

Statistical Mechanics

Eric D'Hoker

*Department of Physics and Astronomy,
University of California, Los Angeles, CA 90095, USA*

15 September 2012

Contents

1	Introduction	6
1.1	Brief history	7
1.2	Units and physical constants	10
2	Statistical Physics and Thermodynamics	11
2.1	Micro-states and macro-states	11
2.1.1	Counting micro-states; entropy	11
2.2	Statistical and thermal equilibrium	12
2.3	The thermodynamic limit, extensive and intensive variables	15
2.4	Work and Heat	15
2.4.1	The first law of thermodynamics	16
2.4.2	The second law of thermodynamics	16
2.5	Relations between thermodynamic variables	17
2.6	Non-interacting microscopic constituents: the ideal gas	18
2.7	The simplest ideal gas	19
2.7.1	Counting micro-states in the simplest ideal gas	20
2.8	Mixing entropy and the Gibbs paradox	21
2.9	Indistinguishability of identical particles	22
2.10	Thermodynamic transformations	23
3	Classical Statistical Ensembles	25
3.1	Classical mechanics	25
3.2	Time averaging	26
3.3	Statistical ensembles	27
3.4	The density function	28
3.5	The Liouville theorem	29
3.6	Equilibrium ensembles and distribution functions	31
3.7	The uniform ensemble	31
3.8	The micro-canonical ensemble	32
3.9	The canonical ensemble	32
3.10	Deriving the canonical from the micro-canonical ensemble	34
3.11	The Gibbs and the grand-canonical ensembles	36
4	Applications of Classical Statistical Ensembles	38
4.1	Application I: The Maxwell distribution	38
4.2	Application II: Magnetism in classical statistical mechanics	38
4.3	Application III: Diatomic ideal gasses (classical)	39
4.4	Quantum vibrational modes	41

4.5	Quantum rotational modes	43
5	Quantum Statistical Ensembles	44
5.1	Quantum Mechanics	44
5.2	Mixed quantum states	44
5.3	Subsystems of a quantum system in a pure state	45
5.4	The density matrix	46
5.5	Ensemble averages and time-evolution	48
5.6	The density matrix for a subsystem of a pure state	48
5.7	Statistical entropy of a density matrix	50
5.8	The uniform and micro-canonical ensembles	52
5.9	Construction of the density matrix in the canonical ensemble	52
5.10	Generalized equilibrium ensembles	54
6	Applications of the canonical ensemble	55
6.1	The statistics of paramagnetism	55
6.2	Non-relativistic Boltzmann ideal gas	56
6.3	Van der Waals equation of state	57
6.4	The Mayer cluster expansion	59
7	Systems of indistinguishable quantum particles	61
7.1	FD and BE quantum permutation symmetries	61
7.2	BE and FD statistics for N identical free particles	62
7.3	Boltzmann statistics rederived	64
7.4	Fermi-Dirac statistics	64
7.5	Bose-Einstein statistics	66
7.6	Comparing the behavior of the occupation numbers	67
8	Ideal Fermi-Dirac Gases	68
8.1	The simplest ideal Fermi-Dirac gas	68
8.2	Entropy, specific heat, and equation of state	69
8.3	Corrections to the Boltzmann gas	70
8.4	Zero temperature behavior	71
8.5	Low temperature behavior: The Sommerfeld expansion	72
8.6	Pauli paramagnetism of ideal gasses	74
8.7	Landau diamagnetism	75
8.8	White dwarfs	77

9 Bose-Einstein statistics	81
9.1 Black body radiation	81
9.2 Cosmic micro-wave background radiation	83
9.3 Thermodynamic variables for Bose-Einstein ideal gasses	84
9.4 Bose-Einstein condensation	86
9.5 Behavior of the specific heat	86
10 Phase coexistence: thermodynamics	88
10.1 Conditions for phase equilibrium	88
10.2 Latent heat	90
10.3 Clausius-Clapeyron equation	90
10.4 Example of the Van der Waals gas-liquid transition	91
10.5 The Maxwell construction	92
11 Phase transitions: statistical mechanics	94
11.1 Classification of phase transitions	94
11.2 The Ising Model	95
11.3 Exact solution of the 1-dimensional Ising Model	96
11.4 Ordered versus disordered phases	97
11.5 Mean-field theory solution of the Ising model	99
12 Functional integral methods	101
12.1 Path integral representation for the partition function	101
12.2 The classical = high temperature limit	104
12.3 Integrating out momenta	105

Bibliography

Course textbook

- *Statistical Mechanics*,
R.K. Pathria, Elsevier Publishing, (2011).

Classics

- *Thermodynamics and Statistical Mechanics*,
A. Sommerfeld, Academic Press, 1956;
- *Thermodynamics and Statistics*,
E. Fermi, Phoenix Science Series (1966);
- *Statistical Mechanics, An advanced course with problems and solutions*,
R. Kubo, North Holland, 1971;
- *Statistical Physics, Part 1*,
L.D. Landau and E.M. Lifschitz, Pergamon Press, Third edition (1982);
- *Statistical Mechanics, A set of Lectures*,
R.P. Feynman, Benjamin Cummings Publishing (1982);
- *Statistical Thermodynamics*,
E. Schrödinger, Dover Publications (1989).

Standard textbooks

- *Fundamentals of statistical and thermal physics*,
F. Reif, McGraw Hill;
- *From Microphysics to Macrophysics*,
Methods and Applications of Statistical Mechanics, Vol 1,
R. Balian. Springer (2007);
- *Introduction to Statistical Physics*,
K. Huang, CRC Press, Second Edition (2010);
- *Statistical Physics of Particles*,
M. Kardar, Cambridge University Press (2010);
- *Statistical Physics of Fields*,
M. Kardar, Cambridge University Press (2007).

1 Introduction

The goal of statistical mechanics is to explain the physical properties of macroscopic systems in terms of the dynamics of its microscopic constituents.

A macroscopic system is one that contains a large number of microscopic constituents; for example, 12 grams of pure Carbon ^{12}C contains 1 mole or 6.022141×10^{23} Carbon atoms, and constitutes a macroscopic system. As such, almost all of the matter that we encounter in everyday life is macroscopic.

Microscopic constituents may be atoms, such as Carbon atoms in the above example. They may, however, also be larger than atoms and consist of molecules each one of which is composed of several or many atoms. It is also possible to consider the microscopic constituents to be smaller than atoms, and work with nuclei and electrons, as would be appropriate for ionized gases, or even smaller constituents such as quarks and gluons in a quark plasma environment. The choice of which particles to consider as microscopic constituents is guided by the length-scales and time-scales on which the system is considered.

The dynamical laws which individual microscopic constituents obey are generally known. If the microscopic constituents are taken to be electrons and nuclei, and the effects of the weak, strong, and gravitational forces can be neglected, then the dynamics of electromagnetism remains as the only relevant interaction between the microscopic constituents. Almost all matter encountered naturally on Earth should result, in one way or the other, from the dynamics of the electro-magnetic interactions between these constituents.

The physical properties of the corresponding macroscopic system show, however, tremendous qualitative and quantitative differences. The qualitative distinctions are usually referred to as *phases of matter*. In gasses, the microscopic constituents interact only weakly. In both solids and liquids the interactions are strong and produce various degrees of spatial ordering. In crystals and quasi-crystals, only discrete translations and rotations survive as symmetries. Customary liquids are homogeneous and isotropic phases (with unbroken translation and rotation symmetries). Liquid crystals share certain properties with liquids, others with solids. For example, in a nematic phase molecules are aligned along a spatial direction in a translationally invariant way, while in a smectic phase there is ordering along layers but molecules need not be pointing along the same direction. Electric and magnetic properties provide further and finer distinctions between possible phases, giving rise to conductors, semi-conductors, super-conductors, insulators, paramagnets, ferromagnets, and so on.

The goal of statistical mechanics is to explain quantitatively the physical properties of these phases in terms of the dynamics of the underlying microscopic constituents. The time-scales and length-scales over which we seek this information are typically macroscopic, on the order of seconds and centimeters. The degrees of freedom corresponding to time-scales and length-scales on the order of atomic processes will be treated effectively, or *statistically*.

1.1 Brief history

The development of statistical mechanics was driven by attempts to understand thermodynamics from a dynamical microscopic point of view. Thermodynamics is a phenomenological (or purely macroscopic) theory used to describe the transfer of heat, work, and chemical constituents in and out of a macroscopic system. The development of thermodynamics goes back to the early 19-th century, at a time when the concept of energy, and its conservation, were still ill-understood and controversial. Sadi Carnot (1796-1832) uncovered some of the physical principles that underly the functioning of steam-engines, and was led to distinguish between reversible and irreversible processes. James Joule (1818-1889) extended the concept of energy by include heat. In doing so, he clarified conservation of energy, and laid the foundation for the first law of thermodynamics. Rudolf Clausius (1822-1888) clarified the role of energy conservation in the Carnot-cycle, introduced the concept of entropy, and first stated the second law of thermodynamics. The third law of thermodynamics was proposed in 1905 by Walther Nernst (1864-1941).

The development of statistical mechanics also grew out of the kinetic theory of gasses. Daniel Bernoulli (1700-1782) accounted for the pressure exerted by a gas on the walls of its container in terms of collisions of molecules against the wall. The concepts of *mean free path* and *collision rate* were introduced by Clausius, and incorporated by James Clerk Maxwell (1831-1879) into a precise theory for the statistical distribution of velocities obeyed by the molecules of a gas in equilibrium, the so-called *Maxwell distribution*. Ludwig Boltzmann (1844-1906) generalized the Maxwell distribution to include external forces, thereby replacing the purely kinetic energy of the Maxwell distribution by the total energy in the famous Boltzmann factor $e^{-E/kT}$. The Boltzmann transport equation describes the kinetic dynamics of gasses out of equilibrium; it may be used to study how equilibrium is reached in time. The range of applicability of these kinetic theories is limited by the fact that its microscopic constituents are invariably assumed to be weakly (or non-) interacting.

The kinetic theories of gasses provide a statistical molecular interpretation of thermodynamic quantities and phenomena. In particular, Boltzmann's famous relation between the entropy S of a macro-state and the number Ω of micro-states accessible to this macro-state,

$$S = k \ln \Omega$$

provides one of the fundamental links between a thermodynamic quantity S , and a statistical mechanical quantity Ω . Understanding exactly what is being counted by Ω requires introducing the concept of an *ensemble*. The kinetic theory led Maxwell and Boltzmann to introducing ensembles for gasses of non-interacting molecules.

Josiah Willard Gibbs (1839-1903) held the first professorship in theoretical physics in the United States, at Yale University. Although his work on statistical mechanics (a term he

coined) and on vector calculus (which he invented) were carried out in relative isolation, its influence on 20-th century physics has been profound. Amongst many other contributions, Gibbs proposed to think of *ensembles* in an abstract but generally applicable formalism, which allows for the counting of Ω in a systematic statistical framework. Twenty-five years before the discovery of quantum mechanics, Gibbs showed that counting molecules as *distinguishable constituents* leads to the *Gibbs paradox*. The paradox evaporates as soon as constituents are treated quantum mechanically, and counted as indistinguishable. The Gibbs approach to ensembles will be explained in detail and used extensively during this course.

The discovery of quantum mechanics through the analysis of the spectrum of black body radiation by Max Planck (1858-1947) revolutionized physics in general and statistical mechanics in particular. Remarkably, the statistical counting formalism of Gibbs lent itself perfectly to incorporating the changes brought about by the quantum theory. These fundamental changes are as follows,

1. Quantum states are discrete, countable entities;
2. Identical quantum constituents (they could be electrons, protons, atoms, or molecules) are *indistinguishable* from one another and must be counted as such;
3. Integer spin particles obey Bose-Einstein statistics;
4. Half-integer spin particles obey Fermi-Dirac statistics.

Satyendra Bose (1894-1974) pioneered the quantum counting for photons, while Albert Einstein (1879-1955) extended the method to (integer spin) atoms and molecules. Enrico Fermi (1901-1954) and Paul Dirac (1902-1984) invented the modification needed to treat electrons, protons, and other spin 1/2 particles. A systematic revision of the principles of statistical mechanics, required by quantum mechanics, was pioneered by Lev Landau (1908-1968) and John von Neumann (1903-1957), who introduced the tool of the *density matrix*. Landau developed a systematic theory of second order phase transitions, and obtained general predictions for the *critical behavior* near the phase transition point. Landau also invented a general theory of quantum liquids well-suited to describing liquids of bosonic or fermionic excitations, and applied it successfully to explaining the superfluidity of liquid ^4He .

In the second half of the 20-th century, many developments of statistical mechanics went hand in hand with progress in quantum field theory. The functional integral of Dirac and Feynman provided a general starting point for the statistical mechanics of fields. Statistical mechanical methods, such as block spin methods, were applied to the problem of renormalization in quantum field theory, while the renormalization group methods of field theory provided systematic approximation methods in statistical mechanics. Kenneth Wilson (1936-present) fully developed the interplay between quantum field theory and statistical mechanics to derive a physical understanding of renormalization, and to usher in a detailed and calculable approach to the theory of critical phenomena and second order phase transitions.

In the 1970's, Jacob Bekenstein (1947-present) put forward a surprising and bold hypothesis, namely that black holes, when considered quantum mechanically, must have a definite *entropy*. Further evidence for this hypothesis was supplied by Stephen Hawking (1942-present), who showed that the objects we refer to classically as black holes actually emit radiation due to quantum effects. The spectrum of Hawking radiation is that of a black body, which implies that a black hole must have temperature in addition to entropy, and thus behaves as a thermodynamic objects. Intuitively, the assignment of an entropy makes physical sense when one considers that micro-states living inside the black hole horizon are shielded away from us and are thus accessible to us only in the form of a macro-state. From a statistical mechanics point of view, the question thus arises as to what these micro-states inside the event horizon are. For very special types of black holes, the answer is provided by string theory, and the micro-states are found to be certain *branes*, namely spatially extended objects of various dimensions that generalize membranes and strings. But the nature of the micro-states for a regular black hole, like Schwarzschild is not understood to date.

1.2 Units and physical constants

The following values of physical constants may be useful.

Planck's constant	$\hbar = 1.0546 \times 10^{-27} \text{ erg} \times \text{sec}$
speed of light	$c = 2.9979 \times 10^{10} \text{ cm} \times \text{sec}^{-1}$
Newton's constant	$G_N = 6.6720 \times 10^{-8} \text{ cm}^3 \times g^{-1} \times \text{sec}^{-2}$
Boltzmann's constant	$k = 1.3807 \times 10^{-16} \text{ erg} \times \text{K}^{-1}$
	$k = 8.6175 \times 10^{-5} \text{ eV} \times \text{K}^{-1}$
electron charge	$e = 1.6022 \times 10^{-19} \text{ Coulomb (in MKS units)}$
electron mass	$m_e = 9.11 \times 10^{-28} \text{ g} = 0.511 \text{ MeV}/c^2$
Bohr radius	$r_B = \hbar/(m_e c) = 0.53 \text{ \AA}$
Avogadro's number	$N_A = 6.022141 \times 10^{23}$
ideal gas constant	$R = 8.32 \text{ J/K}$

The following conversion relations between physical units may be useful.

1 \AA	$= 10^{-8} \text{ cm}$
1 erg	$= 1 \text{ g} \times \text{cm}^2 \times \text{sec}^{-2}$
1 erg	$= 6.241 \times 10^{11} \text{ eV}$
1 erg	$= 10^{-7} \text{ J}$
1 eV	$= 1.6022 \times 10^{-19} \text{ J}$
1 eV	$= 1.6022 \times 10^{-12} \text{ erg}$
1 eV	$= 1.1605 \times 10^4 \text{ K}$
1 cal	$= 4.184 \text{ J}$
1 Btu	$= 1055 \text{ J}$

2 Statistical Physics and Thermodynamics

In this section, we shall introduce some of the basic methods of statistical mechanics and their implications for thermodynamics in a relatively informal way. In the subsequent section, more complete definitions and applications of statistical ensembles will be introduced and applied. In as much as possible, we shall discuss the cases of classical mechanics and quantum physics in parallel.

2.1 Micro-states and macro-states

Consider a system of an extremely large number N of microscopic constituents (often referred to as *particles*), for example on the order of Avogadro's number $\sim 10^{23}$. A micro-state corresponds to a precise and complete specification of each and every one of these particles.

From the point of view of classical mechanics of point-like particles, a micro-state will be described by the generalized positions $q_i(t)$ and their velocities $\dot{q}_i(t)$ for $i = 1, \dots, 3N$ in the Lagrangian formulation or the generalized momenta $p_i(t)$ and positions $q_i(t)$ in Hamiltonian mechanics. From the point of view of quantum mechanics, a microscopic state corresponds to a vector $|\psi\rangle$ in the Hilbert space \mathcal{H} of the quantum mechanical system. In a coordinate basis, the state $|\psi\rangle$ is represented by a wave function $\psi(q_1, \dots, q_{3N})$ though additional quantum numbers, such as spin, may have to be included as well. In both the classical and the quantum mechanical cases it would be utterly impossible to completely specifying a micro-state, as doing so would require supplying on the order of $N \sim 10^{23}$ numerical entries.

A macro-state, either from the point of view of classical or quantum mechanics, is specified by a relatively small number of thermodynamic variables. Exactly which variables are specified will depend on the nature of the physical system. The geometrical volume V in which the system is constrained to evolve is often used to specify a macro-state. The total energy E of a system plays a special role in view of the general principle of energy conservation, and is almost always used to specify macro-states. In fact, any quantity that is conserved on general principle may be used to specify macro-states. These include total momentum, total angular momentum, total electric charge, total magnetization, and total particle number (in situations where the dynamics of the system conserves particle number). Their conservation guarantees their exact matching between the micro-states and the macro-states.

2.1.1 Counting micro-states; entropy

Many different micro-states will correspond to one single macro-state. Given the thermodynamic variables specifying a macro-state, we denote by Ω the number of micro-states in which the given macro-state can be realized. If a macro-state is specified by its internal energy E , then the number of micro-states in terms of which this macro-state can be realized will be

a function of E , and denoted $\Omega(E)$. If the total volume V is used in addition, Ω becomes a function of two variables $\Omega(E, V)$, and so on. The more thermodynamic variables we use, the more precise our specification of the macro-state will be. In practice, macro-states will be described by just a few thermodynamic variables. Thus, specifying a macro-state will always amount to *omitting a large amount of information about the system*.

The key logical connection between statistical mechanics and thermodynamics is made by Boltzmann's formula for entropy S ,

$$S = k \ln \Omega \tag{2.1}$$

in terms of the number of micro-states Ω accessible to a given macro-state. For macro-states specified by energy and volume, for example, both sides will be functions of E, V . The dependence of $\Omega(E, V)$ and thus $S(E, V)$ on E and V is governed by the specific dynamical laws of the system. In classical mechanics as well as in quantum physics these laws are encapsulated in the microscopic degrees of freedom, and the precise form of the Hamiltonian in terms of these degrees of freedom. The exact computation of Ω and S is possible only for non-interacting systems. Yet, non-interacting systems provide approximations to a wealth of real physical problems, and we will carry out the calculations of Ω and S there shortly.

2.2 Statistical and thermal equilibrium

From common experience and detailed experiments we know that an *isolated system*, namely one which has no contacts or exchanges with the outside, will settle into a preferred state of *thermal equilibrium* after a sufficiently long time has passed. More precisely, statistical equilibrium may be defined as the thermodynamic state in which the given macro-state may be realized in terms of the largest possible number of micro-states. In other words, *statistical equilibrium* is achieved for maximal disorder. As a result, the state of statistical equilibrium may be quantitatively characterized as the macro-state with maximal entropy.

To see how this works, consider two macroscopic systems $\mathcal{S}_1, \mathcal{S}_2$ parametrized by the thermodynamic variables E_1, V_1, E_2, V_2 . Without spelling out the precise microscopic dynamics at work here, we shall simply assume that the numbers of micro-states accessible to each one of the above macro-states are given by $\Omega_1(E_1, V_1)$ and $\Omega_2(E_2, V_2)$ respectively. We now bring these two macro-systems in contact with one another, as shown in the figure 1 below.

In a first stage, we keep the volumes V_1, V_2 fixed, but allow for energy exchange between \mathcal{S}_1 and \mathcal{S}_2 , for example by placing a heat-conducting screen in between them. The total system will be denoted by \mathcal{S} , and is assumed to be isolated. By energy conservation, the total energy E of the system is given by

$$E = E_1 + E_2 \tag{2.2}$$

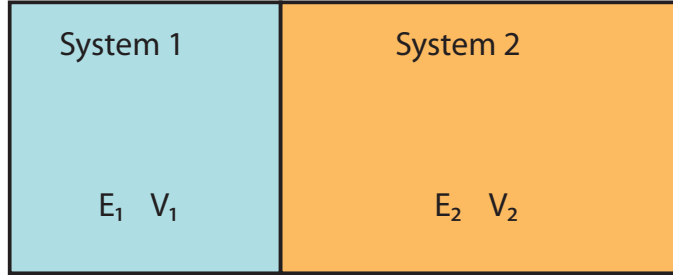


Figure 1: Two macroscopic systems \mathcal{S}_1 and \mathcal{S}_2 in contact.

and is fixed since \mathcal{S} is assumed to be isolated. In general, the number of accessible micro-states $\Omega(E)$ of \mathcal{S} at energy E (recall the volumes are fixed) depends on the interactions between \mathcal{S}_1 and \mathcal{S}_2 and is not readily calculable. We shall consider here an idealized situation where the interactions between \mathcal{S}_1 and \mathcal{S}_2 are neglected (which can be justified in part since they occur through a surface and have small effects in the bulk). As a result, the system \mathcal{S} at energy E may be in any of the macro-states (E_1, E_2) with $E_1 + E_2 = E$, with the following corresponding number of micro-states,

$$\Omega_0(E_1, E_2) = \Omega_1(E_1)\Omega_2(E_2) \quad (2.3)$$

We see that the partition of \mathcal{S} into the subsystems \mathcal{S}_1 and \mathcal{S}_2 which are in mutual thermal contact provides us with one macroscopic parameter which is not yet fixed, namely the energy E_1 of the subsystem \mathcal{S}_1 . The state of equilibrium will be achieved when Ω_0 is maximal for given E . In the process of evolving towards equilibrium, the free parameter E_1 will adjust to produce the maximum possible value for $\Omega_0(E_1, E_2)$, so that we have,

$$\frac{\partial \ln \Omega_0(E_1, E_2)}{\partial E_1} = \frac{\partial \ln \Omega_1(E_1)}{\partial E_1} - \frac{\partial \ln \Omega_2(E_2)}{\partial E_2} = 0 \quad (2.4)$$

Partial derivatives are used to make it clear that all other thermodynamic variables, such as V and N , are being kept constant. Recognizing $k \ln \Omega$ as the entropy by Boltzmann's equation (2.1), and the derivative of S with respect to E as inverse of the *absolute temperature* T ,

$$\frac{1}{T} = \frac{\partial S(E)}{\partial E} \quad (2.5)$$

the equilibrium condition (2.4) may be re-expressed in terms of the temperature T of the two systems,

$$T_1 = T_2 \quad (2.6)$$

namely, at equilibrium, the temperatures of the two subsystems coincide.

Extending the above derivation, assume that the systems \mathcal{S}_1 and \mathcal{S}_2 are not only in thermal contact, but that also the wall separating them is movable, so that the volumes V_1 and V_2 may vary, all the while keeping the total volume V of the system fixed,

$$V = V_1 + V_2 \quad (2.7)$$

In addition, let us assume that one or several further conserved quantities N , such as momentum, angular momentum, or the conserved numbers of various species of particles, may also be exchanged between the two systems, all the while keeping their total value fixed,

$$N = N_1 + N_2 \quad (2.8)$$

By the same arguments as were used for energy exchange, equilibrium of the system \mathcal{S} will be characterized by the following equalities,

$$\begin{aligned} \frac{\partial \ln \Omega_1(E_1, V_1, N_1)}{\partial V_1} - \frac{\partial \ln \Omega_2(E_2, V_2, N_2)}{\partial V_2} &= 0 \\ \frac{\partial \ln \Omega_1(E_1, V_1, N_1)}{\partial N_1} - \frac{\partial \ln \Omega_2(E_2, V_2, N_2)}{\partial N_2} &= 0 \end{aligned} \quad (2.9)$$

In terms of entropy S , temperature T , pressure P , and chemical potential μ , defined by,

$$\begin{aligned} \left. \frac{\partial S(E, V, N)}{\partial E} \right|_{V, N} &= \frac{1}{T} \\ \left. \frac{\partial S(E, V, N)}{\partial V} \right|_{E, N} &= \frac{P}{T} \\ \left. \frac{\partial S(E, V, N)}{\partial N} \right|_{E, V} &= -\frac{\mu}{T} \end{aligned} \quad (2.10)$$

the equilibrium conditions become,

$$\begin{aligned} T_1 &= T_2 \\ P_1 &= P_2 \\ \mu_1 &= \mu_2 \end{aligned} \quad (2.11)$$

which constitute, of course, the well-known equations of thermodynamical equilibrium. Thus equilibrium guarantees that T, P, μ are constant throughout the system.

These relations may be further generalized to the case of more than two subsystems in thermal contact. If \mathcal{S} may be subdivided into three subsystems $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, which are all in mutual thermal contact, then the equilibrium conditions between $\mathcal{S}_1, \mathcal{S}_2$ on the one hand, and $\mathcal{S}_2, \mathcal{S}_3$ on the other hand, will imply that the subsystems \mathcal{S}_1 and \mathcal{S}_3 are also in equilibrium with one another. This fact is referred to as the 0-th law of thermodynamics in Kardar.

2.3 The thermodynamic limit, extensive and intensive variables

Putting together multiple identical copies of a given macro-system allows us to scale the energy E , volume V , and particle number N , as well as any other conserved quantities, by a common scale factor λ ,

$$E \rightarrow \lambda E \qquad V \rightarrow \lambda V \qquad N \rightarrow \lambda N \qquad (2.12)$$

This scaling will be reliable in the limit where boundary and surface effects can be neglected, an approximation which is expected to hold in the limit where $\lambda \rightarrow \infty$. This limit is referred to as the *thermodynamic limit*.

Thermodynamic variables which scale linearly with λ are referred to as *extensive*, while those which are untransformed under λ are referred to as *intensive*. For example, the energy density E/V , and the number density N/V are intensive. Given that thermodynamic equilibrium sets equal T , P , and μ across the bulk of a substance, these quantities should be intensive. As a result of the defining relations in (2.10), the entropy is expected to be an extensive thermodynamic variable.

2.4 Work and Heat

The partial derivative relations of (2.10) may equivalently be expressed in differential form,

$$dE = TdS - PdV + \mu dN \qquad (2.13)$$

and may be viewed as a fundamental relation of thermodynamics. Carefully note what this formula means. It gives the relation between the changes in thermodynamic variables for *neighboring states of thermodynamic equilibrium* of two subsystems in contact with one another, the whole system \mathcal{S} being isolated, and not undergoing any change in total entropy. Such changes are referred to as *reversible* since no total entropy is produced.

In classical and quantum mechanics, it is often convenient to separate contributions to energy into kinetic from potential, even though only the sum of the two is conserved. The potential energy may be viewed as summarizing the change in the Hamiltonian due to changes in the *external conditions*, such as the volume, or an electric potential. In thermodynamics, the changes dE in the internal energy which result from changes in the external conditions, such as a change in volume V or in the conserved quantity N are referred to as the *work δW done by the system*. In the case at hand, the contributions to the work are given by,

$$\delta W = PdV - \mu dN \qquad (2.14)$$

Since the parameters V and N are not statistical in nature, but are supplied as external conditions, this relation holds for reversible as well as irreversible processes, as long as the changes produced (in time) are such that the system remains in equilibrium throughout.

All changes in the internal energy which are not work will be lumped together as *heat*, and denoted by δQ . Thus, heat transfer may be thought of as associated with changes in the kinetic energy of the system.

2.4.1 The first law of thermodynamics

The *first law of thermodynamics* reflects the principle of conservation of energy,

$$dE = \delta Q - \delta W \quad (2.15)$$

The minus sign in front of work is a matter of convention, and results from the fact that we have defined the work done *by* the system (as opposed to the work done *on* the system).

The notations δQ and δW are used here to stress the following important distinction with dE, dS, dV etc. The thermodynamic variables V, E, N, S, \dots are all *thermodynamic state functions*, and depend only on the (equilibrium) macro-state under consideration. Thus, the corresponding differentials dV, dE, dN, dS, \dots are exact differentials, and their integration depends only on the initial and final macro-states of the system. Heat and work are not, in general, thermodynamic state functions, and integrals of the differentials $\delta Q, \delta W$ do depend on the path followed to go from one macro-state of the system to another. The differentials δQ and δW are not closed (and thus not exact); the notation δ is used to indicate this (Clausius introduced this distinction, and used the notation \bar{d} rather than δ).

When the thermodynamic changes are *reversible*, we may combine formulas (2.14) and (2.15) which are generally valid, with formula (2.13) which is valid only for reversible processes, to obtain a relation between heat and entropy,

$$\delta Q = TdS \quad (2.16)$$

In view of its very derivation, this relation will hold *only for reversible processes*. For irreversible processes, the relation does not hold. Instead, we need to appeal to the second law of thermodynamics to clarify the situation.

2.4.2 The second law of thermodynamics

The *second law of thermodynamics* may be formulated in a number of equivalent ways.

