Statistical Mechanics I

Peter S. Riseborough

November 15, 2011

Contents

1 Introduction

Real materials are composed of a huge numbers of particles. For example, one cubic centimeter of copper or one liter of water contains about 10^{23} atoms. The enormity of the number of degrees of freedom prevents one from be able to either determine or store the initial conditions let alone from solving the equations of motion. Hence, a detailed microscopic description appears impossible. Nevertheless, the equilibrium states of such materials can be defined by a relative few macroscopic quantities, such as temperature, pressure or volume, etc. These quantities reflect the collective properties of the constituents of the material but still can be measured quite directly by macroscopic means. Likewise, certain non-equilibrium states of the material can also described by a few easily measured quantities such as the voltage drop across or the electrical current flowing through an electrical circuit element. Often simple "laws" emerge between the macroscopic quantities that describe the properties of these complex systems. The subject of Thermodynamics is devoted to revealing relations, sometimes expected and sometimes unexpected, between the macroscopic quantities describing materials. Statistical Mechanics provides statistically based methods which bridge the gap between the physics of the individual particles that comprise the materials and the simple thermodynamic "laws" that describe the macroscopic properties of many-particle systems.

2 Thermodynamics

Thermodynamics is a branch of science that does not assert new fundamental principles but, instead predicts universal relations between measurable quantities that characterize macroscopic systems. Specifically, thermodynamics involves the study of macroscopic coordinates which can be expressed in terms of an extremely large number of microscopic degrees of freedom, and that describe the macroscopic states of systems.

2.1 The Foundations of Thermodynamics

Macroscopic measurements have the attributes that they involve large numbers of microscopic degrees of freedom (such at the positions and momenta of 10^9 atoms) and, are measured over extremely long time scales compared with the time scales describing the microscopic degrees of freedom (of the order 10^{-7} seconds). In general, for sufficiently large systems and when averaged over sufficiently long time scales, the fluctuations of the macroscopic variables are extremely small and so only the average values need be retained.

Typical macroscopic variables are the internal energy U, the number of particles N and the volume of the system V. The internal energy U is a precisely defined quantity, which in the absence of interactions between the system and its environment, is also a conserved quantity. For systems which contain particles that do not undergo reactions, the number of particles N is also a well-defined

and conserved quantity. The volume V of a system can be measured, to within reasonable accuracy. A measurement of the volume usually occurs over time-scales that are long enough so that the long wavelength fluctuations of the atomic positions of the systems boundaries or walls are averaged over.

Thermodynamic measurements are usual indirect, and usually involve externally imposed constraints. Systems which are constrained such that they cannot exchange energy, volume or number of particles with their environments are said to be closed. Since one usually only measures changes in the internal energy of a system ΔU, such measurements necessarily involve the system's environment and the assumption that energy is conserved. Such measurements can be performed by doing electrical or mechanical work on the system, and preventing energy in the form of heat from flowing into or out of the system. The absence of heat flow is ensured by utilizing boundaries which are impermeable to heat flow. Boundaries which are impermeable to heat flow are known as adiabatic. Alternatively, one may infer the change in internal energy of a system by putting it in contact with other systems that are monitored and that have been calibrated so that their changes in internal energy can be found. For any process, the increase in the internal energy ΔU can be expressed as the sum of the heat absorbed by the system ΔQ and the work done on the system ΔW.

$$\Delta U \;=\; \Delta Q \;+\; \Delta W \tag{1}$$

where N is being held constant. This equation is a statement of the conservation of energy. This basic conservation law in its various forms, forms an important principle of thermodynamics.

2.2   Thermodynamic Equilibrium

Given a macroscopic system, experience shows that this system will evolve to a state in which the macroscopic properties are determined by intrinsic factors and not by any external influences that had been previously exerted on the system. The final states, by definition, are independent of time and are known as equilibrium states.

Postulate I

It is postulated that, in equilibrium, the macroscopic states of a system can be characterized by a set of macroscopic variables. These variable may include variables taken from the set $\{U, V, N\}$ together with any other macroscopic variables that must be added to the set in order to describe the equilibrium state uniquely.

For example, in a ferromagnet this set may be extended by adding the total magnetic moment M of the sample. Thus, for a ferromagnet one might specify the equilibrium state by the macroscopic variables $\{U, N, V, M\}$. Another example is given by the example of a gas containing r different species of atoms, in which case the set of macroscopic variables should be extended to include

the number of atoms for each species $\{N_1, N_2, \ldots N_r\}$. Due to the constraint $N = \sum_{i=1}^{r} N_i$, the total number N of atoms should no longer be considered as an independent variable.

The set of variables $\{U, V, N, M, \ldots\}$ are extensive variables, since they scale with the size of the system. This definition can be made more precise as follows: Consider a homogeneous system that is in thermal equilibrium. The value of the variable X for the equilibrated system is denoted by X_0. Then the variable X is extensive if, when one considers the system as being composed of λ identical subsystems ($\lambda > 1$), the value of the variable X for each subsystem is equal to $\lambda^{-1} X_0$. This definition assumes that the subsystems are sufficiently large so that the fluctuations δX of X are negligibly small.

The extensive variables $\{U, V, N\}$ that we have introduced, so far, all have mechanical significance. There are extensive variables that only have thermodynamic significance, and these variables can can also be used to characterize equilibrium states. One such quantity is the entropy S.

Postulate II

The entropy S is defined only for equilibrium states, and takes on a value which is uniquely defined by the state. That is, S is a single-valued function $S(U, V, N)$ of the mechanical extensive variables. The entropy has the property that it is maximized in an equilibrium state, with respect to the variation of hypothetical internal constraints. The constraint must be designed so that, in the absence of the constraint, the system is free to select any one of a number of states each of which may be realized in the presence of the constraint. If the hypothetical internal constraint characterized by the variable x is imposed on the system, then the entropy of the system depends on the constraint through x and can be denoted by $S(x)$. The maximum value of the entropy of the unconstrained system S is given by the maximum value of $S(x)$ found when x is varied over all possible values.

The function $S(E, V, N)$ for a system is known as the fundamental relation, since all conceivable thermodynamic information on the system can be obtained from it.

Postulate III

The entropy of a system is not only an extensive variable, but also the entropy of a composite system is the sum of the entropies of its components. The entropy is a continuous, differentiable and a monotonically increasing function of the entropy.

Postulate III ensures that when the absolute temperature T is defined for an equilibrium state, then T will be positive.

Postulate IV

The entropy of a system vanishes in the limit where

$$\left(\frac{\partial S}{\partial U}\right)_{V,N} \rightarrow \infty \tag{2}$$

The above condition identifies a state for which the absolute temperature approaches the limiting value $T \rightarrow 0$.

Postulate IV is equivalent to Nernst's postulate that the entropy takes on a universal value when $T \rightarrow 0$. The above form of the postulate defines the universal value of the entropy to be zero.

2.3 The Conditions for Equilibrium

The above postulates allows the conditions for equilibrium to be expressed in terms of the derivatives of the entropy. Since, entropy is an extensive quantity, its derivatives with respect to other extensive quantities will be intensive. That is, the derivatives are independent of the size of the system. The derivatives will be used to define intensive thermodynamic variables.

First, we shall consider making use of the postulate that entropy is a single-valued monotonic increasing function of energy. This implies that the equation for the entropy

$$S = S(U, V, N) \tag{3}$$

can be inverted to yield the energy as a function of entropy

$$U = U(S, V, N) \tag{4}$$

This inversion may be difficult to do if one is presented with a general expression for the function, but if the function is presented graphically this is achieved by simply interchanging the axes. Consider making infinitesimal changes of the independent extensive variables (S, V, N), then the energy $U(S, V, N)$ will change by an amount given by

$$dU = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS + \left(\frac{\partial U}{\partial V}\right)_{S,N} dV + \left(\frac{\partial U}{\partial N}\right)_{S,V} dN \tag{5}$$

The three quantities

$$\left(\frac{\partial U}{\partial S}\right)_{V,N}$$
$$\left(\frac{\partial U}{\partial V}\right)_{S,N}$$
$$\left(\frac{\partial U}{\partial N}\right)_{S,V} \tag{6}$$

are intensive since if a system in equilibrium is considered to divided into λ identical subsystems the values of these parameters for each of the subsystems is the same as for the combined system. These three quantities define the energy intensive parameters. A quantity is intensive if its value is independent of the scale of the system, that is, its value is independent of the amount of matter used in its measurement. The intensive quantities can be identified as follows:

By considering a process in which S and N are kept constant and V is allowed to change, one has

$$dU = \left(\frac{\partial U}{\partial V}\right)_{S,N} dV \tag{7}$$

which, when considered in terms of mechanical work ΔW, leads to the identification

$$\left(\frac{\partial U}{\partial V}\right)_{S,N} = -P \tag{8}$$

where P is the mechanical pressure. Likewise, when one considers a process in which V and N are kept constant and S is allowed to change, one has

$$dU = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS \tag{9}$$

which, when considered in terms of heat flow ΔQ, leads to the identification

$$\left(\frac{\partial U}{\partial S}\right)_{V,N} = T \tag{10}$$

where T is the absolute temperature. Finally, on varying N, one has

$$dU = \left(\frac{\partial U}{\partial N}\right)_{S,V} dN \tag{11}$$

which leads to the identification

$$\left(\frac{\partial U}{\partial N}\right)_{S,V} = \mu \tag{12}$$

where μ is the chemical potential. Thus, one obtains a relation between the infinitesimal changes of the extensive variables

$$dU = T\, dS - P\, dV + \mu\, dN \tag{13}$$

This is an expression of the conservation of energy.

Direct consideration of the entropy, leads to the identification of the entropic intensive parameters. Consider making infinitesimal changes of the independent

8

extensive variables (U, V, N), then the entropy $S(U, V, N)$ will change by an amount given by

$$dS = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \qquad (14)$$

The values of the coefficients of the infinitesimal quantities are the entropic intensive parameters. By a suitable rearrangement of eqn(13) as

$$dS = \frac{1}{T} dU + \frac{P}{T} dV - \frac{\mu}{T} dN \qquad (15)$$

one finds that the entropic intensive variables are given by

$$\left(\frac{\partial S}{\partial U}\right)_{V,N} = \frac{1}{T}$$

$$\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{P}{T}$$

$$\left(\frac{\partial S}{\partial N}\right)_{U,V} = -\frac{\mu}{T} \qquad (16)$$

where T is the absolute temperature, P is the pressure and μ is the chemical potential.

The conditions for equilibrium can be obtained from Postulate II and Postulate III, that the entropy is maximized in equilibrium and is additive. We shall consider a closed system composed of two systems in contact. System 1 is described by the extensive parameters $\{U_1, V_1, N_1\}$ and system 2 is described by $\{U_2, V_2, N_2\}$. The total energy $U_T = U_1 + U_2$, volume $V_T = V_1 + V_2$ and number of particles $N_T = N_1 + N_2$ are fixed. The total entropy of the combined system S_T is given by

$$S_T = S_1(U_1, V_1, N_1) + S_2(U_2, V_2, N_2) \qquad (17)$$

and is a function of the variables $\{U_1, V_1, N_1, U_2, V_2, N_2\}$.

Heat Flow and Temperature

If one allows energy to be exchanged between the two systems, keeping the total energy $U_T = U_1 + U_2$ constant, then

$$dS_T = \left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,N_1} dU_1 + \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,N_2} dU_2 \qquad (18)$$

Since U_T is kept constant, one has $dU_1 = -dU_2$. Therefore, the change in the total entropy is given by

$$dS_T = \left[\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,N_1} - \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,N_2} \right] dU_1 \qquad (19)$$
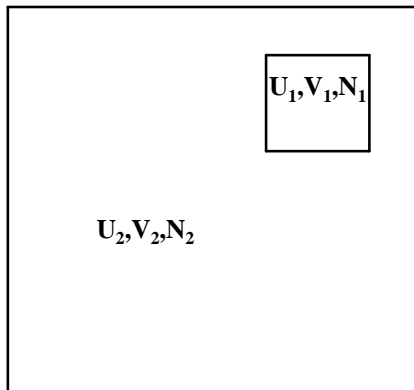
Figure 1: An isolated system composed of two subsystems.

Furthermore, in equilibrium S_T is maximized with respect to the internal partitioning of the energy, so one has

$$dS_T = 0 \tag{20}$$

For this to be true, independent of the value of dU_1, one must satisfy the condition

$$\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1, N_1} = \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2, N_2} \tag{21}$$

or, equivalently

$$\frac{1}{T_1} = \frac{1}{T_2} \tag{22}$$

Thus, the condition that two systems, which can only exchange internal energy by heat flow, are in thermal equilibrium is simply the condition that the temperatures of the two systems must be equal, $T_1 = T_2$.

Let us consider the same closed system, but one in which the two bodies are initially not in thermal contact with each other. Since the two systems are isolated, they are in a state of equilibrium but may have different temperatures. However, if the two systems are put in thermal contact, the adiabatic constraint is removed and they will no longer be in thermal equilibrium. The system will evolve, by exchanging energy in the form of heat, between the two systems and a new equilibrium state will be established. The new equilibrium state, obtained by removing the internal constraint will have a larger entropy. Hence, for the two equilibrium states which differ infinitesimally in the partitioning of the energy, $dS_T > 0$ and

$$dS_T = \left[\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1, N_1} - \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2, N_2} \right] dU_1 > 0 \tag{23}$$

or

$$\left[\frac{1}{T_1} - \frac{1}{T_2} \right] dU_1 > 0 \tag{24}$$

This inequality shows that heat flows from systems with higher temperatures to systems with lower temperatures, in agreement with expectations.

Work and Pressure

Consider a system composed of two sub-systems, which are in contact that can exchange energy and also exchange volume. System 1 is described by the extensive parameters $\{U_1, V_1, N_1\}$ and system 2 is described by $\{U_2, V_2, N_2\}$. The total energy is fixed as is the total volume. The energy and volumes of the sub-systems satisfy

$$\begin{aligned} U_T &= U_1 + U_2 \\ V_T &= V_1 + V_2 \end{aligned} \tag{25}$$

and N_1 and N_2 are kept constant. For an equilibrium state, one can consider constraints that result in different partitionings of the energy and volume. The entropy of the total system is additive

$$S_T = S_1(U_1, V_1, N_1) + S_2(U_2, V_2, N_2) \tag{26}$$

The infinitesimal change in the total entropy S_T found by making infinitesimal changes in U_1 and V_1 is given by

$$dS_T = \left[\left(\frac{\partial S_1}{\partial U_1} \right)_{V_1, N_1} - \left(\frac{\partial S_2}{\partial U_2} \right)_{V_2, N_2} \right] dU_1 + \left[\left(\frac{\partial S_1}{\partial V_1} \right)_{U_1, N_1} - \left(\frac{\partial S_2}{\partial V_2} \right)_{U_2, N_2} \right] dV_1 \tag{27}$$

since $dU_1 = -dU_2$ and $dV_1 = -dV_2$. Thus, on using the definitions for the intensive parameters of the sub-systems, one has

$$dS_T = \left[\frac{1}{T_1} - \frac{1}{T_2} \right] dU_1 + \left[\frac{P_1}{T_1} - \frac{P_2}{T_2} \right] dV_1 \tag{28}$$

Since the equilibrium state is that in which S_T is maximized with respect to the variations dU_1 and dV_1, one has $dS_T = 0$ which leads to the conditions

$$\begin{aligned} \frac{1}{T_1} &= \frac{1}{T_2} \\ \frac{P_1}{T_1} &= \frac{P_2}{T_2} \end{aligned} \tag{29}$$

Hence, the pressure and temperature of two the sub-systems are equal in the equilibrium state.

Furthermore, if the systems are initially in their individual equilibrium states but are not in equilibrium with each other, then they will ultimately come into

11

thermodynamic equilibrium with each other. If the temperatures of the two subsystems are equal but the initial pressures of the two systems are not equal, then the change in entropy that occurs is given by

$$dS_T = \left(P_1 - P_2 \right) \frac{dV_1}{T} > 0 \tag{30}$$

Since $dS_T > 0$, one finds that if $P_1 > P_2$ then $dV_1 > 0$. That is, the system at higher pressure will expand and the system at lower pressure will contract.

Matter Flow and Chemical Potential

The above reasoning can be extended to a system with fixed total energy, volume and number of particles, which is decomposed into two sub-systems that exchange energy, volume and number of particles. Since $dU_1 = - dU_2$, $dV_1 = - dV_2$ and $dN_1 = - dN_2$, one finds that an infinitesimal change in the extensive variables yields to an infinitesimal change in the total entropy, which is given by

$$
\begin{aligned}
dS_T &= \left[\left(\frac{\partial S_1}{\partial U_1} \right)_{V_1,N_1} - \left(\frac{\partial S_2}{\partial U_2} \right)_{V_2,N_2} \right] dU_1 \\
&+ \left[\left(\frac{\partial S_1}{\partial V_1} \right)_{U_1,N_1} - \left(\frac{\partial S_2}{\partial V_2} \right)_{U_2,N_2} \right] dV_1 \\
&+ \left[\left(\frac{\partial S_1}{\partial N_1} \right)_{U_1,V_1} - \left(\frac{\partial S_2}{\partial N_2} \right)_{U_2,V_2} \right] dN_1 \\
&= \left[\frac{1}{T_1} - \frac{1}{T_2} \right] dU_1 + \left[\frac{P_1}{T_1} - \frac{P_2}{T_2} \right] dV_1 - \left[\frac{\mu_1}{T_1} - \frac{\mu_2}{T_2} \right] dN_1
\end{aligned}
\tag{31}
$$

Since the total entropy is maximized in equilibrium with respect to the internal constrains, one has $dS_T = 0$ which for equilibrium in the presence of a particle exchange process yields the condition

$$\frac{\mu_1}{T_1} = \frac{\mu_2}{T_2} \tag{32}$$

On the other hand, if the systems initially have chemical potentials that differ infinitesimally from each other, then

$$dS_T = \left(\mu_2 - \mu_1 \right) \frac{dN_1}{T} > 0 \tag{33}$$

Hence, if $\mu_2 > \mu_1$ then $dN_1 > 0$. Therefore, particles flow from regions of higher chemical potential to regions of lower chemical potential.

Thus, two systems which are allowed to exchange energy, volume and particles have to satisfy the conditions

$$
\begin{aligned}
T_1 &= T_2 \\
P_1 &= P_2 \\
\mu_1 &= \mu_2
\end{aligned}
\tag{34}
$$

if they are in equilibrium.

2.4   The Equations of State

The fundamental relation $S(U, V, N)$ or alternately $U(S, V, N)$ provides a complete thermodynamic description of a system. From the fundamental relation one can derive three equations of state. The expressions for the intensive parameters are equations of state

$$
\begin{aligned}
T &= T(S, V, N) \\
P &= P(S, V, N) \\
\mu &= \mu(S, V, N)
\end{aligned}
\tag{35}
$$

These particular equations relate the intensive parameters to the independent extensive parameters. If all three equations of state are not known, then one has an incomplete thermodynamic description of the system.

If one knows all three equations of state, one can construct the fundamental relation and, hence, one has a complete thermodynamic description of the system. This can be seen by considering the extensive nature of the fundamental relation and it behavior under a change of scale by s. The fundamental equation is homogeneous and is of first order so

$$
U(sS, sV, SN) = s\, U(S, V, N)
\tag{36}
$$

Differentiating the above equation w.r.t s yields

$$
\left(\frac{\partial U}{\partial sS}\right)_{sV,sN}\left(\frac{dsS}{ds}\right) + \left(\frac{\partial U}{\partial sV}\right)_{sS,sN}\left(\frac{dsV}{ds}\right) + \left(\frac{\partial U}{\partial sN}\right)_{sS,sV}\left(\frac{dsN}{ds}\right) = U(S,V,N)
$$

$$
\left(\frac{\partial U}{\partial sS}\right)_{sV,sN} S + \left(\frac{\partial U}{\partial sV}\right)_{sS,sN} V + \left(\frac{\partial U}{\partial sN}\right)_{sS,sV} N = U(S,V,N)
$$

which, on setting $s = 1$, yields the Euler Equation

$$
\left(\frac{\partial U}{\partial S}\right)_{V,N} S + \left(\frac{\partial U}{\partial V}\right)_{S,N} V + \left(\frac{\partial U}{\partial N}\right)_{S,V} N = U
\tag{37}
$$

which, when expressed in terms of the intensive parameters, becomes

$$
T\,S - P\,V + \mu\,N = U
\tag{38}
$$

In the entropy representation, one finds the Euler equation in the form

$$\frac{1}{T}\, U \;+\; \frac{P}{T}\, V \;-\; \frac{\mu}{T}\, N \;=\; S \tag{39}$$

which has exactly the same content as the Euler equation found from the energy representation. From either of these equations it follows that knowledge of the three equations of state can be used to find the fundamental relation.

The three intensive parameters cannot be used as a set of independent variables. This can be seen by considering the infinitesimal variations of the Euler Equation

$$dU \;=\; T\, dS \;+\; S\, dT \;-\; P\, dV \;-\; V\, dP \;+\; \mu\, dN \;+\; N\, d\mu \tag{40}$$

and comparing it with the form of the first law of thermodynamics

$$dU \;=\; T\, dS \;-\; P\, dV \;+\; \mu\, dN \tag{41}$$

This leads to the discovery that the infinitesimal changes in the intensive parameters are related by the equation

$$0 \;=\; S\, dT \;-\; V\, dP \;+\; N\, d\mu \tag{42}$$

which is known as the Gibbs-Duhem relation. Thus, for a one component system, there are only two independent intensive parameters, i.e. there are only two thermodynamic degrees of freedom.

2.5 Thermodynamic Processes

Not all processes, that conserve energy, represent real physical processes. Since if the system is initially in a constrained equilibrium state, and an internal constraint is removed, then the final equilibrium state that is established must have a higher entropy.

A quasi static processes, is one that proceeds sufficiently slowly that its trajectory in thermodynamic phase space can can be approximated by a dense set of equilibrium states. Thus, at each macroscopic equilibrium state one can define an entropy $S_j \;=\; S(U_j, V_j, N_j, X_j)$. The quasi-static process is a temporal succession of equilibriums states, connected by non-equilibrium states. Since, for any specific substance, an equilibrium state can be characterized by $\{U, V, N, X\}$, a state can be represented by a point on a hyper-surface $S = S(U, V, N, X)$ in thermodynamic configuration space. The cuts of the hyper-surface at constant U are concave. The quasi-static processes trace out an almost continuous line on the hyper-surface. Since individual quasi-static processes are defined by sequence of equilibrium states connected by non-equilibrium states, the entropy cannot decrease along any part of the sequence if it is to represent a possible process, therefore, $S_{j+1} \;\geq\; S_j$. Thus, an allowed quasi-static process must follow a path on the hyper-surface which never has a segment on which S decreases. A reversible process is an allowed quasi-static

process in which the overall entropy difference becomes infinitesimally small. Hence, a reversible process must proceed along a contour of the hyper-surface which has constant entropy. Therefore, reversible process occur on a constant entropy cut of the hyper-surface. The constant entropy cuts of the hyper-surface are convex.

Adiabatic Expansion of Electromagnetic Radiation in a Cavity

Consider a spherical cavity of radius R which contains electromagnetic radiation. The radius of the sphere expands slowly at a rate given by $\frac{dR}{dt}$. The spectral component of wavelength λ contained in the cavity will be changed by the expansion. The change occurs through a change of wavelength $d\lambda$ that occurs at reflection with the moving boundary. Since the velocity $\frac{dR}{dt}$ is much smaller than the speed of light C, one only needs to work to keep terms first-order in the velocity. A single reflection through an angle θ produces a Doppler
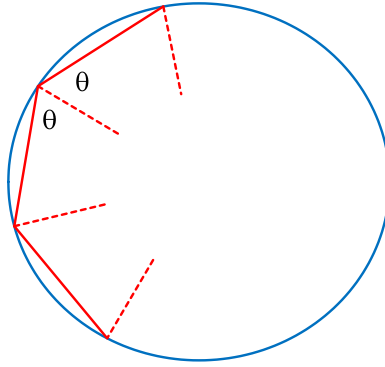
Figure 2: Electromagnetic radiation being reflected through an angle θ from the walls of a slowly expanding spherical electromagnetic cavity.

shift of the radiation by an amount given by

$$d\lambda \;=\; 2\;\frac{\lambda}{c}\;\frac{dR}{dt}\;\cos\theta \tag{43}$$

The ray travels a distance $2\;R\;\cos\theta$ between successive reflections. Hence, the time between successive reflections is given by

$$\frac{c}{2\;R\;\cos\theta} \tag{44}$$

Thus, the rate at which the wavelength changes is given by

$$\frac{d\lambda}{dt} \;=\; \frac{\lambda}{R}\;\frac{dR}{dt} \tag{45}$$

15

On integrating the above equation, one finds that

$$\frac{\lambda}{R} = \text{Constant} \tag{46}$$

Therefore, the wavelength scales with the radius. Quantum mechanically, each state evolves adiabatically so no transitions occur. The wavelength scales with the radius such as to match the boundary condition.

The equation of state for the electromagnetic radiation is

$$P = \frac{1}{3}\frac{U}{V} \tag{47}$$

so for adiabatic expansion (with $\mu = 0$ and $dN = 0$), one has

$$dU = -P\,dV \tag{48}$$

which leads to

$$\frac{dU}{U} = -\frac{1}{3}\frac{dV}{V} \tag{49}$$

Hence, for adiabatic expansion, one has

$$U\,V^{\frac{1}{3}} = \text{Constant} \tag{50}$$

Putting this together with Stefan's law

$$U = \sigma\,V\,T^4 \tag{51}$$

one finds that

$$\sigma\,T^4\,V^{\frac{4}{3}} = \text{Constant} \tag{52}$$

or $R\,T = \text{const}$. Thus, the temperature of the cavity decreases inversely with the radius of the cavity as R increases. Furthermore, since λ scales with R, one finds that the density of each spectral component must scales so that $\lambda\,T$ is constant.

2.6   Thermodynamic Potentials

The Postulate II is the entropy maximum principle which can be restated as the entropy of a system is maximized in equilibrium (at fixed total energy) with respect to variations of an internal parameter. This is a consequence of the concave geometry of the constant energy cuts of the hyper-surface in thermodynamic configuration space. We have formulated thermodynamics in terms of entropy and the extensive parameters. Thermodynamics has an equivalent formulation in terms of the energy and the extensive parameters. In this alternate formulation, the entropy maximum principle is replaced by the energy minimum principle. The energy minimum principle states that the energy is minimized in an equilibrium state with respect to variations of an internal parameter (for fixed values of the entropy). This is a consequence of the convex nature of

the constant entropy cuts of the hyper-surface in thermodynamic configuration space. The statements of the entropy maximum and the energy minimum principles are equivalent as can be seen from the following mathematical proof.

Equivalence of the Entropy Maximum and Energy Minimum Principles

The entropy maximum principle can be stated as

$$
\begin{aligned}
\left(\frac{\partial S}{\partial X}\right)_U &= 0 \\
\left(\frac{\partial^2 S}{\partial X^2}\right)_U &< 0
\end{aligned}
\tag{53}
$$

where X is an internal parameter. From the chain rule, it immediately follows that the energy is an extremum in equilibrium since

$$
\begin{aligned}
\left(\frac{\partial U}{\partial X}\right)_S &= -\frac{\left(\frac{\partial S}{\partial X}\right)_U}{\left(\frac{\partial S}{\partial U}\right)_X} \\
&= -T\left(\frac{\partial S}{\partial X}\right)_U \\
&= 0
\end{aligned}
\tag{54}
$$

Hence, it follows from the entropy maximum principle that the energy is an extremum.

The energy is extremum is a minimum, this follows by re-writing the second derivative of S as

$$
\left(\frac{\partial^2 S}{\partial X^2}\right)_U = \left[\frac{\partial}{\partial X}\left(\frac{\partial S}{\partial X}\right)_U\right]_U
\tag{55}
$$

and designating the internal derivative by A, i.e. let

$$
A = \left(\frac{\partial S}{\partial X}\right)_U
\tag{56}
$$

so the entropy maximum principle requires that

$$
\left(\frac{\partial^2 S}{\partial X^2}\right)_U = \left(\frac{\partial A}{\partial X}\right)_U < 0
\tag{57}
$$

If we consider A to be a function of (X, U) instead of (X, S), i.e. $A(X, S) = A(X, S(X, U))$, then

$$
\left(\frac{\partial A}{\partial X}\right)_U = \left(\frac{\partial A}{\partial X}\right)_S + \left(\frac{\partial A}{\partial S}\right)_X \left(\frac{\partial S}{\partial X}\right)_U
\tag{58}
$$

17

where the last term vanishes because of the entropy maximum principle. Hence,

$$\left(\frac{\partial A}{\partial X}\right)_U = \left(\frac{\partial A}{\partial X}\right)_S \tag{59}$$

Thus, we have

$$\left(\frac{\partial^2 S}{\partial X^2}\right)_U = \left[\frac{\partial}{\partial X}\left(\frac{\partial S}{\partial X}\right)_U\right]_S < 0 \tag{60}$$

Using the chain rule, the innermost partial derivative can be re-written as

$$\left(\frac{\partial S}{\partial X}\right)_U = -\left(\frac{\partial U}{\partial X}\right)_S \left(\frac{\partial S}{\partial U}\right)_X \tag{61}$$

Hence, on substituting this into the maximum principle, one has

$$
\begin{aligned}
\left(\frac{\partial^2 S}{\partial X^2}\right)_U &= -\left[\frac{\partial}{\partial X}\left(\frac{\partial U}{\partial X}\right)_S \left(\frac{\partial S}{\partial U}\right)_X\right]_S \\
&= -\left(\frac{\partial^2 U}{\partial X^2}\right)_S \left(\frac{\partial S}{\partial U}\right)_X - \left(\frac{\partial U}{\partial X}\right)_S \left[\frac{\partial}{\partial X}\left(\frac{\partial S}{\partial U}\right)_X\right]_S
\end{aligned} \tag{62}
$$

The last term vanishes since we have shown that the energy satisfies an extremum principle. Therefore, one has

$$
\begin{aligned}
\left(\frac{\partial^2 S}{\partial X^2}\right)_U &= -\left(\frac{\partial^2 U}{\partial X^2}\right)_S \left(\frac{\partial S}{\partial U}\right)_X \\
&= -\frac{1}{T}\left(\frac{\partial^2 U}{\partial X^2}\right)_S \\
&< 0
\end{aligned} \tag{63}
$$

Thus, since $T > 0$, we have

$$\left(\frac{\partial^2 U}{\partial X^2}\right)_S > 0 \tag{64}$$

so the energy satisfies the minimum principle if the entropy satisfies the maximum principle. The proof also shows that the energy minimum principle implies the entropy maximum principle, so the two principles are equivalent.

Sometimes it is more convenient to work with the intensive parameters rather than the extensive parameters. The intensive parameters are defined in terms of the partial derivatives of the fundamental relation $S(U, V, N)$ or equivalently $U(S, V, N)$. Taking partial derivatives usually leads to a loss of information, in the sense that a function can only be re-created from its derivative by integration up to a constant (or more precisely a function) of integration. Therefore, to avoid loss of information, one changes extensive variables to intensive variables by performing Legendre transformations.

18

Legendre Transformations

The Legendre transformation relies on the property of concavity of $S(E, V, N)$ and is introduced so that one can work with a set of more convenient variables, such as T instead of S or P instead of V. This amounts to transforming from an extensive parameter to its conjugate intensive parameter which is introduced as a derivative.

The Legendre transformation is introduced such that the change of variables is easily invertable. Instead of considering the convex function $y = y(x)$ [1] being given by the ordered pair (x, y) for each x, one can equally describe the curve by an envelope of a family of tangents to the curve. The tangent is a straight line

$$y = p\,x + \psi(p) \tag{65}$$

with slope p and has a y-axis intercept denoted by $\psi(p)$. Due to the property of convexity, for each value of p there is a unique tangent to the curve. Hence, we have replaced the sets of pairs (x, y) with a set of pairs (p, ψ). The set of pairs (p, ψ) describes the same curve and has the same information as the set of pairs (x, y).
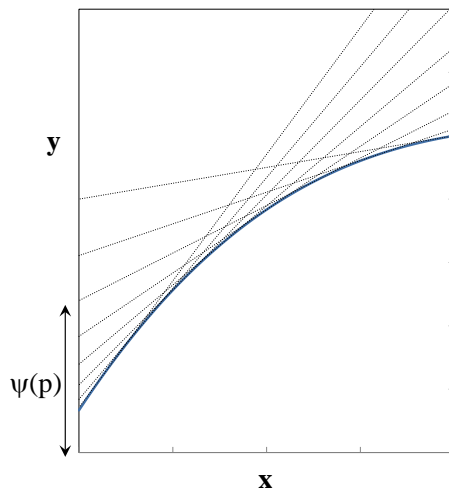
Figure 3: A concave function $y(x)$ is specified by the envelope of a family of tangents with slopes p and y-axis intercepts $\psi(p)$.

Given a curve in the form of $y = y(x)$, one can find $\psi(p)$ by taking the derivative to yield

$$p = \frac{dy}{dx} \tag{66}$$

---

[1] The convexity and concavity of a function implies that the second derivative of the function has a specific sign. All that we shall require is that the second derivative of the function does not go to zero in the interval of x that is under consideration.

which specifies the slope p of the tangent line at the tangent point x. The above equation can be inverted to yield $x = x(p)$ and, hence, one can obtain $y(p)$ from $y = y(x(p))$. Then, the y-axis intercept of the tangent can be found as a function of p from

$$\psi(p) \;=\; y(p) \;-\; p\,x(p) \tag{67}$$

The function $\psi(p)$ is the Legendre transform of $y(x)$. The quantity $\psi(p)$ contains exactly the same information as $y(x)$ but depends on the variable p instead of the x variable.

The inverse transform can be found by constructing (x, y) from (p, ψ). First the point x at which a tangent with slope p touches the curve is found. Second, after inverting $x(p)$ to yield $p(x)$, one finds $y(x)$ from

$$y \;=\; p(x) \;+\; \psi(p(x)) \tag{68}$$

The point of tangency x is found by considering a tangent with slope p and a neighboring tangent with slope $p + dp$. The tangent is described by

$$y \;=\; p\,x \;+\; \psi(p) \tag{69}$$

which is valid everywhere on the tangent including the point of tangency which we denote by (x, y). The neighboring tangent which has an infinitesimally different slope $p + dp$ is described by a similar equation, but has a point of tangency $(x + dx, y + dy)$ that differs infinitesimally from (x, y). To first-order in the infinitesimals, one finds the coordinates describing the separation of the two points of tangency are related by

$$dy \;=\; p\,dx \;+\; (\,x \;+\; \frac{d\psi}{dp}\,)\,dp \tag{70}$$

However, since the two neighboring points of tangency lie on the same curve and because the slope of the tangent is p, one has

$$dy \;=\; p\,dx \tag{71}$$

Thus, we find that the x-coordinate of the point of tangency is determined by the equation

$$x \;=\; -\,\frac{d\psi}{dp} \tag{72}$$

The abscissa is given by

$$y(x) \;=\; \psi \;+\; x\,p \tag{73}$$

in which p has been expressed in terms of x. This is the inverse Legendre transformation.

The inverse Legendre transformation should be compared to the Legendre transformation

$$\psi(p) \;=\; y \;-\; x\,p \tag{74}$$

in which x has been expressed in terms of p via inversion of

$$p \ = \ \frac{dy}{dx} \tag{75}$$

Thus, the relation between (x, y) and (p, ψ) is, apart from a minus sign, symmetrical between the Legendre and inverse Legendre transformations.

The Helmholtz Free-Energy F

The Helmholtz Free-Energy is denoted by F is a function of the variables (T, V, N) and is obtained by performing a Legendre transform on the energy $U(S, V, N)$. The process involves defining the temperature T via the derivative

$$T \ = \ \left(\frac{\partial U}{\partial S}\right)_{V,N} \tag{76}$$

and then defining a quantity F via

$$F \ = \ U \ - \ T \, S \tag{77}$$

The definition of T is used to express S as a function of T. Then eliminating S from the two terms in the above expression for F, yields the Helmholtz Free-Energy $F(T, V, N)$.

One can show that F does not depend on S by considering an infinitesimal transformation

$$dF \ = \ dU \ - \ S \, dT \ - \ T \, dS \tag{78}$$

and then by substituting the expression

$$dU \ = \ T \, dS - \ P \, dV \ + \ \mu \, dN \tag{79}$$

obtained from $U(S, V, N)$ and the definition of the energetic extensive parameters. Substitution of the expression for dU into dF yields

$$dF \ = \ - \ S \, dT \ - \ P \, dV \ + \ \mu \, dN \tag{80}$$

which shows that F only varies with T, V and N. It does not vary as dS is varied. Thus F is a function of the variables (T, V, N). Furthermore, we see that S can be found from F as a derivative

$$S \ = \ - \ \left(\frac{\partial F}{\partial T}\right)_{V,N} \tag{81}$$

The Helmholtz Free-Energy has the interpretation that it represents the work done on the system in a process carried out at constant T (and N). This can be seen from the above infinitesimal form of dF since, under the condition that $dT \ = \ 0$, one has

$$dF \ = \ - \ P \, dV \tag{82}$$

The inverse transform is given found by starting from $F(T, V, N)$ and expressing S as

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \tag{83}$$

This equation is used to express T as a function of S, i.e. $T = T(S)$. The quantity U is formed via

$$U = F + T S \tag{84}$$

Elimination of T in favour of S in both terms leads to $U(S, V, N)$ the energy.

The Enthalpy $\mathcal{H}$

The enthalpy is denoted by $\mathcal{H}$ and is a function of the variables (S, P, N). It is obtained by a Legendre transform on $U(S, V, N)$ which eliminates the extensive variable V and introduces the intensive variable P. The pressure P is defined by the equation

$$-P = \left(\frac{\partial U}{\partial V}\right)_{S,N} \tag{85}$$

and then one forms the quantity $\mathcal{H}$ via

$$\mathcal{H} = U + P V \tag{86}$$

which on inverting the equation expressing pressure as a function of V and eliminating V from $\mathcal{H}$, one obtains the enthalpy $\mathcal{H}(S, P, N)$.

The enthalpy is a function of (S, P, N) as can be seen directly from the infinitesimal variation of $\mathcal{H}$. Since

$$d\mathcal{H} = dU + V\, dP + P\, dV \tag{87}$$

and as

$$dU = T\, dS - P\, dV + \mu\, dN \tag{88}$$

one finds that

$$d\mathcal{H} = T\, dS + V\, dP + \mu\, dN \tag{89}$$

which shows that $\mathcal{H}$ only varies when S, P and N are varied. The above infinitesimal relation also shows that

$$V = \left(\frac{\partial \mathcal{H}}{\partial P}\right)_{S,N} \tag{90}$$

The enthalpy has the interpretation that it represents the heat flowing into a system in a process at constant pressure (and constant N). This can be seen from the expression for the infinitesimal change in $\mathcal{H}$ when $dP = 0$

$$d\mathcal{H} = T\, dS \tag{91}$$

which is recognized as an expression for the heat flow into the system.

The inverse Legendre transform of $\mathcal{H}(S, P, N)$ is $U(S, V, N)$ and is performed by using the relation

$$V = \left(\frac{\partial \mathcal{H}}{\partial P} \right)_{S,N} \tag{92}$$

to express P as a function of V. On forming U via

$$U = \mathcal{H} - P V \tag{93}$$

and eliminating P from U, one has the energy $U(S, V, N)$.

The Gibbs Free-Energy G

The Gibbs Free-Energy $G(T, P, N)$ is formed by making two Legendre transformation on $U(S, V, N)$ eliminating the extensive variables S and V and introducing their conjugate intensive parameters T and P. The process starts with $U(Sv, N)$ and defines the two intensive parameters T and $-P$ via

$$T = \left(\frac{\partial U}{\partial S} \right)_{V,N}$$

$$-P = \left(\frac{\partial U}{\partial V} \right)_{S,N} \tag{94}$$

The quantity G is formed via

$$G = U - T S + P V \tag{95}$$

which on eliminating S and V leads to the Gibbs Free-Energy $G(T, P, N)$.

On performing infinitesimal variations of S, T, V, P and N, one finds the infinitesimal change in G is given by

$$dG = dU - T\, dS - S\, dT + P\, dV + V\, dP \tag{96}$$

which on eliminating dU by using the equation

$$dU = T\, dS - P\, dV + \mu\, dN \tag{97}$$

leads to

$$dG = -S\, dT + V\, dP + \mu\, dN \tag{98}$$

This confirms that the Gibbs Free-Energy is a function of T, P and N, $G(T, P, N)$. It also shows that

$$-S = \left(\frac{\partial G}{\partial T} \right)_{P,N}$$

$$V = \left(\frac{\partial G}{\partial P} \right)_{T,N} \tag{99}$$

The inverse (double) Legendre transform of $G(T, P, N)$ yields $U(S, V, N)$ and is performed by expressing the extensive parameters as

$$
\begin{aligned}
-S &= \left(\frac{\partial G}{\partial T}\right)_{P,N} \\
V &= \left(\frac{\partial G}{\partial P}\right)_{T,N}
\end{aligned}
\tag{100}
$$

and using these to express T in terms of S and P in terms of V. The energy is formed via

$$
U = G + T S - P V
\tag{101}
$$

and eliminating T and P in favour of S and V, to obtain $S(U, V, N)$.

The Grand-Canonical Potential Ω

The Grand-Canonical Potential $\Omega(T, V, \mu)$ is a function of T, V and μ. It is obtained by making a double Legendre transform on $U(S, V, N)$ which eliminates S and N and replaces them by the intensive parameters T and μ. This thermodynamic potential is frequently used in Statistical Mechanics when working with the Grand-Canonical Ensemble, in which the energy and number of particles are allowed to vary as they are exchanged with a thermal and particle reservoir which has a fixed T and a fixed μ.

The double Legendre transformation involves the two intensive parameters defined by

$$
\begin{aligned}
T &= \left(\frac{\partial U}{\partial S}\right)_{V,N} \\
\mu &= \left(\frac{\partial U}{\partial N}\right)_{S,V}
\end{aligned}
\tag{102}
$$

The quantity Ω is formed as

$$
\Omega = U - T S - \mu N
\tag{103}
$$

elimination of the extensive variables S and N leads to $\Omega(T, V, \mu)$, the Grand-Canonical Potential.

The infinitesimal change in Ω is given by

$$
d\Omega = dU - T\, dS - S\, dT - \mu\, dN - N\, d\mu
\tag{104}
$$

which, on substituting for dU, leads to

$$
d\Omega = -S\, dT - P\, dV - N\, d\mu
\tag{105}
$$

The above equation confirms that Ω only depends on the variables T, V and μ. Furthermore, this relation also shows that

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

$$- N = \left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} \tag{106}$$

The inverse (double) transformation uses the two relations

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

$$- N = \left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} \tag{107}$$

to express T and μ in terms of S and N. The quantity U is formed via

$$U = \Omega + T S + \mu N \tag{108}$$

which on eliminating T and μ leads to the energy as a function of S, V and N, i.e. $U(S, V, N)$.

As examples of the use of thermodynamic potentials, we shall consider the processes of Joule Free Expansion and the Joule-Thomson Throttling Process.

Joule Free Expansion

Consider a closed system which is composed of two chambers connected by a valve. Initially, one chamber is filled with gas and the second chamber is evacuated. The valve connecting the two chambers is opened so that the gas can expand into the vacuum.

The expansion process occurs at constant energy, since no heat flows into the system and no work is done in expanding into a vacuum. Hence, the process occurs at constant U.

Due to the expansion the volume of the gas changes by an amount ΔV and, therefore, one might expect that the temperature of the gas may change by an amount ΔT. For a sufficiently small change in volume ΔV, one expects that ΔT and ΔV are related by

$$\Delta T = \left(\frac{\partial T}{\partial V}\right)_{U,N} \Delta V \tag{109}$$

where U is being kept constant.

On applying the chain rule, one finds

$$\Delta T = - \left[\left(\frac{\partial U}{\partial V}\right)_{T,N} \Big/ \left(\frac{\partial U}{\partial T}\right)_{V,N} \right] \Delta V \tag{110}$$

However, from the expression for the infinitesimal change in U

$$dU = T \, dS - P \, dV \tag{111}$$

25

one finds that the numerator can be expressed as

$$\left(\frac{\partial U}{\partial V}\right)_{T,N} = T \left(\frac{\partial S}{\partial V}\right)_{T,N} - P \tag{112}$$

whereas the denominator is identified as

$$\left(\frac{\partial U}{\partial T}\right)_{V,N} = C_V \tag{113}$$

which is the specific heat at constant volume.

The quantity proportional to

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} \tag{114}$$

is not expressed in terms of directly measurable quantities. It can be expressed as a derivative of pressure by using a Maxwell relation. We note that the quantity should be considered as a function of V, which is being varied and is also a function of T which is being held constant. Processes which are described in terms of the variables V and T can be described by the Helmholtz Free-Energy $F(T, V, N)$, for which

$$dF = - S\ dT - P\ dV + \mu\ dN \tag{115}$$

The Helmholtz Free-Energy is an analytic function of T and V, therefore, it satisfies the Cauchy-Riemann condition

$$\left(\frac{\partial^2 F}{\partial V \partial T}\right)_N = \left(\frac{\partial^2 F}{\partial T \partial V}\right)_N \tag{116}$$

which on using the infinitesimal form of dF to identify the inner partial differentials of F, yields the Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \left(\frac{\partial P}{\partial T}\right)_{V,N} \tag{117}$$

Hence, the temperature change and volume change that occur in Joule Free Expansion are related via

$$\Delta T = - \left[\frac{T \left(\frac{\partial P}{\partial T}\right)_{V,N} - P}{C_V} \right] \Delta V \tag{118}$$

which can be evaluated with the knowledge of the equation of state.

Since the expansion occurs at constant energy, one finds from

$$dU = T\ dS - P\ dV = 0 \tag{119}$$

that

$$\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{P}{T} \tag{120}$$

Thus, the entropy increases on expansion, as it should for an irreversible process.

Joule-Thomson Throttling Process

The Joule-Thomson throttling process involves the constant flow of fluid through a porus plug. The flowing fluid is adiabatically insulated so that heat cannot flow into or out of the fluid. The temperature and pressure of the fluid on either side of the porus plug are uniform but are not equal

$$T_1 \neq T_2 \tag{121}$$
$$P_1 \neq P_2 \tag{122}$$

Thus, a pressure drop ΔP defined by

$$\Delta P = P_1 - P_2 \tag{123}$$

and temperature drop ΔT defined by

$$\Delta T = T_1 - T_2 \tag{124}$$

occur across the porus plug.
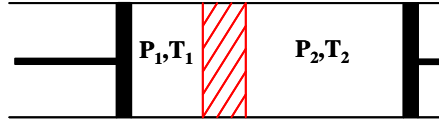
Figure 4: A fluid is confined in a cylindrical tube and two pistons (solid black objects). The pistons force the fluid through the porus plug (orange hatched region). In this process the pressure and temperature on each side of the plug are kept constant but not equal.

The Joule-Thomson process is a process for which the enthalpy $\mathcal{H}$ is constant. This can be seen by considering a fixed mass of fluid as it flows through the plug. The pump that generates the pressure difference can, hypothetically, be replaced by two pistons. Consider the volume of fluid contained in the volume V_1 between the piston and the plug, as having internal energy U_1. When this volume of gas has been pushed through the plug, the piston has performed an amount of work $P_1 V_1$. The piston on the other side of the porus plug performs a negative amount of work equal to $- P_2 V_2$ when the gas occupies the volume V_2 between the piston and the plug. The change in internal energy is given by

$$U_2 - U_1 = P_1 V_1 - P_2 V_2 \tag{125}$$

This implies that

$$U_1 + P_1 V_1 = U_2 + P_2 V_2 \tag{126}$$

or the enthalpy $\mathcal{H}$ of the fluid is constant in the throttling process.

