\langle \hat{A} \rangle_{mc} = \text{Tr}(\hat{\rho}_{mc} \hat{A}) = \frac{1}{g} \sum_{\alpha=1}^g \langle \alpha | \hat{A} | \alpha \rangle, \quad (9.1.2)$$

where $\frac{1}{g}$ represents the equal probability of each microstate within the energy shell; we also used $|\alpha\rangle$ to denote the eigenstates $|\psi_{j,\alpha}\rangle$ for simplicity.

We can also define the quantum analog of the entropy in the microcanonical ensemble, from the Boltzmann formula:

$$\begin{aligned} S &= -k_B \langle \log \hat{\rho}_{mc} \rangle_{mc} = -k_B \text{Tr}(\hat{\rho}_{mc} \log \hat{\rho}_{mc}) \\ &= -k_B \sum_{\alpha} \langle \alpha | \left[\frac{1}{g} \sum_{\beta} |\beta\rangle \langle \beta| \log \left(\frac{1}{g} \sum_{\gamma} |\gamma\rangle \langle \gamma| \right) \right] |\alpha\rangle \\ &= -k_B \sum_{\alpha} \frac{1}{g} \log \left(\frac{1}{g} \right) \langle \alpha | \alpha \rangle = -k_B \frac{g}{g} \log \left(\frac{1}{g} \right) = k_B \log g, \end{aligned}$$

where again g is the number of accessible microstates at energy E and we have used the completeness relations $\sum_{\alpha} |\alpha\rangle \langle \alpha| = \mathbb{I}$. Everything else being equal, a higher degeneracy g leads to a higher entropy S , reflecting the greater number of accessible microstates for the system. This aligns with the classical interpretation of entropy as a measure of the number of microstates corresponding to a given macrostate.

9.2 | Canonical Ensemble

In the canonical ensemble, we consider a quantum system in thermal equilibrium with a heat bath at a fixed temperature T . The system can exchange energy with the bath, leading to fluctuations in its energy levels. The number of particles N and the volume V of the system remain constant.

Now we cannot restrict the system to a single energy level, as in the microcanonical ensemble; instead, the system can occupy various energy levels E_j with probabilities determined by the Boltzmann factors: for each level E_j the system has a probability proportional to $e^{-\beta E_j}$ of being observed in that level.

Now the density operator for the canonical ensemble is given by

$$\hat{\rho}_c = \frac{1}{Z_N} \sum_j e^{-\beta E_j} \hat{\mathbb{P}}_j = \frac{1}{Z_N} e^{-\beta \hat{H}_N}, \quad (9.2.1)$$

where $Z_N = \text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N})$ is the partition function, ensuring the normalization of the density operator. Let's compute this results explicitly:

$$e^{-\beta \hat{H}_N} = \sum_n \frac{(-\beta \hat{H}_N)^n}{n!} = \sum_n \frac{(-\beta)^n}{n!} \left(\sum_j E_j \hat{\mathbb{P}}_j \right)^n,$$

practically we need to compute the n -th power of the Hamiltonian:

$$\hat{H}_N^2 = \left(\sum_i E_i \hat{\mathbb{P}}_i \right) \left(\sum_j E_j \hat{\mathbb{P}}_j \right) = \sum_{i,j} E_i E_j \hat{\mathbb{P}}_i \hat{\mathbb{P}}_j = \sum_j E_j^2 \hat{\mathbb{P}}_j,$$

so we can generalize to $\hat{H}_N^n = \sum_j E_j^n \hat{\mathbb{P}}_j$; thus we have

$$e^{-\beta \hat{H}_N} = \sum_j \left(\sum_n \frac{(-\beta E_j)^n}{n!} \right) \hat{\mathbb{P}}_j = \sum_j e^{-\beta E_j} \hat{\mathbb{P}}_j.$$

We can find the partition function Z_N by requiring the normalization of the density operator:

$$\text{Tr}_{\mathcal{H}_N}(\hat{\rho}_c) = 1 = \frac{1}{Z_N} \text{Tr}\left(e^{-\beta \hat{H}_N}\right) \implies Z_N = \text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N}).$$

We can also reduce the expression for the partition function to a more familiar one using the spectral decomposition of the Hamiltonian:

$$Z_N = \text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N}) = \sum_j e^{-\beta E_j} \text{Tr}_{\mathcal{H}_N}(\hat{\mathbb{P}}_j) = \sum_j g_j e^{-\beta E_j}, \quad (9.2.2)$$

where $g_j = \text{Tr}_{\mathcal{H}_N}(\hat{\mathbb{P}}_j)$ is the degeneracy of the energy level E_j .

The expectation value of an observable \hat{A} in the canonical ensemble is given by

$$\langle \hat{A} \rangle_c = \text{Tr}_{\mathcal{H}_N}(\hat{\rho}_c \hat{A}) = \frac{1}{Z_N} \sum_j e^{-\beta E_j} \text{Tr}_{\mathcal{H}_N}(\hat{\mathbb{P}}_j \hat{A}). \quad (9.2.3)$$

Now we can compute the quantum analog of the Helmholtz free energy F in the canonical ensemble, in order to relate thermodynamic quantities to the partition function:

$$Z_N = e^{-\beta F} \implies F = -\frac{1}{\beta} \log Z_N.$$

From the Helmholtz free energy, we can derive other thermodynamic quantities. The internal energy E instead, is given by

$$\begin{aligned} E &= \langle \hat{H}_N \rangle_c = \text{Tr}_{\mathcal{H}_N}(\hat{\rho}_c \hat{H}_N) = \frac{1}{Z_N} \text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N} \hat{H}_N) \\ &= \frac{-1}{Z_N} \frac{\partial}{\partial \beta} \text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N}) = -\frac{\partial}{\partial \beta} \log Z_N. \end{aligned}$$

The entropy S in the canonical ensemble can be computed using the relation

$$\begin{aligned} S &= -k_B \langle \log \hat{\rho}_c \rangle_c = -k_B \text{Tr}_{\mathcal{H}_N}(\hat{\rho}_c \log \hat{\rho}_c) = -k_B \text{Tr}_{\mathcal{H}_N} \left(\frac{e^{-\beta \hat{H}_N}}{Z_N} (-\beta \hat{H}_N - \log Z_N) \right) \\ &= -k_B \left(-\beta \text{Tr}_{\mathcal{H}_N} \left(\frac{\hat{H}_N e^{-\beta \hat{H}_N}}{Z_N} \right) + \frac{\beta F}{Z_N} \text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N}) \right) \\ &= -k_B (-\beta E + \beta F) = \frac{E - F}{T}, \end{aligned}$$

since $\log Z_N = -\beta F$ and $\text{Tr}_{\mathcal{H}_N}(e^{-\beta \hat{H}_N}) = Z_N$. This result is consistent with the thermodynamic, confirming the validity of our quantum statistical mechanics framework.

9.3 | Grand Canonical Ensemble

In the grand canonical ensemble, we consider a quantum system that can exchange both energy and particles with a reservoir. The system is characterized by a fixed temperature T , volume V , and chemical potential μ . The number of particles N in the system can fluctuate, so we work in the Fock space

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{H}_{S/A}^{(N)}.$$

The Hamiltonian operator \hat{H} in the grand canonical ensemble acts on the Fock space, and it is the combination of the Hamiltonians for different particle numbers acting on their respective Hilbert spaces; thus we have the number operator \hat{N} that commutes with the Hamiltonian at fixed N (canonical Hamiltonian \hat{H}_N), while the full Hamiltonian \hat{H} will respect

$$\hat{H} |\psi_{j,\alpha}\rangle^{(N)} = E_j^{(N)} |\psi_{j,\alpha}\rangle^{(N)},$$

where the ON base $\{|\psi_{j,\alpha}\rangle^{(N)}\}_{j,\alpha,N}$ spans the Fock space, with α indexing the degeneracy of each energy level $E_j^{(N)}$ for a fixed number of particles N . Thus we have projectors in the Fock space $\hat{\mathbb{P}}_j^{(N)} = \sum_{\alpha=1}^{g_j^{(N)}} |\psi_{j,\alpha}\rangle^{(N)} \langle \psi_{j,\alpha}|^{(N)}$ which are N dependent and orthogonal in each index. Finally we can express the Hamiltonian in terms of its spectral decomposition:

$$\hat{H} = \sum_N \sum_j E_j^{(N)} \hat{\mathbb{P}}_j^{(N)}.$$

The density operator for a mixed state in the grand canonical ensemble is given by

$$\hat{\rho}_{gc} = \sum_N \sum_j p_j^{(N)} \hat{\mathbb{P}}_j^{(N)},$$

where, since we are mimicking the classical ensemble, the probabilities $p_j^{(N)}$ are determined by the Boltzmann factors, taking into account both energy and particle number: $p_j^{(N)} \propto e^{-\beta(E_j^{(N)} - \mu N)}$. Thus we can write the density operator as

$$\hat{\rho}_{gc} = \frac{1}{Z} \sum_N \sum_j e^{-\beta(E_j^{(N)} - \mu N)} \hat{\mathbb{P}}_j^{(N)} = \frac{1}{Z} e^{-\beta(\hat{H} - \mu \hat{N})}, \quad (9.3.1)$$

where we need to make some considerations:

- $\sum_N \left(\sum_j e^{-\beta(E_j^{(N)})} \hat{\mathbb{P}}_j^{(N)} \right) = e^{-\beta \hat{H}}$, since we are summing over all possible particle numbers and energy levels, reconstructing the full Hamiltonian operator acting on the Fock space (the computation is almost identical to the canonical case);
- $\sum_N \left(\sum_j e^{\beta \mu N} \hat{\mathbb{P}}_j^{(N)} \right) = e^{\beta \mu \hat{N}}$, since the number operator \hat{N} acts on the Fock space and counts the number of particles in each state (again, it is the same computation as before);
- the normalization factor $\frac{1}{Z}$ is the grand partition function, ensuring that the density operator is properly normalized.

We will spend some words on the computation of the grand partition function Z : it is defined by

the unitary condition on the density operator trace

$$\begin{aligned}\mathrm{Tr}_{\mathcal{F}}(\hat{\rho}_{gc}) &= \sum_{N=0}^{\infty} \mathrm{Tr}_{\mathcal{H}_N} \left(\frac{1}{\mathcal{Z}} e^{-\beta(\hat{H}-\mu\hat{N})} \right) = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} e^{\beta\mu N} \mathrm{Tr}_{\mathcal{H}_N}(e^{-\beta\hat{H}_N}) \\ &= \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N Z_N, = 1\end{aligned}$$

where we used the fact that the trace over the Fock space can be decomposed into traces over the fixed particle number Hilbert spaces \mathcal{H}_N , and we recognized the canonical partition function $Z_N = \mathrm{Tr}_{\mathcal{H}_N}(e^{-\beta\hat{H}_N})$ and the fugacity $z = e^{\beta\mu}$. Therefore, we find that the grand partition function is given by

$$\mathcal{Z} = \sum_{N=0}^{\infty} z^N Z_N = \sum_{N=0}^{\infty} z^N \mathrm{Tr}_{\mathcal{H}_N}(e^{-\beta\hat{H}_N}) = \mathrm{Tr}_{\mathcal{F}}(e^{-\beta(\hat{H}-\mu\hat{N})}), \quad (9.3.2)$$

The expectation value of an observable \hat{A} in the grand canonical ensemble is given by

$$\langle \hat{A} \rangle_{gc} = \mathrm{Tr}_{\mathcal{F}}(\hat{\rho}_{gc}\hat{A}) = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \mathrm{Tr}_{\mathcal{H}_N}(e^{-\beta\hat{H}_N}\hat{A}), \quad (9.3.3)$$

where the observable considered must act on the Fock space and commute with the number operator \hat{N} on each \mathcal{H}_N .

From all this tools, we can compute the quantum analog of the grand potential Ω in the grand canonical ensemble:

$$\mathcal{Z} = e^{-\beta\Omega} \implies \Omega = -\frac{1}{\beta} \log \mathcal{Z}.$$

For the other thermodynamic quantities, we could derive them from the grand potential, but it is more instructive to show how to compute them directly from the grand partition function, at least for the internal energy E and the grand canonical entropy S .

The internal energy E in the grand canonical ensemble is given by

$$\begin{aligned}E - \mu N &= \langle \hat{H} - \mu\hat{N} \rangle_{gc} = \mathrm{Tr}_{\mathcal{F}}(\hat{\rho}_{gc}(\hat{H} - \mu\hat{N})) = \frac{1}{\mathcal{Z}} \mathrm{Tr}_{\mathcal{F}}(e^{-\beta(\hat{H}-\mu\hat{N})}(\hat{H} - \mu\hat{N})) \\ &= \frac{-1}{\mathcal{Z}} \frac{\partial}{\partial\beta} \mathrm{Tr}_{\mathcal{F}}(e^{-\beta(\hat{H}-\mu\hat{N})}) = -\frac{\partial}{\partial\beta} \log \mathcal{Z}.\end{aligned}$$

The entropy S in the grand canonical ensemble can be computed as

$$\begin{aligned}S &= -k_B \langle \log \hat{\rho}_{gc} \rangle_{gc} = -k_B \mathrm{Tr}_{\mathcal{F}}(\hat{\rho}_{gc} \log \hat{\rho}_{gc}) = -k_B \mathrm{Tr}_{\mathcal{F}} \left(\frac{e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathcal{Z}} (-\beta(\hat{H} - \mu\hat{N}) - \log \mathcal{Z}) \right) \\ &= -k_B \left(-\beta \mathrm{Tr}_{\mathcal{F}} \left(\frac{(\hat{H} - \mu\hat{N})e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathcal{Z}} \right) + \frac{\beta\Omega}{\mathcal{Z}} \mathrm{Tr}_{\mathcal{F}}(e^{-\beta(\hat{H}-\mu\hat{N})}) \right) \\ &= -k_B (-\beta(E - \mu N) + \beta\Omega) = \frac{E - \mu N - \Omega}{T}.\end{aligned}$$

10 | Quantum Gases

Quantum gases are systems of non-interacting or weakly interacting particles that obey quantum statistics. Depending on the nature of the particles, they can be classified as either bosons or fermions, leading to different statistical behaviors.

A quantum gas can be described using the grand canonical ensemble, where the number of particles can fluctuate, and the system is characterized by a fixed temperature T , volume V , and chemical potential μ ; the Hamiltonian operator for a quantum gas of non-interacting particles can be expressed, for a fixed N , as

$$\hat{H} = \sum_i^N \hat{\mathcal{O}}_i = \sum_i^N \left(\frac{\hat{p}_i^2}{2m} + U(\hat{q}_i) \right),$$

where $\hat{\mathcal{O}}_i$ is the single-particle operator, \hat{p}_i and \hat{q}_i are the momentum and position operators of the i -th particle, m is the mass of the particles, and $U(\hat{q}_i)$ is the potential energy operator acting on the position of the i -th particle.

In second quantization, we can express the Hamiltonian operator in terms of creation and annihilation operators \hat{a}_α^\dagger and \hat{a}_α for the single-particle states:

$$\hat{H} = \sum_{\alpha\beta} t_{\alpha\beta} \hat{a}_\alpha^\dagger \hat{a}_\beta,$$

where the ladder operators satisfy the appropriate commutation or anticommutation relations depending on whether the particles are bosons or fermions: the nature of the particles determines the algebra of the creation and annihilation operators, which in turn affects the statistical properties of the quantum gas. Looking at the ON base for the single-particle Hamiltonian $\hat{\mathcal{O}}$, we can deduce a more useful form for our hamiltonian, diagonalized in terms of single-particle energy eigenstates $|e_\alpha\rangle$ with eigenvalues ϵ_α :

$$\hat{\mathcal{O}}|e_\alpha\rangle = \epsilon_\alpha |e_\alpha\rangle,$$

so that the full Hamiltonian operator can be expressed as

$$\hat{H} = \sum_\alpha \epsilon_\alpha \hat{a}_\alpha^\dagger \hat{a}_\alpha,$$

where \hat{a}_α^\dagger and \hat{a}_α are the creation and annihilation operators for the single-particle state $|e_\alpha\rangle$, forming the number operator $\hat{n}_\alpha = \hat{a}_\alpha^\dagger \hat{a}_\alpha$ that counts the number of particles in the state $|e_\alpha\rangle$.