1. **Clausius:** *No thermodynamic process is possible whose sole result is the transfer of heat from a cooler to a hotter body.*
2. **Kelvin:** *No thermodynamic process is possible whose sole result is the complete conversion of heat into work.*

These statements may be translated into an inequality with the help of a theorem by Clausius (proven, for example, in Kardar.) In a general thermodynamic process (reversible or irreversible), the entropy change dS for given heat transfer δQ and temperature will be larger than $\delta Q/T$, so that for any cyclic thermodynamic process we have,

$$\oint \frac{\delta Q}{T} \leq 0 \quad (2.17)$$

Here, δQ is the heat increment supplied to the system at temperature T . For any reversible processes we have equality, while for any irreversible process we have strict inequality.

2.5 Relations between thermodynamic variables

Next, we derive a number of immediate but fundamental thermodynamic relations from (2.13), an equation we repeat here for convenience,

$$dE = TdS - PdV + \mu dN \quad (2.18)$$

(1) Applying the Euler equation for homogeneous functions of degree 1 to the entropy in the thermodynamic limit, and using the defining equations of (2.10), we derive the relation,

$$E = TS - PV + \mu N \quad (2.19)$$

Taking the total differential of this relation, and then eliminating dE with the help of (2.18) gives the *Gibbs-Duhem relation*,

$$SdT - VdP + Nd\mu = 0 \quad (2.20)$$

(2) Next, we provide alternative formulas for T , P , and μ as derivatives of the internal energy E rather than of the entropy S . The derivations are standard, and extend to many other thermodynamic relations. It is immediate from (2.18) that we alternatively have,

$$T = \left(\frac{\partial E}{\partial S} \right)_{V,N} \quad P = - \left(\frac{\partial E}{\partial V} \right)_{S,N} \quad \mu = \left(\frac{\partial E}{\partial N} \right)_{S,V} \quad (2.21)$$

Homogeneity of E now leads to the same equation as we found in (2.19).

(3) To change independent variables from S, V, N to T, V, N we use the free energy F ,

$$F = E - TS \quad (2.22)$$

in terms of which equation (2.18) takes the form,

$$dF = -SdT - PdV + \mu dN \quad (2.23)$$

It follows that S , P , and μ may be represented as derivatives of F with respect to the variables respectively T , V , and N ,

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V,N} \quad P = - \left(\frac{\partial F}{\partial V} \right)_{T,N} \quad \mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} \quad (2.24)$$

(4) The specific heat functions C_V and C_P are defined as follows,

$$C_V = T \left(\frac{\partial S}{\partial T} \right)_{V,N} \quad C_P = T \left(\frac{\partial S}{\partial T} \right)_{P,N} \quad (2.25)$$

Again, alternative formulas are available in terms of derivatives of the internal energy, or free energy. We quote the alternative formulas for C_V and C_P ,

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{V,N} \quad C_P = \left(\frac{\partial E}{\partial T} \right)_{P,N} + P \left(\frac{\partial V}{\partial T} \right)_{P,N} \quad (2.26)$$

2.6 Non-interacting microscopic constituents: the ideal gas

In a macroscopic system of non-interacting microscopic constituents, the number N of particles is conserved, as is the total energy E of the system. One refers to such a system as an *ideal gas* since, generally, gasses are characterized by weakly interacting particles.

Without any detailed model or calculations, and using scaling arguments alone, one can deduce the dependence of the number of states $\Omega(E, V, N)$ on V (always in the thermodynamic limit where $V \rightarrow \infty$). Assuming that the size of the particles is negligible compared to V , we see that introducing a single particle in a volume V produces a number of states proportional to V (doubling V will double the number of states). Now, since the particles are non-interacting and of negligible size, introducing a second particle will just multiply the number of 1-particle states by V etc. Thus, we have,

$$\begin{aligned} \Omega(E, V, N) &\sim V^N \\ S(E, V, N) &\sim N \ln V \end{aligned} \quad (2.27)$$

Computing the pressure, using the second relation in (2.10), we find $P = NkT/V$. It is customary to express this relation in terms of the *gas constant* $R = N_A k$, defined in terms of k and Avogadro's number N_A , and takes the form of the ideal gas law,

$$PV = nRT = NkT \quad (2.28)$$

where n is the number of moles of the gas.

2.7 The simplest ideal gas

Not all ideal gasses are the same. This is due to the fact that the number of degrees of freedom of each microscopic constituent by itself may vary. Each microscopic constituent in any ideal gas will have three translational degrees of freedom, namely the three components of momentum. Clearly, the simplest ideal gas corresponds to the case where these are the only degrees of freedom of each microscopic constituent. More generally, the atoms and molecules in a general ideal gas will have also rotational and internal degrees of freedom, with corresponding quantum numbers. In this section, we shall treat only the simplest ideal gas, and postpone treatment of more complicated cases.

To also extract the dependence on energy and particle number, we need a more detailed model, specifying in particular whether the gas is relativistic or not, and quantum mechanical or classical. We shall work with the non-relativistic quantum version, for non-interacting particles confined to a square box of size L and volume $V = L^3$, with Dirichlet boundary conditions (or equivalently an infinite positive potential outside the box). The micro-states of the system are specified by the quantum state of each particle $i = 1, \dots, N$ by three non-negative integers $n^i = (n_x^i, n_y^i, n_z^i)$. The energy of a particle in this state is given by,

$$\varepsilon(n_x^i, n_y^i, n_z^i) = \frac{\pi^2 \hbar^2}{2mL^2} \left((n_x^i)^2 + (n_y^i)^2 + (n_z^i)^2 \right) \quad (2.29)$$

The energy of the system in a micro-state specified by the n^i is given by,

$$E = \sum_{i=1}^N \varepsilon(n_x^i, n_y^i, n_z^i) \quad (2.30)$$

Thus, specifying a macro-state by E, V , and N then precisely determines $\Omega(E, V, N)$ as the number of solutions $n_r \geq 0$ for $r = 1, \dots, 3N$ to the equation,

$$\sum_{r=1}^{3N} n_r^2 = \frac{2mEL^2}{\pi^2 \hbar^2} \quad (2.31)$$

We see that the counting of states depends on N , and on the combination $EL^2 = EV^{2/3}$ but not on E and V separately. As a result, the functional dependence of S can be simplified, since it may be expressed in terms of a function s of only 2 variables,

$$S(E, V, N) = s(EV^{2/3}, N) \quad (2.32)$$

Therefore, the derivatives of S with respect to E and V are related to one another,

$$E \frac{\partial S}{\partial E} \Big|_{V,N} - \frac{3}{2} V \frac{\partial S}{\partial V} \Big|_{E,N} = 0 \quad (2.33)$$

so that one derives the following formulas,

$$P = \frac{2}{3} \frac{E}{V} \qquad E = \frac{3}{2} NkT \qquad (2.34)$$

The second relation reproduces the classic result that the average thermal energy per particle is $\frac{3}{2}kT$, or the average energy per degree of freedom is $\frac{1}{2}kT$. Finally, combining the earlier result of (2.27) with that of (2.32), we derive the complete dependence on E as well,

$$S(E, V, N) = N \ln \left(\frac{VE^{3/2}}{N^{5/2}} \right) + \tilde{s}(N) \qquad (2.35)$$

where $\tilde{s}(N)$ depends only on N . If S is to be an extensive quantity, then $\tilde{s}(N)$ should be a linear function of N , given by $\tilde{s}(N) = s_0N$ for some constant s_0 .

2.7.1 Counting micro-states in the simplest ideal gas

Next, we investigate how the actual counting of micro-states for the non-relativistic ideal gas proceeds, and which results we obtain. The combinatorial problem is well-posed in (2.31). It may be solved in the approximation where $mEL^2/\hbar^2 \gg 1$ by computing the volume of the positive n_r “quadrant” of a sphere in $3N$ dimensions, of radius ν defined by,

$$\nu^2 = \frac{2mEV^{2/3}}{\pi^2\hbar^2} \qquad (2.36)$$

The volume of the $n - 1$ dimensional sphere of unit radius (embedded in n dimensional flat Euclidean space) is given by,¹

$$V_{S^{n-1}} = \frac{2\pi^{n/2}}{\Gamma(n/2)} \qquad (2.37)$$

As a result, the number of micro-states is given by,

$$\Omega(E, V, N) \approx \frac{1}{2^{3N}} \frac{2\pi^{3N/2}\nu^{3N-1}}{\Gamma(3N/2)} \approx \frac{2}{\Gamma(3N/2)} \left(\frac{mEV^{2/3}}{2\pi\hbar^2} \right)^{3N/2} \qquad (2.38)$$

In the second approximation, we have used $N \gg 1$ is large to drop the -1 in the exponent of ν . Using the Sterling formula for the Γ -function evaluated at large argument $n \gg 1$,

$$\ln \Gamma(n + 1) \approx n \ln(n) - n + \frac{1}{2} \ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right) \qquad (2.39)$$

¹This formula may be obtained by evaluating an n -dimensional spherically symmetric Gaussian integral in two different ways: first as a product of identical 1-dimensional Gaussian integrals; second by expressing the integral in spherical coordinates, and factoring out the volume $V_{S^{n-1}}$.

and converting the result of (2.38) to a formula for the entropy, we find,

$$S(E, V, N) \approx kN \ln \left[\frac{V}{N} \left(\frac{mE}{3\pi\hbar^2 N} \right)^{3/2} \right] + \frac{3}{2}kN + kN \ln N \quad (2.40)$$

Note that the argument of the logarithm is properly dimensionless, and forms an intensive combination of E, V, N . The first two terms are extensive contributions, as one would expect for the total entropy, and their dependence on E and V is in accord with the results derived on general scaling grounds in equation (2.35). The last term is consistent with the form of (2.35) as well, but fails to be properly extensive. It is useful to re-express E in terms of T using the second equation in (2.34), and we find,

$$S(E, V, N) = kN \ln \left[\frac{V}{N} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} \right] + \frac{3}{2}kN + kN \ln N \quad (2.41)$$

The failure of the last term to be properly extensive is referred to as the *Gibbs paradox*.

2.8 Mixing entropy and the Gibbs paradox

The non-extensive formula for the entropy leads to a paradox for the following reasons. Consider again the situation of the two systems \mathcal{S}_1 and \mathcal{S}_2 depicted in figure 1, with gasses such that the masses of their microscopic constituents are equal $m_1 = m_2 = m$. We begin by assuming that each system is isolated and in thermal equilibrium, and that their temperatures coincide $T_1 = T_2 = T$. We shall also assume that their pressures coincide so that, by the ideal gas law (2.28), the number densities must also be equal, $N_1/V_1 = N_2/V_2 = \rho$. We will leave their chemical potentials unconstrained.

First, consider bringing the systems \mathcal{S}_1 and \mathcal{S}_2 into thermal contact, by allowing exchanges of energy and volume, but not particle number. Introduction or removal of thermal contact are then reversible processes, and no increase in entropy is generated, since the preparation of the systems guarantees thermal equilibrium of \mathcal{S} , as long as no particles are being exchanged.

Second, consider the gasses in \mathcal{S}_1 and \mathcal{S}_2 to be composed of identical particles, and allowing exchange also of particle number. Again, the introduction or removal of thermal contact are reversible processes with no associated entropy increases.

Third, consider the gas in \mathcal{S}_1 to be composed of one species of particles, but the gas in \mathcal{S}_2 to be composed of another species, which happen to have equal masses (this can happen approximately to rather high precision in Nature; for example with the proton and the neutron, or with isobaric nuclei of different elements such as ^{40}K and ^{40}Ca). Now, we know that allowing the two gasses to mix will increase the entropy. One way to see this is that each species of particles can have its separate chemical potential, and these chemical potentials

are different for the two isolated systems \mathcal{S}_1 and \mathcal{S}_2 prior to thermal contact. Thus, upon producing thermal contact, the total system will not be in thermal equilibrium.

Let us explore the predictions of formula (2.41), in order to compare the entropies S_α of the systems \mathcal{S}_α for $\alpha = 1, 2$,

$$S_\alpha = kN_\alpha \ln \left[\frac{1}{\rho} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} \right] + kN_\alpha \left(\frac{3}{2} + \ln N_\alpha \right) \quad (2.42)$$

with the entropy S of the total system \mathcal{S} ,

$$S = k(N_1 + N_2) \ln \left[\frac{1}{\rho} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} \right] + k(N_1 + N_2) \left(\frac{3}{2} + \ln(N_1 + N_2) \right) \quad (2.43)$$

The mixing entropy increase $S - S_1 - S_2$ may be evaluated using the Stirling formula,

$$S - S_1 - S_2 = k \ln \frac{(N_1 + N_2)!}{N_1! N_2!} \quad (2.44)$$

a result which clearly cries out for a combinatorial interpretation.

2.9 Indistinguishability of identical particles

When Gibbs articulated the above paradox in 1875, quantum mechanics was still 25 years away, and he did not have the factor of \hbar in the expression for the entropy. The paradox existed even then, and Gibbs proposed an ad hoc remedy by dividing the number of accessible micro-states by the combinatorial factor $N!$.

We now know the reason for this factor: in quantum mechanics, identical particles are *indistinguishable*. Proper counting of indistinguishable particles requires the inclusion of the factor $N!$, though we also know, of course, that quantum mechanics will impose further modifications on this counting, depending on whether we are dealing with fermions or bosons. The result is the Sackur-Tetrode formula,

$$S = kN \ln \left(\frac{V}{N} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} \right) + \frac{5}{2}kN \quad (2.45)$$

It was our initial treatment of the identical particles as distinguishable (namely each one being labeled by (p_i, q_i)), that produced the paradox.

The heat capacities C_V, C_P for the ideal gas may be evaluated using nothing more than the ideal gas laws of (2.34), and we find,

$$C_V = \frac{3}{2}Nk \quad C_P = \frac{5}{2}Nk \quad (2.46)$$

Evaluating the internal entropy as a function of entropy allows us to compute the chemical potential for the ideal gas, and we find,

$$\begin{aligned} E &= \frac{3\pi\hbar^2 N^{5/3}}{mV^{2/3}} \exp\left(\frac{2S}{3kN} - \frac{5}{3}\right) \\ \mu &= -kT \ln\left(\frac{V}{N} \left(\frac{mkT}{2\pi\hbar^2}\right)^{3/2}\right) \end{aligned} \quad (2.47)$$

Note that the ubiquitous combination,

$$\lambda = \left(\frac{2\pi\hbar^2}{mkT}\right)^{1/2} \quad (2.48)$$

has dimension of length, and is usually referred to as the *thermal wavelength*.

Recasting the formula for the entropy (2.45) in terms of the density $\rho = N/V$ and the thermal wavelength λ , we find,

$$S = -kN \ln(\rho\lambda^3) + \frac{5}{2}kN \quad (2.49)$$

Since the entropy must be positive by definition, we see that the above formula is valid only when the density is small compared with the scale set by the thermal wavelength

$$\rho\lambda^3 \ll 1 \quad (2.50)$$

This is the condition for a Boltzmann (or sufficiently dilute) gas. For $\rho\lambda^3 \geq 1$, we will need to appeal to the full quantum statistics of Bose-Einstein or Fermi-Dirac statistic instead.

2.10 Thermodynamic transformations

A thermodynamic transformation corresponds to a change of the macroscopic state of a system. If the thermodynamic variables are, for example, P, V, N , then any change in these variables produces a thermodynamic transformation. Special thermodynamic transformations are often of interest, and usually correspond to keeping one or another thermodynamic variable fixed. If a state is initially in equilibrium, then a thermodynamic transformation can be brought about only by changing the external conditions on the system.

We shall distinguish the following thermodynamic transformations,

1. *Quasi-static*: if the external conditions change slowly enough to keep the system in equilibrium throughout the transformation process.

2. *Reversible*: a transformation producing no increase in total entropy. (A reversible process is quasi-static, but the converse does not hold.)
3. *Adiabatic*: a transformation in which only the external conditions on the system are changed, but no heat is transferred: $\delta Q = 0$. As a result, we have $dE = -\delta W$. (Note that Landau and Lifshytz use terminology where the term stands for what we understand to be reversible adiabatic; see next entry.)
4. *Reversible adiabatic*: combining the two properties, one shows that the entropy of the system remains unchanged. (For the ideal gas, this means that $VT^{2/3}$ or equivalently $PV^{5/3}$ are constant, for fixed particle number N .)
5. *Isothermal*: a process at constant temperature. (For an ideal gas, the internal energy is then constant, and the product PV is constant, for fixed particle number N .)

3 Classical Statistical Ensembles

In the preceding section, we have succeeded in counting the number of micro-states for an ideal gas of non-interacting non-relativistic particles confined to a box. As soon as interactions are turned on, however, an exact counting of micro-states in the thermodynamic limit can be carried out only very rarely. Therefore, a general formalism of the statistical approach to arbitrary systems is needed which, in particular, will permit a systematic calculation of perturbative approximations.

The general formalism of statistical mechanics is based on the theory of *statistical ensembles*, a concept which dates back to Maxwell and Boltzmann for gasses, and which was articulated in wide generality by Gibbs. We shall begin with ensembles in classical mechanics, where the formulation may be set and understood in more familiar terrain, and introduce the density function. Ensembles in the quantum case will be introduced in the subsequent section, where we shall also distinguish between pure and mixed ensembles, and introduce the density matrix.

3.1 Classical mechanics

The Hamiltonian formulation of classical mechanics will be used to describe a system of N classical particles. Points in phase space \mathcal{P} may be described by generalized coordinates (p_i, q_i) with $i = 1, \dots, s$. For point particles without internal degrees of freedom in three space dimensions, we have $s = 3N$; when further internal degrees of freedom are present, such as rotational or vibrational degrees of freedom, we will have $s > 3N$. A point in phase space completely specifies a mechanical state at a given time, and corresponds to a micro-state. Phase space \mathcal{P} is the space of all possible micro-states of the classical system.

Dynamics is governed by a classical Hamiltonian $H(p, q; t) = H(p_1, \dots, p_s, q_1, \dots, q_s; t)$. The time evolution of a micro-state, namely the evolution in time t of its N particles, is given by the time evolution of the generalized coordinates via Hamilton equations,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad \dot{q}_i = \frac{\partial H}{\partial p_i} \qquad (3.1)$$

where the dot denotes the derivative with respect to t . Time-evolution produces a *flow* on phase space. When the Hamiltonian H has no explicit time-dependence, and is of the form $H(p, q)$, the total energy of the system is conserved. We may then consistently restrict attention to dynamics at fixed total energy E ,

$$H(p, q) = E \qquad (3.2)$$

for each value of E . The corresponding subspaces \mathcal{P}_E of phase space do not intersect one another for different values of E , and the family of spaces \mathcal{P}_E provides a *foliation* of \mathcal{P} ,

depicted schematically in Figure 2. Some of these leaves may be empty sets; this happened for example for negative E with the harmonic oscillator Hamiltonian, which is non-negative throughout \mathcal{P} .

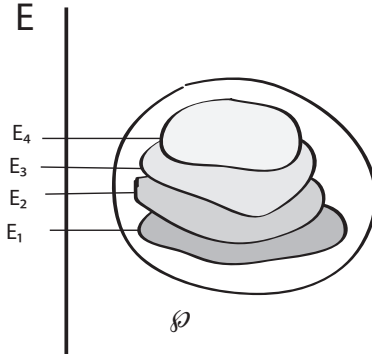


Figure 2: Foliation of phase space \mathcal{P} by the conserved energy variable E . Leaves \mathcal{P}_E at different energies do not intersect.

3.2 Time averaging

For a very large number N of interacting particles, it will be impossible to specify the initial conditions and to solve the evolution equations for individual micro-states. Instead, we shall be interested in predicting statistically averaged quantities.

One conceivable average is over long periods of time. Consider a mechanical function $f(p, q)$ on phase space, and follow its evolution in time $f(p(t), q(t))$ under Hamilton's equations (3.1). Since we cannot predict $p(t)$ and $q(t)$, we cannot predict $f(p(t), q(t))$ either. Calculating the average of f over long periods of time,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_t^{t+T} dt' f(p(t'), q(t')) \quad (3.3)$$

may offer a better prospect. Although the time-averaging approach may seem natural, it has a number of drawbacks. The time-averaging of (3.3) may depend on the initial time t , and on the initial conditions imposed on the time-evolution. Ideally, such data will be washed out for large enough T provided we have strong enough interactions. Unfortunately, characterizing the behavior of a system over long time intervals is very complicated, and the question whether the initial conditions will be or are will not be washed out for large enough T introduces a difficult dynamical complication. In particular, during its time evolution, the system may or may not reach all of the allowed phase space, a problem that is the subject of the *ergodic theorem* of dynamical systems. In fact, for free particles (say on a torus with

period boundary conditions), or for systems with maximal numbers of conserved quantities, we know that not all of the phase space allowed by total energy will be reached.

3.3 Statistical ensembles

The Gibbs approach to statistical mechanics via the use of ensembles greatly simplifies and clarifies the manner in which we take averages. The approach also enjoys better physical motivation, as well as much wider applicability due to its more formal setting (an excellent combination for making progress).

Let us begin with the physical motivation in the case of classical mechanics. When we wish to make statistical predictions, for example about the macroscopic system of one liter of water, we are interested not so much in predicting the behavior of a single specific one-liter bottle of water (whose initial conditions or micro-state we would not know precisely anyway). Instead, we are rather interested in making predictions about *any liter of water with the same composition, and macroscopic variables*. In particular, we do not prescribe the initial conditions of the system beyond giving its macroscopic variables. Therefore, from the outset, we are interested in considering together all possible micro-states to which a given macro-state has access, and averaging over them. To summarize, in addition to time averaging of just one specific system, we will take an average also over all possible similar systems, with the same Hamiltonian, but with different initial conditions. This procedure makes physical sense because the initial conditions on the systems were not specified for a given macro-state anyway.

Let us now provide a precise set-up. Following Gibbs, we consider not one system, but rather a collection of \mathcal{N} systems, the dynamics of each one of these systems being governed by the same Hamiltonian $H(p, q; t)$, but whose initial conditions will be different. Recall that, at any given time, a single system is characterized by a single point in phase space \mathcal{P} . Thus, at any given time, the collection of \mathcal{N} systems will be characterized by a *cloud* of \mathcal{N} points in phase space, each point corresponding precisely to one system. Such a collection (or cloud) of phase space points is referred to as an *ensemble*. Each point (or system) in the ensemble evolves in time according to the same Hamiltonian $H(p, q; t)$.

It is important to point out that, by construction, *the different systems collected in an ensemble do not interact with one another*.

It cannot be emphasized enough that a system composed of N similar interacting elements (or similar interacting subsystems, such as for a gas of N interacting identical molecules) *is not an ensemble of independent systems*, and should not be confused with an ensemble.²

²To emphasize this distinction, we shall denote the number of particles or subsystems by N , but the number of non-interacting systems collected in an ensemble by \mathcal{N} .

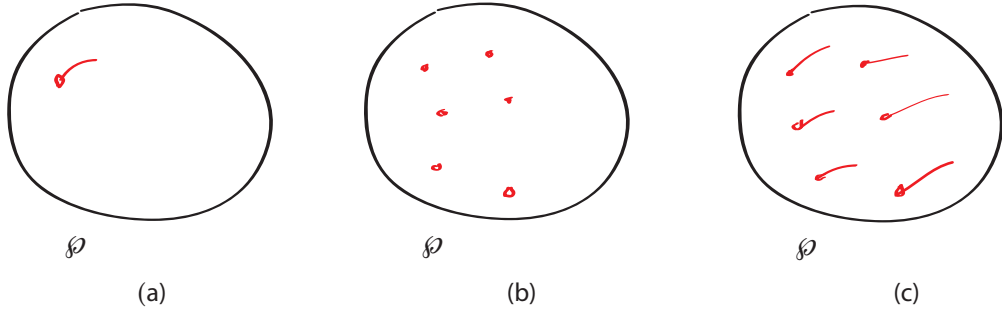


Figure 3: Systems in phase space \mathcal{P} : (a) a single system corresponds to a point, its time-evolution gives a curve in \mathcal{P} ; (b) a cloud of points defines an ensemble; gives the time evolution of an ensemble.

Instead, when a system is composed of N interacting molecules, the Gibbs formulation of statistical mechanics will require \mathcal{N} copies of the entire system of N molecules.

We conclude by noting that in a system composed of N *identical non-interacting constituents*, each constituent can be used as a system in its own right, the total system then being equivalent to an ensemble of $\mathcal{N} = N$ such non-interacting subsystems. This is why our earlier treatment of non-interacting molecules in an ideal gas produces the same result as will be gotten from a treatment by ensemble theory.

3.4 The density function

Consider a mechanical system parametrized by generalized coordinates (p_i, q_i) , $i = 1, \dots, s$ on phase space \mathcal{P} , whose dynamics is governed by a Hamiltonian $H(p, q; t)$. An ensemble of systems is then represented, at a fixed time t , by a *cloud* of \mathcal{N} points in \mathcal{P} . In the statistical approach, we shall consider \mathcal{N} extremely large, in such a way that the cloud of points may be well-approximated by a continuous distribution,

$$dw = \rho(p, q; t) dpdq \qquad dpdq = \prod_{i=1}^s dp_i dq_i \qquad (3.4)$$

where $dpdq$ stands for the canonical volume form on phase space. The interpretation is that dw counts the number of points of the ensemble contained within an infinitesimal volume $dpdq$ surrounding the point (p, q) at time t . Equivalently, the number of phase space points \mathcal{N}_D of the ensemble contained in a finite domain D , and the total number of points \mathcal{N} in the entire ensemble, are given by,

$$\mathcal{N}_D = \int_D dpdq \rho(p, q; t) \qquad \mathcal{N} = \int_{\mathcal{P}} dpdq \rho(p, q; t) \qquad (3.5)$$

The *ensemble average* $\bar{f}(t)$ of any function $f(p, q)$ on phase space is defined by

$$\bar{f}(t) = \frac{1}{\mathcal{N}} \int_{\mathcal{P}} dpdq \rho(p, q; t) f(p, q) \quad (3.6)$$

Since all ensemble averages are unchanged under multiplication of ρ by a positive constant, we will throughout use the *normalized distribution function* $\rho(p, q; t)/\mathcal{N}$ instead of ρ itself, and continue to use the notation $\rho(p, q; t)$ to signify the normalized distribution function. In the standard language of probability theory, the normalized distribution function is then a *probability distribution*.

The time-averaging procedure of (3.3) on a *single system* reaching equilibrium, may be extended to a time-averaging procedure on the entire ensemble. To see how this works, consider an infinitesimal volume of size $dpdq$ surrounding a point (p, q) in phase space. At time t , various individual points in the cloud of phase space points representing the ensemble will be inside the volume $dpdq$. This number is given by $\mathcal{N} dpdq \rho(p, q; t)$, for a normalized distribution function ρ . Thus, the time-average of (3.3) may be expressed as follows,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_t^{t+T} dt' f(p(t'), q(t')) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_t^{t+T} dt' \int_{\mathcal{P}} dpdq \rho(p, q; t') f(p, q) \quad (3.7)$$

All dependence on T and t is now contained in the t' -integral over the density function. Clearly, equilibrium will be attained for large enough T provided the combination

$$\frac{1}{T} \int_t^{t+T} dt' \rho(p, q; t') \quad (3.8)$$

tends towards a finite, t -independent limit as $T \rightarrow \infty$. This will be realized provided the following equilibrium condition holds at all times t ,

$$\frac{\partial \rho(p, q; t)}{\partial t} = 0 \quad (3.9)$$

Note that the general density function $\rho(p, q; t)$ introduced through the method of ensembles need not correspond to a system in equilibrium. The added advantage of the ensemble formulation is that we have a clear and clean characterization of equilibrium given in terms of the density function by condition (3.9).

3.5 The Liouville theorem

The Liouville theorem expresses the fact that the number of material points in a Hamiltonian system is conserved. This conservation applies to the phase space points describing the systems of an ensemble. We characterize an ensemble by a normalized density function

$\rho(p, q; t)$, so that the number of ensemble points inside an infinitesimal volume $dpdq$ of phase space is given by $\mathcal{N} dpdq \rho(p, q; t)$. The number of ensemble points \mathcal{N}_D contained inside a domain D (of dimension $2s$), is given by the integral over the density function as in (3.5),

$$\mathcal{N}_D(t) = \mathcal{N} \int_D dpdq \rho(p, q; t) \quad (3.10)$$

Since ρ depends explicitly on t , the number \mathcal{N}_D will in general depend on t as well. Conservation of the number of points is expressed by the conservation equation, which relates the density function ρ to the current density $(\rho\dot{p}_i, \rho\dot{q}_i)$ on all of phase space,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^s \left(\frac{\partial}{\partial p_i} (\rho\dot{p}_i) + \frac{\partial}{\partial q_i} (\rho\dot{q}_i) \right) = 0 \quad (3.11)$$

where the sum of partial derivatives gives the “divergence” of the current $(\rho\dot{p}_i, \rho\dot{q}_i)$. Using Hamilton’s equations to eliminate \dot{p}_i and \dot{q}_i , and using the relation,

$$\frac{\partial}{\partial p_i} \dot{p}_i + \frac{\partial}{\partial q_i} \dot{q}_i = -\frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} = 0 \quad (3.12)$$

we find,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^s \left(\frac{\partial \rho}{\partial p_i} \dot{p}_i + \frac{\partial \rho}{\partial q_i} \dot{q}_i \right) = 0 \quad (3.13)$$

This conservation equation may be recast in two equivalent ways. First, given that ρ is a function of p, q, t , we see that (3.13) expresses its total time-independence,

$$\frac{d\rho}{dt} = 0 \quad (3.14)$$

as the Hamiltonian system evolves in time according to Hamilton’s equations. This means that the density does not change in time, so that the ensemble evolves as an *incompressible fluid* does. Alternatively, (3.13) may be expressed with the help of Poisson brackets $\{\cdot, \cdot\}$. We normalize the generalized coordinates canonically by $\{p_i, p_j\} = \{q_i, q_j\} = 0$ and $\{q_i, p_j\} = \delta_{ij}$. Equations (3.13) and (3.14) are then equivalent to,

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0 \quad (3.15)$$

We stress that the above results hold for any distribution function, whether it describe an ensemble in equilibrium or out of equilibrium.