For sufficiently small changes in pressure, the temperature drop is related to the pressure drop by

$$\Delta T = \left(\frac{\partial T}{\partial P}\right)_{\mathcal{H}} \Delta P \tag{127}$$

where the enthalpy $\mathcal{H}$ is being kept constant.

On applying the chain rule, one finds

$$\Delta T = -\left[\left(\frac{\partial \mathcal{H}}{\partial P}\right)_{T,N} \Big/ \left(\frac{\partial \mathcal{H}}{\partial T}\right)_{P,N} \right] \Delta P \tag{128}$$

However, from the expression for the infinitesimal change in $\mathcal{H}$

$$d\mathcal{H} = T \, dS + V \, dP \tag{129}$$

one finds that the numerator can be expressed as

$$\left(\frac{\partial \mathcal{H}}{\partial P}\right)_{T,N} = T \left(\frac{\partial S}{\partial P}\right)_{T,N} + V \tag{130}$$

whereas the denominator is identified as

$$\left(\frac{\partial \mathcal{H}}{\partial T}\right)_{P,N} = C_P \tag{131}$$

which is the specific heat at constant pressure.

The quantity proportional to

$$\left(\frac{\partial S}{\partial P}\right)_{T,N} \tag{132}$$

is not expressed in terms of directly measurable quantities. It can be expressed as a derivative of volume by using a Maxwell relation. We note that the quantity should be considered as a function of P, which is being varied and is also a function of T which is being held constant. Processes which are described in terms of the variables P and T can be described by the Gibbs Free-Energy $G(T, P, N)$, for which

$$dG = -S \, dT + V \, dP + \mu \, dN \tag{133}$$

The Gibbs Free-Energy is an analytic function of T and P, therefore, it satisfies the Cauchy-Riemann condition

$$\left(\frac{\partial^2 G}{\partial P \partial T}\right)_N = \left(\frac{\partial^2 G}{\partial T \partial P}\right)_N \tag{134}$$

28

which on using the infinitesimal form of dG to identify the inner partial differentials of G, yields the Maxwell relation

$$- \left(\frac{\partial S}{\partial P} \right)_{T,N} = \left(\frac{\partial V}{\partial T} \right)_{P,N} \tag{135}$$

Hence, the pressure change and volume change that occur in the Joule-Thomson process are related via

$$\Delta T = \left[\frac{T \left(\frac{\partial V}{\partial T} \right)_{P,N} - V}{C_P} \right] \Delta P \tag{136}$$

which can be evaluated with the knowledge of the equation of state.

Since the expansion occurs at constant enthalpy, one finds from

$$d\mathcal{H} = T \, dS + V \, dP = 0 \tag{137}$$

that

$$\left(\frac{\partial S}{\partial P} \right)_{\mathcal{H},N} = - \frac{V}{T} \tag{138}$$

Thus, the entropy increases for the irreversible Joule-Thomson process only if the pressure drops across the porus plug.

The description of the above processes used two of the Maxwell's relations. We shall give a fuller description of these relations below:

Maxwell Relations

The Maxwell Relations are statements about the analyticity of the thermodynamic potentials. The Maxwell relations are expressed in the form of an equality between the mixed second derivatives when taken in opposite order. If $B(x,y)$ is a thermodynamic potential which depends on the independent variables x and y, then analyticity implies that

$$\left(\frac{\partial^2 B}{\partial x \partial y} \right) = \left(\frac{\partial^2 B}{\partial y \partial x} \right) \tag{139}$$

The Maxwell relations for the four thermodynamic potentials which we have considered are described below:

The Internal Energy $U(S,V,N)$

Since the infinitesimal change in the internal energy is written as

$$dU = T \, dS - P \, dV + \mu \, dN \tag{140}$$

29

one has the three Maxwell relations

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

$$\left(\frac{\partial T}{\partial N}\right)_{S,V} = \left(\frac{\partial \mu}{\partial S}\right)_{V,N}$$

$$-\left(\frac{\partial P}{\partial N}\right)_{S,V} = \left(\frac{\partial \mu}{\partial V}\right)_{S,N} \tag{141}$$

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

Since the infinitesimal change in the Helmholtz Free-Energy is written as

$$dF = -S\ dT - P\ dV + \mu\ dN \tag{142}$$

one finds the relations

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

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

$$-\left(\frac{\partial P}{\partial N}\right)_{T,V} = \left(\frac{\partial \mu}{\partial V}\right)_{T,N} \tag{143}$$

The Enthalpy $\mathcal{H}(S,P,N)$

Since the infinitesimal change in the enthalpy is written as

$$d\mathcal{H} = T\ dS + V\ dP + \mu\ dN \tag{144}$$

one has

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

$$\left(\frac{\partial T}{\partial N}\right)_{S,P} = \left(\frac{\partial \mu}{\partial S}\right)_{P,N}$$

$$\left(\frac{\partial V}{\partial N}\right)_{S,P} = \left(\frac{\partial \mu}{\partial P}\right)_{S,N} \tag{145}$$

The Gibbs Free-Energy $G(T,P,N)$

Since the infinitesimal change in the Gibbs Free-Energy is written as

$$dG \; = \; - \; S \; dT \; + \; V \; dP \; + \; \mu \; dN \qquad (146)$$

one has

$$- \left(\frac{\partial S}{\partial P} \right)_{T,N} \; = \; \left(\frac{\partial V}{\partial T} \right)_{P,N}$$

$$- \left(\frac{\partial S}{\partial N} \right)_{T,P} \; = \; \left(\frac{\partial \mu}{\partial T} \right)_{P,N}$$

$$\left(\frac{\partial V}{\partial N} \right)_{T,P} \; = \; \left(\frac{\partial \mu}{\partial P} \right)_{T,N} \qquad (147)$$

The Grand-Canonical Potential $\Omega(T, V, \mu)$

Since the infinitesimal change in the Grand-Canonical Potential is written as

$$d\Omega \; = \; - \; S \; dT \; - \; P \; dV \; - \; N \; d\mu \qquad (148)$$

one finds the three Maxwell relations

$$- \left(\frac{\partial S}{\partial V} \right)_{T,\mu} \; = \; - \left(\frac{\partial P}{\partial T} \right)_{V,\mu}$$

$$- \left(\frac{\partial S}{\partial \mu} \right)_{T,V} \; = \; - \left(\frac{\partial N}{\partial T} \right)_{V,\mu}$$

$$- \left(\frac{\partial P}{\partial \mu} \right)_{T,V} \; = \; - \left(\frac{\partial N}{\partial V} \right)_{T,\mu} \qquad (149)$$

The Nernst Postulate

The Nernst postulate states that as $T \rightarrow 0$, then $S \rightarrow 0$. This postulate may not be universally valid. It can be motivated by noting that the specific heat C_V is positive, which implies that the internal energy U is a monotonically increasing function of temperature T. Conversely, if T decreases then U should decrease monotonically. Therefore, U should approach its smallest value as $T \rightarrow 0$ and the system should be in a quantum mechanical ground state. The ground state is unique if it is non-degenerate or, in the case where the ground state has a spontaneously broken symmetry, may have the degeneracy associated with the broken symmetry. In either case, since the entropy is proportional to the logarithm of the degeneracy, one expects the entropy at $T = 0$ to be a minimum. For degeneracies which are not exponential in the size of the system, the entropy is not extensive and, therefore, can be considered as being effectively zero in the thermodynamic limit $N \rightarrow 0$. This assumption might not be valid

for the case of highly frustrated systems such as ice or spin glasses, since these systems remain highly degenerate as $T \to 0$.

Classically, the entropy can only be defined up to an additive constant. Since classical states are continuous and, therefore, the "number of states" depends on the choice of the measure. Because of this, the classical version of Nernst's postulate states that the entropy reaches a universal minimum value in the limit $T \to 0$. Therefore, Walther Nernst's initial 1906 formulation was that the $T = 0$ isotherm is also an isentrope[2]. Max Planck's 1911 restatement of the postulate gave a value of zero to the entropy at $T = 0$. This restatement is frequently attributed to Simon [3].

Nernst's postulate has a number of consequences. For example, the specific heat vanishes as $T \to 0$. This follows if S approaches zero with a finite derivative, then

$$C_v = T \left(\frac{\partial S}{\partial T} \right)_V \to 0 \qquad \text{as} \quad T \to 0 \tag{150}$$

Likewise,

$$C_P = T \left(\frac{\partial S}{\partial T} \right)_P \to 0 \qquad \text{as} \quad T \to 0 \tag{151}$$

The thermal expansion coefficient also vanishes as $T \to 0$, as the Maxwell relation

$$\left(\frac{\partial V}{\partial T} \right)_{P,N} = - \left(\frac{\partial S}{\partial P} \right)_{T,N} \tag{152}$$

shows that

$$\left(\frac{\partial V}{\partial T} \right)_{P,N} \to 0 \qquad \text{as} \quad T \to 0 \tag{153}$$

Hence the coefficient of volume expansion α defined by

$$\alpha = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_{P,N} \to 0 \qquad \text{as} \quad T \to 0 \tag{154}$$

vanishes as $T \to 0$.

Likewise, from the Maxwell relation

$$\left(\frac{\partial P}{\partial T} \right)_{V,N} = \left(\frac{\partial S}{\partial V} \right)_{T,N} \tag{155}$$

one discovers that

$$\left(\frac{\partial P}{\partial T} \right)_{V,N} \to 0 \qquad \text{as} \quad T \to 0 \tag{156}$$

[2]W. Nernst, "Über die Beziehung zwischen Wärmeentwicklung und maximaler Arbeit bei kondensierten Systemen", Ber. Kgl. Pr. Akad. Wiss. **52**, 933-940, (1906).

[3]F. Simon and F. Lange, "Zur Frage die Entropie amorpher Substanzen", Zeit. für Physik, **38**, 227-236 (1926)

In the limit $T \to 0$, the difference between the specific heats at constant pressure and constant volume vanish with a higher power of T than the power of T with which the specific heats vanish.

From the above formula, one realizes that the classical ideal gas does not satisfy Nernst's postulate. However, quantum mechanical ideal gasses do satisfy Nernst's postulate.

Another consequence of the Nernst postulate is that the absolute zero temperature cannot be attained by any means. More precisely, it is impossible by any procedure, no matter how idealized, to reduce the temperature of any system to the absolute zero in a finite number of operations. First we shall consider the final step of a finite process. Cooling a substance below a bath temperature usually requires an adiabatic stage, since otherwise heat would leak from the bath to the system and thereby increase its temperature. Suppose that, by varying a parameter X from X_1 to X_2, one adiabatically cools a system from a finite temperature T_2 to a final temperature T_1. Then the adiabaticity condition requires

$$S(T_1, X_1) \;=\; S(T_2, X_2) \tag{157}$$

Furthermore, if we reduce the system's final temperature T_2 to zero, the right-hand side vanishes according to Simon's statement of Nernst's principle. Thus, we require

$$S(T_1, X_1) \;=\; 0 \tag{158}$$

which is impossible for real systems for which S is expected to only approach its minimal value in the limit $T \to 0$. Hence, this suggests that the final stages of the process must involve infinitesimal temperature differences. Such a processes is illustrated by a sequence of processes composed of adiabatic expansions between a high pressure P_1 and a low pressure P_2 followed by isothermal contractions between P_2 and P_1. The internal energy and temperature is lowered during the adiabatic expansion stages. The curves of entropy versus temper-
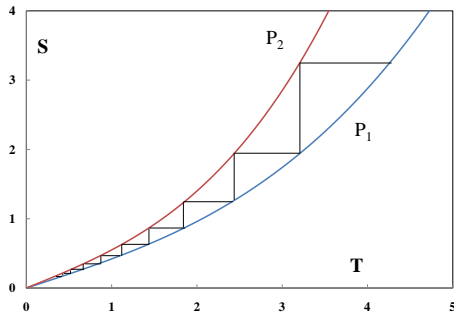
Figure 5: The unattainability of $T = 0$ is illustrated by a substance which undergoes a series of stages composed of an adiabatic expansion followed by an isothermal compression. An infinite number of stages would be required to reach $T = 0$.

ature for the different pressures must approach each other as $T \to 0$. Hence,

both the magnitudes of temperature changes and entropy changes decrease in the successive stages as T approaches zero. Therefore, absolute zero can only be attained for this sequence in the limit of an infinite number of stages. For these two example, the unattainability of absolute zero can simply be understood by noting that the adiabat becomes the isotherm as $T \to 0$.

Extremum Principles for Thermodynamic Potentials

The Energy Minimum Principle states that the equilibrium value of any unconstrained internal parameter X minimizes $U(S, V, X)$ for a fixed value of S. It can be stated in terms of the first-order and second-order infinitesimal changes

$$
\begin{aligned}
dU &= 0 \\
d^2U &\geq 0
\end{aligned}
\tag{159}
$$

where S is held constant.

This principle can be formulated in terms of a composite system which is composed of a system and a reservoir, for which the total energy and entropy are defined by

$$
\begin{aligned}
U_T &= U + U_R \\
S_T &= S + S_R
\end{aligned}
\tag{160}
$$

The energy minimum principle applied to the combined system becomes

$$
\begin{aligned}
dU + dU_R &= 0 \\
d^2(U + U_R) &\geq 0
\end{aligned}
\tag{161}
$$

where¡ since S_T is constant, $dS_R = -dS$. We also note that, if the reservoir is sufficiently larger than the system one may set $d^2U_R = 0$, in which case the second line simplifies to

$$
d^2U \geq 0
\tag{162}
$$

For a system in thermal contact with a reservoir at constant temperature T, the infinitesimal change in internal energy of the reservoir is given by the heat it absorbs

$$
dU_R = T\, dS_R = -T\, dS
\tag{163}
$$

Hence, one has

$$
\begin{aligned}
dU + dU_R &= 0 \\
dU - T\, dS &= 0
\end{aligned}
\tag{164}
$$

which, if T is being held constant, leads to

$$d(\ U\ -\ T\ S\ )\ =\ 0$$
$$dF\ =\ 0 \qquad (165)$$

where F is defined as

$$F\ =\ U\ -\ T\ S \qquad (166)$$

Hence, the quantity F satisfies an extremum principle for processes at constant T. For a sufficiently large reservoir, one may set $d^2 U_R\ \approx\ 0$. This can be seen by examining the second-order change due to a fluctuation, say of the entropy. For this particular case,

$$
\begin{aligned}
d^2 U_R\ &=\ \left(\frac{\partial^2 U_R}{\partial S_R{}^2}\right)(dS_R)^2 \\
&=\ \left(\frac{\partial^2 U_R}{\partial S_R{}^2}\right)(dS)^2 \\
&=\ \frac{T}{C_R}\ (dS)^2 \qquad (167)
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
d^2 U\ &=\ \left(\frac{\partial^2 U}{\partial S^2}\right)(dS)^2 \\
&=\ \frac{T}{C}\ (dS)^2 \qquad (168)
\end{aligned}
$$

Therefore, if $C_R\ \gg\ C$, one has $d^2 U\ \gg\ d^2 U_R$. Applying this type of consideration to the fluctuations of any set of extensive variables leads to the same conclusion. The extremum principle is a minimum principle since

$$
\begin{aligned}
d^2 U\ &=\ d^2(\ U\ -\ T\ S\ ) \\
&\geq\ 0 \qquad (169)
\end{aligned}
$$

where the first line holds, since T is being held constant and since S is an independent variable, so the last term can only contribute to a first-order change $T\ dS$. Thus, one has the condition

$$d^2 F\ \geq\ 0 \qquad (170)$$

If F is reinterpreted in terms of the Helmholtz Free-Energy $F(T, V, N)$, this leads to the Helmholtz Minimum Principle. The Helmholtz Minimum Principle states that, for a system being held at constant temperature T, the equilibrium value of unconstrained internal parameter minimizes $F(T, V, X)$.

For a system in thermal contact with a pressure reservoir of pressure P, the infinitesimal change in internal energy of the reservoir is equal to the work done on it

$$dU_R\ =\ -\ P\ dV_R\ =\ P\ dV \qquad (171)$$

Hence, one has

$$
\begin{aligned}
dU \; + \; dU_R &= \; 0 \\
dU \; + \; P \; dV &= \; 0
\end{aligned}
\tag{172}
$$

which, if P is being held constant, leads to

$$
\begin{aligned}
d(\, U \; + \; P \; V \,) &= 0 \\
d\mathcal{H} &= 0
\end{aligned}
\tag{173}
$$

where $\mathcal{H}$ is defined as

$$
\mathcal{H} \; = \; U \; + \; P \; V
\tag{174}
$$

Hence, the quantity $\mathcal{H}$ satisfies an extremum principle for process at constant P. The extremum principle is a minimum principle since

$$
\begin{aligned}
d^2 U \; &= \; d^2(\, U \; + \; P \; V \,) \\
&\geq \; 0
\end{aligned}
\tag{175}
$$

where the first line holds since P is being held constant and V is an independent variable. Thus, one has the condition

$$
d^2\mathcal{H} \; \geq \; 0
\tag{176}
$$

The Enthalpy Minimum Principle states that, for a system being held at constant pressure P, the equilibrium value of unconstrained internal parameter minimizes $\mathcal{H}(S, P, X)$.

For a system in thermal contact with a reservoir at constant temperature T and constant pressure P

$$
dU_R \; = \; T \; dS_R \; - \; P \; dV_R \; = \; - \; T \; dS \; + \; P \; dV
\tag{177}
$$

Hence, one has

$$
\begin{aligned}
dU \; + \; dU_R &= \; 0 \\
dU \; - \; T \; dS \; + \; P \; dV &= \; 0
\end{aligned}
\tag{178}
$$

which, if T and P are being held constant, leads to

$$
\begin{aligned}
d(\, U \; - \; T \; S \; + \; P \; V \,) &= 0 \\
dG &= 0
\end{aligned}
\tag{179}
$$

where G is defined as

$$
G \; = \; U \; - \; T \; S \; + \; P \; V
\tag{180}
$$

Hence, the quantity G satisfies an extremum principle for process at constant T and P. The extremum principle is a minimum principle since

$$
\begin{aligned}
d^2 U \; &= \; d^2(\, U \; - \; T \; S \; + \; P \; V \,) \\
&\geq \; 0
\end{aligned}
\tag{181}
$$

36

where the first line holds since T and P are being held constant and since S and V are independent variables. Thus, one has the condition

$$d^2 G \ \geq \ 0 \qquad (182)$$

The Gibbs Minimum Principle states that, for a system being held at constant temperature T and pressure P, the equilibrium value of unconstrained internal parameter minimizes $G(T, P, X)$.

A perhaps clearer, but less general, derivation of the minimum principle for thermodynamic potentials can be found directly from the entropy maximum principle. As an example of a minimum principle for a thermodynamic potential, consider a closed system composed of a system and reservoir which are in thermal contact. The entropy of the combined system S_T is given by

$$S_T(U, V, N : U_T, V_T, N_T) \ = \ S(U, V, N) \ + \ S_R(U_T - U, V_T - V, N_T - N) \quad (183)$$

We shall consider the Taylor expansion of S_T in powers of U, and we shall assume that the reservoir is much bigger than the system so that the terms involving higher-order derivatives are negligibly small

$$
\begin{aligned}
S_T(U, V, N : U_T, V_T, N_T) \ &= \ S(U, V, N) \ + \ S_R(U_T, V_T - V, N_T - N) \ - \ \frac{U}{T_R} \ + \ \dots \\
&= \ S_R(U_T, V_T - V, N_T - N) \ + \ \left(\frac{T_R \, S(U, V, N) \ - \ U}{T_R} \right)
\end{aligned}
$$
$$(184)$$

where terms of the order N^2/N_R have been neglected. We note that the term in the round parenthesis is of order N and contains all the information about the subsystem of interest. The entropy maximum principle applied to the combined system then implies that, in equilibrium, one must have

$$\left(\frac{\partial S}{\partial U} \right)_{V,N} \ = \ \frac{1}{T_R} \qquad (185)$$

where T_R is the temperature of the thermal reservoir defined by the partial derivative $\frac{\partial S_R(U_T)}{\partial U_T}$. Also one has

$$\left(\frac{\partial^2 S}{\partial U^2} \right)_{V,N} \ \leq \ 0 \qquad (186)$$

Now consider the convex generalized thermodynamic function $\tilde{F}(U : T_R, V, N)$, previously identified in the expression for S_T, which is defined by

$$\tilde{F}(U : T_R, V, N) \ = \ U \ - \ T_R \, S(U, V, N) \qquad (187)$$

for some constant T_R. The first two derivatives of $\tilde{F}$ w.r.t. U are given by

$$\left(\frac{\partial \tilde{F}}{\partial U} \right)_{V,N} \ = \ 1 \ - \ T_R \left(\frac{\partial S}{\partial U} \right)_{V,N} \qquad (188)$$

37

and

$$\left(\frac{\partial^2 \tilde{F}}{\partial U^2}\right)_{V,N} = -T_R \left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N} \tag{189}$$

which shows that, if the parameter T_R is identified with the temperature T of the system, then $\tilde{F}$ satisfies a minimum principle and that the minimum value of $\tilde{F}$ is given by the Helmholtz Free-Energy $F(T,V,N)$.

2.7 Thermodynamic Stability

The condition of thermodynamic stability imposes the condition of convexity or concavity on the thermodynamic functions characterizing the system. Consider two identical systems in thermal contact. The entropy maximum principle holds for the combined system, of energy $2\ U$, volume $2\ V$ and a total number of particles $2\ N$. For the combined system to be stable against fluctuations of the energy, the entropy function must satisfy the inequality

$$S(2U, 2V, 2N) \geq S(U + \Delta U, V, N) + S(U - \Delta U, V, N) \tag{190}$$

for any value of ΔU. Due to the extensive nature of the entropy, this inequality can be re-written as

$$2\ S(U, V, N) \geq S(U + \Delta U, V, N) + S(U - \Delta U, V, N) \tag{191}$$

Geometrically, the inequality expresses the fact that any chord joining two points on the curve $S(U)$ must lie below the curve. Such a curve is known as a concave curve. In the limit $\Delta U \to 0$, one obtains the weaker stability condition
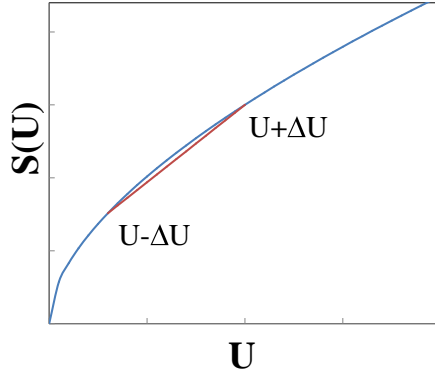
Figure 6: A concave curve representing $S(U)$. Any chord connecting two points on $S(U)$ must lie below the curve.

$$0 \geq \left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N} \tag{192}$$

This condition must hold if the macroscopic state of the system characterized by U, V, N is an equilibrium state. This condition can be re-stated as

$$
\begin{aligned}
0 &\geq \left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N} \\
&\geq -\frac{1}{T^2}\left(\frac{\partial T}{\partial U}\right)_{V,N} \\
&\geq -\frac{1}{T^2}\frac{1}{C_{V,N}} \qquad (193)
\end{aligned}
$$

Thus, for a system to be stable, its heat capacity at constant volume must be positive. This implies that the energy is a monotonically increasing function of temperature at constant volume.

Likewise, if the energy and volume are allowed to fluctuate, the condition for stability becomes

$$
2\,S(U,V,N) \geq S(U + \Delta U, V + \Delta V, N) + S(U - \Delta U, V - \Delta V, N) \quad (194)
$$

which can be expanded to yield

$$
0 \geq \left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N}\Delta U^2 + 2\left(\frac{\partial^2 S}{\partial U \partial V}\right)_{N}\Delta U\,\Delta V + \left(\frac{\partial^2 S}{\partial V^2}\right)_{U,N}\Delta V^2 \quad (195)
$$

The right hand side of this inequality can be expressed as the sum of two terms

$$
0 \geq \frac{1}{\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N}}\left[\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N}\Delta U + \left(\frac{\partial^2 S}{\partial U \partial V}\right)_{N}\Delta V\right]^2 + \left[\left(\frac{\partial^2 S}{\partial V^2}\right)_{U,N} - \frac{\left(\frac{\partial^2 S}{\partial U \partial V}\right)_{N}^2}{\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N}}\right]\Delta V^2
$$

$$(196)$$

This leads to two weak conditions for stability, which are

$$
0 \geq \left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N} \qquad (197)
$$

and

$$
0 \geq \left[\left(\frac{\partial^2 S}{\partial V^2}\right)_{U,N} - \frac{\left(\frac{\partial^2 S}{\partial U \partial V}\right)_{N}^2}{\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N}}\right] \qquad (198)
$$

The last condition can be re-stated as

$$
\left(\frac{\partial^2 S}{\partial V^2}\right)_{U,N}\left(\frac{\partial^2 S}{\partial U^2}\right)_{V,N} \geq \left(\frac{\partial^2 S}{\partial U \partial V}\right)_{N}^2 \qquad (199)
$$

which is a condition on the determinant of the matrix of the second-order derivatives. The two by two matrix is a particular example of a Hessian Matrix which,

more generally, is an N by N matrix of the second-order derivatives of a function of N independent variables. The Hessian is the determinant of the Hessian matrix. The Hessian describes the local curvature of the function. Although the above two conditions have been derived for two identical subsystems, they can be applied to any macroscopic part of a homogeneous system since thermodynamic quantities are uniformly distributed throughout the system.

Stability Conditions for Thermodynamic Potentials

The energy satisfies a minimum principle, which is reflected in the behavior of the thermodynamic potentials. Therefore, the convexity of the thermodynamic potentials can be used to obtain stability conditions for the thermodynamic potentials.

The energy $U(S, V, N)$ satisfies a minimum principle. For a system composed of two identical subsystems each with entropy S, volume V and number of particles N, the condition for equilibrium under interchange of entropy and volume is given by

$$U(S + \Delta S, V + \Delta V, N) \ + \ U(S - \Delta S, V - \Delta V, N) \ > \ 2\, U(S, V, N) \quad (200)$$

For stability against entropy fluctuations, one has

$$\left(\frac{\partial^2 U}{\partial S^2}\right)_{V,N} \quad \geq \quad 0$$

$$\left(\frac{\partial T}{\partial S}\right)_{V,N} \geq 0 \tag{201}$$

which leads to the condition $C_V \geq 0$ i.e. the specific heat at constant volume is always positive. Stability against volume fluctuations leads to

$$\left(\frac{\partial^2 U}{\partial V^2}\right)_{S,N} \quad \geq \quad 0$$

$$-\left(\frac{\partial P}{\partial V}\right)_{S,N} \geq 0 \tag{202}$$

Thus, the entropy is a convex function of the extensive variables and the convexity leads to stability conditions against fluctuations of the extensive variables which always have the same signs. However, stability against fluctuations of both S and V leads to a more complex and less restrictive condition

$$\left(\frac{\partial^2 U}{\partial V^2}\right)_{S,N} \left(\frac{\partial^2 U}{\partial S^2}\right)_{V,N} \geq \left(\frac{\partial^2 U}{\partial S \partial V}\right)_{N}^{2} \tag{203}$$

This can be shown to lead to the condition

$$-\left(\frac{\partial P}{\partial V}\right)_{T,N} \geq 0 \tag{204}$$

i.e. an increase in volume at constant temperature is always accompanied by a decrease in pressure.

The extension of the stability conditions to the thermodynamic potentials involves some consideration of the properties of the Legendre transform. It will be seen that the thermodynamic potentials are convex functions of their extensive variables but are concave functions of their intensive variables.

Consider a function $y(x)$ which satisfies a minimum condition. The Legendre transform of $y(x)$ is $\psi(p)$. One notes that the Legendre transform and inverse Legendre transform introduces the conjugate variables x and p via

$$p = \left(\frac{\partial y}{\partial x} \right) \tag{205}$$

and

$$x = - \left(\frac{\partial \psi}{\partial p} \right) \tag{206}$$

These relations lead to

$$\frac{\partial p}{\partial x} = \left(\frac{\partial^2 y}{\partial x^2} \right) \tag{207}$$

and

$$\frac{\partial x}{\partial p} = - \left(\frac{\partial^2 \psi}{\partial p^2} \right) \tag{208}$$

Thus, on equating the expressions for $\frac{dp}{dx}$, one has

$$\frac{\partial p}{\partial x} = \left(\frac{\partial^2 y}{\partial x^2} \right) = - \frac{1}{\left(\frac{\partial^2 \psi}{\partial p^2} \right)} \tag{209}$$

which shows that the sign of the second derivative w.r.t. the conjugate variables changes under the Legendre transform. Therefore, the condition for stability against fluctuations in x when expressed in terms of the thermodynamic potential y has the opposite sign to the condition for stability of fluctuations in p when expressed in terms of ψ. The stability condition for fluctuations of the other variables (which are not involved in the Legendre transform) have the same sign for both y and ψ.

The Helmholtz Free-Energy $F(T, V, N)$ is derived from the Legendre transform of $U(S, V, N)$ by eliminating the extensive variable S in favour of the intensive variable T. The condition for stability against temperature fluctuations is expressed in terms of $F(T, V, N)$ as

$$\left(\frac{\partial^2 F}{\partial T^2} \right)_{V,N} \leq 0 \tag{210}$$

which has the opposite sign as the stability conditions against entropy fluctuations when expressed in terms of $U(S, V, N)$. Stability against volume fluctuations leads to

$$\left(\frac{\partial^2 F}{\partial V^2} \right)_{T,N} \geq 0 \tag{211}$$

which has the same sign as the stability conditions against volume fluctuations when expressed in terms of U.

The stability condition for the enthalpy $\mathcal{H}(S, P, N)$ against entropy fluctuations is given by

$$\left(\frac{\partial^2 \mathcal{H}}{\partial S^2}\right)_{P,N} \geq 0 \tag{212}$$

which has the same sign as the stability conditions against entropy fluctuations when expressed in terms of $U(S, V, N)$. Stability against pressure fluctuations leads to

$$\left(\frac{\partial^2 \mathcal{H}}{\partial P^2}\right)_{S,N} \leq 0 \tag{213}$$

which has the opposite sign as the stability conditions against volume fluctuations when expressed in terms of U.

The Gibbs Free-Energy involves a double Legendre transform of U, so both stability conditions have opposite signs. The condition for stability against temperature fluctuations is expressed in terms of $G(T, P, N)$ as

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_{P,N} \leq 0 \tag{214}$$

which has the opposite sign as the stability conditions against entropy fluctuations when expressed in terms of $U(S, V, N)$. Stability against pressure fluctuations leads to the condition

$$\left(\frac{\partial^2 G}{\partial P^2}\right)_{T,N} \leq 0 \tag{215}$$

which has the opposite sign as the stability condition against volume fluctuations when expressed in terms of U.

The stability against volume fluctuations of a system held at constant temperature is expressed in terms of the second derivative of the Helmholtz Free-Energy as

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} \geq 0 \tag{216}$$

This can be related to the inequality

$$\left(\frac{\partial^2 U}{\partial V^2}\right)_{S,N} \left(\frac{\partial^2 U}{\partial S^2}\right)_{V,N} \geq \left(\frac{\partial^2 U}{\partial S \partial V}\right)_{N}^{2} \tag{217}$$

describing the stability condition obtained the energy minimum principle. This can be proved by noting that the infinitesimal change in F shows that

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = -\left(\frac{\partial P}{\partial V}\right)_{T,N} \tag{218}$$

The derivative of P with respect to V at constant T can be expressed as a Jacobian

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = -\frac{\partial(P,T)}{\partial(V,T)} \tag{219}$$

Since we wish to express the inequality in terms of the energy, one should change variables from V and T to S and V. This can be achieved using the properties of the Jacobian

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = -\frac{\partial(P,T)}{\partial(S,V)}\frac{\partial(S,V)}{\partial(V,T)} \tag{220}$$

On using the antisymmetric nature of the Jacobian one can recognize that the second factor is a derivative of S with respect to T, with V being held constant

$$
\begin{aligned}
\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} &= \frac{\partial(P,T)}{\partial(S,V)}\frac{\partial(V,S)}{\partial(V,T)} \\
&= \left(\frac{\partial S}{\partial T}\right)_{V,N}\frac{\partial(P,T)}{\partial(S,V)} \\
&= \left(\frac{\partial S}{\partial T}\right)_{V,N}\left[\left(\frac{\partial P}{\partial S}\right)_V\left(\frac{\partial T}{\partial V}\right)_S - \left(\frac{\partial T}{\partial S}\right)_V\left(\frac{\partial P}{\partial V}\right)_S\right]
\end{aligned}
\tag{221}
$$

where the expression for the Jacobian has been used to obtain the last line. On recognizing that P and T are the energy intensive parameters, one can write

$$
\begin{aligned}
\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} &= \left(\frac{\partial S}{\partial T}\right)_{V,N}\left[\left(\frac{\partial P}{\partial S}\right)_V\left(\frac{\partial T}{\partial V}\right)_S - \left(\frac{\partial T}{\partial S}\right)_V\left(\frac{\partial P}{\partial V}\right)_S\right] \\
&= \left(\frac{\partial S}{\partial T}\right)_{V,N}\left[-\frac{\partial}{\partial S}\left(\frac{\partial U}{\partial V}\right)_S\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial S}\right)_V + \frac{\partial}{\partial S}\left(\frac{\partial U}{\partial S}\right)_V\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial V}\right)_S\right] \\
&= \left(\frac{\partial S}{\partial T}\right)_{V,N}\left[\left(\frac{\partial^2 U}{\partial S^2}\right)_V\left(\frac{\partial^2 U}{\partial V^2}\right)_S - \left(\frac{\partial^2 U}{\partial S\partial V}\right)^2\right]
\end{aligned}
\tag{222}
$$

where the last line has been obtained by using the analyticity of U. Finally, one can write

$$
\begin{aligned}
\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} &= \left(\frac{\partial S}{\partial T}\right)_{V,N}\left[\left(\frac{\partial^2 U}{\partial S^2}\right)_V\left(\frac{\partial^2 U}{\partial V^2}\right)_S - \left(\frac{\partial^2 U}{\partial S\partial V}\right)^2\right] \\
&= \frac{\left[\left(\frac{\partial^2 U}{\partial S^2}\right)_V\left(\frac{\partial^2 U}{\partial V^2}\right)_S - \left(\frac{\partial^2 U}{\partial S\partial V}\right)^2\right]}{\left(\frac{\partial T}{\partial S}\right)_{V,N}} \\
&= \frac{\left[\left(\frac{\partial^2 U}{\partial S^2}\right)_V\left(\frac{\partial^2 U}{\partial V^2}\right)_S - \left(\frac{\partial^2 U}{\partial S\partial V}\right)^2\right]}{\left(\frac{\partial^2 U}{\partial S^2}\right)_{V,N}}
\end{aligned}
\tag{223}
$$

43

which relates the stability condition against volume fluctuations at constant T to the stability condition for fluctuations in S and V.

Homework:

Prove the stability condition

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_P \left(\frac{\partial G}{\partial P^2}\right)_T - \left(\frac{\partial^2 G}{\partial T \partial P}\right)^2 \geq 0 \tag{224}$$

Physical Consequences of Stability

The convexity of F with respect to V has been shown to lead to the condition

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = -\left(\frac{\partial P}{\partial V}\right)_{T,N} \geq 0 \tag{225}$$

which can be expressed as

$$\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N} = \frac{1}{V \, \kappa_T} \geq 0 \tag{226}$$

Hence, the isothermal compressibility κ_T defined by

$$\kappa_T = -V \left(\frac{\partial P}{\partial V}\right)_T \geq 0 \tag{227}$$

must always be positive. Likewise, the concavity of F with respect to T leads to the stability condition

$$C_V \geq 0 \tag{228}$$

that the specific heat at constant volume C_V must always be positive.

Another consequence of thermodynamic stability is that a thermodynamic system which is composed of parts that are free to move with respect to each other, will become unstable if the temperature is negative. The entropy of the α-th component of the system is a function of the internal energy, that is U_α minus the kinetic energy $\frac{p_\alpha^2}{2m_\alpha}$. Since the entropies are additive, the total entropy is given by

$$S = \sum_\alpha S_\alpha \left(U_\alpha - \frac{p_\alpha^2}{2m_\alpha}\right) \tag{229}$$

but is subject to the constraint that the total momentum is conserved

$$\sum_\alpha \underline{p}_\alpha = \underline{P} \tag{230}$$

The entropy has to be maximized subject to the constraint. This can be performed by using Lagrange's method of undetermined multipliers. Thus, Φ is to be maximized with respect to $\underline{p}_\alpha$, where

$$\Phi = \sum_\alpha \left[S_\alpha \left(U_\alpha - \frac{p_\alpha^2}{2m_\alpha} \right) + \underline{\lambda} \cdot \underline{p}_\alpha \right] \tag{231}$$

Maximizing Φ w.r.t. $\underline{p}_\alpha$ leads to the equation

$$
\begin{aligned}
0 &= -\left(\frac{\partial S_\alpha}{\partial U_\alpha} \right) \frac{\underline{p}_\alpha}{m_\alpha} + \underline{\lambda} \\
&= -\frac{1}{T} \frac{\underline{p}_\alpha}{m_\alpha} + \underline{\lambda}
\end{aligned}
\tag{232}
$$

which leads to the velocities of each component being the same. Thus, no independent internal macroscopic linear motions are allowed in an equilibrium state.

For the stationary state to be stable against the momentum fluctuations of the α-th part, one requires that

$$-\frac{1}{m_\alpha\, T} \leq 0 \tag{233}$$

Therefore, stability against break-up of the system requires that $T \geq 0$.

Homework:

Prove the two equalities

$$C_P - C_V = T\, V\, \frac{\alpha^2}{\kappa_T} \tag{234}$$

and

$$\frac{\kappa_S}{\kappa_T} = \frac{C_V}{C_P} \tag{235}$$

Hence, prove that the stability conditions imply the inequalities

$$C_P \geq C_V \geq 0 \tag{236}$$

and

$$\kappa_T \geq \kappa_S \geq 0 \tag{237}$$

The above conditions for stability are necessary but not sufficient to establish that the equilibrium is completely stable, since a state may decrease its entropy when there are infinitesimally small fluctuations in its macroscopic parameters, but its entropy may increase if the deviations of the parameters have large values. Such states are known as metastable states. A system which is in a metastable state will remain there until a sufficiently large fluctuation occurs that will take the system into a new state that is more stable.
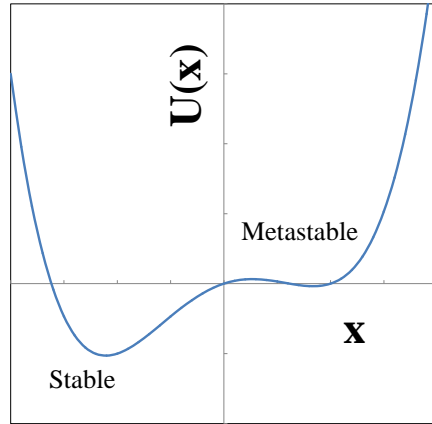
Figure 7: A curve of the internal energy $U(X)$ versus an internal variable X for a system which exhibits a stable and a metastable state.

3   Foundations of Statistical Mechanics

Statistical Mechanics provides us with:

(i) A basis for the first-principles calculations of thermodynamic quantities and transport coefficients of matter in terms of the dynamics of its microscopic constituents.

(ii) A physical significance for entropy.

3.1   Phase Space

In general, phase space Γ is the space of a set of ordered numbers which describes the microscopic states of a many-particle system. For a classical system, one can describe the state of a system by a set of continuously varying variables corresponding to the generalized momenta and generalized coordinates of each particle. However, for quantum systems, the Heisenberg uncertainty principle forbids one to know the momentum and position of any single particle precisely. In this case, the quantum states of a particle can be proscribed by specifying the eigenvalues of a mutually commuting set of operators representing physical observables. The eigenvalues can be either continuous or discrete. Thus, the phase space for a quantum system can either consist of a set of discrete numbers or can consist of a set of continuous numbers, as in the classical case.

Classical Phase Space

A microscopic state of a classical system of particles can be described by

proscribing all the microscopic coordinates and momenta describing the internal degrees of freedom.

For a classical system of N particles moving in a three-dimensional space, the state of one particle, at any instant of time, can be specified by proscribing the values of the three coordinates (q_1, q_2, q_3) and the values of the three canonically conjugate momenta (p_1, p_2, p_3).

The state of the many-particle system, at one instant of time, is proscribed by specifying the values of $3N$ coordinates q_i, $(i \in \{1, 2, 3, \ldots 3N\})$ and the values of $3N$ canonically conjugate momenta p_i, $(i \in \{1, 2, 3, \ldots 3N\})$. The space composed of the ordered set of $6N$ components of the coordinates and momenta is the phase space of the N particle system. This phase space has $6N$ dimensions.

Distinguishable Particles

For distinguishable particles for which each particle can be given a unique label, each point in phase space represents a unique microscopic state.

Indistinguishable Particles

By contrast, for indistinguishable particles it is not admissible to label particles. The material is invariant under all permutations of the sets of labels assigned to each of the N particles. There are $N!$ such permutations for the N particle system, and each one of these $N!$ permutations can be built by successively permuting the two sets of (six) labels assigned to pairs of particles. To be sure, the permutation of a particle described by the values of the ordered set of variables $\{q_1, q_2, q_3, p_1, p_2, p_3\}$ and a second particle described by the values of the ordered set $\{q_4, q_5, q_6, p_4, p_5, p_6\}$ is achieved by the interchange of the values $\{q_1, q_2, q_3, p_1, p_2, p_3\} \leftrightarrow \{q_4, q_5, q_6, p_4, p_5, p_6\}$. Any of these $N!$ permutations of the sets of labels assigned the N particles, has the action of transforming one point in phase space to a different point. Since it is not permissable to label indistinguishable particles, the resulting $N!$ different points in phase space must represent the same physical state.

The Number of Microscopic States.

Given the correspondence between points in phase space and microscopic states of the system, it is useful to introduce a measure of the number of microscopic states of a system N_Γ. One such measure is proportional to the volume of accessible phase space. Consider an infinitesimal volume element of phase space, defined by the conditions that the generalized momenta p_i lie in the intervals given by

$$\mathcal{P}_i + \Delta p_i > p_i > \mathcal{P}_i \tag{238}$$

and the coordinates q_i are restricted to the intervals

$$\mathcal{Q}_i + \Delta q_i > q_i \ > \ \mathcal{Q}_i \tag{239}$$

for all i. This infinitesimal volume element $\Delta\Gamma$ is given by

$$\Delta\Gamma \ = \ \prod_{i=1}^{3N} \left(\ \Delta p_i \ \Delta q_i \ \right) \tag{240}$$

The infinitesimal volume element of phase space $\Delta\Gamma$ has dimensions of $\Delta p^{3N} \ \Delta q^{3N}$. To turn this into a dimensionless quantity, one has to divide by a quantity with
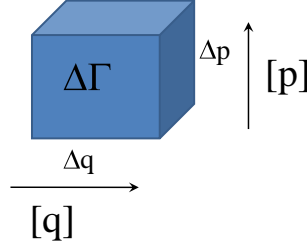
Figure 8: An infinitesimal hyper-cubic volume of phase space $\Delta\Gamma = \Delta p^{3N}\Delta q^{3N}$.

dimensions of $(\text{Action})^{3N}$. Although, any quantity with dimensions of the action would do, it is convenient, to use $2\,\pi\,\hbar$ as the measure for the action. With this particular choice, the dimensionless measure of the volume of phase space is given by

$$\frac{\Delta\Gamma}{(\ 2\ \pi\ \hbar\ )^{3N}} \ = \ \prod_{i=1}^{3N} \left(\ \frac{\Delta p_i \ \Delta q_i}{2\ \pi\ \hbar} \ \right) \tag{241}$$

The identification of $\hbar$ with Planck's constant is convenient since it allows one to make a connection with the number of quantum states, within the quasi-classical limit. The Heisenberg uncertainty principle dictates that the uncertainty in the momentum and position of a single-particle (wave-packet) state cannot be determined to better than $\Delta p_i \ \Delta q_i \ > \ 2\,\pi\,\hbar$. Hence, it appears to be reasonable to define the volume of phase space occupied by a single-particle state as $(\ 2\ pi\ \hbar)^3$ and so the dimensionless measure for the number of states for a single-particle system would be given by

$$\prod_{i=1}^{3} \left(\ \frac{\Delta p_i \ \Delta q_i}{2\ \pi\ \hbar} \ \right) \tag{242}$$

and consequently, the measure of the distinct microscopic states is given by

$$N_\Gamma \ = \ \frac{\Delta\Gamma}{(\ 2\ \pi\ \hbar\ )^{3N}} \ = \ \prod_{i=1}^{3N} \left(\ \frac{\Delta p_i \ \Delta q_i}{2\ \pi\ \hbar} \ \right) \tag{243}$$

for a system of N distinguishable particles. If the particles are indistinguishable, the number of distinct microscopic states N_Γ is defined as

$$N_\Gamma \;=\; \frac{\Delta\Gamma}{N! \, (\, 2 \, \pi \, \hbar \,)^{3N}} \tag{244}$$

where we have divided by $N!$ which is the number of permutations of the N sets of particle labels.

3.2 Trajectories in Phase Space

As time evolves, the system is also expected to evolve with time. For a classical system, the time evolution of the coordinates and momenta are governed by Hamilton's equations of motion, and the initial point in phase space will map out a trajectory in the $6N$ dimensional phase space. A closed system, where no time-dependent external fields are present, the Hamiltonian is a function of the set of $3N$ generalized momenta and the $3N$ generalized coordinates $H(\{p_i, q_i\})$ and has no explicit time dependence. The rate of change of $\{p_i, q_i\}$ where $i \in \{1, 2, 3, \ldots 3N\}$ are given by the set of Hamilton's equations of motion

$$
\begin{aligned}
\frac{dp_i}{dt} &= \{ \, p_i \, , \, H \, \}_{PB} \;=\; -\,\frac{\partial H}{\partial q_i} \\
\frac{dq_i}{dt} &= \{ \, q_i \, , \, H \, \}_{PB} \;=\; +\,\frac{\partial H}{\partial p_i}
\end{aligned}
\tag{245}
$$

where P.B. denotes the Poisson Bracket. The Poisson Bracket of two quantities A and B is defined as the antisymmetric quantity

$$\{ \, A \, , \, B \, \}_{PB} \;=\; \sum_{i=1}^{3N} \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i}\frac{\partial A}{\partial p_i} \right) \tag{246}$$

The trajectory originating form a specific point in phase space will be given by the solution of Hamilton's equations of motion, where the initial conditions correspond to the values of the $6N$ variables at the initial point.

===

Example: Motion of a single particle in One-Dimension.

A particle of mass m moving in one dimension in the presence of a potential energy $V(q)$ is described by the Hamiltonian

$$H \;=\; \frac{p^2}{2\,m} \;+\; V(q) \tag{247}$$

The motion of the particle is described Hamilton's equations of motion which simply reduce to the form

$$
\begin{aligned}
\frac{dp}{dt} &= -\frac{\partial V}{\partial q} \\
\frac{dq}{dt} &= \frac{p}{m}
\end{aligned}
\tag{248}
$$

as is expected.

===

The time dependence of any physical quantity $A(\{p_i, q_i\} : t)$ can be evaluated by evaluating it on the trajectory followed in phase space Γ. Hamilton's equation of motion have the consequence that the total derivative of any quantity $A(\{p_i, q_i\} : t)$ can be found from the Poisson Bracket equation of motion

$$
\begin{aligned}
\frac{dA}{dt} &= \sum_{i=1}^{3N} \left(\frac{dq_i}{dt} \frac{\partial A}{\partial q_i} + \frac{dp_i}{dt} \frac{\partial A}{\partial p_i} \right) + \frac{\partial A}{\partial t} \\
&= \sum_{i=1}^{3N} \left(\frac{dH}{dp_i} \frac{\partial A}{\partial q_i} - \frac{dH}{dq_i} \frac{\partial A}{\partial p_i} \right) + \frac{\partial A}{\partial t} \\
&= \{ A , H \}_{PB} + \frac{\partial A}{\partial t}
\end{aligned}
\tag{249}
$$

The first term describes the implicit time dependence of A and the second term describes its explicit time dependence.