Thus we can define

$$\hat{H} - \mu \hat{N} = \sum_\alpha (\epsilon_\alpha - \mu) \hat{n}_\alpha,$$

where $\hat{N} = \sum_{\alpha} \hat{n}_{\alpha}$ is the total number operator for the system. The density operator for the quantum gas in the grand canonical ensemble is given by

$$\hat{\rho}_{gc} = \frac{1}{Z} e^{-\beta(\hat{H}-\mu\hat{N})} = \frac{1}{Z} e^{-\beta \sum_{\alpha} (\epsilon_{\alpha} - \mu) \hat{n}_{\alpha}},$$

and the grandcanonical partition function can be computed as

$$\mathcal{Z} = \text{Tr}_{\mathcal{F}}(e^{-\beta(\hat{H}-\mu\hat{N})}) = \text{Tr}_{\mathcal{F}}\left(e^{-\beta \sum_{\alpha} (\epsilon_{\alpha} - \mu) \hat{n}_{\alpha}}\right),$$

where we can construct the Fock space \mathcal{F} as the tensor product of the single-particle state spaces:

$$|n_1, n_2, \dots\rangle = C(\hat{a}_1^{\dagger})^{n_1} (\hat{a}_2^{\dagger})^{n_2} \cdots |0\rangle,$$

with C a normalization constant and $|0\rangle$ the vacuum state. The occupation numbers n_{α} indicate how many particles occupy the single-particle state $|e_{\alpha}\rangle$. Now we can compute the grand partition function exploiting the trace

$$\begin{aligned} \mathcal{Z} &= \sum_{n_1=0}^{n_{\max}} \sum_{n_2=0}^{n_{\max}} \cdots \langle n_1, n_2, \dots | e^{-\beta \sum_{\alpha} (\epsilon_{\alpha} - \mu) \hat{n}_{\alpha}} | n_1, n_2, \dots \rangle \\ &= \sum_{n_1, n_2, \dots} e^{-\beta \sum_{\alpha} (\epsilon_{\alpha} - \mu) n_{\alpha}} \langle n_1, n_2, \dots | n_1, n_2, \dots \rangle \\ &= \sum_{n_1, n_2, \dots} \prod_{\alpha} e^{-\beta(\epsilon_{\alpha} - \mu)n_{\alpha}} = \prod_{\alpha} \left(\sum_{n_{\alpha}=0}^{n_{\max}} e^{-\beta(\epsilon_{\alpha} - \mu)n_{\alpha}} \right), \end{aligned}$$

where we swapped the sums and products since the indices are independent (note how the sum changed index). The upper limit n_{\max} of the sums over occupation numbers depends on the nature of the particles:

- for bosons, $n_{\max} = \infty$, since multiple bosons can occupy the same quantum state;
- for fermions, $n_{\max} = 1$, due to the Pauli exclusion principle, which states that no two fermions can occupy the same quantum state simultaneously.

10.1 | Quantum Statistics

We can now derive the specific forms of the grand partition function and related thermodynamic quantities for fermionic and bosonic quantum gases. We need to compute the sum over occupation numbers for each case, rememnbering that we cannot accept divergences to infinity in the partition function, since it must be finite and well defined (imagine trying to compute the TD limit).

Fermionic case. For fermions, the occupation numbers n_α can only take values 0 or 1. Therefore, the sum over occupation numbers for each single-particle state α gives the following result grand partition function:

$$\mathcal{Z}_F = \prod_{\alpha} \left(\sum_{n_\alpha=0}^1 e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \right) = \prod_{\alpha} \left(1 + e^{-\beta(\epsilon_\alpha - \mu)} \right).$$

From the grand partition function, we can derive the grand potential Ω_F for the fermionic quantum gas, in order to relate thermodynamic quantities to the partition function:

$$\Omega_F = -\frac{1}{\beta} \log \mathcal{Z}_F = -\frac{1}{\beta} \sum_{\alpha} \log \left(1 + e^{-\beta(\epsilon_\alpha - \mu)} \right).$$

Now it is time to compute the average occupation number $\langle n_\alpha \rangle$ for each single-particle state α in the fermionic quantum gas. This quantity represents the expected number of particles occupying the state $|e_\alpha\rangle$, and it's the core of **Fermi-Dirac statistics**:¹

$$\langle n_\alpha \rangle_F = \frac{1}{e^{\beta(\epsilon_\alpha - \mu)} + 1}. \quad (10.1.1)$$

Bosonic case. For bosons, the occupation numbers n_α can take any non-negative integer value (0, 1, 2, ...). Therefore, the sum over occupation numbers for each single-particle state α gives the following result for the grand partition function:

$$\mathcal{Z}_B = \prod_{\alpha} \left(\sum_{n_\alpha=0}^{\infty} e^{-\beta(\epsilon_\alpha - \mu)n_\alpha} \right) = \prod_{\alpha} \frac{1}{1 - e^{-\beta(\epsilon_\alpha - \mu)}},$$

where we used the formula for the sum of a geometric series.² Since the geometric series $\sum_n x^n = \frac{1}{1-x}$ converges only for values of $x < 1$, we get a prescription on the exponent:

$$e^{-\beta(\epsilon_\alpha - \mu)} < 1 \iff \epsilon_\alpha - \mu > 0 \iff \mu < \min_{\alpha} \{\epsilon_\alpha\},$$

meaning that, since we can take $\min_{\alpha} \{\epsilon_\alpha\} = 0$, the chemical potential μ must be negative for the bosonic quantum gas to ensure convergence of the grand partition function; this condition has important physical implications, especially when considering phenomena like **Bose-Einstein condensation**. From the grand partition function, we can derive the grand potential Ω_B for the bosonic quantum gas:

$$\Omega_B = -\frac{1}{\beta} \log \mathcal{Z}_B = \frac{1}{\beta} \sum_{\alpha} \log \left(1 - e^{-\beta(\epsilon_\alpha - \mu)} \right).$$

¹The computation will be done after some result from the Bosonic case.

²It is crucial that the series converges, it's not guaranteed for all values of μ and ϵ_α , and surely it's not only a way to reconduce to a prettier mathematical expression, as we will highlight.

Finally we get to derive the **Bose-Einstein statistics**, computing the average occupation number $\langle n_\alpha \rangle$ for each single-particle state α in the bosonic quantum gas:

$$\langle n_\alpha \rangle_B = \frac{1}{e^{\beta(\epsilon_\alpha - \mu)} - 1}. \quad (10.1.2)$$

It is good practice to compact the expressions for the two different statistics in a single formula: from the grand partition function

$$\mathcal{Z}_{B/F} = \prod_\alpha \left(1 \mp e^{-\beta(\epsilon_\alpha - \mu)} \right)^{\mp 1}, \quad (10.1.3)$$

where the upper sign refers to bosons and the lower sign to fermions. For the grandpotential we have

$$\Omega_{B/F} = \pm \frac{1}{\beta} \sum_\alpha \log \left(1 \mp e^{-\beta(\epsilon_\alpha - \mu)} \right), \quad (10.1.4)$$

and finally we are able to compute explicitly the average occupation number for both statistics at the same time:

$$\begin{aligned} \langle \hat{n}_\alpha \rangle_{gc}^{B/F} &= \text{Tr}_{\mathcal{F}}(\rho_{gc} \hat{n}_\alpha) = \frac{1}{\mathcal{Z}_{B/F}} \text{Tr}_{\mathcal{F}} \left(e^{-\beta(\hat{H} - \mu \hat{N})} \hat{n}_\alpha \right) \\ &= \frac{1}{\mathcal{Z}_{B/F}} \text{Tr}_{\mathcal{F}} \left(e^{-\beta \sum_\gamma (\epsilon_\gamma - \mu) \hat{n}_\gamma} \hat{n}_\alpha \right) \\ &= \frac{-1}{\beta \mathcal{Z}_{B/F}} \text{Tr}_{\mathcal{F}} \left(\frac{\partial}{\partial \epsilon_\alpha} e^{-\beta \sum_\gamma (\epsilon_\gamma - \mu) \hat{n}_\gamma} \right) \\ &= \frac{-1}{\beta} \frac{\partial}{\partial \epsilon_\alpha} \log \mathcal{Z}_{B/F} = \frac{\partial}{\partial \epsilon_\alpha} \Omega_{B/F} = \frac{1}{e^{\beta(\epsilon_\alpha - \mu)} \mp 1}, \end{aligned}$$

where again the upper sign refers to bosons and the lower sign to fermions, recovering equations (10.1.1) and (10.1.2). We can now interpret these results: the average occupation number $\langle n_\alpha \rangle$ gives the expected number of particles occupying the single-particle state $|e_\alpha\rangle$ at thermal equilibrium. The difference in the denominators reflects the fundamental distinction between bosons and fermions:

- For bosons, the minus sign in the denominator allows for the possibility of multiple particles occupying the same quantum state, leading to phenomena like Bose-Einstein condensation, where a macroscopic number of bosons can occupy the ground state at low temperatures.
- For fermions, the plus sign in the denominator enforces the Pauli exclusion principle, which prohibits multiple fermions from occupying the same quantum state. This leads to the characteristic filling of energy levels up to the Fermi energy at absolute zero temperature.

It is also interesting to analyze the behavior of these distributions compared to the classical Maxwell-Boltzmann distribution.

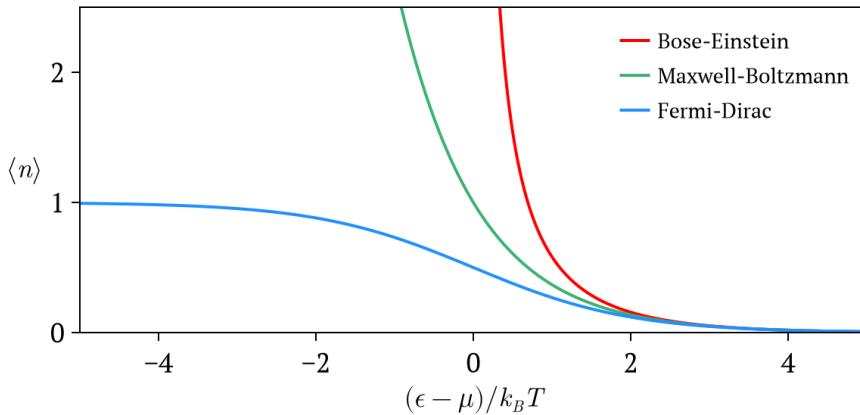


Figure 10.1: Comparison of quantum statistics (Fermi-Dirac and Bose-Einstein) with classical Maxwell-Boltzmann statistics. The plot illustrates how the average occupation number $\langle n_\alpha \rangle$ varies with energy ϵ_α for different temperatures T and chemical potentials μ . At high energies or low densities, all three distributions converge, reflecting classical behavior. However, at low energies or high densities, quantum effects become significant, leading to deviations from classical predictions. Notably, the Bose-Einstein distribution allows for a divergence in occupation number as energy approaches the chemical potential, indicating the onset of Bose-Einstein condensation. In contrast, the Fermi-Dirac distribution approaches a maximum occupation of 1 due to the Pauli exclusion principle.

Let us compute the internal energy E of the quantum gas using the average occupation numbers derived above:

$$\begin{aligned} E &= \langle \hat{H} \rangle_{gc}^{B/F} = \text{Tr}_{\mathcal{F}}(\rho_{gc} \hat{H}) = \frac{1}{Z_{B/F}} \text{Tr}_{\mathcal{F}} \left(e^{-\beta(\hat{H}-\mu\hat{N})} \hat{H} \right) \\ &= -\frac{\partial}{\partial \beta} \log Z_{B/F}|_z, \end{aligned}$$

where the subscript $|_z$ indicates that the derivative is taken at constant fugacity $z = e^{\beta\mu}$. Expanding this expression, we have

$$\begin{aligned} E &= \frac{1}{Z_{B/F}} \text{Tr}_{\mathcal{F}} \left(e^{-\beta \sum_\alpha (\epsilon_\alpha - \mu) \hat{n}_\alpha} \sum_\alpha \epsilon_\alpha \hat{n}_\alpha \right) \\ &= \sum_\alpha \epsilon_\alpha \left(\frac{1}{Z_{B/F}} \text{Tr}_{\mathcal{F}} \left(e^{-\beta(\hat{H}-\mu\hat{N})} \hat{n}_\alpha \right) \right) \\ &= \sum_\alpha \epsilon_\alpha \langle \hat{n}_\alpha \rangle_{gc}^{B/F} = \sum_\alpha \frac{\epsilon_\alpha}{e^{\beta(\epsilon_\alpha - \mu)} \mp 1}. \end{aligned}$$

So the internal energy of the quantum gas can be expressed as a sum over all single-particle states, weighted by their respective energies and average occupation numbers according to either Bose-Einstein or Fermi-Dirac statistics

$$\langle \hat{H} \rangle_{gc}^{B/F} = \sum_\alpha \frac{\epsilon_\alpha}{e^{\beta(\epsilon_\alpha - \mu)} \mp 1}, \quad (10.1.5)$$

where we have not made any assumptions about the specific form of the single-particle energy levels ϵ_α ; this expression is quite general and can be applied to various quantum gas systems, including those confined in different potentials or geometries. Let us now specialize to the case of non-relativistic 3D quantum gases confined in a cubic box of volume $V = L^3$ with periodic boundary conditions.

10.2 | Non Relativistic 3D Quantum Gases

For a non-relativistic quantum gas confined in a cubic box of volume $V = L^3$ with periodic boundary conditions, the single-particle energy levels ϵ_α can be expressed in terms of the kinetic energy of the particles

$$\epsilon_{\mathbf{k}} = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m},$$

from the solution of the Schrödinger equation for a free particle in a box:

$$-\frac{\hbar^2 \nabla^2}{2m} |\psi\rangle = \epsilon |\psi\rangle.$$

The allowed wavevectors \mathbf{k} are quantized due to the boundary conditions, leading to discrete energy levels given by

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

so that the energy levels can be rewritten as

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2) \implies \epsilon_{\mathbf{n}} = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2).$$

The key thermodynamic quantities for the quantum gas can be computed using the expressions derived earlier, substituting the specific form of the energy levels $\epsilon_{\mathbf{n}}$: from the grand partition function

$$\mathcal{Z}_{B/F} = \prod_{\mathbf{n}} \left(1 \mp e^{-\beta(\epsilon_{\mathbf{n}} - \mu)}\right)^{\mp g},$$

where g is the discrete degeneracy factor accounting for internal degrees of freedom (e.g., spin), we can derive the grand potential, internal energy, and average number of particles:

$$\begin{aligned} \Omega_{B/F} &= \pm \frac{g}{\beta} \sum_{\mathbf{n}} \log \left(1 \mp e^{-\beta(\epsilon_{\mathbf{n}} - \mu)}\right), \\ E_{B/F} &= g \sum_{\mathbf{n}} \frac{\epsilon_{\mathbf{n}}}{e^{\beta(\epsilon_{\mathbf{n}} - \mu)} \mp 1}, \\ N_{B/F} &= g \sum_{\mathbf{n}} \frac{1}{e^{\beta(\epsilon_{\mathbf{n}} - \mu)} \mp 1}. \end{aligned}$$

If we now consider the thermodynamic limit, where the volume V goes to infinity while keeping the particle density N/V constant, we can replace the sums over discrete states with integrals over continuous momentum space.

$$\sum_{\mathbf{k}} \rightarrow \sum_{\mathbf{n}} \xrightarrow{TD} (?)$$

In order to perform this substitution, we need to determine the density of states in momentum space. The strategy is to look at the volume element in \mathbf{k} -space corresponding to each quantum state

$$\Delta k_j = \frac{2\pi}{L} ((n_j + 1) - n_j) = \frac{2\pi}{L} \implies \Delta k_x \Delta k_y \Delta k_z = \left(\frac{2\pi}{L}\right)^3 = \frac{(2\pi)^3}{V},$$

which in the thermodynamic limit becomes infinitesimal $L \rightarrow \infty \implies \Delta k_j \rightarrow 0$. Therefore, the number of states in a volume element d^{3k} in \mathbf{k} -space is given by

$$\sum_{\mathbf{n}} \frac{\hbar^2}{2m} \left(\frac{2\pi}{L}\right)^3 (\Delta n_x)(\Delta n_y)(\Delta n_z) \longrightarrow \sum_{\mathbf{k}} \frac{V}{(2\pi)^3} (\Delta k_x)(\Delta k_y)(\Delta k_z),$$

where we know that $(\Delta n_x)(\Delta n_y)(\Delta n_z) = 1$ since we are counting each discrete state and $(\Delta k_x)(\Delta k_y)(\Delta k_z)$ is the infinitesimal volume d^3k element in \mathbf{k} -space. Thus we pass from the sum over discrete states to an integral over continuous momentum space as

$$\sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3k.$$

We can use polar coordinates in momentum space, since the energy levels depend only on the magnitude of the wavevector $k = |\mathbf{k}|$, leading to

$$\int d^3k = 4\pi \int_0^\infty k^2 dk \implies \sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} 4\pi \int_0^\infty k^2 dk = \frac{V}{2\pi^2} \int_0^\infty k^2 dk.$$

The final step is to change the integration variable from k to the energy ϵ , using the relation $\epsilon = \frac{\hbar^2 k^2}{2m}$, which gives

$$k^2 dk \rightarrow \left(\frac{2m}{\hbar^2}\right)^{3/2} \frac{\epsilon^{1/2}}{2} d\epsilon,$$

so that we can express the sum over states as

$$\sum_{\mathbf{k}} \rightarrow \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \frac{1}{2} \int_0^\infty \epsilon^{1/2} d\epsilon = V \int g(\epsilon) d\epsilon.$$

Here, we have introduced the **density of states** $g(\epsilon)$ for a non-relativistic 3D quantum gas, defined as

$$g(\epsilon) = \frac{g}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \epsilon^{1/2} = A\epsilon^{1/2}, \quad (10.2.1)$$

which gives the number of available quantum states per unit energy interval at energy ϵ , where we have included the discrete degeneracy factor g from the internal degrees of freedom. For a generic system, the strategy to compute the density of states $g(\epsilon)$ involves finding the relationship between the energy levels and the quantum numbers characterizing the states, then determining how many states fall within a given energy range (compute the volume in quantum number space corresponding to that energy range).