3.6 Equilibrium ensembles and distribution functions

The construction of equilibrium ensembles, and their associated distribution function, is guided by two key principles. The first is the condition for equilibrium in (3.9). The second is *Boltzmann's assumption of equal a priori probabilities*. We explain each these below.

If a distribution function ρ describes an ensemble in equilibrium, then by (3.9) and (3.15), it must satisfy the following conditions which are equivalent to one another,

$$\frac{\partial \rho}{\partial t} = 0 \qquad \{\rho, H\} = 0 \qquad (3.16)$$

Both relations provide a concrete characterization of equilibrium density functions, and thus of ensembles in equilibrium. We shall concentrate on conservative systems, in which case H has no explicit time dependence. The second equation in (3.16) then instructs us that ρ is a conserved quantity. Generically, a conservative Hamiltonian exhibits only a single conserved quantity, namely the total mechanical energy $E = H(p, q)$, so that ρ will be a function of the Hamiltonian only. When further conserved quantities, such as total momentum, total angular momentum, particle number, electric charge, exist, then ρ will also depend on the corresponding conserved mechanical functions.

The *principle of equal a priori probabilities* postulates that, if a certain macro-state has been completely specified by a certain number of thermodynamic variables, and is accessible by a number Ω of micro-states, then the probability for finding the system in any one of these micro-states is the same for all micro-states, and thus equal to $1/\Omega$. As this probability distribution does not favor any one micro-state above another, it is referred to as an *unbiased probability distribution*.

3.7 The uniform ensemble

In the uniform ensemble, $\rho(p, q)$ is taken to be constant, i.e. independent of p, q, t . This means that all systems in the ensemble have the same statistical probability, in accord with the Boltzmann principle of equal a priori probabilities. In view of (3.16), the uniform distribution function characterizes an equilibrium ensemble. Ensemble averages are obtained as integrals over phase space by,

$$\bar{f} = \int_{\mathcal{P}} dpdq f(p, q) / \int_{\mathcal{P}} dpdq \qquad (3.17)$$

It is readily seen that these expressions will be finite and make physical sense only if the volume of phase space is finite, for example when \mathcal{P} is compact. For most systems of physical interest, this will not be the case. Thus, in classical mechanics, this ensemble is not that useful. The uniform ensemble will be useful later, however, when considering certain quantum mechanical systems.

3.8 The micro-canonical ensemble

In the micro-canonical ensemble, the total (or internal) energy E is used to specify the macro-states completely. All micro-states with total energy E have equal weight, in accord with the Boltzmann principle of equal a priori probabilities. The motion of every system is restricted to the subspace \mathcal{P}_E of energy E of the full phase space \mathcal{P} , and the density function ρ_E for energy E is supported on \mathcal{P}_E with uniform weight. As a result, the normalized distribution function for the micro-canonical ensemble may be written down explicitly,

$$\rho_E(p, q) = \frac{1}{\mathcal{N}(E)} \delta(H(p, q) - E) \quad \mathcal{N}(E) = \int_{\mathcal{P}} dpdq \delta(H(p, q) - E) \quad (3.18)$$

The micro-canonical ensemble average of a phase space function $f(p, q)$ is given by,

$$\bar{f} = \int_{\mathcal{P}} dpdq \rho_E(p, q) f(p, q) \quad \int_{\mathcal{P}} dpdq \rho_E(p, q) = 1 \quad (3.19)$$

and the statistical entropy of the ensemble is given by,

$$S(E) = k \ln \Omega(E) \quad \Omega(E) = \mathcal{N}(E)/\mathcal{N}_0 \quad (3.20)$$

The need for the presence of a non-purely-mechanical normalization factor \mathcal{N}_0 in the entropy formula was already made clear by considering the simplest ideal gas, which showed that \mathcal{N}_0 has a quantum component involving \hbar . It is clear that this factor is needed also from the consideration of dimensions: the total number of micro-states $\Omega(E)$ is dimensionless, but $\mathcal{N}(E)$, defined in (3.18) has dimensions. Note that $\Omega(E)$ is defined in an intrinsic geometric manner, and invariant under canonical transformations on p, q .

In practice, it will often be more convenient, and more physical, to define the micro-canonical ensemble for a finite range of energies, instead of strictly at one value of energy. This treatment will be especially valuable in the quantum case, where energy levels are usually discrete. To do so, we define a shell of energies, in the range $[E, E + \Delta]$, with

$$\rho_{E,\Delta}(p, q) = \frac{1}{\Omega(E, \Delta)} \begin{cases} 1 & \text{if } E \leq H(p, q) \leq E + \Delta \\ 0 & \text{otherwise} \end{cases} \quad (3.21)$$

The ensemble averages are defined accordingly as in (3.19) with ρ_E replaced with $\rho_{E,\Delta}$.

3.9 The canonical ensemble

The canonical ensemble is the most standard one. As a result, a number of different derivations are available, and will be discussed here.

The canonical ensemble may be obtained from the micro-canonical ensemble by changing variables from fixed internal energy E to fixed absolute temperature T . The density function for the canonical ensemble $\rho(p, q; T)$ is the Laplace transform of the density function $\rho_E(p, q)$ for the micro-canonical ensemble, and is given by,

$$\rho(p, q; T) = \frac{1}{Z(T)} e^{-H(p, q)/kT} \quad (3.22)$$

The normalization factor $Z(T)$ is referred to as the *partition function*, and is given by,

$$Z(T) = \int_{\mathcal{P}} dpdq e^{-H(p, q)/kT} \quad (3.23)$$

Statistical averages in the canonical ensemble are obtained by,

$$\bar{f} = \frac{1}{Z(T)} \int_{\mathcal{P}} dpdq e^{-H(p, q)/kT} f(p, q) \quad (3.24)$$

The canonical ensemble is more useful in carrying out calculations than the micro-canonical ensemble. The three formulas above are fundamental throughout statistical mechanics.

To make contact with other familiar thermodynamic functions, we proceed as follows. By definition, the ensemble average of the Hamiltonian $H(p, q)$ gives the internal energy E . It may be expressed via a derivative of Z with respect to T ,

$$E = \bar{H} = kT^2 \frac{\partial}{\partial T} \ln Z(T) \quad (3.25)$$

To compare this result with formulas in thermodynamics, we express it in terms of the free energy F , with the help of $F = E - TS$ and $S = -\partial F/\partial T$ to obtain,

$$E = -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T} \right) \quad (3.26)$$

Comparison of (3.6) and (3.5) shows that the free energy is given by,

$$F(T) = -kT \ln \frac{Z(T)}{Z_0} \quad (3.27)$$

which is one of the most fundamental formulas of correspondence between statistical mechanics and thermodynamics, in the canonical ensemble. The constant Z_0 is not determined by classical statistical mechanics. Its presence shifts the entropy by a constant. We know from the example of the simplest ideal gas that Z_0 is quantum mechanical in origin. More importantly, one needs the constant there to make the argument of the logarithm to be properly dimensionless. Finally, we have already established in the ideal gas calculation

that the quantum indistinguishability of identical particles requires the presence of an extra combinatorial factor $N!$. Thus, the correct constant is found to be,

$$Z_0 = (2\pi\hbar)^s N! \quad (3.28)$$

This result may be shown in all generality by first using the functional integral formulation of the quantum problem and then taking its classical limit.

An alternative derivation makes use of the fact that the normalized density function is a probability distribution. As appropriate for the canonical ensemble, we assume that the density function only depends on energy, namely on the Hamiltonian of the system. Now consider two independent systems (each system having its own phase space), with Hamiltonians H_1, H_2 , and corresponding density functions $\rho(H_1)$ and $\rho(H_2)$. The Hamiltonian of the total system is then $H = H_1 + H_2$. The probability distributions $\rho(H_1)$ and $\rho(H_2)$ are probabilistically independent. Therefore, as always with probability, the probability distribution of the combined system $\rho(H)$ must be the product,

$$\rho(H) = \rho(H_1) \cdot \rho(H_2) \quad (3.29)$$

But since we have $H = H_1 + H_2$, this implies that the dependence of $\rho(H)$ on H must be exponential. This gives the Canonical distribution with the Boltzmann factor.

3.10 Deriving the canonical from the micro-canonical ensemble

The constructing of the preceding section is somewhat formal. Here, we present another derivation of the distribution function (3.22) which provides more physical insight. Keeping a system (denoted here by \mathcal{S}) in equilibrium at a fixed temperature T may be realized physically by putting the system \mathcal{S} in thermal contact with a very large system \mathcal{S}' , which is assumed to be in equilibrium at temperature T . The system \mathcal{S}' provides a *heat bath* or *thermal reservoir* with which \mathcal{S} can freely exchange energy. A pictorial representation of the system is given in Figure 4. We keep the total energy E_{tot} of the combined system fixed, so that the set-up used early on for the micro-canonical ensemble applies. We shall assume here that the micro-states of the system \mathcal{S} are discrete, may be labelled by a discrete index n , and have discrete energy levels E_n .

The total number $\Omega_{\text{tot}}(E_{\text{tot}})$ of micro-states to which the combined system has access is then given by,

$$\Omega_{\text{tot}}(E_{\text{tot}}) = \sum_E \Omega(E)\Omega'(E') \quad (3.30)$$

where E and E' are the energies of the systems \mathcal{S} and \mathcal{S}' respectively, and the total energy is given by $E_{\text{tot}} = E + E'$. Labeling all the micro-states of \mathcal{S} by the discrete label n , we have

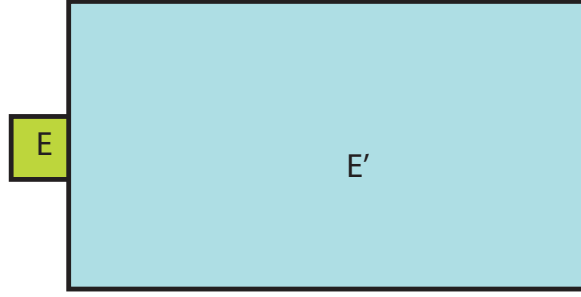


Figure 4: A small macro-systems \mathcal{S} in thermal equilibrium with a large heat bath \mathcal{S}' .

the following explicit formula for $\Omega(E)$,

$$\Omega(E) = \sum_n \delta_{E, E_n} \quad (3.31)$$

where δ_{E, E_n} is the Kronecker delta. It is important to stress that the sum over n is to run over all distinct states, not just over all allowed energies; in particular, there may be several states of \mathcal{S} with energy E_n , and the degeneracy of the energy level E_n must be properly included in the sum. Using (3.31) and (3.31), formula (3.30) may be cast in terms of a sum over all micro-states of \mathcal{S} ,

$$\Omega_{\text{tot}}(E_{\text{tot}}) = \sum_n \Omega'(E_{\text{tot}} - E_n) = \sum_n \exp \left\{ \frac{1}{k} S'(E_{\text{tot}} - E_n) \right\} \quad (3.32)$$

Now, we use the fact that the heat bath system \mathcal{S} is huge compared with \mathcal{S} , so that we can assume $E_n/E_{\text{tot}} \ll 1$, and expand in this quantity,

$$S'(E_{\text{tot}} - E_n) = S'(E_{\text{tot}}) - E_n \frac{\partial S'(E_{\text{tot}})}{\partial E_{\text{tot}}} + \frac{1}{2} E_n^2 \frac{\partial^2 S'(E_{\text{tot}})}{\partial E_{\text{tot}}^2} + \mathcal{O}(E_n^3) \quad (3.33)$$

Now, by definition, the first derivative,

$$\frac{\partial S'(E_{\text{tot}})}{\partial E_{\text{tot}}} = \frac{1}{T} \quad (3.34)$$

gives the temperature T of the heat bath. Since $S'(E_{\text{tot}})$ is extensive, the second derivative term is of order E_n/E_{tot} and may be neglected compared to the first derivative term. Thus, the total number of states becomes,

$$\Omega_{\text{tot}}(E_{\text{tot}}) = e^{S'(E_{\text{tot}})/k} \sum_n e^{-E_n/kT} \quad (3.35)$$

Using now the principle of equal a priori probabilities, we know that every state n with energy E_n is equally likely, so we find the number of states of the total system accessible to micro-state n to be given by

$$e^{S'(E_{\text{tot}})/k} e^{-E_n/kT} \quad (3.36)$$

As a result, the probability of finding the whole system in a state such that \mathcal{S} is in the micro-state n is given by the ratio,

$$\rho_n = \frac{e^{-E_n/kT}}{Z(T)} \quad Z(T) = \sum_n e^{-E_n/kT} \quad (3.37)$$

Note that in these probabilities, all references to the heat bath has vanishes, except for its temperature. In classical mechanics, the possible micro-states are labelled by the generalized coordinates (p, q) , the corresponding energy being given by the Hamiltonian $H(p, q)$. The probability becomes the normalized density function, so that we recover (3.22).

3.11 The Gibbs and the grand-canonical ensembles

The Gibbs ensemble is a generalization of the canonical ensemble in which time-independent external generalized forces may applied to the system, and work may be exchanged between systems in contact. Each type of generalized force f has a conjugate generalized displacement x , together forming *thermodynamic paired variables*. The variable x is a function of phase space. Most frequently x is of the form $x = x(p, q)$, though when x is the number of particles N , or the volume V of space, the functional dependence is somewhat unusual. Classic examples of *thermodynamic paired variables* are listed in Table 1.

System	generalized force	f	generalized displacement	x, X
Fluid	pressure	$-P$	volume	V
Film	surface tension	σ	area	A
String	string tension	κ	length	L
Magnetic	magnetic field	\mathbf{B}	magnetization	\mathbf{M}
Electric	electric field	\mathbf{E}	charge	q
			polarization	\mathbf{P}
chemical	chemical potential	μ	particle number	N

Table 1: Paired thermodynamic variables

We shall assume that the generalized forces f on the system are kept fixed in time, as is required for an equilibrium ensemble. The work done *on the system* by a generalized

displacement x is then linear in x , and given by $W = fx$. The corresponding change in the Hamiltonian is given by $H(p, q) \rightarrow H(p, q) - fx(p, q)$. (If several generalized forces f_α are applied, we will instead have a sum $\sum_\alpha f_\alpha x_\alpha(p, q)$.) The distribution function is given by,

$$\rho(p, q; T, f) = \frac{1}{Z(T, f)} e^{-(H(p, q) - fx(p, q))/kT} \quad (3.38)$$

The normalization factor $Z(T, f)$ is referred to as the *grand canonical partition function*,

$$Z(T, f) = \int_{\mathcal{P}} dpdq e^{-(H(p, q) - fx(p, q))/kT} \quad (3.39)$$

Statistical averages in the canonical ensemble are obtained by,

$$\bar{f} = \frac{1}{Z(T, f)} \int_{\mathcal{P}} dpdq e^{-(H(p, q) - fx)kT} f(p, q) \quad (3.40)$$

In particular, the ensemble average of the generalized displacement variable x will be denoted by $X = \bar{x}$. The Gibbs free energy is defined by,

$$G(T, f) = E - TS - fX \quad (3.41)$$

and is related to the partition function by,

$$G(T, f) = -kT \ln Z(T, f) \quad (3.42)$$

When the exchange of particles is permitted, the corresponding Gibbs ensemble is sometimes referred to as the *grand-canonical ensemble*.

4 Applications of Classical Statistical Ensembles

4.1 Application I: The Maxwell distribution

The density function in the canonical ensemble may be readily used to derive the Maxwell distribution for the momenta or velocities in a classical ideal gas. The particles being non-interacting, we can focus on the velocity distribution of a single particle, the case of N particles being given by multiplying the probabilities and partition functions for single-particles. The single-particle partition function Z_1 is given by,

$$Z_1(T) = \int \frac{d^3\mathbf{p} d^3\mathbf{q}}{(2\pi\hbar)^3} e^{-H(p,q)/kT} \quad (4.1)$$

The Hamiltonian is that for a free non-relativistic particle $H(p, q) = \mathbf{p}^2/2m$ with no dependence on \mathbf{q} . Thus, the \mathbf{q} -integral may be carried out and yields the space-volume V . Changing variables from momentum to velocity, $\mathbf{p} = m\mathbf{v}$, we have,

$$Z_1(T) = \frac{m^3V}{(2\pi\hbar)^3} \int d^3\mathbf{v} e^{-m\mathbf{v}^2/2kT} \quad (4.2)$$

It is straightforward to evaluate the integral, and we find,

$$Z_1(T) = \frac{m^3V}{(2\pi\hbar)^3} \times \left(\frac{2\pi kT}{m} \right)^{3/2} \quad (4.3)$$

As a result, the probability distribution for one particle is given by,

$$\rho_{\text{Max}}(v) = \left(\frac{m}{2\pi kT} \right)^{3/2} 4\pi v^2 e^{-mv^2/2kT} \quad \int_0^\infty dv \rho_{\text{Max}}(v) = 1 \quad (4.4)$$

This is the Maxwell distribution. It is readily checked that the average velocity $\langle \mathbf{v} \rangle$ vanishes, while the average velocity square, and thus the average energy, are given by,

$$\langle \mathbf{v}^2 \rangle = \int_0^\infty dv \rho_{\text{Max}}(v) v^2 = \frac{3kT}{m} \quad \langle H \rangle = \frac{3}{2}kT \quad (4.5)$$

as expected from the equipartition theorem.

4.2 Application II: Magnetism in classical statistical mechanics

Magnetic fields couple to matter in two ways; first, to the currents of moving electric charges; second to intrinsic magnetic dipole moments. The first effect is diamagnetic, the second paramagnetic. The intrinsic magnetic moments of elementary particles is a purely quantum

effect, at the same level as their spin. The magnetic coupling to electric charge is governed by the minimally coupled Hamiltonian for charged particles,

$$H(p, q) = \sum_{i=1}^N \frac{1}{2m_i} (\mathbf{p}_i - e_i \mathbf{A}_i(q))^2 + U(q) \quad (4.6)$$

Here, $U(q)$ is any potential contribution which depends only on the generalized positions q ; $\mathbf{A}_i(q)$ is the vector potential acting on particle i ; and e_i, m_i are respectively the electric charge and mass of particle i . Note that the vector potential may include the effects of an external magnetic field, as well as the effects of internal magnetic interactions. We begin by computing the partition function,

$$Z(T) = \frac{1}{(2\pi\hbar)^{3N} N!} \int \prod_{i=1}^N d^3 \mathbf{p}_i d^3 \mathbf{q}_i e^{-H(p,q)/kT} \quad (4.7)$$

It may be decomposed as follows,

$$Z(T) = \frac{1}{(2\pi\hbar)^{3N} N!} \int \prod_{i=1}^N d^3 \mathbf{q}_i e^{-U(q)/kT} \prod_{i=1}^N \int d^3 \mathbf{p}_i \exp \left\{ -\frac{(\mathbf{p}_i - e_i \mathbf{A}_i(q))^2}{2m_i kT} \right\} \quad (4.8)$$

Each integration measure $d^3 \mathbf{p}_i$ is invariant under translation by an arbitrary vector, which may depend on i . In particular, it is invariant under the shifts,

$$\mathbf{p}_i \rightarrow \mathbf{p}'_i = \mathbf{p}_i - e_i \mathbf{A}_i(q) \quad (4.9)$$

Performing these changes of variables, we obtain,

$$Z(T) = \frac{1}{(2\pi\hbar)^{3N} N!} \int \prod_{i=1}^N d^3 \mathbf{q}_i e^{-U(q)/kT} \prod_{i=1}^N \int d^3 \mathbf{p}_i \exp \left\{ -\frac{\mathbf{p}_i^2}{2m_i kT} \right\} \quad (4.10)$$

All magnetic effects have vanished from the partition function. This result is referred to as the *Bohr - Van Leeuwen theorem*. In particular, all thermodynamic functions are independent of external magnetic fields. Thus, there is no diamagnetism in classical statistical mechanics. In as far as intrinsic magnetic moments are purely quantum effects, there are strictly speaking then also no paramagnetic effects classically.

4.3 Application III: Diatomic ideal gasses (classical)

The simplest ideal gas model only included translational degrees of freedom, as would be suitable for point-like microscopic constituents. Single atoms can of course also have excited

quantum states, but these start to be relevant only at relatively high temperatures. For example, the first excited state of Hydrogen is 13.6 eV above the ground state, corresponding to a temperature of 10^5 K . Polyatomic molecules do, however, have low lying extra degrees of freedom, such as rotational and vibrational, which are relevant at room temperatures.

Consider a simple model of an ideal gas composed of diatomic molecules. The atoms in each molecule can rotate in the two directions perpendicular to the symmetry axis of the molecule with moment of inertia I , while we will neglect the moment of inertia of rotations about the symmetry axis. The atoms in each molecule can also vibrate by changing their relative distance, and we shall model these vibrations by a harmonic oscillator of frequency ω . The translational, rotational and vibrational degrees of freedom are all decoupled from one another, so that their Hamiltonians add up, and their phase space measures factorize. Thus, the partition function of an individual molecule (recall that the gas is ideal, so the total partition function is given by the N -th power and a prefactor of $1/N!$) factorizes,

$$Z_1(T) = Z_{\text{tranl}} Z_{\text{rot}} Z_{\text{vib}} \quad (4.11)$$

The factor Z_{tranl} was already computed, and is given by (4.3).

The rotational degrees of freedom may be parametrized by the angles on the sphere (θ, ϕ) , and the associated Lagrangian is given by,

$$L = \frac{I}{2}(\dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta) \quad (4.12)$$

The conjugate momenta are $p_\theta = I\dot{\theta}$ and $p_\phi = I\dot{\phi} \sin^2 \theta$, and the Hamiltonian is given by,

$$H_{\text{rot}} = \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I \sin^2 \theta} \quad (4.13)$$

Carefully including the measure with the standard normalization, we find,

$$Z_{\text{rot}} = \int \frac{d\theta d\phi dp_\theta dp_\phi}{(2\pi\hbar)^2} e^{-\beta H_{\text{rot}}} = \frac{2IkT}{\hbar^2} \quad (4.14)$$

Using (3.16), we derive the contribution from the rotational energy per molecule,

$$E_{\text{rot}} = kT \quad (4.15)$$

in accord with the equipartition theorem for 2 rotational degrees of freedom.

Finally, there is a single vibrational degrees of freedom between the positions of the two atoms, which we will denote by x , and whose Hamiltonian is given by

$$H_{\text{vib}} = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2 \quad (4.16)$$

from which we deduce,

$$Z_{\text{vib}} = \frac{kT}{\hbar\omega} \qquad E_{\text{vib}} = kT \qquad (4.17)$$

The factor of 2 in this contribution compared with the naive expectation from equipartition is due to the presence of the potential energy. Adding all up, we find the internal energy of the gas as follows,

$$E = N(E_{\text{transl}} + E_{\text{rot}} + E_{\text{vib}}) = \frac{7}{2}NkT \qquad (4.18)$$

In particular, the specific heat is found to be

$$C_V = \frac{7}{2}Nk \qquad (4.19)$$

This prediction is very simple and quite distinct from the one for mono-atomic gasses. So, one immediate question is: does (4.19) agree with experiment? For most gasses at room temperature, the answer is no; instead their specific heat is more in line with the predictions of a mono-atomic gas. The explanation lies, even at room temperature, in quantum effects, which we shall study in the next section.

4.4 Quantum vibrational modes

Consider the vibrational effects first. Classically, the oscillator can carry any small amount of energy above that of its rest position. But this is not so in quantum mechanics, since the energy levels of the oscillator are discrete,

$$\varepsilon_n = \left(\frac{1}{2} + n\right) \hbar\omega \qquad n = 0, 1, 2, 3, \dots \qquad (4.20)$$

Thus, the minimal amount of energy (or quantum) above the ground state $\hbar\omega$ is discrete. This introduces a characteristic temperature scale in the problem,

$$T_{\text{vib}} = \frac{\hbar\omega}{k} \qquad (4.21)$$

As a consequence, we expect that at temperatures below T_{vib} , the vibrational mode will be strongly suppressed by the Boltzmann factor, and will not contribute to the specific heat.

We shall now work out the quantum contribution from the vibrational modes quantitatively. The partition function is given by,

$$Z_{\text{vib}}(\beta) = \sum_{n=0}^{\infty} e^{-\beta\varepsilon_n} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \qquad (4.22)$$

The free energy and internal energy for the vibrational mode are given by,

$$\begin{aligned} F_{\text{vib}} &= -kT \ln Z_{\text{vib}} = \frac{1}{2} \hbar \omega + kT \ln (1 - e^{-\beta \hbar \omega}) \\ E_{\text{vib}} &= -\frac{\partial}{\partial \beta} \ln Z_{\text{vib}} = \hbar \omega \left(\frac{1}{2} + \frac{1}{e^{\beta \hbar \omega} - 1} \right) \end{aligned} \quad (4.23)$$

The specific heat is then readily evaluated, as we find,

$$C_{\text{vib}} = \frac{\partial E_{\text{vib}}}{\partial T} = \frac{T_{\text{vib}}^2}{T^2} \frac{e^{T_{\text{vib}}/T}}{(e^{T_{\text{vib}}/T} - 1)^2} \quad (4.24)$$

The contribution to the total specific heat from the N vibrational modes is simply $C_V = NC_{\text{vib}}$. The behavior of this function may be inferred from its asymptotic behavior,

$$\begin{aligned} T \gg T_{\text{vib}} & \quad C_V \approx kN \\ T \ll T_{\text{vib}} & \quad C_V \approx kN \frac{T_{\text{vib}}^2}{T^2} e^{-T_{\text{vib}}/T} \end{aligned} \quad (4.25)$$

which is clearly being reflected in a numerical plot of the function given in Figure 5. The large T contribution corresponds to the classical behavior found in the previous section. Below T_{vib} , the vibrational contribution to the specific heat turns off quickly.

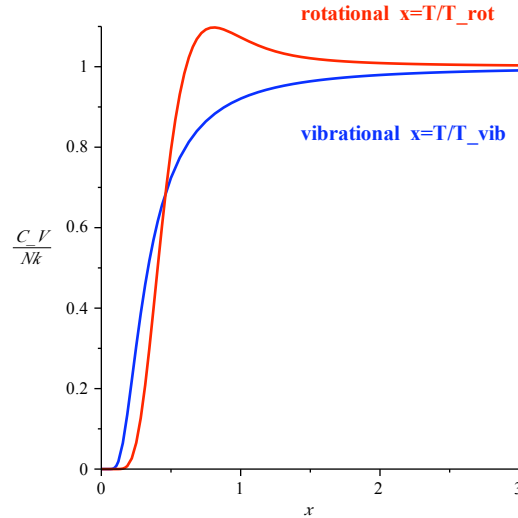


Figure 5: Contribution to the specific heat from a quantum vibrational mode.

4.5 Quantum rotational modes

A similar story applies to the rotational modes. Their contribution to the specific heat will be cut off at a characteristic temperature T_{rot} as well since the energy levels of the quantum rotor are again discrete. The energy levels of a (simplified) spherical model are given by $\varepsilon_\ell = \ell(\ell + 1)\hbar^2/2I$ with degeneracy $2\ell + 1$, so that the partition function is given by,

$$Z_{\text{rot}} = \sum_{\ell=0}^{\infty} (2\ell + 1) e^{-\ell(\ell+1)T_{\text{rot}}/T} \quad T_{\text{rot}} = \frac{\hbar^2}{2Ik} \quad (4.26)$$

Analytically, this sum would be given by an elliptic function, but we can easily plot the function numerically, as shown in Figure 5. It is useful to get some order of magnitude estimated for the transition temperatures. For molecular Hydrogen, H_2 , we have,

$$\begin{aligned} T_{\text{vib}} &\approx 1500K & \omega &\approx 2 \times 10^{14} \times s^{-1} \\ T_{\text{rot}} &\approx 100 K & I &\approx 0.46 \times 10^{-40} \text{ erg} \times s^2 \end{aligned} \quad (4.27)$$

These order of magnitudes demonstrate that quantum effects are relevant at relatively low as well as at relatively high temperatures.

5 Quantum Statistical Ensembles

In this section, we shall define ensembles in quantum mechanics, distinguish between pure and mixed ensembles, and introduce the density matrix. We have set up the classical case in such a way that it will resemble most closely the treatment of the quantum case. The uniform, micro-canonical and canonical ensembles are then introduced for the quantum case.

5.1 Quantum Mechanics

In quantum mechanics, physical states are represented by vectors in a Hilbert space \mathcal{H} , and denoted by bras and kets following Dirac. Two vectors $|\psi\rangle$ and $|\psi'\rangle$ correspond to the same physical state if and only if there exists a non-zero complex number λ such that $|\psi'\rangle = \lambda|\psi\rangle$. Observables are represented by self-adjoint linear operators on \mathcal{H} . A state $|\phi_i\rangle$ has a definite measured value a_i for an observable A provided $|\phi_i\rangle$ is an eigenstate of A with eigenvalue a_i . Let $|\psi\rangle$ be an arbitrary state in \mathcal{H} , and let $\{|\varphi_i\rangle\}$ denote a set of orthonormal states. Then, the probability P for measuring the state $|\psi\rangle$ to be in one of the states $|\varphi_i\rangle$ is given by,

$$P = |\langle\varphi_i|\psi\rangle|^2 \qquad \langle\psi|\psi\rangle = 1 \qquad (5.1)$$

Finally, in the Schrödinger picture of quantum mechanics, observables are time-independent and states $|\psi(t)\rangle$ evolve in time according to the Schrödinger equation associated with a Hamiltonian H , which is itself an observable,

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \qquad (5.2)$$

Although these statements certainly provide the basic principles of quantum mechanics, not all physical states of systems encountered in Nature can be properly described by vectors in a Hilbert space. To make the distinction, one refers to a state described by a vector in Hilbert space as a *pure state*. Instances where the vector description is not adequate include,

1. Incoherent mixtures of pure states;
2. Subsystems of a system in a pure quantum state.

Understanding both of these cases will turn out to be of fundamental importance in quantum statistical mechanics, and we shall begin by discussing them in some detail.