If a quantity B has no explicit time dependence and the Poisson Bracket of B and H are zero, then B is conserved.

$$
\begin{aligned}
\frac{dB}{dt} &= \{ B , H \}_{PB} + \frac{\partial B}{\partial t} \\
&= \{ B , H \}_{PB} \\
&= 0
\end{aligned}
\tag{250}
$$

where the first two lines follow from our stated assumptions. Since the total derivative governs the change of B as the system flows through phase space, B is conserved. As an example, since our Hamiltonian does not explicitly depend on time, the Poisson Bracket equation of motion shows that the total derivative of the Hamiltonian w.r.t. time is zero. Explicitly, the equation of motion for H is given by

$$
\begin{aligned}
\frac{dH}{dt} &= \{ H , H \}_{PB} + \frac{\partial H}{\partial t} \\
&= \{ H , H \}_{PB} \\
&= 0
\end{aligned}
\tag{251}
$$

where the second line follows from the absence of any explicit time dependence and the last line follows from the antisymmetric nature of the Poisson Bracket. Hence, the energy is a constant of motion for our closed system. That is, the energy is constant over the trajectory traversed in phase space.
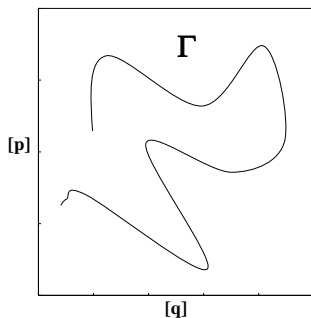
Figure 9: A microscopic state of macroscopic system is described point in phase space and, as it evolves, maps out a very complex trajectory which is governed by Hamilton's equations of motion.

The equations of motion allows us to follow the time evolution of a point in phase space, i.e. the evolution of the microscopic state of the system. The trajectory in phase space may be extremely complicated and rapidly varying. For example in a collision between two neutral molecules, the change in momentum almost exclusively occurs when the separation between the two molecules is of the order of the molecular size. This should be compared with the length scale over which the momentum of each the pair of molecules is constant, which is given by the ratio of the molecular size divided by the distance travelled by a molecule between its successive collisions (i.e. the mean free path). This ratio is usually quite large for dilute gasses. If these distances are scaled by the molecular velocities, one concludes that the momentum of the particles changes rapidly at the collisions and so the trajectory in phase space changes abruptly. The same distance ratio also implies that the particular form of a trajectory is extremely sensitive to the initial conditions, since a small change in initial conditions determines whether or not a particular collision will occur. The sensitivity to initial conditions and the complexity of the trajectories in phase space prohibit both analytic solution and also numerical solution for realistic materials. Numerical solution is prohibited due to the enormity of the requirements for storing the initial conditions, let alone for implementing the numerical solution of the equations of motion. Despite the complexity of trajectories in phase space, and their sensitivity to initial conditions, the trajectories do have some important common features.

The trajectory of a closed system cannot intersect with itself. This is a consequence of Hamilton's equations of motion completely specifying the future

51

motion of a system, if the set of initial conditions are given and since H has no explicit time dependence. Thus, a trajectory cannot cross itself, since if there was then Hamilton's equations would lead to an indeterminacy at the point of intersection. That is, there would be two possible solutions of Hamilton's equations of motion, if the systems initial conditions placed it at the crossing point. This is not possible. However, it is possible that a trajectory closes up on itself and forms a closed orbit.

Secondly, the trajectories only occupy a portion of phase space for which the constants of motion are equal to their initial values.

3.3 Conserved Quantities and Accessible Phase Space

If a system has a set of conserved quantities, then the trajectory followed by the system is restricted to a generalized "surface" in the $6N$ dimensional phase space, on which the conserved quantities take on their initial values. The set of points on the generalized "surface" is known as the accessible phase space Γ_a.

For a classical system where only the energy is conserved and has the initial value E, the points in the accessible phase space is given by the set of points $\{p_i, q_i\}$ that satisfy the equation

$$H(\{p_i, q_i\}) \;=\; E \tag{252}$$

or if the energy is only known to within an uncertainty of ΔE, then the accessible phase space is given by the set of points that satisfy the inequality

$$E \;+\; \Delta E \;>\; H(\{p_i, q_i\}) \;>\; E \tag{253}$$

==

Example: A One-Dimensional Classical Harmonic Oscillator

The Hamiltonian for a particle of mass m constrained to move in one dimension, subject to a harmonic restoring force, is described by the Hamiltonian

$$H \;=\; \frac{p^2}{2\,m} \;+\; \frac{m\,\omega_0^2}{2}\,q^2 \tag{254}$$

The phase space Γ of this system corresponds to the entire two-dimensional plane (p, q). If the energy is known to lie in an interval of width ΔE around E, then the accessible phase space Γ_a is determined by

$$E \;+\; \Delta E \;>\; \frac{p^2}{2\,m} \;+\; \frac{m\,\omega_0^2}{2}\,q^2 \;>\; E \tag{255}$$

The "surfaces" of constant energy[4] are in the form of ellipses in phase space, with semi-major and semi-minor axes given by the turning points

$$p_{max} = \sqrt{2 \ m \ E} \tag{256}$$

and

$$q_{max} = \sqrt{\frac{2 \ E}{m \ \omega_0^2}} \tag{257}$$

The ellipse encloses an area of phase space which is given by

$$\pi \ p_{max} \ q_{max} = 2 \ \pi \ \frac{E}{\omega_0} \tag{258}$$

Therefore, the accessible phase space Γ_a forms an area enclosed between two ellipses, one ellipse with energy $E + \Delta E$ and another with energy E. Thus, the
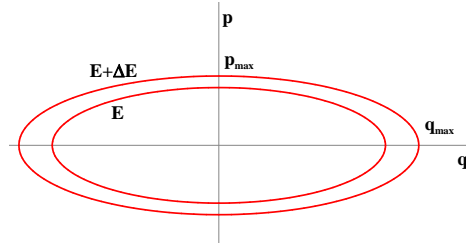
Figure 10: The accessible area of phase space of a one-dimensional Harmonic oscillator is the area enclosed by the two ellipses.

area of accessible phase space is found as

$$\Gamma_a = 2 \ \pi \ \frac{\Delta E}{\omega_0} \tag{259}$$

On diving by $2 \ \pi \ \hbar$ we can turn Γ_a into a measure of the number of microscopic states accessible to the system N_Γ, we find

$$N_\Gamma = \frac{\Delta E}{\hbar \ \omega_0} \tag{260}$$

This is a measure of the number of different states accessible to the system, and can be interpreted quantum mechanically as the number of different quantum states which correspond to the energy within the accuracy ΔE that has been specified. The result N_Γ is just the uncertainty in the number of quanta in the system.

---

[4]In this case the "volume" of phase space is an infinite two-dimensional area and, if the energy is specified precisely, the "area" of accessible phase space is a line.

==

In the most general case where there are several other conserved quantities $B_j(\{p_i, q_i\})$ (say there are M in number) which have specific values B_j, the accessible phase space will consist of the points in phase space where the "surfaces" $B_j = B_j(\{p_i, q_i\})$ corresponding to the conserved quantities intersect. That is, the accessible phase space corresponds to the points which are consistent with the values of all the M conserved quantities B_j

$$
\begin{aligned}
E &= H(\{p_i, q_i\}) \\
B_j &= B_j(\{p_i, q_i\})
\end{aligned}
\tag{261}
$$

for all $j \in 1, 2, \ldots M$. In all cases, the physical trajectories of the system are restricted to move within the accessible region of phase space.

3.4   Macroscopic Measurements and Time Averages

The measurement of thermodynamic quantities usually represents a relatively slow process when compared to microscopic time scales. Furthermore, the measurement involves the participation of many of the systems degrees of freedom. This implies that a macroscopic measurement of a quantity A corresponds to a time-average of a quantity $A(\{p_i, q_i\})$ over a trajectory in phase space, over a long period of time. Furthermore, the quantity $A(\{p_i, q_i\})$ must involve many degrees of freedom of the system. For a long period of time T, the macroscopic quantity is given by

$$
\overline{A} = \frac{1}{T} \int_0^T dt \, A(\{p_i(t), q_i(t)\})
\tag{262}
$$

where $A(\{p_i(t), q_i(t)\})$ varies with time, as the microscopic state changes with time. That is, the set of momenta and coordinates $\{p_i(t), q_i(t)\}$ are considered to be implicit functions of time and are obtained by solving Hamilton's equations using the initial data.

As an example, consider the pressure on a container wall which encloses a dilute gas. The pressure P is defined as the force per unit area. The force F is averaged over a time long compared with the time between molecular collisions with the wall. The force is is given by the rate of change of momentum of the molecules impinging on the wall. The force due to a molecular collision occurs over the time-scale which corresponds to the time in which the molecule is in close proximity to the wall. On introducing a short-ranged interaction between the particle and the wall, one finds that the instantaneous force exerted on the wall is given by

$$
F_3(t) = \left(\frac{dV(q_3)}{dq_3} \right) \bigg|_{q_3(t)}
\tag{263}
$$

where $V(q_3)$ is a short-ranged potential due to the interaction of the particle with the wall. Therefore, the instantaneous pressure is given by

$$P(t) = \frac{1}{A} \sum_{i=1}^{N} \left(\frac{dV(q_{3i})}{dq_{3i}} \right) \Bigg|_{q_{3i}(t)} \tag{264}$$

where A is the area of the wall. The instantaneous pressure would have the appearance of a sparse sequence of delta-like functions. The thermodynamic pressure is given by the time-average over an interval T in which many collisions occur

$$\begin{aligned} P &= \frac{1}{T} \int_0^T dt\ P(t) \\ &= \frac{1}{T\,A} \int_0^T dt \sum_{i=1}^{N} \frac{dV(q_{3i})}{dq_{3i}} \Bigg|_{q_{3i}(t)} \end{aligned} \tag{265}$$

This result is of the form that we are considering. If the time average is over a long enough time interval, the result should be representative of the equilibrium state in which P does not change with time.

The process of time averaging over long intervals is extremely convenient since it circumvents the question of what microscopic initial conditions should be used. For sufficiently long times, the same average would be obtained for many point on the trajectory. Thus, the long time average is roughly equivalent to an average with a statistical distribution of microscopic initial conditions.

3.5 Ensembles and Averages over Phase Space

The time-average of any quantity over the trajectory in phase space can be replaced by an average over phase space, in which the different volumes are weighted with a distribution function $\rho(\{p_i, q_i\} : t)$. The distribution function may dependent on the point of phase space $\{p_i, q_i\}$, and may also depend on the time t.

Conceptually, the averaging over phase space may be envisaged by introducing an ensemble composed of a very large number of identical systems each of which have the same set of values for their measured conserved quantities and all systems must represent the same macroscopic equilibrium state. Although the different systems making up the ensemble correspond to the same macroscopic equilibrium state, the systems may correspond to different microstates. The concept of "Ensemble Averaging" was first introduced by Maxwell in 1879 and developed more fully by Gibbs in 1909.

There are infinitely many possible choices of ensembles, one trivial example is that each system in the ensemble corresponds to the same initial microstate.

Another example corresponds to taking all the different points of a trajectory of one microscopic state as the initial states of the ensemble. A frequently used ensemble corresponds to distributing the probability density equally over all points in phase space compatible with the measured quantities of the macroscopic state.

The Probability Distribution Function

The probability distribution function $\rho(\{p_i, q_i\} : t)$ could, in principle, be measured by measuring the microstates of the systems composing an ensemble at time t and determining the relative number of systems which are found in microstates in the volume $d\Gamma$ of phase space around the point $\{p_i, q_i\}$. In the limit that the number of systems in the ensemble goes to infinity, this ratio reduces to a probability. The probability $dp(t)$ is expected to be proportional to the volume of phase space $d\Gamma$. Therefore, we expect that

$$dp(t) \;=\; \rho(\{p_i, q_i\} : t)\, d\Gamma \qquad (266)$$

where $\rho(\{p_i, q_i\} : t)$ is the probability distribution function. The probability distribution function is only finite for the accessible volume of phase space. Since probabilities are non-negative, then so is the probability distribution function. Furthermore, since the probabilities are defined to be normalized to unity, the probability distribution function must also be normalized

$$
\begin{aligned}
1 \;&=\; \int dp(t) \\
&=\; \int d\Gamma\; \rho(\{p_i, q_i\} : t)
\end{aligned}
\qquad (267)
$$

for all times t. For a macroscopic system, the integration over $d\Gamma$ may be restricted to the volume of available phase space with any loss of generality.

Ensemble Averages

Once $\rho(\{p_i, q_i\} : t)$ has been determined, the measured value of any physical quantity $A(\{p_i, q_i\} : t)$ of a system in a macroscopic state at time t can then be represented by an ensemble average. The ensemble average is the average over phase space weighted by the probability distribution function

$$\overline{A}(t) \;=\; \int d\Gamma\; A(\{p_i, q_i\} : t)\; \rho(\{p_i, q_i\} : t) \qquad (268)$$

If all the different points of a trajectory of one microscopic state is taken to define the initial states of the ensemble, the the ensemble averages will coincide with the long-time average for the microscopic state. At the other extreme, if each system in the ensemble corresponds to the same initial microstate, then

the ensemble average of a quantity at any time t will simply correspond to the value of the quantity for the microstate at time t.

The fundamental problem of statistical mechanics is to find the probability distribution function for the ensemble that describes measurements on the macroscopic equilibrium states of physical systems most closely. We shall examine the equations that determine the time-dependence of the probability distribution function in the next section.

3.6   Liouville's Theorem

Liouville's Theorem concerns how the probability distribution function for finding our N-particle system in some volume element of phase space at time t varies with time.

Since the probability is normalized and since the states of a system evolve on continuous trajectories in phase space, the probability density must satisfy a continuity equation. Consider a volume element $d\Gamma$ of phase space, the number of systems in the ensemble that occupy this volume element is proportional to

$$\rho(\{p_i, q_i\} : t) \; d\Gamma \tag{269}$$

and the increase of the number of systems in this volume element that occurs in the time interval dt is proportional to

$$\left(\rho(\{p_i, q_i\} : t + dt) \; - \; \rho(\{p_i, q_i\} : t) \right) d\Gamma \; \approx \; \frac{\partial}{\partial t} \rho(\{p_i, q_i\} : t) \; d\Gamma \; dt \tag{270}$$

where we have used the Taylor expansion to obtain the right hand side of the equation. Due to the continuous nature of the trajectories, the increase in the number of trajectories in the volume must be due to system trajectories which cross the surface of our $6N$-dimensional volume. That is, the net increase must be due to an excess of the flow across the bounding surfaces into the volume over the flow out of the volume.

Consider the infinitesimal volume of phase space $d\Gamma$ where the i-th coordinate is restricted to be between q_i and $q_i + \Delta q_i$ and the i-th generalized momentum is restricted to be between p_i and $p_i + \Delta p_i$. The volume element $d\Gamma$ is given by

$$d\Gamma \;\; = \;\; \prod_{i=1}^{3N} \left(\Delta q_i \; \Delta p_i \right) \tag{271}$$

The pair of opposite surfaces defined by the coordinates q_i and $q_i + \Delta q_i$ have $6N - 1$ dimensions and have an area given by

$$\prod_{j=1,j\neq i}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \tag{272}$$
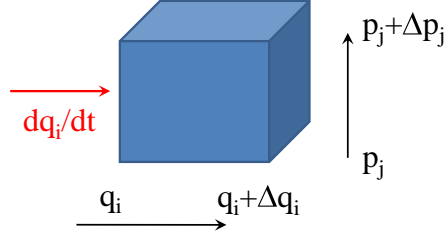
Figure 11: An infinitesimal hyper-cubic element of phase of dimensions $\Delta\Gamma = \prod_{i=1}^{3N} \Delta q_i \Delta p_i$. In time interval dt, the probability density within a distance $\frac{dq_i}{dt}$ perpendicular to the bounding surface at q_i is swept into the volume.

Trajectories which enter or leave the volume element $d\Gamma$ must cross one of its $6N$ boundaries.

Flow In Across a Surface

All the systems of the ensemble in microstates within a distance $\dot{q}_i \, dt$ behind the surface at q_i will enter $d\Gamma$ in time dt. That is, the ensemble systems in the volume $\prod_{j=1,j\neq i}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \, \dot{q}_i(\{p_i, q_i\}) \, dt$ will enter $d\Gamma$ in the time interval dt. The number of systems in this volume is proportional to

$$\prod_{j=1,j\neq i}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \, dt \, \dot{q}_i(\{p_i, q_i\}) \, \rho(\{p_i, q_i\} : t) \tag{273}$$

Flow Out Across a Surface

All the systems in the ensemble with microstates that are within a distance $\dot{q}_i \, dt$ behind the surface at $q_i + \Delta q_i$ will leave $d\Gamma$ in time dt. The number of systems in this volume is proportional to

$$\prod_{j=1,j\neq i}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \, dt \, \dot{q}_i(\{p_i, q_i + \Delta q_i\}) \, \rho(\{p_i, q_i + \Delta q_i\} : t) \tag{274}$$

where the velocity and density must be evaluated at the position of the second surface.

The Net Flow into the Volume

The net flow into $d\Gamma$ from a pair of coordinate surfaces is given by the difference of the flow crossing the coordinate surface entering the volume and

the flow crossing the opposite surface thereby leaving the volume

$$
\prod_{j=1,j\neq i}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \ dt \left(\ \dot{q}_i(\{p_i,q_i\}) \ \rho(\{p_i,q_i\} : t) \ - \ \dot{q}_i(\{p_i,q_i+\Delta q_i\}) \ \rho(\{p_i,q_i+\Delta q_i\} : t) \ \right)
$$

$$
\approx \ - \prod_{j=1}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \ dt \ \frac{\partial}{\partial q_i} \left(\ \dot{q}_i(\{p_i,q_i\}) \ \rho(\{p_i,q_i\} : t) \ \right) \tag{275}
$$

where we have Taylor expanded in powers of Δq_i. Likewise, the net flow into $d\Gamma$ from the pair of momentum surfaces at p_i and $p_i + \Delta p_i$ is given by

$$
\prod_{j=1}^{3N} \Delta q_j \prod_{j=1,j\neq i}^{3N} \Delta p_j \ dt \left(\ \dot{p}_i(\{p_i,q_i\}) \ \rho(\{p_i,q_i\} : t) \ - \ \dot{p}_i(\{p_i+\Delta p_i,q_i\}) \ \rho(\{p_i+\Delta p_i,q_i\} : t) \ \right)
$$

$$
\approx \ - \prod_{j=1}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \ dt \ \frac{\partial}{\partial p_i} \left(\ \dot{p}_i(\{p_i,q_i\}) \ \rho(\{p_i,q_i\} : t) \ \right) \tag{276}
$$

On summing over all the $6N$ surfaces, one finds that the net increase of the number of ensemble systems in the volume $d\Gamma$ that occurs in time dt due to their flowing across all its boundaries is proportional to

$$
- \prod_{j=1}^{3N} \Delta q_j \prod_{j=1}^{3N} \Delta p_j \ dt \left[\frac{\partial}{\partial q_i} \left(\ \dot{q}_i(\{p_i,q_i\}) \ \rho(\{p_i,q_i\} : t) \ \right) + \frac{\partial}{\partial p_i} \left(\ \dot{p}_i(\{p_i,q_i\}) \ \rho(\{p_i,q_i\} : t) \ \right) \right]
$$

$$
\tag{277}
$$

The Continuity Equation

On equating the net increase of the probability in the infinitesimal volume element with the net probability flowing into the volume, one can cancel the factors of dt and $d\Gamma$. Hence, one finds that the probability density satisfies the linear partial differential equation

$$
\frac{\partial \rho}{\partial t} \ + \ \sum_{i=1}^{3N} \left[\ \frac{\partial}{\partial q_i} \left(\ \dot{q}_i \ \rho \ \right) \ + \ \frac{\partial}{\partial p_i} \left(\ \dot{p}_i \ \rho \ \right) \right] \ = \ 0 \tag{278}
$$

On expanding the derivatives of the products one obtains

$$
\frac{\partial \rho}{\partial t} \ + \ \sum_{i=1}^{3N} \left[\ \frac{\partial \dot{q}_i}{\partial q_i} \ \rho \ + \ \dot{q}_i \ \frac{\partial \rho}{\partial q_i} \ + \ \frac{\partial \dot{p}_i}{\partial p_i} \ \rho \ + \ \dot{p}_i \ \frac{\partial \rho}{\partial p_i} \ \right] \ = \ 0 \tag{279}
$$

The above expression simplifies on using Hamilton's equations of motion

$$
\dot{q}_i \ = \ \frac{\partial H}{\partial p_i} \tag{280}
$$

$$
\dot{p}_i \ = \ - \ \frac{\partial H}{\partial q_i} \tag{281}
$$

59

so one obtains

$$\frac{\partial \dot{q}_i}{\partial q_i} = \frac{\partial^2 H}{\partial q_i \partial p_i} \qquad (282)$$

$$\frac{\partial \dot{p}_i}{\partial p_i} = -\frac{\partial^2 H}{\partial p_i \partial q_i} \qquad (283)$$

On substituting these two relations in the equation of motion for ρ, the pair of second-order derivatives cancel and one finally obtains Liouville's equation

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left[\dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i} \right] = 0 \qquad (284)$$

That is, the total derivative of ρ vanishes

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left[\dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i} \right] = 0 \qquad (285)$$

The total derivative is the derivative of ρ evaluated on the trajectory followed by the system. Hence, ρ is constant along the trajectory. Therefore, Liouville's theorem states that ρ flows like an incompressible fluid.
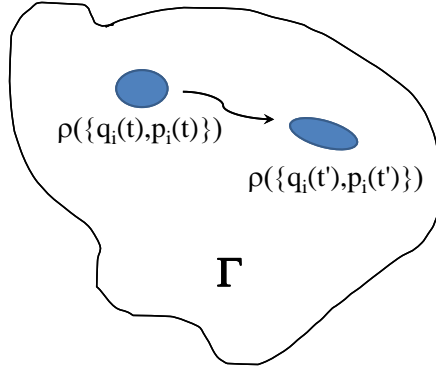
Figure 12: The time evolution of an inhomogeneous probability density $\rho(\{q_i, p_i\})$ satisfies a continuity equation.

On substituting Hamilton's equations for the expressions for $\dot{q}_i$ and $\dot{p}_i$ in Liouville's theorem, one recovers the Poisson Bracket equation of motion for ρ

$$\begin{aligned}
\frac{d\rho}{dt} &= \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left[\frac{\partial H}{\partial p_i} \frac{\partial \rho}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial \rho}{\partial p_i} \right] \\
&= \frac{\partial \rho}{\partial t} + \left\{ \rho , H \right\}_{PB} = 0
\end{aligned} \qquad (286)$$

which is in a form suitable for Canonical Quantization, in which case ρ and H should be replaced by operators and the Poisson Bracket by a commutator times an imaginary number.

Liouville's theorem is automatically satisfied for any ρ which has no explicit t-dependence and can be expressed in terms of the constants of motion. Specifically, when ρ is initially uniform over the accessible phase space, Liouville's theorem ensures that it will remain constant. To be sure, if the distribution satisfies

$$\frac{\partial \rho}{\partial p_i} \;=\; 0 \quad \forall\, i \tag{287}$$

and

$$\frac{\partial \rho}{\partial q_i} \;=\; 0 \quad \forall\, i \tag{288}$$

for all points $\{p_i, q_i\}$ within the accessible volume of phase space (defined by $H(\{p_i, q_i\}) = E$ and any other relevant conservation laws) then Liouville's theorem yields that

$$\frac{\partial \rho}{\partial t} \;=\; 0 \tag{289}$$

==

Example: A Particle in a One-Dimensional Box.

We shall consider an example that illustrates how a probability density thins and folds as time evolves. The example also shows that for sufficiently large times, the probability distribution is finely divided and distributed over the volume of accessible phase space.

We shall consider an ensemble of systems. Each system is composed of a single particle that is confined in a one-dimensional box of length L. When the particle is not in contact with the walls, the Hamiltonian reduces to

$$H(p, q) \;=\; \frac{p^2}{2\,m} \tag{290}$$

The energies of the systems in the ensemble are bounded by

$$E \;+\; \Delta E \;>\; H(p, q) \;>\; E \tag{291}$$

which restricts the momenta to the two intervals

$$p_{max} \;>\; p \;>\; p_{min} \quad \text{and} \quad -\,p_{min} \;>\; p \;>\; -\,p_{max} \tag{292}$$

The coordinates are restricted to the interval

$$\frac{L}{2} \;>\; q \;>\; -\,\frac{L}{2} \tag{293}$$

Thus, the volume of accessible phase space consists of two two-dimensional strips.

The probability distribution $\rho(p, q : t)$ evolves according to Liouville's theorem

$$\frac{\partial \rho}{\partial t} + \frac{\partial H}{\partial p} \frac{\partial \rho}{\partial q} - \frac{\partial H}{\partial q} \frac{\partial \rho}{\partial p} = 0 \tag{294}$$

which for volumes contained within the spatial boundaries reduces to

$$\frac{\partial \rho}{\partial t} + \frac{p}{m} \frac{\partial \rho}{\partial q} = 0 \tag{295}$$

This equation has the general solution

$$\rho(p, q : t) = A\left(q - \frac{p\, t}{m} \right) B(p) \tag{296}$$

which is valid everywhere except at the locations of the walls. In the general solution A and B are arbitrary functions which must be fixed by the boundary conditions.

We shall adopt the initial condition that the probability distribution function has the form

$$\rho(p, q : 0) = \delta(q)\, B(p) \tag{297}$$

which initially confines all the particles in the ensemble to the center $q = 0$. The momentum distribution function $B(p)$ is evenly distributed over the allowed range

$$B(p) = \frac{1}{2\,(\,p_{max}\, -\, p_{min}\,)} \left[\Theta(p-p_{max}) - \Theta(p-p_{min}) + \Theta(p+p_{max}) - \Theta(p+p_{min}) \right] \tag{298}$$

For sufficiently short times, short enough so that the particles in the ensemble have not yet made contact with the walls, the solution is of the form

$$\rho(p, q : t) = \delta\left(q - \frac{p}{m} t \right) B(p) \tag{299}$$

which has the form of two segments of a line. The slope of the line in phase space is given by m/t. For small times the segments are almost vertical, but the slope increases as t increases. The increase in slope is caused by the dispersion of the velocities, and causes the length of the line to increase. The increase in the length of the line does not affect the normalization, which is solely determined by $B(p)$. At a time T_1 some particles in the ensemble will first strike the walls, that is the line segments in available phase space will extend to $q = \pm \frac{L}{2}$. This first happens when

$$T_1 = \frac{L\, m}{2\, p_{max}} \tag{300}$$

62
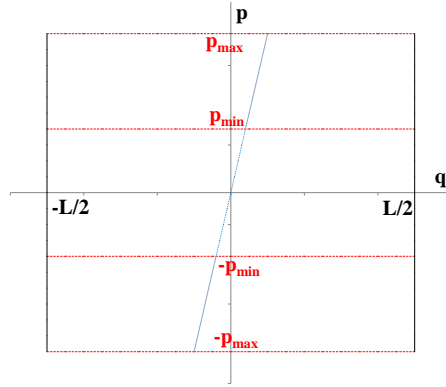
Figure 13: The regions where the probability density for an ensemble of systems composed of a particle in a box is non-zero, at short times, is shown by the solid portion of the blue line. The slope of the line is caused by the dispersion in the velocities. The accessible phase space is enclosed by the red dashed lines between p_{max} and p_{min}, and a similar region in the lower half space.

For times greater than T_1 some of the ensemble's particles will be reflected from the walls. The solution of Liouville's equation can be found by the method of im-
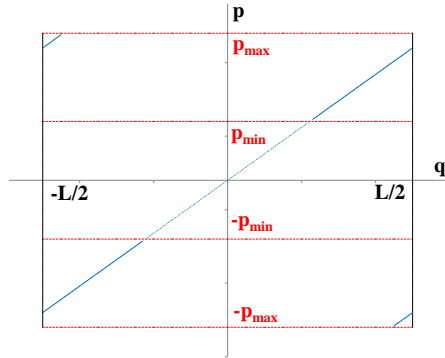
Figure 14: The regions where the probability density for an ensemble of particles in boxes is non-zero, for times slightly greater than the times of the first collision, is shown by the solid portion of the blue line. The two small line segments in the upper left-hand and lower right-hand portion of accessible phase space represents the region of the probability density for systems where the particle has been reflected.

ages. That is, the reflected portion of the probability density can be thought of as originating from identical systems with identical initial conditions except that they are obtained by spatially reflecting our system at its boundaries $q = \pm \frac{L}{2}$. The reflection requires that $B(p) \to B(-p)$ in the image. The probability dis-

tribution emanating from these other systems will enter the volume of available our available phase space at time T_1 which will represent the reflected portion of the probability distribution function. The probability distribution that leaves
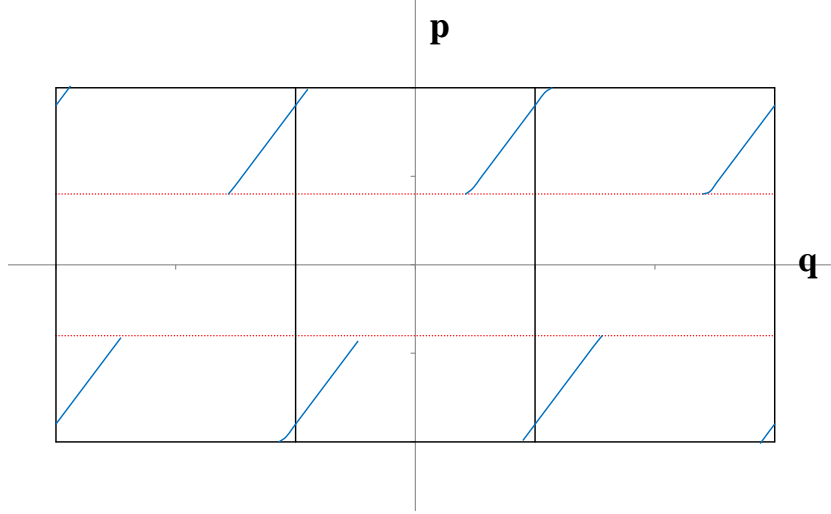
Figure 15: The extended phase space produced by reflecting the central area across its boundaries. In this extended system, the reflected probability density is simply represented by the free evolution of the initial probabilities of the image systems.

our system, represents the reflected portion of the probability distribution for the neighboring systems. Thus, we are mentally extending the region of accessible phase space in the spatial direction. The solution just after the first reflection has occurred, but for times before any system has experienced two reflections is given by

$$\rho(p, q : t) \;=\; \sum_{n=-1}^{n=1} \delta\left(q \;-\; n\,L \;-\; \frac{p}{m}\,t \right) B(p) \tag{301}$$

where q is restricted to the interval $\frac{L}{2} \;>\; q \;>\; -\frac{L}{2}$. The folding of the distribution does not affect its normalization.

For larger times, for which any system in the ensemble has undergone multiple reflections, the set of systems must be periodically continued along the spatial axis. That is, we must consider multiple images of our system. The probability distribution valid at any time obviously has the form

$$\rho(p, q : t) \;=\; \sum_{n=-\infty}^{\infty} \delta\left(q \;-\; n\,L \;-\; \frac{p}{m}\,t \right) B(p) \tag{302}$$

64
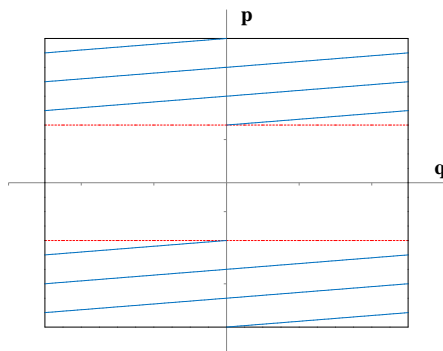
Figure 16: The regions where the probability density for a particle in a box is non-zero, for large times, is shown by the solid blue lines. For large times, particles in the ensemble have experienced different numbers of collisions and is spread over many line segments.

where q is still restricted to the interval of length L. The probability distribution is non-zero on a set of parallel line segments with slope m/t. The line segments are separated by a "distance" $(mL)/t$ along the momentum direction. For sufficiently large times, the slope of the lines will be small and they will be closely spaced. In conclusion, for sufficiently large times, we have shown that the probability distribution will be finely divided and spread throughout the volume of accessible phase space.

==

The Time Dependence of Averages

Liouville's theorem shows that the time dependence of any quantity $A(\{p_1, q_i\})$ (with no explicit t dependence) also follows from the Poisson Bracket equations. This can be seen by first multiplying Liouville's equation by $A(\{p_1, q_i\})$ and then integrating over phase space.

$$0 \; = \; \int d\Gamma \; A(\{p_1, q_i\}) \; \frac{\partial \rho}{\partial t} \; + \; \int d\Gamma \; A(\{p_1, q_i\}) \left\{ \rho \, , \, H \right\}_{PB} \qquad (303)$$

The derivatives of ρ w.r.t the variables $\{p_i, q_i\}$ that occur in the Poisson Bracket term can be removed by integrating by parts. That is, on noting that ρ vanishes on the boundaries of the integration, then integration by parts yields

$$0 \; = \; \int d\Gamma \; A \; \frac{\partial \rho}{\partial t} \; - \; \sum_{i=1}^{3N} \int d\Gamma \; \rho \left[\frac{\partial}{\partial q_i} \left(A \; \frac{\partial H}{\partial p_i} \right) \; - \; \frac{\partial}{\partial p_i} \left(A \; \frac{\partial H}{\partial q_i} \right) \right] \quad (304)$$

65

The derivatives of the products can be expanded to yield

$$\int d\Gamma \, A \, \frac{\partial \rho}{\partial t} \; = \; \sum_{i=1}^{3N} \int d\Gamma \, \rho \left[\frac{\partial A}{\partial q_i} \frac{\partial H}{\partial p_i} + A \frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial H}{\partial q_i} - A \frac{\partial^2 H}{\partial p_i \partial q_i} \right]$$
(305)

The terms proportional to the second derivative of the Hamiltonian cancel, leading to

$$\begin{aligned}
\frac{d\overline{A}}{dt} &= \int d\Gamma \, A \, \frac{\partial \rho}{\partial t} \\
&= \sum_{i=1}^{3N} \int d\Gamma \, \rho \left[\frac{\partial A}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial H}{\partial q_i} \right] \\
&= \int d\Gamma \, \rho \left\{ A \, , \, H \right\}_{PB}
\end{aligned}$$
(306)

which equates the time-derivative of the average of A with the average of the Poisson Brackets.

The above equation has the consequence that for a macroscopic equilibrium state the condition that the average of any quantity A that has no explicit t-dependence should be independent of time

$$\frac{d\overline{A}}{dt} \; = \; \int d\Gamma \, \frac{\partial \rho}{\partial t} \, A \; = \; 0$$
(307)

where the entire volume of the integration is fixed. (Note that in this expression, the total derivative has a different meaning from before since the integration volume element is considered as being held fixed.) The requirement of the time-independence of any quantity A in equilibrium necessitates that the Poisson Bracket of ρ and H must vanish. This can be achieved if ρ only depends on H and any other conserved quantities.

3.7   The Ergodic Hypothesis

In proving Liouville's theorem, we noted that

$$\frac{\partial \dot{q}_i}{\partial q_i} \; = \; \frac{\partial^2 H}{\partial q_i \partial p_i}$$
(308)

$$\frac{\partial \dot{p}_i}{\partial p_i} \; = \; - \frac{\partial^2 H}{\partial p_i \partial q_i}$$
(309)

This has the consequence that if one follows the flow of the systems of the ensemble with microstates contained in a specific volume of phase space $d\Gamma$ at time t then at time t' the set of microstates will have evolved to occupy a volume of phase space $d\Gamma'$ such that

$$d\Gamma \; = \; d\Gamma'$$
(310)

This can be seen by considering the product of the canonically conjugate pairs of infinitesimal momenta and coordinates at time t

$$dp_i \ dq_i \tag{311}$$

At time $t + dt$ the time evolution will have mapped the ends of these intervals onto new intervals such that the lengths of the new intervals are given by

$$dp_i' \ = \ dp_i \left(1 \ + \ \frac{\partial \dot{p}_i}{\partial p_i} \ dt \right) \tag{312}$$

and

$$dq_i' \ = \ dq_i \left(1 \ + \ \frac{\partial \dot{q}_i}{\partial q_i} \ dt \right) \tag{313}$$

Therefore, the product of the new intervals is given by

$$dp_i \ dq_i \left[1 \ + \ \left(\frac{\partial \dot{q}_i}{\partial q_i} \ + \ \frac{\partial \dot{p}_i}{\partial p_i} \right) dt \ + \ O(dt^2) \right] \tag{314}$$

which since

$$\frac{\partial \dot{q}_i}{\partial q_i} \ + \ \frac{\partial \dot{p}_i}{\partial p_i} \ = 0 \tag{315}$$

leaves the product invariant, to first-order in dt. Hence, since

$$d\Gamma' \ = \ \prod_{i=1}^{3N} dp_i' \ dq_i' \tag{316}$$

the size of the volume element occupied by the microstates is invariant, i.e. $d\Gamma = d\Gamma'$. This does not imply that the shape of the volume elements remains unchanged, in fact they will become progressively distorted as time evolves. For most systems for which the trajectories are very sensitive to the initial conditions, the volume elements will be stretched and folded, resulting in the volume being finely divided and distributed over the accessible phase space.

The initial formulation of the Ergodic Hypothesis was introduced by Boltzmann[5] in 1871. A modified form of the hypothesis asserts that if the volume of accessible phase space is finite, then given a sufficiently long time interval, the trajectories of the microstates initially contained in a volume element $d\Gamma$ will come arbitrarily close to every point in accessible phase space. If this hypothesis is true, then a long-time average of an ensemble containing states initially in $d\Gamma$ will be practically equivalent to an average over the entire volume of accessible phase space with a suitable probability density. That is, the Ergodic Hypothesis leads one to expect that the equation

$$\overline{A} \ = \ \frac{1}{T} \int_0^T \ dt \ A(\{p_i(t), q_i(t)\}) \ = \ \int d\Gamma \ A(\{p_i, q_i\}) \ \rho(\{p_i, q_i\}) \tag{317}$$

---

[5]L. Boltzmann, "Einige allgemeninen Sätze über das Wärmegleichgewicht", Wien Ber. 63, 670-711 (1871).

holds for some $\rho(\{p_i, q_i\})$ (the Ergodic Distribution) at sufficiently large times T.

The Ergodic Theorem.

The Ergodic Theorem (due to J. von Neumann[6], and then improved on by Birkhoff[7] in the 1930's) states that the time-average of a quantity A along a trajectory that is initially located at any point in phase space, then in the limit as the time goes to infinity one has:

(i) the time-average converges to a limit.

(ii) that limit is equal to the weighted average of the quantity over accessible phase space. That is, the trajectory emanating from any initial point resembles the whole of the accessible phase space.

The Ergodic Theorem has been proved for collisions of hard spheres and for motion on the geodesics on surfaces with constant negative curvature. Ergodicity can also be demonstrated for systems through computer simulations. The Ergodic Theorem has similar implications as a weaker theorem which is known as the Poincaré Recurrence Theorem.

Poincaré's Recurrence Theorem

The Poincaré Recurrence Theorem[8] states that most systems will, after a sufficiently long time, return to a state very close to their initial states. The Poincaré Recurrence Time T_R is the time interval that has elapsed between the initial time and the time when the systems recur. The theorem was first proved by Henri Poincaré in 1890.

The proof is based on the two facts:

(i) The phase trajectories of a closed system do not intersect.

(ii) The infinitesimal volume of a phase space is conserved under time evolution.

Consider an arbitrarily small neighbourhood around any initial point in accessible phase space and follow the volume's trajectory as the microstates evolve with time. The volume "sweeps out" a tube in phase space as it moves. The tube can never cross the regions that have been already "swept out", since trajectories in phase space do not intersect. Hence, as the accessible phase space is a compact manifold, the total volume available for future motion without recurrence will decrease as the time increases. If the tube has not already returned to

---

[6]J. von Neumann, "Physical Applications of the Ergodic Hypothesis", Proc. Natl. Acad. Sci. 18, 263-266,(1932).

[7]G.D. Birkhoff, "Proof of the ergodic theorem", Proc. Natl. Acad. Sci. 17, 656-660 (1930).

[8]H. Poincaré, "Sur les courbes définies par une équation différentielle", Oeuvres, 1, Paris, (1892).
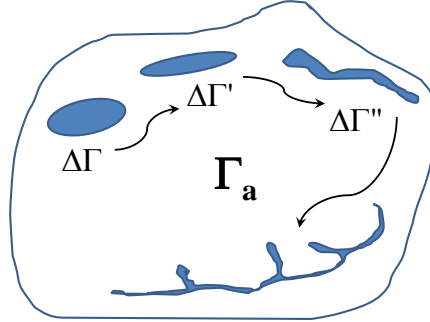
Figure 17: A schematic description of Poincaré's Recurrence Theorem. Under time-evolution, a region $\Delta\Gamma$ of phase space sweeps out a trajectory in phase space. At each instant of time, the region occupies an equal volume of the accessible phase space Γ_a, so that $\Delta\Gamma = \Delta\Gamma' = \Delta\Gamma$". After a time T_R, the trajectory of the region will come arbitrarily close to the initial region.

the initial neighborhood (in which case recurrence has already occurred), then since the total volume of accessible phase space is finite, in a finite time T_R all the volume of accessible phase space must be exhausted. At that time, the only possibility is that the phase tube returns to the neighbourhood of the initial point.

Quod Erat Demonstrandum (QED).

Thus, the trajectory comes arbitrarily close to itself at a later time T_R. If the trajectories don't repeat and form closed orbits they must densely fill out all the available phase space. However, if Poincaré recurrence occurs before the entire volume of available phase space is swept out, some of it may remain unvisited. Liouville's theorem implies that the density of trajectories is uniform in the volume of accessible phase space that is visited. The Recurrence Time T_R is expected to be extremely larger than the time scale for any macroscopic measurement, in which case the Recurrence Theorem cannot be used to justify replacing time averages with ensemble averages.

A simple example of the Ergodic Theorem is given by the One-dimensional Harmonic Oscillator. The Double-Well Potential exhibits a region where ergodicity applies, but for low energies the motion may become constrained to one well in which case ergodicity does not apply.

If the Ergodic Theorem holds then, for sufficiently large times T_R, the time-average of any quantity A

$$\overline{A} \;=\; \frac{1}{T_R} \int_0^{T_R} dt \; A(\{p_i(t), q_i(t)\}) \tag{318}$$

represents the measured value of a macroscopic quantity and the trajectory passes arbitrarily close to every point in phase space. If the system's trajectory dwells in the volume of phase space $\Delta\Gamma$ for time Δt, then the ratio

$$\frac{\Delta t}{T_R} \tag{319}$$

has a definite limit which defines the probability that, if the system is observed at some instant of time, it will be found to have a microscopic state in $\Delta\Gamma$.

There are a number of systems which are known not to obey the Ergodic Hypothesis. These include integrable systems, or nearly integrable systems. An integrable system has a number of conservation laws B_i equal to half the number of dimensions of phase space. Furthermore, each pair of conserved quantities must be in involution

$$\left\{ B_i , B_j \right\}_{PB} = 0 \tag{320}$$

These sets of conservation laws reduce the trajectory of an integrable system to motion on a $3N$-dimensional surface embedded in the $6N$-dimensional phase space. Furthermore, due to the involution condition, the normals to the surface of constant B_i lay in the surface and define a coordinate grid on the surface which does not have a singularity. The singularities of coordinate grids define the topology of the surface. For example, a coordinate grid on the surface of a sphere has two singularities (one at each pole), whereas a coordinate grid on a doughnut does not. The trajectories of an integrable system are confined to surfaces that have the topology of $3N$-dimensional tori. Different initial conditions will lead to different tori that are nested within phase space. The motion on the torus can be separated into $3N$ different types of periodic modes. Since these modes have different frequencies the motion is quasi-periodic. It is the separability of the coordinates that makes integrability a very special property. The very existence of $3N$ conserved quantities implies that most of the conserved quantities are microscopic and their values are not directly measurable by macroscopic means. Hence, they should not be considered as restricting the available phase space for the macroscopic state. Thus, it should be no surprise that integrable systems are generally considered to be non-ergodic.

===

Example: Circular Billiards

A simple example of an integrable system is given by the circular billiard. In this case, a particle is free to move within a circular area of the plane. The particle is confined to the area of radius R since it is specularly reflected by the perimeter. The spatial paths followed by the particle consists of a succession of cords. Whenever a cord meets the perimeter, the angle α between the cord and the perimeter's tangent is the same for each reflection. The phase space is
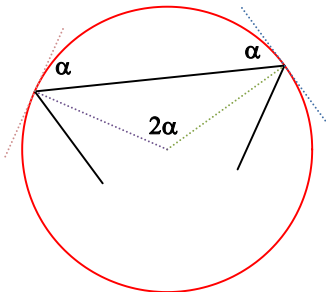
Figure 18: The basic geometry of the scattering for circular billiards.

four-dimensional. However, there are two constant of motion, the energy E and the angular momentum p_φ. The angle φ satisfies the equation of motion

$$\dot{\varphi} \;=\; \frac{p_\varphi}{m \; r^2} \tag{321}$$

The radial motion is described by the pair of equations

$$\dot{r} \;=\; \frac{p_r}{m}$$

$$\dot{p}_r \;=\; \frac{p_\varphi^2}{m \; r^3} \;-\; V_0 \; \delta(r - R) \tag{322}$$

where $V_0 \to \infty$. The second equation can be integrated once w.r.t. t by using an integrating factor of p_r on the left and $m \, \dot{r}$ on the right, which introduces a constant of motion which is the (conserved) energy. The remaining integration w.r.t. t leads to the solution for $r(t)$. From this one sees that the radial coordinate $r(t)$ performs anharmonic oscillations between the turning point at a

$$a \;=\; \frac{p_\varphi}{\sqrt{2 \; E}} \tag{323}$$

where $0 \; < \; a \; < \; R$, and the radius R. The ratio of the period for a $2 \, \pi$ rotation of φ to the period of the radial motion may not be a rational number. Hence, in general the motion is quasi-periodic. Also since the turning point a is a constant, the paths in the space of Cartesian coordinates are obviously excluded from a circular region of radius a centered on the origin.

===

When a system does not have $3N$ conserved quantities, the system is non-integrable. The trajectories have much fewer restrictions and extend to higher dimensions in phase space. The trajectories are more sensitive to the initial conditions, so the trajectories are inevitably chaotic.

71
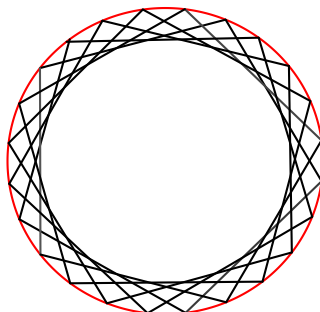
Figure 19: A spatial path traced out by the billiard ball over a long time interval.

The Kolmogorov-Arnold-Moser (KAM) theorem indicates that there is a specific criterion which separates ergodic from non-ergodic behaviour.

The trajectories of an integrable system are confined to a doughnut-shaped surface in phase space, an invariant torus. If the integrable system is subjected to different initial conditions, its trajectories in phase space will trace out different invariant tori. Inspection of the coordinates of an integrable system shows that the motion is quasi-periodic. The KAM theorem specifies the maximum magnitude of a small non-linear perturbation acting an a system (which when non-perturbed is integrable) for which the quasi-periodic character of the orbits is still retained. For larger magnitudes of the perturbation, some invariant tori are destroyed and the orbits become chaotic so ergodicity can be expected to hold. The KAM Theorem was first outlined by Andrey Kolmogorov[9] in 1954. It was rigorously proved and extended by Vladimir Arnol'd[10] (1963) and by Jürgen Moser[11] (1962).