For a generic thermodynamic quantity that can be expressed as a sum over states $\sum_{\alpha} f(\epsilon_{\alpha})$, we can now write it as an integral over energy using the density of states:

$$\sum_{\alpha} f(\epsilon_{\alpha}) \longrightarrow V \int_0^\infty g(\epsilon) f(\epsilon) d\epsilon,$$

so that we can express the grand potential, internal energy, and average number of particles for the non-relativistic 3D quantum gas in the thermodynamic limit as³

$$\begin{aligned} \Omega_{B/F} &= \pm \frac{V}{\beta} \int_0^\infty g(\epsilon) \log \left(1 \mp e^{-\beta(\epsilon-\mu)} \right) d\epsilon, \\ E_{B/F} &= V \int_0^\infty g(\epsilon) \frac{\epsilon}{e^{\beta(\epsilon-\mu)} \mp 1} d\epsilon, \\ N_{B/F} &= V \int_0^\infty g(\epsilon) \frac{1}{e^{\beta(\epsilon-\mu)} \mp 1} d\epsilon. \end{aligned}$$

These integrals can be evaluated using special functions, such as the polylogarithm function or power series, to obtain explicit expressions for the thermodynamic quantities of interest in terms of temperature, chemical potential, and volume.

Now, by some developments of these expressions, we can get interesting results and laws.

³Remember that the internal degeneracy factor g has already been included in the density of states $g(\epsilon)$.

Equation of state. If we integrate by part the expression for the grandpotential $\Omega_{B/F}$, we get

$$\begin{aligned}\Omega_{B/F} &= \pm \frac{AV}{\beta} \left[\frac{2}{3} \epsilon^{\frac{3}{2}} \log \left(1 \mp e^{-\beta(\epsilon-\mu)} \right) \Big|_0^\infty - \int_0^\infty \frac{2}{3} \epsilon^{\frac{3}{2}} \frac{\pm \beta e^{-\beta(\epsilon-\mu)}}{1 \mp e^{-\beta(\epsilon-\mu)}} d\epsilon \right] \\ &\quad \lim_{\epsilon \rightarrow \infty} \log(1 \mp 0) = 0 \quad \text{and} \quad \epsilon^{\frac{3}{2}} \log(1 \mp e^{\beta\mu}) \Big|_{\epsilon=0} = 0, \\ A \frac{2}{3} \epsilon^{\frac{3}{2}} &= \frac{2}{3} \epsilon g(\epsilon), \\ &= \mp \frac{V}{\beta} (\pm \beta) \frac{2}{3} \int_0^\infty g(\epsilon) \frac{\epsilon e^{-\beta(\epsilon-\mu)}}{1 \mp e^{-\beta(\epsilon-\mu)}} d\epsilon = -\frac{2}{3} V \int_0^\infty g(\epsilon) \frac{\epsilon}{e^{\beta(\epsilon-\mu)} \mp 1} d\epsilon = -\frac{2}{3} E_{B/F},\end{aligned}$$

where we used the expression for the internal energy $E_{B/F}$ in the last step. Thus we have derived an important relation between the grand potential and the internal energy for a non-relativistic 3D quantum gas: since the pressure p is related to the grand potential by $p = -\frac{\Omega}{V}$, we can express the equation of state for the quantum gas as

$$pV = \frac{2}{3} E_{B/F}.$$

10.2.1 | Power Series Expansion

Now we have expressions for the thermodynamic quantities as integrals over energy, we can evaluate these integrals using power series expansions. We will analyze the average number of particles $N_{B/F}/V = n_{B/F}$ and the pressure $p_{B/F}$ in detail

$$\begin{aligned}n_{B/F} &= A \int_0^\infty \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} \mp 1} d\epsilon, \\ p_{B/F} &= \frac{2}{3} A \int_0^\infty \frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)} \mp 1} d\epsilon.\end{aligned}$$

The idea is to perform a change of variable, since these expressions are impossible to evaluate directly written in this form (we will not get a simple solution, but a form from which it will be easy at least to compute the limits in temperature). Let us define the fugacity $z = e^{\beta\mu}$ and perform the change

$$\beta\epsilon = x^2 \implies \beta d\epsilon = 2x dx,$$

which will let us find a power series and also gaussian integrals, as we will show. Thus, it is possible to rewrite the expressions for the particle density and pressure as

$$\begin{aligned}n_{B/F} &= C \int_0^\infty \frac{zx^2}{e^{x^2} \mp z} dx = C \int_0^\infty \frac{zx^2 e^{-x^2}}{1 \mp ze^{-x^2}} dx, \\ p_{B/F} &= \frac{2}{3\beta} C \int_0^\infty \frac{zx^4}{e^{x^2} \mp z} dx = \frac{2}{3\beta} C \int_0^\infty \frac{zx^4 e^{-x^2}}{1 \mp ze^{-x^2}} dx,\end{aligned}$$

where we have defined the constant⁴

$$C = \frac{4g}{\sqrt{\pi}} \frac{1}{\lambda_T^3}, \quad \lambda_T = \sqrt{\frac{2\pi\hbar^2}{mk_B T}}.$$

In both integrals we find the term $\frac{ze^{-x^2}}{e^{-x^2} \mp z}$, which we can expand as a power series in ze^{-x^2} paying attention to the convergence conditions and the signs:

⁴It comes from the change of variable, which left us with $\frac{2A}{\beta^{3/2}} = 2 \frac{g}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \beta^{-3/2}$.

- for bosons, we have

$$\frac{ze^{-x^2}}{1 - ze^{-x^2}} = ze^{-x^2} \sum_{l=0}^{\infty} (ze^{-x^2})^l = \sum_{l=0}^{\infty} (ze^{-x^2})^{l+1},$$

which converges for $|ze^{-x^2}| < 1$, always true since $z < 1$ for bosons;

- for fermions, we have

$$\frac{ze^{-x^2}}{1 + ze^{-x^2}} = ze^{-x^2} \sum_{l=0}^{\infty} (-1)^l (ze^{-x^2})^l = \sum_{l=0}^{\infty} (-1)^l (ze^{-x^2})^{l+1},$$

which converges for all $z \notin [-\infty, -1]$ for the alternating series, but since for fermions $z > 0$ it always converges;

- we can thus write a single expression for both cases:

$$\frac{ze^{-x^2}}{e^{-x^2} \mp z} = \sum_{l=0}^{\infty} (\pm 1)^l (ze^{-x^2})^{l+1}.$$

Substituting this series expansion into the integrals for $n_{B/F}$ and $p_{B/F}$, we get

$$\begin{aligned} n_{B/F} &= C \sum_{l=0}^{\infty} (\pm 1)^l z^{l+1} \int_0^{\infty} x^2 e^{-(l+1)x^2} dx, \\ p_{B/F} &= \frac{2}{3\beta} C \sum_{l=0}^{\infty} (\pm 1)^l z^{l+1} \int_0^{\infty} x^4 e^{-(l+1)x^2} dx. \end{aligned}$$

The integrals appearing in these expressions are standard gaussian integrals, which can be computed using the formula

$$\int_0^{\infty} x^{2n} e^{-ax^2} dx = \frac{(2n-1)!!}{2^{n+1}} \sqrt{\frac{\pi}{a^{2n+1}}}, \quad a > 0,$$

where $(x)!! = (x)(x-2)\cdots 3 \cdot 1$ is the double factorial⁵. Applying this formula to our integrals, we find

$$\begin{aligned} \int_0^{\infty} x^2 e^{-(l+1)x^2} dx &= \frac{\sqrt{\pi}}{4} \frac{1}{(l+1)^{3/2}}, \\ \int_0^{\infty} x^4 e^{-(l+1)x^2} dx &= \frac{3\sqrt{\pi}}{8} \frac{1}{(l+1)^{5/2}}. \end{aligned}$$

Substituting these results back into the expressions for $n_{B/F}$ and $p_{B/F}$, we obtain

$$\begin{aligned} n_{B/F} &= C \sum_{l=1}^{\infty} (\pm 1)^l z^{l+1} \frac{\sqrt{\pi}}{4} \frac{1}{(l+1)^{3/2}} = \frac{g}{\lambda_T^3} \sum_{l=0}^{\infty} (\pm 1)^l \frac{z^{l+1}}{(l+1)^{3/2}}, \\ p_{B/F} &= \frac{2}{3\beta} C \sum_{l=1}^{\infty} (\pm 1)^l z^{l+1} \frac{3\sqrt{\pi}}{8} \frac{1}{(l+1)^{5/2}} = \frac{g}{\beta \lambda_T^3} \sum_{l=0}^{\infty} (\pm 1)^l \frac{z^{l+1}}{(l+1)^{5/2}}. \end{aligned}$$

These series expansions provide a way to compute the particle density and pressure of the quantum gas in terms of the fugacity z and temperature T . They are particularly useful for analyzing the behavior of the system in different regimes, such as low or high temperatures, and for comparing

⁵Starting from $2n-1$ we know it is an odd number and thus the product goes down by 2 each time until the unity.

the properties of bosonic and fermionic gases. We can make the notation lighter by defining the functions

$$\begin{aligned} f_\nu(z) &= \sum_{l=0}^{\infty} (-1)^l \frac{z^{l+1}}{(l+1)^\nu} \quad (\text{Fermions}), \\ g_\nu(z) &= \sum_{l=0}^{\infty} \frac{z^{l+1}}{(l+1)^\nu} \quad (\text{Bosons}), \end{aligned} \quad (10.2.2)$$

so that we can rewrite the expressions for the particle density and pressure as

$$\begin{aligned} n_{B/F} &= \frac{g}{\lambda_T^3} \begin{cases} f_{3/2}(z), & (\text{Fermions}); \\ g_{3/2}(z), & (\text{Bosons}); \end{cases} \\ p_{B/F} &= \frac{g}{\beta \lambda_T^3} \begin{cases} f_{5/2}(z), & (\text{Fermions}); \\ g_{5/2}(z), & (\text{Bosons}). \end{cases} \end{aligned}$$

10.2.2 | Classical Limit

We can now analyze the classical limit of the quantum gas, by stopping at the first order ($l = 0$) in the power series expansions derived above: this is justified because in the classical limit the fugacity z is much smaller than 1 ($z \ll 1$), so higher-order terms in the series become negligible. Thus, for the density we have

$$n = \frac{g}{\lambda_T^3} z \implies z = \frac{n \lambda_T^3}{g},$$

which can be substituted back into the expressions for the pressure to obtain

$$p = \frac{g}{\beta \lambda_T^3} z = \frac{g}{\beta \lambda_T^3} \frac{n \lambda_T^3}{g} = \frac{n}{\beta} = nk_B T,$$

recovering the classical ideal gas law $pV = Nk_B T$.

This limit corresponds to the regime where quantum effects are negligible, and the behavior of the gas can be described by classical statistics. The classical limit is typically valid when the thermal wavelength $\lambda_T = \sqrt{\frac{2\pi\hbar^2}{mk_B T}}$ is much smaller than the average interparticle distance, which occurs at **high temperatures or low densities**:

$$z \ll 1 \implies \text{high } T \text{ or low } n.$$

In this regime, both bosonic and fermionic quantum gases behave like classical ideal gases, and the differences in their quantum statistics become irrelevant: we can show this result explicitly by looking at the average occupation numbers for bosons and fermions in the limit $z \ll 1$:

$$\langle n_\alpha \rangle_{gc}^{B/F} = \frac{1}{e^{\beta(\epsilon_\alpha - \mu)} \mp 1} \sim e^{-\beta(\epsilon_\alpha - \mu)} \quad \text{for } z \ll 1,$$

which is the Maxwell-Boltzmann distribution for classical particles.

A final remark can be made about the expression of the chemical potential in the classical limit: since $z = e^{\beta\mu} \ll 1$, we have $\beta\mu \ll 0$ and thus we can recover

$$n = \frac{g}{\lambda_T^3} e^{\beta\mu} \implies \mu = -\frac{3}{2} k_B T \log \left(\frac{mk_B T}{2\pi\hbar^2} \left(\frac{n}{g} \right)^{\frac{2}{3}} \right) \xrightarrow{T \rightarrow \infty} -\infty,$$

which shows that in the classical limit the chemical potential diverges to negative infinity as the temperature increases.

Semiclassical expansion. If we now consider the next order in the power series expansions for the particle density and pressure, we can derive corrections⁶ to the ideal gas law that account for quantum effects. Including the second term ($l = 1$) in the series, we have

$$n_{B/F} = \frac{g}{\lambda_T^3} \left(z \pm \frac{z^2}{2^{3/2}} \right),$$

$$p_{B/F} = \frac{g}{\beta \lambda_T^3} \left(z \pm \frac{z^2}{2^{5/2}} \right).$$

Solving for z in terms of n , we find

$$\frac{z^2}{2\sqrt{2}} \pm z \mp \frac{n\lambda_T^3}{g} = 0 \implies z = \sqrt{2} \left(\mp 1 \pm \sqrt{1 \pm \sqrt{2} \frac{n\lambda_T^3}{g}} \right),$$

we solved the second order equation: the first and third signs are due to Bosonic or fermionic nature, the second to the choice of solution of the quadratic equation. Expanding the square root for small $n\lambda_T^3/g$, since $z \ll 1$ implies small n , we can use the Taylor expansion of the radical $\sqrt{1+x} \sim 1 + \frac{x}{2} - \frac{x^2}{8} + \dots$ for $x \ll 1$, leading to

$$z_{B/F} \sim \sqrt{2} \left[\mp 1 \pm \left(1 \pm \frac{1}{\sqrt{2}} \frac{n\lambda_T^3}{g} - \frac{1}{4} \left(\frac{n\lambda_T^3}{g} \right)^2 \right) \right] = \frac{n\lambda_T^3}{g} \mp \frac{1}{2\sqrt{2}} \left(\frac{n\lambda_T^3}{g} \right)^2.$$

If we now substitute this expression for z back into the pressure, and keep terms up to second order in $n\lambda_T^3/g$, we find

$$p_{B/F} = \frac{g}{\beta \lambda_T^3} \left[\frac{n\lambda_T^3}{g} \mp \frac{1}{2\sqrt{2}} \left(\frac{n\lambda_T^3}{g} \right)^2 \pm \frac{1}{4\sqrt{2}} \left(\frac{n\lambda_T^3}{g} \right)^2 \right] = nk_B T \left[1 \mp \frac{n_{B/F}}{4\sqrt{2}} \frac{\lambda_T^3}{g} \right].$$

This expression represents the semiclassical expansion of the equation of state for a quantum gas, showing how quantum statistics modify the ideal gas law. The correction term $\mp \frac{n_{B/F}}{4\sqrt{2}} \frac{\lambda_T^3}{g}$ indicates:

- For bosons (upper sign), the pressure is reduced compared to the classical ideal gas, reflecting the tendency of bosons to cluster together in the same quantum state.
- For fermions (lower sign), the pressure is increased due to the Pauli exclusion principle, which prevents multiple fermions from occupying the same quantum state, leading to a higher effective pressure.
- These corrections become significant at low temperatures or high densities, where quantum effects dominate the behavior of the gas.

We get the idea of quantum gases being *attractive* or *repulsive* depending on their statistics, leading to different macroscopic behaviors compared to classical gases: it is not a real interaction between particles (we are studying free gases) but an effective behavior due to the underlying quantum statistics. Without interactions, bosons still tend to bunch together while fermions repel each other due to the Pauli exclusion principle, but these are statistical effects which manifest in the macroscopic properties of the gas, since they influence how particles distribute themselves among the available quantum states. We have to remember that these corrections are valid only in the semiclassical regime, where quantum effects are present but not dominant, hence the fugacity z is still much smaller than 1 along with the density n .

Thus the quantum corrections to the classical ideal gas law becomes small when:

⁶We are considering semiclassical limit, called like that because the statistic is almost classical, but we are in a regime where we cannot ignore contributions due to the quantum nature of the considered particles.

1. g is very large (high internal degeneracy): this reduces the effect of quantum statistics by providing more available states for the particles to occupy;⁷
2. the thermal wavelength λ_T is very small (high temperature);⁸
3. the particle density n is very small (low density, again leading to classical statistics).

⁷Fermions cannot occupy the same state, but if we have a very large degeneracy in the energy levels, then we can have a lot of fermions at the same energy, thus mimicking the Bosons behavior.

⁸We are speaking of "small" and "large" in relative terms: we intend that λ_T is small compared to the Plank constant and other characteristic lengths of the system.