5.2 Mixed quantum states

The interference properties of quantum mechanics may be illustrated using the physical examples of polarized beams of photons or of spin 1/2 particles. In either set-up, the first

piece of apparatus used is usually a polarizer, responsible for filtering a definite polarization out of an unpolarized beam. Particles in a polarized beam are pure quantum states. In particular, a polarized beam of Silver atoms in the Stern-Gerlach experiment should be viewed as an ensemble of spins which are all in the same quantum state. This states in the beam are said to be *coherent* and to form a *pure ensemble*.

What we have not yet provided in quantum mechanics is a mathematical description for a beam which is unpolarized, or is said to be in a *mixed state*. The defining property of an *unpolarized beam of spin 1/2 particles* is that measurement of the spin operator $\mathbf{n} \cdot \mathbf{S}$ along any direction \mathbf{n} , with $\mathbf{n}^2 = 1$, gives $\pm\hbar/2$ with equal probabilities, namely 50% each.

We begin by showing that an unpolarized beam *cannot* be described mathematically by a vector in Hilbert space. Let us assume the contrary, and represent the particles in the unpolarized beam by a vector $|\psi\rangle$. The probability $P_{\mathbf{n}}$ for observing the spins in a quantum state $|\mathbf{n}, +\rangle$, which is the eigenstate of $\mathbf{n} \cdot \mathbf{S}$ with eigenvalue $+\hbar/2$, would then be given by, $P_{\mathbf{n}} = |\langle \mathbf{n}, + | \psi \rangle|^2$. To compute this probability, we express the normalized states $|\mathbf{n}, +\rangle$ and $|\psi\rangle$ in a fixed orthonormal basis $|\pm\rangle$ of the Hilbert space, parametrized as follows,

$$\begin{aligned} |\mathbf{n}, +\rangle &= \cos\theta e^{i\psi}|+\rangle + \sin\theta e^{-i\psi}|-\rangle \\ |\psi\rangle &= \cos\alpha e^{i\beta}|+\rangle + \sin\alpha e^{-i\beta}|-\rangle \end{aligned} \quad (5.3)$$

The probability $P_{\mathbf{n}}$ is then given by,

$$P_{\mathbf{n}} = \cos^2\theta \cos^2\alpha + \sin^2\theta \sin^2\alpha + 2\cos\theta \sin\theta \cos\alpha \sin\alpha \cos(2\beta - 2\psi) \quad (5.4)$$

For an unpolarized beam, $P_{\mathbf{n}}$ must be independent of both θ and ψ . Independence of ψ requires $\sin 2\alpha = 0$ (since θ is to be arbitrary). Its solutions $\alpha = 0, \pi/2$ respectively give $P_{\mathbf{n}} = \cos^2\theta$ and $P_{\mathbf{n}} = \sin^2\theta$, neither of which can be independent of θ . We conclude that an unpolarized beam cannot be described mathematically by a state in Hilbert space.

5.3 Subsystems of a quantum system in a pure state

Consider a quantum system whose Hilbert space \mathcal{H} is the tensor product of two Hilbert spaces \mathcal{H}_A and \mathcal{H}_B . A simple example is provided by A and B being both two-state systems. Here, \mathcal{H}_A and \mathcal{H}_B may be generated by the ortho-normal states $|\pm, A\rangle$ and $|\pm, B\rangle$, in which the operators S_A^z and S_B^z are diagonal and take eigenvalues $\pm\hbar/2$. Consider a pure state $|\psi\rangle$ of the system \mathcal{H} ; for example, the singlet state is given by,

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+, A\rangle \otimes |-, B\rangle - |-, A\rangle \otimes |+, B\rangle) \quad (5.5)$$

Now suppose that we measure observables in the subsystem A only, such as for example S_A^z . As such, this question does not quite make sense yet, because S_A^z is an operator in \mathcal{H}_A , but

not in \mathcal{H} . Since we make no observations on system B , the only natural way to extend S_A^z to an operator of \mathcal{H} is by using the identity matrix I_B in \mathcal{H}_B ,

$$S_A^z \rightarrow \hat{S}_A^z = S_A^z \otimes I_B \quad (5.6)$$

The expectation value of \hat{S}_A^z in the pure quantum state $|\psi\rangle$ is then given as follows,

$$\langle\psi|\hat{S}_A^z|\psi\rangle = \frac{1}{2}(\langle+, A|S_A^z|+, A\rangle + \langle-, A|S_A^z|-, A\rangle) = 0 \quad (5.7)$$

using the orthonormality of the states $|\pm, B\rangle$. In fact, by the same reasoning, one shows that we also have $\langle\psi|\mathbf{n} \cdot \mathbf{S}|\psi\rangle = 0$ for any \mathbf{n} .

Remarkably, we have established that observing a subsystem of a total system which is in a pure quantum state produces quantum expectation values which are identical to those of a mixture of pure quantum states. Thus, the problems of representing mixtures of pure quantum states and the problem of observing only a subsystem are closely related to one another, and neither of them permits a description by vectors in Hilbert space alone.

5.4 The density matrix

A new mathematical tool is needed to describe quantum systems in a mixture, or observed only by the observables of a subsystem. This formalism was introduced by Landau and von Neumann in 1927, and the key object is referred to as the *density matrix*.

A *pure ensemble*, by definition, is a collection of physical systems such that every member of the ensemble is characterized by the same element $|\psi\rangle$ in Hilbert space.

A *mixed ensemble*, by definition, is a collection of physical systems such that a fraction w_1 of the members is characterized by a pure state $|\psi_1\rangle$, a fraction w_2 of the members is characterized by a pure state $|\psi_2\rangle$, and so on. We shall assume that each pure state ket $|\psi_i\rangle$ is normalized, but the vectors $|\psi_1\rangle, |\psi_2\rangle, \dots$ do not need to be mutually orthogonal to one another. Thus, a mixed ensemble is characterized by N vectors $|\psi_i\rangle$ for $i = 1, \dots, N$, each one representing a pure state, and entering into the mixed state with a *population fraction* $w_i \geq 0$, which quantitatively indicates the proportion of state $|\psi_i\rangle$ in the mixture. The population fractions are normalized by,

$$\sum_{i=1}^N w_i = 1 \quad (5.8)$$

A mixture is an *incoherent superposition of pure states* which means that all relative phase information of the pure states must be lost in the mixture. This is achieved by superimposing, not the states $|\psi_i\rangle$ in Hilbert space, but rather the projection operators $|\psi_i\rangle\langle\psi_i|$ associated

with each pure state. The *density matrix* ρ for an ensemble of N pure states $|\psi_i\rangle$ incoherently superimposed with population fractions w_i is defined by

$$\rho = \sum_{i=1}^N |\psi_i\rangle w_i \langle\psi_i| \quad (5.9)$$

Since the superpositions involved here are incoherent, the weights w_i may be thought of as classical probability weights assigned to each population. We stress that *the various different pure states $|\psi_i\rangle$ are not required to be orthogonal to one another*, since it should certainly be possible to superimpose pure states which are not orthogonal to one another.

The following properties of the density matrix immediately result,

1. Self-adjointness : $\rho^\dagger = \rho$;
2. Unit trace : $\text{Tr}(\rho) = 1$;
3. Non-negative : $\langle\psi|\rho|\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$;
4. A state corresponding to ρ is pure (rank 1) if and only if ρ is a projection operator.

Conversely, any operator ρ in \mathcal{H} satisfying properties 1, 2, 3 above is of the form of (5.9) with normalization (5.8). The above three conditions guarantee that ρ can be diagonalized in an orthonormal basis $|\phi_i\rangle$, with real non-negative eigenvalues p_i ,

$$\rho = \sum_i |\phi_i\rangle p_i \langle\phi_i| \quad \sum_i p_i = 1 \quad (5.10)$$

Note that $|\phi_i\rangle$ need not be proportional to $|\psi_i\rangle$, and p_i need not coincide with w_i . Thus, a given density matrix will have several equivalent representations in terms of pure states. To check property 4, we note that if ρ corresponds to a pure state, it is rank 1, and thus a projection operator in view of property 2. Conversely, substituting the expression (5.10) into the projection operator relation $\rho^2 = \rho$ requires $p_i(p_i - 1) = 0$ for all i . The condition is solved by either $p_i = 0$ or $p_i = 1$, but $\text{Tr}(\rho) = 1$ requires that p_i equal 1 for a single i .

Using the above properties of density matrices, we construct the general density matrix for a 2-state system for which the Hilbert space is 2-dimensional. The density matrix is then a 2×2 matrix and may be expressed in terms of the Pauli matrices $\sigma = (\sigma^1, \sigma^2, \sigma^3)$,

$$\rho = (I + \mathbf{a} \cdot \sigma)/2 \quad (5.11)$$

where \mathbf{a} is a real 3-vector. Self-adjointness and normalization are guaranteed by (5.11), while positivity requires $|\mathbf{a}| \leq 1$. Pure states precisely correspond to the boundary $|\mathbf{a}| = 1$, while any density matrix with $|\mathbf{a}| < 1$ corresponds to a mixed state which is not pure. The ensemble average of the spin operator $\mathbf{n} \cdot \mathbf{S}$ in the direction \mathbf{n} is readily calculated,

$$\langle\mathbf{n} \cdot \mathbf{S}\rangle = \text{Tr}(\rho \mathbf{n} \cdot \mathbf{S}) = \mathbf{n} \cdot \mathbf{a} \frac{\hbar}{2} \quad (5.12)$$

The above formula confirms that the state ρ is unpolarized for $\mathbf{a} = 0$, partially polarized for $0 < |\mathbf{a}| < 1$, and pure for $|\mathbf{a}| = 1$. This gives a nice geometrical representation of the space of density matrices for the two-state system.

5.5 Ensemble averages and time-evolution

In a pure normalized state $|\psi\rangle$, we defined the *expectation value of an observable* A by the matrix element $\langle\psi|A|\psi\rangle = \text{Tr}(A|\psi\rangle\langle\psi|)$, a quantity that gives the quantum mechanical weighed probability average of the eigenvalues of A . In a mixed state, these quantum mechanical expectation values must be further weighed by the population fraction of each pure state in the mixture. One denotes this double quantum and statistical average by either \bar{A} , or by $\langle A \rangle$, and one defines the *ensemble average* of the observable A by

$$\bar{A} = \langle A \rangle = \sum_i w_i \langle\psi_i|A|\psi_i\rangle = \text{Tr}(\rho A) \quad (5.13)$$

It follows from the properties of the density matrix that the ensemble average of any self-adjoint operator is real. Note that formula (5.13) holds whether ρ corresponds to a pure state or a mixture.

The time-evolution of the density matrix may be derived from the Schrödinger equation for pure states. In the Schrödinger picture, the time evolution of a quantum state $|\psi(t)\rangle$ is governed by a self-adjoint Hamiltonian operator H on \mathcal{H} via the Schrödinger equation (5.2). Assuming that the population fractions w_i of (5.9), or equivalently the p_i of (5.10), do not change in time during this evolution, the density matrix $\rho(t)$ will obey the following time evolution equation,

$$i\hbar \frac{d}{dt} \rho(t) = [H, \rho(t)] \quad (5.14)$$

To prove this formula, one first derives the time evolution equation for $|\psi(t)\rangle\langle\psi(t)|$ and then takes the weighed average with the time-indepenent population fractions. This time evolution may be solved for in terms of the unitary evolution operator $U(t) = e^{-itH/\hbar}$, by

$$\rho(t) = U(t)\rho(0)U(t)^\dagger \quad (5.15)$$

The normalization condition $\text{Tr}(\rho) = 1$ is automatically preserved under time evolution.

5.6 The density matrix for a subsystem of a pure state

Next, we return to the problem of observing a subsystem with Hilbert space \mathcal{H}_A of a total system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $|\alpha_i, A\rangle$ be an orthonormal basis in \mathcal{H}_A , and

$|\beta_j, B\rangle$ an orthonormal basis in \mathcal{H}_B with $i = 1, \dots, \dim \mathcal{H}_A$ and $j = 1, \dots, \dim \mathcal{H}_B$. The states $|\alpha_i, A\rangle \otimes |\beta_j, B\rangle$ then form an orthonormal basis in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, with the above ranges for i, j . A normalized pure quantum state $|\psi\rangle$ in \mathcal{H} may be expressed in this basis by,

$$|\psi\rangle = \sum_{i,j} \psi_{ij} |\alpha_i, A\rangle \otimes |\beta_j, B\rangle \quad \sum_{i,j} |\psi_{ij}|^2 = 1 \quad (5.16)$$

where ψ_{ij} are complex coefficients. Let us now compute the purely quantum mechanical expectation value of an observable \mathcal{A} of the subsystem \mathcal{H}_A in the pure state $|\psi\rangle \in \mathcal{H}$. Making no observations in subsystem B may be represented by extending the operator \mathcal{A} on \mathcal{H}_A to an operator $\hat{\mathcal{A}}$ on \mathcal{H} by letting,

$$\mathcal{A} \rightarrow \hat{\mathcal{A}} = \mathcal{A} \otimes I_B \quad (5.17)$$

where I_B is the identity operator in \mathcal{H}_B . This is a generalization of a similar extension we used for the two-state system in (5.6). Evaluating the expectation value of $\hat{\mathcal{A}}$ in the pure state $|\psi\rangle$ of \mathcal{H} , we find,

$$\langle \psi | \hat{\mathcal{A}} | \psi \rangle = \sum_{i,j} \sum_{i',j'} \psi_{ij}^* \psi_{i'j'} \langle \alpha_i, A | \otimes \langle \beta_j, B | \mathcal{A} \otimes I_B | \alpha_{i'}, A \rangle \otimes | \beta_{j'}, B \rangle \quad (5.18)$$

Using the orthogonality relation $\langle \beta_j, B | \beta_{j'}, B \rangle = \delta_{j,j'}$, and introducing the linear operator ρ_A on \mathcal{H}_A , defined by,

$$\rho_A = \sum_{i,j,i'} |\alpha_{i'}, A\rangle \psi_{ij} \psi_{ij}^* \langle \alpha_i, A | \quad (5.19)$$

we see that we have

$$\langle \psi | \hat{\mathcal{A}} | \psi \rangle = \text{Tr}_{\mathcal{H}_A} (\rho_A \mathcal{A}) \quad (5.20)$$

The operator ρ_A is self-adjoint, satisfies $\text{Tr}_{\mathcal{H}_A} (\rho_A) = 1$, and is positive. Thus, ρ_A qualifies as a density matrix. The density matrix ρ_A is a projection operator and corresponds to a pure state of \mathcal{H}_A if and only if it has rank 1, which in turn requires the matrix ψ_{ij} to have rank 1. In all other cases, ρ_A corresponds to a non-trivial mixture.

A quicker, but more formal derivation of the above calculation makes use of the relation,

$$\text{Tr}_{\mathcal{H}} \mathcal{O} = \text{Tr}_{\mathcal{H}_A} (\text{Tr}_{\mathcal{H}_B} \mathcal{O}) \quad (5.21)$$

valid for any operator in \mathcal{O} on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, as well as the relation $\langle \psi | \hat{\mathcal{A}} | \psi \rangle = \text{Tr}_{\mathcal{H}} (\hat{\mathcal{A}} |\psi\rangle \langle \psi|)$. Applying both of these relations gives (5.20) with,

$$\rho_A = \text{Tr}_{\mathcal{H}_B} (|\psi\rangle \langle \psi|) \quad (5.22)$$

It is readily checked that the right hand formula for ρ_A reproduces (5.19).

5.7 Statistical entropy of a density matrix

The density matrix for an ensemble of n pure (orthonormal) states $|\phi_i\rangle$ occurring with probability p_i in the ensemble is given by,

$$\rho = \sum_{i=1}^n |\phi_i\rangle p_i \langle \phi_i| \qquad \sum_{i=1}^n p_i = 1 \qquad (5.23)$$

When all probabilities are equal and given by $p_i = 1/n$, the density matrix corresponds to the uniform ensemble, and the entropy is given by Boltzmann's formula $S = k \ln n$. Let us now compute the entropy when the probabilities are not all equal.

To do so, we follow Gibbs again. Instead of considering just one system, consider a very large number \mathcal{N} of identical systems, and let us compute the entropy not just for one system, but for all \mathcal{N} of them. By the definition of what probability means for large \mathcal{N} , we then see that of those \mathcal{N} systems, $\mathcal{N}_i = p_i \mathcal{N}$ will be in state $|\phi_i\rangle$, with

$$\sum_{i=1}^n \mathcal{N}_i = \mathcal{N} \qquad \mathcal{N}_i \geq 0 \qquad (5.24)$$

The number of micro-states that can realize a macro-state with the \mathcal{N}_i systems in state $|\phi_i\rangle$ is then given by a standard combinatorial formula,

$$\Omega_{\mathcal{N}} = \frac{\mathcal{N}!}{\mathcal{N}_1! \cdots \mathcal{N}_n!} \qquad (5.25)$$

This expression and the corresponding entropy $\xi_{\mathcal{N}}$ for the ensemble of \mathcal{N} systems may be recast in terms of the probabilities p_i and the total number of copies \mathcal{N} , and we find,

$$S_{\mathcal{N}} = k \ln \Omega_{\mathcal{N}} = k \left(\ln \mathcal{N}! - \sum_{i=1}^n \ln (p_i \mathcal{N})! \right) \qquad (5.26)$$

Since \mathcal{N} is very large, we may use the Sterling formula $m! \approx m \ln m - m$, and find,

$$S_{\mathcal{N}} = k \left(\mathcal{N} \ln \mathcal{N} - \mathcal{N} - \sum_{i=1}^n \left\{ p_i \mathcal{N} \ln (p_i \mathcal{N}) - p_i \mathcal{N} \right\} \right) \qquad (5.27)$$

The term $\mathcal{N} \ln \mathcal{N} - \mathcal{N}$ cancels since the p_i sum to 1, and we are left with,

$$S_{\mathcal{N}} = -k \mathcal{N} \sum_{i=1}^n p_i \ln p_i \qquad (5.28)$$

We find, that the entropy for the \mathcal{N} systems of the ensemble is simply proportional to \mathcal{N} , as should be expected. Thus, it makes sense to extract the entropy S for a single system by dividing $S_{\mathcal{N}}$ by \mathcal{N} and setting $S = S_{\mathcal{N}}/\mathcal{N}$, so that we find,

$$S(p_1, \cdots, p_n) = -k \sum_{i=1}^n p_i \ln p_i \qquad (5.29)$$

or equivalently in terms of the density matrix directly, we have,

$$S(\rho) = -k \text{Tr}(\rho \ln \rho) \quad (5.30)$$

Setting all probabilities equal $p_i = 1/n$ for the uniform ensemble, we recover $S = k \ln n$.

Some basic properties of the statistical entropy are as follows.

1. The above construction of the statistical entropy does not assume equilibrium.
2. Positivity, $S(p_1, \dots, p_n) \geq 0$;
3. The minimum $S = 0$ is attained when all probability assignments are 0, except for a single entry $p_j = 1$. The entropy vanishes if and only if ρ corresponds to a pure state.
4. The maximum $S_{\max} = k \ln n$ is attained when all probabilities are equal, $p_i = 1/n$.
5. **Invariance** under conjugation of the density operator by a unitary transformation. In particular, the entropy is invariant under time evolution, under the assumption that the probabilities p_i remain unchanged in time;
6. **Additivity** upon combination of two subsystems which are *statistically uncorrelated*. Let the systems be described by Hilbert spaces \mathcal{H}_a and \mathcal{H}_b , with density operators ρ_a and ρ_b respectively, then the full Hilbert space is $\mathcal{H}_{ab} = \mathcal{H}_a \otimes \mathcal{H}_b$ and the density matrix for the combined system is $\rho_{ab} = \rho_a \otimes \rho_b$. The entropy is then additive,

$$S(\rho_{ab}) = S(\rho_a) + S(\rho_b) \quad (5.31)$$

7. **Subadditivity** upon dividing a system with Hilbert space \mathcal{H}_{ab} and density operator ρ_{ab} into two subsystems with Hilbert spaces \mathcal{H}_a and \mathcal{H}_b , and density matrices $\rho_a = \text{Tr}_{\mathcal{H}_b}(\rho_{ab})$ and $\rho_b = \text{Tr}_{\mathcal{H}_a}(\rho_{ab})$ which are *statistically correlated*. The full density operator ρ_{ab} is *not the tensor product of ρ_a and ρ_b* , in view of the non-trivial statistical correlations between the two subsystems. Instead, one only has a strict inequality,

$$S(\rho_{ab}) \leq S(\rho_a) + S(\rho_b) \quad (5.32)$$

The proofs of these properties will be developed in problem sets.

Entropy may be given a meaning beyond traditional statistical mechanics. In developing a theory of information around 1948, Claude Shannon was led to a generalized notion of entropy that characterizes the amount of *missing information* for a given ensemble. In the case of information theory, the ensembles consist of messages, sent in words and sentences. To make contact with the previous sections, a message may be viewed as a mixture of a certain number of letters and words, occurring with certain probabilities p_i . Shannon was led precisely to the entropy of (5.29) to characterize quantitatively the missing information.

5.8 The uniform and micro-canonical ensembles

The ensembles defined in classical statistical mechanics may be generalized to the quantum case. In this section, we shall briefly discuss the uniform and micro-canonical ensembles, but spend most of our attention on the canonical ensemble whose usefulness is the greatest.

In the uniform ensemble, no macro-state information is given. By the principle of a priori equal probabilities, the density matrix is proportional to the identity operator in the full Hilbert space \mathcal{H} of the quantum system, $\rho = I_{\mathcal{H}}/\Omega$, where Ω is the total number of states which is equal to the dimension of \mathcal{H} . Clearly, the uniform density matrix is time-independent and corresponds to an equilibrium ensemble. The corresponding entropy is $S = k \ln \Omega$.

In the micro-canonical ensemble, the total energy of the system is fixed to be E . The density matrix has support only on eigenstates $|E_i, \alpha_i\rangle$ of the Hamiltonian with energy $E_i = E$, the α_i denoting degeneracies of the energy level E_i . In view of the principle of equal a priori probabilities, the weights of all the states $|E, \alpha\rangle$ for various values of α are all the same, and given by the total number of states $\Omega(E)$ at energy E . Thus, the density matrix may be expressed in terms of the projection operator P_E onto states of energy E ,

$$\rho = \frac{P_E}{\Omega(E)} = \frac{1}{\Omega(E)} \sum_{\alpha} |E, \alpha\rangle \langle E, \alpha| \quad (5.33)$$

The corresponding entropy is $S(E) = k \ln \Omega(E)$.

5.9 Construction of the density matrix in the canonical ensemble

In the canonical ensemble, the temperature T is fixed. This may be achieved by putting the system in contact with another very large system (or heat bath), and letting the combined system reach equilibrium. Internal energy E may then be exchanged with the heat bath, and its average value must then be determined so as to achieve temperature T . In terms of the density matrix, we have the equations,

$$\text{Tr}(\rho) = 1 \qquad \text{Tr}(\rho H) = E \quad (5.34)$$

The normalization constraint has been made explicit here in order to be able to liberate this condition on ρ . Equilibrium will be achieved by maximizing the entropy, subject to the above constraints. Extremization subject to constraints is carried out mathematically with the help of Lagrange multipliers, one for each constraint. Thus, we shall extremize the combination, $S(\rho) - \alpha \text{Tr} \rho - k\beta \text{Tr}(\rho H)$, and set,

$$\begin{aligned} 0 &= \delta S(\rho) - \alpha \delta \text{Tr} \rho - k\beta \delta \text{Tr}(\rho H) \\ &= -k \delta \text{Tr}(\rho \ln \rho) - \alpha \text{Tr}(\delta \rho) - k\beta \text{Tr}(\delta \rho H) \end{aligned} \quad (5.35)$$

The first term is calculated with the help of the mathematical identity,

$$\delta \text{Tr}(\rho \ln \rho) = \text{Tr}(\delta \rho \ln \rho + \delta \rho) \quad (5.36)$$

Putting all together, we find the following relation for ρ ,

$$0 = \text{Tr} \delta \rho \left(\ln \rho + (k + \alpha)/k I_{\mathcal{H}} + \beta H \right) \quad (5.37)$$

The operators ρ , $\delta \rho$, $I_{\mathcal{H}}$, and H are all self-adjoint as well. Satisfying equation for all $\delta \rho$ then requires that $\ln \rho + \beta H$ is proportional to the identity or,

$$\rho = \frac{e^{-\beta H}}{Z(\beta)} \quad (5.38)$$

The parameter β must be identified with inverse temperature $\beta = 1/kT$. The partition function $Z(\beta)$ is determined by the normalization condition $\text{Tr} \rho = 1$, and is thus given by,

$$Z(\beta) = \text{Tr} \left(e^{-\beta H} \right) \quad (5.39)$$

where the trace extends over the full Hilbert space. The internal energy may now be computed directly from $Z(\beta)$ by,

$$E(\beta) = \text{Tr}(\rho H) = \frac{1}{Z(\beta)} \text{Tr} \left(H e^{-\beta H} \right) = -\frac{\partial \ln Z(\beta)}{\partial \beta} \quad (5.40)$$

and the free energy is found to be,

$$F(\beta) = -kT \ln Z(\beta) \quad (5.41)$$

Finally, we show that the thermodynamic definition of the entropy coincides with its statistical definition. To this end, we start with the thermodynamic relation $E = F - TS$ to obtain the entropy with the help of (5.40) and (5.41),

$$S = \frac{1}{T} \text{Tr}(\rho H) + k \ln Z \quad (5.42)$$

Taking the logarithm of the operator equation (5.38), we have $\ln \rho = -\beta H - \ln Z$. Using this formula to eliminate H in (5.42), we find,

$$S = -k \text{Tr}(\rho \ln \rho) \quad (5.43)$$

reproducing the statistical definition of entropy.

5.10 Generalized equilibrium ensembles

The derivation of the Boltzmann weights and associated density matrix corresponds to the canonical ensemble, in which only the energy of the system is kept constant. In the grand canonical ensemble, both the energy and the number of particles in the system is kept constant. More generally, we consider an ensemble in which the ensemble average of a number of commuting observables A_i , $i = 1, \dots, K$ is kept constant. To compute the associated density operator ρ of this ensemble, we extremize with respect to variations in ρ the entropy, under the constraint that the ensemble averages $\text{Tr}(\rho A_i)$ are kept constant. Using again Lagrange multipliers β_i , $i = 1, \dots, K$, we extremize

$$-\text{Tr}(\rho \ln \rho) - \sum_{i=1}^K \beta_i \text{Tr}(\rho A_i) \quad (5.44)$$

Upon enforcing the normalization $\text{Tr}(\rho) = 1$, this gives,

$$\begin{aligned} \rho &= \frac{1}{Z} \exp \left\{ - \sum_{i=1}^K \beta_i A_i \right\} \\ Z &= \text{Tr} \left(\exp \left\{ - \sum_{i=1}^K \beta_i A_i \right\} \right) \end{aligned} \quad (5.45)$$

In the grand canonical ensemble, for example, these quantities are

$$\begin{aligned} \rho &= \frac{1}{Z} e^{-\beta H + \mu N} \\ Z &= \text{Tr} \left(e^{-\beta H + \mu N} \right) \end{aligned} \quad (5.46)$$

where N is the number operator and μ is the chemical potential. Other observables whose ensemble averages are often kept fixed in this way are electric charge, baryon number, electron number etc.

6 Applications of the canonical ensemble

We shall now illustrate the use of the canonical (and micro-canonical) distributions on some simple, but physically important systems.

6.1 The statistics of paramagnetism

Consider a system of N non-interacting magnetic dipoles with individual magnetic moment μ in the presence of an external magnetic field \mathbf{B} . The magnetic dipoles may be those of elementary particles, such as the electron, proton or neutron, in which case they are quantum effects, along with the spin of these particles. The corresponding Hamiltonian is given by,

$$H = - \sum_{i=1}^N \vec{\kappa}_i \cdot \mathbf{B} \quad (6.1)$$

Quantum mechanically,³ the magnetic moment is proportional to the spin \mathbf{s} of the particle (or more generally to the total angular),

$$\vec{\kappa}_i = g\kappa \frac{\mathbf{s}_i}{\hbar} \quad \kappa = \frac{\hbar e}{2m_e c} \quad (6.2)$$

where κ is the *Bohr magneton* of the dipole, given in terms of its mass m_e of the electron, and the basic unit electric charge e (namely the charge of the electron), and the *Landé factor* g . For the electron, we have $g \approx 2$. For a given type of particle, the total spin s will be fixed, and given by the eigenvalue $\hbar^2 s(s+1)$ of \mathbf{s}^2 . The eigenvalues of $\vec{\kappa} \cdot \mathbf{B}$ are then given by,

$$g\kappa B m \quad m = -s, -s+1, \dots, s-1, s \quad (6.3)$$

Since the particles are non-interacting, the partition function Z is given by the N -th power of the single particle partition function Z_1 , which in turn is given by,

$$Z = (Z_1)^N \quad Z_1 = \sum_{m=-s}^s e^{\beta g \kappa B m} \quad (6.4)$$

The sum over m is geometric, and may be carried out analytically,

$$Z_1 = \frac{\sinh((2s+1)x)}{\sinh x} \quad x = \frac{1}{2} \beta g \kappa B \quad (6.5)$$

³The classical treatment is parallel to the classical treatment of the electric dipole in an electric field, which was solved in problem set 3.