3.8 Equal a priori Probabilities

The Hypothesis of Equal a priori Probabilities was an assumption assigns equal probabilities to equal volumes of phase space. This hypothesis, first introduced by Boltzmann, assumes that the probability density for ensemble averaging over phase space is uniform. The hypothesis is based on the assumption that dynamics does not preferentially bias some volume elements of available phase space

---

[9]A.N. Kolmogorov, "On Conservation of Conditionally Periodic Motions for a Small Change in Hamilton's Function." Dokl. Akad. Nauk SSSR 98, 527-530, (1954).

[10]V.I. Arnol'd, "Proof of a Theorem of A. N. Kolmogorov on the Preservation of Conditionally Periodic Motions under a Small Perturbation of the Hamiltonian." Uspehi Mat. Nauk 18, 13-40, (1963).

[11]J. Moser, "On Invariant Curves of Area-Preserving Mappings of an Annulus." Nachr. Akad. Wiss. Gōttingen Math.-Phys. Kl. II, 1-20, (1962).

$d\Gamma$ of other elements with equal volumes. This hypothesis is consistent with Liouville's theorem which ensures that an initially uniform probability distribution will remain uniform at all later times. The Equal a priori Hypothesis is equivalent to assuming that the in many consecutive rolling a dice, that each of the six faces of a dice will have an equal probability of appearing.

If the Ergodic Hypothesis holds then, for sufficiently large times T_R, the time-average of any quantity A

$$\overline{A} \;=\; \frac{1}{T_R} \int_0^{T_R} dt\; A(\{p_i(t), q_i(t)\}) \tag{324}$$

represents the measured value of a macroscopic quantity and the trajectory passes arbitrarily close to every point in phase space. If the system's trajectory dwells in the volume of phase space $\Delta\Gamma$ for time Δt, then the ratio

$$\frac{\Delta t}{T_R} \tag{325}$$

has a definite limit which defines the probability that, if the system is observed at some instant of time, it will be found to have a microscopic state in $\Delta\Gamma$. However, the Hypothesis of Equal a priori Probabilities assigns the probability density ρ for a system to be found in the volume $\Delta\Gamma$ of accessible phase space to a constant value given by the normalization condition

$$\rho \;=\; \frac{1}{\Gamma_a} \tag{326}$$

where Γ_a is the entire volume of accessible phase space. The requirement of the equality of the time-average and ensemble average requires that the two probabilities must be equal

$$\frac{\Delta t}{T_R} \;=\; \frac{\Delta\Gamma}{\Gamma_a} \tag{327}$$

Hence, the Ergodic Hypothesis when combined with the Hypothesis of Equal a priori probabilities requires that the trajectory must spend equal times in equal volumes of phase space.

==

Example: The One-Dimensional Harmonic Oscillator.

We shall show that for the one-dimensional harmonic oscillator that the time Δt spent in some volume $\Delta\Gamma \;=\; \Delta p\,\Delta q$ of its two-dimensional phase space is proportional to the volume. That is, the trajectory spends equal time in equal volumes.

The Hamiltonian is expressed as as

$$H(p,q) \;=\; \frac{p^2}{2\,M} + \frac{M\,\omega_0^2}{2}\,q^2 \tag{328}$$

73

The equations of motion have the form

$$\frac{dp}{dt} = -M \omega_0^2 q$$
$$\frac{dq}{dt} = \frac{p}{M} \tag{329}$$

The equations of motion for the one-dimensional Harmonic Oscillator can be integrated to yield

$$p(t) = M \omega_0 A \cos(\omega_0 t + \phi)$$
$$q(t) = A \sin(\omega_0 t + \phi) \tag{330}$$

where the amplitude A and initial phase ϕ are constants of integration. The Hamiltonian is a constant of motion, and the accessible phase space is given by

$$E + \Delta E > H(p, q) > E \tag{331}$$

which leads to the constraint on the amplitude

$$E + \Delta E > \frac{M \omega_0^2}{2} A^2 > \Delta E \tag{332}$$

From the solution one finds that the orbits are closed and form ellipses in phase
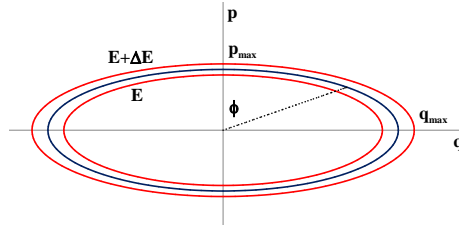
Figure 20: A typical trajectory for the one-dimensional Classical Harmonic Oscillator is shown in blue. The initial phase ϕ is assumed to be unknown. The energy is known to within ΔE so the accessible phase space Γ_a is the area enclosed between the two ellipses.

space, which pass arbitrary close to every point in accessible phase space. The Poincaré recurrence time T_R is given by

$$T_R = \frac{2 \pi}{\omega_0} \tag{333}$$

Consider an element of phase space $\Delta\Gamma = \Delta p \, \Delta q$ where $\Delta q \ll \Delta p$. The trajectory will spend a time Δt in this volume element where

$$\Delta t = \frac{\Delta q}{|\dot{q}|}$$
$$= \frac{\Delta q \, M}{|p|} \tag{334}$$

74

Now, the extent of the volume of phase space at (p, q) is determined from the energy spread

$$\Delta E \;=\; \frac{p \, \Delta p}{M} \;+\; M \, \omega_0^2 \, q \, \Delta q \tag{335}$$

Since we have assumed that $\Delta q \;\ll\; \Delta p$, the spread in energy is related to Δp via

$$\Delta E \;=\; \frac{\mid p \mid \Delta p}{M} \tag{336}$$

On substituting for M/p into the expression for Δt we obtain

$$\Delta t \;=\; \frac{\Delta q \, \Delta p}{\Delta E} \tag{337}$$

However, as we have already shown, ΔE is related to the volume of accessible phase space Γ_a via

$$\Gamma_a \;=\; 2 \, \pi \, \frac{\Delta E}{\omega_0} \tag{338}$$

Therefore,

$$\begin{aligned}
\Delta t \;&=\; \frac{\Delta q \, \Delta p}{\Gamma_a} \left(\frac{2 \, \pi}{\omega_0} \right) \\
&=\; \frac{\Delta \Gamma}{\Gamma_a} \, T_R
\end{aligned} \tag{339}$$

Hence, we have shown that
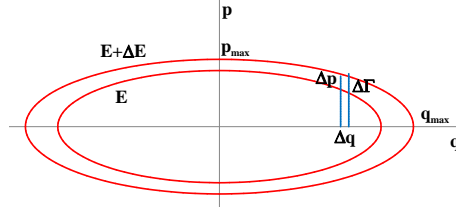
Figure 21: The trajectory crosses an element $\Delta \Gamma$ of accessible phase space with a narrow width Δq in time Δt, the height Δp of the element is determined form the uncertainty in the energy ΔE.

$$\frac{\Delta t}{T_R} \;=\; \frac{\Delta \Gamma}{\Gamma_a} \tag{340}$$

which shows that the trajectory spends equal times in equal volumes of phase space.

This relation is independent of the assumed shape of the volume element since if we considered a volume for which $\Delta q \;\gg\; \Delta p$ then Δt is given by

$$\begin{aligned}
\Delta t \;&=\; \frac{\Delta p}{\mid \dot{p} \mid} \\
&=\; \frac{\Delta p}{M \, \omega_0^2 \, \mid q \mid}
\end{aligned} \tag{341}$$

75

However, in this case the extent of the volume of accessible phase space at the point (p, q) is determined from the energy spread

$$\Delta E \;=\; M \; \omega_0^2 \mid q \mid \Delta q \tag{342}$$

Therefore, one has

$$\Delta t \;=\; \frac{\Delta q \; \Delta p}{\Delta E} \tag{343}$$

which, on relating ΔE to the volume of accessible phase space Γ_a, leads to the same relation

$$\frac{\Delta t}{T_R} \;=\; \frac{\Delta \Gamma}{\Gamma_a} \tag{344}$$

This shows the result probably does not depend on the specific shape of the volume of accessible phase space $\Delta \; \Gamma$.

This example also illustrates how the average of a property of a system with unknown initial conditions phases (in this case the initial phase ϕ) can be thought of either as a time average or as an ensemble average.

===

The hypothesis of equal a priori probabilities does provide a reasonable basis for calculating the equilibrium thermodynamic properties of a large number of physical systems. This anecdotal evidence provides justification for its use. However, one is lead to suspect that ρ is not really uniform but instead is finely dispersed throughout the volume of accessible phase space. In our discussion of the Micro-Canonical Ensemble and everything that follows from it. we shall be assuming that the Hypothesis of Equal a priori Probabilities is valid.

3.9   The Physical Significance of Entropy

A system can only make a transition from one macroscopic equilibrium state to another if the external conditions are changed. A change in external conditions, without supplying energy to the system, can be achieved by removing a constraint on the system. The removal of a constraint usually results in an increase in N_Γ the number of microscopic states available to the system. It is convenient to introduce a measure of the number of N_Γ which is extensive, or additive. Since N_Γ is multiplicative, $\ln N_\Gamma$ is additive and represents a measure of the number of microscopic states corresponding to the macroscopic equilibrium state. The removal of a constraint has the effect that $\ln N_\Gamma$ increases, as does the thermodynamic entropy. Therefore, this argument suggest that the entropy may be defined by

$$S \;=\; k_B \; \ln N_\Gamma \tag{345}$$

in which case, the entropy is a measure of the dispersivity of the distribution of microscopic states. The factor of k_B (Boltzmann's constant) is required to give

Table 1: Percentage frequency of occurrence of the letters in English language texts.

a	b	c	d	e
8.17	1.49	2.78	4.25	12.70

f	g	h	i	j
2.23	2.01	6.09	6.97	0.15

k	l	m	n	o
0.77	4.02	2.41	6.75	7.51

p	q	r	s	t
1.93	0.09	5.99	6.33	9.06

u	v	w	x	y
2.76	0.98	2.36	0.15	1.97

z	-	-	-	-
0.07	-	-	-	-

the entropy the same dimensions as the thermodynamic entropy.

Information Theory

Shannnon[12] has rigorously proved that the information content S of a probability distribution function of a random process with M possible outcomes is given by

$$S \;=\; - \sum_{i=1}^{M} p_i \; \ln p_i \tag{346}$$

where p_i is the probability of the i-th outcome.

Consider sending a message consisting of an ordered string of N values of the possible outcomes. The outcomes can be considered similar to the letters of an alphabet in which case the message is a word containing N letters. Like languages, the letters don't occur with equal frequency, for example in English language texts the letter e appears most frequently, and those who play the game "scrabble" know that the letters q and z occur very infrequently.

The total possible number of messages of length N is just M^N. However, not all messages occur with equal probability, since if the outcome i occurs with a small probability p_i messages in which the outcome i occurs a significant number of times have very small probabilities of appearing. In the analogy with words of N letters, some allowed words occur so infrequently that they are never listed in a dictionary.

---

[12]C.E. Shannon, "A Mathematical Theory of Communication." Bell System Tech. J., 27, 379- 423, 623656, (1948).

A typical message of length N could be expected to contain the outcome i an average of $N_i = Np_i$ times. Hence, one can determine the approximate number of times N_i each outcome i will occur in a typical message. Since the N_i are fixed, the set of typical messages merely differ in the order that these outcomes are listed. The number of these typical messages D_N can be found from the number of different ways of ordering the outcomes

$$ D_N \;=\; \frac{N!}{\prod_{i=1}^{M} N_i!} \tag{347} $$

Hence, the dictionary of typical N-character messages (N-letter words) contains D_N entries. We could index each message in the dictionary by a number. Suppose we wish to transmit a message, instead of transmitting the string of characters of the message we could transmit the index which specifies the place it has in the dictionary. If we were to transmit this index using a binary code, then allowing for all possible messages, one would have to transmit a string binary digits of length given by

$$ \log_2 D_N \;\approx\; \log_2 \left(\frac{N!}{\prod_{i=1}^{M} N_1!} \right) $$
$$ \approx\; -N \sum_{i=1}^{M} p_i \, \log_2 p_i \tag{348} $$

where the last line has been obtained by using Stirling's formula (valid for large Np_i). For an uniform probability distribution, this number would be just $N \, \log_2 M$. The difference in these numbers, divided by N is the information content of the probability distribution function. Shannon's Theorem proves this rigorously.

The Entropy

We shall describe the entropy of a macroscopic state as a phase space average

$$ S \;=\; -k_B \int d\Gamma \, \rho \, \ln(\rho \, \Gamma_0) \tag{349} $$

where the factor of Γ_0 has been introduced to make the argument of the logarithm dimensionless. It is convenient to express Γ_0 for a system of N indistinguishable particles moving in a three-dimensional space as

$$ \Gamma_0 \;=\; N! \, (\, 2 \, \pi \, \hbar \,)^N \tag{350} $$

since the introduction of this factor and the use of the equal a priori hypothesis results in the expression

$$ S \;=\; k_B \, \ln N_\Gamma \tag{351} $$

if one identifies the number of accessible microstates as

$$ N_\Gamma \;=\; \frac{\Gamma_a}{\Gamma_0} \tag{352} $$

A different choice of Γ_0 will result in the entropy being defined up to an additive constant.

The Entropy and Equilibrium States

The assumption of equal a priori probabilities is only a simplification of the widely held belief that a system's physical trajectory follows an intricate path which changes rapidly and is finely spread across the volume of accessible phase space. The corresponding physical distribution function will evolve with respect to time, according to Liouville's theorem. If ρ is expected to describe an equilibrium state S should not evolve.

We shall show that entropy defined by

$$S(t) \;=\; -\, k_B \int d\Gamma \; \rho(\{p_i, q_i\} : t) \; \ln \rho(\{p_i, q_i\} : t) \;-\; k_B \; \ln \Gamma_0 \qquad (353)$$

is time independent. The last term is an additive constant added to make the argument of the logarithm dimensionless and has no affect on our deliberations. The time derivative of the entropy is given by

$$
\begin{aligned}
\frac{dS}{dt} \;&=\; -\, k_B \int d\Gamma \left(\frac{\partial \rho}{\partial t} \; \ln \rho \;+\; \frac{\partial \rho}{\partial t} \right) \\
&=\; -\, k_B \int d\Gamma \; \frac{\partial \rho}{\partial t} \left(\ln \rho \;+\; 1 \right) \qquad (354)
\end{aligned}
$$

Using Liouville's theorem reduces this to

$$
\begin{aligned}
\frac{dS}{dt} \;&=\; k_B \int d\Gamma \left\{ \rho \,,\, H \right\} \left(\ln \rho \;+\; 1 \right) \\
&=\; -\, k_B \int d\Gamma \; \sum_{i=1}^{3N} \left[\frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial q_i} \;-\; \frac{\partial \rho}{\partial q_i} \frac{\partial H}{\partial p_i} \right] \left(\ln \rho \;+\; 1 \right)
\end{aligned}
$$

$$(355)$$

The terms linear in the derivatives of ρ can be transformed into factors of ρ, by integrating by parts. This yields

$$\frac{dS}{dt} \;=\; k_B \int d\Gamma \; \sum_{i=1}^{3N} \rho \left[\frac{\partial}{\partial p_i} \left(\frac{\partial H}{\partial q_i} \left(\ln \rho \;+\; 1 \right) \right) \;-\; \frac{\partial}{\partial q_i} \left(\frac{\partial H}{\partial p_i} \left(\ln \rho \;+\; 1 \right) \right) \right]$$

$$(356)$$

since the boundary terms vanish. On expanding the derivatives of the terms in the round parentheses, one finds that some terms cancel

$$\frac{dS}{dt} \;=\; k_B \int d\Gamma \left[\rho \, \frac{\partial^2 H}{\partial p_i \, \partial q_i} \left(\ln \rho \;+\; 1 \right) \;+\; \frac{\partial H}{\partial q_i} \frac{\partial \rho}{\partial p_i} \right.$$

79

$$- \rho \; \frac{\partial^2 H}{\partial q_i \; \partial p_i} \left(\ln \rho \; + \; 1 \right) \; - \; \frac{\partial H}{\partial p_i} \frac{\partial \rho}{\partial q_i} \; \Bigg]$$

$$= \; k_B \; \int d\Gamma \; \left[\; \frac{\partial H}{\partial q_i} \frac{\partial \rho}{\partial p_i} \; - \; \frac{\partial H}{\partial p_i} \frac{\partial \rho}{\partial q_i} \; \right] \tag{357}$$

which on integrating by parts yields

$$\frac{dS}{dt} \; = \; k_B \; \int d\Gamma \; \sum_{i=1}^{3N} \left[\; - \; \rho \left(\frac{\partial^2 H}{\partial p_i \; \partial q_i} \right) \; + \; \rho \left(\frac{\partial^2 H}{\partial p_i \; \partial q_i} \right) \; \right]$$

$$= \; 0 \tag{358}$$

Hence, the entropy of a state with a time-dependent probability density is constant.

From the above discussion, it is clear that the entropy of a system can only change if the Hamiltonian of the system is modified, such as by removing an external constraint. A removal of an external constraint will increase the volume of phase space to Γ_a' . Furthermore, the assumption of equal a prior probabilities implies that in the final equilibrium state the probability density ρ' will be uniformly spread over the increased available phase space. From the normalization condition

$$1 \; = \; \int d\Gamma' \; \rho' \tag{359}$$

one finds that in the final state, the probability distribution function is given by

$$\rho' \; = \; \frac{1}{\Gamma_a'} \tag{360}$$

Since the entropy is given by

$$S \; = \; - \; k_B \; \int d\Gamma \; \rho \; \ln \left(\; \rho \; \Gamma_0 \; \right) \tag{361}$$

where Γ_0 is constant and since $\Gamma_a' \; > \; \Gamma_a$, the entropy will have increased by an amount given by

$$\Delta S \; = \; k_B \; \ln \left(\frac{\Gamma_a'}{\Gamma_a} \right) \tag{362}$$

as expected from thermodynamics.

==

Example: Joule Free Expansion

We shall consider the Joule Free Expansion of an ideal gas. The gas is initially enclosed by a container of volume V, but the a valve is opened so that the gas can expand into an adjacent chamber which initially contained a vacuum.

The volume available to the gas in the final state is V'. Since the adjacent chamber is empty, no work is done in the expansion.

The Hamiltonian for an idea gas can be represented by

$$H = \sum_{i=1}^{3N} \frac{p_i^2}{2\,m} \tag{363}$$

so the sum of the squares of the is restricted by E. The volume of accessible phase space is given by

$$\Gamma_a = \prod_{i=1}^{3N} \left\{ \int dp_i \int dq_i \right\} \delta\left(E - \sum_{i=1}^{3N} \frac{p_i^2}{2\,m} \right) \tag{364}$$

The integrations for the spatial coordinates separates from the integration over the momenta. The integration over the three spatial coordinates for each particle produces a factor of the volume. The integration over the momenta will produce a result which depends on the energy $f(E)$ which is independent of the volume. Hence, the expression for the available phase space has the form

$$\Gamma_a = V^N\, f(E) \tag{365}$$

On recognizing that the particles are indistinguishable, one finds that the entropy is given by

$$\begin{aligned} S &= k_B \,\ln\frac{\Gamma_a}{\Gamma_0} \\ &= N\,k_B\,\ln V + k_B\,\ln f(E) - k_B\,\ln N! - N\,k_B\ln(\,2\,\pi\,\hbar\,) \end{aligned} \tag{366}$$

where Γ_0 is the measure of phase space that is used to define a single microscopic state. Thus, the change in entropy is given by

$$\Delta S = N\,k_B\,\ln\left(\frac{V'}{V}\right) \tag{367}$$

The same result may be obtained from thermodynamics. Starting from the expression for infinitesimal change of the internal energy

$$dU = T\,dS - P\,dV + \mu\,dN \tag{368}$$

and recognizing that Joule Free expansion is a process for which

$$\begin{aligned} dU &= 0 \\ dN &= 0 \end{aligned} \tag{369}$$

Therefore, one has

$$dS \;=\; \frac{P}{T} \; dV \tag{370}$$

and on using the equation of state for the ideal gas $P \; V \;=\; N \; k_B \; T$ one finds

$$dS \;=\; N \; k_B \; \frac{dV}{V} \tag{371}$$

which integrates to yield

$$\Delta S \;=\; N \; k_B \; \ln \left(\frac{V'}{V} \right) \tag{372}$$

Hence, the expression for the change in entropy derived by using Statistical Mechanics is in agreement with the expression derived by using Thermodynamics.

==

4 The Micro-Canonical Ensemble

4.1 Classical Harmonic Oscillators

Consider a set of N classical harmonic oscillators described by the Hamiltonian H

$$H = \sum_{i=1}^{dN} \left[\frac{p_i^2}{2\,m} + \frac{m\,\omega_0^2\,q^2}{2} \right] \qquad (373)$$

The above Hamiltonian may be adopted as a model of the vibrational motion of the atoms in a solid. However, it is being assumed that the vibrations of the atoms are independent and are harmonic and all the oscillators have the same frequency. Furthermore, our treatment will be based on classical mechanics.

We shall consider the system in the Micro-Canonical Ensemble, where the energy is determined to within an uncertainty ΔE

$$E > H > E - \Delta E \qquad (374)$$

The volume of accessible phase space Γ_a is given by the integral

$$\Gamma_a = \prod_{i=1}^{dN} \left\{ \int_{-\infty}^{\infty} dq_i \int_{-\infty}^{\infty} dp_i \right\} \left[\Theta(E - H) - \Theta(E - \Delta E - H) \right] \quad (375)$$

where $\Theta(x)$ is the Heaviside step function. On transforming the coordinates to $\tilde{p}_i = m\,\omega_0\,q_i$, the Hamiltonian can be written in the form

$$H = \frac{1}{2\,m} \sum_{=1}^{dN} \left[p_i^2 + \tilde{p}_i^2 \right] \qquad (376)$$

so the accessible phase space is defined by the inequalities

$$2\,m\,E > \sum_{i=1}^{dN} \left[p_i^2 + \tilde{p}_i^2 \right] > 2\,m\,(\,E - \Delta E\,) \qquad (377)$$

so

$$\Gamma_a = \prod_{i=1}^{dN} \left\{ \frac{1}{m\,\omega_0} \int_{-\infty}^{\infty} d\tilde{p}_i \int_{-\infty}^{\infty} dp_i \right\} \left[\Theta(E - H) - \Theta(E - \Delta E - H) \right] \quad (378)$$

Thus, the area of accessible phase space is proportional to the volume enclosed between two $2dN$-dimensional hyperspheres of radii $\sqrt{2\,m\,E}$ and $\sqrt{2\,m\,(\,E - \Delta E\,)}$. Therefore, we need to evaluate the volume enclosed by a hypersphere.

The Volume of a d-dimensional Hypersphere.

The volume of a d-dimensional hypersphere of radius R has the form

$$V_d(R) = C\, R^d \tag{379}$$

where C is a constant that depends on the dimensionality d. The constant can be determined by comparing two methods of evaluating the integral I_d, given by

$$I_d = \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \dots \int_{-\infty}^{\infty} dx_d \, \exp\left[-\sum_{i=1}^{d} x_i^2 \right] \tag{380}$$

In a d-dimensional Cartesian coordinate system, the integral can be evaluated as a product of d identical integrals

$$
\begin{aligned}
I_d &= \prod_{i=1}^{d} \left\{ \int_{-\infty}^{\infty} dx_i \, \exp\left[-x_i^2 \right] \right\} \\
&= \left\{ \sqrt{\pi} \right\}^d \\
&= \pi^{\frac{d}{2}} \tag{381}
\end{aligned}
$$

Alternatively, one may evaluate the integral in hyperspherical polar coordinates as

$$I_d = S_d \int_0^{\infty} dr\, r^{d-1} \, \exp\left[-r^2 \right] \tag{382}$$

where the radial coordinate is defined by

$$r^2 = \sum_{i=1}^{d} x_i^2 \tag{383}$$

and S_d is the surface area of a d-dimensional unit sphere. This integral can be re-written in term of the variable $t = r^2$ as

$$I_d = \frac{S_d}{2} \int_0^{\infty} dt\, t^{\frac{d-2}{2}} \, \exp\left[-t \right] \tag{384}$$

The integration is evaluated as

$$I_d = \frac{S_d}{2} \, \Gamma(\tfrac{d}{2}) \tag{385}$$

where $\Gamma(n+1) = n!$ is the factorial function. On equating the above two expressions, one obtains the equality

$$\frac{S_d}{2} \, \Gamma(\frac{d}{2}) = \pi^{\frac{d}{2}} \tag{386}$$

Hence, we find that the surface area of a unit d-dimensional sphere, S_d, is given by

$$S_d = 2 \, \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \tag{387}$$

Using this, one finds that the volume of a d-dimensional sphere of radius R is given by

$$
\begin{aligned}
V_d(R) &= S_d \int_0^R dr \, r^{d-1} \\
&= S_d \, \frac{1}{d} \, R^d \\
&= \frac{2}{d} \, \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \, R^d \\
&= \frac{\pi^{\frac{d}{2}}}{\frac{d}{2}\Gamma(\frac{d}{2})} \, R^d \\
&= \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)} \, R^d \tag{388}
\end{aligned}
$$

which is our final result.

The Volume of Accessible Phase Space

The volume of accessible phase space Γ_a is proportional to the volume enclosed by two $2dN$-dimensional hyperspheres of radius $\sqrt{2\,m\,E}$ and $\sqrt{2\,m\,(E-\Delta E)}$. Using the above results, one finds

$$
\Gamma_a = \frac{\pi^{dN}}{\Gamma(dN+1)} \left(\frac{1}{m\,\omega_0}\right)^{dN} \left[(2\,m\,E)^{dN} - (2\,m\,(E-\Delta E))^{dN}\right] \tag{389}
$$

where the factor of

$$
\left(\frac{1}{m\,\omega_0}\right)^{dN} \tag{390}
$$

is the Jacobian for the coordinate transformations. The number of accessible microstates N_Γ is then defined as

$$
\begin{aligned}
N_\Gamma &= \frac{\Gamma_a}{(2\,\pi\,\hbar\,)^{dN}} \\
&= \frac{1}{\Gamma(dN+1)} \left(\frac{E}{\hbar\,\omega_0}\right)^{dN} \left[1 - \left(1 - \frac{\Delta E}{E}\right)^{dN}\right] \tag{391}
\end{aligned}
$$

The second factor in the square brackets is extremely small when compared to unity since the term in the parenthesis is less than unity and the exponent is extremely large. Therefore, it can be neglected

$$
N_\Gamma \approx \frac{1}{\Gamma(dN+1)} \left(\frac{E}{\hbar\,\omega_0}\right)^{dN} \tag{392}
$$

This implies that, for sufficiently high dimensions, the volume of the hypersphere is the same as the volume of the hypershell.

Derivation of Stirling's Approximation

The Gamma function is defined by the integral

$$\Gamma(n+1) = \int_0^\infty dx \; x^n \; \exp\left[-x \right] \tag{393}$$

which, for integer n coincides with $n!$. This can be verified by repeated integration by parts

$$
\begin{aligned}
\Gamma(n+1) &= \int_0^\infty dx \; x^n \; \exp\left[-x \right] \\
&= -\int_0^\infty dx \; x^n \; \frac{\partial}{\partial x} \exp\left[-x \right] \\
&= -x^n \; \exp\left[-x \right]\Big|_0^\infty + n \int_0^\infty dx \; x^{n=1} \; \exp\left[-x \right] \\
&= n \int_0^\infty dx \; x^{n-1} \; \exp\left[-x \right] \\
&= n \; \Gamma(n)
\end{aligned}
\tag{394}
$$

which together with

$$\Gamma(1) = \int_0^\infty dx \; \exp\left[-x \right] = 1 \tag{395}$$

leads to the evaluation of the integral as

$$\Gamma(n+1) = \int_0^\infty dx \; x^n \; \exp\left[-x \right] = n! \tag{396}$$

for integer n.

Stirling's approximation to $\ln n!$ can be obtained by evaluating the integral using the method of steepest descents.

$$n! = \int_0^\infty dx \; \exp\left[-x + n \ln x \right] \tag{397}$$

The extremal value of $x = x_c$ is found from equating the derivative of the exponent to zero

$$-1 + \frac{n}{x_c} = 0 \tag{398}$$

This yields $x_c = n$. On expanding the integrand to second order in $(x - x_c)$, one has

$$n! \approx \int_0^\infty dx \; \exp\left[-x_c + n \ln x_c \right] \exp\left[-\frac{n}{x_c^2} (x - x_c)^2 \right] \tag{399}$$

86

On extending the lower limit of the integration to $-\infty$, one obtains the approximation

$$
\begin{aligned}
n! &\approx \int_{-\infty}^{\infty} dx \, \exp\left[\, - \, x_c \, + \, n \, \ln x_c \, \right] \, \exp\left[\, - \, \frac{n}{2 \, x_c^2} \, (\, x \, - \, x_c \,)^2 \, \right] \\
&= \sqrt{\frac{2 \, \pi \, x_c^2}{n}} \, \exp\left[\, - \, x_c \, + \, n \, \ln x_c \, \right]
\end{aligned}
\tag{400}
$$

This is expected to be valid for sufficiently large n. On setting $x_c \, = \, n$, one has

$$
n! \; \approx \; \sqrt{2 \, \pi \, n} \, \exp\left[\, - \, n \, + \, n \, \ln n \, \right]
\tag{401}
$$

Stirling's approximation is obtained by taking the logarithm, which yields

$$
\ln n! \, = \, n \, \ln n \, - \, n \, + \, \frac{1}{2} \, \ln(\, 2 \, \pi \, n \,)
\tag{402}
$$

Stirling's approximation will be used frequently throughout this course.

The Entropy

The entropy S is given by

$$
\begin{aligned}
S \; &= \; k_B \, \ln N_\Gamma \\
&= \; d \, N \, k_B \, \ln\left(\frac{E}{\hbar \, \omega_0} \right) \, - \, k_B \, \ln(dN)!
\end{aligned}
\tag{403}
$$

The logarithm of $N!$ can be approximated for large N by Stirling's approximation. This can be quickly re-derived by noting that

$$
\ln N! \, = \, \ln N \, + \, \ln(N - 1) \, + \, \ln(N - 2) \, + \, \ln 2 \, + \, \ln 1
\tag{404}
$$

For large N, the sum on the right hand side can be approximated by an integral

$$
\begin{aligned}
\ln N! \; &\approx \; \int_0^N dx \, \ln x \\
&\approx \; x \, (\, \ln x \, - \, 1 \,)\Big|_0^N \\
&\approx \; N \, (\, \ln N \, - \, 1 \,)
\end{aligned}
\tag{405}
$$

which results in Stirling's approximation

$$
\ln N! \, = \, N \, (\, \ln N \, - \, 1 \,)
\tag{406}
$$

Using Stirling's approximation in the expression for the entropy S, one obtains

$$
\begin{aligned}
S(E, N) \; &= \; d \, N \, k_B \, \ln\left(\frac{E}{\hbar \, \omega_0} \right) \, - \, k_B \, d \, N \, (\, \ln(dN) \, - \, 1 \,) \\
&= \; d \, N \, k_B \, \ln\left(\frac{E}{d \, N \, \hbar \, \omega_0} \right) \, + \, k_B \, d \, N
\end{aligned}
\tag{407}
$$

which shows that the entropy is an extensive monotonically increasing function of E. This is the fundamental relation. In the Micro-Canonical Ensemble, the energy E is the thermodynamic energy U.

The temperature is defined by the derivative

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_N \tag{408}$$

which yields

$$\frac{1}{T} = \frac{d N k_B}{U} \tag{409}$$

Hence, we find that the internal energy U is given by

$$U = d N k_B T \tag{410}$$

which shows that each degree of freedom carries the thermodynamic energy $k_B T$. The specific heat at constant volume is then found as

$$C_V = d N k_B \tag{411}$$

which is Dulong and Petit's law[13]. Dulong and Petit's law describes the high-temperature specific heat of solids quite well, but fails at low temperatures where the quantum mechanical nature of the solid manifests itself.

4.2 An Ideal Gas of Indistinguishable Particles

The Hamiltonian for an ideal gas is written as the sum of the kinetic energies

$$H = \sum_{i=1}^{dN} \frac{p_i^2}{2 m} \tag{412}$$

The gas is contained in a volume V with linear dimensions L, such that

$$V = L^d \tag{413}$$

where d is the number of dimensions of space. In the Micro-Canonical Ensemble, the energy is constrained to an interval of width ΔE accorsing to the inequality

$$E > H > E - \Delta E \tag{414}$$

The volume of accessible phase space Γ_a is given by the multiple integral

$$\Gamma_a = \prod_{i=1}^{dN} \left\{ \int_0^L dq_i \int_{-\infty}^{\infty} dp_i \right\} \left[\Theta(E - H) - \Theta(E - \Delta E - H) \right] \tag{415}$$

---

[13]A.-T. Petit and P.-L. Dulong, "Recherches sur quelques points importants de la Théorie de la Chaleur", Annales de Chimie et de Physique 10, 395-413 (1819).

The integration over the coordinates can be performed, leading to the expression

$$
\begin{aligned}
\Gamma_a \;&=\; \prod_{i=1}^{dN} \left\{ L \int_{-\infty}^{\infty} dp_i \right\} \Big[\Theta(E - H) \;-\; \Theta(E - \Delta E - H) \Big] \\
&=\; L^{dN} \prod_{i=1}^{dN} \left\{ \int_{-\infty}^{\infty} dp_i \right\} \Big[\Theta(E - H) \;-\; \Theta(E - \Delta E - H) \Big] \\
&=\; V^{N} \prod_{i=1}^{dN} \left\{ \int_{-\infty}^{\infty} dp_i \right\} \Big[\Theta(E - H) \;-\; \Theta(E - \Delta E - H) \Big] \quad (416)
\end{aligned}
$$

The step functions constrain the momenta such that

$$
2\, m\, E \;>\; \sum_{i=1}^{dN} p_i^2 \;>\; 2\, m\, (\, E \;-\; \Delta E\,) \qquad (417)
$$

Thus, the integration over the momenta is equal to the volume contained between two dN-dimensional hyperspheres of radii $\sqrt{2\,m\,E}$ and $\sqrt{2\,m\,(\,E\;-\;\Delta E\,)}$. Using the expressions for the volume of a dN dimensional hypersphere

$$
V_{dN}(R) \;=\; \frac{\pi^{\frac{dN}{2}}}{\Gamma(\frac{dN}{2} + 1)}\, R^{dN} \qquad (418)
$$

which yields

$$
\begin{aligned}
\Gamma_a \;&=\; V^{N} \frac{\pi^{\frac{dN}{2}}}{\Gamma(\frac{dN}{2} + 1)} \left(\, 2\, m\, E\, \right)^{\frac{dN}{2}} \left[1 \;-\; \left(1 \;-\; \frac{\Delta E}{E} \right)^{\frac{dN}{2}} \right] \\
&=\; V^{N} \frac{\pi^{\frac{dN}{2}}}{\Gamma(\frac{dN}{2} + 1)} \left(\, 2\, m\, E\, \right)^{\frac{dN}{2}} \left(\, 1 \;-\; \exp\left[-\, \frac{d\, N \Delta E}{2\, E} \right]\, \right) (419)
\end{aligned}
$$

Since

$$
1 \;\gg\; \exp\left[-\, \frac{d\, N \Delta E}{2\, E} \right] \qquad (420)
$$

the volume of accessible phase space is given by

$$
\Gamma_a \;=\; V^{N} \frac{\pi^{\frac{dN}{2}}}{\Gamma(\frac{dN}{2} + 1)} \left(\, 2\, m\, E\, \right)^{\frac{dN}{2}} \qquad (421)
$$

However, for an ideal gas of identical particles, we have to take into account that specifying all the momenta p_i and coordinates q_i of the N particles provides us with too much information. Since the N particles are identical, we cannot distinguish between two points of phase space that differ only by the interchange of identical particles. There are $N!$ points corresponding to the different labelings of the particles. These $N!$ points represent the same microstates of the system. To only account for the different microstates, one must divide the volume of

accessible phase space by $N!$. Hence, the number of microscopic states N_Γ is given by

$$
\begin{aligned}
N_\Gamma &= \frac{\Gamma_a}{N! \ (\ 2 \ \pi \ \hbar \ )^{dN}} \\
&= \frac{V^N}{N! \ \Gamma(\frac{dN}{2} + 1)} \ \left(\frac{m \ E}{2 \ \pi \ \hbar^2} \right)^{\frac{dN}{2}}
\end{aligned}
\tag{422}
$$

The entropy S is given by the expression

$$
\begin{aligned}
S &= k_B \ \ln N_{\Gamma_a} \\
&= k_B \ \ln \left[\frac{V^N}{N! \ \Gamma(\frac{dN}{2} + 1)} \ \left(\frac{m \ E}{2 \ \pi \ \hbar^2} \right)^{\frac{dN}{2}} \right] \\
&= k_B \ \ln \left[\frac{V^N}{N! \ (\frac{dN}{2})!} \ \left(\frac{m \ E}{2 \ \pi \ \hbar^2} \right)^{\frac{dN}{2}} \right]
\end{aligned}
\tag{423}
$$

On using Stirling's formulae

$$
\ln N! = N \ (\ \ln N \ - \ 1 \ )
\tag{424}
$$

valid for large N, one finds

$$
S = N \ k_B \ \ln \left(\frac{V}{N} \right) + \frac{dN}{2} \ k_B \ \ln \left(\frac{m \ E}{2 \ \pi \ \hbar^2 \ \frac{d \ N}{2}} \right) + N \ k_B + \frac{d}{2} \ N \ k_B
\tag{425}
$$

or

$$
S = N \ k_B \ \ln \left[\ \left(\frac{V}{N} \right) \left(\frac{m \ E}{d \ \pi \ \hbar^2 \ N} \right)^{\frac{d}{2}} \ \right] + \left(\frac{d + 2}{2} \right) \ N \ k_B
\tag{426}
$$

On identifying E with U, the thermodynamic energy, one has the fundamental relation $S(U, V, N)$ from which all thermodynamic quantities can be obtained.

The intensive quantities can be obtained by taking the appropriate derivatives. For example, the temperature is found from

$$
\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V,N}
\tag{427}
$$

which yields

$$
\frac{1}{T} = \frac{d \ N \ k_B}{2 \ U}
\tag{428}
$$

Hence, we have recovered the equation of state for an ideal gas

$$
U = \frac{d \ N}{2} \ k_B \ T
\tag{429}
$$

90

Likewise, the pressure is given by

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{U,N} \tag{430}$$

which yields

$$\frac{P}{T} = \frac{N \, k_B}{V} \tag{431}$$

which is the ideal gas law. The chemical potential μ is found from

$$-\frac{\mu}{T} = \left(\frac{\partial S}{\partial N} \right)_{U,V} \tag{432}$$

which yields

$$-\frac{\mu}{T} = k_B \, \ln \left[\left(\frac{V}{N} \right) \left(\frac{m \, U}{d \, \pi \, \hbar^2 \, N} \right)^{\frac{d}{2}} \right] - \left(\frac{d+2}{2} \right) k_B + \left(\frac{d+2}{2} \right) k_B \tag{433}$$

Since

$$\frac{U}{d \, N} = \frac{k_B \, T}{2} \tag{434}$$

one has

$$-\frac{\mu}{T} = k_B \, \ln \left[\left(\frac{V}{N} \right) \left(\frac{m \, k_B \, T}{2 \, \pi \, \hbar^2} \right)^{\frac{d}{2}} \right] \tag{435}$$

This can be re-written as

$$\frac{\mu}{T} = k_B \, \ln P + f(T) \tag{436}$$

where P is the pressure and $f(T)$ is a function of only the temperature T.

On substituting the equation of state

$$U = \frac{d \, N}{2} \, k_B \, T \tag{437}$$

into the expression for the entropy, one finds

$$S = N \, k_B \, \ln \left[\left(\frac{V}{N} \right) \left(\frac{m \, k_B \, T}{2 \, \pi \, \hbar^2} \right)^{\frac{d}{2}} \right] + \left(\frac{d+2}{2} \right) N \, k_B \tag{438}$$

This is the Sackur-Tetrode formula for the entropy of an ideal gas.

$$\lambda = \frac{h}{(\, 2 \, \pi \, m \, k_B \, T \,)^{\frac{1}{2}}} \tag{439}$$

The SackurTetrode equation was derived independently by Hugo Martin Tetrode and Otto Sackur, using Maxwell-Boltzmann statistics in 1912.

Note that the factor

$$(\ 2 \ \pi \ m \ k_B \ T \ )^{\frac{1}{2}} \tag{440}$$

has the character of an average thermal momentum of a molecule. We can define λ via

$$\lambda \ = \ \frac{h}{(\ 2 \ \pi \ m \ k_B \ T \ )^{\frac{1}{2}}} \tag{441}$$

as a thermal de Broglie wave length associated with the molecule. The entropy can be re-written as

$$S \ = \ N \ k_B \ \ln \left[\frac{V}{N \ \lambda^d} \right] \ + \ \left(\frac{d \ + \ 2}{2} \right) \ N \ k_B \tag{442}$$

which shows that the entropy S is essentially determined by the ratio of the volume per particle to the volume λ^d associated with the thermal de Broglie wavelength. The classical description is approximately valid in the limit

$$\frac{V}{N \ \lambda^d} \ \gg \ 1 \tag{443}$$

where the uncertainties in particle positions are negligible compared with the average separation of the particles. When the above inequality does not apply, quantum effects become important.

The Momentum Distribution of an atom in an Ideal Gas

The probability that a particle has momentum of magnitude $|\underline{p}|$ can be obtained using the Micro-Canonical Ensemble. The probability is found from the probability distribution by integrating over the coordinates of all the particles and integrating over all the momenta of the other particles. Thus, we find the momentum probability distribution function $P(|\underline{p}|)$ via

$$
\begin{aligned}
P(|\underline{p}|) \ = \ & \frac{1}{\Gamma_a} \ \prod_{i=1}^{dN} \left\{ \int_0^L dq_i \right\} \ \prod_{i=d+1}^{dN} \left\{ \int_{-\infty}^{\infty} dp_i \right\} \\
& \times \left[\Theta \left(2 \ m \ E \ - \ \sum_{i=1}^{dN} p_i^2 \right) \ - \ \Theta \left(2 \ m \ (\ E \ - \ \Delta E \ ) \ - \ \sum_{i=1}^{dN} p_i^2 \right) \right]
\end{aligned}
\tag{444}
$$

which is evaluated as

$$P(|\underline{p}|) \ = \ \frac{\Gamma(\frac{dN}{2} + 1)}{\pi^{\frac{d}{2}} \ \Gamma(\frac{d(N-1)}{2} + 1)} \ \frac{\left(\ 2 \ m \ E \ - \ |\underline{p}|^2 \ \right)^{\frac{d(N-1)}{2}}}{\left(\ 2 \ m \ E \ \right)^{\frac{dN}{2}}}$$

$$
= \quad \frac{\Gamma(\frac{dN}{2}+1)}{\pi^{\frac{d}{2}} \, \Gamma(\frac{d(N-1)}{2}+1)} \; \frac{\left(1 \; - \; \frac{\mid \underline{p} \mid^2}{2 \, m \, E} \right)^{\frac{d(N-1)}{2}}}{\left(2 \, m \, E \right)^{\frac{d}{2}}}
$$

$$
\approx \quad \left(\frac{d \, N}{2 \, \pi} \right)^{\frac{d}{2}} \; \frac{\exp\left[\; - \; \frac{d(N-1)}{2} \; \frac{\mid \underline{p} \mid^2}{2 \, m \, E} \right]}{\left(2 \, m \, E \right)^{\frac{d}{2}}}
$$

$$
\approx \quad \left(\frac{d \, N}{4 \, \pi \, m \, E} \right)^{\frac{d}{2}} \; \exp\left[\; - \; \frac{d(N-1)}{2} \; \frac{\mid \underline{p} \mid^2}{2 \, m \, E} \right] \tag{445}
$$

Which is the desired result. On using the thermodynamic relation for the energy

$$
U \; = \; \frac{d \, N}{2} \; k_B \, T \tag{446}
$$

one obtains the Maxwell distribution

$$
P(\mid \underline{p} \mid) \; = \; \left(2 \, \pi \, m \, k_B \, T \right)^{-\frac{d}{2}} \; \exp\left[\; - \; \frac{\mid \underline{p} \mid^2}{2 \, m \, k_B \, T} \right] \tag{447}
$$

which is properly normalized.

4.3   Spin One-half Particles

A system of spin one-half particles is described by a discrete not continuous phase space. The discreteness, is due to the discrete nature of the quantum mechanical eigenvalues.
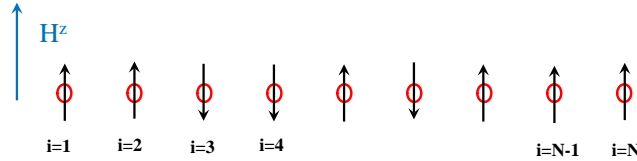
Figure 22: A set of N spins, in the presence of a uniform applied magnetic field H^z directed along the z-axis. The spins are quantized along the z-direction, so their S^z components are given by $\pm \hbar/2$.

Consider a set of N spin one-half particles, in an applied magnetic field. The particles may be either aligned parallel or anti-parallel to the applied magnetic

field H^z. The Hamiltonian describing the spins is given by

$$H = -g\,\mu_B \sum_{i=1}^{N} S_i^z\,H^z \tag{448}$$

where $S^z = \pm \frac{1}{2}$. If one defines the total magnetic moment as

$$M^z = g\,\mu_B \sum_{i=1}^{N} S_i^z \tag{449}$$

one finds that the energy is determined by the magnetization via

$$H = -M^z\,H^z \tag{450}$$

Therefore, if the energy has a fixed value E, the accessible microstates are determined by the fixed value of the magnetization M^z. We shall introduce the dimensionless magnetization as

$$m = M^z/\mu_B \tag{451}$$

Hence, for a fixed energy there are $(N+m)/2$ spin-up particles and $(N-m)/2$ spin-down particles. The number of ways of selecting $(N+m)/2$ particles out of N particles as being spin up is given by

$$\frac{N!}{\left(\frac{N-m}{2}\right)!} \tag{452}$$

since there are N ways of selecting the first particle as being spin up, $(N-1)$ ways of selecting the second particle as being spin up, etc. This process continues until the $(N+m)/2$ spin-up particle is chosen and this can be selected in $(N+1-(N+m)/2)$ ways. Since the number of choices is multiplicative, the product of the number of choices give the result above. However, not all of these choices lead to independent microstates. Interchanges of the $(N+m)/2$ spin up particles between themselves lead to identical microstates. There are $((N+m)/2)!$ such interchanges. The total number of discrete microstates with magnetization m is found by diving the above result by $((N+m)/2)!$. The end result is N_Γ given by

$$N_\Gamma = \frac{N!}{\left(\frac{N+m}{2}\right)!\,\left(\frac{N-m}{2}\right)!} \tag{453}$$

The entropy S is found from

$$S = k_B\,\ln N_\Gamma \tag{454}$$

which is evaluated as

$$
\begin{aligned}
S &= k_B \, \ln \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} \\
&= k_B \, \ln N! \; - \; k_B \, \ln \left(\frac{N+m}{2}\right)! \; - \; k_B \, \ln \left(\frac{N-m}{2}\right)! \\
&= k_B \, N \, \ln N \; - \; k_B \left(\frac{N+m}{2}\right) \ln \left(\frac{N+m}{2}\right) \; - \; k_B \left(\frac{N-m}{2}\right) \ln \left(\frac{N-m}{2}\right)
\end{aligned}
\tag{455}
$$

where we have used Stirling's approximation in the last line. Hence, the entropy has been expressed in terms of n and m or equivalently in terms of E and N, This is the fundamental relation, from which we may derive all thermodynamic quantities.
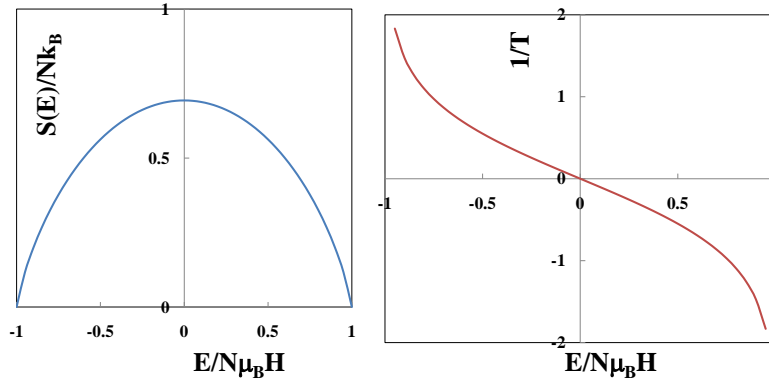
Figure 23: The entropy $S(E)$ as a function of entropy E for a model of a system of spins in a magnetic field is shown in (a). Since the energy is bounded from above, the entropy is not a monotonically increasing function of E. This has the consequence that T can become negative when there is population inversion, as is indicated in (b).