10.3 | Fermi-Dirac Gas

As the temperature decreases towards absolute zero, the fugacity $z = e^{\beta\mu}$ increases, challenging the validity of simple virial approximations. For Bosons, z approaches unity from below ($z \rightarrow 1^-$), pushing the expansion to the edge of its radius of convergence. For Fermions, z can diverge to infinity ($z \gg 1$), exceeding the radius of convergence of standard power series. Consequently, we cannot truncate the expansion at the first few terms; we must instead rely on the full definitions of the Fermi-Dirac and Bose-Einstein functions:

$$f_\nu(z) = \sum_{l=0}^{\infty} (-1)^l \frac{z^{l+1}}{(l+1)^\nu} \quad (\text{Fermions}),$$

$$g_\nu(z) = \sum_{l=0}^{\infty} \frac{z^{l+1}}{(l+1)^\nu} \quad (\text{Bosons}).$$

While the Bosonic series $g_\nu(z)$ remains valid up to $z = 1$, the Fermionic series $f_\nu(z)$ is only convergent for $|z| \leq 1$, requiring analytical continuation (such as the Sommerfeld expansion) in the degenerate regime.

Absolute convergence. We can study the absolute convergence of the series expansion (for both the Bose-Einstein functions $g_\nu(z)$ and the Fermi-Dirac functions $f_\nu(z)$) by considering the sum of the absolute values of its terms. Since the alternating sign $(\pm 1)^l$ vanishes inside the modulus, we obtain:

$$\sum_{l=0}^{\infty} \left| (\pm 1)^l \frac{z^{l+1}}{(l+1)^\nu} \right| = \sum_{l=0}^{\infty} \frac{|z|^{l+1}}{(l+1)^\nu}.$$

Applying the standard ratio test,⁹ this series converges absolutely if and only if:

$$\lim_{l \rightarrow \infty} \left| \frac{z^{l+2}}{z^{l+1}} \left(\frac{l+1}{l+2} \right)^\nu \right| \sim |z| < 1.$$

While for Bosons this series is always convergent (they always satisfy $|z| < 1$, since $\mu \leq 0$), for Fermions the fugacity z can take any positive value, including values greater than 1 when the chemical potential μ is positive. Therefore, we have absolute convergence only for $|z| < 1$. For $z > 1$, the series may or may not be *simply convergent*,¹⁰ and it can be necessary to evaluate the functions $f_\nu(z)$ using the integral representation (analytic continuation) rather than this power series. Usually, for fermions at very low temperatures, the *Sommerfeld expansion* is used to approximate these functions for large z (as an asymptotic technique used to approximate integrals involving the Fermi-Dirac distribution by expanding in powers of the temperature T), but this is beyond the scope of this course.

Thus, for fermions in the classical regime the series converges absolutely for all physically relevant values of z , while the problem of convergence arises only in the degenerate regime at low temperatures when z becomes large.

⁹The standard ratio test consists in taking the limit of the absolute value of the fraction a_{n+1}/a_n , and then checking if this limit is less than 1 for convergence.

¹⁰Simple convergence (also known as *pointwise convergence*) means that the sequence of functions converges to a limit for each fixed point x individually, but the rate of convergence may vary depending on x .

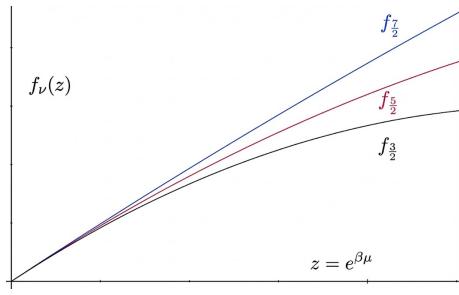


Figure 10.2: Plot of the Fermi-Dirac power series $f_\nu(z)$ for different values of ν in the classical limit. If the series is well defined and convergent for all $z \geq 0$, then it is positive, continuous and monotonically increasing. For z near 0 the series assumes a linear behavior, while for large z it grows more slowly.

We are looking at the limit for small temperatures $T \rightarrow 0$, such that $\beta \rightarrow \infty$ and thus $z = e^{\beta\mu} \rightarrow \infty$ if $\mu > 0$: recalling the integral expressions for the particle density and energy density (since they are valid for all temperatures in general)¹¹

$$\begin{aligned} n &= \frac{N}{V} = A \int_0^\mu d\epsilon \epsilon^{1/2} \frac{1}{e^{\beta(\epsilon-\mu)} + 1} = \int_0^\infty d\epsilon g(\epsilon) \langle n(\epsilon) \rangle_{gc}^{FD}, \\ \epsilon &= \frac{E}{V} = A \int_0^\mu d\epsilon \epsilon^{3/2} \frac{1}{e^{\beta(\epsilon-\mu)} + 1} = \int_0^\infty d\epsilon \epsilon g(\epsilon) \langle n(\epsilon) \rangle_{gc}^{FD}, \end{aligned}$$

where it is crucial to know that the *density of states* $g(\epsilon)$ has unique definitions valid for the system under consideration (remember we are in the non-relativistic 3D quantum gas setting). Now it is of maximum interest to understand how $\langle n(\epsilon) \rangle_{gc}^{FD}$ behaves in this limit: plotting the function for $\beta \rightarrow \infty$, and we find that it behaves as a *Heaviside step function*

$$\langle n(\epsilon) \rangle_{gc}^{FD} = \frac{1}{e^{\beta(\epsilon-\mu)} + 1} = \begin{cases} 1, & \epsilon < \mu; \\ 1/2, & \epsilon = \mu; \\ 0, & \epsilon > \mu. \end{cases}$$

This is the famous **Fermi-Dirac distribution at zero temperature**, which shows that all states with energy below the chemical potential μ are fully occupied, while those above μ are empty. The chemical potential at zero temperature is known as the **Fermi energy** ϵ_F , which defines the highest occupied energy level in the system at absolute zero.

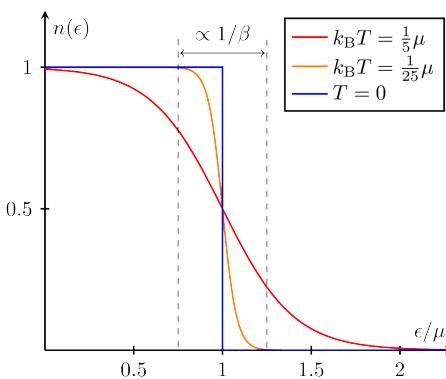


Figure 10.3: Plot of the Fermi-Dirac distribution at zero temperature, which resembles a Heaviside step function. All states with energy below the chemical potential μ (the Fermi energy ϵ_F) are fully occupied, while those above are empty; while we let T increase the step function smooths out as thermal excitations allow some particles to occupy higher energy states.

Thus, we can compute the integrals for n and ϵ at zero temperature as

$$\begin{aligned} n &= \int_0^{\mu=\epsilon_F} g(\epsilon) \cdot 1 d\epsilon = A \int_0^\mu \epsilon^{1/2} d\epsilon = \frac{2}{3} A \epsilon_F^{3/2}, \\ \epsilon &= \frac{1}{n} \int_0^{\mu=\epsilon_F} g(\epsilon) \epsilon \cdot 1 d\epsilon = \frac{A}{n} \int_0^\mu \epsilon^{3/2} d\epsilon = \frac{2}{5} \frac{A}{n} \epsilon_F^{5/2} = \frac{3}{5} \epsilon_F. \end{aligned}$$

¹¹Remember for the following computation that the discrete internal degeneracy is included in A or $g(\epsilon)$.

We can also find an explicit expression for the Fermi energy ϵ_F in terms of the particle density n :

$$\epsilon_F = \left(\frac{3n}{2A} \right)^{2/3}, \quad (10.3.1)$$

so that we can control the behavior of the system at zero temperature acting on the density of particles.

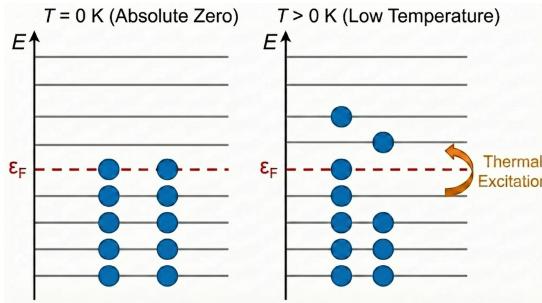
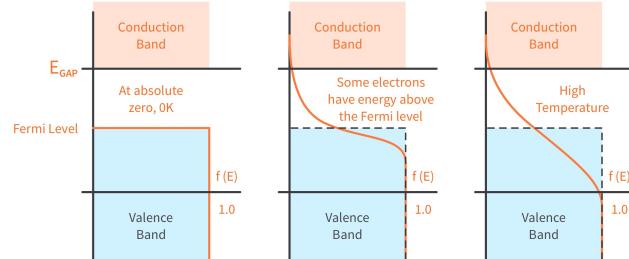


Figure 10.4: Occupation of energy levels by electrons in a metal at zero temperature, illustrating the filling up to the Fermi energy ϵ_F . As temperature increases, electrons can be thermally excited to higher energy states above ϵ_F .

We can study how metals behave under these conditions, since the conduction electrons in metals can be modeled as a Fermi gas. At low temperatures, the electrons fill up energy levels up to the Fermi energy, leading to unique electronic properties, such as electrical conductivity and heat capacity, which differ significantly from classical predictions. We can dope the metal to change the electron density n , which in turn modifies the Fermi energy ϵ_F and thus the electronic properties of the material.

Figure 10.5: Band diagram of a semiconductor showing the Fermi level in relation to the conduction band and valence band.



We can also compute the grand potential Ω_F at zero temperature using the relation $\Omega_F = -\frac{2}{3}E_F$ with the intent to find the pressure of the Fermi gas at absolute zero:

$$\Omega_F = -\frac{2}{3}E = -pV, \quad p = \frac{2}{3}\frac{E}{V} = \frac{2}{3}n\epsilon = \frac{2}{3}n\left(\frac{3}{5}\epsilon_F\right) = \frac{2}{5}n\epsilon_F > 0.$$

For any perfect classical gas we have a zero pressure at zero temperature, while here we have a finite pressure due to the Pauli exclusion principle, which prevents all fermions from occupying the lowest energy state. This leads to a degeneracy pressure that supports the system against collapse, which is particularly important in astrophysical contexts such as white dwarf stars: the electron degeneracy pressure arising from the Fermi gas of electrons counteracts gravitational collapse, stabilizing the star.

Fermi temperature. With an augment in temperature, the sharp occupation of energy levels at the Fermi energy smooths out, leading to thermal excitations around $k_B T$ of electrons near the Fermi level, affecting the electronic properties of the material. It is thus really useful to define a Fermi temperature T_F associated to the Fermi energy:

$$k_B T_F = \epsilon_F \implies T_F = \frac{\epsilon_F}{k_B},$$

which provides a temperature scale below which quantum effects become significant for the fermi gas. At temperatures much lower than T_F , the gas exhibits quantum behavior, with non negligible behaviours due to statistical properties, while at temperatures much higher than T_F , it behaves more classically

$$T_F = \frac{1}{k_B} \left(\frac{6\pi^2}{g} \right)^{2/3} \frac{\hbar^2}{2m} n^{2/3}.$$

This fact is not arbitrary: it comes from the comparison between the thermal energy $k_B T$ and the Fermi energy ϵ_F , in the denominator of the Fermi-Dirac distribution. When $k_B T \ll \epsilon_F$, the thermal excitations are small compared to the energy scale set by the Fermi energy, leading to quantum behavior. Conversely, when $k_B T \gg \epsilon_F$, thermal excitations dominate, and the gas behaves more classically (i.e. the occupation of energy levels follows the Maxwell-Boltzmann distribution):

$$\frac{1}{e^{\beta(\epsilon-\mu)} + 1} \xrightarrow{k_B T \gg \epsilon_F} e^{-\beta(\epsilon-\mu)},$$

considering ϵ near the Fermi energy.

We can derive an approximate value for a metal with $n \sim 10^{22/23}$ electron/cm³, $g = 2$ (for spin degeneracy) and m equal to the electron mass:

$$T_F \sim 10^{4-5} K,$$

so that at room temperature ($\sim 300K$) the electron gas in metals can be treated as a degenerate Fermi gas, with quantum effects dominating its behavior. For stars like white dwarfs, where the electron density is extremely high ($n \sim 10^{30}$ electron/cm³), the Fermi temperature can reach values on the order of $10^{11} K$ or higher (where the star temperature is much lower, around 10^7 Kelvin), indicating that the electron gas remains degenerate even at very high temperatures.

10.4 | Bose-Einstein Gas

We can now analyze the behavior of a Bose gas at low temperatures, where quantum effects become significant and lead to phenomena such as Bose-Einstein condensation. Starting from the expressions for the particle density and the pressure:

$$n = \frac{g}{\lambda_T^3} g_{3/2}(z),$$

$$p = \frac{g}{\beta \lambda_T^3} g_{5/2}(z),$$

where again we have defined the Bose-Einstein functions $g_\nu(z)$ as a power series in the fugacity z :

$$g_\nu(z) = \sum_{l=0}^{\infty} \frac{z^{l+1}}{(l+1)^\nu}.$$

We can study the behavior of these functions as z approaches 1 (low temperatures) and as z is very small compared to 1 (high temperatures). These are the two regimes of interest for the Bose gas: the quantum degenerate regime and the classical limit. We already discussed the classical and semiclassical limit in 10.2.2, thus we will now focus on the quantum degenerate regime where characteristic effects of bosons emerge.

To analyze the behavior of the **Bose gas in the quantum degenerate regime**, we need to consider the full power series expansion for the Bose-Einstein functions $g_\nu(z)$: we need to check convergence of the series as z approaches 1 (low temperatures). The series converges for all $z < 1$, but as z approaches 1, the terms in the series become larger and larger, leading to a divergence at $z = 1$: we cannot simply truncate the series at a finite number of terms, as we did in the classical limit.

We will consider both the particle density and pressure expressions, so we need the series expansion for both $g_{3/2}(z)$ and $g_{5/2}(z)$. Starting from the general expression

$$g_\nu(z) = \sum_{l=0}^{\infty} \frac{z^{l+1}}{(l+1)^\nu},$$

we have to analyze $\nu = 3/2$ and $\nu = 5/2$:

1. For $\nu = 3/2$:

$$g_{3/2}(z) = \sum_{l=0}^{\infty} \frac{z^{l+1}}{(l+1)^{3/2}}.$$

TODO: insert plot of $g_{\nu}(z)$ for $\nu=3/2$ and $5/2$.

As z approaches 1, the series diverges, indicating that the particle density n also diverges. In particular

- For $|z| < 1$, the series is **absolutely convergent**;
- For $|z| = 1$, the series is **simply convergent**;
- For $|z| > 1$, the series is **divergent**.

These properties should not surprise us, since bosons can have $\mu(T) \leq 0$ and thus $z \leq 1$: the only physically relevant case is the first one, where the series converges absolutely.

2. For $\nu = 5/2$:

$$g_{5/2}(z) = \sum_{l=0}^{\infty} \frac{z^{l+1}}{(l+1)^{5/2}}.$$

Similar considerations apply here: as z approaches 1, the series diverges, indicating that the pressure p also diverges. The series diverges as $|z|$ grows, but it is defined for all values of the fugacity, even for $|z| > 1$.

We now intend to study the dependence of the **chemical potential** μ on the temperature T in this quantum degenerate regime.

- At high temperatures, in the classical limit, $\mu \rightarrow -\infty$ as $T \rightarrow \infty$, leading to $z \ll 1$.
- As the temperature decreases, $\mu \rightarrow 0^-$, leading to an increase in the fugacity z towards 1.
- The chemical potential is a monotonically decreasing function of temperature

$$\frac{\partial \mu}{\partial T} \leq 0,$$

approaching 0 from below as T approaches a critical temperature T_c ; this is true because a thermodynamic potential is expressed like that and verifies this property.

TODO: insert plot of mu vs T for Bose gas.

Thus we can define the critical temperature T_c for Bose-Einstein condensation as the temperature at which the chemical potential μ reaches 0, leading to $z = 1$; there are two possible cases:

- $T_c = 0$ so that $\mu(T_c) = 0 \implies$ we are in the 2D-Bose gas case, where no condensation occurs at finite temperature (we will show this later in the exercises, as well as the case of 2D Bose gas trapped in a harmonic potential, which instead shows condensation at finite temperature);
- $T_c \neq 0$ so that $\mu(T_c) = 0 \implies$ we are in the 3D-Bose gas case, where condensation occurs at a finite temperature (below T_c).

Setting our analysis in the 3D case, we can find an expression for the critical temperature T_c by setting $\mu = 0$ (thus $z = 1$) in the expression for the particle density:

$$n = \frac{g}{\lambda_{T_c}^3} g_{3/2}(1) = \frac{g}{\lambda_{T_c}^3} \zeta(3/2),$$

where we have used the Riemann zeta function $\zeta(\nu) = \sum_{l=0}^{\infty} \frac{1}{(l+1)^{\nu}} = g_{\nu}(1)$ to express $g_{3/2}(1)$. Solving for T_c , we find

$$T_{BEC} = \frac{2\pi\hbar^2}{mk_B} \left(\frac{n}{g\zeta(3/2)} \right)^{2/3}.$$

This expression shows that the critical temperature for **Bose-Einstein condensation** depends on the particle density n , the mass of the bosons m , and the internal degeneracy g .