The corresponding free energy F , internal energy E , and magnetization M are given by,

$$\begin{aligned} F &= -kTN \ln Z_1 \\ E &= -N \frac{\partial \ln Z_1}{\partial \beta} \\ M &= \frac{N}{\beta} \frac{\partial \ln Z_1}{\partial B} \end{aligned} \quad (6.6)$$

Of special interest is the magnetization, which may be recast as follows,

$$M = \frac{1}{2}Ng\kappa \left((2s+1) \frac{\text{ch}(2s+1)x}{\text{sh}(2s+1)x} - \frac{\text{ch}x}{\text{sh}(x)} \right) \quad (6.7)$$

For large x (corresponding to large B and/or small T), the magnetization saturates at the value $sNg\kappa$, while for small x (corresponding to small B and/or large T), the magnetization follows the *Curie law*,

$$M = \frac{1}{3}s(s+1)Ng^2\kappa^2 \frac{B}{kT} \quad (6.8)$$

As expected, it behaves linearly in B for small B , and tends to zero for large T .

6.2 Non-relativistic Boltzmann ideal gas

We return once more to the system of an ideal gas of N particles, in the approximation of low density where Boltzmann statistics may be used. At higher densities, one will need to appeal to Fermi-Dirac or Bose-Einstein statistics instead. The microscopic constituents in an ideal gas are mutually non-interacting, so that the partition function is simply the N -th power of the partition function for a single particle. The single-particle Hamiltonian takes the following form,

$$H_1 = H_{\text{transl}} + H_{\text{int}} \quad H_{\text{transl}} = \frac{\mathbf{p}^2}{2M} \quad (6.9)$$

Here, H_{transl} corresponds to the center of mass motion for total momentum \mathbf{p} and total mass M , while H_{int} corresponds to the internal degrees of freedom, in which we include rotational and vibrational modes. These two Hamiltonians commute with one another, and may be simultaneously diagonalized. Putting the system in a large cubic box of linear size L , and volume $V = L^3$ with Dirichlet boundary conditions, the energy levels are then given by,

$$E_{\text{transl}} = \frac{\hbar^2}{2ML^2} (n_x^2 + n_y^2 + n_z^2) \quad n_x, n_y, n_z \geq 0 \quad (6.10)$$

and $E_{\text{int}} = \varepsilon_n$, where ε_n are the 1-particle energy levels of a single particle. The levels ε_n are clearly independent of V and N . Thus, the partition function decomposes as follows,

$$Z = (Z_{\text{transl}})^N (Z_{\text{int}})^N \quad (6.11)$$

The translational part was already computed earlier, and we have,

$$F_{\text{transl}} = -NkT + NkT \ln \left(\frac{N\lambda(T)^3}{V} \right) \quad \lambda^2 = \frac{2\pi\hbar^2}{MkT} \quad (6.12)$$

where $\lambda(T)$ is the thermal wavelength introduced in (2.48). In all generality, all that we can state about the internal part Z_{int} is that it only depends on T , but not on V or N ,

$$F_{\text{int}}(T) = -kT \ln Z_{\text{int}} \quad Z_{\text{int}} = \sum_n e^{-\beta\varepsilon_n} \quad (6.13)$$

We conclude that a general non-relativistic Boltzmann ideal gas has the following free energy,

$$F = NF_{\text{int}}(T) - NkT + NkT \ln \left(\frac{N\lambda(T)^3}{V} \right) \quad (6.14)$$

Using the standard thermodynamic relations of (2.24), we find,

$$\begin{aligned} E &= \frac{3}{2}NkT - N \frac{\partial}{\partial \beta} (\beta F_{\text{int}}) \\ S &= -NF'_{\text{int}} + \frac{5}{2}Nk - Nk \ln \left(\frac{N\lambda(T)^3}{V} \right) \\ PV &= NkT \\ \mu &= F_{\text{int}} + kT \ln \left(\frac{N\lambda(T)^3}{V} \right) \end{aligned} \quad (6.15)$$

We observe as general properties of any ideal gas that,

- the internal energy density E/N depends only on T , and not on the density N/V ;
- the law $PV = NkT$ holds for all ideal gasses.

6.3 Van der Waals equation of state

In real gasses, interactions between atoms do occur, and must be taken into account. We shall here treat the case of a classical mono-atomic gas of identical particles, so that the only degrees of freedom are the translational ones. Thus, the Hamiltonian for N particles will be,

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + U \quad (6.16)$$

Generally, U will be a function of all variables at once. In a mono-atomic gas, U will be a function only of the positions of the atoms. In the approximation of a dilute real gas, however, one may assume that the interactions occur only between pairs of particles, neglecting interactions between triples, quadruples etc. In the partition function, the momentum and coordinate integrals factorize. The momentum integrals give the contribution for the ideal gas, and may be factored out. Thus, we find,

$$Z = Z_{\text{ideal}} Z_U \quad (6.17)$$

where the ideal gas part is given by

$$Z_{\text{ideal}} = \frac{V^N}{N!} \left(\frac{mkT}{2\pi\hbar^2} \right)^{3N/2} \quad (6.18)$$

while the interaction part is given by,

$$Z_U = 1 + \frac{1}{V^N} \prod_{i=1}^N \int d^3\mathbf{q}_i \left(e^{-U(\mathbf{q})/kT} - 1 \right) \quad (6.19)$$

We have rearranged this partition function to expose the value for the ideal gas when $U = 0$. There are $N(N-1)/2 \sim N^2/2$ combinatorial ways in which the two-body interaction can occur amongst N bodies, so that within this approximation, we find,

$$Z_U = 1 + \frac{N^2}{2V^2} \int d^3\mathbf{q}_1 d^3\mathbf{q}_2 \left(e^{-U(\mathbf{q}_1, \mathbf{q}_2)/kT} - 1 \right) \quad (6.20)$$

If U depends only on the relative distances between the particles, then this formula may be further simplified, and we find,

$$Z_U = 1 - \frac{N^2}{V} B(T) \quad B(T) = \frac{1}{2} \int d^3\mathbf{q} \left(1 - e^{-U(\mathbf{q})/kT} \right) \quad (6.21)$$

where $B(T)$ depends only on T and the interaction, but not on V , under the assumption that the interaction is sufficiently short-ranged.

If the interaction is everywhere weak compared to the temperature scale, then we may use the approximation $e^{-U/kT} \ll 1$, to derive the following expression for the free energy,

$$F = -NkT + NkT \ln \left(\frac{N\lambda(T)^3}{V} \right) + kT \frac{N^2 B(T)}{V} \quad (6.22)$$

The pressure $P = -\partial F/\partial V$ is given by,

$$PV = NkT \left(1 + \frac{NB(T)}{V} \right) \quad (6.23)$$

This approximation is not very realistic for real gasses though. Instead, at room temperatures, there is a core overlap of radius r_0 where the interaction is very strong (compared to T), but negligible outside of this core. Thus, we may then use the following approximate formula for B ,

$$B(T) = 2\pi \int_0^{r_0} r^2 dr + \frac{2\pi}{kT} \int_{r_0}^{\infty} r^2 dr U \quad (6.24)$$

The first term is independent of T , while the second is inversely proportional to it. For an attractive potential $U < 0$, we have,

$$B(T) = b - \frac{a}{T} \quad a, b > 0 \quad (6.25)$$

The corresponding free energy is found to be,

$$F = -NkT + NkT \ln \left(\frac{N\lambda(T)^3}{V} \right) + \frac{N^2}{V} (bkT - a) \quad (6.26)$$

Computing the pressure, we find,

$$P = \frac{kNT}{V} + \frac{N^2}{V^2} (bkT - a) \quad (6.27)$$

or

$$\left(P + a \frac{N^2}{V^2} \right) = \frac{kNT}{V} + \frac{N^2}{V^2} bkT \approx \frac{NkT}{V - Nb} \quad (6.28)$$

which gives the Van der Waals equation for a real gas,

$$\left(P + a \frac{N^2}{V^2} \right) (V - Nb) = NkT \quad (6.29)$$

under the assumption that $V \gg Nb$.

6.4 The Mayer cluster expansion

More generally, we can perform a more systematic expansion for a system in which the interaction potential is a sum of a single to-dody potential acting between any pair of identical particles. We shall denote this potential by $U(r_{ij})$ where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, and we use \mathbf{r} here rather than \mathbf{q} for the position vectors. The partition function

$$Z = \frac{1}{N!(2\pi\hbar)^N} \prod_{i=1}^N \int d^3\mathbf{p}_i d^3\mathbf{r}_i e^{-\beta H} \quad (6.30)$$

factors as follows,

$$Z = Z_{\text{ideal}} Z_U \quad Z_U = \frac{1}{V^N} \prod_{i=1}^N \int d^3 \mathbf{r}_i e^{-\beta \sum_{j < k} U(r_{jk})} \quad (6.31)$$

where the ideal gas partition function was given in (??). Defining the following function,

$$f_{jk} = e^{-\beta U(r_{jk})} - 1 \quad (6.32)$$

we may recast Z_U as follows,

$$Z_U = \frac{1}{V^N} \prod_{i=1}^N \int d^3 \mathbf{r}_i \prod_{j < k} (1 + f_{jk}) \quad (6.33)$$

The Mayer expansion is obtained by expanding the product in powers of f , and we have,

$$Z_U = \frac{1}{V^N} \prod_{i=1}^N \int d^3 \mathbf{r}_i \left(1 + \sum_{j < k} f_{jk} + \sum_{j < k, m < n} f_{jk} f_{mn} + \dots \right) \quad (6.34)$$

The first term gives 1, the second term gives

$$\frac{N^2}{2V} \int d^3 \mathbf{r} f(r) \quad (6.35)$$

and so on. The higher order terms admit a diagrammatic expansion.

7 Systems of indistinguishable quantum particles

Identical quantum particles (or systems of particles such as nuclei, atoms or molecules) are not only indistinguishable; they obey specific *quantum permutation symmetries*. In 3 space-dimensions (and higher), the only two types of permutation symmetries allowed by local Poincaré invariant quantum field theory correspond to either Bose-Einstein or Fermi-Dirac, respectively for particles (or systems) either of integer spin or of half odd-integer spin. This correspondence will be heavily used, but not be proven here, as it is properly the subject of quantum field theory.

7.1 FD and BE quantum permutation symmetries

Concretely, we consider a system of N identical particles (or systems), which we will label by an integer $n = 1, \dots, N$. We denote the Hilbert space of a single particle by \mathcal{H} , and label a basis of quantum states $|\nu\rangle = |\nu^1, \dots, \nu^A\rangle$ in \mathcal{H} by an array ν of A quantum numbers ν^α with $\alpha = 1, \dots, A$, corresponding to a maximal set of commuting observables \mathcal{O}^α in \mathcal{H} ,

$$\mathcal{O}^\alpha |\nu^1, \dots, \nu^A\rangle = \nu^\alpha |\nu^1, \dots, \nu^A\rangle \quad (7.1)$$

The Hilbert space \mathcal{H}_N of the N indistinguishable particles is then given by the tensor product of N copies of \mathcal{H} ,

$$\mathcal{H}_N = \underbrace{\mathcal{H} \otimes \mathcal{H} \otimes \dots \otimes \mathcal{H}}_N \quad (7.2)$$

The observables \mathcal{O}^α on \mathcal{H} may be naturally extended to observables \mathcal{O}_n^α on \mathcal{H}_N , with the help of the identity operator I in \mathcal{H} , and we have,

$$\mathcal{O}_n^\alpha = \underbrace{I \otimes \dots \otimes I}_{n-1} \otimes \mathcal{O}^\alpha \otimes \underbrace{I \otimes \dots \otimes I}_{N-n-1} \quad (7.3)$$

The quantum numbers labeling the states of \mathcal{H}_N may be taken to be those of the maximal set of commuting observables $\{\mathcal{O}_n^\alpha\}$ of \mathcal{H}_N , with $n = 1, \dots, N$ and $\alpha = 1, \dots, A$. Thus, the basis states of \mathcal{H}_N may be labeled by,

$$|\nu_1\rangle \otimes |\nu_2\rangle \otimes \dots \otimes |\nu_N\rangle \quad (7.4)$$

We stress that ν_n here stands for the full array $\nu_n = \nu_n^1, \nu_n^2, \dots, \nu_n^A$. The action of the operators \mathcal{O}_i^α in this basis may be read off from their definition,

$$\mathcal{O}_n^\alpha |\nu_1\rangle \otimes |\nu_2\rangle \otimes \dots \otimes |\nu_N\rangle = \nu_n^\alpha |\nu_1\rangle \otimes |\nu_2\rangle \otimes \dots \otimes |\nu_N\rangle \quad (7.5)$$

The action of a permutation σ amongst these N particles is defined by its action on each particle, and may be expressed in terms of the quantum numbers ν_n . The quantum permutation symmetry allowed for a physical state $|\nu_1, \nu_2, \dots, \nu_N\rangle \in \mathcal{H}_N$ of N identical particles with quantum numbers ν_n for $n = 1, \dots, N$, is either one of the following,

$$|\nu_{\sigma(1)}, \nu_{\sigma(2)}, \dots, \nu_{\sigma(N)}\rangle = \begin{cases} +|\nu_1, \nu_2, \dots, \nu_N\rangle & \text{Bose-Einstein} \\ (-)^\sigma |\nu_1, \nu_2, \dots, \nu_N\rangle & \text{Fermi-Dirac} \end{cases} \quad (7.6)$$

where $(-)^{\sigma}$ denotes the signature of the permutation σ . The states $|\nu_1, \nu_2, \dots, \nu_N\rangle$ may be constructed explicitly by symmetrizing (BE) or anti-symmetrizing (FD) the states of (7.4),

$$|\nu_1, \nu_2, \dots, \nu_N\rangle = \frac{1}{\sqrt{N!}} \sum_{\sigma \in \mathcal{S}_N} \delta_\sigma |\nu_{\sigma(1)}, \nu_{\sigma(2)}, \dots, \nu_{\sigma(N)}\rangle \quad (7.7)$$

where

$$\delta_\sigma = \begin{cases} 1 & \text{Bose-Einstein} \\ (-)^{\sigma} & \text{Fermi-Dirac} \end{cases} \quad (7.8)$$

and \mathcal{S}_N is the set of all permutations of N particles.

This implies that two identical fermions cannot occupy the same 1-particle quantum state; i.e. two identical fermions cannot have the same quantum numbers. Two or more identical bosons, however, may freely occupy the same quantum state.

7.2 BE and FD statistics for N identical free particles

The BE and FD quantum permutation symmetry conditions were formulated on the micro-states of the theory. To derive the consequences of these microscopic conditions on macroscopic states of indistinguishable particles, and on the counting of the number of micro-states to which a macro-state has access, we need to develop some new counting techniques. In this section, we limit study to that of non-interacting particles.

Since the particles are non-interacting, the Hamiltonian is just the sum of the N one-particle Hamiltonians $H^{(1)}$. Following the notation of the previous section, it is given by,

$$H = \sum_{n=1}^N H_n \quad H_n = \underbrace{I \otimes \dots \otimes I}_{n-1} \otimes H^{(1)} \otimes \underbrace{I \otimes \dots \otimes I}_{N-n-1} \quad (7.9)$$

We denote the eigenstates of $H^{(1)}$ by $|\alpha\rangle$, and the corresponding eigenvalue by ε_α and use α as a collective label for all quantum numbers of the one-particle states (α was denoted by ν

in the preceding section). A macro-state with total number of particles N , total energy E , and N_α particles in micro-state $|\alpha\rangle$ then satisfies,

$$\sum_{\alpha} N_{\alpha} = N \qquad \sum_{\alpha} N_{\alpha} \varepsilon_{\alpha} = E \qquad (7.10)$$

In the thermodynamic limit, the energy levels of $H^{(1)}$ will become closely spaced, and approach a continuum distribution.

To impose the quantum permutation symmetry on the micro-states, we appeal to the *coarse-graining procedure*, and divide the (basis of states in the) Hilbert space \mathcal{H} of all states into discrete cells, labelled by an integer $i = 1, 2, 3, \dots$. Physically, this may be done by ordering the states according to the energy levels of a single particle $H^{(1)}$ and/or any other conserved quantum number. The number of (distinct) states in cell i is denoted by G_i . In the thermodynamic limit, the spectrum of $H^{(1)}$ becomes closely spaced, so that we may assume $G_i \gg 1$ for all i . We denote the number of particles in cell i by N_i , and assume that $N_i \gg 1$. Thus, each cell by itself may be viewed as a macroscopic subsystem of the whole macro-state. The set-up is schematically represented in Figure 6, where the set of basis states of \mathcal{H} is divided into cells i (two cells are separated by a long vertical dash), each cell having G_i quantum states (each state is indicated with a short vertical dash), and N_i particles (each particle is indicated with a dot above an available micro-state).

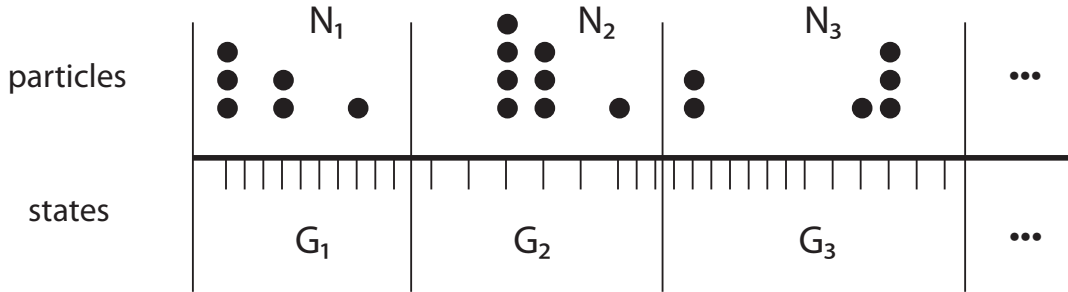


Figure 6: Cell decomposition of the spectrum of available micro-states.

In practice, it will be assumed that the cells are “small enough” so that the energy and/or other conserved quantum numbers remain constant throughout a given cell. Finally, it must be stressed that any physically observable quantity should be independent of the precise coarse graining mesh that has been applied to the system. One such physical quantity is the *mean occupation number* n_i for each cell, defined by,

$$n_i = \frac{N_i}{G_i} \qquad (7.11)$$

The number Ω_i of micro-states in each cell, and the associated contribution to the entropy S_i , are related to the total number of micro-states Ω and the total entropy by,

$$\Omega = \prod_i \Omega_i \qquad S = k \sum_i \ln \Omega_i \qquad (7.12)$$

The object is to compute the numbers Ω_i of micro-states in each cell i , and the associated total energy E and total particle number N , given by,

$$\begin{aligned} N &= \sum_i N_i = \sum_i G_i n_i \\ E &= \sum_i N_i \varepsilon_i = \sum_i G_i n_i \varepsilon_i \end{aligned} \qquad (7.13)$$

where ε_i is the average energy of the one-particle states in cell i .

Equilibrium will be achieved by maximizing the entropy with respect to the occupation numbers n_i , while keeping fixed the sizes of the cells G_i , the total energy E and, when appropriate also the total number of particles N . This may be achieved with the help of Lagrange multipliers β and $\beta\mu$, as usual,

$$\delta_{n_i} (S/k - \beta E + \beta\mu N) = 0 \qquad (7.14)$$

Note that the sizes G_i of the cells should be viewed as artifacts of the coarse-graining, and therefore should not enter into any physical quantities.

7.3 Boltzmann statistics rederived

As a warm-up, we could use this construction to re-derive Boltzmann statistics. The number of micro-states Ω_i available to cell i is given by $G_i^{N_i}$ for distinguishable particles, and $G_i^{N_i}/N_i!$ for indistinguishable particles. In the approximation of $G_i \gg 1$, the entropy is given by,

$$S = -k \sum_i G_i (n_i \ln n_i + n_i) \qquad (7.15)$$

Maximizing S according to (7.14), keeping G_i, E, N fixed, we recover the Boltzmann distribution $n_i = e^{-\beta(\varepsilon_i - \mu)}$. Note that this result is independent of G_i as anticipated.

7.4 Fermi-Dirac statistics

For Fermi-Dirac statistics, at most one particle can occupy a given quantum state. Thus, the number of micro-states accessible to the macroscopic subsystem of a single cell is,

$$\Omega_i = \frac{G_i!}{N_i!(G_i - N_i)!} \qquad (7.16)$$

In the approximation of large G_i , the total entropy is given by,

$$S = -k \sum_i G_i \left(n_i \ln n_i + (1 - n_i) \ln(1 - n_i) \right) \quad (7.17)$$

Note that this derivation of the entropy does not appeal to any equilibrium arguments, it is simply based on coarse graining and quantum statistics counting.

To obtain the equilibrium distribution for the occupation numbers n_i , we use again (7.14), for fixed G_i, E and, where appropriate, fixed N ,

$$\sum_i G_i \delta n_i \left(-\ln n_i + \ln(1 - n_i) - \beta \varepsilon_i + \beta \mu \right) = 0 \quad (7.18)$$

Setting this variation to zero for each δn_i gives an equation for each n_i ,

$$n_i = \frac{1}{e^{\beta(\varepsilon_i - \mu)} + 1} \quad (7.19)$$

where β and μ are related to E, N by,

$$E = \sum_i \frac{G_i \varepsilon_i}{e^{\beta(\varepsilon_i - \mu)} + 1} \quad N = \sum_i \frac{G_i}{e^{\beta(\varepsilon_i - \mu)} + 1} \quad (7.20)$$

We see that G_i naturally has the interpretation as the degeneracy of the energy level i . The partition function is readily deduced, and we find,

$$Z = \prod_i \left(1 + e^{-\beta(\varepsilon_i - \mu)} \right)^{G_i} \quad (7.21)$$

In the grand-canonical ensemble, the partition function is directly related to the Gibbs free energy G by the relation $G = -kT \ln Z$ of (3.42). For our case, this quantity is given by,

$$G = E - TS - \mu N = -PV \quad (7.22)$$

where we have used the homogeneity relation $E = TS - PV + \mu N$ of (2.19). In terms of the partition function, we thus find,

$$PV = kT \sum_i G_i \ln \left(1 + e^{-\beta(\varepsilon_i - \mu)} \right) \quad (7.23)$$

which precisely corresponds to filling up each micro-state with at most one particle. This is a very convenient relation, which gives directly the *equation of state*.

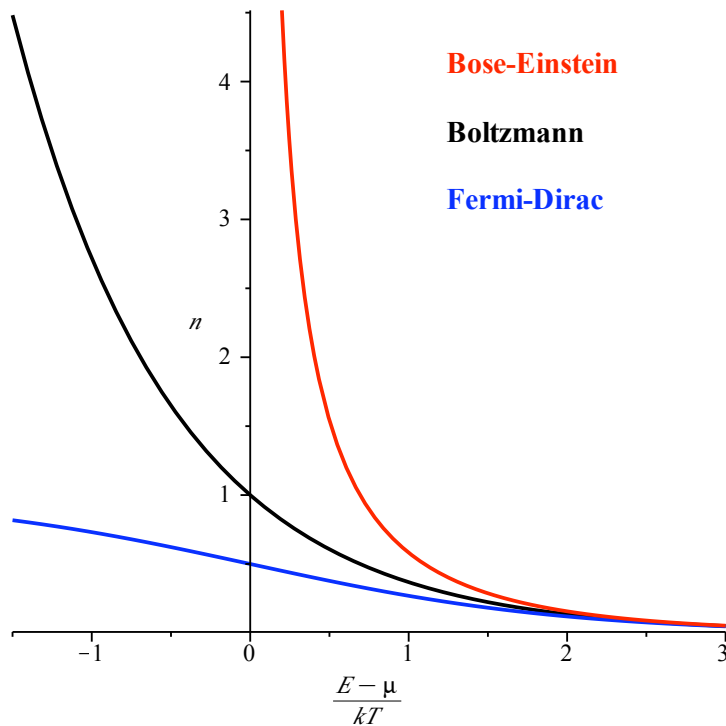


Figure 7: Comparison of the occupation numbers for quantum statistics.

7.5 Bose-Einstein statistics

For Bose-Einstein statistics, an arbitrary number of particles may occupy any given quantum state. Therefore, the number of accessible micro-states is given by,

$$\Omega_i = \frac{(G_i + N_i - 1)!}{N_i! (G_i - 1)!} \quad (7.24)$$

In the approximation of large G_i , the total entropy is given by,

$$S = -k \sum_i G_i (n_i \ln n_i - (1 + n_i) \ln(1 + n_i)) \quad (7.25)$$

Again, this derivation does not require the system to be in equilibrium.

To obtain the equilibrium distribution, we again use (7.14), and obtain,

$$n_i = \frac{1}{e^{\beta(\varepsilon_i - \mu)} - 1} \quad (7.26)$$

where β and μ are related to E, N by,

$$E = \sum_i \frac{G_i \varepsilon_i}{e^{\beta(\varepsilon_i - \mu)} - 1} \quad N = \sum_i \frac{G_i}{e^{\beta(\varepsilon_i - \mu)} - 1} \quad (7.27)$$

The free energy and partition function are readily deduced, and we find,

$$\begin{aligned}
 F &= kT \sum_i G_i \ln \left(1 - e^{-\beta(\varepsilon_i - \mu)} \right) \\
 Z &= \prod_i \left(\sum_{n=0}^{\infty} e^{-\beta n(\varepsilon_i - \mu)} \right)^{G_i}
 \end{aligned}
 \tag{7.28}$$

which corresponds to filling up each micro-state with an arbitrary number of particles.

7.6 Comparing the behavior of the occupation numbers

Boltzmann, Fermi-Dirac, and Bose-Einstein statistics produce qualitatively different behaviors for the occupation numbers, as shown in Figure 7. Bose-Einstein statistics is consistent only for $\mu \leq 0$, while Boltzmann and Fermi-Dirac are defined for all ranges of μ . Both FD and BE statistics asymptote to Boltzmann in the limit of large $(E - \mu)/kT$, which is the limit of large energy levels, or better even, of

$$n \ll 1 \tag{7.29}$$

This was the limit in which the Sackur-Tetrode formula of the simplest ideal gas was valid. Thus, we see that FD and BE statistics will produce modifications thereof when the occupation numbers are not small.

8 Ideal Fermi-Dirac Gases

In this section, we shall discuss ideal gasses of particles obeying Fermi-Dirac statistics. Applications to the physics of electrons in metals, their magnetic properties, low energy behavior, white dwarfs and neutron stars, will be treated in relative detail.

Recall the basic equations for an ideal FD gas from equations (7.20) and (7.23),

$$\begin{aligned} N &= \sum_i \frac{1}{e^{\beta(\varepsilon_i - \mu)} + 1} \\ E &= \sum_i \frac{\varepsilon_i}{e^{\beta(\varepsilon_i - \mu)} + 1} \\ -G = PV = kT \ln Z &= kT \sum_i \ln \left(1 + e^{-\beta(\varepsilon_i - \mu)} \right) \end{aligned} \quad (8.1)$$

where we take the sum over all states labelled by i , including their degeneracies, and corresponding energy ε_i . Here, Z stands for the grand canonical partition function. The Gibbs free energy G , which is the natural thermodynamic potential in the grand canonical ensemble, is related to these quantities by $G = -PV$. As we are assuming that the gas is ideal, no mutual interactions take place; the Hamiltonian is just that of the free relativistic or non-relativistic particle, and ε_i are the free particle energies.

8.1 The simplest ideal Fermi-Dirac gas

For particles whose energy depends only on translational degrees of freedom, we may set,

$$\sum_i \rightarrow \frac{gV}{(2\pi\hbar)^3} \int d^3\mathbf{p} \quad (8.2)$$

where a degeneracy factor g has been included to account for the number of internal degrees of freedom of each particle. For example, when this degree of freedom is spin s , we have $g = 2s + 1$. The standard thermodynamic functions are then given as follows,

$$\begin{aligned} N &= \frac{gV}{(2\pi\hbar)^3} \int d^3\mathbf{p} \frac{1}{e^{\beta(\mathbf{p}^2/2m - \mu)} + 1} \\ E &= \frac{gV}{2m(2\pi\hbar)^3} \int d^3\mathbf{p} \frac{\mathbf{p}^2}{e^{\beta(\mathbf{p}^2/2m - \mu)} + 1} \\ PV &= kT \frac{gV}{(2\pi\hbar)^3} \int d^3\mathbf{p} \ln \left(1 + e^{-\beta(\mathbf{p}^2/2m - \mu)} \right) \end{aligned} \quad (8.3)$$

It is a property of the simplest ideal gas thermodynamics that temperature may be scaled out of all these integrals. To do so, we express the quantities in terms of the composite

variable z , referred to as the *fugacity*, and we will continue to use the notation $\lambda = \lambda(T)$ for the thermal wavelength,

$$z = e^{\beta\mu} \qquad \lambda = \left(\frac{2\pi\hbar^2}{mkT} \right)^{\frac{1}{2}} \quad (8.4)$$

Expressing μ in terms of z , and p in terms of x defined by $p^2 = 2mkTx$, we have the following alternative expressions,

$$\frac{N}{V} = \frac{g}{\lambda^3} f_{3/2}(z) \qquad E = \frac{3}{2}PV \qquad \frac{P}{kT} = \frac{g}{\lambda^3} f_{5/2}(z) \quad (8.5)$$

The Fermi-Dirac functions $f_\nu(z)$ depend only upon the fugacity z , and are defined as follows,

$$f_\nu(z) \equiv \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1} dx}{z^{-1}e^x + 1} \quad (8.6)$$

From (8.4), we have $0 < z < \infty$. More generally, in the regime $0 < \text{Re}(z) < \infty$, the integral is absolutely convergent for $\text{Re}(\nu) > 0$, and may be analytically continued to all $\nu \in \mathbf{C}$. The functions have the following expansion in powers of z ,

$$f_\nu(z) = \sum_{n=1}^{\infty} (-)^{n+1} \frac{z^n}{n^\nu} \quad (8.7)$$

The series is absolutely convergent only for $|z| \leq 1$ and $1 < \nu$, and gives the small z behavior of the functions, or equivalently for $\mu/kT \ll -1$. From the expansion in (8.7), it is immediate that we have the following recursion relation,

$$z \frac{\partial}{\partial z} f_\nu(z) = f_{\nu-1}(z) \quad (8.8)$$

which may also be derived directly from the integral representation, and holds for all $z \in \mathbf{C}$. The function $f_\nu(1)$ is related to the Riemann ζ -function $\zeta(\nu)$,

$$f_\nu(1) = \left(1 - \frac{1}{2^{\nu-1}} \right) \zeta(\nu) \qquad \zeta(\nu) = \sum_{n=1}^{\infty} \frac{1}{n^\nu} \quad (8.9)$$

where $\zeta(\nu)$ is well-known to be meromorphic throughout $\nu \in \mathbf{C}$, with a single pole at $\nu = 1$.