On identifying the fixed energy with the thermodynamic energy, one may use the definition of temperature

$$
\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_N
\tag{456}
$$

or

$$
\begin{aligned}
\frac{1}{T} &= \left(\frac{\partial S}{\partial m}\right)_N \left(\frac{\partial m}{\partial U}\right) \\
&= -\left(\frac{\partial S}{\partial m}\right)_N \frac{1}{\mu_B \, H_z}
\end{aligned}
$$

95

Therefore, one has

$$
\begin{aligned}
\frac{\mu_B \; H^z}{k_B \; T} &= \frac{1}{2} \; \ln\left(\frac{N+m}{2}\right) - \frac{1}{2} \; \ln\left(\frac{N-m}{2}\right) \\
&= \frac{1}{2} \; \ln\left(\frac{N+m}{N-m}\right) \quad\quad (458)
\end{aligned}
$$

This can be exponentiated to yield

$$
\exp\left[\frac{2 \; \mu_B \; H^z}{k_B \; T}\right] = \left(\frac{N+m}{N-m}\right) \quad\quad (459)
$$

which can be solved for m as

$$
\begin{aligned}
m &= N \; \left(\frac{\exp[\frac{2 \; \mu_B \; H^z}{k_B \; T}] \; - \; 1}{\exp[\frac{2 \; \mu_B \; H^z}{k_B \; T}] \; + \; 1}\right) \\
&= N \; \tanh\left(\frac{\mu_B \; H^z}{k_B \; T}\right) \quad\quad (460)
\end{aligned}
$$

Hence, the magnetization is an odd function of H^z and saturates for large fields and low temperatures at $\pm N$. Finally, we obtain the expression for the thermal average of the internal energy

$$
U = - N \; \mu_B \; H^z \; \tanh\left(\frac{\mu_B \; H^z}{k_B \; T}\right) \quad\quad (461)
$$

which vanishes as the square of the field H^z in the limit of zero applied field, since the Hamiltonian is linear in H^z and since the magnetization is expected to vanish linearly as H^z vanishes.

Zero Applied Field

We shall now determine the magnetization probability distribution function in the limit of zero applied magnetic field. The spins of the particles may either be aligned parallel or anti-parallel to the axis of quantization. There are a total of 2^N possible microstates. Hence, for zero applied field

$$
N_\Gamma = 2^N \quad\quad (462)
$$

Since all microstates are assumed to occur with equal probabilities, the probability of finding a system with magnetization m is given by

$$
\begin{aligned}
P(m) &= \frac{1}{N_\Gamma} \; \frac{N!}{\left(\frac{N+m}{2}\right)! \; \left(\frac{N-m}{2}\right)!} \\
&= \left(\frac{1}{2}\right)^N \; \frac{N!}{\left(\frac{N+m}{2}\right)! \; \left(\frac{N-m}{2}\right)!} \quad\quad (463)
\end{aligned}
$$

which is normalized to unity. On using the more accurate form of Stirling's approximation that we found using the method of steepest descents

$$
\begin{aligned}
\ln N! &\approx \frac{1}{2} \ln(2 \pi N) + N \ln N - N \\
&\approx \frac{1}{2} \ln(2 \pi) + (N + \frac{1}{2}) \ln N - N
\end{aligned}
\tag{464}
$$

in $\ln P(m)$, one obtains

$$
\begin{aligned}
\ln P(m) &\approx - N \ln 2 - \frac{1}{2} \ln(2 \pi) + (N + \frac{1}{2}) \ln N \\
&\quad - \frac{N + 1}{2} (1 + \frac{m}{N + 1}) \ln \frac{N}{2} (1 + \frac{m}{N}) \\
&\quad - \frac{N + 1}{2} (1 - \frac{m}{N + 1}) \ln \frac{N}{2} (1 - \frac{m}{N})
\end{aligned}
\tag{465}
$$

On expanding in powers of m, the expression simplifies to

$$
\ln P(m) \approx - \frac{1}{2} \ln N - \frac{1}{2} \ln(2 \pi) + \ln 2 - \frac{m^2}{2 N} + \ldots
\tag{466}
$$

Hence, one finds that the magnetization probability distribution function $P(m)$ is approximated by a Gaussian distribution

$$
P(m) \approx \sqrt{\frac{2}{\pi N}} \exp \left[- \frac{m^2}{2 N} \right]
\tag{467}
$$

Therefore, the most probable value of the magnetization is $m = 0$ and the width of the distribution is given by $\sqrt{N}$. This is small compared with the total range of the possible magnetization which is $2 N$. Most of the microstates correspond to zero magnetization. This can be seen as total number of available microstates is given by

$$
2^N
\tag{468}
$$

and since the number of states with zero magnetization is given by

$$
\frac{N!}{(N/2)! \, (N/2)!} \sim \sqrt{\frac{2}{\pi N}} \, 2^N
\tag{469}
$$

Thus, this implies that, for $H^z = 0$, the relative size of the fluctuations in the magnetization is small.

4.4 The Einstein Model of a Crystalline Solid

The Einstein Model of a crystalline solid considers the normal modes of vibrations of the lattice to be quantized, and it assumes that the frequencies of the all the normal modes are identical and equal to ω_0. It is a reasonable approximation for the optical phonon modes in a solid. For a solid with N unit cells,
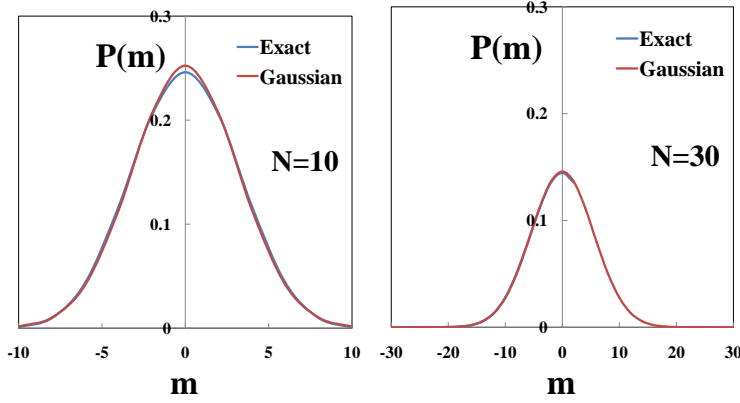
Figure 24: The exact probability distribution $P(m)$ of the magnetic moment m for a system of N spins, and the approximate Gaussian distribution. After scaling m with the size of the system N, the width of the distribution decreases on increasing N.

where there are p atoms per unit cell, one expects there to be $N' = 3(p-1)N$ optic modes. The remaining $3N$ modes are expected to be acoustic modes.

Consider a set of N' quantum mechanical harmonic oscillators in the Micro-Canonical Ensemble. Each oscillator has the same frequency ω_0. The total energy E is given by the sum of the energies of each individual quantum oscillator

$$E = \sum_{i=1}^{N'} \hbar \omega_0 \left(n_i + \frac{1}{2} \right) \tag{470}$$

where n_i is the number of quanta in the i-th oscillator. The possible values of n_i are the set of 0, 1, 2, ... , ∞. The last term in the round parenthesis is the zero-point energy of the i-th oscillator.

If we subtract the zero-point energy for each quantum oscillator, then the energy E_{exc} available to distribute amongst the N' quantum mechanical harmonic oscillators is given by

$$E_{exc} = E - N' \left(\frac{\hbar \omega_0}{2} \right) \tag{471}$$

The excitation energy E_{exc} is to be distributed amongst the N' quantum oscillators

$$E_{exc} = \sum_{i=1}^{N'} \hbar \omega_0 \ n_i \tag{472}$$

98

The total number of quanta Q available to the entire system is given by

$$\frac{E_{exc}}{\hbar\,\omega_0} = \sum_{i=1}^{N'} n_i = Q \tag{473}$$

which acts as a restriction on the possible sets of values of n_i. Each possible distribution of the Q quanta is described by a set of integer values, $\{n_i\}$, which uniquely describes a microstate of the system. In any allowed microstate the values of $\{n_i\}$ are restricted so that the number of quanta add up to Q.

There are Q quanta which must be distributed between the N' oscillators. We shall count all the possible ways of distributing the Q quanta among the N' oscillators. Let us consider each oscillator as a box and each quanta as a marble. Eventually, the marbles are to be considered as being indistinguishable, so interchanging any number of marbles will lead to the same configuration. We shall temporarily suspend this assumption and instead assume that the marbles could be tagged. Later, we shall restore the assumption of indistinguishability.
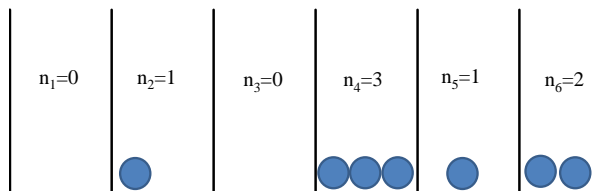
Figure 25: One microscopic state of a system in which Q quanta have been distributed amongst N' oscillators ($Q = 7, N' = 6$).

The Number of Distinguishable Ways

The number of ways of putting Q marbles in N' boxes can be found by arranging the boxes in a row. In this case, a box shares a common wall with its neighboring boxes so there are $N' + 1$ walls for the N' boxes. If one considers both the walls and marbles as being distinguishable objects, then in any distribution of the marbles in the boxes there are $Q + N' + 1$ objects in a row. If there are n_i marbles between two consecutive walls, then that box contains n_i marbles. If there are two consecutive walls in a distribution, then that box is empty. However, the first object and the last object are always walls, so really there are only $Q + N' - 1$ objects that can be re-arranged. Therefore, the total number of orderings can be found from the number of ways of arranging $Q + N' - 1$ objects in a row. This can be done in

$$(Q + N' - 1)! \tag{474}$$

number of ways. This happens since there are $Q + N' - 1$ ways of selecting the first object. After the first object has been chosen, there are $Q + N' - 2$ objects that remain to be selected, so there are only $Q + N' - 2$ ways of selecting the second object. Likewise, there are $Q + N' - 3$ ways of choosing the third object, and this continues until only the last object is unselected, in which case there is only one possible way of choosing the last object. The number of possible arrangements is given by the product of the number of ways of making each independent choice. Thus, we have found that there are $(Q + N' - 1)!$ possible ways of sequencing or ordering $(Q + N' - 1)$ distinguishable objects.

The Number of Indistinguishable Ways

We do need to consider the walls as being indistinguishable and also the marbles should be considered as indistinguishable. If we permute the indistinguishable walls amongst themselves, the ordering that results is identical to the initial ordering. There are $(N' - 1)!$ ways of permuting the $N' - 1$ walls amongst themselves. Hence, we should divide by $(N' - 1)!$ to only count the number of orderings made by placing the marbles between indistinguishable walls. Likewise, if one permutes the Q indistinguishable marbles, it leads to an identical ordering, and there are $Q!$ such permutations. So we have over-counted the number of orderings by $Q!$, and hence we also need to divide our result by a factor of $Q!$. Therefore, the total number of inequivalent ways N_Γ of distributing Q indistinguishable marbles in N' boxes is given by

$$N_\Gamma = \frac{(N' + Q - 1)!}{(N' - 1)! \; Q!} \tag{475}$$

This is equal to the total number of microstates N_Γ consistent with having a total number of quanta Q distributed amongst N' oscillators.

The Entropy

In the Micro-Canonical ensemble, the entropy S is given by the logarithm of the number of accessible microstates N_Γ

$$S = k_B \; \ln \; N_\Gamma \tag{476}$$

On substituting the expression for N_Γ, one obtains

$$\begin{aligned} S &= k_B \; \ln \left(\frac{(N' + Q - 1)!}{(N' - 1)! \; Q!} \right) \\ &= k_B \left[\; \ln(N' + Q - 1)! \; - \; \ln(N' - 1)! \; - \; \ln Q! \; \right] \end{aligned} \tag{477}$$

On using Stirling's approximation

$$\ln N! \approx N \left(\ln \; N \; - \; 1 \right) \tag{478}$$

100

valid for large N, for all three terms, after some cancellation one has

$$S \approx k_B \left[(N'+Q-1) \ln(N'+Q-1) - (N'- 1) \ln(N' - 1) - Q \ln Q \right] \quad (479)$$

which is valid for large Q and N'. It should be recalled that $Q = (E_{exc}/\hbar \omega_0)$, so S is a function of the total energy E. The above relation between the entropy and the total energy is the same as the relation between the entropy and the thermodynamic energy U. The expression for S in terms of U is the "Fundamental Relation" for the thermodynamics of the model.
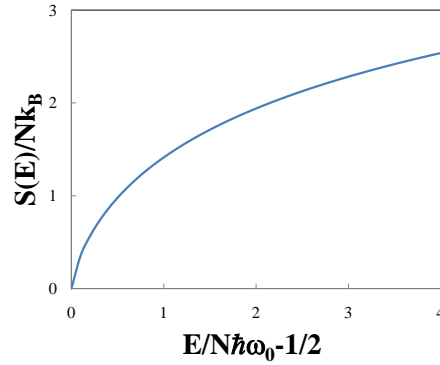
Figure 26: The entropy $S(E)$ versus the dimensionless excitation energy, for the Einstein model for the specific heat a solid.

We shall now consider the system to be in thermal equilibrium with a thermal reservoir held at temperature T. The temperature is defined by

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_N \quad (480)$$

which yields

$$\begin{aligned}
\frac{1}{T} &= \left(\frac{\partial S}{\partial Q} \right) \left(\frac{\partial Q}{\partial U} \right) \\
&= k_B \left[\ln(N' + Q - 1) - \ln Q \right] \left(\frac{\partial Q}{\partial U} \right) \\
&= \frac{k_B}{\hbar \omega_0} \left[\ln(N' + Q - 1) - \ln Q \right] \\
&= \frac{k_B}{\hbar \omega_0} \ln \left(\frac{N' + Q - 1}{Q} \right) \\
&= \frac{k_B}{\hbar \omega_0} \ln \left(\frac{\hbar \omega_0 (N' - 1) + U_{exc}}{U_{exc}} \right) \quad (481)
\end{aligned}$$

101

where it is now understood that the energy is the thermodynamic value U that is determined by T. On multiplying by $\hbar\,\omega_0/k_B$ and then exponentiating the equation, one finds

$$\exp\left[\frac{\hbar\,\omega_0}{k_B\,T}\right] \;=\; \left(\frac{\hbar\,\omega_0\,(\,N'\,-\,1\,)\,+\,U_{exc}}{U_{exc}}\right) \tag{482}$$

or on multiplying through by U_{exc}

$$U_{exc}\,\exp\left[\frac{\hbar\,\omega_0}{k_B\,T}\right] \;=\; \hbar\,\omega_0\,(\,N'\,-\,1\,)\,+\,U_{exc} \tag{483}$$

This equation can be solved to yield U_{exc} as a function of T

$$U_{exc} \;=\; \frac{\hbar\,\omega_0\,(\,N'\,-\,1\,)}{\exp[\,\frac{\hbar\,\omega_0}{k_B\,T}\,]\,-\,1} \tag{484}$$

We can neglect the term 1 compared with N' since, in our derivation we have assumed that N' is very large. Since

$$U_{exc} \;=\; \sum_{i=1}^{N'}\,\hbar\,\omega_0\,\overline{n}_i \tag{485}$$

we have found that the thermodynamic average number of quanta $\overline{n}$ of energy $\hbar\omega_0$ in each quantum mechanical harmonic oscillator is given by

$$\overline{n} \;=\; \frac{1}{\exp[\,\frac{\hbar\omega_0}{k_B\,T}\,]\,-\,1} \tag{486}$$

If we were to include the zero point energy, then the total thermodynamic energy is given by

$$\begin{aligned}
U \;&=\; \sum_{i=1}^{N'}\,\hbar\,\omega_0\,\left(\,\overline{n}_i\,+\,\frac{1}{2}\,\right)\\
&=\; \sum_{i=1}^{N'}\,\frac{\hbar\,\omega_0}{2}\,\left(\,\frac{2}{\exp[\,\frac{\hbar\,\omega_0}{k_B\,T}\,]\,-\,1}\,+\,1\,\right)\\
&=\; \sum_{i=1}^{N'}\,\frac{\hbar\,\omega_0}{2}\,\left(\frac{\exp[\,\frac{\hbar\,\omega_0}{k_B\,T}\,]\,+\,1}{\exp[\,\frac{\hbar\omega_0}{k_B\,T}\,]\,-\,1}\right)\\
&=\; N'\,\frac{\hbar\,\omega_0}{2}\,\coth\left(\frac{\hbar\,\omega_0}{2\,k_B\,T}\right)
\end{aligned} \tag{487}$$

The specific heat of the Einstein model can be found from

$$C_V \;=\; \left(\frac{\partial U}{\partial T}\right)_V \tag{488}$$
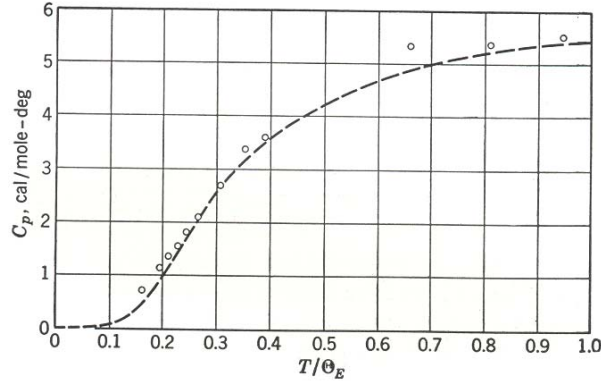
Figure 27: The specific heat of diamond compared with the results of the Einstein Model. The parameter $\Theta_E = \hbar \, \omega_0 / k_B$ is a characteristic temperature that has been assigned the value $\Theta_E = 1320$ K. [After A. Einstein, Ann. Physik **22**, 180-190 (1907).]

which yields

$$C_V \; = \; N' \, k_B \; \left(\frac{\hbar \, \omega_0}{k_B \, T} \right)^2 \; \frac{\exp[\; \frac{\hbar \, \omega_0}{k_B \, T} \;]}{(\; \exp[\; \frac{\hbar \, \omega_0}{k_B \, T} \;] \; - \; 1 \;)^2} \tag{489}$$

The specific heat tends to $N'k_B$ for temperatures $k_B \, T \; \gg \; \hbar \, \omega_0$, as is expected classically. However, at low temperatures, defined by $k_B \, T \; \ll \; \hbar \, \omega_0$, the specific heat falls to zero exponentially

$$C_V \; \approx \; N' \, k_B \; \left(\frac{\hbar \, \omega_0}{k_B \, T} \right)^2 \; \exp \left[\; - \; \frac{\hbar \, \omega_0}{k_B \, T} \; \right] \tag{490}$$

Therefore, the specific heat vanishes in the limit $T \to 0$ in accordance with Nernst's law. However, the specific heat of most materials deviate from the prediction of the Einstein model at low temperatures.

4.5 Vacancies in a Crystal

Consider a crystal composed of N identical atoms arranged in a periodic lattice. If an atom is on a proper atomic site, then it has an energy which we shall define to have a constant value denoted by $- \, \epsilon$. If an atom moves to an interstitial site, it has an energy of zero. This is because it may diffuse to the surface and escape from the crystal. Alternatively, the excitation energy required to unbind an atom from its site and, thereby create a vacancy is given by ϵ. We are considering the number of vacancies to be much smaller than the number of lattice sites, so that we can neglect the possibility that two vacancies sit on neighboring lattice sites, so we can neglect any effects of their

103
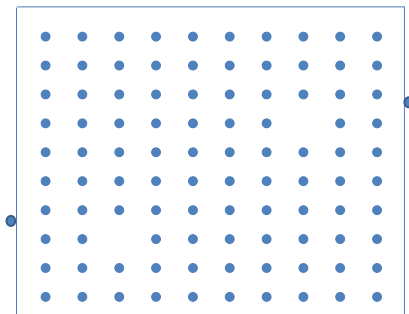
Figure 28: A schematic depiction of a crystalline solid composed of N atoms which contains N_v vacancies.

interactions. The number of possible vacancies n_i on the single lattice site i, is restricted to have values zero or one. That is, there either is a vacancy or there is not. Also, we should note that the total number of vacancies is not conserved.

Let us consider a lattice with N_v vacancies. This state has an energy which is greater than the energy of a perfect lattice by the amount $U = N_v \, \epsilon$. We should note that the vacancies are indistinguishable, since if we permute them, the resulting state is identical. The number of distinct ways of distributing N_v indistinguishable vacancies on N lattice sites is given by

$$N_\Gamma = \left(\frac{N!}{N_v! \, (\, N \, - \, N_v \,)!} \right) \tag{491}$$

This is just the number of ways of choosing the lattice sites for N_v distinguishable vacancies $N!/(N - N_v)!$, divided by the number of permutations of the vacancies $N_v!$. Dividing by $N_v!$ then just counts the vacancies as if they were indistinguishable.

In the Micro-Canonical Ensemble, the entropy is given by

$$S = k_B \, \ln N_\Gamma \tag{492}$$

which on using Stirling's approximation, yields

$$S \approx k_B \left[N \, \ln N \, - \, N_v \, \ln \, N_v \, - \, (\, N \, - \, N_v \,) \, \ln(\, N \, - \, N_v \,) \right] \tag{493}$$

which is a function of U since $U = N_v \, \epsilon$.

The energy can be expressed in terms of temperature, by using the relation

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_N \tag{494}$$

since the entropy is a function of energy. This yields

$$
\begin{aligned}
\frac{1}{T} &= \left(\frac{\partial S}{\partial N_v} \right) \left(\frac{\partial N_v}{\partial U} \right) \\
&= k_B \left[\ln(N - N_v) - \ln N_v \right] \left(\frac{\partial N_v}{\partial U} \right) \\
&= \frac{k_B}{\epsilon} \left[\ln(N - N_v) - \ln N_v \right] \\
&= \frac{k_B}{\epsilon} \ln \left(\frac{N - N_v}{N_v} \right)
\end{aligned}
\tag{495}
$$

After multiplying by ϵ/k_B and exponentiating, the expression can be inverted to give the number of vacancies

$$
N_v = \frac{N}{\exp[\frac{\epsilon}{k_B T}] + 1}
\tag{496}
$$

which shows that the average number of thermally excited vacancies on a site is given by

$$
\frac{N_v}{N} = \frac{1}{\exp[\frac{\epsilon}{k_B T}] + 1}
\tag{497}
$$

The thermodynamic energy U at a temperature T is given by the expression

$$
U = \frac{N \epsilon}{\exp[\frac{\epsilon}{k_B T}] + 1}
\tag{498}
$$

At low temperatures, $\epsilon \gg k_B T$, this reduces to zero exponentially.

$$
U = N \epsilon \exp\left[- \frac{\epsilon}{k_B T} \right]
\tag{499}
$$

At high temperatures (where the approximate model is not valid) half the lattice sites would host vacancies.

The specific heat due to the formation of vacancies is given by the expression

$$
C = N k_B \left(\frac{\epsilon}{2 k_B T} \right)^2 \mathrm{sech}^2 \left(\frac{\epsilon}{2 k_B T} \right)
\tag{500}
$$

which vanishes exponentially at low T as is characteristic of systems with excitation gaps in their excitation spectra. At high temperatures, the specific heat vanishes as the inverse square of T, characteristic of a system with an energy spectrum bounded from above. This form of the specific heat is known as a Schottky anomaly or Schottky peak. The above expression has been derived from the configurational entropy of the vacancies. In real materials, there will also be a vibrational entropy since vacancies will cause local phonon modes to form.
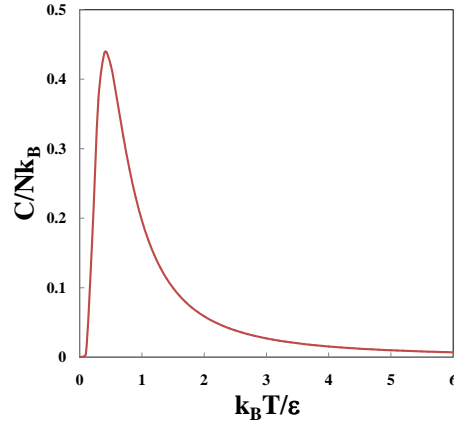
Figure 29: The Schottky specific heat versus temperature of a model of vacancies in a crystalline solid composed of N atoms.

5   The Canonical Ensemble

The Canonical Ensemble describes a closed system that is divided into two parts, each with a fixed number of particles and fixed volumes. However, the two subsystems can exchange energy with each other. One subsystem is the system which is the focus of our interest. The second subsystem is assumed to be much larger than the system of interest and is known as the environment. The properties of the environment will not be of direct interest and its main role will be to act as a thermal reservoir which absorbs or supplies energy to the system of interest. The distribution function for the subsystem of interest can be derived from the Micro-Canonical Probability Distribution Function for the total system.

5.1   The Boltzmann Distribution Function

The total energy of the complete system E_T is partitioned into the energy of our subsystem E and that of the thermal reservoir E_R

$$E_T \; = \; E_R \; + \; E \tag{501}$$

where the interaction energy between the system and the environment has been assumed to be negligible. The infinitesimal volume element of total phase space $d\Gamma_T$ is also assumed to be factorizable in terms of the products of the volume elements of the thermal reservoir $d\Gamma_R$ with the volume element of our subsystem $d\Gamma$. This assumes that every degree of freedom for the total system can be uniquely assigned either to the thermal reservoir or to the system of interest. Hence, we assume that

$$d\Gamma_T \; = \; d\Gamma_R \; d\Gamma \tag{502}$$

106

The probability dp_T of finding the total system in the volume element of phase space $d\Gamma$ is described by the constant Micro-Canonical Distribution Function ρ_{mc}

$$dp_T = \rho_{mc} \ d\Gamma_R \ d\Gamma \tag{503}$$

The probability dp for finding the subsystem in the phase space volume element $d\Gamma$ associated with the energy $H = E$ is found by integrating over all the phase space of the reservoir, consistent with the reservoir having the energy $H_R = E_T - E$. Hence,

$$dp = \rho_{mc} \ \Gamma_R(E_T - E) \ d\Gamma \tag{504}$$

where $\Gamma_R(E_T - E)$ is the volume of phase space accessible to the reservoir. However, the entropy of the reservoir is related to the volume of its accessible phase space via

$$\Gamma_R(E_T - E) = \Gamma_{R,0} \ \exp\left[\ S_R(E_T - E)/k_B \ \right] \tag{505}$$

where $\Gamma_{R,0}$ is the volume of phase space associated with one microscopic state of the reservoir. Hence, the probability density for the system of interest is given by

$$\left(\frac{dp}{d\Gamma}\right) = \rho_{mc} \ \Gamma_{R,0} \ \exp\left[\ S_R(E_T - E)/k_B \ \right] \tag{506}$$

where ρ_{mc} is just a constant. The energy of our subsystem E is much smaller than the energy of total system, E_T since the energy is extensive and the thermal reservoir is much larger than our subsystem. Hence, it is reasonable to assume that the reservoir's entropy can be Taylor expanded in powers of E and also that the second and higher-order terms can be neglected. That is

$$S_R(E_T - E) = S_R(E_T) - \left.\frac{\partial S_R(E)}{\partial E}\right|_{E_T} E + \ldots \tag{507}$$

but one recognizes that the derivative of the reservoir's entropy w.r.t. energy is defined as the inverse of the temperature T of the thermal reservoir

$$\left.\frac{\partial S_R(E)}{\partial E}\right|_{E_T} = \frac{1}{T} \tag{508}$$

Hence, the probability distribution function ρ_c for finding the system in some region of its phase space, as described in the Canonical Ensemble, depends on the energy of the system E via

$$\left(\frac{dp}{d\Gamma}\right) \Gamma_0 = \rho_{mc} \ \Gamma_{R,0} \ \Gamma_0 \ \exp\left[\ S_R(E_T)/k_B \ \right] \exp\left[\ - \ \beta \ E \ \right]$$

$$= \frac{1}{Z} \ \exp\left[\ - \ \beta \ E \ \right] \tag{509}$$

where Z is a constant and Γ_0 is the volume of phase space of the system which is used to define a single microscopic state. The factor $\rho_{mc} \, \Gamma_{R,0} \, \Gamma_0$ is a dimensionless constant which is independent of the specific point of the system's phase space, as is the first exponential factor. It is to be recalled that the region of phase space $d\Gamma$ under consideration corresponds to a specific value of the system's energy E, hence one can express the Canonical Probability Distribution Function as

$$\rho_c(\{p_i, q_i\}) \, \Gamma_0 \;=\; \frac{1}{Z} \, \exp\left[\, - \, \beta \, H(\{p_i, q_i\}) \, \right] \tag{510}$$

which depends on the point $\{p_i, q_i\}$ of the system's phase space only via the value of the system's Hamiltonian $H(\{p_i, q_i\})$. The dimensionless normalization constant Z is known as the Canonical Partition Function. The normalization condition

$$
\begin{aligned}
1 \;&=\; \int d\Gamma \, \left(\frac{dp}{d\Gamma} \right) \\
&=\; \int d\Gamma \, \rho_c(\{p_i, q_i\}) \\
&=\; \int \frac{d\Gamma}{\Gamma_0} \, \frac{1}{Z} \, \exp\left[\, - \, \beta \, H(\{p_i, q_i\}) \, \right]
\end{aligned} \tag{511}
$$

can be used to express the Canonical Partition Function Z as a weighted integral over the entire phase space of our system

$$Z \;=\; \int \frac{d\Gamma}{\Gamma_0} \, \exp\left[\, - \, \beta \, H(\{p_i, q_i\}) \, \right] \tag{512}$$

where the weighting function depends exponentially on the Hamiltonian H. Hence, in the Canonical Ensemble, the only property of the environment that actually appears in the distribution function is the temperature T. The distribution function $\rho_c(\{p_i, q_i\})$ is known as the Boltzmann Distribution Function. In the Canonical Ensemble, averages of quantities $A(\{p_i, q_i\})$ belonging solely to the system are evaluated as

$$\overline{A} \;=\; \frac{1}{Z} \, \int \frac{d\Gamma}{\Gamma_0} \, A(\{p_i, q_i\}) \, \exp\left[\, - \, \beta \, H(\{p_i.q_i\}) \, \right] \tag{513}$$

where the range of integration runs over all the phase space of our system, irrespective of the energy of the element of phase space. In the Canonical Distribution Function, the factor that depends exponentially on the Hamiltonian replaces the restriction used in the Micro-Canonical Ensemble where integration only runs over regions of phase space which corresponds to a fixed value of the energy E.

The Relation between the Canonical Partition Function and the Helmholtz Free-Energy

If the partition function is known, it can be used directly to yield the thermodynamic properties of the system. This follows once the partition function has been related to the Helmholtz Free-Energy $F(T, V, N)$ of our system via

$$Z = \exp\left[- \beta F \right] \tag{514}$$

This identification can be made by recalling that the partition function is related to the Micro-Canonical Distribution Function ρ_{mc} and the entropy of the thermal reservoir with energy E_T via

$$
\begin{aligned}
\frac{1}{Z} &= \Gamma_{T,0} \, \rho_{mc} \, \exp\left[S_R(E_T)/k_B \right] \\
&= \frac{\Gamma_{T,0}}{\Gamma_T(E_T)} \, \exp\left[S_R(E_T)/k_B \right]
\end{aligned} \tag{515}
$$

where the products of the volumes of phase space representing one microscopic state of the reservoir $\Gamma_{R,0}$ and one microscopic state of the subsystem Γ_0 has been assumed to be related to the volume of phase space $\Gamma_{T,0}$ representing one microscopic state of the total system by the equation $\Gamma_{R,0} \, \Gamma_0 = \Gamma_{T,0}$. The second line follows from the relation between the Micro-Canonical Distribution Function of the total system with the entropy evaluated at E_T. However, for the total system, one has

$$
\begin{aligned}
\frac{\Gamma_{T,0}}{\Gamma_T(E_T)} &= \exp\left[- S_T(E_T)/k_B \right] \\
&= \exp\left[- (S_R(E_T - U) + S(U))/k_B \right]
\end{aligned} \tag{516}
$$

where we have used the fact that the thermodynamic value of the entropy is extensive and the thermodynamic entropy of the subsystem is evaluated at the thermodynamic value of its energy U. (One expects from consideration of the maximization of the entropy that the thermodynamic energy U should be equal to the most probable value of the energy. However, as we shall show, the thermodynamic energy also coincides with the average energy $\overline{E}$.) On combining the above two expressions, one finds that

$$\frac{1}{Z} = \exp\left[- (S_R(E_T - U) - S_R(E_T) + S(U))/k_B \right] \tag{517}$$

which on Taylor expanding the first term in the exponent in powers of the relatively small average energy of the system U yields

$$\frac{1}{Z} = \exp\left[\beta U - S(U)/k_B \right] \tag{518}$$

where the higher-order terms in the expansion have been assumed negligible. Since the Helmholtz Free-Energy of the system is described as a Legendre transformation of the system's energy $U(S, V, N)$

$$F = U - T S(U) \tag{519}$$

then F is a function of the variables (T, V, N). Hence, since recognizes that the Canonical Partition Function is related to the Helmholtz Free-Energy F of the subsystem of interest via

$$Z = \exp\left[-\beta F \right] \tag{520}$$

it is also a function of the variable (T, V, N). For thermodynamic calculations, it is more convenient to recast the above relation into the form

$$F = -k_B T \ln Z \tag{521}$$

The Equality between the Average and the Thermodynamic Energies.

The above analysis is completed by identifying the thermodynamic energy U with the average energy $\overline{E}$. First we shall note that within the Canonical Ensemble, the average energy $\overline{E}$ is defined as

$$\overline{E} = \frac{1}{Z} \int \frac{d\Gamma}{\Gamma_0} H(\{p_i, q_i\}) \exp\left[-\beta H(\{p_i.q_i\}) \right] \tag{522}$$

which can be re-written as a logarithmic derivative of Z w.r.t. β, since the numerator of the integrand is recognized as the derivative of Z w.r.t. β

$$\begin{aligned}
\overline{E} &= -\frac{1}{Z} \frac{\partial}{\partial \beta} \left(\int \frac{d\Gamma}{\Gamma_0} \exp[-\beta H(\{p_i.q_i\})] \right) \\
&= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \\
&= -\frac{\partial \ln Z}{\partial \beta} \tag{523}
\end{aligned}$$

However, $\ln Z$ is also related to the value of the Helmholtz Free-Energy F, so one has

$$\begin{aligned}
\overline{E} &= \frac{\partial}{\partial \beta} (\beta F) \\
&= F + \beta \frac{\partial F}{\partial \beta} \\
&= F - T \frac{\partial F}{\partial T} \\
&= F + T S \tag{524}
\end{aligned}$$

where F is the Helmholtz Free-Energy of thermodynamics and the thermodynamic entropy S has been introduced via

$$S = -\left(\frac{\partial F}{\partial T} \right)_{V,N} \tag{525}$$

110

Hence, since the Free-Energy and the thermodynamic energy are related via $F = U - TS$, one finds that

$$\overline{E} = U \qquad (526)$$

This shows the thermodynamic energy U coincides with the average energy $\overline{E}$ when calculated in the Canonical Ensemble.

5.2 The Equipartition Theorem

5.3 The Ideal Gas

An ideal gas of N particles moving in a d-dimensional space is described by the Hamiltonian

$$H_N = \sum_{i=1}^{dN} \frac{p_i^2}{2\,m} \qquad (527)$$

and the particles are constrained to move within a hypercubic volume with linear dimensions L.

The partition function Z_N is given by

$$Z_N = \frac{1}{N!\,(\,2\,\pi\,\hbar\,)^{dN}} \prod_{i=1}^{dN} \left\{ \int_{-\infty}^{\infty} dp_i \int_0^L dq_i \right\} \exp\left[\,-\,\beta\,H_N\,\right] \qquad (528)$$

Since the Hamiltonian is the sum of independent terms, the expression for Z_N can be expressed as a product of dN terms

$$
\begin{aligned}
Z_N &= \frac{1}{N!\,(\,2\,\pi\,\hbar\,)^{dN}} \prod_{i=1}^{dN} \left\{ \int_{-\infty}^{\infty} dp_i\,\exp\left[\,-\,\frac{\beta\,p_i^2}{2\,m}\,\right] \int_0^L dq_i \right\} \\
&= \frac{1}{N!\,(\,2\,\pi\,\hbar\,)^{dN}} \prod_{i=1}^{dN} \left\{ \left(\frac{2\,\pi\,m}{\beta}\right)^{\frac{1}{2}} L \right\} \\
&= \frac{1}{N!\,(\,2\,\pi\,\hbar\,)^{dN}} \left\{ \left(\frac{2\,\pi\,m}{\beta}\right)^{\frac{1}{2}} L \right\}^{dN} \\
&= \frac{V^N}{N!} \left(\frac{m\,k_B\,T}{2\,\pi\,\hbar^2}\right)^{\frac{dN}{2}} \qquad (529)
\end{aligned}
$$

On introducing the thermal de Broglie wave length λ, via

$$\lambda = \frac{h}{(\,2\,\pi\,m\,k_B\,T\,)^{\frac{1}{2}}} \qquad (530)$$

one finds

$$Z_N = \frac{1}{N!} \left(\frac{V}{\lambda^d}\right)^N \qquad (531)$$

Thermodynamic quantities can be obtained by recalling that the Helmholtz Free-Energy is given by

$$F = -k_B \, T \, \ln Z_N \tag{532}$$

and by using Stirling's approximation

$$\ln N! = N \, \ln N - N \tag{533}$$

which yields

$$F = -N \, k_B \, T \, \ln \left(\frac{e \, V}{N \, \lambda^d} \right) \tag{534}$$

One can find all other thermodynamic functions from F. Thus, one can obtain the entropy from

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V,N} \tag{535}$$

as

$$S = N \, k_B \, \ln \left(\frac{V \, e}{N \, \lambda^d} \right) + \frac{d}{2} \, N \, k_B \tag{536}$$

which is the Sackur-Tetrode formula.

It is quite simple to show that the chemical potential μ is given by

$$\mu = k_B \, T \, \ln \frac{N \, \lambda^d}{V} \tag{537}$$

The condition under which the classical description is a expected to be a reasonable approximation is given by

$$\frac{V}{N \, \lambda^d} \gg 1 \tag{538}$$

Hence, we discover that the classical approximation is expected to be valid whenever

$$\exp \left[-\beta \, \mu \right] \gg 1 \tag{539}$$

5.4 The Entropy of Mixing

The entropy of mixing is associated with the factor of $N!$ needed to describe the microstates available to a gas of identical particles. This factor is required to make the entropy extensive so that on changing scale by a factor of s we have

$$S(sE, sV, sN) = s \, S(E, V, N) \tag{540}$$

The $N!$ is also needed to make the expression for the chemical potential intensive.

Consider a container partitioned off into two volumes V_1 and V_2. The containers hold N_1 and N_2 gas molecules respectively, and assume that the

molecules have the same masses and the gasses are kept at the same temperature (or average energy per particle). Then consider removing the partition. If the gas molecules in the two containers are indistinguishable, then in the Micro-Canonical Ensemble the entropy of the final state is given by

$$S_{indis} = k_B \ln \left[\frac{\Gamma_a}{(N_1 + N_2)! (2 \pi \hbar)^{d(N_1+N_2)}} \right] \tag{541}$$

which corresponds to dividing the enlarged accessible phase space Γ_a by a factor of $(N_1 + N_2)!$ to avoid over-counting the number of microstates. Equivalently, in the Canonical Ensemble the partition function Z is given by

$$Z_{indis} = \frac{\left(\frac{V_1 + V_2}{\lambda^d} \right)^{N_1+N_2}}{(N_1 + N_2)!} \tag{542}$$

However, if the molecules are distinguishable, the accessible phase space of the final state is the same as that for indistinguishable particles. However, it should be divided by $N_1! \, N_2!$ corresponding to the number of permutations of like molecules. In this case, the final state entropy is given by the expression

$$S_{dis} = k_B \ln \left[\frac{\Gamma_a}{(N_1! \, N_2!) (2 \pi \hbar)^{d(N_1+N_2)}} \right] \tag{543}$$

or equivalently

$$Z_{dis} = \frac{\left(\frac{V_1 + V_2}{\lambda^d} \right)^{N_1+N_2}}{(N_1! \, N_2!)} \tag{544}$$

Since in this case the final state consists of a mixture of distinct gasses, the entropy of the mixture must be larger than the entropy of the mixture of identical gasses. That is, it is expected that work would have to be expended to separate the distinct molecules. The entropy of mixing is defined as

$$S_{mix} = S_{indis} - S_{dis} \tag{545}$$

and since Γ_a are identical, it is found to be given by

$$
\begin{aligned}
S_{mix} &= k_B \ln(N_1 + N_2)! - k_B \ln(N_1! \, N_2!) \\
&= (N_1 + N_2) k_B \ln(N_1 + N_2) - N_1 \, k_B \ln N_1 - N_2 \, k_B \ln N_2 \\
&= - N_1 \, k_B \ln \frac{N_1}{N_1 + N_2} - N_2 \, k_B \ln \frac{N_2}{N_1 + N_2} \\
&= - (N_1 + N_2) k_B \left[\frac{N_1}{N_1 + N_2} \ln \frac{N_1}{N_1 + N_2} + \frac{N_2}{N_1 + N_2} \ln \frac{N_2}{N_1 + N_2} \right]
\end{aligned}
\tag{546}
$$

which has a form reminiscent of Shannon's entropy.

5.5 The Einstein Model of a Crystalline Solid

We shall revisit the Einstein model of a Crystalline Solid, in the Canonical Ensemble. The Hamiltonian of N' harmonic oscillators with frequency ω_0 takes the form

$$\hat{H} = \sum_{i=1}^{N'} \hbar\,\omega_0 \left(\hat{n}_i + \frac{1}{2} \right) \tag{547}$$

in the number operator representation. The set of possible eigenvalues of $\hat{n}_i$ are the integer values $0, 1, 2, 3, \ldots, \infty$. In this occupation number representation, the partition function $Z_{N'}$ is given by the trace

$$\begin{aligned}
Z_{N'} &= \ \text{Trace} \ \exp\left[-\,\beta\,\hat{H} \right] \\
&= \ \text{Trace} \ \exp\left[-\,\beta \sum_{i=1}^{N'} \hbar\,\omega_0\,(n_i + \frac{1}{2}) \right] \tag{548}
\end{aligned}$$

where the Trace is the sum over all the set of quantum numbers n_i for each oscillator. Hence, on recognizing that the resulting expression involves the sum of a geometric series, we have

$$\begin{aligned}
Z_{N'} &= \ \prod_{i=1}^{N'} \left\{ \sum_{n_i=0}^{\infty} \exp\left[-\,\beta\,\hbar\,\omega_0\,(n_i + \frac{1}{2}) \right] \right\} \\
&= \ \prod_{i=1}^{N'} \left\{ \frac{\exp[\,-\,\frac{\beta\,\hbar\,\omega_0}{2}\,]}{1 \ - \ \exp[\,-\,\beta\,\hbar\,\omega_0\,]} \right\} \\
&= \ \left[2 \ \sinh \frac{\beta\,\hbar\,\omega_0}{2} \right]^{-N'} \tag{549}
\end{aligned}$$

where each normal mode gives rise to an identical factor. The Free-Energy is given by

$$F = \ N'\,k_B\,T \ \ln\left(2 \ \sinh \frac{\beta\,\hbar\,\omega_0}{2} \right) \tag{550}$$

The entropy S is found from

$$S = -\left(\frac{\partial F}{\partial T} \right) \tag{551}$$

which yields

$$S = -\frac{F}{T} + N'\,\frac{\hbar\,\omega_0}{2\,T} \ \coth \frac{\beta\,\hbar\,\omega_0}{2} \tag{552}$$

However, since $F = U - T\,S$, one finds the internal energy U is given by

$$U = \ N'\,\frac{\hbar\,\omega_0}{2} \ \coth \frac{\beta\,\hbar\,\omega_0}{2} \tag{553}$$

This result is the same as that which was previously found using the Micro-Canonical Ensemble.

5.6 Vacancies in a Crystal

The Hamiltonian describing vacancies in a crystal can be described by

$$H_N = \sum_{i=1}^{N} \epsilon \, n_i \tag{554}$$

where the number of vacancies on site i is defined by n_i. The number of vacancies at site i only n_i has two possible values of unity or zero, since there either is a vacancy or there is not. The partition function Z_N is given by

$$Z_N = \text{Trace} \, \exp \left[\, - \, \beta \, H_N \, \right] \tag{555}$$

which is evaluated as

$$
\begin{aligned}
Z_N &= \text{Trace} \, \exp \left[\, - \, \beta \, \sum_{i=1}^{N} \epsilon \, n_i \, \right] \\
&= \text{Trace} \, \prod_{i=1}^{N} \exp \left[\, - \, \beta \, \epsilon \, n_i \, \right]
\end{aligned} \tag{556}
$$

The Trace runs over all the set of possible values of n_i for each oscillator. Thus

$$
\begin{aligned}
Z_N &= \text{Trace} \, \prod_{i=1}^{N} \left\{ \, \exp \left[\, - \, \beta \, \epsilon \, n_i \, \right] \, \right\} \\
&= \prod_{i=1}^{N} \left\{ \, \sum_{n_i=0}^{1} \exp \left[\, - \, \beta \, \epsilon \, n_i \, \right] \, \right\} \\
&= \prod_{i=1}^{N} \left\{ \, 1 \, + \, \exp[\, - \, \beta \, \epsilon \,] \, \right\} \\
&= \left(\, 1 \, + \, \exp[\, - \, \beta \, \epsilon \,] \, \right)^{N}
\end{aligned} \tag{557}
$$

which leads to the expression for the Free-Energy

$$F = - \, N \, k_B \, T \, \ln \left(\, 1 \, + \, \exp[\, - \, \beta \, \epsilon \,] \, \right) \tag{558}$$

Hence, from thermodynamics, one finds that the energy U is given by

$$
\begin{aligned}
U &= N \, \epsilon \, \frac{\exp[\, - \, \beta \, \epsilon \,]}{1 \, + \, \exp[\, - \, \beta \, \epsilon \,]} \\
&= N \, \epsilon \, \frac{1}{\exp[\, \beta \, \epsilon \,] \, + \, 1}
\end{aligned} \tag{559}
$$

which is identical to the expression that was found using the Micro-Canonical Ensemble.

5.7 Quantum Spins in a Magnetic Field

Consider a set of N quantum spins with magnitude S. We shall set $\hbar = 1$ for convenience. The spins interact with a magnetic field H^z through the Zeeman interaction

$$\hat{H}_{int} = - \sum_{i=1}^{n} g \, \mu_B \, H^z \, \hat{S}_i^z \tag{560}$$

where the $\hat{S}_i^z$ have eigenvalues m_i where $S \geq m \geq -S$.