10.4.1 | Bose-Einstein Condensation

Below this critical temperature, a macroscopic number of bosons occupy the ground state, leading to the phenomenon of Bose-Einstein condensation, which has been observed experimentally in systems such as ultracold atomic gases. This happens at temperatures comprised in the range of nanoKelvins to microKelvins, since the chemical potential remains equal to zero in this whole range of temperatures below T_{BEC} .

So the occupation number for temperatures below T_{BEC} can be expressed as

$$n(T < T_{BEC}) = \frac{g}{\lambda_T^3} \zeta(3/2).$$

We have to pay attention reading the consequences of this result: as temperature decreases below T_{BEC} and approaches absolute zero, $\lambda_T = \left(\frac{2\pi\hbar^2}{mk_B T}\right)^{1/2} \xrightarrow{T \rightarrow 0} \infty$ diverges, leading to a decrease in the particle density n until the absolute zero is reached:

$$n \xrightarrow{\lambda_T \rightarrow \infty} 0.$$

Particles cannot disappear as temperature decreases, so we have to conclude that the expression for the particle density above only accounts for particles in excited states (with $\epsilon > 0$), while a macroscopic number of particles condense into the ground state (with $\epsilon = 0$). If we consider, indeed, the number of particles in the ground state for $\epsilon = 0$, we can write

$$\langle n_0 \rangle_{gc}^B = \frac{1}{e^{\beta(0-\mu)} - 1} = \frac{1}{e^{-\beta\mu} - 1} = \frac{z}{1-z} \xrightarrow{z \rightarrow 1} \infty,$$

indicating that as μ approaches 0 (i.e., as z approaches 1), the occupation number of the ground state diverges, leading to a macroscopic occupation of this state. But we can be more precise: the total number of particles N in the system can be fixed, so it cannot diverge statistically; we can express it as the sum of particles in the ground state N_0 and particles in excited states N_{exc} : starting again from the number of states in the momentum space

$$N = \sum_{\mathbf{k}} n(\epsilon_{\mathbf{k}}), \quad \epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}}{2m}, \quad \mathbf{k} = \frac{2\pi}{L} (n_x, n_y, n_z),$$

if we compute the TD limit the number of particles becomes

$$N = \int_0^\infty d\epsilon g(\epsilon) \langle n(\epsilon) \rangle_{gc}^B = N_0 + \int_{\epsilon>0} d\epsilon g(\epsilon) \langle n(\epsilon) \rangle_{gc}^B = N_0 + N_{exc}.$$

This is the crucial point: since $\langle n(0) \rangle_{gc}^B \Big|_{\mu=0} \rightarrow \infty$ we formally need to separate the contribution of the ground state (with $\epsilon = 0$) from the contributions of all excited states (with $\epsilon > 0$), even if the integral over all energies wasn't seeing the divergence: it is only one point and computing the integral does not capture the divergence. Thus, we can express the number of particles as

$$N = N_0 + \sum_{\mathbf{k}} N_{\epsilon_{\mathbf{k}}},$$

N_0 being the number of particles in the ground state, and the sum being over all excited states. In the TD limit, the number of particles divided by the volume becomes¹²

$$n = n_0(T) + \frac{1}{V} \int_0^\infty d\epsilon g(\epsilon) \langle n(\epsilon) \rangle_{gc}^B,$$

an equivalently, after substituting the expression for the density of states and the Bose-Einstein distribution, we find

$$n = n_0(T) + \frac{g}{\lambda_T^3} g_{3/2}(z) = n_0(T) + n_n(T). \quad (10.4.1)$$

since we call the number of particles in excited states $n_n(T)$, the "**normal**" component of the gas. If we now take the limit $\mu \rightarrow 0$ (thus $z \rightarrow 1$) we find

$$n = n_0(T) + \frac{g}{\lambda_T^3} \zeta(3/2),$$

¹²Where the integral is again computed over all values of ϵ , since the ground state contribution has already been separated out formally and, as we said, the integral does not see the divergence practically.

where $\zeta(3/2)$ is the Riemann zeta function, equal to $g_{3/2}(1)$. We can derive $\zeta(3/2)$ by computing $n(T_{BEC})$: at the critical temperature T_{BEC} we can assume $n_0 = 0$, since condensation is just starting, so that

$$n = \frac{g}{\lambda_{T_{BEC}}^3} \zeta(3/2) \implies \zeta(3/2) = \frac{n \lambda_{T_{BEC}}^3}{g}.$$

If we compute the normal density for temperatures below T_{BEC} , the number of particles in excited states is given by

$$n_n(T < T_{BEC}) = \frac{g}{\lambda_T^3} \zeta(3/2) = \frac{g}{\lambda_T^3} \frac{n \lambda_{T_{BEC}}^3}{g} = n \left(\frac{T}{T_{BEC}} \right)^{3/2}, \quad (10.4.2)$$

indicating that not all particles can be accommodated in excited states. We have made no assumptions on the nature and behaviour of $n_0(T)$, but now that particles are disappearing from the excited states ($\frac{T}{T_{BEC}} < 1$), we are led to believe that a macroscopic occupation of the ground state is happening, so that the number of particles in the ground state becomes:

$$n_0 = n - n_n = n \left[1 - \left(\frac{T}{T_{BEC}} \right)^{3/2} \right]. \quad (10.4.3)$$

This expression shows that as the temperature decreases below T_{BEC} , the number of particles in the ground state N_0 increases, reaching a maximum of N at absolute zero. This phenomenon is the hallmark of Bose-Einstein condensation, where a significant fraction of bosons occupy the lowest energy state, leading to unique macroscopic quantum phenomena such as superfluidity and coherence.

Bose-Einstein Condensation as a Phase Transition

Thus our refined expression for the density of particles in a Bose gas becomes

$$n = n_0(T) + n_n(T).$$

Now we can compute the two components separately for the two regimes of temperature:

- For $T \geq T_{BEC}$, there is no condensation, so all particles are in excited states:

$$T \geq T_{BEC} \begin{cases} n_0(T) = 0, \\ n_n(T) = \frac{g}{\lambda_T^3} g_{3/2}(z) = n(T). \end{cases}$$

In particular, when $T = T_{BEC}$, we have $z = 1$ and thus

$$n_n(T_{BEC}) = \frac{g}{\lambda_{T_{BEC}}^3} \zeta(3/2) = n(T_{BEC}).$$

- For $T < T_{BEC}$, condensation occurs, so we have a macroscopic occupation of the ground state:

$$T < T_{BEC} \begin{cases} n_0(T) = n \left[1 - \left(\frac{T}{T_{BEC}} \right)^{3/2} \right], \\ n_n(T) = \frac{g}{\lambda_T^3} \zeta(3/2) = n \left(\frac{T}{T_{BEC}} \right)^{3/2}. \end{cases}$$

Thus we can consider $n_0(T)$ as an order parameter for the **Bose-Einstein condensation phase transition**, which is zero above T_{BEC} and non-zero below it, indicating the onset of macroscopic quantum coherence in the system. There will be a SSB of the gauge symmetry associated to the phase of the condensate wavefunction, leading to unique physical properties such as superfluidity and coherence.

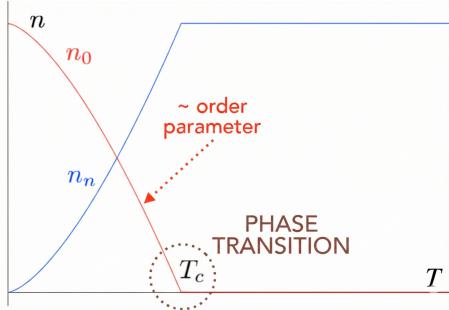


Figure 10.6: Order parameter $n_0(T)/n$ for the Bose-Einstein condensation phase transition as a function of temperature T . The order parameter is zero above the critical temperature T_{BEC} and increases as the temperature decreases below T_{BEC} , indicating the onset of macroscopic occupation of the ground state.

We will learn more about the implications of this phase transition by studying the thermodynamical quantities of the Bose gas in the two temperature regimes.

Remark. We will see in the next few sections that this phase transition has somehow its own thermodynamical signatures, as discontinuities in the specific heat and other thermodynamic quantities at the critical temperature T_{BEC} which are not solely riconducible to one specific order of PT. This phase transition is a purely quantum mechanical effect, arising from the indistinguishability and bosonic nature of the particles, and has no classical counterpart since no interactions are needed.

10.4.2 | Thermodynamical Quantities

Now we can compute the pressure of the Bose gas and relate it to the energy through the grand potential Ω_B . Recalling the expression for the pressure

$$p = \frac{g}{\beta \lambda_T^3} g_{5/2}(z),$$

and since the grand potential is related to the pressure by $\Omega = -pV$ and to the energy by $\Omega = -\frac{2}{3}E$, we can write

$$E = \frac{3}{2}pV = \frac{3}{2} \frac{gV}{\beta \lambda_T^3} g_{5/2}(z).$$

Notably, the energy expression does not explicitly depend on the condensate fraction $n_0(T)$, since the ground state does not contribute to the energy in this idealized model. The energy is determined solely by the particles in excited states, which follow the Bose-Einstein distribution.

We can now compute the energy per particle E/N in the two temperature regimes, considering that the total number of particles contributing to the energy N_n is

$$n_n = \frac{N_n}{V} = \begin{cases} \frac{g}{\lambda_T^3} g_{3/2}(z) = n, & T > T_{BEC}, \\ \frac{g}{\lambda_T^3} \zeta(3/2) = n \left(\frac{T}{T_{BEC}} \right)^{3/2}, & T < T_{BEC}. \end{cases}$$

Thus we have

$$u = \frac{E}{N_n} = \begin{cases} \frac{3}{2}k_B T \frac{g_{5/2}(z)}{g_{3/2}(z)}, & T > T_{BEC}, \\ \frac{3}{2}k_B T \frac{\zeta(5/2)}{\zeta(3/2)}, & T < T_{BEC}. \end{cases} \quad (10.4.4)$$

Note that we have used the expression for the normal component $n_n(T)$ to relate the energy to the total number of particles N : only the normal component contributes to the energy and we are computing the expression for $u = \frac{E}{N} V$ accordingly. We should also note that at the critical temperature $T = T_{BEC}$, the energy per particle is continuous:

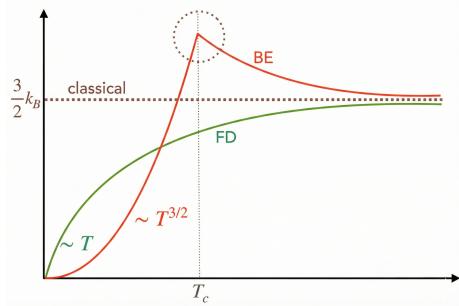
$$u\Big|_{T=T_{BEC}} = \frac{3}{2}k_B T_{BEC} \frac{\zeta(5/2)}{\zeta(3/2)},$$

since $g_{5/2}(1) = \zeta(5/2)$ and $g_{3/2}(1) = \zeta(3/2)$. Thus the energy per particle does not exhibit a discontinuity at the critical temperature, indicating that the Bose-Einstein condensation transition is a continuous phase transition in terms of energy.

If we consider the expression for the specific heat at constant volume c_V , we can derive it from the energy expression by differentiating with respect to temperature to find a typical II order PT discontinuity:¹³

$$c_V = \frac{C_V}{N} = \left(\frac{\partial u}{\partial T} \right)_V.$$

Calculating c_V in the two temperature regimes, we find that it exhibits a discontinuity at the critical temperature T_{BEC} , characteristic of a second-order phase transition, and it goes to zero as T approaches absolute zero, in accordance with the third law of thermodynamics (fixing the problem we had with the classical ideal gas).



Finally, the **entropy** S of the Bose gas can be computed using the thermodynamic relation

$$S = \frac{E - \Omega - \mu N}{T}.$$

So we can compute the three terms separately:

- The energy E is given by the expressions derived above for the two temperature regimes.
- The grand potential Ω is related to the energy by $\Omega = -(2/3)E$, with E given by the previous expression.
- The chemical potential μ is zero for $T < T_{BEC}$ and negative for $T > T_{BEC}$, approaching zero from below as T approaches T_{BEC} from above, while N is the total number of particles in the system.

¹³Since C_V is already a first derivative of a thermodynamical quantity, the energy: II order PT are characterized by discontinuities at second or higher derivatives (as here we found the discontinuity in the derivative of the first derivative of the energy).

Substituting these expressions into the entropy formula, we can derive

$$S = \begin{cases} \frac{gV}{\lambda_T^3} k_B \left(\frac{5}{2} g_{5/2}(z) - g_{3/2}(z) \log z \right), & T > T_{BEC}, \\ \frac{gV}{\lambda_T^3} k_B \frac{5}{2} \zeta(5/2), & T < T_{BEC}. \end{cases}$$

Now the entropy per particle $S/N_n = s$ can be computed in both temperature regimes, revealing how the entropy behaves across the Bose-Einstein condensation transition:

$$s = \frac{S}{N_n} = \begin{cases} \frac{5}{2} k_B \frac{g_{5/2}(z)}{g_{3/2}(z)} - k_B \log z, & T > T_{BEC}, \\ \frac{5}{2} k_B \frac{\zeta(5/2)}{\zeta(3/2)}, & T < T_{BEC}. \end{cases}$$

Here we have used the expression for the normal component N_n to relate the entropy to the total number of particles N , since we know $s(0) = 0$ for the condensate component.

We can also compute the entropy per particle precisely at the critical temperature $T = T_{BEC}$:

$$s \Big|_{T=T_{BEC}} = \frac{5}{2} k_B \frac{\zeta(5/2)}{\zeta(3/2)},$$

while at the critical temperature we have zero particles in the ground state, so all particles contribute to the entropy. As temperature decreases below T_{BEC} , the entropy $S(T < T_c)$ decreases due to the increasing occupation of the ground state, in which particles stop to contribute to the system entropy,¹⁴ reflecting the onset of macroscopic quantum coherence in the system. If we compute the entropy change per particle at the transition temperature:

$$q = T_{BEC} \Delta s = T_{BEC} s(T_{BEC}) - s(0) = \frac{5}{2} k_B T_{BEC} \frac{\zeta(5/2)}{\zeta(3/2)} \neq 0,$$

since $s(0) = 0$; this indicates a *finite entropy change* at the transition temperature T_{BEC} with non zero **latent heat**, characteristic of a first-order phase transition.

Usually PT are induced by interactions among particles, while here we have a PT in an ideal gas: this is possible because of the SSB of the gauge symmetry associated to the phase of the condensate wavefunction, which leads to a new ordered phase with lower entropy. This is a peculiar feature of Bose-Einstein condensation, where the transition is driven by quantum statistics rather than interactions. It is a strange PT indeed, since we have a **specific heat discontinuity**¹⁵ at T_{BEC} (typical of second-order PT), but we also have a **finite entropy change** at the transition (typical of first-order PT). This is a unique characteristic of the Bose-Einstein condensation phenomenon.

10.4.3 | Spontaneous Symmetry Breaking in BE Condensation

The hamiltonian of a non-interacting Bose gas is invariant under a global $U(1)$ gauge transformation, which corresponds to multiplying the bosonic field operators by a phase factor $e^{i\theta}$:

$$\hat{\psi}(\mathbf{r}) \rightarrow e^{i\theta} \hat{\psi}(\mathbf{r}), \quad \hat{\psi}^\dagger(\mathbf{r}) \rightarrow e^{-i\theta} \hat{\psi}^\dagger(\mathbf{r}).$$

¹⁴We would have noticed this behaviour from the expression of the entropy per particle, which shows only the contribution of the normal component.

¹⁵ C_V is already a first derivative and the discontinuity is found in its derivative: it is a discontinuity in second derivatives of TD potentials, typical of a continuous PT.

This symmetry reflects the conservation of particle number in the system, since it is originated from the number operator $\hat{N} = \sum_{\mathbf{k}} \hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k})$.

$$U(\theta) = e^{i\theta\hat{N}} \implies \begin{cases} U^\dagger(\theta)\hat{a}(\mathbf{k})U(\theta) = e^{i\theta}\hat{a}(\mathbf{k}), \\ U^\dagger(\theta)\hat{a}^\dagger(\mathbf{k})U(\theta) = e^{-i\theta}\hat{a}^\dagger(\mathbf{k}), \end{cases}$$

passing from the infinitesimal transformation $U(\theta) \sim \mathbb{I} + i\theta\hat{N} + O(\theta^2)$. When Bose-Einstein condensation occurs below the critical temperature T_{BEC} , a macroscopic number of bosons occupy the ground state, where

$$n_0 = \langle \hat{a}_0^\dagger \hat{a}_0 \rangle = \langle \hat{a}_0^\dagger \rangle \langle \hat{a}_0 \rangle \begin{cases} = 0, & \text{if } T > T_{BEC} \text{ since } \langle \hat{a}_0 \rangle = 0; \\ \neq 0, & \text{if } T < T_{BEC} \text{ since } \langle \hat{a}_0 \rangle \neq 0. \end{cases}$$

This non-zero expectation value of the annihilation operator \hat{a}_0 under the critical temperature indicates that the ground state has a well-defined phase, breaking the global $U(1)$ gauge symmetry of the hamiltonian. We know that $N_{\mathbf{k}}$ is invariant, since

$$N'_{\mathbf{k}} = \hat{a}'^\dagger(\mathbf{k})\hat{a}'(\mathbf{k}) = e^{-i\theta}\hat{a}^\dagger(\mathbf{k})e^{i\theta}\hat{a}(\mathbf{k}) = \hat{a}^\dagger(\mathbf{k})\hat{a}(\mathbf{k}) = N_{\mathbf{k}},$$

but, as we have seen, this is not valid for the ground state below T_{BEC} , where $\langle \hat{a}_0 \rangle \neq 0$. Thus the system chooses a specific phase for the condensate wavefunction, leading to spontaneous symmetry breaking (SSB). This SSB has profound implications for the physical properties of the Bose-Einstein condensate, including effects of quantum coherence.