8.2 Entropy, specific heat, and equation of state

The free energy $F = E - TS = -PV + \mu N$ may be expressed by,

$$F = -PV + \mu N = NkT \left(\ln(z) - \frac{f_{5/2}(z)}{f_{3/2}(z)} \right) \quad (8.10)$$

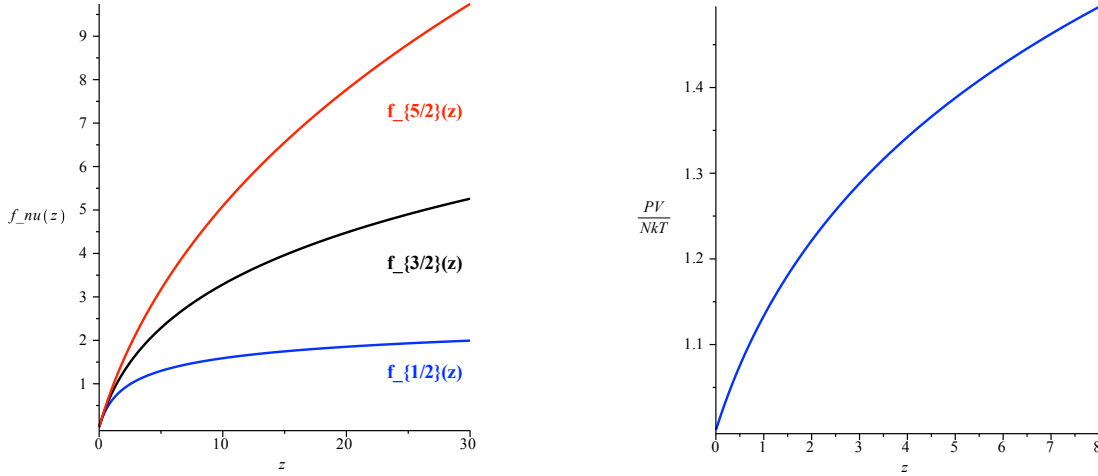


Figure 8: Left panel: numerical plot of the Fermi-Dirac functions $f_{1/2}(z)$, $f_{3/2}(z)$, $f_{5/2}(z)$. Right panel: the equation of state relation in terms of z .

from which the entropy is obtained by $S = (E - F)/T$,

$$S = Nk \left(\frac{5}{2} \frac{f_{5/2}(z)}{f_{3/2}(z)} - \ln(z) \right) \quad (8.11)$$

To compute the specific heat C_V (which is defined at constant V and N), the change in the fugacity with temperature (as N is being kept constant) must be taken into account,

$$\frac{1}{z} \left(\frac{\partial z}{\partial T} \right)_V = -\frac{3}{2T} \frac{f_{3/2}(z)}{f_{1/2}(z)} \quad (8.12)$$

Re-expressing E in terms of PV , and PV in terms of f using (8.5), we find,

$$\frac{C_V}{Nk} = \frac{15}{4} \frac{f_{5/2}(z)}{f_{3/2}(z)} - \frac{9}{4} \frac{f_{3/2}(z)}{f_{1/2}(z)} \quad (8.13)$$

The equation of state is obtained by forming the combination,

$$\frac{PV}{NkT} = \frac{f_{5/2}(z)}{f_{3/2}(z)} \quad (8.14)$$

This function is plotted against z in Figure 8.

8.3 Corrections to the Boltzmann gas

In the regime of *low density* compared to the thermal wavelength scale,

$$N\lambda^3/V = g f_{3/2}(z) \ll 1 \quad (8.15)$$

we have $z \ll 1$ (see Figure 8 for the numerical behavior of $f_{3/2}(z)$), and we recover the Boltzmann gas with,

$$z \approx \frac{N\lambda^3}{gV} \quad f_\nu(z) \approx z \quad (8.16)$$

In this limit the equation of state reduces to $PV = NkT$, and the internal energy, entropy and specific heat are given by their Boltzmann expressions, with $E = 3PV/2$ as well as,

$$E = \frac{3}{2}kT \quad S = \frac{5}{2}Nk - Nk \ln \left(\frac{N\lambda^3}{gV} \right) \quad C_V = \frac{3}{2}Nk \quad (8.17)$$

Using the expansion for small z of (8.7), we easily obtain the leading corrections to the Boltzmann expressions,

$$\begin{aligned} PV &= NkT \left(1 + \frac{N\lambda^3}{4\sqrt{2}gV} \right) \\ S &= Nk \left(\frac{5}{2} + \frac{N\lambda^3}{8\sqrt{2}gV} \right) - Nk \ln \left(\frac{N\lambda^3}{gV} \right) \\ C_V &= \frac{3}{2}Nk \left(1 - \frac{N\lambda^3}{8\sqrt{2}gV} \right) \end{aligned} \quad (8.18)$$

The increase in the pressure, at fixed temperature, is the result of the exclusion principle.

8.4 Zero temperature behavior

The Fermi-Dirac distribution exhibits a drastic simplification in the limit of low temperature, clearly apparent from Figure 9. In the limit $T = 0$, the distribution is a step function,

$$\lim_{T \rightarrow 0^+} \frac{1}{e^{\beta(\varepsilon - \mu)} + 1} = \begin{cases} 1 & \text{for } \varepsilon < \mu \\ 0 & \text{for } \varepsilon > \mu \end{cases} \quad (8.19)$$

The chemical potential $\mu = \varepsilon_F$ of the system at $T = 0$ is referred to as the *Fermi energy*. The Pauli exclusion principle underlies the physical interpretation of the distribution: as the lowest energy quantum states are filled up, additional particles must be placed in successively higher energy states, up till the Fermi energy. The sharp cut-off is the result of having exactly zero temperature. The corresponding surface of momenta \mathbf{p} satisfying the relation,

$$\varepsilon_F = \frac{\mathbf{p}^2}{2m} \quad p_F = |\mathbf{p}| \quad (8.20)$$

is referred to as the *Fermi surface*. The magnitude p_F of the momenta lying on the Fermi surface is constant, and is referred to as the *Fermi momentum*. Strictly at $T = 0$, the various thermodynamic functions may be evaluated in terms of the parameters m, V, ε_F , and we find,

$$\begin{aligned} N &= \frac{gV}{(2\pi\hbar)^3} \int_0^{p_F} 4\pi p^2 dp = \frac{4\pi gV p_F^3}{3(2\pi\hbar)^3} \\ E &= \frac{gV}{(2\pi\hbar)^3} \int_0^{p_F} 4\pi p^2 dp \frac{p^2}{2m} = \frac{4\pi gV p_F^5}{10m(2\pi\hbar)^3} \end{aligned} \quad (8.21)$$

Energy density and pressure are found to be,

$$\frac{E}{N} = \frac{3}{5}\varepsilon_F \quad P = \frac{2}{5}\frac{N}{V}\varepsilon_F \quad (8.22)$$

Using the expression for the free energy $F = E - TS = -PV + \mu N$, the entropy is found to vanish. This is in accord with the fact that the system at $T = 0$ is expected to be in its (unique) microscopic ground state.

Immediate applications of the low temperature thermodynamics of a gas of fermions is to massive gravitational bodies in which the pressure generated by the exclusion principle is compensated by the force of gravity at equilibrium. This is the case of neutron stars, and of the hypothetical quark stars and strange stars. Later on, we shall discuss in some detail the case of white dwarfs where electrons are the key players.

8.5 Low temperature behavior: The Sommerfeld expansion

For low temperatures, the Fermi-Dirac distribution is close to a step function, but shows that some states with energies less than the Fermi energy ε_F are empty, while some states with energies above ε_F are filled. The spread in these energies is set by the temperature kT , as is manifest from Figure 9. In practice, it is only the contribution of these degrees of freedom that will be important in the low temperature thermodynamics of the system. In particular, the number of degrees of freedom contributing to the specific heat should decrease as T decreases, and should no longer be given by the classical constant value $C_V \approx Nk$. Indeed, we shall find that $C_V \sim T$ for small T .

A systematic expansion for the thermodynamic functions at low but finite temperature is somewhat delicate. It corresponds to a regime where $z \gg 1$, and is provided by the *Sommerfeld expansion*. Following Pathria, we set $z = e^\xi$ with $\xi = \beta\mu$, and begin by evaluating a generalized form of the Fermi-Dirac functions,

$$f[\phi] = \int_0^\infty dx \frac{\phi(x)}{e^{x-\xi} + 1} \quad z = e^\xi \quad (8.23)$$

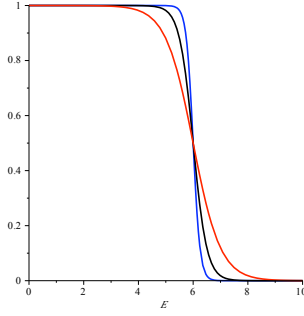


Figure 9: Low temperature behavior of the Fermi-Dirac occupation number for $\mu = 6$, and $kT = 0.125, 0.25$, and 1 respectively in blue, black, and red.

for any function $\phi(x)$ with polynomial growth in $|x|$. By separating the integration region at the point $x = \xi$, we readily obtain (after suitable shifts of variables in each region),

$$f[\phi] = \int_0^\xi dx \phi(x) + \int_0^\infty dx \frac{\phi(\xi + x) - \phi(\xi - x)}{e^x + 1} + \int_\xi^\infty dx \frac{\phi(x)}{e^x + 1} \quad (8.24)$$

The first two terms produce power law dependences in large ξ , while the last term is suppressed by exponentials $\mathcal{O}(e^{-\xi})$, and may be neglected compared to the first two. Expanding $\phi(\xi \pm x)$ in a power series in x around $x = 0$, and using the formula, for the value $\nu = 2n + 2$, we have the following final formula,

$$F[\phi] = \int_0^\xi dx \phi(x) + \sum_{n=0}^{\infty} 2 \left(1 - \frac{1}{2^{2n+1}}\right) \zeta(2n + 2) \phi^{(2n+1)}(\xi) + \mathcal{O}(e^{-\xi}) \quad (8.25)$$

or, for the first few orders, we have,

$$F[\phi] = \int_0^\xi dx \phi(x) + \frac{\pi^2}{6} \phi^{(1)}(\xi) + \frac{7\pi^4}{360} \phi^{(3)}(\xi) + \dots \quad (8.26)$$

where $\phi^{(2n+1)}(\xi)$ denotes the derivative of order $2n + 1$ evaluated at the point ξ . Using this formula with $\phi(x) = x^{\nu-1}/\Gamma(\nu)$, we find,

$$f_\nu(z) = \frac{\xi^\nu}{\Gamma(\nu + 1)} \left(1 + \nu(\nu - 1) \frac{\pi^2}{6\xi^2} + \nu(\nu - 1)(\nu - 2)(\nu - 3) \frac{7\pi^4}{360\xi^4} + \dots\right) + \mathcal{O}(e^{-\xi}) \quad (8.27)$$

The functions we need are given as follows,

$$f_{1/2}(z) = \frac{\xi^{1/2}}{\Gamma(3/2)} \left(1 - \frac{\pi^2}{24\xi^2} - \frac{7\pi^4}{384\xi^4} + \dots\right) + \mathcal{O}(e^{-\xi})$$

$$\begin{aligned}
f_{3/2}(z) &= \frac{\xi^{3/2}}{\Gamma(5/2)} \left(1 + \frac{\pi^2}{8\xi^2} + \frac{7\pi^4}{640\xi^4} + \dots \right) + \mathcal{O}(e^{-\xi}) \\
f_{5/2}(z) &= \frac{\xi^{5/2}}{\Gamma(7/2)} \left(1 + \frac{5\pi^2}{8\xi^2} - \frac{7\pi^4}{384\xi^4} + \dots \right) + \mathcal{O}(e^{-\xi})
\end{aligned} \tag{8.28}$$

Recall that this expansion holds in $\xi = \beta\mu = \mu/kT$ for large ξ , i.e. for small T at fixed μ .

Corrections to various thermodynamic functions may now be evaluated with the help of the general formulas (8.5), (8.11), and (8.13). We begin by expressing the chemical potential μ at finite temperature in terms of the Fermi energy ε_F , by comparing the formulas,

$$\begin{aligned}
\frac{N}{V} &= \frac{4\pi g(2m\varepsilon_F)^{3/2}}{3(2\pi\hbar)^3} \\
\frac{N}{V} &= \frac{g}{\lambda^3} f_{3/2}(z) = \frac{g}{\lambda^3 \Gamma(5/2)} \left(\frac{\mu}{kT} \right)^{3/2} \left(1 + \frac{\pi^2}{8\xi^2} \right)
\end{aligned} \tag{8.29}$$

We find to this order of approximation,

$$\varepsilon_F = \mu \left(1 + \frac{\pi^2}{12\xi^2} \right) \qquad \mu = \varepsilon_F \left(1 - \frac{\pi^2}{12} \left(\frac{kT}{\varepsilon_F} \right)^2 \right) \tag{8.30}$$

In terms of T and ε_F , the energy per particle, entropy, and specific heat are given by,

$$\begin{aligned}
\frac{E}{N} &= \frac{3}{5} \varepsilon_F \left(1 + \frac{5\pi^2}{12} \left(\frac{kT}{\varepsilon_F} \right)^2 \right) \\
\frac{S}{Nk} &= \frac{C_V}{Nk} = \frac{\pi^2 kT}{2\varepsilon_F}
\end{aligned} \tag{8.31}$$

8.6 Pauli paramagnetism of ideal gasses

Spin 1/2 particles have intrinsic magnetic dipole moments, and therefore produce paramagnetic effects in a gas or liquid. An external magnetic field splits each energy level in two, with effective energies,

$$\varepsilon = \frac{\mathbf{p}^2}{2m} \pm \kappa B \tag{8.32}$$

where $\kappa = e\hbar/2mc$ is the Bohr magneton. Since the shifts $\pm\kappa B$ enter in combination with the chemical potential for the spin 1/2 particles, the total Gibbs free energy is given by,

$$G(T, V, \mu, B) = G_0(T, V, \mu + \kappa B) + G_0(T, V, \mu - \kappa B) \tag{8.33}$$

where G_0 is the Gibbs free energy of the simplest spinless ideal gas, and is given by,

$$G_0(T, V, \mu) = -\frac{kTV}{(2\pi\hbar)^3} \int d^3\mathbf{p} \ln \left(1 + e^{-\beta(\mathbf{p}^2/2m - \mu)} \right) = -\frac{kTV}{\lambda^3} f_{5/2}(e^{\beta\mu}) \quad (8.34)$$

The paramagnetic susceptibility, per unit volume and at zero external magnetic field, is then obtained as follows,

$$\chi_{\text{para}} = -\frac{1}{V} \left. \frac{\partial^2 G}{\partial B^2} \right|_{B=0} = -2 \frac{\kappa^2}{V} \frac{\partial^2 G_0}{\partial \mu^2} = 2 \frac{\kappa^2}{V} \left. \frac{\partial N}{\partial \mu} \right|_{T,V} \quad (8.35)$$

At zero temperature, we may use the first relation in (8.21) for $g = 1$, and expressed in terms of the chemical potential,

$$N = V \frac{(2m\mu)^{3/2}}{3\pi^2\hbar^3} \quad (8.36)$$

and we find,

$$\chi_{\text{para}} = \frac{\kappa^2 m p_F}{\pi^2 \hbar^3} = \frac{3}{2} \frac{\kappa^2}{\varepsilon_F} \frac{N}{V} \quad (8.37)$$

Finite temperature corrections may be deduced using the full expression for G_0 . At high temperatures, we recover, of course, the Curie-Weiss law already derived using Boltzmann statistics in section 6.

8.7 Landau diamagnetism

The energy levels of a charged particle in the presence of a constant magnetic field are arranged into *Landau levels*, labelled by an integer $n = 0, 1, 2, \dots$, and given by,

$$\varepsilon(p_z, n) = \frac{p_z^2}{2m} + \frac{eB\hbar}{mc} \left(n + \frac{1}{2} \right) \quad (8.38)$$

where p_z is the momentum along the direction of the magnetic field. Particles with a magnetic moment require an extra term κB term, but at small magnetic fields, the case we are most interested in here, this effect is paramagnetic and may be treated independently along the lines of the preceding section.

Some care is needed in correctly normalizing the summation over Landau levels. For weak magnetic fields, the Landau levels are closely spaced and degenerate into the continuous p_x, p_y spectrum at $B = 0$. We may use this correspondence to normalize the sum, as follows. The

number of quantum states between two successive Landau energy levels $\varepsilon(p_z, n + 1)$ and $\varepsilon(p_z, n)$ is $eB\hbar/mc$; computing the same from the classical measure, we find,

$$\int_{\varepsilon(p_z, n)}^{\varepsilon(p_z, n+1)} \frac{dx dy dp_x dp_y}{(2\pi\hbar)^2} = \frac{L_x L_y eB}{2\pi\hbar c} \quad (8.39)$$

where L_x and L_y are the linear dimensions of a square box in the plane perpendicular to \mathbf{B} . This factor provides the correct normalization for the calculation of the grand canonical partition function, actually for all values of the field B ,

$$G = -kT \ln Z = -kT \sum_{n=0}^{\infty} \frac{L_x L_y eB}{2\pi\hbar c} \int \frac{dz dp_z}{2\pi\hbar} \ln \left(1 + e^{-\beta(\varepsilon(p_z, n) - \mu)} \right) \quad (8.40)$$

Since we are interested in weak magnetic fields, $eB\hbar/mc \ll kT$, we may use an approximate evaluation of the sum over n , by using the Euler-McLaurin formula to first order,

$$\sum_{n=0}^{\infty} f\left(n + \frac{1}{2}\right) = \int_0^{\infty} dx f(x) + \frac{1}{24} f'(0) \quad (8.41)$$

up to higher derivative corrections (which will be accompanied by higher powers of B). The contribution of the integral is proportional to

$$eB \int_0^{\infty} dx \ln \left(1 + e^{-\beta(p_z^2/2m + eB\hbar/mcx - \mu)} \right) \quad (8.42)$$

By changing integration variables $eBx \rightarrow x$, we see that the integral is in fact independent of B , and for our purposes of studying magnetic properties, will be immaterial. We abbreviate its contribution to G as $G_{B=0}$. The remaining contribution is readily evaluated, and we find,

$$G = G_{B=0} + \frac{1}{24} \frac{V e^2 B^2 \hbar}{(2\pi\hbar)^2 m c^2} \int dp_z \frac{1}{e^{\beta(p_z^2/2m - \mu)} + 1} \quad (8.43)$$

Changing variables $p_z = \sqrt{2mkT} \sqrt{x}$ and $z = e^{\beta\mu}$, we may express the result in terms of a Fermi-Dirac function, and the Bohr magneton $\kappa = e\hbar/2mc$,

$$G = G_{B=0} + \frac{mV}{24\pi^2 \hbar^3} \kappa^2 B^2 (2mkT)^{\frac{1}{2}} \sqrt{\pi} f_{1/2}(z) \quad (8.44)$$

The diamagnetic susceptibility per unit volume is defined by,

$$\chi_{\text{diam}} = -\frac{1}{V} \left. \frac{\partial^2 G}{\partial B^2} \right|_{T, \mu} \quad (8.45)$$

and is found to be given by

$$\chi_{\text{diam}} = -\frac{m}{12\pi^2 \hbar^3} \kappa^2 (2mkT)^{\frac{1}{2}} \sqrt{\pi} f_{1/2}(z) \quad (8.46)$$

For the Boltzmann regime, with $z \ll 1$, and $z \approx N\lambda^3/V$, we get,

$$\chi_{\text{diam}} \approx -\frac{N}{V} \frac{\kappa^2}{3kT} \quad (8.47)$$

In the low temperature regime where $z \gg 1$, we have $f_{1/2}(z) \approx 2(\mu/\pi kT)^{1/2}$, and we find,

$$\chi_{\text{diam}} = -\frac{N\kappa^2}{4V\varepsilon_F} \quad (8.48)$$

Assembling the contributions from the paramagnetic and diamagnetic parts, we find for large temperature,

$$\chi = \chi_{\text{para}} + \chi_{\text{diam}} = \frac{N}{VkT} \left(\kappa^2 - \frac{1}{3}\kappa^2 \right) = \frac{2N\kappa^2}{3VkT} \quad (8.49)$$

Thus, at high temperatures, gasses tend to be paramagnetic. The same is generally true at low temperatures.

8.8 White dwarfs

The Pauli exclusion principle prevents electrons in atoms from collapsing into the lowest atomic quantum state, and it also prevents certain types of large astronomical bodies from collapsing under the forces of their gravitational attraction. The oldest system where this mechanism was applied is the case of *white dwarfs*. More recent systems include neutron stars, and possibly quark stars and strange stars. Here, we shall discuss in detail only the case of white dwarfs.

The energy radiated by the sun is generated by the nuclear reactions of converting Hydrogen into Deuterium and then into Helium. The reaction pretty much ends when this conversion is complete, as the ${}^4\text{He}$ nucleus has larger binding energy per nucleon than its three subsequent elements, Lithium ${}^7\text{Li}$, Beryllium ${}^9\text{Be}$ and Boron ${}^{11}\text{B}$. At that point the star runs out of thermonuclear fuel, and may be expected to collapse under its gravitational weight. But this is not what happens, at least for certain sufficiently small stars, referred to as *white dwarfs*. They remain stable in size and mass, and glow relatively faintly. The typical numbers are as follows,

$$\begin{aligned} M &\approx 10^{33} \text{ g} \\ \rho &\approx 10^7 \text{ g/cm}^3 \\ T &\approx 10^7 \text{ K} \sim 10^3 \text{ eV} \end{aligned} \quad (8.50)$$

where M is the total mass of star, ρ its mass density, and T its temperature. At this temperature, all ${}^4\text{He}$ is completely ionized, so we have a plasma of N electrons and $N/2$ Helium nuclei. The relations between the total mass M and volume V are as follows,

$$M = N(m + 2m_p) \approx 2m_p N \qquad \frac{N}{V} = \frac{\rho}{2m_p} \quad (8.51)$$

The electron density N/V allows us to compute the Fermi momentum by,

$$\frac{N}{V} = \frac{2p_F^3}{3\pi^2\hbar^3} = \frac{\rho}{2m_p} \quad (8.52)$$

The ratio of the Fermi momentum to mc is given by,

$$\frac{p_F}{mc} = \left(\frac{3\rho}{16\pi m_p} \right)^{1/3} \frac{2\pi\hbar}{mc} \quad (8.53)$$

Substituting the above data for the characteristics of a white dwarf, we get $p_F/mc \sim 100$, so that the electrons are in fact relativistic. The Fermi energy and temperature $T_F = \varepsilon_F/k$ are respectively found to be $\varepsilon_F \sim 10^6$ eV, and $T_F \approx 10^{10}$ K. Thus the actual temperature of the white dwarfs is small compared to the Fermi temperature, and the electron gas is fully degenerate. The electrons produce the largest contribution to the pressure against gravitation, and so we shall neglect the contribution of the nuclei.

To study the equilibrium conditions quantitatively, we compute the pressure of this relativistic gas of electrons. Using the Gibbs ensemble, we have,

$$P = \frac{gkT}{(2\pi\hbar)^3} \int d^3\mathbf{p} \ln \left(1 + e^{-\beta(\varepsilon(p)-\mu)} \right) \quad (8.54)$$

where the relativistic energy is given by,⁴

$$\varepsilon(p) = \sqrt{p^2c^2 + m^2c^4} \quad (8.55)$$

In the degenerate limit, β is large and positive, and the logarithmic term vanishes when $\varepsilon(p) > \mu$. The contribution for $\varepsilon(p) < \mu$ is dominated by the exponential part of the argument, so that we obtain,

$$P = \frac{g}{(2\pi\hbar)^3} \int d^3\mathbf{p} (\mu - \varepsilon(p)) \theta(\mu - \varepsilon(p)) \quad (8.56)$$

⁴Equivalently, one could use for $\varepsilon(p)$ only the relativistic kinetic energy; the difference amounts to a shift of μ by the rest energy mc^2 .

Carrying out the angular integration, and changing variables to

$$p = mc \sinh \psi \qquad p_F = mc \sinh \psi_F \qquad \mu = \varepsilon_F = \varepsilon(p_F) \qquad (8.57)$$

we have,

$$P = \frac{4\pi g m^4 c^5}{(2\pi\hbar)^3} \int_0^{\psi_F} d\psi \cosh \psi \sinh^2 \psi (\cosh \psi_F - \cosh \psi) \qquad (8.58)$$

The integral is readily computed, and we find,

$$P = \frac{\pi g m^4 c^5}{6(2\pi\hbar)^3} A\left(\frac{p_F}{mc}\right) \qquad (8.59)$$

where the function A is given by,

$$A(x) = x(2x^2 - 3)\sqrt{1+x^2} + 3 \ln(x + \sqrt{x^2 + 1}) \qquad (8.60)$$

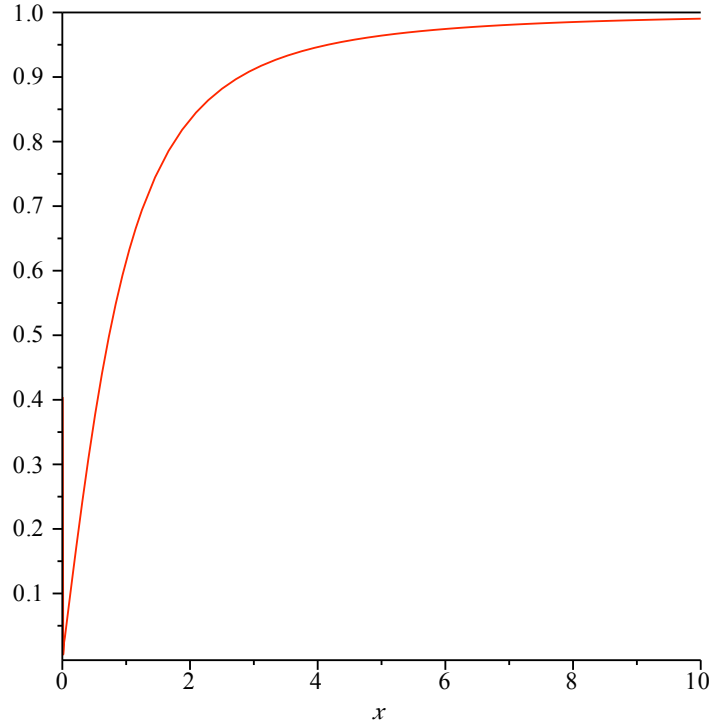


Figure 10: The function $A(x)/2x^4$ versus x .

Next, we need to balance the pressure of the electron gas against the gravitational pull on them. In terms of the work done by each, we have,

$$\begin{aligned}\delta E_P &= -P 4\pi R^2 dR \\ \delta E_g &= \frac{\alpha}{4\pi} \frac{GM^2}{R^2} dR\end{aligned}\tag{8.61}$$

Here, α is a *fudge factor* of order unity introduced in order to account for the fact that the mass distribution in a real star will not be uniform. At equilibrium, these two must balance one another, so we must have,

$$P = \frac{\alpha}{4\pi} \frac{GM^2}{R^4}\tag{8.62}$$

We now relate the Fermi momentum to the radius, by using the formula for the mass density $\rho = 3M/4\pi R^3$, and we find,

$$x = \frac{p_F}{mc} = \frac{R_0}{R} \qquad R_0 = \frac{2\pi\hbar}{mc} \left(\frac{9M}{64\pi^2 m_p} \right)^{1/3}\tag{8.63}$$

The equilibrium relation then takes the form,

$$\frac{A(x)}{x^4} = 6\pi\alpha \frac{G}{\hbar c} M^2 \left(\frac{8m_p}{9\pi M} \right)^{4/3}\tag{8.64}$$

or after some simplifications,

$$M = \frac{9}{64m_p^2} \left(\frac{3\pi}{\alpha^3} \right)^{1/2} \left(\frac{\hbar c}{G} \right)^{3/2} \left(\frac{A(x)}{2x^4} \right)^{3/2}\tag{8.65}$$

Now, it is easy to see, for example from Figure 10, that the function $A(x)$ satisfies,

$$A(x) \leq 2x^4\tag{8.66}$$

for all $0 < x < \infty$. As a result, there is a maximum mass that a white dwarf can have, whose value is given by the *Chandrasekar limit*, which to our approximation is given by,

$$M_c = \frac{9}{64m_p^2} \left(\frac{3\pi}{\alpha^3} \right)^{1/2} \left(\frac{\hbar c}{G} \right)^{3/2}\tag{8.67}$$

In practice, the limiting mass is not much larger than the mass of the sun: $M_c \approx 1.44M_\odot$. We conclude that our sun will end its life as a white dwarf.