Sice the spins do not interact with themselves, the partition function factorizes as

$$Z = \left\{ \sum_{m=-S}^{S} \exp\left[\beta \, g \, \mu_B \, H^z \, m \right] \right\}^N$$

$$= \left\{ \frac{\exp\left[\beta \, g \, \mu_B \, H^z \, (S + \frac{1}{2}) \right] - \exp\left[- \beta \, g \, \mu_B \, H^z \, (S + \frac{1}{2}) \right]}{\exp\left[\beta \, g \, \mu_B \, H^z \, \frac{1}{2} \right] - \exp\left[- \beta \, g \, \mu_B \, H^z \, \frac{1}{2} \right]} \right\}^N$$

$$= \left\{ \frac{\sinh\left[\beta \, g \, \mu_B \, H^z \, (S + \frac{1}{2}) \right]}{\sinh\left[\beta \, g \, \mu_B \, H^z \, \frac{1}{2} \right]} \right\}^N \tag{561}$$

The Free-Energy is given by

$$F = - k_B \, T \, \ln Z \tag{562}$$

which is evaluated as

$$F = - N \, k_B \, T \, \ln \sinh\left[\beta \, g \, \mu_B \, H^z \, (S + \frac{1}{2}) \right] + N \, k_B \, T \, \ln \sinh\left[\beta \, g \, \mu_B \, H^z \, \frac{1}{2} \right] \tag{563}$$

This can be expressed as

$$F = - N \, g \, \mu_B \, S \, H^z - N \, k_B \, T \, \ln\left(1 - \exp[- \beta \, g \, \mu_B \, H^z \, (2 \, S + 1)] \right)$$

$$+ N \, k_B \, T \, \ln\left(1 - \exp[- \beta \, g \, \mu_B \, H^z] \right) \tag{564}$$

Using thermodynamics, one can obtain the entropy. At high temperatures, the entropy saturates at

$$S \to k_B \, \ln(2 \, S + 1) \tag{565}$$

From the entropy and F, one can find the energy U which is given by

$$U = N \, g \, \mu_B \, H^z \left(\frac{1}{2} \coth\left[\beta \, g \, \mu_B \, H^z \, \frac{1}{2} \right] - (S + \frac{1}{2}) \coth\left[\beta \, g \, \mu_B \, H^z \, (S + \frac{1}{2}) \right] \right) \tag{566}$$

The internal energy saturates at

$$U = - N g \mu_B S H^z \tag{567}$$

in the low temperature limit, $T \rightarrow 0$ where the spins are completely aligned with the field. The internal energy vanishes in the high temperature limit, where the different spin orientations have equal probabilities.

Homework:

Determine the magnetization M^z defined by

$$M^z = - \left(\frac{\partial F}{\partial H^z} \right) \tag{568}$$

and the susceptibility $\chi^{z,z}$ which is defined as

$$\chi^{z,z} = \left(\frac{\partial M^z}{\partial H^z} \right) \tag{569}$$

Find the zero field limit of the susceptibility.

5.8 Interacting Ising Spin One-half Systems

Consider a one-dimensional array of spins interacting via the Ising Hamiltonian[14] given by

$$\hat{H} = - \sum_i J S_i^z S_{i+1}^z \tag{570}$$

The operator $\hat{S}^z$ has 2 possible eigenvalues which are $-\frac{\hbar}{2}, \frac{\hbar}{2}$. The interaction J couples the z-components of $N-1$ pairs of nearest neighbor spins. We shall assume that the interaction J has a positive value, so that the lowest energy configuration is ferromagnetic in which all the spins are aligned parallel to each other.

The partition function is given by

$$Z = \text{Trace} \exp \left[- \beta \hat{H} \right]$$

$$= \text{Trace} \prod_{i=1}^{N-1} \left\{ \exp \left[\beta J \hat{S}_i^z \hat{S}_{i+1}^z \right] \right\} \tag{571}$$

which is the product of factors arising from each sequential pair-wise interaction. The factors $\exp[\beta J \hat{S}_i^z \hat{S}_{i+1}^z]$ arising from an interaction can be re-written as

$$\exp \left[\beta J \hat{S}_i^z \hat{S}_{i+1}^z \right] = \left(\frac{1}{2} + \frac{2 S_i^z S_{i+1}^z}{\hbar^2} \right) \exp \left[+ \beta \frac{J \hbar^2}{4} \right]$$

---

[14] E. Ising, "Beitrag zur Theorie des Ferromagnetismus", Z. Phys. 31, 253-258, (1925).

$$+ \left(\frac{1}{2} - \frac{2 \, S_i^z \, S_{i+1}^z}{\hbar^2} \right) \exp \left[- \beta \, \frac{J \, \hbar^2}{4} \right]$$

$$= \left[\cosh \frac{\beta \, J \, \hbar^2}{4} + \frac{4}{\hbar^2} \, \hat{S}_i^z \, \hat{S}_{i+1}^z \, \sinh \frac{\beta \, J \, \hbar^2}{4} \right]$$

(572)

since they are to be evaluated on the space where $S_i^z \, S_{i+1}^Z = \pm \frac{\hbar^2}{4}$. Thus

$$Z = \text{Trace} \prod_{i=1}^{N-1} \left\{ \cosh \frac{\beta \, J \, \hbar^2}{4} + \frac{4}{\hbar^2} \, \hat{S}_i^z \, \hat{S}_{i+1}^z \, \sinh \frac{\beta \, J \, \hbar^2}{4} \right\}$$

(573)

The trace can be evaluated as a sum over all possible values of the spin eigenvalues

$$\text{Trace} \equiv \prod_{i=1}^{N} \left\{ \sum_{S_i^z = -\frac{\hbar}{2}}^{\frac{\hbar}{2}} \right\}$$

(574)

The trace runs over all the 2^N possible microstates of the system. The trace can be evaluated, by noting that the summand in the expression for the partition function only contains one factor which depends on $\hat{S}_1^z$

$$\left[\cosh \frac{\beta \, J \, \hbar^2}{4} + \frac{4}{\hbar^2} \, \hat{S}_1^z \, \hat{S}_2^z \, \sinh \frac{\beta \, J \, \hbar^2}{4} \right]$$

(575)

The terms odd in $\hat{S}_1^z$ cancel when taking the trace. Hence, the trace over S_1^z contributes a multiplicative factor of

$$2 \, \cosh \frac{\beta \, J \, \hbar^2}{4}$$

(576)

to the partition function, where the factor of two comes from the two spin directions. After the trace over S_1^z has been performed, only the factor

$$\left[\cosh \frac{\beta \, J \, \hbar^2}{4} + \frac{4}{\hbar^2} \, \hat{S}_2^z \, \hat{S}_3^z \, \sinh \frac{\beta \, J \, \hbar^2}{4} \right]$$

(577)

depends on S_2^z. On taking the trace over S_2^z, the last term in this factor vanishes and the trace contributes a second multiplicative factor of $\cosh \frac{\beta \, J \, \hbar^2}{4}$ to Z. Each of the $N-1$ interactions contributes a factor of

$$2 \, \cosh \frac{\beta \, J \, \hbar^2}{4}$$

(578)

to the partition function. The trace over the last spin, produces a multiplicative factor of 2 to Z. Hence, the partition function is given by

$$Z = 2 \left(2 \, \cosh \frac{\beta \, J \, \hbar^2}{4} \right)^{N-1}$$

(579)

118

The Free-Energy F is given by

$$F = - k_B \, T \, \ln Z \tag{580}$$

which is evaluated as

$$F = - N \, k_B \, T \, \ln 2 - (N - 1) \, k_B \, T \, \ln \cosh \frac{\beta \, J \, \hbar^2}{4} \tag{581}$$

The entropy S is found from

$$S = - \left(\frac{\partial F}{\partial T} \right) \tag{582}$$

which yields

$$S = N \, k_B \, \ln 2 + (N - 1) \, k_B \, \ln \cosh \frac{\beta \, J \, \hbar^2}{4} - (N - 1) \, k_B \, \frac{\beta \, J \, \hbar^2}{4} \, \tanh \frac{\beta \, J \, \hbar^2}{4} \tag{583}$$

The entropy is seen to reach the value $N \, k_B \, \ln 2$ appropriate to non-interacting spins in the limit $\beta \to 0$ and reaches the value of $k_B \ln 2$ in the limit $T \to 0$. The internal energy U is found from the relation

$$F = U - T \, S \tag{584}$$

as

$$U = - (N - 1) \, \frac{J \, \hbar^2}{4} \, \tanh \frac{\beta \, J \, \hbar^2}{4} \tag{585}$$

The energy vanishes in the limit $\beta \to 0$ and saturates to the minimal value of $- (N - 1) \, J \, \frac{\hbar^2}{4}$ appropriate to the $(N - 1)$ pair-wise interaction between completely aligned spins in the low temperature limit $T \to 0$. Hence, the ground state is two-fold degenerate and corresponds to minimizing the energy by the spins aligning so that either they are all up or they are all down. While at high temperatures, the system is dominated by the entropy which is maximized by randomizing the spin directions.

5.9   Density of States of Elementary Excitations

Consider normal modes of excitation that extend throughout the a hypercubic volume $V = L^d$. If the excitations satisfy an isotropic dispersion relation of the form

$$\hbar \, \omega = \hbar \, \omega(k) \tag{586}$$

where $\omega(k)$ is a monotonically increasing function of k, then this relation can be inverted to yield

$$k = k(\omega) \tag{587}$$

Since the normal modes are confined to the system, the normal modes wave functions must vanish on the walls of the system at $x_i = 0$ and $x_i = L$, for $i = 1, 2, \ldots, d$. If the wave functions have the form

$$\underline{\Psi}_\alpha(\underline{r}) \;=\; \frac{\underline{\epsilon}_\alpha(\underline{k})}{\sqrt{V}} \; \sin \underline{k} \cdot \underline{r} \tag{588}$$

for each polarization α, the allowed wave vectors satisfy the d-boundary conditions

$$k_i \, L \;=\; n_i \, \pi \tag{589}$$

for positive integer values of n_i. Thus, the allowed values of $\underline{k}$ are quantized and can be represented by a vector $\underline{n}$ in n-space

$$\underline{n} \;=\; \frac{L \, \underline{k}}{\pi} \tag{590}$$

which has positive integer components n_1 , n_2 , n_3 , $\ldots$, n_d. In n-space, each normal mode with polarization α is represented by a point with positive integer coordinates. Therefore, the normal modes per polarization form a lattice of points arranged on a hyper-cubic lattice with lattice spacing unity. In the segment composed of positive integers, there is one normal mode for each unit volume of the lattice.
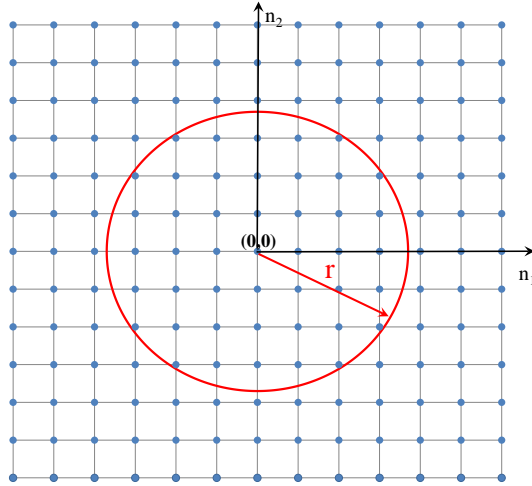
Figure 30: A pictorial representation of n-space in two-dimensions. Each state corresponds to a point (n_1, n_2) for positive integer values of the n's. In the positive quadrant, there is one state per unit area. The states with energy less than E are located in an area of the positive quadrant enclosed by a circular arc of radius r given by $r = Lk(E)/\pi$.

Due to the monotonic nature of $\omega(k)$, the number of excitations, per polarization, with energies less than ω, $N(\omega)$, is given by the number of lattice points $\underline{n}$ which satisfy the inequality

$$| \underline{n} | \leq \frac{L \ k(\omega)}{\pi} \tag{591}$$

or more explicitly

$$\sqrt{\sum_{i=1}^{d} n_i^2} \leq \frac{L \ k(\omega)}{\pi} \tag{592}$$

Since the segment of n-space with positive integers,

$$
\begin{aligned}
n_1 &\geq 0 \\
n_2 &\geq 0 \\
&\cdots \\
n_{d-1} &\geq 0 \\
n_d &\geq 0
\end{aligned}
\tag{593}
$$

is a fraction of $\frac{1}{2^d}$ of the entire volume of n-space, the number of normal modes with energy less than ω is given by $\frac{1}{2^d}$ of the volume enclosed by a radius

$$r = \frac{L \ k(\omega)}{\pi} \tag{594}$$

where we have recalled that there is one normal mode for each unit cell in n-space and that each cell has a volume 1^d. Hence, on dividing the expression for the volume of a hypersphere of radius r by 2^d, one finds

$$
\begin{aligned}
N(\omega) &= \frac{1}{2^d} \frac{S_d}{d} \left(\frac{L \ k(\omega)}{\pi} \right)^d \\
&= \frac{S_d}{d} \left(\frac{L \ k(\omega)}{2 \ \pi} \right)^d
\end{aligned}
\tag{595}
$$

This assumes that no points lie on the bounding surface of the hypersphere, or if they do their numbers are negligible. The surface area of a unit dimensional hypersphere is given by

$$S_d = \frac{2 \ \pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \tag{596}$$

so

$$N(\omega) = \frac{2 \ \pi^{\frac{d}{2}} \ V}{d \ \Gamma(\frac{d}{2})} \left(\frac{k(\omega)}{2 \ \pi} \right)^d \tag{597}$$

The number of excitations, per polarization, with energy less that $\hbar \omega$ can be expressed as an integral of the density of states $\rho(\omega)$, per polarization, defined as

$$\rho(\omega) \; = \; \sum_{\underline{k}} \delta\left(\omega \; - \; \omega(k) \right) \tag{598}$$

as

$$
\begin{aligned}
N(\omega) \; &= \; \int_{-\infty}^{\omega} d\omega' \; \rho(\omega') \\
&= \; \int_{-\infty}^{\omega} d\omega' \; \sum_{\underline{k}} \delta\left(\omega' \; - \; \omega(k) \right) \\
&= \; \sum_{\underline{k}} \Theta\left(\omega \; - \; \omega(k) \right)
\end{aligned}
\tag{599}
$$

where Θ is the Heaviside step function. The step function restricts the summation to the number of normal modes with frequencies less than ω, which are counted with weight unity. Thus, the density of states per polarization can be found from $N(\omega)$ by taking the derivative w.r.t. ω

$$\rho(\omega) \; = \; \frac{d}{d\omega} \; N(\omega) \tag{600}$$

Hence, we find that the density of states can be represented by

$$\rho(\omega) \; = \; \frac{V \; \pi^{\frac{d}{2}}}{\pi \; \Gamma(\frac{d}{2})} \left(\frac{k(\omega)}{2 \; \pi} \right)^{d-1} \frac{dk(\omega)}{d\omega} \tag{601}$$

The total density of states is given by the sum of the density of states for each polarization.

Homework:

Find the density of states for particles moving in a three-dimensional space obeying the dispersion relation

$$\omega \; = \; c \; k^n \qquad \text{for } n \; > \; 0 \tag{602}$$

5.10 The Debye Model of a Crystalline Solid

Consider a crystalline solid consisting of N atoms on a d-dimensional lattice. The model considers the lattice vibrations as being isotropic sound waves. The sound waves are characterized by their wave vectors $\underline{k}$ and by their polarizations. The vibrational modes consist of N longitudinal modes and $(d-1)N$ transverse modes. The dispersion relations for the modes will be denoted by $\omega_\alpha(\underline{k})$. The Hamiltonian is given by

$$\hat{H} \; = \; \sum_{\underline{k},\alpha} \hbar \; \omega_\alpha(\underline{k}) \; (\; n_{\underline{k},\alpha} \; + \; \frac{1}{2} \;) \tag{603}$$

where $n_{\underline{k},\alpha}$ is a integer quantum number.

The partition function Z is given by

$$
\begin{aligned}
Z \; &= \; \prod_{\underline{k},\alpha} \Big\{ \; \sum_{n_{\underline{k},\alpha}=0}^{\infty} \; \Big\} \; \exp\Big[\; - \; \beta \; \sum_{\underline{k},\alpha} \; \hbar \; \omega_\alpha(\underline{k}) \; (\; n_{\underline{k},\alpha} \; + \; \frac{1}{2} \;)\; \Big] \\
&= \; \prod_{\underline{k},\alpha} \Big\{ \; \sum_{n_{\underline{k},\alpha}=0}^{\infty} \; \exp\Big[\; - \; \beta \; \hbar \; \omega_\alpha(\underline{k}) \; (\; n_{\underline{k},\alpha} \; + \; \frac{1}{2} \;)\; \Big] \; \Big\} \\
&= \; \prod_{\underline{k},\alpha} \Big\{ \; \frac{\exp[\; -\; \frac{1}{2}\; \beta\; \hbar\; \omega_\alpha(\underline{k})\;]}{1 \; - \; \exp[\; - \; \beta\; \hbar\; \omega_\alpha(\underline{k})\;]} \; \Big\} \qquad (604)
\end{aligned}
$$

where we have performed the sum over a geometric series. The Free-Energy F is given by

$$
\begin{aligned}
F \; &= \; - \; k_B \; T \; \ln Z \\
&= \; k_B \; T \; \sum_{\underline{k},\alpha} \; \ln\Big[\; \exp[\; +\; \frac{1}{2}\; \beta\; \hbar\; \omega_\alpha(\underline{k})\;] \; - \; \exp[\; -\; \frac{1}{2}\; \beta\; \hbar\; \omega_\alpha(\underline{k})\;] \; \Big] \\
&= \; k_B \; T \; \int_{-\infty}^{\infty} \; d\omega \; \rho(\omega) \; \ln\Big[\; \exp[\; +\; \frac{1}{2}\; \beta\; \hbar\; \omega\;] \; - \; \exp[\; -\; \frac{1}{2}\; \beta\; \hbar\; \omega\;] \; \Big] \\
&= \; \int_{-\infty}^{\infty} \; d\omega \; \rho(\omega) \; \Big[\; \frac{\hbar\; \omega}{2} \; + \; k_B \; T \; \ln\Big(\; 1 \; - \; \exp[\; -\; \beta\; \hbar\; \omega\;] \; \Big) \; \Big] \qquad (605)
\end{aligned}
$$

where we have introduced the density of states $\rho(\omega)$ via

$$
\rho(\omega) \; = \; \sum_{\underline{k},\alpha} \; \delta(\; \omega \; - \; \omega_{\underline{k},\alpha}\;) \qquad (606)
$$

Since the density of states from the different polarizations is additive, one has

$$
\rho(\omega) \; = \; \frac{V \; S_d}{(\; 2\; \pi\;)^d} \; \Big(\; \frac{1}{c_L^d} \; + \; \frac{(d-1)}{c_T^d} \; \Big) \; \omega^{d-1} \qquad (607)
$$

where the dispersion relation for the longitudinal modes is given by $\omega \; = \; c_L \; k$ and the dispersion relation for the $(d-1)$ transverse modes is given by $\omega \; = \; c_T \; k$.

Since the lattice vibrations are only defined by the motion of point particles arranged on a lattice, there is an upper limit on the wave vectors $\underline{k}$ and, hence, a maximum frequency. The maximum frequency ω_D is determined from the condition that the total number of normal modes is dN. Thus,

$$
\int_0^{\omega_D} \; \rho(\omega) \; = \; d\; N \qquad (608)
$$

which yields

$$
\frac{V \; S_d}{d\; (\; 2\; \pi\;)^d} \; \Big(\; \frac{1}{c_L^d} \; + \; \frac{(d-1)}{c_T^d} \; \Big) \; \omega_D^d \; = \; d\; N \qquad (609)
$$

123
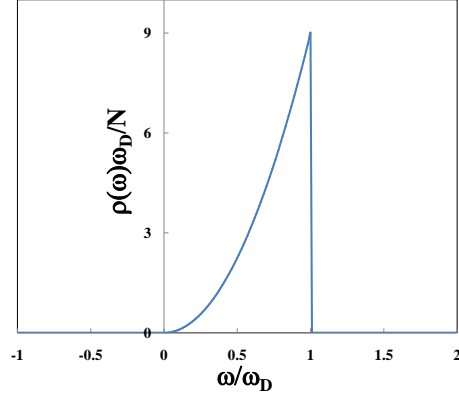
Figure 31: The density of states for the Debye model of a three-dimensional solid containing N atoms, with an upper cut-off frequency ω_D.

Hence, we may write the density of states as

$$\rho(\omega) \;=\; d^2 \; N \; \frac{\omega^{d-1}}{\omega_D^d} \tag{610}$$

for $\omega_D \;\geq\; \omega \;\geq\; 0$, ad is zero otherwise.

The Free-Energy is given in terms of the density of states as

$$F \;=\; \int_{-\infty}^{\infty} d\omega \; \rho(\omega) \left[\frac{\hbar \, \omega}{2} \;+\; k_B \, T \; \ln \left(1 \;-\; \exp[\, -\, \beta \, \hbar \, \omega \,] \right) \right] \tag{611}$$

and the entropy S is found from

$$S \;=\; - \left(\frac{\partial F}{\partial T} \right)_V \tag{612}$$

The internal energy is found from F and S as

$$U \;=\; \int_0^{\omega_D} d\omega \; \rho(\omega) \; \hbar \, \omega \left(\frac{1}{2} \;+\; \frac{1}{\exp[\, \beta \, \hbar \, \omega \,] \;-\; 1} \right) \tag{613}$$

The specific heat at constant volume is found from

$$C_V \;=\; \left(\frac{\partial E}{\partial T} \right)_V \tag{614}$$

which yields

$$C_V \;=\; k_B \; \beta^2 \int_0^{\omega_D} d\omega \; \rho(\omega) \; \hbar^2 \, \omega^2 \; \frac{\exp[\, \beta \, \hbar \, \omega \,]}{(\, \exp[\, \beta \, \hbar \, \omega \,] \;-\; 1 \,)^2} \tag{615}$$

124

or

$$C_V = \frac{\hbar^2}{k_B \, T^2} \int_0^{\omega_D} d\omega \; \rho(\omega) \; \omega^2 \; \frac{\exp[\; \beta \, \hbar \, \omega \;]}{(\; \exp[\; \beta \, \hbar \, \omega \;] \; - \; 1 \;)^2} \tag{616}$$

On substituting for the density of states, one finds

$$C_V = \frac{\hbar^2}{k_B \, T^2} \frac{d^2 \, N}{\omega_D^d} \int_0^{\omega_D} d\omega \; \omega^{d+1} \; \frac{\exp[\; \beta \, \hbar \, \omega \;]}{(\; \exp[\; \beta \, \hbar \, \omega \;] \; - \; 1 \;)^2} \tag{617}$$

The specific heat can be evaluated in two limits. In the high temperature limit, $k_B \, T \; \gg \; \hbar \, \omega_D$, then $k_B \, T \; \gg \; \hbar \, \omega$. In this limit, one can expand the integrand in powers of $\beta \, \hbar \, \omega$, which leads to

$$
\begin{aligned}
C_V &= k_B \; \frac{d^2 \, N}{\omega_D^d} \int_0^{\omega_D} d\omega \; \omega^{d-1} \\
&\approx d \, N \, k_B
\end{aligned}
\tag{618}
$$

Thus, at high temperatures, the Debye model of a solid reproduces Dulong and Petit's law.

At low temperatures, $k_B \, T \; \ll \; \hbar \, \omega_D$, one can introduce a dimensionless variable

$$x = \beta \, \hbar \, \omega \tag{619}$$

The maximum frequency ω_D corresponds to a maximum value x_D

$$x_D \; \gg \; 1 \tag{620}$$

The specific heat can be expressed as

$$C_V = k_B \; \frac{d^2 \, N}{x_D^d} \int_0^{x_D} dx \; x^{d+1} \; \frac{\exp[\; x \;]}{(\; \exp[\; x \;] \; - \; 1 \;)^2} \tag{621}$$

or on extending the upper limit to infinity

$$
\begin{aligned}
C_V &= k_B \; (\; k_B \, T \;)^d \; \frac{d^2 \, N}{(\; \hbar \, \omega_D \;)^d} \int_0^{\infty} dx \; x^{d+1} \; \frac{\exp[\; x \;]}{(\; \exp[\; x \;] \; - \; 1 \;)^2} \\
&= k_B \; (\; k_B \, T \;)^d \; \frac{d^2 \, N}{(\; \hbar \, \omega_D \;)^d} \; (\; d \; + \; 1 \;) \int_0^{\infty} dx \; \frac{x^d}{(\; \exp[\; x \;] \; - \; 1 \;)}
\end{aligned}
\tag{622}
$$

where the last line has been found through integration by parts. The integration produces a constant $\Gamma(\; d \; + \; 1 \;) \; \xi(d+1)$, so the end result is

$$C_V = N \, k_B \; \left(\frac{k_B \, T}{\hbar \, \omega_D}\right)^d \; d^2 \; \Gamma(\; d \; + \; 2 \;) \; \xi(d+1) \tag{623}$$

Hence, the specific heat at low temperatures is proportional to T^d, which is in accord with experimental observation.

For $d = 3$, one finds

$$C_V = N k_B \left(\frac{k_B T}{\hbar \omega_D} \right)^3 216 \frac{\pi^4}{90} \qquad (624)$$

5.11 Electromagnetic Cavities

Electromagnetic waves can flow through a vacuum, therefore an empty cavity can support electromagnetic modes. The normal modes can be represented by their wave vectors $\underline{k}$ and by their two possible polarizations. However, unlike the sound waves in a solid, there is no upper cut off for the wave vector of an electromagnetic wave. The Hamiltonian is given by

$$\hat{H} = \sum_{\underline{k},\alpha} \hbar \, \omega_\alpha(\underline{k}) \, (\, n_{\underline{k},\alpha} \, + \, \frac{1}{2} \,) \qquad (625)$$

where $n_{\underline{k},\alpha}$ is a quantum number (the number of photons) which has the allowed values of $0, 1, 2, 3, \ldots, \infty$.

The partition function Z is given by

$$
\begin{aligned}
Z \; &= \; \prod_{\underline{k},\alpha} \left\{ \sum_{n_{\underline{k},\alpha}=0}^{\infty} \right\} \exp\left[\, - \, \beta \sum_{\underline{k},\alpha} \hbar \, \omega_\alpha(\underline{k}) \, (\, n_{\underline{k},\alpha} \, + \, \frac{1}{2} \,) \right] \\
&= \; \prod_{\underline{k},\alpha} \left\{ \sum_{n_{\underline{k},\alpha}=0}^{\infty} \exp\left[\, - \, \beta \, \hbar \, \omega_\alpha(\underline{k}) \, (\, n_{\underline{k},\alpha} \, + \, \frac{1}{2} \,) \right] \right\} \\
&= \; \prod_{\underline{k},\alpha} \left\{ \frac{\exp[\, - \, \frac{1}{2} \, \beta \, \hbar \, \omega_\alpha(\underline{k}) \,]}{1 \, - \, \exp[\, - \, \beta \, \hbar \, \omega_\alpha(\underline{k}) \,]} \right\} \qquad (626)
\end{aligned}
$$

and the Free-Energy is given by

$$
\begin{aligned}
F \; &= \; - \, k_B \, T \, \ln Z \\
&= \; k_B \, T \sum_{\underline{k},\alpha} \ln\left[\, \exp[\, + \, \frac{1}{2} \, \beta \, \hbar \, \omega_\alpha(\underline{k}) \,] \, - \, \exp[\, - \, \frac{1}{2} \, \beta \, \hbar \, \omega_\alpha(\underline{k}) \,] \, \right] \\
&= \; k_B \, T \int_{-\infty}^{\infty} d\omega \, \rho(\omega) \, \ln\left[\, \exp[\, + \, \frac{1}{2} \, \beta \, \hbar \, \omega \,] \, - \, \exp[\, - \, \frac{1}{2} \, \beta \, \hbar \, \omega \,] \, \right] \\
&= \; \int_{-\infty}^{\infty} d\omega \, \rho(\omega) \left[\, \frac{\hbar \, \omega}{2} \, + \, k_B \, T \, \ln\left(1 \, - \, \exp[\, - \, \beta \, \hbar \, \omega \,] \right) \right] \qquad (627)
\end{aligned}
$$

where the density of states $\rho(\omega)$ is given by

$$\rho(\omega) = \sum_{\underline{k},\alpha} \delta(\, \omega \, - \, \omega_{\underline{k},\alpha} \,) \qquad (628)$$

The first term in the Free-Energy represents the (infinite) zero-point energy of the electromagnetic modes. It is divergent, because the electromagnetic cavity can support modes of arbitrarily high frequency. Divergences due to the presence of modes with arbitrarily large frequencies are known as ultra-violet divergences. Since only excitation energies are measured, the zero-point energy can usually be ignored. However, if the boundaries of the cavity are changed, there may be a measurable change in the zero-point energy of the cavity such as found in the Casimir effect[15]. That is, although it may be reasonable to speculate that the divergence in the zero-point energy may merely reflect our ignorance of the true physics at ultra-short distances, the zero-point energy cannot be dismissed since it does have some physical reality.

The density of states for the $(d - 1)$ transverse electromagnetic modes can be described by

$$
\rho(\omega) = (d - 1) \frac{S_d \, V \, \omega^{d-1}}{(2 \, \pi \,)^d \, c^d} \tag{629}
$$

Hence, the Free-Energy is given by the integral

$$
F = (d - 1) \frac{S_d \, V}{(2 \, \pi \,)^d \, c^d} \int_0^\infty d\omega \, \omega^{d-1} \left[\frac{\hbar \, \omega}{2} + k_B \, T \, \ln \left(1 - \exp[- \beta \, \hbar \, \omega \,] \right) \right] \tag{630}
$$

The internal energy is given by

$$
\begin{aligned}
U &= F + T \, S \\
&= F - T \left(\frac{\partial F}{\partial T} \right)
\end{aligned} \tag{631}
$$

which leads to

$$
U = (d - 1) \frac{S_d \, V}{(2 \, \pi \, c \,)^d} \int_0^\infty d\omega \, \omega^{d-1} \left[\frac{\hbar \, \omega}{2} + \frac{\hbar \, \omega}{\exp[\, \beta \, \hbar \, \omega \,] \, - \, 1} \right] \tag{632}
$$

The first term is divergent and represents the zero-point energy, The second term, ΔU, represents the energy of thermally excited photons. The second term can be evaluated by changing variable to x defined by

$$
x = \beta \, \hbar \, \omega \tag{633}
$$

Thus

$$
\Delta U = \frac{4 \, (\, d \, - \, 1 \,) \, \pi^{\frac{d+2}{2}}}{\Gamma(\frac{d}{2})} \, V \, \hbar \, c \left(\frac{k_B \, T}{2 \, \pi \, \hbar \, c} \right)^{d+1} \int_0^\infty dx \, \frac{x^d}{\exp[\, x \,] \, - \, 1} \tag{634}
$$

[15] H.B.G. Casimir, "On the attraction between two perfectly conducting plates", Proc. Kon. Nederland. Akad. Wetensch. B51, 793 (1948).

which leads ro

$$\Delta U = \frac{4 \ (\ d \ - \ 1 \ ) \ \pi^{\frac{d+2}{2}}}{\Gamma(\frac{d}{2})} \ V \ \hbar \ c \ \left(\frac{k_B \ T}{2 \ \pi \ \hbar \ c} \right)^{d+1} \Gamma(d+1) \ \xi(d+1) \qquad (635)$$

In three-dimensions, the thermal energy of the cavity is given by

$$\Delta U = \frac{\pi^2 \ k_B^4 \ T^4 \ V}{15 \ \hbar^3 \ c^3} \qquad (636)$$

since $3! \ \xi(4) = \frac{\pi^4}{15}$. This result is closely related to the Stefan-Boltzmann law which describes the energy flux radiating from an electromagnetic cavity held at temperature T.

The energy flux escaping through a wall of the cavity is given by the energy that passes through a unit area, per unit time. In a unit time, a photon travels a distance $c \ \cos\theta$ perpendicular to the wall, where θ denotes the angle subtended by the photon's velocity to the normal to the wall. Thus, the number of photons incident on a unit area of the wall, in a unit time, is given by the integral

$$\mathcal{F}_N = \int \frac{d\Omega}{S_d} \ c \ \cos\theta \int_0^\infty d\omega \ \frac{N_\omega}{V} \qquad (637)$$

where the integration over the angle θ runs from 0 to $\frac{\pi}{2}$ as the flux only includes light that is traveling towards the wall. In this expression $\int \ d\omega \ N_\omega/V$ is the photon density. Therefore, the energy flux is given by

$$\begin{aligned}
\mathcal{F}_E &= \int \frac{d\Omega}{S_d} \ c \ \cos\theta \int_0^\infty d\omega \ \hbar \ \omega \ \frac{N_\omega}{V} \\
&= \frac{S_{d-1}}{S_d} \int_0^{\frac{\pi}{2}} d\theta \ \sin^{d-2}\theta \ c \ \cos\theta \int_0^\infty d\omega \ \hbar \ \omega \ \frac{N_\omega}{V} \\
&= \frac{c \ S_{d-1}}{(\ d \ - \ 1 \ ) \ S_d} \int_0^\infty d\omega \ \hbar \ \omega \ \frac{N_\omega}{V} \qquad (638)
\end{aligned}$$

The density of photons in a frequency interval $d\omega$ is given by

$$\frac{N_\omega}{V} \ d\omega = \frac{(\ d \ - \ 1 \ ) \ S_d}{(\ 2 \ \pi \ c \ )^d} \ \frac{\omega^{d-1}}{\exp[\ \beta \ \hbar \ \omega \ ] \ - \ 1} \ d\omega \qquad (639)$$

Hence, the energy flux is given by

$$\begin{aligned}
\mathcal{F}_E &= \frac{c \ S_{d-1}}{(\ 2 \ \pi \ c \ )^d} \int_0^\infty d\omega \ \frac{\hbar \ \omega^d}{\exp[\ \beta \ \hbar \ \omega \ ] \ - \ 1} \\
&= 2 \ \pi \ \hbar \ c^2 \ S_{d-1} \ \Gamma(d+1) \ \xi(d+1) \ \left(\frac{k_B \ T}{2 \ \pi \ \hbar \ c} \right)^{d+1} \qquad (640)
\end{aligned}$$

For three dimensions, this reduces to

$$\mathcal{F}_E = \frac{\pi^2 \ k_B^4 \ T^4}{60 \ \hbar^3 \ c^2} \qquad (641)$$

which is the Stefan-Boltzmann law inferred from experiment[16] and then deduced theoretically[17]

$$\mathcal{F}_E \; = \; \sigma \; T^4 \tag{642}$$

where Stefan's constant is given by

$$\sigma \; = \; \frac{\pi^2 \; k_B^4}{60 \; \hbar^3 \; c^2} \tag{643}$$

Homework:

Show that the thermal energy, per unit volume, of electromagnetic radiation with frequency ω in the range $d\omega$ is given by

$$\frac{U_\omega}{V} \; d\omega \; = \; \frac{\hbar \; \omega^3}{\pi \; c^3} \left(\; \exp[\; \beta \; \hbar \; \omega \;] \; - \; 1 \; \right)^{-1} d\omega \tag{644}$$

The spectrum of emitted radiation from a perfect emitter is a universal function of temperature which was first devised by Planck[18]. Show that at high temperatures ($k_B \; T \; \gg \; \hbar \; \omega$) it reduces to the Rayleigh-Jeans Law[19]

$$\frac{U_\omega}{V} \; d\omega \; \approx \; \frac{\omega^2}{\pi \; c^3} \; k_B \; T \; d\omega \tag{645}$$

in which each mode has an energy $k_B \; T$.

Homework:

Since the Universe is expanding it is not in thermal equilibrium. The density of matter is so small that the microwave background can be considered to be decoupled from the matter. Therefore, the universe can be considered as an electromagnetic cavity filled with microwave radiation that obeys the Planck distribution with a temperature of approximately 2.725 K [20]. Because of the expansion of the universe, the photons are Doppler shifted. The Doppler shift

---

[16]J. Stefan, "Über die Beziehung zwischen der Wärmestrahlung und der Temperatur", Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften, Wien, Bd. 79, 391-428, (1879).

[17]L. Boltzmann, "Ableitung des Stefan'schen Gesetzes, betreffend die Abhängigkeit der Wärmestrahlung von der Temperatur aus der electromagnetischen Lichttheori", Ann. Physik, 22, 291-294 (1884).

[18]M. Planck, "Über das Gesetz der Energieverteilung im Normalspectrum" Annalen der Physik, 3, 553-563 (1901).

[19]Lord Rayleigh, "Remarks on the Complete Theory of Radiation", Phil. Mag. 49, 539-540 (1900).
J.H. Jeans, "On the Partition of Energy between Matter and Æther", Phil. Mag. 10, 91-98, (1905).

[20]D.J. Fixen, E.S. Cheng, J.M. Gales, J.C. Mather, R.A. Shafer and E.L. Wright, "The Cosmic Ray Background from the full COBE-FIRAS data set." The Astrophysical Journal, 473, 576-587 (1996).
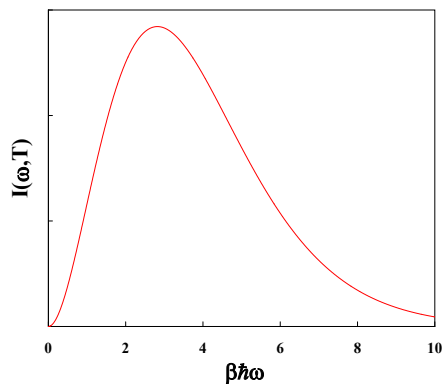
Figure 32: The spectrum of electromagnetic radiation emitted from a cavity held at a temperature T.

reduces the photons' energies and squeezes the frequencies ω to new frequencies ω'.

(i) If the linear-dimension of the universe is changed from L to L', how does the frequency of a photon change?

(ii) Show that the spectral density of thermally activated photons $N_\omega \, d\omega$ retains its form after the expansion but that the temperature is changed from T to T'. What is the final temperature T' in terms of the initial temperature T and L'/L?

(iii) Assume that expansion of the universe is adiabatic, and then use thermodynamics to determine the change in temperature of the background microwave radiation when the linear-dimension L of the universe is increased to L'.

Mathematical Appendix: Evaluation of the Riemann zeta function.

Consider the function

$$f(k) = \int_0^\infty dx \, \frac{\sin(kx)}{\exp[x] - 1} \tag{646}$$

where the integrand is seen to be finite at $x = 0$. Its Taylor expansion can be written as

$$
\begin{aligned}
f(k) &= \sum_{n=0}^\infty (-1)^n \, \frac{k^{2n+1}}{(2n+1)!} \int_0^\infty dx \, \frac{x^{2n+1}}{\exp[x] - 1} \\
&= \sum_{n=0}^\infty (-1)^n \, k^{2n+1} \, \xi(2n+2)
\end{aligned}
\tag{647}
$$

130

so $f(k)$ can be regarded as the generating function for the Riemann zeta functions. The value of the coefficient of k^{2n+1} is simply related to the Riemann zeta function $\xi(2n+2)$. As $\sin(kx)$ is the imaginary part of $\exp[ikx]$, one can re-write the integral as the imaginary part of a related complex function

$$f(k) = \lim_{\eta \to 0} \Im m \int_\eta^\infty dx \, \frac{\exp[ikx]}{\exp[x] - 1} \tag{648}$$

where, due to the finite value of η, the integration avoids the pole at the origin. The real function $f(k)$ can be evaluated by considering an integral of the related complex function over a contour C

$$\oint_C dz \, \frac{\exp[ikz]}{\exp[z] - 1} \tag{649}$$

The integrand has simple poles at the points $2\pi n i$ with integer n on the imaginary axis, and has residues $\exp[-2\pi k n]$ at these points. The contour
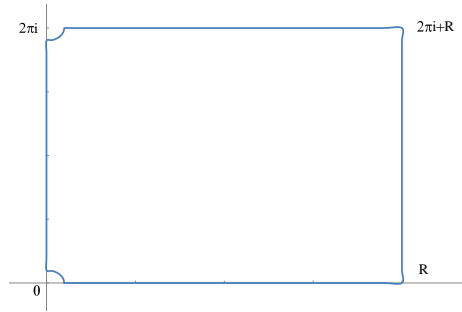
Figure 33: The contour of integration which avoids the poles on the imaginary axis.

C runs from η to R along the real axis, then to $R + 2\pi i$ along the imaginary axis, then the contour runs back parallel to the real axis to $\eta + 2\pi i$. Then, to avoid the pole at $2\pi i$, the contour follows a clockwise circle of radius η centered on $2\pi i$ from $\eta + 2\pi i$ to the point $-i\eta + 2\pi i$. The contour then runs down the imaginary axis from $-i\eta + 2\pi i$ to $i\eta$, and finally returns to η, by following a quarter circle of radius η centered on zero, thereby avoiding the pole at zero. The integral will be evaluated in the limit where $R \to \infty$ and $\eta \to 0$.

Since there are no poles enclosed by the integration contour, Cauchy's theorem yields

$$\oint_C dz \, \frac{\exp[ikz]}{\exp[z] - 1} = 0 \tag{650}$$

In the limit $R \to \infty$, the contribution from the segment from R to $R + 2\pi i$ tends to zero, as the integrand vanishes due to the denominator. The

integrations over the segments parallel to the real axis from η to R and from $R + 2\pi i$ to $\eta + 2\pi i$ can be combined to yield

$$\left[\, 1 \;-\; \exp[\,-\,2\,\pi\,k\,]\,\right] \int_\eta^\infty dx\; \frac{\exp[ikx]}{\exp[x]\,-\,1} \tag{651}$$

which has an imaginary part that is related to $f(k)$. The integrations over the quarter circles about the simple poles are both clockwise and are given by $-\,i\,\frac{\pi}{2}$ times the residues at the poles, and can be combined to yield

$$-\,i\,\frac{\pi}{2}\,\left[\, 1 \;+\; \exp[\,-\,2\,\pi\,k\,]\,\right] \tag{652}$$

The remaining contribution runs from $-\,i\,\eta\;+\;2\,\pi\,i$ to $i\,\eta$

$$\begin{aligned}
i\int_{-\eta+2\pi}^{\eta} dy\; \frac{\exp[\,-\,k\,y\,]}{\exp[\,i\,y\,]\,-\,1} \\
=\; -\int_{\eta}^{-\eta+2\pi} dy\; \frac{\exp[\,-\,(\,2\,k\,+\,i\,)\,\frac{y}{2}\,]}{2\,\sin\frac{y}{2}} \\
=\; -\frac{1}{2}\int_{\eta}^{-\eta+2\pi} dy\; \exp[\,-\,k\,y\,]\left(\,\cot\frac{y}{2}\,-\,i\,\right)
\end{aligned} \tag{653}$$

which, in the limit $\eta \to 0$, has an imaginary part that is given by

$$\frac{1}{2}\int_0^{2\pi} dy\; \exp[\,-\,k\,y\,] \;=\; \frac{1}{2}\left(\frac{\exp[\,-\,2\,\pi\,k\,]\,-\,1}{-\,k}\right) \tag{654}$$

If one now take the imaginary part of the entire integral of eqn(650) and take the limit $\eta \to 0$, one obtains

$$\left[\,1-\exp[\,-\,2\,\pi\,k\,]\,\right]f(k)-\frac{\pi}{2}\left[\,1+\exp[\,-\,2\,\pi\,k\,]\,\right]+\frac{1}{2\,k}\left[\,1-\exp[\,-\,2\,\pi\,k\,]\,\right]=0 \tag{655}$$

On rearranging the equation, $f(k)$ is found to be given by

$$f(k) \;=\; \frac{\pi}{2}\,\coth\pi\,k\;-\;\frac{1}{2\,k} \tag{656}$$

Since the series expansion of $\coth(\pi k)$ is given by

$$\coth\pi\,k \;=\; \frac{1}{\pi\,k}\;+\;\frac{\pi\,k}{3}\;-\;\frac{1}{45}\,(\,\pi\,k\,)^3\;+\;\frac{2}{945}\,(\,\pi\,k\,)^5\;+\;\ldots \tag{657}$$

then

$$f(k) \;=\; \frac{\pi^2}{6}\,k\;-\;\frac{\pi^4}{90}\,k^3\;+\;\frac{\pi^6}{945}\,k^5\;+\;\ldots \tag{658}$$

so the values of the Riemann zeta function are given by

$$\begin{aligned}
\xi(2) &\;=\; \frac{\pi^2}{6} \\
\xi(4) &\;=\; \frac{\pi^4}{90} \\
\xi(6) &\;=\; \frac{\pi^6}{945}
\end{aligned} \tag{659}$$

132

etc.

5.12 Energy Fluctuations

The Canonical Distribution Function can be used to calculate the entire distributions of most physical quantities of the system. However, for most applications it is sufficient to consider the average values $\overline{A}$ and the moments of the fluctuation $\overline{\Delta A^n}$ where the fluctuation is defined as

$$\Delta A \; = \; A \; - \; \overline{A} \tag{660}$$

The average fluctuation $\overline{\Delta A}$ is identically zero since

$$\overline{\Delta A} \; = \; \overline{A} \; - \; \overline{A} \; = \; 0 \tag{661}$$

However, the mean squared fluctuation is given by

$$\begin{aligned}
\overline{\Delta A^2} \; &= \; \overline{(\,\overline{A} \; - \; \overline{A}\,)^2} \\
&= \; \overline{A^2} \; - \; \overline{A}^2
\end{aligned} \tag{662}$$

which can be non-zero.

In the Canonical Ensemble, the energy is no longer fixed. However, the average value of the energy $\overline{E}$ was found to be equal to the thermodynamic value U, as

$$\begin{aligned}
\overline{E} \; &= \; \frac{1}{Z} \left(\int \frac{d\Gamma}{\Gamma_0} \, H \, \exp\left[\, - \, \beta \, H \, \right] \right) \\
&= \; - \, \frac{1}{Z} \, \frac{\partial}{\partial \beta} \left(\int \frac{d\Gamma}{\Gamma_0} \, \exp\left[\, - \, \beta \, H \, \right] \right) \\
&= \; - \, \frac{\partial \ln Z}{\partial \beta} \\
&= \; \frac{\partial}{\partial \beta} \left(\, \beta \, F \, \right) \\
&= \; U
\end{aligned} \tag{663}$$

The mean squared fluctuation in the energy can be expressed as

$$\begin{aligned}
\overline{\Delta E^2} \; &= \; \frac{1}{Z} \left(\int \frac{d\Gamma}{\Gamma_0} \, H^2 \, \exp\left[\, - \, \beta \, H \, \right] \right) \; - \; \frac{1}{Z^2} \left(\int \frac{d\Gamma}{\Gamma_0} \, H \, \exp\left[\, - \, \beta \, H \, \right] \right)^2 \\
&= \; \frac{1}{Z} \, \frac{\partial^2 Z}{\partial \beta^2} \; - \; \frac{1}{Z^2} \left(\frac{\partial Z}{\partial \beta} \right)^2 \\
&= \; \frac{\partial}{\partial \beta} \left(\frac{1}{Z} \, \frac{\partial Z}{\partial \beta} \right) \\
&= \; \frac{\partial^2 \ln Z}{\partial \beta^2}
\end{aligned} \tag{664}$$

It should be noted that the mean squared energy fluctuation can also be expressed as a derivative of the average energy w.r.t. β

$$\overline{\Delta E^2} \;=\; -\,\frac{\partial}{\partial \beta}\left(\,\overline{E}\,\right) \tag{665}$$

Therefore, on expressing this as a derivative w.r.t T, we find that

$$\overline{\Delta E^2} \;=\; k_B\,T^2\,\left(\frac{\partial \overline{E}}{\partial T}\right)_{V,N} \tag{666}$$

Hence, the mean squared energy fluctuation can be expressed in terms of the specific heat at constant volume

$$\overline{\Delta E^2} \;=\; k_B\,T^2\,C_{V,N} \tag{667}$$

From this we deduce that the relative magnitude of the energy fluctuations given by the dimensionless quantity

$$\frac{\overline{\Delta E^2}}{\overline{E}^2} \tag{668}$$

is of the order of $1/N$ since

$$\frac{\overline{\Delta E^2}}{\overline{E}^2} \;=\; \frac{k_B\,T^2\,C_{V,N}}{\overline{E}^2} \;\sim\; \frac{1}{N} \tag{669}$$

where the similarity follows since C_V is extensive and proportional to N as is $\overline{E}$. Thus, on taking the square root, one sees that the relative magnitude of the root mean squared (rms) fluctuation in the energy vanishes in the thermodynamic limit, since

$$\frac{\Delta E_{rms}}{\overline{E}} \;=\; \frac{\sqrt{k_B\,T^2\,C_{V,N}}}{\overline{E}} \;\sim\; \frac{1}{\sqrt{N}} \tag{670}$$

Therefore, the relative fluctuations of the energy are negligible in the thermodynamic limit. This suggest the reason why quantities calculated with the Canonical Distribution Function agree with those found with the Micro-Canonical Distribution Function.