10.5 | Fenomenology of BEC and Experimental Realizations

We are going to discuss some further details on Bose-Einstein condensation, including its physical interpretation, experimental realizations, and implications for quantum coherence and superfluidity. For starter we have to distinguish two regimes from the scale lengths involved:

- The **microscopic scale**, characterized by the thermal wavelength λ_T , which represents the typical de Broglie wavelength of particles at temperature T :

$$\lambda_T = \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{1/2}.$$

At low temperatures, λ_T becomes large, indicating that quantum effects become significant and particles exhibit wave-like behavior.

- The **macroscopic scale**, characterized by the interparticle distance d , which represents the average distance between particles in the gas:

$$d = n^{-1/3},$$

where n is the particle density. At high temperatures, the wavelength λ_T is much smaller than the interparticle distance d , indicating that particles behave classically and their wavefunctions are localized (like narrow deltas in space).

When the temperature decreases, λ_T increases and can become comparable to or larger than d , leading to significant overlap of particle wavefunctions and the emergence of quantum coherence. This is the regime where Bose-Einstein condensation occurs, as particles can occupy the same quantum state and form a macroscopic condensate.

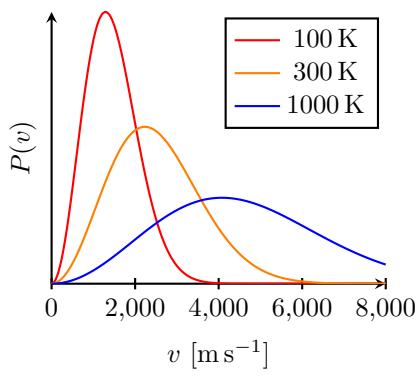


Figure 10.7: Maxwell-Boltzmann velocity distribution for a classical ideal gas at different temperatures. As temperature increases, the distribution broadens and shifts to higher velocities, indicating a wider range of speeds and higher average kinetic energy.

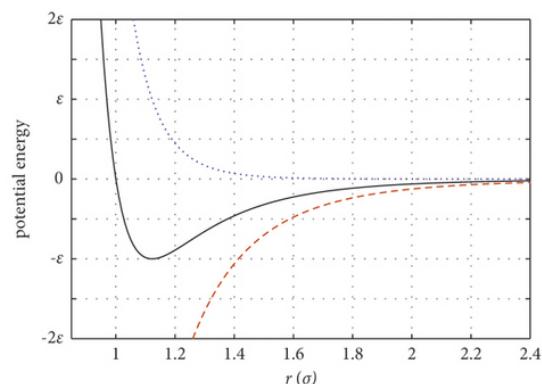


Figure 10.8: Lennard-Jones potential has a repulsive core at short distances and an attractive tail at longer distances, with a minimum at the equilibrium separation r_m . The characteristic length scale d represents the average interparticle distance in a gas.

For example we can consider Helium and Neon:

Element	Temperature (K)	d (nm)	λ_{dB} (nm)
${}^4\text{He}$	4	0.265	0.4
Ne	25	0.296	0.07

Table 10.1: Properties of Helium-4 and Neon, for comparison of characteristic lengths and thermal wavelengths: at the given temperatures, Helium-4 has a thermal wavelength larger than the interparticle distance, indicating significant quantum effects and the possibility of Bose-Einstein condensation, while Neon has a much smaller thermal wavelength, indicating classical behavior.

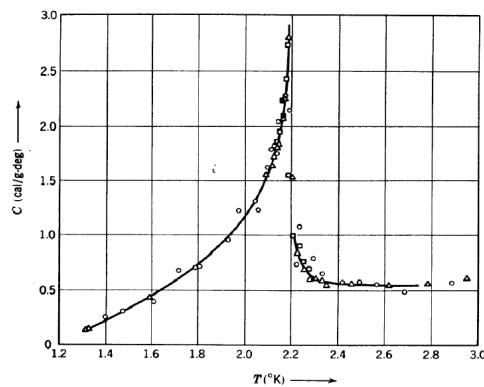
Example (Superfluidity in liquid helium-4). In 1938, Allen and Misener, and independently Kapitza, discovered that liquid helium-4 exhibits superfluidity below a critical temperature of approximately $T_c \sim 2.17\text{ K}$. This phenomenon is characterized by the ability of the liquid to flow without viscosity, allowing it to pass through narrow channels and over obstacles without resistance. We briefly discussed it in the PT chapter (phase diagram in Fig. 6.2).

If we were to measure the specific heat capacity of liquid helium-4 as a function of temperature, we would observe a sharp peak at the critical temperature T_c , known as the lambda point due to its characteristic shape. This peak indicates a superfluid phase transition, where the system undergoes a change in its thermodynamic properties. Below T_c , the specific heat capacity decreases, reflecting the reduced number of available excitations in the superfluid phase.

The superfluidity of liquid helium-4 cannot be fully explained by the ideal Bose gas model, as interactions between helium atoms in the liquid phase (at that temperature helium is liquid) play a crucial role in the emergence of superfluid behavior; we would need a really rarefied gas of helium-4 atoms to observe BEC directly. Furthermore we see a divergence in the specific heat, which was not predicted by the ideal Bose gas model.

In experiments with ultracold atomic gases, researchers have successfully created Bose-Einstein condensates of helium-4 atoms by cooling them to temperatures on the order of nanokelvins, allowing for the direct observation of BEC phenomena.

Figure 10.9: Specific heat capacity of liquid helium-4 as a function of temperature, showing a sharp peak at the lambda point ($T_c \sim 2.17\text{ K}$), indicating the superfluid phase transition. Below T_c , the specific heat decreases, reflecting the reduced number of excitations in the superfluid phase.



Example (BEC in ultracold atomic gases). In 1995, Eric Cornell and Carl Wieman at JILA (University of Colorado) and Wolfgang Ketterle at MIT successfully created Bose-Einstein condensates (BEC) in ultracold atomic gases, marking a significant milestone in experimental physics. They used laser cooling and evaporative cooling techniques to cool a dilute gas of Rubidium-87 atoms (Cornell and Wieman) and Sodium-23 atoms (Ketterle) to temperatures on the order of nanokelvins ($T_{BEC} \sim 10\text{ nK} - 1\text{ } \mu\text{K}$), allowing the atoms to occupy the same

quantum state and form a macroscopic condensate (see Fig. 10.12).

The mathematical description of BEC was achieved completely around 1924-25 by Satyendra Nath Bose and Albert Einstein, but it took about 70 years to realize it experimentally due to the extreme conditions required to achieve such low temperatures and high densities. The experimental realization of BEC in ultracold atomic gases provided a platform for studying quantum phenomena on a macroscopic scale, leading to insights into superfluidity, quantum coherence, and many-body physics.

10.5.1 | How to Reach NanoKelvins?

Achieving temperatures in the nanokelvin range requires sophisticated cooling techniques that go beyond traditional methods. Here are some of the key techniques used to reach such ultra-low temperatures:

- **Laser cooling:** This technique uses the momentum of photons to slow down and cool atoms. By tuning the frequency of laser light slightly below an atomic transition, atoms moving towards the laser absorb photons and experience a recoil that slows them down. This process, known as Doppler cooling, can bring atomic gases down to microkelvin temperatures.

We would need to use multiple laser beams arranged in a configuration known as *optical molasses* to achieve effective cooling in three dimensions. The atoms are trapped in a region where the laser beams intersect, allowing for efficient cooling.

This technique is often the first step in cooling atoms before applying more advanced methods to reach even lower temperatures, since it would likely reach only **microkelvin temperatures**.

- **Magnetic evaporation:** This technique involves trapping atoms in a magnetic field and selectively removing the most energetic atoms from the trap. By gradually lowering the depth of the magnetic trap, the hottest atoms escape, leaving behind a colder sample (like evaporation at room temperature with a cup of coffee). This process can be repeated multiple times, lowering further and further the magnetic trap depth to achieve temperatures in the **nanokelvin range**.

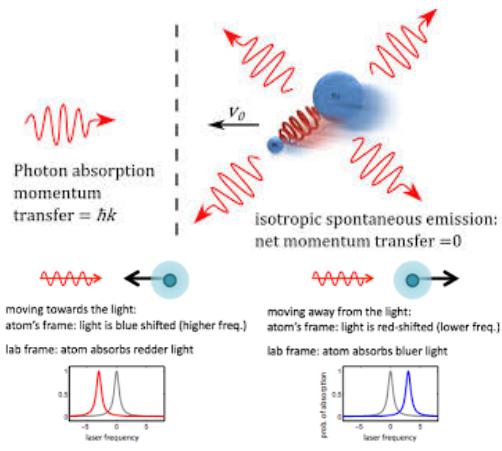
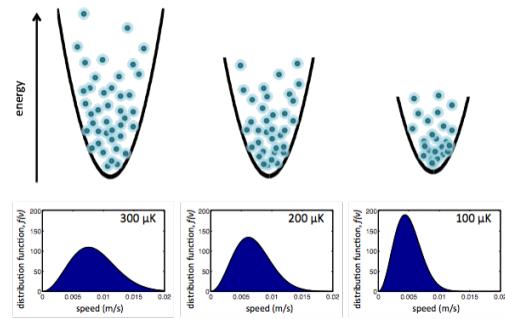


Figure 10.10: Schematic of laser cooling using optical molasses (top) and Doppler cooling mechanism (bottom). The atoms absorb photons from the laser beams, experiencing a recoil that reduces their velocity and temperature, leading to cooling down to microkelvin temperatures; the Doppler effect ensures that atoms moving towards the laser absorb more photons, resulting in a net cooling effect.

Figure 10.11: Schematic of magnetic evaporation cooling. Atoms are trapped in a magnetic field, and the depth of the trap is gradually lowered, allowing the most energetic atoms to escape. This selective removal of hot atoms leads to a colder sample, enabling temperatures in the nanokelvin range to be achieved.



10.5.2 | How to Observe BEC?

When we are at high temperatures, the atomic cloud behaves like a classical gas, with atoms moving randomly and independently, as localized particles with a Maxwell-Boltzmann distribution of velocities. As we cool the gas down to temperatures near T_{BEC} , we can observe the onset of Bose-Einstein condensation through several experimental signatures:

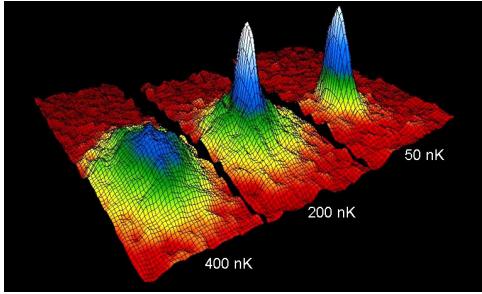


Figure 10.12: Atomic clouds of Rubidium-87 at different temperatures, showing the transition from a classical gas to a Bose-Einstein condensate as the temperature decreases below T_{BEC} . The images illustrate the density distribution of the atomic cloud, with a sharp peak forming at the center of the trap as the condensate forms.

- **Density distribution:** Above T_{BEC} , the density distribution of the atomic cloud is relatively uniform, reflecting the classical behavior of the gas. As we approach T_{BEC} , a sharp peak in the density distribution appears at the center of the trap, indicating the formation of a condensate. Below T_{BEC} , this peak becomes more pronounced, with a significant fraction of atoms occupying the lowest energy state.
- **Momentum distribution:** The momentum distribution of atoms can be measured using time-of-flight imaging. Above T_{BEC} , the momentum distribution is broad and Gaussian-like, characteristic of a thermal gas. As we cool below T_{BEC} , a narrow peak emerges in the momentum distribution, corresponding to the condensate fraction with low momentum: this is a direct signature of macroscopic occupation of the ground state.
- **BEC release:** When the trap confining the atomic cloud is suddenly turned off, the condensate expands differently compared to the thermal component of the gas. The condensate exhibits a characteristic expansion pattern, often referred to as "inverted parabolic" expansion, which can be observed in time-of-flight images. This behavior is distinct from the isotropic expansion of the thermal cloud.

But there are other experimental phenomena that can be observed in a Bose-Einstein condensate, such as:

- **Interference patterns:** When two condensates are allowed to overlap, they can produce interference patterns due to their coherent wavefunctions. This phenomenon is a direct

demonstration of the macroscopic quantum coherence of the condensate. We said that the BEC PT is characterized by the SSB of the gauge symmetry associated to the phase of the condensate wavefunction: thus, if we have two condensates with different phases, when they overlap the difference in phase will lead to constructive and destructive interference patterns, which can be observed experimentally.

- **Vortices:** In rotating Bose-Einstein condensates (realised with rotating magnetic fields), quantized vortices can form, which are topological defects characterized by a phase singularity in the condensate wavefunction. The observation of vortices provides evidence of superfluidity in the condensate. Practically we do not have a single classical fluid rotating, but a lattice of quantized vortices localized almost as a grid in the condensate.

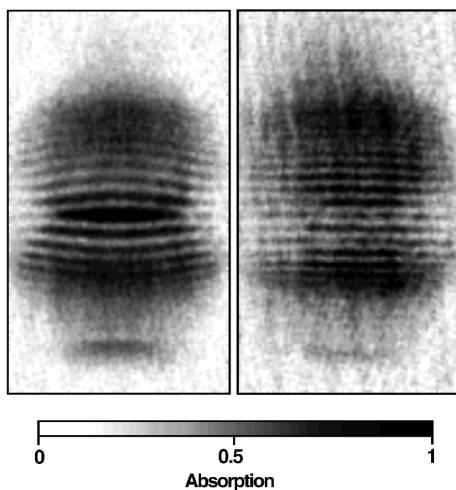


Figure 10.13: Interference patterns observed when two Bose-Einstein condensates overlap, arising from the constructive and destructive interference of the condensate wavefunctions with different phases.

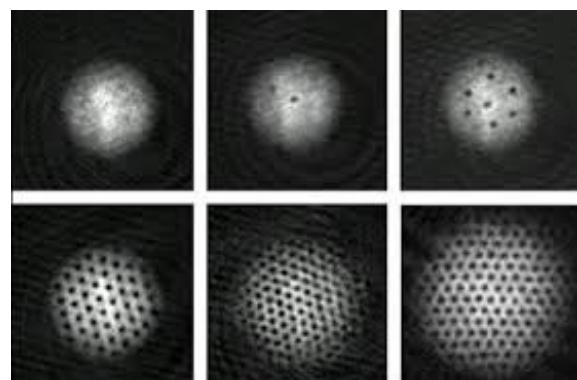


Figure 10.14: Quantized vortices observed in a rotating Bose-Einstein condensate indicating superfluidity. The vortices are characterized by phase singularities in the condensate wavefunction and form a lattice structure within the condensate. The different stages show the formation and arrangement of vortices as the rotation rate of the condensate is increased.

Cold Atoms in Artificial Lattice

By using **optical lattices** created by interfering laser beams, we can simulate periodic potentials for ultracold atoms. This allows us to study phenomena such as the superfluid to Mott insulator transition, where the interplay between kinetic energy and interaction energy leads to different quantum phases. In the superfluid phase, atoms can move freely across the lattice, while in the Mott insulator phase, atoms are localized at individual lattice sites due to strong interactions.

Since the atoms can change minimum energy states by tunneling between lattice sites, we can control the tunneling rate by adjusting the depth of the optical lattice potential. This tunability allows us to explore different regimes of quantum behavior, from weakly interacting superfluids to strongly correlated Mott insulators.

Atoms in optical lattices can also interact with each other through controlled interactions, since they can be made to collide at specific lattice sites. By tuning the interaction strength using Feshbach resonances, we can study the effects of interactions on the quantum phases of the system, leading to rich phase diagrams and novel quantum states.

We can also explore topological phases by engineering specific lattice geometries and interactions. For example, by creating honeycomb or kagome lattices, we can realize systems that exhibit topologically protected edge states and non-trivial band structures. These systems can host exotic quasiparticles such as Majorana fermions and anyons, which have potential applications in quantum computing.