9 Bose-Einstein statistics

In this section, we shall discuss ideal gasses of particles obeying Bose-Einstein statistics. Applications to Planck's black body radiation, microwave background radiation, and Bose-Einstein condensation will be given.

Recall the basic equations for an ideal BR gas from equations (7.26), (7.27), and (7.28) are given by,

$$\begin{aligned} N &= \sum_i \frac{1}{e^{\beta(\varepsilon_i - \mu)} - 1} \\ E &= \sum_i \frac{\varepsilon_i}{e^{\beta(\varepsilon_i - \mu)} - 1} \\ -G = PV = kT \ln Z &= -kT \sum_i \ln \left(1 - e^{-\beta(\varepsilon_i - \mu)} \right) \end{aligned} \quad (9.1)$$

where we take the sum over all states labelled by i , including their degeneracies, and corresponding energy ε_i . We shall normalize the ground state energy to zero, so that the Bose-Einstein distribution requires,

$$\mu \leq 0 \quad (9.2)$$

Here, Z stands for the grand canonical partition function. The Gibbs free energy G is related to these quantities by $G = -PV$. As we are assuming that the gas is ideal, no mutual interactions take place; the Hamiltonian is just that of the free relativistic or non-relativistic particle, and ε_i are the free particle energies.

9.1 Black body radiation

Black body radiation refers to a system of photons which are in equilibrium at temperature T with some weakly interacting quantity of matter, usually the material walls of the box or oven in which the photons are being observed. Photons without matter present interact too weakly (through scattering of light by light) to achieve equilibrium. It is key to observe that *the number of photons N is not conserved*. Hence, the chemical potential associated with the number of photons must vanish. In other words, there is no potential energy associated with the *number* of photons created, only with their total kinetic energy. Thus, in the above formulas, we must set

$$\mu = 0 \quad (9.3)$$

and we will really be dealing with the canonical ensemble, where G is actually the Helmholtz free energy. Finally, for photons, the energies are given by,

$$\varepsilon = \hbar\omega \quad (9.4)$$

where $\omega = |\mathbf{k}|c$. Since a photon has two degrees of polarization, we have $g = 2$, and obtain the following formulas for energy and pressure,

$$\begin{aligned} E &= \frac{2V}{(2\pi\hbar)^3} \int d^3\mathbf{p} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \\ PV &= -kT \frac{2V}{(2\pi\hbar)^3} \int d^3\mathbf{p} \ln(1 - e^{-\beta\hbar\omega}) \end{aligned} \quad (9.5)$$

Expressing momentum $\mathbf{p} = \hbar\mathbf{k}$, performing the angular integration, and then transforming $|\mathbf{k}|$ into frequency ω , we have,

$$\frac{2V d^3\mathbf{p}}{(2\pi\hbar)^3} \rightarrow \frac{V\omega^2 d\omega}{\pi^2 c^3} \quad (9.6)$$

The energy density and pressure distribution may be,

$$\begin{aligned} \frac{E}{V} &= \int_0^\infty \frac{\omega^2 d\omega}{\pi^2 c^3} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \\ P &= -kT \int_0^\infty \frac{\omega^2 d\omega}{\pi^2 c^3} \ln(1 - e^{-\beta\hbar\omega}) \end{aligned} \quad (9.7)$$

The famous Planck black body frequency distribution curve is shown in Figure 11. The maximum in the energy distribution function is attained at the maximum of the function $\omega^3/(e^{\beta\hbar\omega} - 1)$, which is given approximately by $\hbar\omega_{\max} = 2.822 kT$.

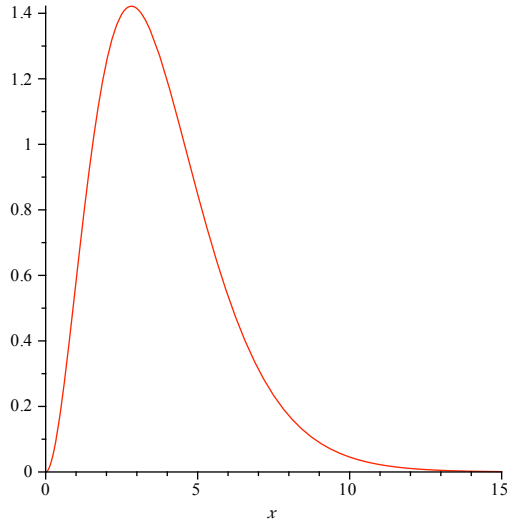


Figure 11: The normalized energy distribution function $x^3/(e^x - 1)$ versus $x = \hbar\omega/kT$.

The temperature dependence may be easily extracted by changing integration variables, $\beta\hbar\omega = x$. The integrals may be performed by expanding the integrands in powers of the exponentials e^{-x} , and we find,

$$\begin{aligned} \frac{E}{V} &= \frac{\pi^2(kT)^4}{15c^3\hbar^3} & \int_0^\infty \frac{dx x^3}{e^x - 1} &= 6\zeta(4) \\ P &= \frac{\pi^2(kT)^4}{45c^3\hbar^3} & \int_0^\infty dx x^2 \ln(1 - e^{-x}) &= -2\zeta(4) \end{aligned} \quad (9.8)$$

where $\zeta(z)$ is the Riemann zeta function, and we have $\zeta(4) = \pi^4/90$. It readily follows that internal energy and pressure are related by,

$$E = 3PV \quad (9.9)$$

The Helmholtz free energy is given by $F = E - TS = -PV$ (since $\mu = 0$), we can also compute the entropy, and we find,

$$\frac{S}{V} = \frac{4\pi^2(kT)^3}{45c^3\hbar^3} \quad (9.10)$$

One immediate application is to the Stefan-Boltzmann law of black body radiation. A small hole made in the vessel that contains the black body radiation will let a small amount of the radiation out. While the system is now not strictly at equilibrium any more, it is sufficiently close to equilibrium that we can continue to use equilibrium formulas. In particular, the rate of radiation energy flow per unit area, is given by,

$$\frac{c E}{4V} = \sigma T^4 \quad \sigma = \frac{\pi^2 k^4}{60c^3\hbar^3} \quad (9.11)$$

where σ is the Stefan-Boltzmann constant.

9.2 Cosmic micro-wave background radiation

The Cosmic Micro-wave Background (CMB) radiation was predicted by Robert Dicke on theoretical grounds and discovered experimentally, by accident, by Penzias and Wilson, who were awarded the Nobel Prize in 1978. The universe is permeated by an electro-magnetic radiation, obeying the Planck black body radiation spectral distribution to 1 part in 10^5 , and corresponding to a temperature of $T_{\text{CMB}} = 2.725$ K. This is a very low temperature compared to any nuclear or atom energy scales. Where does it come from ?

Assuming the Big Bang took place, then early in the evolution of the Universe, various particle species are all interacting relatively strongly, which allows for local equilibrium to occur. As the Universe cooled down, one crosses the *combination* temperature T_{comb} . Above

T_{comb} , nucleons and electrons form a plasma which interacts with the photons, and below which ionized nuclei and electrons *combine*. The typical ionization energy of atoms and molecules is on the order of $T_{\text{comb}} \sim 1\text{eV} \sim 10^4$ K. At temperatures below T_{comb} , photons do not directly interact with charged particles any more, but rather with electrically neutral atoms and molecules. As a result, below T_{comb} , photons are not in equilibrium any more with atoms and molecules, and *decouple*. So, below T_{comb} , photons remain as an isolated free gas.

The fact that we observe today CMB radiation at a much lower temperature is explained by the expansion of the universe. Assuming that this expansion is *adiabatic*, namely entropy is conserved, then we can read off the relation between temperature and volume of the universe from formula (9.10). Thus, we know how the volume V of the present Universe is related to the volume V_{comb} of the Universe at the time of combination:

$$\frac{V}{V_{\text{comb}}} = \frac{T_{\text{comb}}^3}{T_{\text{CMB}}^3} \quad (9.12)$$

Given the temperatures, this number evaluates to 5×10^{10} . The corresponding red-shift factor is $\gamma \sim 3000$, as may be deduced by comparing the average photon energies.

9.3 Thermodynamic variables for Bose-Einstein ideal gasses

We shall now consider an ideal gas of non-relativistic particles obeying BE statistics, for which particle number is conserved (unlike in the case of photons, where photon number was not conserved). Using the standard energy-momentum relation $\varepsilon(\mathbf{p}) = \mathbf{p}^2/2m$, it is clear that the ground state energy vanishes, so that the chemical potential needs to obey $\mu \leq 0$. As a result, the fugacity $z = e^{\beta\mu}$ obeys $z \leq 1$, and the Bose-Einstein functions,

$$g_\nu(z) \equiv \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{dx x^{\nu-1}}{z^{-1}e^x - 1} \quad (9.13)$$

are completely described by the series expansion,

$$g_\nu(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^\nu} \quad (9.14)$$

since this series converges absolutely for $|z| \leq 1$ as long as $\nu > 1$.

Special care needs to be exerted in translating the discrete quantum spectrum of a BE gas in a box of finite size into formulas for continuous distributions as V becomes large. The reason is that the ground state of the 1-particle Hamiltonian, namely for $\mathbf{p} = 0$, can be occupied by a small or by a macroscopically large number of particles, without extra cost in

energy. This can happen when z is very close to 1, since then the occupation number of the $\varepsilon = 0$ state can become large,

$$\frac{1}{z^{-1}e^{\beta\varepsilon} - 1} = \frac{z}{1 - z} \quad \varepsilon = 0 \quad (9.15)$$

The quantum states at finite energy are not affected by this phenomenon, and the customary translation into continuum states may be applied there. For the internal energy, the zero energy states do not contribute, and we have,

$$\frac{E}{V} = \frac{4\pi}{(2\pi\hbar)^3} \int d^3\mathbf{p} \frac{\mathbf{p}^2/2m}{z^{-1}e^{\beta\mathbf{p}^2/2m} - 1} \quad (9.16)$$

Isolating the $\varepsilon = 0$ contribution to N and P gives the following formulas,

$$\begin{aligned} \frac{N}{V} &= \frac{1}{V} \frac{z}{1 - z} + \frac{4\pi}{(2\pi\hbar)^3} \int d^3\mathbf{p} \frac{1}{z^{-1}e^{\beta\mathbf{p}^2/2m} - 1} \\ P &= -\frac{kT}{V} \ln(1 - z) - \frac{4\pi kT}{(2\pi\hbar)^3} \int d^3\mathbf{p} \ln(1 - ze^{-\beta\mathbf{p}^2/2m}) \end{aligned} \quad (9.17)$$

Next, we change variables, to exhibit the number N_0 of particles in the ground state,

$$N_0 = \frac{z}{1 - z} \quad z = \frac{N_0}{N_0 + 1} \approx 1 - \frac{1}{N_0} \quad (9.18)$$

Assuming that N_0/V tends to a finite value in the thermodynamic limit, we see that the effects of N_0 survives only in the above formulas for the particle number, but drops out of the formula for the pressure, since we have, $V^{-1} \ln(1 - z) \approx V^{-1} \ln N_0 \rightarrow 0$ as $V \rightarrow \infty$ with N_0/V fixed. The final formulas for the total number of particles and pressure are thus,

$$\begin{aligned} N - N_0 &= \frac{4\pi V}{(2\pi\hbar)^3} \int d^3\mathbf{p} \frac{1}{z^{-1}e^{\beta\mathbf{p}^2/2m} - 1} \\ P &= -\frac{4\pi kT}{(2\pi\hbar)^3} \int d^3\mathbf{p} \ln(1 - ze^{-\beta\mathbf{p}^2/2m}) \end{aligned} \quad (9.19)$$

In terms of the BE functions, we obtain,

$$N - N_0 = \frac{V}{\lambda^3} g_{3/2}(z) \quad E = \frac{3}{2}PV \quad P = \frac{kT}{\lambda^3} g_{5/2}(z) \quad (9.20)$$

The equation of state is obtained by eliminating z between the relations

$$\frac{PV}{(N - N_0)kT} = \frac{g_{5/2}(z)}{g_{3/2}(z)} \quad (9.21)$$

and the first relation in (9.20).

9.4 Bose-Einstein condensation

The number $N_e = N - N_0$ of excited states, i.e. states with non-zero energy, is bounded from above by the fact that,

$$g_{3/2}(z) \leq g_{3/2}(1) = \zeta(3/2) \approx 2.612 \quad (9.22)$$

for all values of z in the physical range $0 < z \leq 1$, and we have,

$$N_e = N - N_0 \leq \zeta(3/2)V \left(\frac{mkT}{2\pi\hbar^2} \right)^{3/2} \quad (9.23)$$

As T is lowered, at fixed volume and fixed total number of particles N , the number of excited states N_e must decrease. At some critical temperature T_c , one is forced to populate the states at zero entropy *in a macroscopic way*, namely with N_0/V finite in the large V limit. This critical temperature is obtained by setting $N_0 = 0$, and we find,

$$T_c = \frac{2\pi\hbar^2}{mk} \left(\frac{N}{\zeta(3/2)V} \right)^{2/3} \quad (9.24)$$

To leading approximation, namely by setting $z = 1$, the number of excited and zero energy states are related as follows, for $T < T_c$,

$$\frac{N_0}{N} = 1 - \frac{N_e}{N} = 1 - \left(\frac{T}{T_c} \right)^{3/2} \quad (9.25)$$

9.5 Behavior of the specific heat

A calculation of the specific heat may be carried out in parallel to the one done for a FD gas in section 7.2. For BE gasses, however, we must be careful to carry out the calculation corresponding to the gas phase and the condensed phase separately. In either case, the starting point is,

$$C_V = \left. \frac{\partial E}{\partial T} \right|_V = \frac{3}{2}V \left. \frac{\partial P}{\partial T} \right|_V \quad P = \frac{kT}{\lambda^3} g_{5/2}(z) \quad (9.26)$$

In the gas phase, we have $z < 1$, while in the condensed phase, we set $z = 1$ in view of (9.18).

In the condensed phase, the formula for the pressure simplifies, and we have,

$$P = \frac{kT}{\lambda^3} g_{5/2}(1) \quad (9.27)$$

All reference to the volume has disappeared, so that the specific heat becomes,

$$\frac{C_V}{Nk} = \frac{15}{4} \frac{V}{N\lambda^3} g_{5/2}(1) \sim T^{3/2} \quad T < T_c \quad (9.28)$$

In the gas phase, the calculation will be similar to the one that led to (8.12) for the FD gas, and uses the fact that z depends on T . Thus we have,

$$\left. \frac{\partial P}{\partial T} \right|_V = \frac{5}{2} \frac{k}{\lambda^3} g_{5/2}(z) + \frac{kT}{\lambda^3} g_{3/2}(z) \left(\frac{\partial \ln z}{\partial T} \right)_V \quad (9.29)$$

The derivative of $\ln z$ is computed by differentiating equation $N\lambda^3 = Vg_{3/2}(z)$, and we find,

$$g_{1/2}(z) \left(\frac{\partial \ln z}{\partial T} \right)_V = -\frac{3}{2} \frac{N\lambda^3}{VT} \quad (9.30)$$

Putting all together, the specific heat in the gaseous phase is found to be,

$$\frac{C_V}{Nk} = \frac{15}{4} \frac{g_{5/2}(z)}{g_{3/2}(z)} - \frac{9}{4} \frac{g_{3/2}(z)}{g_{1/2}(z)} \quad T > T_c \quad (9.31)$$

As $z \rightarrow 1$, we have $g_{1/2}(z) \rightarrow \infty$, and so near the transition point, the second term cancels, and we find that the specific heat in the gas phase and in the condensed phase match at the transition point, since we have $N\lambda^3 = Vg_{3/2}(z)$. The slope at $T = T_c$ is discontinuous across the transition. This is a sign of a *phase transition*, as we shall study in detail in subsequent sections.

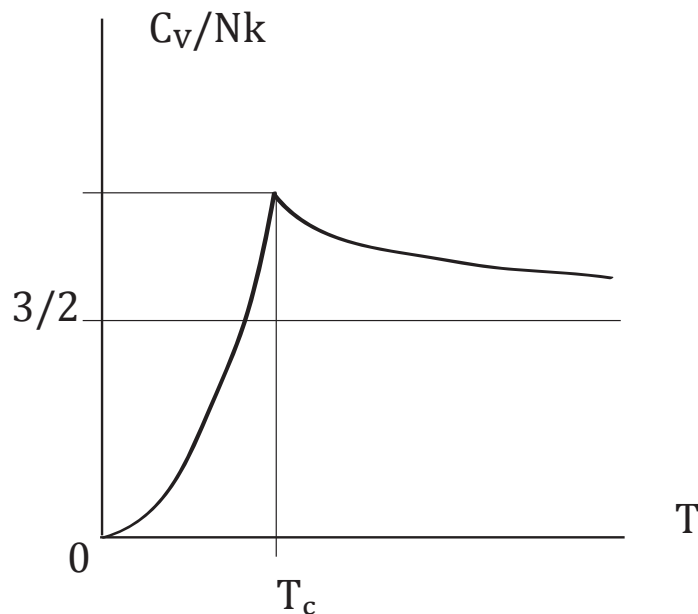


Figure 12: Molecular specific heat curve across the Bose-Einstein condensation transition.

10 Phase coexistence: thermodynamics

A given system or material can exhibit different physically distinct phases, as a function of the ranges of its macroscopic parameters. A common example is that of water, which may exist in its solid phase of ice, its liquid phase, or its vapor phase. Different phases may co-exist in equilibrium. The passage from one phase to another, as thermodynamic variables are changed, is referred to as a *phase transition*. Given the equations of state in both phases, thermodynamics gives an effective formulation of the conditions of equilibrium between different phases, and transitions between them. The underlying microscopic dynamics which is responsible for the system existing in one phase or another, or transitioning between one phase to another is, however, captured only by the application of statistical mechanics.

10.1 Conditions for phase equilibrium

We shall begin with a thermodynamic study of the equilibrium between gasses, liquids and solids, or any other forms of matter. If we have two systems, referred to as \mathcal{S}_1 and \mathcal{S}_2 , such as the two different phases of matter, in equilibrium with one another, then their temperature, pressure and chemical potentials must match,

$$\begin{aligned}T_2 &= T_1 \\P_2 &= P_1 \\ \mu_2 &= \mu_1\end{aligned}\tag{10.1}$$

Given these conditions, it will be most convenient to use not the variables E, V, N but rather with the intensive parameters T, P, μ in terms of which the equilibrium conditions are written. In either phase, these three variables are not independent, so we shall choose T, P and express the chemical potential for each phase as a function of T, P ,

$$\begin{aligned}\mu_1 &= \mu_1(T, P) \\ \mu_2 &= \mu_2(T, P)\end{aligned}\tag{10.2}$$

The chemical potentials μ_1 and μ_2 have different functional forms for the two phases. For given T, P it may, or may not, be possible to set the chemical potentials equal. Thus we conclude that

$$\begin{aligned}\text{if } \mu_1(T, P) &= \mu_2(T, P) && \text{then phases 1 and 2 can coexist at } T, P \\ \text{if } \mu_1(T, P) &\neq \mu_2(T, P) && \text{then phases 1 and 2 cannot coexist at } T, P\end{aligned}\tag{10.3}$$

Generally, the relation $\mu_1(T, P) = \mu_2(T, P)$ will hold on a line segment in the T, P plane, as shown in figure 13.

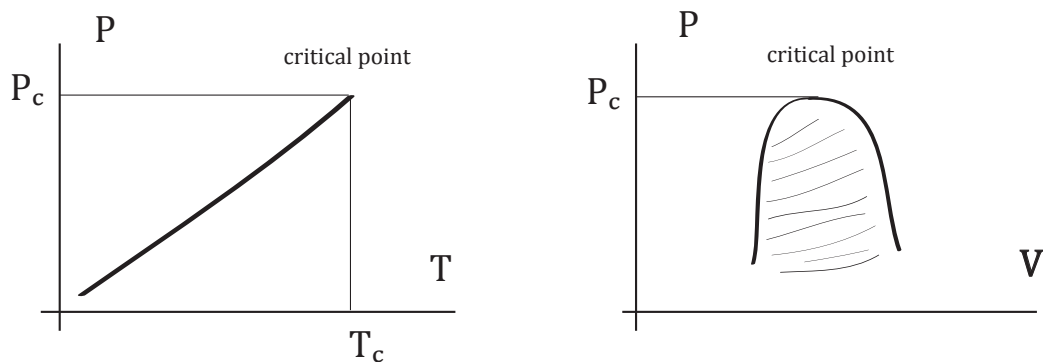


Figure 13: Phase coexistence curve in (T, P) and (V, P) variables ending at a critical point.

In terms of different thermodynamic variables, such as V, P , the phase diagrams takes on a completely different form. Above P_c , we read off from the T, P diagram that there is a unique system for each value of V, P , but this ceases to be the case for $P < P_c$. The two phases below P_c may coexist, but their densities are unequal, leaving the volume only partially determined. We shall analyze this phenomenon with the help of the Van der Waals equation of state.

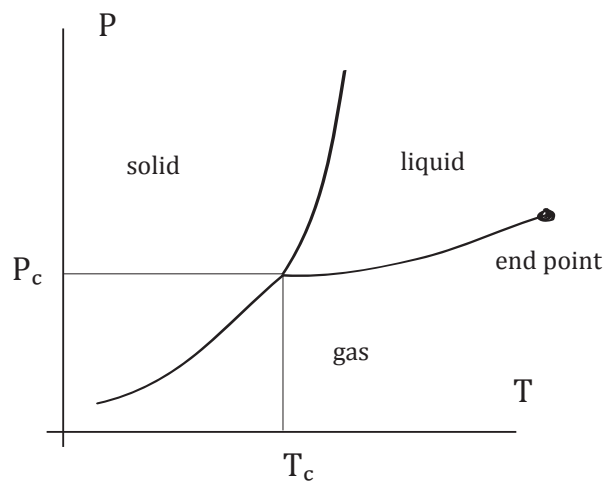


Figure 14: A gas-liquid-solid phase diagram exhibiting a triple point and an end point.

More than two phases can coexist pairwise at various temperatures and pressures. It is also possible for three phases to coexist at a point (generically). In fact this is the case of the standard gas-liquid-solid phase diagram, as show schematically in figure 14. At the triple

point, we have the relations,

$$\mu_{\text{gas}}(T, P) = \mu_{\text{liquid}}(T, P) = \mu_{\text{solid}}(T, P) \quad (10.4)$$

This phase diagram applies, for example, to water.

10.2 Latent heat

The intensive quantities different from T, P, μ are generally discontinuous across a phase separation lines. This include the molecular volume $v = V/N$ and the molecular entropy $s = S/N$. We shall assume that the substance of interest is composed out of a single material, such as water, and that molecules of this substance are transferred from one phase to another only. Thus, the total N is then conserved. Evaluating the relation

$$E = TS - PV + \mu N \quad (10.5)$$

on both phases, and taking into account that T, P, μ coincide for these phases at equilibrium, we find the relation,

$$E_2 - E_1 = T(S_2 - S_1) - P(V_2 - V_1) \quad (10.6)$$

By the first law of thermodynamics, we have

$$E_2 - E_1 + P(V_2 - V_1) = Q_2 - Q_1 = Q \quad (10.7)$$

namely the *latent heat* transfer between the phases. Combining both equations, we find that

$$Q = T(S_2 - S_1) \quad (10.8)$$

This relation shows that the process of phase change is *reversible*, upon release of absorption of the latent heat, since it agrees with the reversible relation $\delta Q = dS/T$ at constant T .

10.3 Clausius-Clapeyron equation

The Clausius-Clapeyron equation gives the shape of the coexistence curve in terms of the differences of molecular volume and entropy. It is a good exercise in thermodynamics gymnastics to work this out. Along the coexistence curve between two phases, we have,

$$\mu_1(T, P) = \mu_2(T, p) \quad (10.9)$$

and this relation determines the dependence between T, P along the coexistence curve. Differentiate this relation with respect to T ,

$$\frac{\partial \mu_1}{\partial T} + \frac{\partial \mu_1}{\partial P} \frac{dP}{dT} = \frac{\partial \mu_2}{\partial T} + \frac{\partial \mu_2}{\partial P} \frac{dP}{dT} \quad (10.10)$$

The quantity of interest to us is dP/dT , as this gives a differential equation for the coexistence curve. To evaluate the other quantities in the equation, we use T, P, N as independent thermodynamic variables, where N will in the end be fixed. The appropriate thermodynamic potential is

$$\Phi = E - TS + PV = \mu N \quad (10.11)$$

sometimes referred to as the Gibbs potential. Its differential is given by,

$$d\Phi = -SdT + VdP + \mu dN \quad (10.12)$$

but the last term not be important here as $dN = 0$. The relation may also be expressed in terms of μ using (10.5) or (10.11), and for constant N we find,

$$d\mu = -sdT + vdP \quad (10.13)$$

where again $s = S/N$ and $v = V/N$. The desired ingredients for (10.10) now immediately follows, and we have,

$$\left. \frac{\partial \mu}{\partial T} \right|_P = -s \quad \left. \frac{\partial \mu}{\partial P} \right|_T = v \quad (10.14)$$

so that,

$$\frac{dP}{dT} = \frac{s_2 - s_1}{v_2 - v_1} \quad (10.15)$$

This is the Clausius-Clapeyron equation.

10.4 Example of the Van der Waals gas-liquid transition

The Van der Waals equation of state,

$$\left(P + \frac{a}{V^2} \right) (V - b) = NkT \quad a, b > 0 \quad (10.16)$$

is rich enough to describe an interacting gas which can condense into a liquid phase. To study how this works, we work in T, P variables, the total number N of molecules being fixed. Given T, P , the equation for the volume V is a polynomial of third degree with real coefficients,

$$PV^3 - (bP + NkT)V^2 + aV - ab = 0 \quad (10.17)$$

The equation always has at least one real root, the additional two roots being either both real, or complex conjugates of one another. For sufficiently large P and T , we may approximate

the equation by $PV^3 - NkTV^2 = 0$ which reduces to the ideal gas equation of state, and the solution is unique. Thus, for T, P sufficiently large, we have two complex conjugate solutions in addition to the single real root. As T, P are lowered, all three real solutions will appear. This will happen when the two roots of the derivative equation coincide, a point we denote by V_c . But V_c must also satisfy the Van der Waals equation, so that the equation must be equivalent to $(V - V_c)^3 = 0$. Multiplying by P_c and expanding gives,

$$P_c V^3 - 3P_c V_c V^2 + 3P_c V_c^2 V - P_c V_c^3 = 0 \quad (10.18)$$

Identifying coefficients with the equation of (10.17) gives,

$$\begin{aligned} bP_c + NkT_c &= 3P_c V_c \\ a &= 3P_c V_c^2 \\ ab &= P_c V_c^3 \end{aligned} \quad (10.19)$$

which is solved uniquely as follows,

$$V_c = 3b \quad P_c = \frac{a}{27b^2} \quad NkT_c = \frac{8a}{27b} \quad (10.20)$$

Eliminating a, b from the Van der Waals equation in favor of T_c, P_c, V_c gives,

$$\left(\frac{P}{P_c} + 3 \frac{V_c^2}{V^2} \right) \left(3 \frac{V}{V_c} - 1 \right) = 8 \frac{T}{T_c} \quad (10.21)$$

Note that this curve is universal, and has no extra free parameters in terms of the normalized thermodynamic variables $T/T_c, P/P_c$, and V/V_c . A schematic plot in V, P variables of its isothermal curves (lines of constant T) are plotted in figure 15.

For $P > P_c$, the pressure is a monotonously decreasing function as V increases at constant T . This is in accord with the physical dynamics of a gas. For $P < P_c$, however, the situation is more complex. As one decreases the volume away from very large volume, the pressure at first increases monotonically, as was the case for $P > P_c$. But as the volume is further reduced, it would appear that there comes a point where the pressure now starts to decrease (on the lowest curve in the left panel of figure 15). This behavior is not consistent with the behavior of a gas. Rather, the tendency to decrease the pressure should be viewed as a symptom of condensation into the liquid phase.

10.5 The Maxwell construction

The actual physical equilibrium curve between the two coexisting phases may be derived by using the Gibbs-Duhem relation,

$$SdT - VdP + Nd\mu = 0 \quad (10.22)$$

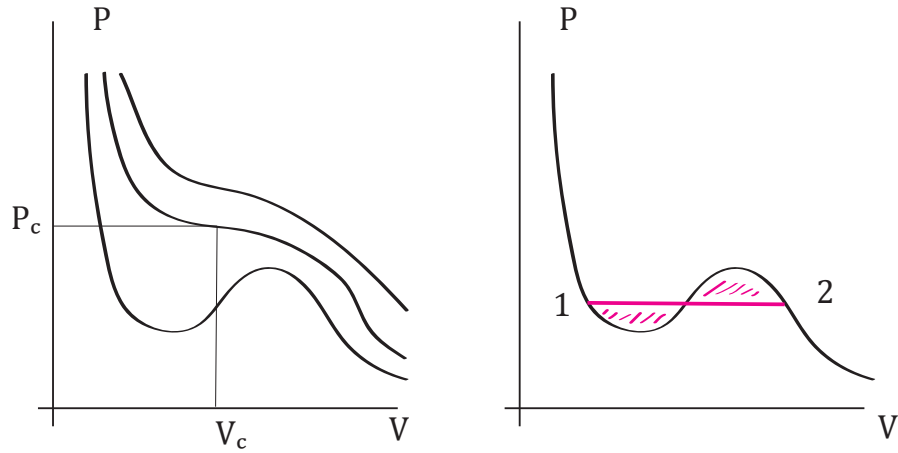


Figure 15: The (V, P) diagram of isotherms for the Van der Waals equation of state on the left panel, and the Maxwell construction for phase equilibrium on the right panel.