The underlying reason for the Canonical and Micro-Canonical Ensemble yielding identical results is that the energy probability distribution function is sharply peaked. The probability density that the system has energy E is given by $P(E)$, where

$$
\begin{aligned}
P(E) \;&=\; \int d\Gamma\,\delta\!\left(E - H(\{p_i, q_i\})\right)\,\rho_c(\{p_i, q_i\}) \\
&=\; \frac{1}{Z}\int \frac{d\Gamma}{\Gamma_0}\,\delta\!\left(E - H(\{p_i, q_i\})\right)\,\exp\!\left[-\,\beta\,H(\{p_i, q_i\})\right] \\
&=\; \frac{1}{Z}\,\frac{\Gamma(E)}{\Gamma_0}\,\exp\!\left[-\,\beta\,E\right]
\end{aligned} \tag{671}
$$

which is the product of an exponentially decreasing function of energy $\exp[-\beta E]$ and $\Gamma(E)$ the volume of accessible phase space for a system in the Micro-Canonical Ensemble with energy E. We recognize that $\Gamma(E)$ is an extremely
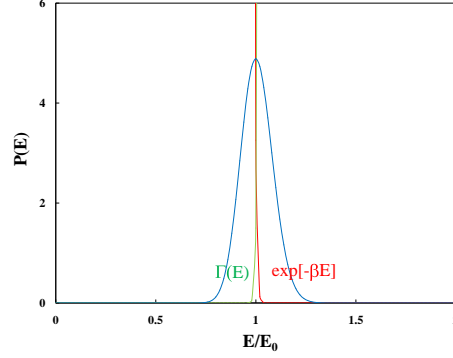
Figure 34: The energy distribution function $P(E)$ in the Canonical Ensemble (shown in blue). The distribution is sharply peaked since it is the product of an exponentially decreasing factor $\exp[-\beta E]$ (red) and a rapidly increasing function $\Gamma(E)$ (green).

rapidly increasing function of E, since for a typical system $\Gamma(E) \sim E^{\alpha N}$ where α ia a number of the order of unity. The most probable value of energy E_{max} can be determined from the condition for the maximum of the energy distribution function

$$\left. \frac{dP(E)}{dE} \right|_{E_{max}} = 0 \tag{672}$$

which leads to

$$\left. \frac{d}{dE} \left(\Gamma(E) \, \exp\left[-\beta \, E \right] \right) \right|_{E_{max}} = 0 \tag{673}$$

on representing $\Gamma(E)$ in terms of the entropy $S(E)$ one finds that the most probable value of the energy is given by the solution for E_{max} of the equation

$$\left. \frac{d}{dE} \left(\exp\left[-\beta \, E + S(E)/k_B \right] \right) \right|_{E_{max}} = 0 \tag{674}$$

or after some simplification

$$-\frac{1}{T} + \left. \frac{\partial S(E)}{\partial E} \right|_{E_{max}} = 0 \tag{675}$$

This equation is satisfied if the temperature T of the thermal reservoir is equal to the temperature of the system. This condition certainly holds true in thermal equilibrium, in which case

$$E_{max} = U \tag{676}$$

where U is the thermodynamic energy. Thus, we find that the most probable value of the energy E_{max} is equal to U, the thermodynamic energy. From our previous consideration, we infer that the most probable value of the energy is also equal to the average value of the energy $\overline{E}$,

$$E_{max} = \overline{E} \tag{677}$$

Thus, the probability distribution function is sharply peaked at the average energy.

The energy probability distribution function $P(E)$ can be approximated by a Gaussian expression, centered on U. This follows by Taylor expanding the exponent of $P(E)$ in powers of $(E - U)$

$$P(E) = \frac{1}{Z} \exp\left[-\beta F + \frac{1}{2 k_B} \left(\frac{d^2 S}{dE^2} \right) \bigg|_U (E - U)^2 + \ldots \right] \tag{678}$$

or on cancelling the factor of Z with $\exp[-\beta F]$, one finds

$$P(E) = \exp\left[\frac{1}{2 k_B} \left(\frac{d^2 S}{dE^2} \right) \bigg|_U (E - U)^2 + \ldots \right] \tag{679}$$

The energy width of the approximate Gaussian distribution is governed by the quantity

$$
\begin{aligned}
-\frac{1}{k_B} \left(\frac{d^2 S}{dE^2} \right) \bigg|_U &= -\frac{1}{k_B} \left(\frac{\partial \frac{1}{T}}{\partial U} \right)_{V,N} \\
&= \frac{1}{k_B T^2} \left(\frac{\partial T}{\partial U} \right)_{V,N} \\
&= \frac{1}{k_B T^2 C_V} \tag{680}
\end{aligned}
$$

Hence, the mean square fluctuations in the energy are given by

$$\overline{\Delta E^2} = k_B T^2 C_V \tag{681}$$

in accordance with our previous calculation. We note that in the thermodynamic limit $N \to \infty$, the energy distribution is so sharply peaked that the energy fluctuations usually can be ignored.

Homework:

Show that for an ideal gas, the energy fluctuations in the Canonical Ensemble are such that

$$\sqrt{\frac{\overline{\Delta E^2}}{\overline{E}}} = \sqrt{\frac{2}{3 N}} \tag{682}$$

136

Hence, the relative energy fluctuations vanish in the thermodynamic limit $N \to \infty$.

Homework:

Calculate $\overline{\Delta E^3}$ for a system in the Canonical Ensemble. Hence, show that

$$\overline{\Delta E^3} \; = \; k_B^2 \, T^2 \, \left[\, T^2 \, \frac{\partial C_v}{\partial T} \; + \; 2 \, T \, C_V \, \right] \tag{683}$$

and evaluate this for an ideal gas.

Homework:

Prove that

$$\overline{\Delta E^n} \; = \; (\, -\, 1 \,)^n \, \frac{\partial^n \ln Z}{\partial \beta^n} \tag{684}$$

generally holds for the Canonical Ensemble. Hence deduce that the higher-order moments of the energy fluctuations are all proportional to N.

5.13 The Boltzmann Distribution from Entropy Maximization

The general expression for entropy in terms of the probability distribution function $\rho_c(\{p_i, q_i\})$ is given by the integral over phase space

$$S \; = \; - \, k_B \, \int d\Gamma \, \left[\, \rho_c(\{p_i, q_i\}) \, \ln \rho_c(\{p_i, q_i\}) \Gamma_0 \, \right] \tag{685}$$

This is trivially true in the Micro-Canonical Ensemble and is also true in the Canonical Ensemble where

$$\Gamma_0 \, \rho_c(\{p_i, q_i\}) \; = \; \frac{1}{Z} \, \exp \left[\, - \, \beta \, H(\{p_i, q_i\}) \, \right] \tag{686}$$

This can be seen by substituting the equation for $\rho_c(\{p_i, q_i\})$ in the expression for S, which leads to

$$S \; = \; k_B \, \int d\Gamma \, \rho_c(\{p_i, q_i\}) \, \left[\, \beta \, H(\{p_i, q_i\}) \; + \; \ln Z \, \right] \tag{687}$$

However, we know that

$$\ln Z \; = \; - \, \beta \, F \tag{688}$$

and the distribution function is normalized

$$\int d\Gamma \, \rho_c(\{p_i, q_i\}) \; = \; 1 \tag{689}$$

so on multiplying by T we find

$$T \ S \ = \ \int d\Gamma \ \left[\ \rho_c(\{p_i, q_i\}) \ H(\{p_i, q_i\}) \ \right] \ - \ F \tag{690}$$

Since the average energy $\overline{E}$ is defined as

$$\overline{E} \ = \ \int d\Gamma \ \left[\ \rho_c(\{p_i, q_i\}) \ H(\{p_i, q_i\}) \ \right] \tag{691}$$

then we end up with an expression for the Helmholtz Free-Energy

$$F \ = \ \overline{E} \ - \ T \ S \tag{692}$$

Finally, since $\overline{E} = U$ (the thermodynamic energy), we have shown that

$$S \ = \ - \ k_B \ \int d\Gamma \ \left[\ \rho_c(\{p_i, q_i\}) \ \ln \rho_c(\{p_i, q_i\})\Gamma_0 \ \right] \tag{693}$$

which we shall regard as the fundamental form of S for any distribution function.

Derivation

Given the above form of S one can derive the Canonical Distribution Function as the distribution function ρ which maximizes the functional $S[\rho]$, subject to the requirements that the average energy is U and that the distribution function is normalized. That is ρ_c must maximize the functional $S[\rho]$ subject to the constraints that

$$\begin{aligned} 1 \ &= \ \int d\Gamma \ \rho(\{p_i, q_i\}) \\ U \ &= \ \int d\Gamma \ \rho(\{p_i, q_i\}) \ H(\{p_i, q_i\}) \end{aligned} \tag{694}$$

The maximization of S subject to the constraints is performed by using Lagrange's method of undetermined multipliers. In this method, one forms the functional $\Phi[\rho]$ defined by

$$\begin{aligned} \Phi[\rho] \ = \ &- \ k_B \ \int d\Gamma \ \left[\ \rho(\{p_i, q_i\}) \ \ln \rho(\{p_i, q_i\})\Gamma_0 \ \right] \\ &+ \gamma \ \left(\ 1 \ - \int d\Gamma \ \rho(\{p_i, q_i\}) \ \right) \\ &+ \delta \ \left(\ U \ - \int d\Gamma \ \rho(\{p_i, q_i\}) \ H(\{p_i, q_i\}) \ \right) \end{aligned} \tag{695}$$

where γ and δ are undetermined numbers. If ρ satisfies the two constraints then $\Phi[\rho]$ is equal to $S[\rho]$, and then maximizing Φ is equivalent to maximizing S.

If $\Phi[\rho]$ is to be maximized by $\rho_c(\{p_i, q_i\})$ then a small change in $\rho(\{p_i, q_i\})$ by $\delta\rho(\{p_i, q_i\})$ should not change Φ. That is, if we set

$$\rho(\{p_i, q_i\}) = \rho_c(\{p_i, q_i\}) + \lambda \ \delta\rho(\{p_i, q_i\}) \tag{696}$$

where $\delta\rho$ is an arbitrary deviation, then the first-order change in Φ, $\delta\Phi^{(1)}$ defined by

$$\Phi[\rho_c + \lambda\delta\rho] - \Phi[\rho_c] = \lambda \ \delta\Phi^{(1)} + O(\lambda)^2 \tag{697}$$

must vanish

$$\delta\Phi^{(1)} = 0 \tag{698}$$

If this condition was not satisfied, then a specific choice for the sign of λ would cause Φ to increase further. Thus, the requirement that $\Phi[\rho]$ is maximized by ρ_c, leads to the condition

$$\int d\Gamma \ \delta\rho(\{p_i, q_i\}) \left[- k_B \ \ln \rho_c(\{p_i, q_i\})\Gamma_0 - k_B - \gamma - \delta \ H(\{p_i, q_i\}) \right] = 0 \tag{699}$$

This integral must vanish for any choice of $\delta\rho$. This can be achieved by requiring that the quantity inside the square brackets vanishes at every point in phase space. That is

$$k_B \ \ln \rho_c(\{p_i, q_i\})\Gamma_0 = - k_B - \gamma - \delta \ H(\{p_i, q_i\}) \tag{700}$$

where γ and δ are undetermined constants. Hence, on exponentiating, we have

$$\rho_c(\{p_i, q_i\})\Gamma_0 = \exp\left[- 1 - \gamma/k_B \right] \exp\left[- \delta \ H(\{p_i, q_i\})/k_B \right] \tag{701}$$

The constants γ and δ are determined by ensuring that the two constraints are satisfied. The conditions are satisfied by requiring that

$$
\begin{aligned}
1 &= \exp\left[- 1 - \gamma/k_B \right] \int \frac{d\Gamma}{\Gamma_0} \exp\left[- \delta \ H(\{p_i, q_i\})/k_B \right] \\
U &= \exp\left[- 1 - \gamma/k_B \right] \int \frac{d\Gamma}{\Gamma_0} H(\{p_i, q_i\}) \exp\left[- \delta \ H(\{p_i, q_i\})/k_B \right]
\end{aligned}
\tag{702}
$$

which then has the effect that ρ_c maximized S. The two constraints suggests that one should rewrite the parameters as

$$\delta = \frac{1}{T} \tag{703}$$

and

$$Z = \exp\left[1 + \gamma/k_B \right] \tag{704}$$

139

In fact, if the form of ρ_c is substituted back into S and one constraint is used to express S in terms of U and the second constraint to produce a constant term (independent of U), then if one demands that

$$\left(\frac{\partial S}{\partial U} \right)_V \;=\; \frac{1}{T} \tag{705}$$

then one finds $\delta = \frac{1}{T}$. Thus, the distribution that maximizes $S[\rho]$ is recognized as being the Boltzmann Distribution Function

$$\rho_c(\{p_i, q_i\}) \Gamma_0 \;=\; \frac{1}{Z} \, \exp\left[\; - \; \beta \, H(\{p_i, q_i\}) \; \right] \tag{706}$$

In summary, we have shown that the Boltzmann Distribution Function maximizes $S[\rho]$ subject to the two constraints

$$1 \;=\; \int d\Gamma \; \rho_c(\{p_i, q_i\})$$
$$U \;=\; \int d\Gamma \; \rho_c(\{p_i, q_i\}) \; H(\{p_i, q_i\}) \tag{707}$$

5.14   The Gibbs Ensemble

The Gibbs Ensemble corresponds to the situation where a closed system is partitioned into two parts. A smaller part of which is our system of interest and the other part comprises its environment. The system and its environment are allowed to exchange energy, and also a partition that separates the system from its environment is free to move. Therefore, the volume of the subsystem can be interchanged with the environment. The total energy E_T is partitioned as

$$E_T \;=\; E \; + \; E_R \tag{708}$$

and the volume is partitioned as

$$V_T \;=\; V \; + \; V_R \tag{709}$$

The probability that the partition, considered by itself, will be found such that the volume of the system is in a range dV around V is assumed to be given by the ratio dV/V_T. The probability dp that the closed system (including the partition) is in the joint volume element $d\Gamma_T$ and dV,

$$dp \;=\; \frac{1}{V_T} \, \rho_{mc} \; d\Gamma_T \; dV \tag{710}$$

On factorizing the infinitesimal volume element of total phase space $d\Gamma_T$ into contributions from the reservoir $d\Gamma_R$ and the system $d\Gamma$, one has

$$dp \;=\; \frac{1}{V_T} \, \rho_{mc} \; d\Gamma_R \; d\Gamma \; dV \tag{711}$$

We are assuming that the phase space $d\Gamma$ is consistent with the position of the partition defining the volume V and also that the system has energy E. The probability dp that the system is in the volume element $d\Gamma$ irrespective of the microstates of the reservoir is obtained by integrating over all of the reservoir's accessible phase space, consistent with the energy $E_T - E$ and volume $V_T - V$. The result is

$$dp = \frac{1}{V_T} \rho_{mc} \Gamma_R(E_T - E, V_T - V) \, d\Gamma \, dV \qquad (712)$$

The Gibbs Probability Distribution Function ρ_G is defined via

$$dp = \left(\frac{dp}{d\Gamma dV} \right) d\Gamma \, dV$$
$$= \rho_G \, d\Gamma \, dV \qquad (713)$$

and is found as

$$\rho_G = \frac{1}{V_T} \rho_{mc} \Gamma_{R,0} \exp\left[S_R(E_T - E, V_T - V)/k_B \right] \qquad (714)$$

However, one also has

$$\rho_{mc} = \frac{1}{\Gamma_{T,0}} \exp\left[- S_T(E_T, V_T)/k_B \right]$$
$$= \frac{1}{\Gamma_{T,0}} \exp\left[- (S_R(E_T - U, V_T - \overline{V}) + S(U, \overline{V}))/k_B \right] \qquad (715)$$

The phase space volumes representing single microscopic states of the total system, reservoir and subsystem are assumed to satisfy the relation

$$\Gamma_{T,0} = \Gamma_{R,0} \Gamma_0 \qquad (716)$$

Hence, we can express the Gibbs Probability Distribution Function as

$$\rho_G \Gamma_0 = \frac{1}{V_T} \exp\left[(S_R(E_T - E, V_T - V) - S_R(E_T - U, V_T - \overline{V}) - S(U, \overline{V}))/k_B \right] \qquad (717)$$

The exponent can be expanded in powers of the energy and volume fluctuations

$$= S_R(E_T - U, V_T - \overline{V})$$
$$+ (U - E) \left(\frac{\partial S_R(E_R, V_T - \overline{V})}{\partial E_R} \right)\Bigg|_{E_T - U}$$
$$+ (\overline{V} - V) \left(\frac{\partial S_R(E_T - U, V_R)}{\partial V_R} \right)\Bigg|_{V_T - \overline{V}} + \ldots$$
$$= S_R(E_T - U, V_T - \overline{V}) + (U - E) \frac{1}{T} + (\overline{V} - V) \frac{P}{T} + \ldots \qquad (718)$$

On substituting this in the expression for ρ_G, one finds that

$$\rho_G \Gamma_0 = \frac{1}{V_T} \exp\left[\beta G \right] \exp\left[- \beta (E + P V) \right] \qquad (719)$$

141

where G is the Gibbs Free-Energy $G(T, P, N)$ of the system

$$G = U - T S + P \overline{V} \tag{720}$$

The only quantities in the Gibbs Distribution Function pertaining to the reservoir is its temperature and pressure. On introducing the Gibbs Partition Function Y via

$$Y = \exp\left[- \beta G \right] \tag{721}$$

one can express the Gibbs Probability Distribution Function as

$$\rho_G \, \Gamma_0 = \frac{1}{V_T} \, \exp\left[- \beta P V \right] \frac{1}{Y} \, \exp\left[- \beta H \right] \tag{722}$$

The Gibbs Partition Function

The normalization condition for the probability distribution function is given by

$$
\begin{aligned}
1 &= \int_0^{V_T} dV \int d\Gamma \, \rho_G \\
&= \int_0^{V_T} dV \, \frac{1}{V_T} \, \exp\left[- \beta P V \right] \frac{1}{Y} \int \frac{d\Gamma}{\Gamma_0} \, \exp\left[- \beta H \right] \\
&= \int_0^{V_T} dV \, \frac{1}{V_T} \, \exp\left[- \beta P V \right] \frac{1}{Y} \, Z(V)
\end{aligned}
\tag{723}
$$

where $Z(V)$ is the Canonical Partition Function. Hence, one finds that the Gibbs Partition Function Y is determined from

$$Y = \int_0^{V_T} dV \, \frac{1}{V_T} \, \exp\left[- \beta P V \right] Z(V) \tag{724}$$

which only involves quantities describing the system. Since the Canonical Partition Function is a function of the variable (T, V, N), the Gibbs partition function is a function of (T, P, N). Once Y has been determined from the above equation, thermodynamic quantities can be evaluated from the Gibbs Free-Energy $G(T, P, N)$ which is expressed in terms of Y as

$$G = - k_B T \, \ln Y \tag{725}$$

==

Example: The Ideal Gas

The Gibbs Partition Function for the ideal gas is given by

$$Y(T,P,N) = \int_0^{V_T} \frac{dV}{V_T} \exp\left[-\beta P V\right] \frac{1}{N!} \prod_{i=1}^{3N} \left\{ \int\int \frac{dp_i \, dq_i}{2 \pi \hbar} \right\} \exp\left[-\beta \sum_{i=1}^{3N} \frac{p_i^2}{2 m}\right]$$

(726)

which on integrating over the particle's coordinates and momenta becomes

$$Y(T,P,N) = \int_0^{V_T} \frac{dV}{V_T} \exp\left[-\beta P V\right] \frac{V^N}{N!} \left(\frac{\sqrt{2 \pi m k_B T}}{2 \pi \hbar}\right)^{3N}$$

(727)

The integral over V is easily evaluated by changing variable to

$$x = \beta P V$$

(728)

and yields

$$Y(T,P,N) = \frac{1}{\beta P V_T} \left(\beta P\right)^{-N} \left(\frac{\sqrt{2 \pi m k_B T}}{2 \pi \hbar}\right)^{3N}$$

(729)

where the factor of $N!$ has cancelled.

On ignoring the non-extensive contributions, the Gibbs Free-Energy is given by

$$\begin{aligned} G &= -k_B T \ln Y \\ &= N k_B T \left[\ln P - \frac{5}{2} \ln(k_B T) + \frac{3}{2} \ln\left(\frac{2 \pi \hbar^2}{m}\right) \right] \end{aligned}$$

(730)

Since the infinitesimal variation in G is given by

$$dG = -S \, dT + \overline{V} \, dP + \mu \, dN$$

(731)

One finds that the average volume is given by

$$\begin{aligned} \overline{V} &= \left(\frac{\partial G}{\partial P}\right)_T \\ &= \frac{N k_B T}{P} \end{aligned}$$

(732)

which is the ideal gas law.

The value of the enthalpy $\mathcal{H} = \overline{E} + P \overline{V}$ can be calculated directly from its average in the Gibbs Distribution

$$\begin{aligned} \mathcal{H} &= \overline{E} + P \overline{V} \\ &= \frac{1}{Y} \int_0^{V_T} \frac{dV}{V_T} \int \frac{d\Gamma}{\Gamma_0} \left(H(\{p_i, q_i\}) + P V \right) \exp\left[-\beta \left(H(\{p_i, q_i\}) + P V \right) \right] \end{aligned}$$

(733)

which can simply be expressed as a derivative w.r.t. β

$$
\begin{aligned}
\mathcal{H} &= -\frac{1}{Y}\left(\frac{\partial Y}{\partial \beta}\right)_P \\
&= -\frac{\partial \ln Y}{\partial \beta} \tag{734}
\end{aligned}
$$

Thus, we find that the enthalpy is given by

$$
\mathcal{H} = \frac{5}{2}\, N\, k_B\, T \tag{735}
$$

from which the specific heat at constant pressure C_P is found as

$$
C_P = \frac{5}{2}\, N\, k_B \tag{736}
$$

as is expected for the ideal gas.

==

5.15   A Flexible Polymer
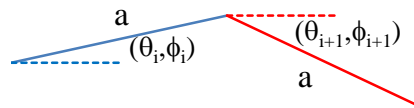
Figure 35: Two successive links of a polymer. The orientation of the i-th monomer is defined by the polar coordinates (θ_i, φ_i) in which the polar axis is defined by displacement vector of the whole polymer which runs from one end to the other.

Consider a polymer made of a large number N monomers of length a. The length of the polymer is variable since, although the monomers are joined end to end, the joints are assumed to be freely flexible. That is, the polymers are joined in way that allows free rotation at the ends. The length L of the polymer is defined by its end to end distance, and this definition of the length also defines a preferred (polar) axis which has the direction of the vector joining the ends. The orientational degrees of freedom of the i-th monomer is given by the polar coordinates (θ_i, φ_i). Hence, the length of the polymer is given by

$$
L = \sum_{i=1}^{N} a\, \cos\theta_i \tag{737}
$$

144

Although there is only one polymer, the fact that it is composed of a very large number of monomers allows one to consider it as being in the thermodynamic limit and to use statistical mechanics effectively.

The Hamiltonian is set equal to zero, since we are assuming that the monomers
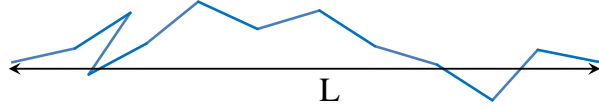
Figure 36: A polymer chain consisting of N links and length L. The polar axis is defined by the orientation of the displacement vector defining the length of the polymer.

are freely jointed and have negligible masses. The partition function $Z(L)$ for the polymer of length L is given by

$$Z(L) = \prod_{i=1}^{N} \left\{ \int_0^{2\pi} d\varphi_i \int_0^{\pi} \sin\theta_i \ d\theta_i \right\} \tag{738}$$

where the integrations are restricted by the constraint

$$L = \sum_{i=1}^{N} a \ \cos\theta_i \tag{739}$$

In this case Z coincides with Γ since $H = 0$.

If a tension $\mathcal{T}$ is applied to the polymer it is described by a Gibbs Distribution Function $Y(T, \mathcal{T})$, defined by

$$Y = \int_0^{Na} dL \ \exp\left[\beta \mathcal{T} L \right] \prod_{i=1}^{N} \left\{ \int_0^{2\pi} d\varphi_i \int_0^{\pi} \sin\theta_i \ d\theta_i \right\} \delta\left(L - \sum_{i=1}^{N} a \ \cos\theta_i \right)$$
$$\tag{740}$$

where the delta function has absorbed the factor of the normalization of the length probability density. The use of the Gibbs Distribution is justified since $-\mathcal{T}$ and L are analogous to P and V. On performing the integral over L, one obtains

$$Y = \prod_{i=1}^{N} \left\{ \int_0^{2\pi} d\varphi_i \int_0^{\pi} \sin\theta_i \ d\theta_i \right\} \exp\left[\beta \mathcal{T} a \sum_{i=1}^{N} \cos\theta_i \right] \tag{741}$$

which is no longer subject to a constraint. The constraint on the length has been replaced by a non-uniform weighting function. The Gibbs Partition Function, Y can be evaluated as

$$Y = \prod_{i=1}^{N} \left\{ \int_0^{2\pi} d\varphi_i \int_0^{\pi} \sin\theta_i \ d\theta_i \ \exp\left[\beta \mathcal{T} a \ \cos\theta_i \right] \right\}$$

$$
\begin{aligned}
&= \prod_{i=1}^{N} \left\{ \int_{0}^{2\pi} d\varphi_i \int_{-1}^{1} d\cos\theta_i \, \exp\left[\beta \, \mathcal{T} \, a \, \cos\theta_i \right] \right\} \\
&= \prod_{i=1}^{N} \left\{ 2\,\pi \int_{-1}^{1} d\cos\theta_i \, \exp\left[\beta \, \mathcal{T} \, a \, \cos\theta_i \right] \right\} \\
&= \prod_{i=1}^{N} \left\{ 2\,\pi \left(\frac{\exp[\,\beta\,\mathcal{T}\,a\,] \,-\, \exp[\,-\,\beta\,\mathcal{T}\,a\,]}{\beta\,\mathcal{T}\,a} \right) \right\} \\
&= \left[2\,\pi \left(\frac{\exp[\,\beta\,\mathcal{T}\,a\,] \,-\, \exp[\,-\,\beta\,\mathcal{T}\,a\,]}{\beta\,\mathcal{T}\,a} \right) \right]^{N} \\
&= \left[4\,\pi \left(\frac{\sinh[\,\beta\,\mathcal{T}\,a\,]}{\beta\,\mathcal{T}\,a} \right) \right]^{N} \tag{742}
\end{aligned}
$$

This has the form expected N for non-interacting monomers, where the only quantity which couples the monomers is the tension across the polymer. The reason for the form is recognized most clearly in the limit $\beta\,\mathcal{T}\,a \to 0$, where the monomers are expected to be distributed uniformly over the unit solid angle 4π. Since, in this limit,

$$
\lim_{\beta\mathcal{T}\to 0} \left(\frac{\sinh[\,\beta\,\mathcal{T}\,a\,]}{\beta\,\mathcal{T}\,a} \right) \to 1 \tag{743}
$$

it is seen that Y reduces to the products of the unit solid angles for each monomer.

The Gibbs Free-Energy G is given by

$$
Y = \exp\left[-\,\beta\,G \right] \tag{744}
$$

where in this case G is given by

$$
G = U \,-\, T\,S \,-\, \mathcal{T}\,\overline{L} \tag{745}
$$

in which $\overline{L}$ is the average length. Since the change in the (thermodynamic) internal energy is given by

$$
dU = T\,dS \,+\, \mathcal{T}\,d\overline{L} \tag{746}
$$

then G is a function of T and $\mathcal{T}$ since

$$
dG = -\,S\,dT \,-\, \overline{L}\,d\mathcal{T} \tag{747}
$$

Furthermore, the average length can be determined from thermodynamics, as a partial derivative of G with respect to the tension $\mathcal{T}$ at constant T

$$
\overline{L} = -\left(\frac{\partial G}{\partial \mathcal{T}} \right)_{T} \tag{748}
$$

or

$$\overline{L} \;=\; k_B\,T\,\left(\frac{\partial \ln Y}{\partial \mathcal{T}}\right)_T \tag{749}$$

which leads to the expression for the average length as a function of tension

$$\overline{L} \;=\; k_B\,T\,N\,\frac{\frac{\partial}{\partial \mathcal{T}}\left(\frac{\sinh[\,\beta\,\mathcal{T}\,a\,]}{\beta\,\mathcal{T}\,a}\right)}{\left(\frac{\sinh[\,\beta\,\mathcal{T}\,a\,]}{\beta\,\mathcal{T}\,a}\right)} \tag{750}$$

This is evaluated as

$$\overline{L} \;=\; N\,a\,\left(\,\coth[\,\beta\,\mathcal{T}\,a\,] \;-\; \frac{1}{\beta\,\mathcal{T}\,a}\,\right) \tag{751}$$

It is seen that the effect of tension is that of extending the length of the polymer,
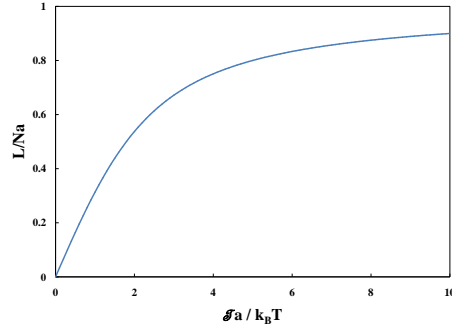
Figure 37: The length-tension relation for a polymer.

whereas the temperature acts to contract the polymer. For small values of the ratio of the tension to temperature, $\beta\,\mathcal{T}\,a \;\ll\; 1$, the relationship becomes approximately linear

$$\overline{L} \;\approx\; N\,a\,\left(\frac{1}{3}\,(\,\beta\,\mathcal{T}\,a\,) \;-\; \frac{1}{45}\,(\,\beta\,\mathcal{T}\,a\,)^3 \;+\; \dots\,\right) \tag{752}$$

Therefore, for small tensions the polymer acts like a rubber band. The polymer will contract as the temperature is increased. However, for large values of the tension $\beta\,\mathcal{T}\,a \;\gg\; 1$ the average length saturates at the value Na as the length is given by

$$L \;\approx\; N\,a\,\left(\,1 \;-\; \frac{1}{\beta\,\mathcal{T}\,a} \;+\; \dots\,\right) \tag{753}$$

with exponentially small corrections. This occurs since, for large ratios of the tension to the temperature, all the segments are aligned parallel.

147

The Helmholtz Free-Energy F is found from

$$F \ = \ G \ + \ \mathcal{T} \, \overline{L} \tag{754}$$

as

$$\begin{aligned} F \ &= \ - \, N \, k_B \, T \ \ln \left(4 \, \pi \ \frac{\sinh[\ \beta \, \mathcal{T} \, a\ ]}{\beta \, \mathcal{T} \, a} \right) \\ &\quad - \, N \, k_B \, T \ + \ N \, a \, \mathcal{T} \ \coth \beta \, \mathcal{T} \, a \end{aligned} \tag{755}$$

where $\mathcal{T}$ should be expressed in terms of $\overline{L}$.

6 The Grand-Canonical Ensemble

The Grand-Canonical Ensemble allows one to thermally average over a system which is able to exchange both energy and particles with its environment. The probability distribution function for the Grand-Canonical Ensemble can be derived by considering the closed system consisting of a subsystem and its larger environment. The probability distribution for the total closed system is calculated in the Micro-Canonical Ensemble. It is assumed that the number of particles and the energy can be uniquely partitioned into contributions from either the subsystem or its environment

$$
\begin{aligned}
E_R &= E + E_R \\
N_T &= N + N_R
\end{aligned}
\tag{756}
$$

Likewise, one assumes that for a given value of N, the infinitesimal phase space volume element can also be uniquely partitioned into factors representing the system and its environment. In this case, the probability dp for finding the entire closed system in a volume element $d\Gamma_T$ of its phase space

$$
dp = \rho_{mc} \, d\Gamma_T
\tag{757}
$$

can be expressed as

$$
dp = \rho_{mc} \, d\Gamma_N \, d\Gamma_{R,N_T-N}
\tag{758}
$$

where the system's phase space element is composed of the contributions from N particles and has energy E, while the reservoir has $N_T - N$ particles and has energy $E_T - E$. Since we are only interested in the probability distribution for finding the system in the a volume element $d\Gamma_N$ corresponding to having N particles and energy E and are not interested in the environment, we shall integrate over the phase space available to the environment. This results in the probability for finding the system in a state with N particles and in a volume of phase space $d\Gamma_N$ with energy E being given by

$$
dp = \rho_{mc} \, \Gamma_{R,N_T-N}(E_T - E) \, d\Gamma_N
\tag{759}
$$

where $\Gamma_{R,N_T-N}(E_T-E)$ is the volume of accessible phase space for the reservoir which has N particles and energy $E_T - E$. The Micro-Canonical Distribution Function ρ_{mc} can be expressed as

$$
\rho_{mc} = \frac{1}{\Gamma_{T,N_T}(E_T)}
\tag{760}
$$

where $\Gamma_{T,N_T}(E_T)$ is the entire volume of phase space accessible to the closed system with energy E_T. Since $\Gamma_{T,N_T}(E_T)$ can be expressed in terms of the total entropy of the closed system $S_T(E_T)$, the Micro-Canonical Distribution Function can be expressed as

$$
\rho_{mc} = \frac{1}{\Gamma_{N_T,0}} \, \exp\left[- \, S_T(E_T, N_T)/k_B \right]
\tag{761}
$$

where $\Gamma_{N_T,0}$ is the volume of phase space which represents one microscopic state of the system with N_T particles. The volume of accessible phase space for the reservoir can also be written in terms of its entropy

$$\Gamma_{R,N_T-N}(E_T - E) \ = \ \Gamma_{N_R,0} \ \exp\left[\ S_R(E_T - E, N_T - N)/k_B \ \right] \qquad (762)$$

where $\Gamma_{N_R,0}$ is the volume of phase space which represents a single microscopic state of the reservoir which contains N_R particles. Hence, the probability dp for finding the N particle system in an infinitesimal volume of phase space $d\Gamma_N$ with energy E is given by

$$
\begin{aligned}
dp_{N,E} \ &= \ \left(\frac{dp}{d\Gamma_N}\right) d\Gamma_N \\
&= \ \exp\left[\ (\ S_R(E_T - E, N_T - N) \ - \ S_T(E_T, N_T) \ )/k_B \ \right] \frac{d\Gamma_N}{\Gamma_{N,0}}
\end{aligned}
$$
$$(763)$$

where we have assumed that

$$\Gamma_{N_T,0} \ = \ \Gamma_{N_R,0} \ \Gamma_{N,0} \qquad (764)$$

Since the environment has been assumed to be much larger than the system both E and N are small compared to E_T and N_T. Therefore, it is reasonable to assume that the entropy of the reservoir can be Taylor expanded in powers of the fluctuations of E from the thermodynamic value U and the fluctuations of N from its thermodynamic value $\overline{N}$.

$$
\begin{aligned}
S_R(E_T - E, N_T - N) \ = \ & S_R(E_T - U, N_T - \overline{N}) \\
& + (U - E) \left(\frac{\partial S_R(E_R, N_T - \overline{N})}{\partial E_R}\right)\Bigg|_{E_T - U} \\
& + (\overline{N} - N) \left(\frac{\partial S_R(E_T - U, N_R)}{\partial N_R}\right)\Bigg|_{N_T - \overline{N}} + \ \dots
\end{aligned}
$$
$$(765)$$

On using the definitions of the thermal reservoir's temperature

$$\left(\frac{\partial S_R(U_R, \overline{N}_R)}{\partial U_R}\right)_{\overline{N}_R} \ = \ \frac{1}{T} \qquad (766)$$

and its chemical potential

$$\left(\frac{\partial S_R(U_R, \overline{N}_R)}{\partial \overline{N}_R}\right)_{U_R} \ = \ - \frac{\mu}{T} \qquad (767)$$

the expansion becomes

$$
\begin{aligned}
S_R(E_T - E, N_T - N) \ = \ & S_R(E_T - U, N_T - \overline{N}) \\
& + \frac{(U - E)}{T} \ - \ \frac{\mu \ (\overline{N} - N)}{T} \ + \ \dots \quad (768)
\end{aligned}
$$

where μ and T are the chemical potential and temperature of the reservoir. The total entropy $S_T(E_T, N_T)$ is extensive and can be decomposed as

$$S_T(E_T, N_T) \; = \; S_R(E_T - U, N_T - \overline{N}) \; + \; S(U, \overline{N}) \tag{769}$$

Thus, the Grand-Canonical Distribution Function can be written as

$$\left(\frac{dp}{d\Gamma}\right)_{N,E} \Gamma_{N,0} \; = \; \exp\left[\, -\beta\,(\,E - \mu\,N\,)\,\right] \exp\left[\,\beta\,(\,U - \mu\,\overline{N}\,) - S(U, \overline{N})/k_B\,\right] \tag{770}$$

or

$$\left(\frac{dp}{d\Gamma}\right)_{N,E} \Gamma_{N,0} \; = \; \exp\left[\, -\,\beta\,(\,E\,-\,\mu\,N\,)\,\right] \exp\left[\,\beta\,\Omega\,\right] \tag{771}$$

where Ω is the Grand-Canonical Potential $\Omega(T, V, \mu)$

$$\Omega \; = \; U \; - \; T\,S \; - \; \mu\,\overline{N} \tag{772}$$

describing the thermodynamics of the system. Once again, we note that the quantities E, N and Ω in the probability distribution function are properties of the system and that the only quantities which describe the environment are the temperature T and the chemical potential μ. The Grand-Canonical Partition function Ξ is defined by

$$\Xi \; = \; \exp\left[\, -\,\beta\,\Omega\,\right] \tag{773}$$

so the Grand-Canonical Distribution Function can be written as

$$\left(\frac{dp}{d\Gamma}\right)_N \Gamma_{N,0} \; = \; \frac{1}{\Xi}\,\exp\left[\, -\,\beta\,(\,H_N\,-\,\mu\,N\,)\,\right] \tag{774}$$

where H_N is the Hamiltonian for the N particle system. The exponential factor containing the Hamiltonian automatically provides different weights for the regions of N particle phase space. The quantity

$$dp_N \; = \; \left(\frac{dp}{d\Gamma}\right)_N d\Gamma_N \tag{775}$$

is the probability for finding the system to have N particles and be in the volume of phase space $d\Gamma_N$. Hence, the Grand-Canonical Probability Distribution Function can be used in determining the average of any function defined on the N particle phase space A_N via

$$\overline{A} \; = \; \sum_{N=0}^{\infty} \int d\Gamma_N \left(\frac{dp}{d\Gamma}\right)_N A_N \tag{776}$$

or, more explicitly

$$\overline{A} \; = \; \frac{1}{\Xi}\,\sum_{N=0}^{\infty} \int \frac{d\Gamma_N}{\Gamma_{N,0}}\,\exp\left[\, -\,\beta\,(\,H_N\,-\,\mu\,N\,)\,\right] A_N \tag{777}$$

151

and involves an integration over the N particle phase space and a summation over all possible particle numbers.

The Grand-Canonical Partition Function

The quantity

$$dp_N = \left(\frac{dp}{d\Gamma} \right)_N d\Gamma_N \qquad (778)$$

is the probability for finding the system to have N particles and also be in the volume of phase space $d\Gamma_N$. The probability p_N for finding the system as having N particles anywhere in its phase space is found by integrating over all $d\Gamma_N$

$$p_N = \frac{1}{\Xi} \int \frac{d\Gamma_N}{\Gamma_{N,0}} \exp \left[- \beta \left(H_N - \mu N \right) \right] \qquad (779)$$

Since the probability p_N must be normalized, one requires that

$$\sum_{N=0}^{\infty} p_N = 1 \qquad (780)$$

since a measurement of the number of particles in the system will give a result which is contained in the set $0, 1, 2, \ldots, \infty$. This normalization condition determines Ξ as being given by

$$
\begin{aligned}
\Xi &= \sum_{N=0}^{\infty} \int \frac{d\Gamma_N}{\Gamma_{N,0}} \exp \left[- \beta \left(H_N - \mu N \right) \right] \\
&= \sum_{N=0}^{\infty} \exp \left[\beta \mu N \right] \int \frac{d\Gamma_N}{\Gamma_{N,0}} \exp \left[- \beta H_N \right] \\
&= \sum_{N=0}^{\infty} \exp \left[\beta \mu N \right] Z_N \qquad (781)
\end{aligned}
$$

which relates the Grand-Canonical Partition Function Ξ to a sum involving the Canonical Partition Functions for the N particle systems Z_N.

Thermodynamics Averages and the Grand-Canonical Ensemble

The above normalization condition can be used to evaluate Ξ and, hence, the Grand-Canonical Potential Ω. One can then evaluate thermodynamic averages directly from $\Omega(T, V, \mu)$, via

$$\Omega = - k_B T \ln \Xi \qquad (782)$$

Thus for example, knowing Ω one can find the average number of particles $\overline{N}$ via the thermodynamic relation

$$\overline{N} = -\left(\frac{\partial \Omega}{\partial \mu}\right)_T \tag{783}$$

which can be expressed as

$$
\begin{aligned}
\overline{N} &= k_B \, T \left(\frac{\partial \ln \Xi}{\partial \mu}\right)_T \\
&= k_B \, T \, \frac{1}{\Xi} \left(\frac{\partial \Xi}{\partial \mu}\right)_T \\
&= k_B \, T \, \frac{1}{\Xi} \sum_{N=0}^{\infty} \left(\frac{\partial \exp[\beta \mu N] \, Z_N}{\partial \mu}\right)_T \\
&= \frac{1}{\Xi} \sum_{N=0}^{\infty} N \, \exp\left[\beta \, \mu \, N\right] Z_N
\end{aligned}
\tag{784}
$$

Hence, p_N defined by

$$p_N = \frac{1}{\Xi} \, \exp\left[\beta \, \mu \, N\right] Z_N \tag{785}$$

appears to be the probability for the system to have N particles, as

$$\overline{N} = \sum_{N=0}^{\infty} N \, p_N \tag{786}$$

Since Z_N is given by

$$Z_N = \int \frac{d\Gamma_N}{\Gamma_{N,0}} \, \exp\left[-\beta \, H_N\right] \tag{787}$$

on substituting in the expression for p_N and combining the exponentials, we can express p_N as

$$p_N = \frac{1}{\Xi} \int \frac{d\Gamma_N}{\Gamma_{N,0}} \, \exp\left[-\beta \, (\, H_N - \mu \, N \,)\right] \tag{788}$$

which is in agreement with our previous identification.

More generally, given the average of a quantity $\overline{A}$ defined by

$$\overline{A} = \frac{1}{\Xi} \sum_{N=0}^{\infty} \int \frac{d\Gamma_N}{\Gamma_{N,0}} \, \exp\left[-\beta \, (\, H_N - \mu \, N \,)\right] A_N \tag{789}$$

and on defining the thermodynamic average of A in the Canonical Ensemble with N particles as

$$\overline{A}_N = \frac{1}{Z_N} \int \frac{d\Gamma_N}{\Gamma_{N,0}} \, \exp\left[-\beta \, H_N\right] A_N \tag{790}$$

one finds that

$$
\begin{aligned}
\overline{A} &= \frac{1}{\Xi} \sum_{N=0}^{\infty} \exp\left[\beta \mu N\right] Z_N \, \overline{A}_N \\
&= \sum_{N=0}^{\infty} p_N \, \overline{A}_N
\end{aligned}
\tag{791}
$$

as is expected.

Homework:

Derive the probability distribution function $\rho_{gc}(N, \{p_i, q_i\}_N)$ for the Grand-Canonical Ensemble by maximizing the entropy subject to the three constraints

$$
\begin{aligned}
\sum_{N=0}^{\infty} \int d\Gamma_N \, \rho_{gc} &= 1 \\
\sum_{N=0}^{\infty} \int d\Gamma_N \, H_N \, \rho_{gc} &= U \\
\sum_{N=0}^{\infty} N \int d\Gamma_N \, \rho_{gc} &= \overline{N}
\end{aligned}
\tag{792}
$$

6.1 The Ideal Gas

The Grand-Canonical Partition function Ξ is given by

$$
\begin{aligned}
\Xi &= \sum_{N=0}^{\infty} \frac{\exp\left[\beta \mu N\right]}{N! \, (2 \pi \hbar)^{dN}} \int d\Gamma_N \, \exp\left[-\beta H_N\right] \\
&= \sum_{N=0}^{\infty} \exp\left[\beta \mu N\right] Z_N
\end{aligned}
\tag{793}
$$

However, for an ideal gas the Canonical partition function $Z)N$ is given by

$$
Z_N = \frac{1}{N!} \left(\frac{V}{\lambda^d}\right)^N
\tag{794}
$$

Therefore, one has

$$
\begin{aligned}
\Xi &= \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{\exp[\beta \mu] V}{\lambda^d}\right)^N \\
&= \exp\left[\frac{\exp[\beta \mu] V}{\lambda^d}\right]
\end{aligned}
\tag{795}
$$

154

This leads to the expression for the Grand-Canonical Potential Ω

$$
\begin{aligned}
\Omega & = -k_B \, T \, \ln \Xi \\
& = -k_B \, T \, \exp \left[\beta \, \mu \right] \frac{V}{\lambda^d}
\end{aligned}
\tag{796}
$$

The average number of particles is given by

$$
\begin{aligned}
\overline{N} & = -\left(\frac{\partial \Omega}{\partial \mu} \right)_{V,T} \\
& = \exp \left[\beta \, \mu \right] \frac{V}{\lambda^d}
\end{aligned}
\tag{797}
$$

Thus, the chemical potential μ is given by the equation

$$
\mu = k_B \, T \, \ln \left(\frac{\overline{N} \, \lambda^d}{V} \right)
\tag{798}
$$

which is identical to the result found by using the Canonical Ensemble. Furthermore, on using the thermodynamic relation

$$
\begin{aligned}
P & = -\left(\frac{\partial \Omega}{\partial V} \right)_{\mu,T} \\
& = \exp \left[\beta \, \mu \right] \frac{k_B \, T}{\lambda^d}
\end{aligned}
\tag{799}
$$

which on combining with the expression for $\overline{N}$, results in the ideal gas law

$$
P = \overline{N} \, \frac{k_B \, T}{V}
\tag{800}
$$

Homework:

Consider an ideal gas of atoms represented by the Grand-Canonical Ensemble.

Show that the probability P_N of finding a subsystem with N atoms is given by

$$
P_N = \frac{1}{N!} \, \overline{N}^N \, \exp \left[-\overline{N} \right]
\tag{801}
$$

where $\overline{N}$ is the average number of atoms.