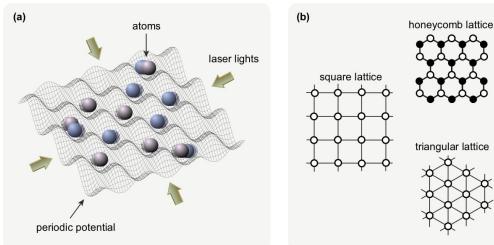


Figure 10.15: Schematic of an optical lattice created by interfering laser beams, forming a periodic potential for trapping ultracold atoms, allowing for the study of quantum phases such as superfluidity and Mott insulator states.

We could also create lattices in higher dimensions (like 4D lattices) using synthetic dimensions, where internal degrees of freedom of atoms (such as spin states) are treated as additional spatial dimensions. This allows us to explore higher-dimensional physics and topological phenomena that are not accessible in traditional three-dimensional systems: we could simulate quantum gravity models and higher-dimensional topological insulators.

10.5.3 | Photons Gas

We are going to study now the case of a gas of photons, which are bosons with zero rest mass and two polarization states (thus $g = 2$); they are ultrarelativistic particles, since they always move at the speed of light c in vacuum. The energy-momentum relation for photons is given by

$$\epsilon \mathbf{p} = c|\mathbf{p}|,$$

where \mathbf{p} is the momentum of the photon.

Photons are unique among bosons because they do not have a conserved particle number: they can be created and destroyed freely, for example through absorption and emission processes. This means that the chemical potential μ for photons is always zero, leading to a fugacity:

$$\mu(T) = 0 \implies z = e^{\beta\mu} = 1, \quad \forall T.$$

We can now compute the density of states for photons, starting from the general expression for the density of states in three dimensions: from the computation for the volume in phase space we have the cumulative number of states with energy less than or equal to ϵ :

$$\begin{aligned} \Sigma(\epsilon) &= \int_{0 \leq c|\mathbf{p}| \leq \epsilon} d\Omega = g \int_{0 \leq c|\mathbf{p}| \leq \epsilon} \frac{d^3x d^3p}{h^3} \\ &= 2 \int_0^{\epsilon/c} dp \frac{V}{h^3} 4\pi p^2 = \frac{V}{(hc)^3} \frac{8\pi}{3} \epsilon^3 = \int_0^\epsilon d\epsilon' g(\epsilon'), \end{aligned}$$

since it is a spherical volume in momentum space with radius $p = \epsilon/c$, and $g = 2$ for the two polarization states of the photon. Differentiating with respect to energy, we find the density of states:

$$\omega(\epsilon) = \frac{\partial}{\partial \epsilon} \Sigma(\epsilon) = \frac{8\pi V}{(hc)^3} \epsilon^2 = g(\epsilon).$$

Using this density of states, we can compute the thermodynamic quantities for the photon gas. The average number of photons N in the system is given by¹⁶

$$N = \int_0^\infty d\epsilon g(\epsilon) \langle n(\epsilon) \rangle_{gc}^B = \int_0^\infty d\epsilon g(\epsilon) \frac{1}{e^{\beta\epsilon} - 1}.$$

Substituting the expression for the density of states, we have

$$N = \frac{8\pi V}{(hc)^3} \int_0^\infty d\epsilon \frac{\epsilon^2}{e^{\beta\epsilon} - 1}.$$

For the energy E of the photon gas, we have

$$E = \int_0^\infty d\epsilon g(\epsilon) \epsilon \langle n(\epsilon) \rangle_{gc}^B = \frac{8\pi V}{(hc)^3} \int_0^\infty d\epsilon \frac{\epsilon^3}{e^{\beta\epsilon} - 1}.$$

For the grand potential Ω , we have

$$\Omega = \frac{1}{\beta} \int_0^\infty d\epsilon g(\epsilon) \log(1 - e^{-\beta\epsilon}) = k_B T \frac{8\pi V}{(hc)^3} \int_0^\infty d\epsilon \epsilon^2 \log(1 - e^{-\beta\epsilon}),$$

but we can integrate by parts

$$\Omega = -\frac{V}{3} \frac{8\pi}{(hc)^3} \int_0^\infty d\epsilon \frac{\epsilon^3}{e^{\beta\epsilon} - 1} = -\frac{1}{3} E.$$

Now the pressure p of the photon gas is given by

$$p = -\frac{\Omega}{V} = \frac{1}{3} \frac{8\pi}{(hc)^3} \int_0^\infty d\epsilon \frac{\epsilon^3}{e^{\beta\epsilon} - 1},$$

so that finally we can compute the equation of state for the photon gas: notably photons exhibit a relativistic equation of state, since they are massless, thus the pressure p and energy density $u = E/V$ of the photon gas are related by

$$p = \frac{1}{3} u, \quad pV = \frac{E}{3}.$$

This relation arises from the fact that photons are ultrarelativistic particles, and it reflects the isotropic nature of the radiation pressure exerted by the photon gas.

In order to explixit the **temperature dependence** of the pressure and energy density, we can evaluate the integrals with a change of variable $x = \beta\epsilon$:

$$\begin{aligned} N &= \frac{8\pi V}{(hc)^3} (k_B T)^3 \int_0^\infty dx \frac{x^2}{e^x - 1}, \\ E &= \frac{8\pi V}{(hc)^3} (k_B T)^4 \int_0^\infty dx \frac{x^3}{e^x - 1}, \\ p &= \frac{8\pi}{(hc)^3} (k_B T)^4 \int_0^\infty dx x^2 \log(1 - e^{-x}). \end{aligned}$$

The integrals can be evaluated using the properties of the Riemann zeta function and the Gamma function, but we are not interested in the exact numerical factors here, we will just note the temperature scaling. The key point is that both the energy E and pressure p scale with the fourth power of temperature, while the number of photons N scales with the third power. We could also derive the temperature dependence of the heat capacity C_V of the photon gas, which is given by

$$C_V = \left(\frac{\partial E}{\partial T} \right)_V \propto T^3,$$

indicating that the heat capacity increases rapidly with temperature, reflecting the increasing number of accessible photon states as the temperature rises.

¹⁶Since the chemical potential is zero, we have $z = 1$ and thus the Bose-Einstein distribution simplifies accordingly.

Spectral Energy Density

The **spectral energy density** $E(\nu, T)$ of the photon gas ($\epsilon = h\nu$) describes how the energy is distributed among different photon energies (or frequencies) at a given temperature T . It is defined as

$$\frac{E(\nu, T)}{V} = \frac{8\pi}{(hc)^3} \int_0^{h\nu} d\epsilon \frac{\epsilon^3}{e^{\beta\epsilon} - 1}.$$

Differentiating with respect to frequency ν , we obtain the spectral energy density per unit frequency:

$$u(\nu, T) = \frac{\partial}{\partial \nu} \left(\frac{E(\nu, T)}{V} \right) = \frac{8\pi h}{c^3} \frac{\nu^3}{e^{\beta h\nu} - 1}.$$

This expression is known as **Planck's law** of blackbody radiation, which describes the distribution of electromagnetic radiation emitted by a blackbody at temperature T . It shows that the spectral energy density depends on both the frequency ν and the temperature T , with a characteristic peak that shifts to higher frequencies as the temperature increases (Wien's displacement law).

We are interested in studying two particular limits of Planck's law:

- **Rayleigh-Jeans limit** ($h\nu \ll k_B T$): In this low-frequency limit, we can approximate the exponential term in the denominator using a Taylor expansion:

$$e^{\beta h\nu} - 1 \approx \beta h\nu.$$

Substituting this approximation into Planck's law, we obtain

$$u(\nu, T) \approx \frac{8\pi h}{c^3} \frac{\nu^3}{\beta h\nu} = \frac{8\pi k_B T}{c^3} \nu^2,$$

which is known as the **Rayleigh-Jeans law**. This law predicts that the spectral energy density increases with the square of the frequency, leading to the so-called "ultraviolet catastrophe" at high frequencies.

- **Wien's limit** ($h\nu \gg k_B T$): In this high-frequency limit, the exponential term in the denominator dominates, and we can approximate it as

$$e^{\beta h\nu} - 1 \approx e^{\beta h\nu}.$$

Substituting this approximation into Planck's law, we obtain

$$u(\nu, T) \approx \frac{8\pi h}{c^3} \nu^3 e^{-\beta h\nu},$$

which is known as **Wien's law**. This law predicts that the spectral energy density decreases exponentially with increasing frequency, reflecting the suppression of high-energy photons at low temperatures.

These two limits highlight the different behaviors of the spectral energy density in the low and high-frequency regimes, providing insights into the nature of blackbody radiation and the underlying quantum mechanics of photons.

11 | Exercises

11.1 | Classical Statistical Mechanics

Phase Space

We work on a phase space \mathcal{M}_N of N particles in d dimensions with coordinates (q_i, p_i) , $i = 1, \dots, N$. We will denote the volume element with:

$$d\Omega = \frac{\prod_i dq_i dp_i}{h^{dN} \xi_N} = \frac{d\mathcal{H} dS_{\mathcal{H}}}{h^{dN} \xi_N}$$

where $\xi_N = N!$ if the particles are indistinguishable while $\xi_N = 1$ if they are distinguishable.

We recall that the density of states $\omega(E)$ is given by

$$\omega(E) \equiv \int_{\mathcal{M}_N} \frac{\prod_i dq_i dp_i}{h^{dN} \xi_N} \delta(\mathcal{H}(q_i, p_i) - E) = \int_{S_E} \frac{dS_{\mathcal{H}}}{h^{dN} \xi_N} = \frac{\partial \Sigma(E)}{\partial E}$$

where $\Sigma(E)$ is the volume of the region of states having energies between 0 and E :

$$\Sigma(E) = \frac{1}{h^{dN} \xi_N} \int_0^E d\mathcal{H} \int_{S_{\mathcal{H}}} dS_{\mathcal{H}}$$

Exercise 1 - Density of States of a Single Particle

1. Describe the phase space \mathcal{M}_1 , specify the Hamiltonian \mathcal{H} for the following systems:
 - (a) A single non-relativistic particle in dimensions $d = 1, 2, 3$, confined in an interval L , an area A or a volume V respectively;
 - (b) A single harmonic oscillator in 1 dimension;
 - (c) A single ultra-relativistic particle in 3 dimensions.
2. Show that the density of states $\omega(E)$ is given by:

Case (a)

$$\omega(E) = \frac{L\sqrt{2m}}{\hbar} E^{-1/2}, \quad d = 1$$

$$\omega(E) = \frac{A\pi 2m}{\hbar^2}, \quad d = 2$$

$$\omega(E) = \frac{V2\pi(2m)^{3/2}}{\hbar^3} E^{1/2}, \quad d = 3$$

Case (b)

$$\omega(E) = \frac{2\pi}{h\omega}$$

Case (c)

$$\omega(E) = \frac{4\pi V}{3c^3 h^3} E^3$$

Solution:

1. Phase space and Hamiltonians

In d dimensions, each particle is described by d coordinates q_i and d conjugate momenta p_i , so that $\mathcal{M}_N = \mathbb{R}^{dN} \times \mathbb{R}^{dN} = \mathbb{R}^{2dN}$. To avoid divergences, we confine the system in a finite volume V (denoted with L or A in $d = 1, 2$), so that $\mathcal{M}_N = V^N \times \mathbb{R}^{dN}$.

Hamiltonians:

- **Case (a)** Free non-relativistic particle:

$$\mathcal{H} = \frac{p^2}{2m}$$

where $p^2 \equiv (p^x)^2$ ($d = 1$), $p^2 \equiv (p^x)^2 + (p^y)^2$ ($d = 2$), etc.

- **Case (b)** Harmonic oscillator in 1 dimension:

$$\mathcal{H} = \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2$$

with $q^2 \equiv (q^x)^2$, $p^2 \equiv (p^x)^2$.

- **Case (c)** Free ultra-relativistic particle in 3 dimensions:

$$\mathcal{H} = c|p|$$

where $|p| \equiv \sqrt{(p^x)^2 + (p^y)^2 + (p^z)^2}$.

2. Density of states $\omega(E)$

Case (a) $d = 1$:

$$\Sigma(E) = \frac{1}{h} \int_L dq \int_{0 \leq |p|^2 / 2m \leq E} dp = \frac{L}{h} \int_{-\sqrt{2mE}}^{\sqrt{2mE}} dp^x = \frac{L 2\sqrt{2m}}{h} E^{1/2}$$

$$\omega(E) = \frac{L \sqrt{2m}}{h} E^{-1/2}$$

Case (a) $d = 2$: Using polar coordinates in p -space:

$$\Sigma(E) = \frac{1}{h^2} \int_A d^2q \int_{0 \leq |p|^2 / 2m \leq E} d^2p = \frac{A 2\pi}{h^2} \int_0^{\sqrt{2mE}} |p| d|p| = \frac{A \pi (2m)}{h^2} E$$

$$\omega(E) = \frac{A \pi (2m)}{h^2}$$

Case (a) $d = 3$: Using spherical coordinates in p -space:

$$\Sigma(E) = \frac{1}{h^3} \int_V d^3q \int_{0 \leq |p|^2 / 2m \leq E} d^3p = \frac{V4\pi}{h^3} \int_0^{\sqrt{2mE}} |p|^2 d|p| = \frac{V4\pi(2m)^{3/2}}{3h^3} E^{3/2}$$

$$\omega(E) = \frac{V2\pi(2m)^{3/2}}{h^3} E^{1/2}$$

Case (b): Using change of coordinates $x = \sqrt{\frac{m}{2}}q$, $y = \sqrt{\frac{1}{2m}}\omega p$:

$$\Sigma(E) = \frac{1}{h} \int_{0 \leq \frac{p^2}{2m} + \frac{m\omega^2}{2} q^2 \leq E} dq dp = \frac{2}{h\omega} \int_{0 \leq x^2 + y^2 \leq E} dx dy = \frac{2\pi}{h\omega} E$$

$$\omega(E) = \frac{2\pi}{h\omega}$$

Case (c): Using spherical coordinates in p -space:

$$\Sigma(E) = \frac{1}{h^3} \int_V d^3q \int_{0 \leq c|p| \leq E} d^3p = \frac{V4\pi}{h^3} \int_0^{E/c} |p|^2 d|p| = \frac{V4\pi}{3c^3 h^3} E^3$$

$$\omega(E) = \frac{V4\pi}{c^3 h^3} E^2$$

Some Integrals

In the following we will often use the Gamma function (or Euler function) which can be defined through an integral representation:

$$\Gamma(x) \equiv \int_0^\infty dt e^{-t} t^{x-1}$$

It satisfies the following useful relations:

$$\Gamma(n+1) = n!, \quad \Gamma(x+1) = x\Gamma(x), \quad \Gamma(1/2) = \sqrt{\pi}$$

In particular, this function is useful to calculate the volume of a sphere of radius r in \mathbb{R}^m , i.e. of the region $0 \leq \sum_{j=1}^m x_j^2 \leq r^2$, which is given by:

$$V_m = \frac{\pi^{m/2}}{\Gamma\left(\frac{m}{2} + 1\right)} r^m$$

Exercise 2 - Density of States of N Particles

1. Describe the phase space \mathcal{M}_N , specify the Hamiltonian \mathcal{H} for:

- (a) A non-relativistic ideal gas in d dimensions: i.e. a gas of non-interacting particles whose kinetic energy is non-relativistic;
- (b) A gas of harmonic oscillators in 1 dimension: i.e. a gas of non-interacting particles whose kinetic energy is non-relativistic and is subject to a harmonic potential with frequency ω .

2. Show that the density of states $\omega(E)$ is given by:

Case (a)

$$\omega(E) = \frac{V^N}{\xi_N \Gamma(dN/2)} \left(\frac{2\pi m}{h^2} \right)^{dN/2} E^{dN/2-1}$$

Case (b)

$$\omega(E) = \frac{1}{\xi_N \Gamma(dN)} \left(\frac{1}{\hbar\omega} \right)^{dN} E^{dN-1}$$

Solution:

1. Phase space and Hamiltonians

See solution of Ex.1 point 1.

2. Density of states $\omega(E)$

Case (a): Using spherical coordinates in momentum space and formula for volume of m -sphere:

$$\Sigma(E) = \frac{1}{h^{dN} \xi_N} \left(\prod_i \int_V d^d q_i \right) \int_{0 \leq \sum_i |p_i|^2 \leq 2mE} \left(\prod_i d^d p_i \right)$$

$$= \frac{V^N}{\xi_N h^{dN}} \frac{1}{\Gamma(dN/2 + 1)} (2\pi m E)^{dN/2} = \frac{V^N}{\xi_N (dN/2) \Gamma(dN/2)} \left(\frac{2\pi m E}{h^2} \right)^{dN/2}$$

$$\omega(E) = \frac{V^N}{\xi_N \Gamma(dN/2)} \left(\frac{2\pi m}{h^2} \right)^{dN/2} E^{dN/2-1}$$

Case (b): Using change of coordinates $p_j = \sqrt{2mE}x_j$ and $q_j = \sqrt{2E/m\omega^2}x_{dN+j}$ for $j = 1, \dots, dN$:

$$\Sigma(E) = \int_{0 \leq \sum_i \left(\frac{p_i^2}{2m} + \frac{m\omega^2}{2} q_i^2 \right) \leq E} \frac{\prod_i d^d q_i d^d p_i}{h^{dN} \xi_N}$$

$$= \frac{1}{h^{dN} \xi_N} (2mE)^{dN/2} \left(\frac{2E}{m\omega^2} \right)^{dN/2} \int_{0 \leq \sum_{j=1}^{2dN} x_j^2 \leq 1} \left(\prod_{j=1}^{2dN} dx_j \right)$$

$$= \frac{1}{\xi_N} \left(\frac{2E}{h\omega} \right)^{dN} \frac{\pi^{dN}}{\Gamma(dN+1)} = \frac{1}{\xi_N (dN) \Gamma(dN)} \left(\frac{2\pi E}{h\omega} \right)^{dN}$$

$$\omega(E) = \frac{1}{\xi_N \Gamma(dN)} \left(\frac{1}{\hbar\omega} \right)^{dN} E^{dN-1}$$

Microcanonical Ensemble

A totally isolated system, at temperature T and volume V , is described by the *microcanonical ensemble*: both the energy E and the number of particles N are conserved.