Along an isotherm, we have $dT = 0$, so the first term drops out. The remaining terms, with constant N , gives us a differential expression $Nd\mu = VdP$. Integrating this relation between the two phases must reflect the equality of the chemical potential μ in the two phases, and thus we must have,

$$0 = N(\mu_2 - \mu_1) = \int_1^2 VdP \quad (10.23)$$

This relation determines a horizontal line in the V, P diagram of the Van der Waals equation of state which is such that the area in V, P space above equals the area below. This is precisely the Maxwell construction for equilibrium between two phases.

11 Phase transitions: statistical mechanics

A statistical mechanics understanding of the existence and coexistence of various phases of matter and the phase transitions between them is what we will pursue in this section. After a brief classification of the different kinds of phase transitions in Nature, we will proceed to studying the Ising model, which captures so many of the key features of phase transitions.

11.1 Classification of phase transitions

In the gas-liquid-solid phase diagram of figure 10, the variables T, P, μ are continuous across the phase coexistence curves, but the molecular volume v and molecular entropy s are discontinuous. When this is the case, the phase transition across the coexistence curve is referred to as *first order*. A first order transition is always accompanied by the exchange of latent heat, and a change of molecular volume, both of which are related to the slope dP/dT by the Clausius-Clapeyron equation.

More generally, the free energy F , or the Gibbs potential G for the grand canonical ensemble, or whichever thermodynamic potential is given directly by the logarithm of the partition function $\ln Z$, is always *continuous*, but its derivatives may exhibit different degrees of discontinuity. *The order of a phase transition is $n \geq 1$ provided the partition function Z and its first $n - 1$ derivatives are continuous, but its derivative of order n is discontinuous.* Thus, in a second order phase transition, the molecular entropy and molecular volume are continuous as well as the free energy and Gibbs potential, but the specific heat is discontinuous, since it is given by two derivatives of the partition function.

Examples of first order transitions include the gas-liquid-solid transitions already discussed. Examples of second order phase transitions include the Curie phase transition in ferro-magnets, and the phase transitions to superfluidity, superconductivity, and so on. Landau developed a general theory of second order phase transitions based on symmetry considerations, and the existence of an order parameter. Generally, second order phase transitions in the presence of an external field, such as a magnetic field, may either disappear or become first order. Conversely, the nature of the first order phase transition along the gas-liquid coexistence curve in figure 10 changes quantitatively, and the latent heat diminishes as one approaches the *end point* of the curve, where the latent heat vanishes. Thus, the phase transition precisely at the end point is second order.

Finally, it pays to pause for a moment and to think about the significance of a phase transition from a purely mathematical point of view. Consider a partition function in the canonical ensemble, given in terms of the energy levels ε_i of the system,

$$Z(\beta) = \sum_i e^{-\beta\varepsilon_i} \tag{11.1}$$

If this sum were finite, then $Z(\beta)$ is an analytic function of β , and in particular all its derivatives will be continuous, to all orders. So, in any system with a finite number of states, no phase transitions can ever occur. Even if there are an infinite number of states available, $Z(\beta)$ will still often be analytic, as is the case for the harmonic oscillator for example. So, one really needs an infinite number of degrees of freedom, usually corresponding to the thermodynamic limit

$$N \rightarrow \infty \qquad V \rightarrow \infty \qquad (11.2)$$

In finite volume V , and with finite N , no phase transitions will occur. But in Nature, N is very large, but not infinite. While this is certainly true, it becomes mathematically more convenient to view a function as discontinuous rather than to consider a function whose slope is $\sim N \sim 10^{23}$. So, to some extent, phase transitions provide a mathematically simplified picture of a very very large system.

11.2 The Ising Model

The Ising model is one of the simplest, and most useful systems in statistical mechanics. Its degrees of freedom are classical spins s_i which can take the values ± 1 for each lattice site i of a d -dimensional lattice, as depicted by arrows in figure 16. It is customary to take the lattice to be “square” (i.e. cubic in 3 dimensions), but the Ising model has been considered on most other types of lattices as well, including triangular, hexagonal, random, Bethe, etc.

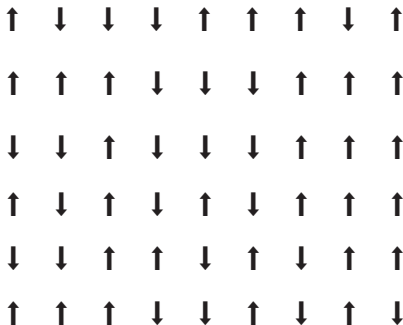


Figure 16: A configuration of spins in the 2-dimensional Ising model.

The Hamiltonian is taken as follows,

$$H = -J \sum_{\langle i,j \rangle} s_i s_j + b \sum_i s_i \qquad (11.3)$$

Here J is the coupling strength, and b is an exterior parameter such as a magnetic field; both of these parameters are real, and could be positive or negative. The symbol $\langle i,j \rangle$ indicates that the sum over the sites i, j is restricted in some way: for example to be *nearest neighbor*.

The Ising model captures binary degrees of freedom, which may capture the interactions between magnetic moments of electrons at sites in the approximation where only the contribution of one of their spatial components is retained (when all components are retained, we have the Heisenberg model instead). It can also capture the degrees of freedom of a binary mixture of substances A and B , where for $s_i = +1$ corresponds to an atom of A while $s_i = -1$ corresponds to an atom of B .

To begin the study of the dynamics of the system, we set the external field $b = 0$. The nature of the ground state of the system then depends on the sign of J .

- The coupling $J > 0$ is referred to as *ferro-magnetic* because the two possible ground states have all spins equal, either all up or all down. Note that the nature of these ground states does not depend on the lattice structure. When thought of as magnetic spins, the system would then be magnetized in either one of its ground state. Of course, thermal fluctuations may wash out the magnetization as temperature is increased. One can show that the spin waves around this ground state obey a non-relativistic dispersion relation $\omega \sim \mathbf{k}^2$.
- The coupling $J < 0$ is referred to as *anti-ferro-magnetic*. The nature of the ground state now depends to some degree on the structure of the lattice. The energy associated with the coupling between two spins i, j included in the sum over $\langle i, j \rangle$ is minimized when s_i and s_j are opposite. But if the coupling is nearest neighbor on a triangular lattice, then it is impossible to satisfy minimum energy for all three bonds on the triangle. The system is said to be *frustrated*. On a d -dimensional square lattice, whose sites are labelled by integers (i_1, i_2, \dots, i_d) , we do have a two-fold degenerate ground state with absolutely minimal energy, given by a perfectly alternating spin assignment,

$$s_{i_1, i_2, \dots, i_d} = (-1)^{i_1 + i_2 + \dots + i_d} s_0 \quad s_0 = \pm 1 \quad (11.4)$$

This state is called the *Néel ground state* of an anti-ferro-magnetic system. One can show that the spin waves around this ground state obey a linear dispersion relation $\omega \sim |\mathbf{k}|$, which is akin to the relation for a massless relativistic particle.

11.3 Exact solution of the 1-dimensional Ising Model

The partition function Z is defined by,

$$Z = \sum_{s_1, s_2, \dots, s_N = \pm 1} e^{\beta \sum_{i=1}^N (s_i s_{i+1} - b s_i)} \quad (11.5)$$

and where we make the system periodic by requiring $s_{N+1} = s_1$. To compute Z , we write it as the trace of a sequential product,

$$Z = \text{tr}(\mathcal{E}_1 \mathcal{E}_2 \cdots \mathcal{E}_{N-1} \mathcal{E}_N)$$

$$\mathcal{E}_i = \exp \left\{ \beta J s_i s_{i+1} + \frac{1}{2} \beta b (s_i + s_{i+1}) \right\} \quad (11.6)$$

The problem now becomes one of 2×2 matrix multiplication. To see this, we work out the matrix elements of \mathcal{E}_i that enter here,

$$\mathbf{T}_{\sigma', \sigma} = \exp \left\{ \beta J \sigma' \sigma + \frac{1}{2} \beta b (\sigma + \sigma') \right\} \quad (11.7)$$

The partition function is then given by

$$Z = \sum_{\{\sigma_i = \pm 1\}} T_{\sigma_1, \sigma_2} T_{\sigma_2, \sigma_3} \cdots T_{\sigma_N, \sigma_1} = \text{tr}(\mathbf{T}^N) \quad (11.8)$$

Written out explicitly, the *transfer matrix* \mathbf{T} is given by

$$\mathbf{T} = \begin{pmatrix} e^{\beta J + \beta b} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J - \beta b} \end{pmatrix} \quad (11.9)$$

Its eigenvalues λ_{\pm} satisfy the equation,

$$\lambda_{\pm}^2 - 2\lambda_{\pm} e^{\beta J} \text{ch}(\beta b) + 2 \text{sh}(2\beta J) = 0 \quad (11.10)$$

which is solved by

$$\lambda_{\pm} = e^{\beta J} \text{ch}(\beta b) \pm \sqrt{e^{2\beta J} \text{sh}^2(\beta b) + e^{-2\beta J}} \quad (11.11)$$

Therefore, the partition function is given by $Z = \lambda_+^N + \lambda_-^N$. for all values of N , β , J , and b . Since we have $\lambda_+ > \lambda_-$ for all values and signs of J , the thermodynamic limit of the system simplifies as we take $N \rightarrow \infty$. The free energy per unit volume is given by,

$$\frac{F}{N} = -kT \lim_{N \rightarrow \infty} \frac{1}{N} \ln (\lambda_+^N + \lambda_-^N) = -kT \ln \lambda_+ \quad (11.12)$$

This function is analytic in β , so there are no phase transitions for any finite value of β . In other words, the system is in the same thermodynamic phase for all temperatures.

11.4 Ordered versus disordered phases

An important qualitative characteristic of the dynamics of statistical magnetic systems is order versus disorder. For magnetic systems, this property may be understood systematically in terms of the magnetization, defined as the thermodynamic average of the spin,

$$m(b) = -\frac{1}{N} \frac{\partial F}{\partial b} \quad (11.13)$$

In the case of the 1-dimensional Ising model this quantity is easily computed for both $J > 0$ or $J < 0$. The eigenvalue λ_+ depends on b through an even function of b , and thus, the magnetization $m(b)$ at zero external magnetic field always vanishes. This result is interpreted as the fact that the spins in the system, on average at all temperatures, point in all directions randomly, so that their total contribution to magnetization vanishes in the bulk.

When can a system be ordered then ? We have seen previously that for $J > 0$, the minimum energy states are

$$\begin{aligned} s_i &= +1 && \text{for all } i \\ s_i &= -1 && \text{for all } i \end{aligned} \tag{11.14}$$

These ground states are mapped into one another by the spin reversal symmetry R of the Hamiltonian for $b = 0$. If both ground states contribute to the partition function, then the total magnetization will get wiped out, and the system will remain in a disordered phase. When N is finite this will always be the case. But when $N \rightarrow \infty$, it is possible for the system to get stuck in one ground state or the other. The reason this only happens for infinite N is that it would then take an infinite number of spin flips to transition between the $s_i = +1$ and $s_i = -1$ states, and this may get energetically impossible. When the system gets stuck in one of its ground states, then $m(0) \neq 0$ and we have spontaneous magnetization, familiar from ferromagnetism below the Curie temperature. The operation of spin reversal, which is a symmetry of the Hamiltonian for $b = 0$ is then NOT a symmetry of the physical system any more, as a definite non-zero value of $m(0)$ is not invariant under R . The symmetry R is said to be *spontaneously broken*, and the system is then in an ordered phase, close to one of its ground states. We have already shown that, for the 1-dimensional Ising model, this phenomenon does not take place.

The 2-dimensional Ising model, however, does exhibit an ordered phase below a critical temperature T_c . This is known since the model was solved exactly by Lars Onsager in 1944, and the critical temperature is known analytically,

$$\text{sh}(2J\beta_c) = 1 \qquad 2J\beta_c = 0.881412 \tag{11.15}$$

The corresponding magnetization was computed by C.N. Yang,

$$\begin{aligned} m(0) &= \left(1 - \frac{1}{\text{sh}^4(2J\beta)}\right)^{1/8} && T < T_c \\ m(0) &= 0 && T > T_c \end{aligned} \tag{11.16}$$

Note that as $T \rightarrow T_c$, the expression $m(0) \sim (T_c - T)^{1/8}$ vanishes and joins continuously with the $T > T_c$ result $m(0) = 0$. The phase transition at $T = T_c$ is actually second order.

The exponent $1/8$ is referred to as a *critical exponent*, in this case of the magnetization order parameter m . Critical exponents tend to be *universal* objects, whose dependence on the detailed short-distance interactions is limited.

Whether the 3-dimensional Ising model allows for an exact solution is one of the great outstanding problems of statistical mechanics. Proposals have been made that the model behaves as a theory of free fermionic random surfaces, but the details have never been conclusive. Numerical studies of the model have shown however that it also admits a phase transition between ordered (low temperature) and disordered (high temperature) phases.

11.5 Mean-field theory solution of the Ising model

Short of an exact solution, we may try to approach the problem for the ferromagnetic $J > 0$ coupling with the help of some drastic simplifications. In mean-field theory, one assume that fluctuations of the spins away from the average value are small, and one determines this average value self-consistently. Our starting point is to recast the Hamiltonian in terms of the average magnetization per spin m which is to be determined. We proceed as follows,

$$s_i s_j = (s_i - m)(s_j - m) + m(s_i + s_j) - m^2 \quad (11.17)$$

The key assumption of mean-field theory is that the statistical average of the first term on the right side may be neglected. The remaining mean-field Hamiltonian is obtained by summing over all pairs $\langle i, j \rangle$. Reformulating this sum in terms of a sum over individual sites may be done by noticing that on a d -dimensional square lattice, the number of bonds from a site is $2d$. The Hamiltonian is then given by,

$$H_{\text{mf}} = JNd m^2 - b_{\text{eff}} \sum_i s_i \quad b_{\text{eff}} = b + 2dJm \quad (11.18)$$

Here, b_{eff} is the *effective magnetic field* felt by each spin s_i due to the mean value of the spins that surround it. The partition function is now easily computed, leaving m as a parameter which remains to be determined self-consistently. We find,

$$Z = e^{JNd m^2} \left(e^{\beta b_{\text{eff}}} + e^{-\beta b_{\text{eff}}} \right)^N \quad (11.19)$$

The average magnetization, computed from Z is the to be set equal to m ,

$$m = \frac{1}{N} \frac{\partial \ln Z}{\partial (\beta b)} = \tanh(\beta b + \beta 2dJm) \quad (11.20)$$

Solving for m at $b = 0$, we see that there is always the solution $m = 0$. When $2dJ\beta < 1$, namely for higher temperatures $2dJ < kT$, then $m = 0$ is the only solution, and we have

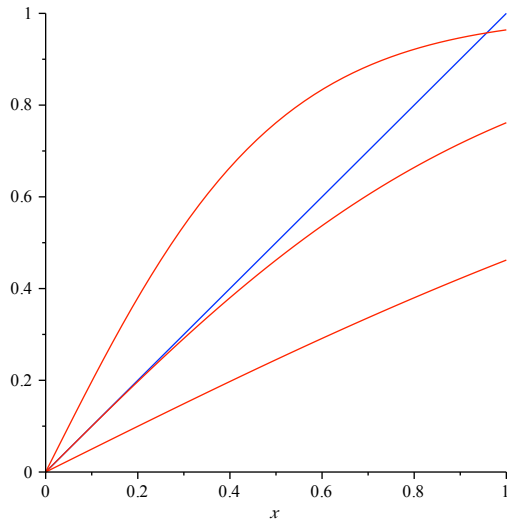


Figure 17: Plot of the curves x in blue, and $\tanh(2x)$, $\tanh(x)$, and $\tanh(x/2)$; only the first produces an intersection with $x \neq 0$.

no spontaneous magnetization. For $2dJ > kT$ on the other hand, we have one solution with non-zero m , as may be seen by inspecting figure 17.

It is straightforward to solve in the approximation where $2dJ\beta \approx 1$, so that m will be small. One finds the following mean-field result for the magnetization near the phase transition point,

$$m \approx \sqrt{6dJk\beta_c^2(T_c - T)^{\frac{1}{2}}} \quad 2dJ\beta_c = 1 \quad (11.21)$$

We see that mean-field theory gives a completely wrong answer for $d = 1$, where there is no phase transition at all in view of our exact result. For $d = 2$, comparison with the Onsager solution shows that β_c is off by a factor of 2, but more importantly, that the critical exponent is off by a factor of 4. So, while mean field theory correctly predicts a phase transition for $d = 2$, the details are not impressive. Numerical simulations show that in three dimensions, we have

$$m \sim (T_c - T)^{0.308} \quad (11.22)$$

so this is closer to the mean-field exponent of 0.5 for $d = 3$. In fact, one can argue that mean-field theory becomes exact in the limit of $d \rightarrow \infty$.

12 Functional integral methods

Functional integral methods may be extended from quantum mechanics to quantum statistical mechanics, where they provide an extremely useful formalism. In this section, we begin by deriving the functional integral for the partition function of a classical mechanical system. We use the path integral representation to study the classical limit, low and high temperature asymptotics, as well as perturbation theory.

12.1 Path integral representation for the partition function

We give the derivation for the case of a system with just one degree of freedom, p, q , represented by quantum operators P, Q via the correspondence principle. The operators obey canonical commutation relations $[Q, P] = i\hbar$, and time-evolution is governed by a self-adjoint, time-independent, quantum Hamiltonian $H = H(P, Q)$. The generalization to N degrees of freedom will be straightforward. The object is to calculate the partition function,

$$Z = \text{Tr} \left(e^{-\beta H} \right) \quad (12.1)$$

The Boltzmann operator $e^{-\beta H}$ may be viewed as an analytic continuation to imaginary time $t \rightarrow -i\beta\hbar$ of the evolution operator $U(t) = e^{-itH/\hbar}$. Thus, the derivation given here for the partition function will show strong parallels, but also some important differences, with the construction of the path integral representation for quantum mechanics.

Recall the respective bases in which the operators Q and P are diagonal,

$$\begin{aligned} Q|q\rangle &= q|q\rangle \\ P|p\rangle &= p|p\rangle \end{aligned} \quad (12.2)$$

Orthonormality and completeness hold as follows,

$$\begin{aligned} \langle q'|q\rangle &= \delta(q - q') & I_{\mathcal{H}} &= \int dq |q\rangle\langle q| \\ \langle p'|p\rangle &= 2\pi\hbar \delta(p - p') & I_{\mathcal{H}} &= \int \frac{dp}{2\pi\hbar} |p\rangle\langle p| \end{aligned} \quad (12.3)$$

where $I_{\mathcal{H}}$ is the identity element in Hilbert space. Translation operators in both bases satisfy

$$\begin{aligned} e^{+iaP/\hbar} Q e^{-iaP/\hbar} &= Q + a & e^{-iaP/\hbar} |q\rangle &= |q + a\rangle \\ e^{-ibQ/\hbar} P e^{+ibQ/\hbar} &= P + b & e^{+ibQ/\hbar} |p\rangle &= |p + b\rangle \end{aligned} \quad (12.4)$$

From these relations, we deduce the values of the mixed matrix elements,

$$\begin{aligned} \langle q|p\rangle &= e^{+iqp/\hbar} \\ \langle p|q\rangle &= e^{-iqp/\hbar} \end{aligned} \quad (12.5)$$

This summarizes all that will be needed to derive the functional integral representation.

To evaluate the trace over the Boltzmann operator in (12.1), we divide β into N intervals of equal length ε such that $\beta = N\varepsilon$, and use the formula,

$$e^{-\beta H} = \left(I_{\mathcal{H}} e^{-\varepsilon H} \right)^N \quad (12.6)$$

We represent $I_{\mathcal{H}}$ using both completeness relations of (12.3) in the following manner,

$$I_{\mathcal{H}} = \int \frac{dp dq}{2\pi\hbar} |q\rangle \langle q|p\rangle \langle p| \quad (12.7)$$

In practice, we will need to use N different pairs of integration variables p_n, q_n with $n = 1, \dots, N$, namely one pair for each inserted identity operator. Working this out, we find,

$$Z = \text{Tr} \left(e^{-\beta H} \right) = \prod_{n=1}^N \int \frac{dp_n dq_n}{2\pi\hbar} \langle q_{n-1} | p_n \rangle \langle p_n | e^{-\varepsilon H} | q_n \rangle \quad (12.8)$$

with the understanding that we set $q_0 = q_N$.

Letting $\varepsilon \rightarrow 0$, we may expand the Boltzmann operator in a power series in ε ,

$$e^{-\beta H} = I_{\mathcal{H}} - \varepsilon H + \mathcal{O}(\varepsilon^2) \quad (12.9)$$

The whole purpose of inserting the completeness relations in both $|p\rangle$ and $|q\rangle$ states is to have a mixed representation in which matrix elements for both kinetic and potential terms in the Hamiltonian may be evaluated explicitly. Consider, for example, a potential-type Hamiltonian of the form $H = P^2/2m + U(Q)$. It is manifest that we have $\langle p|H|q\rangle = \langle p|q\rangle H(p, q)$, where $H(p, q)$ is the classical Hamiltonian. Given a general Hamiltonian, H , we can always order all P operators to the left of all Q operators, using $[Q, P] = i\hbar$. It is in this form that we can use it to *define* the classical Hamiltonian $H(q, p)$ by,

$$\langle p|H|q\rangle = \langle p|q\rangle H(p, q) \quad (12.10)$$

The needed matrix elements are now readily evaluated, and we have,

$$\langle p_n | e^{-\varepsilon H} | q_n \rangle = \langle p_n | q_n \rangle \left(1 - \varepsilon H(p_n, q_n) + \mathcal{O}(\varepsilon^2) \right) = \langle p_n | q_n \rangle e^{-\varepsilon H(p_n, q_n)} \quad (12.11)$$

up to contribution that are of higher order in ε . It remains to evaluate the combination,

$$\langle q_{n-1} | p_n \rangle \langle p_n | q_n \rangle = e^{i(q_{n-1} - q_n)p_n/\hbar} \quad (12.12)$$

Putting all together, we have,

$$\langle q_{n-1} | p_n \rangle \langle p_n | e^{-\varepsilon H(P, Q)} | q_n \rangle = e^{i(q_{n-1} - q_n)p_n/\hbar - \varepsilon H(p_n, q_n)} \quad (12.13)$$

The final step consists of a change of notation,

$$\begin{aligned}\tau_n &= n\varepsilon\hbar = \hbar\beta n/N \\ q_n &= q(\tau_n) \\ p_n &= p(\tau_n)\end{aligned}\tag{12.14}$$

The argument of the exponential may now be recast as follows,

$$\begin{aligned}\frac{i}{\hbar}(q_{n-1} - q_n)p_n - \varepsilon H(p_n, q_n) &= -\frac{\tau_n - \tau_{n-1}}{\hbar} \left(i \frac{q(\tau_{n-1}) - q(\tau_n)}{\tau_{n-1} - \tau_n} p(\tau_n) + H(p(\tau_n), q(\tau_n)) \right) \\ &\approx -\frac{1}{\hbar} \int_{\tau_{n-1}}^{\tau_n} d\tau \left(i\dot{q}(\tau)p(\tau) + H(p(\tau), q(\tau)) \right)\end{aligned}\tag{12.15}$$

Taking $N \rightarrow \infty$, so that $\varepsilon \rightarrow 0$ makes exact the expansion in ε to first order. We introduce the functional integration measure,

$$\int \mathcal{D}p \mathcal{D}q = \lim_{N \rightarrow \infty} \prod_{n=1}^N \int_{\mathbf{R}} \frac{dp(\tau_n) dq(\tau_n)}{2\pi\hbar}\tag{12.16}$$

and notice that the trace condition produces a periodicity requirement on the position coordinate $q(\beta\hbar) = q(0)$. Putting all together,

$$Z = \int \mathcal{D}p \mathcal{D}q \exp \left\{ -\frac{1}{\hbar} \int_0^{\beta\hbar} d\tau \left(i\dot{q}(\tau)p(\tau) + H(p(\tau), q(\tau)) \right) \right\}\tag{12.17}$$

The momenta enter without derivatives in this expression. Without loss of generality, we may impose periodic boundary conditions also on the variable p , so that we have,

$$\begin{aligned}p(\beta\hbar) &= p(0) \\ q(\beta\hbar) &= q(0)\end{aligned}\tag{12.18}$$

Although expression (12.17) involves a factor of i , we are assured from the outset that the partition function must be real.

Generalization to s degrees of freedom p_i, q_i and $i = 1, \dots, s$, with a Hamiltonian $H(p_1, \dots, p_s, q_1, \dots, q_s) = H(p, q)$ is straightforward, and we find,

$$Z = \left(\prod_{i=1}^s \int \mathcal{D}p_i \mathcal{D}q_i \right) \exp \left\{ -\frac{1}{\hbar} \int_0^{\beta\hbar} d\tau \left(i \sum_{i=1}^s \dot{q}_i(\tau)p_i(\tau) + H(p(\tau), q(\tau)) \right) \right\}\tag{12.19}$$

with the periodicity conditions (12.18) enforced on every p_i, q_i .

12.2 The classical = high temperature limit

In view of the periodicity conditions (12.18), we may expand the functions $p(\tau)$ and $q(\tau)$ in an orthogonal basis of periodic functions on the interval $\tau \in [0, \beta\hbar]$,

$$\begin{aligned} q(\tau) &= q + \sum_{n=1}^{\infty} \left(a_n e^{i\tau\omega_n} + a_n^* e^{-i\tau\omega_n} \right) \\ p(\tau) &= p + \sum_{n=1}^{\infty} \left(b_n e^{i\tau\omega_n} + b_n^* e^{-i\tau\omega_n} \right) \end{aligned} \quad (12.20)$$

The *thermal* or *Matsubara frequencies* are given by,

$$\omega_n = \frac{2\pi n}{\beta\hbar} \quad (12.21)$$

and the zero-modes p, q have been explicitly separated.

When $\beta\hbar \rightarrow 0$ the frequencies ω_n for $n > 0$ become large while the zero modes p, q are unaffected. When the Matsubara frequencies are much larger than all the characteristic scales of the Hamiltonian, we may use the following $\beta\hbar$ approximation,

$$\frac{1}{\hbar} \int_0^{\beta\hbar} d\tau H(p(\tau), q(\tau)) \approx \beta H(p, q) \quad (12.22)$$

where only the zero modes p, q are retained on the right side in $H(p, q)$. Separating also the integration measure into their zero-mode and non-zero-mode contributions,

$$\mathcal{D}p \mathcal{D}q = \frac{dp dq}{2\pi\hbar} \mathcal{D}'p \mathcal{D}'q \quad (12.23)$$

and using the fact that the non-zero-modes precisely integrate to 1, namely

$$\int \mathcal{D}'p \mathcal{D}'q \exp \left\{ -\frac{i}{\hbar} \int_0^{\beta\hbar} d\tau p(\tau) \dot{q}(\tau) \right\} = 1 \quad (12.24)$$

we find that the $\beta\hbar \rightarrow 0$ approximation to the partition function is given by,

$$Z \approx \int \frac{dp dq}{2\pi\hbar} e^{-\beta H(p, q)} \quad (12.25)$$

which is the properly normalized classical result we had already encountered earlier.

12.3 Integrating out momenta

Returning to the non-classical case for arbitrary $\beta\hbar$, we may seek to integrate out explicitly the momentum variables, since they enter purely algebraically. When H is quadratic-linear in p , this is always possible. Consider for example the case of a Hamiltonian for a charged particle in a magnetic field and potential $U(\mathbf{q})$,

$$H = \frac{(\mathbf{p} - e\mathbf{A}(\mathbf{q}))^2}{2m} + U(\mathbf{q}) \quad (12.26)$$

The integrand of the exponential may be rearranged as follows,

$$\int_0^{\beta\hbar} d\tau \left(i\dot{\mathbf{q}} \cdot \mathbf{p} + H(\mathbf{p}, \mathbf{q}) \right) = \int_0^{\beta\hbar} d\tau \left(\frac{1}{2m} (\mathbf{p} - e\mathbf{A} + im\dot{\mathbf{q}})^2 + \frac{1}{2}m\dot{\mathbf{q}}^2 + i\dot{\mathbf{q}} \cdot \mathbf{A} + U(\mathbf{q}) \right)$$

The first term produces a Gaussian integral in \mathbf{p} , which may be carried out and gives a factor Z_0 in the partition function which depends only on \hbar , m , and β , but not on the dynamics of the system contained in \mathbf{A} and $U(\mathbf{q})$. Thus, we have,

$$Z = Z_0 \int \mathcal{D}\mathbf{q} \exp \left\{ -\frac{1}{\hbar} \int_0^{\beta\hbar} d\tau \left(\frac{1}{2}m\dot{\mathbf{q}}^2 + i\dot{\mathbf{q}} \cdot \mathbf{A} + U(\mathbf{q}) \right) \right\} \quad (12.27)$$

The object under the τ -integration is the Euclidean Lagrangian. Note the factor of i in front of the electro-magnetic interaction !