We assume that all points on the hypersurface $S_{\mathcal{H}=E} \subset \mathcal{M}_N$, corresponding to the energy E , are equiprobable. Thus the (microcanonical) probability density reads:

$$\rho_{mc}(q_i, p_i) = \frac{\delta(\mathcal{H} - E)}{\omega(E)} \quad (1)$$

where $\omega(E) = \int_{S_E} dS_{\mathcal{H}}$ is the density of states, i.e. the area of the hypersurface $S_{\mathcal{H}=E}$.

Thermodynamics is recovered from the universal Boltzmann function for entropy:

$$S_{mc} \equiv k_B \log \omega(E)$$

which coincides with the thermodynamic entropy S_{th} , in the thermodynamic limit.

In the same limit, the entropy can be equivalently calculated as:

$$S_{mc} \equiv k_B \log \omega(E) = k_B \log \Gamma(E) = k_B \log \Sigma(E)$$

with $\omega(E) = \partial \Sigma / \partial E$ and $\Gamma(E) = \omega(E) \Delta E$.

Exercise 3 - Ideal gas in 3d, microcanonical ensemble

In this exercise we will derive the thermodynamics of an ideal gas from its microcanonical description.

1. Prove that:

$$\log \Gamma(E) = \log \omega(E) + \log \Delta E = \log \Sigma(E) + \log(3N/(2E)) + \log \Delta E$$

Divide the obtained expressions by the total number of particles N , to obtain the entropy per particle s_{mc} , showing that they all coincide, in the thermodynamic limit.

2. Starting from the result obtained in Ex. 2 of Set 1:

$$\Sigma(E) = \frac{V 4\pi (2m)^{3/2}}{3h^3} E^{3/2}$$

show that the entropy $S = k_B \log \Sigma(E)$, in the thermodynamic limit (i.e. neglecting terms that increase slower than N), is given by:

$$S = k_B \begin{cases} \frac{3}{2}N + N \log \left[V \left(\frac{4\pi m E}{3N h^2} \right)^{3/2} \right] & \text{distinguishable particles} \\ \frac{5}{2}N + N \log \left[\frac{V}{N} \left(\frac{4\pi m E}{3N h^2} \right)^{3/2} \right] & \text{indistinguishable particles} \end{cases}$$

Remark. You need to use Stirling formulas: $\log N! = N \log N - N$ for N large and similarly $\log \Gamma(x) \simeq x \log x - x$ for x large.

Are the obtained expressions of S extensive? Discuss the result.

3. Recalling that $TdS = dE + pdV$, calculate:

- $\frac{\partial S}{\partial E}|_V$ and derive the expression for the internal energy:

$$E = \frac{3}{2}Nk_B T$$

- $\frac{\partial S}{\partial V}|_E$ and derive the expression for the equation of state:

$$pV = Nk_B T$$

Solution:

1) The first equality simply follows from $\Gamma(E) = \omega(E)\Delta E$. The second equality can be easily derived from the expressions of $\Sigma(E)$ and $\omega(E) = \frac{d\Sigma(E)}{dE}$. We see that the expressions of $\log \Gamma(E)$, $\log \omega(E)$, $\log \Sigma(E)$ differ by terms such as $\log \Delta E$ and $\log(3N/2)$, which scale sublinearly as $N \rightarrow \infty$. Thus they go to zero when divided by N , in the thermodynamic limit.

2) Starting from $S_{mc} = k_B \log \Sigma(E)$, we get the following.

For distinguishable particles we have to take $\mathcal{N} = 1$, hence

$$S_{\text{dist}} = k_B \left[N \log \left(V \left(\frac{2\pi m E}{h^2} \right)^{3/2} \right) - \log \left(\frac{3N}{2} \right) - \log \Gamma \left(\frac{3N}{2} \right) \right] \quad (11.1.1)$$

For indistinguishable particles we have to take $\mathcal{N} = N!$, hence

$$S_{\text{ind}} = k_B \left[N \log \left(\frac{V}{N} \left(\frac{2\pi m E}{h^2} \right)^{3/2} \right) - \log \left(\frac{3N}{2} \right) - \log \Gamma \left(\frac{3N}{2} \right) - \log N! \right] \quad (11.1.2)$$

The final expressions are recovered by taking into account that, for large N , we can neglect the second term because it does not contribute to the thermodynamic limit, while:

$$\log N! \approx N \log N - N, \quad \log \Gamma(x) \approx x \log x - x \quad \text{for large } x.$$

We can notice that the expression for distinguishable particles is not extensive: if we double the volume V the expression for S changes also by an additional addend $2N \log 2$, while the one for indistinguishable particles is extensive. This is related to the so-called Gibbs paradox, that can be resolved by noticing that in the case of distinguishable particles, we actually have two situations that are not the same from a physical point of view: on one hand, we have two systems of N particles, each in a volume V with particles in one volume that can be distinguished from the ones in the other volume; on the other hand we have $2N$ particles that can all occupy a volume $2V$. If particles are indistinguishable, we cannot say which particle is in which volume and therefore these two situations coincide.

3) The calculation is straightforward, since $\frac{\partial S}{\partial E}$ and $\frac{\partial S}{\partial V}$ give the expressions for the internal energy and the equation of state:

$$E = \frac{3}{2}Nk_B T, \quad pV = Nk_B T.$$

Canonical Ensemble

The canonical ensemble describes a system, with a given volume V and a number of particles N , in thermodynamical equilibrium with an environment at a given temperature T , with which it can exchange energy.

The canonical probability density reads:

$$\rho_C = \frac{1}{Z_N} e^{-\beta \mathcal{H}(q_i, p_i)} \quad (2)$$

where $Z_N = Z_N[V, T]$ is the canonical partition function:

$$Z_N = \int_{\mathcal{M}_N} d\Omega e^{-\beta \mathcal{H}(q_i, p_i)} \quad (3)$$

Thermodynamic potentials can be derived through the formulas:

$$E = \langle \mathcal{H}(q_i, p_i) \rangle_c = -\frac{\partial \ln Z_N}{\partial \beta} \quad (4)$$

$$F = -\frac{1}{\beta} \ln Z_N \quad (5)$$

$$S = \frac{E - F}{T} \quad (6)$$

One can then obtain other thermodynamical quantities, such as:

$$p = -\left. \frac{\partial F}{\partial V} \right|_{T, N} \quad (7)$$

$$\mu = \left. \frac{\partial F}{\partial N} \right|_{T, V} \quad (8)$$

and

$$C_V = T \left. \frac{\partial S}{\partial T} \right|_{V, N} = \left. \frac{\partial E}{\partial T} \right|_{V, N} \quad (9)$$

$$C_p = T \left. \frac{\partial S}{\partial T} \right|_{p, N} = \left. \frac{\partial E}{\partial T} \right|_{p, N} + p \left. \frac{\partial V}{\partial T} \right|_{p, N} \quad (10)$$

Exercise 4 - Ideal gas in 3d, canonical ensemble

In this exercise we will derive the thermodynamics of an ideal gas from its canonical description, supposing particles indistinguishable.

- Recalling that, for a gas of free non-relativistic particles, we have:

$$\mathcal{M}_N = V^N \times \mathbb{R}^{3N}, \quad \mathcal{H} = \sum_i \frac{p_i^2}{2m}$$

show that the canonical partition function Z is given by:

$$Z = \frac{V^N}{N! \lambda_T^{3N}}$$

having defined the thermal wavelength

$$\lambda_T \equiv \sqrt{\frac{h^2}{2\pi m k_B T}}$$

Derive an expression for $\ln Z$ in the large N limit.

Remark. Recall Stirling formula and the formula for Gaussian integral:

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}}$$

2. Derive the internal energy E and the free energy F .
 3. Derive the entropy S , for both the case of distinguishable and indistinguishable particles.
- Calculate the value T_* of the temperature at which S becomes negative. How can you interpret this result?
4. Calculate the pressure p , and derive the equation of state.

How does an isothermal curve look like in the $p - V$ plane?

5. Calculate the chemical potential μ and draw a graph of it as function of temperature T , at a fixed density of particles n .
6. Calculate the specific heats per particle: $c_v = C_v/N$ and $c_p = C_p/N$.

Solution:

1)

$$Z = \frac{1}{h^{3N} N!} \int \prod_{i=1}^N d^3 q_i \int \prod_{i=1}^N d^3 p_i e^{-\beta \sum_i \frac{p_i^2}{2m}} = \frac{V^N}{h^{3N} N!} \left(\int d^3 p e^{-\beta \frac{p^2}{2m}} \right)^N$$

Since we have a Gaussian integral for each component of the three-dimensional momentum p_i of each of the N particles:

$$\int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}} \Rightarrow \int d^3 p e^{-\beta \frac{p^2}{2m}} = \left(\frac{2\pi m}{\beta} \right)^{3/2}$$

Thus, setting $n = N/V$ and defining the thermal wavelength $\lambda_T = \sqrt{\frac{h^2}{2\pi m k_B T}}$:

$$Z = \frac{V^N}{N! \lambda_T^{3N}} \Rightarrow \ln Z = N \ln \left(\frac{V}{N \lambda_T^3} \right) + N$$

2)

$$F = -k_B T \ln Z = -N k_B T \left[\ln \left(\frac{V}{N \lambda_T^3} \right) + 1 \right]$$

$$E = -\frac{\partial \ln Z}{\partial \beta} = \frac{3}{2} N k_B T$$

The latter result coincides with the one obtained from the equipartition theorem, that gives a factor of $k_B T/2$ for each degree of freedom that appears quadratically in the Hamiltonian: here we have $3N$ momenta.

3)

$$S = \frac{E - F}{T} = N k_B \left[\frac{5}{2} + \ln \left(\frac{V}{N \lambda_T^3} \right) \right]$$

The entropy becomes negative for temperatures $T < T_x$, with T_x fixed by the condition:

$$n \lambda_T^3 = e^{5/2} \Rightarrow T_x = \frac{\hbar^2}{2 \pi m k_B} \left(\frac{n}{e^{5/2}} \right)^{2/3}$$

This is of course absurd, signalling that this model is not suited to describe the low temperature limit.

4)

$$p = - \left(\frac{\partial F}{\partial V} \right)_{T,N} = \frac{N k_B T}{V} \Rightarrow pV = N k_B T$$

Note that isothermal curves in the p - V plane are hyperbolas, as shown in the figure below.

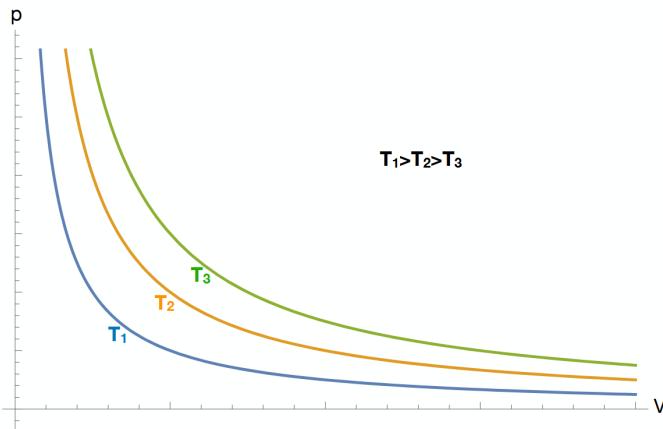


Figure 11.1: Isothermal curves in the p - V plane for different temperatures.

5)

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} = -k_B T \left[\ln \left(\frac{V}{N \lambda_T^3} \right) + 1 \right] = k_B T \ln(n \lambda_T^3) - k_B T$$

As $T \rightarrow 0$, $\mu \rightarrow -\infty$; as $T \rightarrow \infty$, $\mu \rightarrow +\infty$.

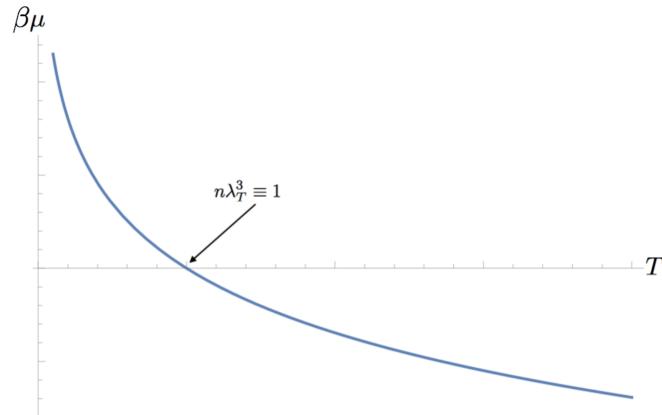


Figure 11.2: The chemical potential μ as function of T , for fixed density n .

6)

$$c_v = \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{3}{2}k_B$$

$$c_p = c_v + \frac{p}{N} \left(\frac{\partial V}{\partial T} \right)_{p,N} = \frac{3}{2}k_B + k_B = \frac{5}{2}k_B$$

Notice that c_v is constant: this is in contradiction with thermodynamic identities that require $c_v \rightarrow 0$ as $T \rightarrow 0$. Again this shows that this model is not suited to describe the low temperature limit.

Exercise 5 - Thermodynamics of a solid

We would like to study, in a first and classical approximation, a model suited to describe a solid: the latter is thought of as composed of individual atoms/molecules arranged in a periodic lattice, each of them oscillating harmonically about its equilibrium position, all with the same frequency ω .

In this model, each particle is described by a coordinate q_i and a conjugate momentum p_i , so that

$$\mathcal{M}_N = \mathbb{R}^{2N},$$

where $q_i = x_i - x_i^{(0)}$ is the displacement of the i -th particle with respect to its oscillation center $x_i^{(0)}$. Thus we have:

$$\mathcal{H} = \sum_i \left(\frac{p_i^2}{2m} + \frac{m\omega^2}{2} q_i^2 \right).$$

Notice that, even if the frequency is the same for all particles, the particles are distinguishable since each of them oscillates about a different and known equilibrium position.

1. Calculate the canonical partition function Z and derive an expression for $\ln Z$ in the large N limit.
2. Derive the internal energy E and the free energy F . Discuss the found expression for E in connection with the equipartition principle.
3. Derive the entropy S and find the value T_* at which it becomes negative. Try to give an interpretation of this result.
4. Calculate the pressure p , showing it is exactly zero. By recalling the microscopic interpretation of pressure, find a physical explanation for this result.
5. Calculate the chemical potential μ and study its limit for $T \rightarrow 0$ and $T \rightarrow \infty$.
6. Calculate the specific heats per particle: $c_v = C_V/N$ and $c_p = C_P/N$.

Solution:

Exercise 6 - Ideal ultra-relativistic gas in 3d

As for the non-relativistic case, each particle is described by 3 coordinates q_i , which are restricted to a finite volume V , and 3 conjugate momenta p_i , so that

$$\mathcal{M}_N = V^N \times \mathbb{R}^{3N}.$$

Now, the energy carried by each particle has to satisfy the relativistic expression:

$$E = \sqrt{c^2 p^2 + m^2 c^4}.$$

The ultra-relativistic limit is obtained by sending $m \rightarrow 0$, so that, when the particles are non-interacting, we have:

$$\mathcal{H} = \sum_i c|p_i|.$$

In the following we will suppose the particles to be indistinguishable.

1. Calculate the canonical partition function Z and derive an expression for $\ln Z$ in the large N limit.
2. Derive the internal energy E and the free energy F . Show that the internal energy E can also be derived via the generalized equipartition theorem. In order to do so, first apply it to each component of the momenta p_x, p_y, p_z , and then use the result to calculate $\langle \mathcal{H} \rangle_c$.
3. Derive the entropy S . Can it become negative?
4. Calculate the pressure p and derive the equation of state. Compare it with that of a non-relativistic ideal gas.
5. Calculate the chemical potential μ and study its limit for $T \rightarrow 0$ and $T \rightarrow \infty$.
6. Calculate the specific heats per particle: $c_v = C_V/N$ and $c_p = C_P/N$.

Solution:

Appendices

A | Notation and Conventions

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

B | Computations and further developments

Sign of complex radical

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.