6.2 Fluctuations in the Number of Particles

In the Grand-Canonical Ensemble the probability of finding the system in a state with N particles in a volume element $d\Gamma_N$ of phase space is given by

$$
\begin{aligned}
dp_{N,E} &= \left(\frac{dp}{d\Gamma} \right)_{N,E} d\Gamma_N \\
&= \frac{1}{\Xi} \exp\left[\beta \mu N \right] \exp\left[- \beta H_N(\{p_i, q_i\}) \right] \frac{d\Gamma_N}{\Gamma_{N,0}} \quad (802)
\end{aligned}
$$

where the Grand-Canonical Partition Function is given by

$$
\Xi = \sum_{N=0}^{\infty} \exp\left[\beta \mu N \right] \int \frac{d\Gamma_N}{\Gamma_{N,0}} \exp\left[- \beta H_N(\{p_i, q_i\}) \right] \quad (803)
$$

The average number of particles $\overline{N}$ is defined by the expression

$$
\overline{N} = \frac{1}{\Xi} \sum_{N=0}^{\infty} N \exp\left[\beta \mu N \right] \int \frac{d\Gamma_N}{\Gamma_{N,0}} \exp\left[- \beta H_N(\{p_i, q_i\}) \right] \quad (804)
$$

where, since we are not interested in the position of the N-particle system in its phase space, we have integrated over $d\Gamma_N$. The above expression can be re-written in terms of a logarithmic derivative of Ξ w.r.t. the product $\beta \mu$. Alternatively, on defining the fugacity z as

$$
z = \exp\left[\beta \mu \right] \quad (805)
$$

one may express $\overline{N}$ as

$$
\overline{N} = \frac{1}{\Xi} \sum_{N=0}^{\infty} N z^N \int \frac{d\Gamma_N}{\Gamma_{N,0}} \exp\left[- \beta H_N(\{p_i, q_i\}) \right] \quad (806)
$$

which can be expressed as the derivative w.r.t. z

$$
\begin{aligned}
\overline{N} &= \frac{1}{\Xi} \left(z \frac{\partial \Xi}{\partial z} \right) \\
&= z \frac{\partial \ln \Xi}{\partial z} \quad (807)
\end{aligned}
$$

The average value $\overline{N^2}$ can also be written as a second derivative w.r.t. $\beta \mu$ or

$$
\begin{aligned}
\overline{N^2} &= \frac{1}{\Xi} \left(z \frac{\partial}{\partial z} + z^2 \frac{\partial^2}{\partial z^2} \right) \Xi \\
&= \frac{1}{\Xi} \left(z \frac{\partial}{\partial z} \right)^2 \Xi \quad (808)
\end{aligned}
$$

Likewise, the average squared fluctuations of N is given by

$$
\begin{aligned}
\overline{\Delta N^2} &= \overline{N^2} - \overline{N}^2 \\
&= \frac{1}{\Xi}\left(z\,\frac{\partial}{\partial z} \right)^2 \Xi - \frac{1}{\Xi^2}\left(z\,\frac{\partial \Xi}{\partial z} \right)^2 \\
&= \left(z\,\frac{\partial}{\partial z} \right)^2 \ln \Xi \\
&= \left(z\,\frac{\partial}{\partial z} \right) \overline{N} \\
&= k_B\,T\,\frac{\partial \overline{N}}{\partial \mu}
\end{aligned}
\tag{809}
$$

The relative fluctuation of the particle number is given by

$$
\frac{\overline{\Delta N^2}}{\overline{N}^2} = \frac{k_B\,T}{\overline{N}^2}\left(\frac{\partial \overline{N}}{\partial \mu} \right)_{V,T}
\tag{810}
$$

which is of order $1/\overline{N}$ and vanishes in the thermodynamic limit.

The above expression for the relative fluctuations is not expressed in terms of quantities that are not easily measurable. However, the factor $(\frac{\partial \overline{N}}{\partial \mu})_T$ can be re-written in terms of $(\frac{\partial V}{\partial P})_T$ which is quite easily measurable. On combining the expression for the infinitesimal change in the Grand-Canonical Potential

$$
d\Omega = -\,S\,dT - P\,dV - \overline{N}\,d\mu
\tag{811}
$$

with

$$
\Omega = -\,P\,V
\tag{812}
$$

one finds the relation

$$
-\,V\,dP = -\,S\,dT - \overline{N}\,d\mu
\tag{813}
$$

For a process at constant T this reduces to a relation between P and μ

$$
V\,dP = \overline{N}\,d\mu
\tag{814}
$$

which on dividing by dV becomes

$$
V\left(\frac{\partial P}{\partial V} \right)_T = \overline{N}\left(\frac{\partial \mu}{\partial V} \right)_T
\tag{815}
$$

This second term in this relation can be expressed in terms of the derivative of the volume per particle

$$
v = \frac{V}{\overline{N}}
\tag{816}
$$

as

$$V \left(\frac{\partial P}{\partial V}\right)_T = \left(\frac{\partial \mu}{\partial v}\right)_T \tag{817}$$

However, the second term can be reinterpreted as the derivative of μ w.r.t. $\overline{N}$ at constant V. Therefore, one obtains

$$V \left(\frac{\partial P}{\partial V}\right)_T = \left(\frac{\partial \overline{N}}{\partial v}\right)_V \left(\frac{\partial \mu}{\partial \overline{N}}\right)_T \tag{818}$$

but since $\overline{N} = V/v$ one has

$$\begin{aligned} V \left(\frac{\partial P}{\partial V}\right)_T &= -\frac{V}{v^2} \left(\frac{\partial \mu}{\partial \overline{N}}\right)_T \\ &= -\frac{\overline{N}^2}{V} \left(\frac{\partial \mu}{\partial \overline{N}}\right)_T \end{aligned} \tag{819}$$

On inverting this relation and substituting this in the expression for the relative fluctuations in the number of particles, one finds that

$$\frac{\overline{\Delta N^2}}{\overline{N}^2} = -\frac{k_B\, T}{V^2} \left(\frac{\partial P}{\partial V}\right)_T \tag{820}$$

Again, one finds that the relative fluctuations of the particle are inversely proportional to the volume and thus vanish in the thermodynamic limit.

Homework:

Show that for an ideal gas

$$\overline{\Delta N^2} = \overline{N} \tag{821}$$

6.3 Energy Fluctuations in the Grand-Canonical Ensemble

The average energy $\overline{E}$ in the Grand-Canonical Ensemble can be represented as a derivative of the logarithm of the Grand-Canonical Partition Function w.r.t. β as

$$\begin{aligned} \overline{E} &= -\frac{1}{\Xi} \left(\frac{\partial \Xi}{\partial \beta}\right)_z \\ &= -\left(\frac{\partial \ln \Xi}{\partial \beta}\right)_{z,V} \end{aligned} \tag{822}$$

where the fugacity z is held constant. Likewise, the mean squared energy is given by

$$\overline{E^2} = \frac{1}{\Xi} \left(\frac{\partial^2 \Xi}{\partial \beta^2}\right)_{z,V} \tag{823}$$

Hence, the mean squared fluctuation of the energy is given by

$$
\begin{aligned}
\overline{\Delta E^2} &= \overline{E^2} - \overline{E}^2 \\
&= \left(\frac{\partial^2 \ln \Xi}{\partial \beta^2} \right)_{z,V} \\
&= - \left(\frac{\partial \overline{E}}{\partial \beta} \right)_{z,V} \\
&= k_B \, T^2 \left(\frac{\partial \overline{E}}{\partial T} \right)_{z,V}
\end{aligned}
\tag{824}
$$

The above relations are similar to the relations found for the Canonical Ensemble, but are different because the derivatives are evaluated at constant N for the Canonical Ensemble and at constant fugacity for the Grand-Canonical Ensemble. Hence, the energy fluctuations are different in the Canonical and the Grand-Canonical Ensembles.

The cause for the difference between the fluctuations in the Grand Canonical and the Canonical Ensembles is not easy to discern from the above expression since the fugacity is difficult to measure. The difference can be made explicit by using thermodynamics, in which case we identify the average energy $\overline{E}$ in the Grand Canonical Ensemble with U. That is, since on holding V fixed

$$
\overline{N} = \overline{N}(T, z)
\tag{825}
$$

one has

$$
U = U(T, \overline{N}(T, z))
\tag{826}
$$

so the infinitesimal variation in U can be expressed as

$$
\begin{aligned}
dU &= \left(\frac{\partial U}{\partial T} \right)_{\overline{N}} dT + \left(\frac{\partial U}{\partial \overline{N}} \right)_{T} d\overline{N} \\
&= \left(\frac{\partial U}{\partial T} \right)_{\overline{N}} dT + \left(\frac{\partial U}{\partial \overline{N}} \right)_{T} \left[\left(\frac{\partial \overline{N}}{\partial T} \right)_{z} dT + \left(\frac{\partial \overline{N}}{\partial z} \right)_{T} dz \right]
\end{aligned}
\tag{827}
$$

Hence, the derivative of U w.r.t. T with z kept constant is given by

$$
\begin{aligned}
\left(\frac{\partial U}{\partial T} \right)_{z,V} &= \left(\frac{\partial U}{\partial T} \right)_{\overline{N},V} + \left(\frac{\partial U}{\partial \overline{N}} \right)_{T,V} \left(\frac{\partial \overline{N}}{\partial T} \right)_{z,V} \\
&= C_{N,V} + \left(\frac{\partial U}{\partial \overline{N}} \right)_{T,V} \left(\frac{\partial \overline{N}}{\partial T} \right)_{z,V}
\end{aligned}
\tag{828}
$$

Therefore, part of the energy fluctuation in the Grand-Canonical Ensemble is the same as the energy fluctuations in the Canonical Ensemble where the number of particles and the volume are fixed and the other contribution originates

from the temperature dependence of the number of particles.

One can obtain more insight into the origin of the energy fluctuations in the Grand-Canonical Ensemble by using a specific thermodynamic relation involving the factor

$$\left(\frac{\partial U}{\partial \overline{N}}\right)_{T,V} \tag{829}$$

and another relation for

$$\left(\frac{\partial \overline{N}}{\partial \overline{T}}\right)_{z,V} \tag{830}$$

On considering the infinitesimal change

$$dU = T\ dS - P\ dV + \mu\ d\overline{N} \tag{831}$$

with constant T and V, then on dividing by $d\overline{N}$ one obtains

$$\left(\frac{\partial U}{\partial \overline{N}}\right)_{T,V} = T\ \left(\frac{\partial S}{\partial \overline{N}}\right)_{T,V} + \mu \tag{832}$$

Substitution of the Maxwell relation

$$\left(\frac{\partial S}{\partial \overline{N}}\right)_{T,V} = -\left(\frac{\partial \mu}{\partial T}\right)_{\overline{N},V} \tag{833}$$

obtained from the analyticity of the Helmholtz Free-Energy $F(T,V,\overline{N})$, yields the first thermodynamic relation

$$\left(\frac{\partial U}{\partial \overline{N}}\right)_{T,V} = \mu - T\ \left(\frac{\partial \mu}{\partial T}\right)_{\overline{N},V} \tag{834}$$

The thermodynamic relation for the factor

$$\left(\frac{\partial \overline{N}}{\partial \overline{T}}\right)_{z,V} \tag{835}$$

is obtained from $\overline{N}(T,V,\mu)$ by expressing $\mu = \mu(T,z)$ so

$$\left(\frac{\partial \overline{N}}{\partial T}\right)_{z,V} = \left(\frac{\partial \overline{N}}{\partial T}\right)_{\mu,V} + \left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V}\ \left(\frac{\partial \mu}{\partial T}\right)_{z,V} \tag{836}$$

The first term on the right-hand side can be re-written yielding

$$\left(\frac{\partial \overline{N}}{\partial T}\right)_{z,V} = -\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V}\ \left(\frac{\partial \mu}{\partial T}\right)_{\overline{N},V} + \left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V}\ \left(\frac{\partial \mu}{\partial T}\right)_{z,V} \tag{837}$$

and on recognizing that μ is related to the fugacity by

$$\mu = k_B\ T\ \ln z \tag{838}$$

then

$$\left(\frac{\partial \mu}{\partial T}\right)_z = \frac{\mu}{T} \tag{839}$$

Hence, we have

$$\left(\frac{\partial \overline{N}}{\partial T}\right)_{z,V} = -\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \left(\frac{\partial \mu}{\partial T}\right)_{\overline{N},V} + \frac{\mu}{T}\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \tag{840}$$

$$= \frac{1}{T}\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \left[\mu - T\left(\frac{\partial \mu}{\partial T}\right)_{\overline{N},V}\right] \tag{841}$$

On substituting the first relation in the above expression one finds

$$\left(\frac{\partial \overline{N}}{\partial T}\right)_{z,V} = \frac{1}{T}\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \left(\frac{\partial U}{\partial \overline{N}}\right)_{T,V} \tag{842}$$

This analysis yields the two equivalent expressions for the energy fluctuations in the Grand-Canonical Ensemble

$$\overline{\Delta E^2} = k_B T^2 C_{N,V} + k_B T \left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \left[\mu - T\left(\frac{\partial \mu}{\partial T}\right)_{\overline{N},V}\right]^2 \tag{843}$$

and

$$\overline{\Delta E^2} = k_B T^2 C_{N,V} + k_B T \left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \left(\frac{\partial U}{\partial \overline{N}}\right)_{T,V}^2$$

$$= k_B T^2 C_{N,V} + \overline{\Delta N^2}\left(\frac{\partial U}{\partial \overline{N}}\right)_{T,V}^2 \tag{844}$$

where, in the last line we have used the equality

$$\overline{\Delta N^2} = k_B T \left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V} \tag{845}$$

The second expression for $\overline{\Delta E^2}$ shows that the mean squared energy fluctuations have two contributions, one originating from the mean squared energy fluctuation with a fixed number of particles and the second contribution comes from the mean squared fluctuations in the particle number where each particle that is exchanged with the reservoir carries with it the energy $(\frac{\partial E}{\partial \overline{N}})_T$.

Homework:

Show that the specific heat at constant N is related to the specific heat at constant μ via the relation

$$C_{V,N} = C_{V,\mu} - T\frac{\left(\frac{\partial \overline{N}}{\partial T}\right)_{\mu,V}^2}{\left(\frac{\partial \overline{N}}{\partial \mu}\right)_{T,V}} \tag{846}$$

7 Quantum Statistical Mechanics

Quantum Statistical Mechanics describes the thermodynamic properties of macroscopically large many-particle quantum systems.

7.1 Quantum Microstates and Measurements

In Quantum Mechanics a microscopic state of a many particle system is represented by a vector in Hilbert space

$$| \Psi >$$

Any two states $| \Psi >$ and $| \Psi >$ in Hilbert space have an inner product, which is given by a complex number

$$< \Phi | \Psi > = < \Psi | \Phi >^* \qquad (847)$$

The states are normalized to unity

$$< \Phi | \Phi > = 1 \qquad (848)$$

A set of states $| n >$ form an orthonormal set if their inner product satisfies

$$< n | m > = \delta_{n.m} \qquad (849)$$

where $\delta_{n,m}$ is the Kronecker delta function. An orthonormal set is complete if any arbitrary state can be expanded as

$$| \Psi > = \sum_n C_n | n > \qquad (850)$$

where the expansion coefficients C_n are complex numbers, which are found as

$$C_n = < n | \Psi > \qquad (851)$$

Thus, the expansion is given by

$$| \Psi > = \sum_n | n > < n | \Psi > \qquad (852)$$

Hence, we have the completeness condition

$$\hat{\mathcal{I}} = \sum_n | n > < n | \qquad (853)$$

Using the completeness condition, the normalization condition can be written as

$$< \Psi | \Psi > = \sum_n < \Psi | n > < n | \Psi > = 1 \qquad (854)$$

162

On inserting a complete set of coordinate eigenstates, this condition reduces to

$$1 \; = \; \prod_{i=1}^{3N} \left\{ \int dq_i \right\} \left| \Psi(q_1, q_2, q_3, \ldots, q_{3N}) \right|^2 \qquad (855)$$

Physical observables $A_j(\{p_i, q_i\})$ are represented by Hermitean operators $A_j(\{\hat{p}_i, \hat{q}_i\})$. If the Poisson Bracket of two classical observables A_j and A_k is represented by

$$\left\{ A_j \; , \; A_k \right\}_{PB}$$

then the Poisson Bracket of two quantum operators is represented by the commutator of the operators divided by $i\,\hbar$

$$\left\{ A_j \; , \; A_k \right\}_{PB} \; = \; \frac{1}{i\,\hbar} \left[\hat{A}_j \; , \; \hat{A}_k \right] \qquad (856)$$

In particular, since the Poisson Bracket for canonically conjugate coordinates and momenta are given by

$$\left\{ p_i \; , \; q_j \right\}_{PB} \; = \; -\,\delta_{i,j} \qquad (857)$$

then, one has the commutation relations

$$\left[\hat{p}_i \; , \; \hat{q}_j \right] \; = \; -\,i\,\hbar\,\delta_{i,j} \qquad (858)$$

The possible values of a measurement of A on a systems are the eigenvalues a_n found from the eigenvalue equation

$$\hat{A} \mid a_n \; > \; = \; a_n \mid a_n \; > \qquad (859)$$

If a measurement of $\hat{A}$ on a system results in the value a_n, then immediately after the measurement the system is definitely known to be in a state which is an eigenstate of A with eigenvalue a_n. The number of eigenstates corresponding to the same eigenvalue a_n is known as the degeneracy $\mathcal{D}_n$. The degenerate eigenstates $\mid a_{n,\alpha} \; >$ can be orthonormalized. Since the observable quantities are represented by Hermitean operators, the eigenstates form a complete set and the eigenvalues a_n are real.

It is possible to know with certainty the simultaneous values of measurements of A_j and A_k on a state if the operators $\hat{A}_j$ and $\hat{A}_k$ commute

$$[\; \hat{A}_j \; , \; \hat{A}_k \;] \; = \; 0 \qquad (860)$$

In which case it is possible to find simultaneous eigenstates of $\hat{A}_j$ and $\hat{A}_k$. A state is completely determined if it is an eigenstate of a maximal set of mutually commuting operators.

If a system is definitely in a state $\mid \Psi >$, it is a pure state. The probability that a measurement of A on the state $\mid \Psi >$ will yield the result a_n is given by

$$P(a_n) \; = \; \sum_{\gamma=1}^{\mathcal{D}_{a_n}} \mid \; < \; a_{n,\gamma} \mid \Psi \; > \; \mid^2 \tag{861}$$

where the sum is over the number of $\mathcal{D}_n$-fold degenerate eigenstates[21] $\mid a_{n,\gamma} >$ that correspond to the eigenvalue a_n. Thus, the average value $\overline{A}$ of the measurement of A on a pure state $\mid \Psi >$ is given by

$$
\begin{aligned}
\overline{A} \;\; &= \;\; \sum_n P_{a_n} \; a_n \\
&= \;\; \sum_{n,\gamma} < \; \Psi \mid a_{n,\gamma} \; > \; a_n \; < \; a_{n,\gamma} \mid \Psi \; > \\
&= \;\; \sum_{n,\gamma} < \; \Psi \mid \hat{A} \mid a_{n,\gamma} \; > < \; a_{n,\gamma} \mid \Psi \; > \\
&= \;\; < \; \Psi \mid \hat{A} \mid \Psi \; >
\end{aligned}
\tag{862}
$$

In the coordinate representation, the average can be expressed as

$$\overline{A} \; = \; \prod_{i=1}^{3N} \left\{ \int \; dq_i \; \right\} \; \Psi^*(q_1, q_2, \ldots, q_{3N}) \; A(\{\hat{p}_i, \hat{q}_i\}) \; \Psi(q_1, q_2, \ldots, q_{3N}) \tag{863}$$

In the time interval in which no measurements are performed, a state $\mid \Psi >$ evolves according to the equation

$$i \; \hbar \; \frac{\partial}{\partial t} \; \mid \Psi \; > \; = \; \hat{H} \mid \Psi \; > \tag{864}$$

Thus, the time evolution of the state $\mid \Psi >$ in our closed system is given by

$$\mid \Psi(t) \; > \; = \; \exp \left[\; - \; \frac{i}{\hbar} \; \hat{H} \; t \; \right] \mid \Psi \; > \tag{865}$$

7.2 The Density Operator and Thermal Averages

A macroscopic state may correspond to numerous microscopic states $\mid \Psi_\alpha >$ which are assigned probabilities p_α. Such macroscopic states are known as

---

[21]The states are assumed to have been orthogonalized, so that the eigenstates of A form a complete orthonormal set.

mixed states. The ensemble average of A is given by the weighted average

$$
\begin{aligned}
\overline{A} &= \sum_\alpha p_\alpha < \Psi_\alpha \mid \hat{A} \mid \Psi_\alpha > \\
&= \sum_\alpha \sum_{n,m} p_\alpha < \Psi_\alpha \mid n >< n \mid \hat{A} \mid m >< m \mid \Psi_\alpha > \\
&= \sum_{n,m} < n \mid \hat{A} \mid m > \sum_\alpha p_\alpha < m \mid \Psi_\alpha >< \Psi_\alpha \mid n > \quad (866)
\end{aligned}
$$

On defining the density operator $\hat{\rho}$ via

$$
\hat{\rho} = \sum_\alpha p_\alpha \mid \Psi_\alpha >< \Psi_\alpha \mid \quad (867)
$$

the average can be represented as

$$
\begin{aligned}
\overline{A} &= \sum_{n,m} < n \mid \hat{A} \mid m >< m \mid \hat{\rho} \mid n > \\
&= \sum_n < n \mid \hat{A}\,\hat{\rho} \mid n > \\
&= \text{Trace } \hat{A}\,\hat{\rho} \quad (868)
\end{aligned}
$$

where the last line defines the Trace over a complete set of states[22].

Since the probabilities p_α are normalized, the density operator satisfies

$$
\begin{aligned}
\text{Trace } \hat{\rho} &= \sum_n < n \mid \hat{\rho} \mid n > \quad &(869) \\
&= \sum_n \sum_\alpha p_\alpha < n \mid \Psi_\alpha >< \Psi_\alpha \mid n > \quad &(870) \\
&= \sum_\alpha p_\alpha \sum_n < \Psi_\alpha \mid n >< n \mid \Psi_\alpha > \quad &(871) \\
&= \sum_\alpha p_\alpha < \Psi_\alpha \mid \Psi_\alpha > \quad &(872) \\
&= \sum_\alpha p_\alpha \quad &(873) \\
&= 1 \quad &(874)
\end{aligned}
$$

Thus, the trace of the density matrix is unity.

The time-dependence of the density operator is found from the time dependence of the basis states

$$
\hat{\rho}(t) = \sum_\alpha p_\alpha \mid \Psi_\alpha(t) >< \Psi_\alpha(t) \mid
$$

---

[22]Gleason's Theorem assures us that the only possible measure of probability on a Hilbert space has the form of a density matrix. [A.M. Gleason,"Measures on the closed subspaces of a Hilbert space", Journal of Mathematics and Mechanics **6**, 885-893, (1957).].

$$
= \sum_{\alpha} p_{\alpha} \, \exp\left[\, - \, \frac{i}{\hbar} \, \hat{H} \, t \, \right] \mid \Psi_{\alpha} > < \, \Psi_{\alpha} \mid \, \exp\left[\, + \, \frac{i}{\hbar} \, \hat{H} \, t \, \right]
$$
$$(875)$$

This shows that the time evolution of the density operator has the form of a unitary transformation. From this, one finds the equation of motion for the density operator is given by

$$
i \, \hbar \, \frac{\partial \hat{\rho}}{\partial t} \; = \; [\, \hat{H} \, , \, \hat{\rho} \,] \tag{876}
$$

or, equivalently

$$
i \, \hbar \, \frac{d \hat{\rho}}{dt} \; = \; i \, \hbar \, \frac{\partial \hat{\rho}}{\partial t} \; + \; [\, \hat{\rho} \, , \, \hat{H} \,] \; = \; 0 \tag{877}
$$

This last expression could have been derived directly from the Poisson equation of motion for the probability density by Canonical Quantization.

In the Micro-Canonical Ensemble, all the states $\mid \Psi_{n,\gamma} >$ in the ensemble must be energy eigenstates belonging to the same energy eigenvalue $E = E_n$

$$
\hat{H} \mid \Psi_{n,\gamma} \; > \; = \; E_n \mid \Psi_{n,\gamma} \; > \tag{878}
$$

The number of these degenerate eigenstates is denoted by N_Γ. Therefore, in an equilibrium state, the probabilities are given by

$$
p_{\gamma} \; = \; \frac{1}{N_\Gamma} \tag{879}
$$

which is equivalent to the hypothesis of equal a priori probabilities. Thus, the density operator in the Micro-Canonical can be written as

$$
\hat{\rho}_{mc} \; = \; \frac{1}{N_\Gamma} \sum_{\gamma=1}^{N_\Gamma} \mid \Psi_{n,\gamma} \; > < \; \Psi_{n,\gamma} \mid \tag{880}
$$

On defining the von Neumann entropy in terms of the density operator by

$$
S \; = \; - \, k_B \, \text{Trace} \, \hat{\rho} \, \ln \hat{\rho} \tag{881}
$$

this is evaluated as

$$
\begin{aligned}
S \; &= \; - \, k_B \, \sum_{\gamma} \, p_{\gamma} \, \ln p_{\gamma} \\
&= \; k_B \, \ln N_\Gamma
\end{aligned} \tag{882}
$$

in agreement with our previous notation. If the energy E of the Micro-Canonical Ensemble corresponds to a non-degenerate state, it is a pure state and has $N_\Gamma = 1$, then the entropy vanishes. This observation is in accordance with the universal constant value of entropy, demanded by Nernst's law in the $T \to 0$, as

166

being defined as being zero.

Since the set of all energy eigenstates are complete, in the Canonical Ensemble the density matrix is given by

$$\hat{\rho}_c = \frac{1}{Z_N} \exp\left[-\beta \hat{H}_N \right] \tag{883}$$

where the partition function is given by the normalization condition on $\hat{\rho}_c$

$$Z_N = \text{Trace} \exp\left[-\beta \hat{H}_N \right] \tag{884}$$

If the trace is evaluated in a complete set of energy eigenstates $\mid \Psi_\alpha >$, the result for the partition function reduces to

$$Z_N = \sum_\alpha \exp\left[-\beta \hat{E}_\alpha \right] \tag{885}$$

where the sum runs over all the degenerate states for each energy.

In the Grand-Canonical Ensemble, one is working in a Hilbert space with a variable number of particles (Fock Space). In this case, one has

$$\hat{\rho}_{gc} = \frac{1}{\Xi} \exp\left[-\beta \left(\hat{H} - \mu \hat{N} \right) \right] \tag{886}$$

in which the Grand Canonical Partition Function is given by

$$\Xi = \text{Trace} \exp\left[-\beta \left(\hat{H} - \mu \hat{N} \right) \right] \tag{887}$$

where $\hat{N}$ is the number operator. If the particle number is conserved so that both $\hat{N}$ and $\hat{H}$ can be diagonalized simultaneously, then the partition function can be reduced to the form

$$\Xi = \sum_{N=0}^{\infty} \exp\left[\beta \mu N \right] Z_N \tag{888}$$

where the N are the eigenvalues of the number operator.

Homework:

Consider a system which has statistically independent sub-systems A and B. Statistically independent systems are defined as having a factorizable density operator

$$\hat{\rho} = \hat{\rho}_A \hat{\rho}_B \tag{889}$$

The von Neumann entropies of the subsystems are defined as

$$S_A = - k_B \text{ Trace}_A \left[\hat{\rho}_A \ln \hat{\rho}_A \right] \tag{890}$$

and

$$S_B = - k_B \text{ Trace}_B \left[\hat{\rho}_B \ln \hat{\rho}_B \right] \tag{891}$$

where the trace over A evaluated as a sum over a complete set of states for the subsystem A, and with a similar definition holds for subsystem B. Show that the total entropy is additive

$$S = S_A + S_B \tag{892}$$

For two subsystems which are not statistically independent, one can define the density operators for the subsystems by the partial traces

$$\hat{\rho}_A = \text{Trace}_B \, \hat{\rho}$$
$$\hat{\rho}_B = \text{Trace}_A \, \hat{\rho} \tag{893}$$

where each trace is evaluated as a sum over a complete set of states for the subsystem. In this case, one can prove that the entropy satisfies the triangle inequality[23]

$$S_A + S_B > S > | S_A - S_B | \tag{894}$$

Find an example of a quantum system in a pure state for which $S = 0$ but, nevertheless, both S_A and S_B are greater than zero.

Homework:

Show that the von Neumann entropy is constant, due to the unitary nature of the time evolution operator.

7.3   Indistinguishable Particles

Any labelling of indistinguishable particles is unphysical, by definition. Since any measurement will A will produce results which are independent of the choice of labelling the operators must be symmetric under any permutation of the labels. Hence, every physical operator $A(\{\underline{p}_i, \underline{r}_i\})$, including the Hamiltonian, must be a symmetric function of the particle position and momenta vectors

---

[23]H. Araki and Elliott H. Lieb, "Entropy Inequalities", Communications in Mathematical Physics, 18, 160-170 (1970).

$(\hat{\underline{p}}_i, \underline{r}_i)$ of the N particles. Any permutation of the set of N particles can be represented in terms of the successive interchanges of pairs of particles. The pair of particles labelled as $(\hat{\underline{p}}_i, \underline{r}_i)$ and $(\hat{\underline{p}}_j, \underline{r}_j)$ are interchanged by the permutation operator $\hat{P}_{i,j}$. The permutation operator $\hat{P}_{i,j}$ is Hermitean and unitary. In the coordinate representation, the permutation operator has the effect

$$\hat{P}_{i,j} \ \Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_i, \ldots, \underline{r}_j, \ldots, \underline{r}_N) \ = \ \Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_j, \ldots, \underline{r}_i, \ldots, \underline{r}_N) \quad (895)$$

Since any physical operator must be invariant under the permutation of any two particles, one has

$$\hat{P}_{i,j} \ A(\{\underline{p}_i, \underline{r}_i\}) \ \hat{P}_{i,j}^{-1} \ = \ A(\{\underline{p}_i, \underline{r}_i\}) \quad (896)$$

Hence, every $\hat{P}_{i,j}$ commutes with every physical operator including the Hamiltonian

$$[\ \hat{P}_{i,j} \ , \ \hat{A} \ ] \ = \ 0 \quad (897)$$

Therefore, the permutation operators can be diagonalized simultaneously together with any complete set of compatible physical operators.

The eigenvalues of the permutation operator are defined as p

$$\hat{P}_{i,j} \mid \Phi_p \ > \ = \ p \mid \Phi_p \ > \quad (898)$$

However, as two successive interchanges of the labels i and j leaves the state unchanged, one has

$$\hat{P}_{i,j}^2 \ = \ \hat{\mathcal{I}} \quad (899)$$

Thus, the eigenvalues of the permutation operators must satisfy

$$\begin{aligned} \hat{P}_{i,j}^2 \mid \Phi_p \ > \ &= \ p^2 \mid \Phi_p \ > \\ &= \ \mid \Phi_p \ > \end{aligned} \quad (900)$$

which leads to the solutions for the eigenvalues

$$p \ = \ \pm \, 1 \quad (901)$$

which are constants of motion. Since the particles are indistinguishable, all pairs particles must have the same value of the eigenvalue p.

Since the real space probability density is given by

$$\mid \Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_i, \ldots, \underline{r}_j, \ldots, \underline{r}_N) \mid^2 \quad (902)$$

is observable, and not the wave function, measurements on two states differing only by the interchange two particle labels will yield results which have identical distributions.

Fermions and Bosons

Particles are known as fermions if their wave functions are antisymmetric under the interchange of any pair of particles, whereas the particles are called bosons if the wave function is symmetric under a single interchange.

In certain two-dimensional systems, states with mixed symmetry can occur[24]. The exotic particles with mixed symmetries are known as anyons, and they obey fractional statistics.

Fermions and Fermi-Dirac Statistics

Particles are fermions, if their wave function is antisymmetric under the interchange of any pair of particles

$$\Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_i, \ldots, \underline{r}_j, \ldots, \underline{r}_N) = -\Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_j, \ldots, \underline{r}_i, \ldots, \underline{r}_N) \quad (903)$$

Examples of fermions are given by electrons, neutrinos, quarks, protons and neutrons, and Helium^3 atoms.

Bosons and Bose-Einstein Statistics

Particles are bosons if their wave function is symmetric under the interchange of any pair of particles

$$\Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_i, \ldots, \underline{r}_j, \ldots, \underline{r}_N) = \Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_j, \ldots, \underline{r}_i, \ldots, \underline{r}_N) \quad (904)$$

Examples of bosons are given by photons, gluons, phonons, and Helium^4 atoms.

One can represent tan arbitrary N-particle state with the requires symmetry as a linear superposition of a complete set of orthonormal N-particle basis states Φ. These many-particle states are composed as a properly symmetrized product of single-particles wave functions $\phi_\alpha(\underline{r})$, which form a complete orthonormal set

$$
\begin{aligned}
\int d^3\underline{r} \; \phi_{\alpha'}^*(\underline{r}) \; \phi_\alpha(\underline{r}) &= \delta_{\alpha',\alpha} \\
\sum_\alpha \phi_\alpha^*(\underline{r}') \; \phi_\alpha(\underline{r}) &= \delta^3(\underline{r} - \underline{r}')
\end{aligned}
\quad (905)
$$

The many-particle basis states $\Phi_{\alpha_1,\alpha_2,\ldots,\alpha_N}(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_N)$ are composed from a symmetrized linear superposition of N single-particle states

$$\Phi_{\alpha_1,\alpha_2,\ldots,\alpha_N}(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_N) = \aleph^{-1} \sum_P (\pm 1)^{n_p} \, \hat{P} \, \phi_{\alpha_1}(\underline{r}_1) \, \phi_{\alpha_2}(\underline{r}_2) \, \ldots \, \phi_{\alpha_N}(\underline{r}_N)$$

$$(906)$$

where the sum runs over all $N!$ permutations of the particle indices and n_p is the order of the permutation. That is, n_p is the number of pairwise interchanges that produce the permutation. For boson wave functions, the positive

---

[24]J.M. Leinaas and J. Myrheim, "On the theory of identical particles". Il Nuovo Cimento B 37, 1-23 (1977).

sign holds, and the minus sign holds for fermions.

The Pauli Exclusion Principle

The Pauli Exclusion Principle states that a single-particle state cannot contain more than one fermion. That is, any physically acceptable many-particle state describing identical fermions $\Phi_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_j\ldots}$ must have $\alpha_i \neq \alpha_j$. Otherwise, if $\alpha_i = \alpha_j$, then one has

$$\hat{P}_{i,j}\,\Phi_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_i\ldots}(\underline{r}_1,\ldots,\underline{r}_i,\ldots,\underline{r}_j,\ldots) \;=\; \Phi_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_i\ldots}(\underline{r}_1,\ldots,\underline{r}_i,\ldots,\underline{r}_j,\ldots) \tag{907}$$

since changing the orders of $\phi_{\alpha_i}(\underline{r}_i)$ and $\phi_{\alpha_i}(\underline{r}_j)$ does not change the sign of the wave function. Furthermore, since the many-particle state is an eigenstate of $\hat{P}_{i,j}$ with eigenvalue -1, one also has

$$\hat{P}_{i,j}\,\Phi_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_i\ldots}(\underline{r}_1,\ldots,\underline{r}_i,\ldots,\underline{r}_j,\ldots) \;=\; -\,\Phi_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_i\ldots}(\underline{r}_1,\ldots,\underline{r}_i,\ldots,\underline{r}_j,\ldots) \tag{908}$$

Therefore, on equating the right-hand sides one finds that

$$\Phi_{\alpha_1,\ldots,\alpha_i,\ldots,\alpha_i\ldots}(\underline{r}_1,\ldots,\underline{r}_i,\ldots,\underline{r}_j,\ldots) \;=\; 0 \tag{909}$$

which indicates that the state where two fermions occupy the same single-particle eigenstate does not exist.

The Occupation Number Representation

Instead of labeling the many-particle basis states by the eigenvalues α_1, α_2, ..., α_N we can specify the number of times each single-particle eigenstate is occupied. The number of times that a specific one-particle state ϕ_α occurs is denoted by n_α which is called the occupation number. Specifying the occupation numbers n_{α_1}, n_{α_2}, ..., uniquely specifies the many-particle states $\Phi_{n_{\alpha_1},n_{\alpha_2},\ldots}$. For a system with N particles, the sum of the occupation numbers is just equal to the total number of particles

$$\sum_\alpha n_\alpha \;=\; N \tag{910}$$

For fermions, the Pauli exclusion principle limits n_α to have values of either 0 or 1. For bosons, n_α can have any positive integer value, including zero.

The orthonormality relation

$$\prod_{i=1}^{N}\left\{\int d^3\underline{r}_i\right\}\Phi^*_{n_{\alpha'_1},n_{\alpha'_2},\ldots}(\underline{r}_1,\underline{r}_2,\ldots,\underline{r}_N)\,\Phi_{n_{\alpha_1},n_{\alpha_2},\ldots}(\underline{r}_1,\underline{r}_2,\ldots,\underline{r}_N) \;=\; \delta_{n_{\alpha'_1},n_{\alpha_1}}\,\delta_{n_{\alpha'_2},n_{\alpha_2}}\,\cdots \tag{911}$$

leads to the identification of the normalization constant $\aleph$. For fermions, one has

$$\aleph \;=\; \sqrt{N!} \tag{912}$$

which is just the number of terms in the wave function. For bosons, the normalization is given by

$$\aleph \;=\; \sqrt{N! \prod_\alpha n_\alpha!} \qquad (913)$$

where it should be noted that 0! is defined as unity.

An arbitrary many-particle state Ψ can be expressed in terms of the set of basis states as

$$\Psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_N) \;=\; \sum_{\{n_\alpha\}} C(n_{\alpha_1}, n_{\alpha_2}, \ldots)\, \Phi_{n_{\alpha_1}, n_{\alpha_2}, \ldots}(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_N) \quad (914)$$

where the expansion coefficients $C(n_{\alpha_1}, n_{\alpha_2}, \ldots)$ play the role of the wavefunction in the occupation number representation.

7.4 "The Spin-Statistics Theorem" and Composite Particles

The Spin-Statistics Theorem has its origins in Quantum Field Theory, and states that fermions have half odd integer spins and that bosons have integer spins. The theorem was first proposed by Markus Fierz[25] and later proved by Wolfgang Pauli[26]. Rather than prove the theorem, we shall be content to show that if the Spin-Statistics holds for elementary particles then it will also hold for composite particles. We shall also outline an observation which might be turned into a proof of the Theorem.

Two indistinguishable composite particles are permuted if all the elementary particles composing one composite particles are interchanged with the corresponding constituent particles of the other. When two identical composite particle each of which is composed of n_F elementary fermions and n_B elementary bosons are interchanged, the wave function will change by a factor of

$$(-1)^{n_F}\,(+1)^{n_B}$$

Hence, if a composite particle contains an odd number of fermions, the composite particle will be a fermion since the wave function of two such identical composite particles is antisymmetric under the interchange of the particles. On the other hand, if a composite particle contains an even number of fermions, the composite particle will be a boson since the wave function of two such identical composite particles will be symmetric under the interchange of the particles.

The above result is consistent with the application of the Spin-Statistics Theorem. A composite particle containing n_F fermions and n_B bosons will

[25]M. Fierz, "Über die relativistische Theorie kräftefreier Teilchen mit beliebigem Spin", Helvetica Physica Acta 12, 3-37, (1939).

[26]W. Pauli, "The Connection Between Spin and Statistics", Phys. Rev. 58, 716-722 (1940).
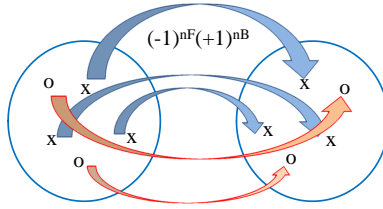
Figure 38: Composite particles are interchanged when their constituent particles are interchanged. For a system with n_F fermions and n_B bosons, the interchange changes the wave function representing the pair of composite particles by a factor of $(-1)^{n_F}(+1)^{n_B}$.

have a spin composed of n_F half-odd integers and n_B integers. If n_F is odd, the resulting spin will be a half-odd integer, whereas if n_F is even the resulting spin will be integer.

Thus, a composite particle containing an odd number of fermions n_F will have a half-odd integer spin and the wave function of a pair of identical composite particles with odd n_F will be antisymmetric under their interchange. Likewise, a composite particle containing an even number of fermions will have an integer spin, and the wave function of a pair of identical composite particles with even n_F will be symmetric under the interchange of the composite particles. Hence, the Spin-Statistics Theorem will be true for identical composite particles if it is true for their elementary constituents.

===

Example: The Isotopes of He

He^3 has two protons, a neutron and two electrons. Therefore, He^3 is a fermion.

He^4 has an extra neutron. Thus, it contains two protons, two neutrons and two electrons. Therefore, He^4 is a boson.

The difference in the quantum statistics of the two isotopes results in their having very different properties at low temperatures, although they are chemically similar. For example, their phase diagrams are very different and He^4 exhibits the phenomenon of superfluidity.

===

The idea behind the Spin-Statistics Theorem is that field operators must

transform under Lorentz transformations according to the spin of the particles that they describe. In a non-relativistic theory, the boost generators generates rotations. It is noted that rotations about an angle 2π are equivalent to the identity for fields with integer spins, such as scalar or vector fields. However, this is not true for odd half integer spin fields represented by spinors.

Consider the products of two fields $\hat{\Psi}$, with arbitrary spin, given by

$$\hat{R}(\pi) \ \hat{\Psi}(\underline{r}) \ \hat{\Psi}(-\underline{r}) \tag{915}$$

This product creates (or annihilates) two particles with spins that are rotated by π relative to each other. Now consider a rotation of this configuration by π around the origin. Under this rotation, the two points $\underline{r}$ and $-\underline{r}$ switch places, and the spins of the two field are rotated through an additional angle of π. Thus, the two product of the field operators transform as

$$\hat{R}(2\pi) \ \hat{\Psi}(-\underline{r}) \ \hat{R}(\pi) \ \hat{\Psi}(\underline{r}) \tag{916}$$

which for integer spins the above product is equivalent with

$$\hat{\Psi}(-\underline{r}) \ \hat{R}(\pi) \ \hat{\Psi}(\underline{r}) \tag{917}$$

and for half-integer spins, the product is equivalent to

$$- \ \hat{\Psi}(-\underline{r}) \ \hat{R}(\pi) \ \hat{\Psi}(\underline{r}) \tag{918}$$

Therefore, the particles associated with two spinful operator fields can be interchanged by a rotation, but must also involve a change of sign for half-integer spin fields.

A proof of the Spin-Statistics Theorem based on this observation would require that the following assumptions hold true:

(i) The theory has a Lorentz invariant Lagrangian.
(ii) The vacuum is Lorentz invariant.
(iii) The particle is a localized excitation which is not connected to any other object.
(iv) The particle is propagating so it has a finite mass.
(v) The particle is a real excitation, so that states involving the particle have positive definite norms.

Homework

Determine whether the isotope Rb^{87} is a Boson or a Fermion particle.

7.5   Second Quantization

Second Quantization for Bosons

174

For bose particles, one can introduce a set of operators $\hat{a}_\alpha^\dagger$ with single-particle quantum numbers α which are defined by

$$\hat{a}_\alpha^\dagger \; \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) \;=\; \sqrt{n_\alpha \,+\, 1} \;\; \Phi_{n_{\alpha_1},\ldots,n_\alpha+1,\ldots}(\underline{r}_1,\ldots,\underline{r}_{N+1})$$
(919)

and the Hermitean conjugate operator is found to satisfy

$$\hat{a}_\alpha \; \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) \;=\; \sqrt{n_\alpha} \;\; \Phi_{n_{\alpha_1},\ldots,n_\alpha-1,\ldots}(\underline{r}_1,\ldots,\underline{r}_{N-1})$$
(920)

The creation operator $\hat{a}_\alpha^\dagger$ adds an additional particle to the quantum level α, and $\hat{a}_\alpha$ annihilates a particle in the state α. If in the initial state Φ, $n_\alpha = 0$ then application of the annihilation operator yields zero.

One can define the number operator $\hat{n}_\alpha$ as the combination

$$\hat{n}_\alpha \;=\; \hat{a}_\alpha^\dagger \; \hat{a}_\alpha$$
(921)

which has the eigenvalue equation

$$\hat{n}_\alpha \; \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) \;=\; n_\alpha \; \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N)$$
(922)

where the eigenvalue n_α is the occupation number for the single-particle state. Hence, the number operator represents a measurement of the number of particles in a quantum level α. The total number of particles in the system N corresponds to the eigenvalues of the operator

$$
\begin{aligned}
\hat{N} \;&=\; \sum_\alpha \hat{n}_\alpha \\
&=\; \sum_\alpha \hat{a}_\alpha^\dagger \; \hat{a}_\alpha
\end{aligned}
$$
(923)

The creation and annihilation operators satisfy the commutation relations

$$[\; \hat{a}_\alpha \;,\; \hat{a}_{\alpha'}^\dagger \;] \;=\; \delta_{\alpha,\alpha'}$$
(924)

and

$$
\begin{aligned}
[\; \hat{a}_\alpha^\dagger \;,\; \hat{a}_{\alpha'}^\dagger \;] \;&=\; 0 \\
[\; \hat{a}_\alpha \;,\; \hat{a}_{\alpha'} \;] \;&=\; 0
\end{aligned}
$$
(925)

as can be easily shown. For example, the diagonal elements of the first commutator yields

$$
\begin{aligned}
\left(\hat{a}_\alpha \, \hat{a}_\alpha^\dagger \;-\; \hat{a}_\alpha^\dagger \, \hat{a}_\alpha \right) \mid \Phi_{n_{\alpha_1}\ldots n_\alpha\ldots} > \;&=\; \left((\, n_\alpha \,+\, 1 \,) \;-\; n_\alpha \right) \mid \Phi_{n_{\alpha_1}\ldots n_\alpha\ldots} > \\
&=\; 1 \mid \Phi_{n_{\alpha_1}\ldots n_\alpha\ldots} >
\end{aligned}
$$
(926)

175

and the off-diagonal elements yield zero.

On defining the vacuum state as $| \ 0 \ >$, one can create any arbitrary basis state by the repeated action of the creation operators

$$| \ \Phi_{n_{\alpha_1}, n_{\alpha_2} \ldots} \ > \ = \ \prod_{\alpha} \left\{ \frac{(\ \hat{a}_{\alpha}^{\dagger} \ )^{n_{\alpha}}}{\sqrt{n_{\alpha}!}} \right\} | \ 0 \ > \qquad (927)$$

where the product runs over all the single-particle quantum numbers α.

Any operator can also be expressed in terms of the creation and annihilation operators. For example, the Hamiltonian of a system of interacting particles can be written as

$$\hat{H} \ = \ \sum_{\alpha;\alpha'} < \ \alpha' \ | \ \hat{H}_0 \ | \ \alpha \ > \ \hat{a}_{\alpha'}^{\dagger} \ \hat{a}_{\alpha} + \frac{1}{2!} \sum_{\alpha,\beta;\alpha',\beta'} < \ \alpha' \ \beta' \ | \ \hat{V}_{int} \ | \ \alpha \ \beta \ > \ \hat{a}_{\alpha'}^{\dagger} \ \hat{a}_{\beta'}^{\dagger} \ \hat{a}_{\beta} \ \hat{a}_{\alpha}$$

$$\qquad (928)$$

where

$$< \ \alpha' \ | \ \hat{H}_0 \ | \ \alpha \ > \ = \ \int d^3\underline{r} \ \phi_{\alpha'}^*(\underline{r}) \left[\frac{\hat{p}^2}{2 \ m} \ + \ V_0(\underline{r}) \right] \phi_{\alpha}(\underline{r}) \qquad (929)$$

are the matrix elements of the single-particle energy and

$$< \ \alpha' \ \beta' \ | \ \hat{V}_{int} \ | \ \alpha \ \beta \ > \ = \ \int d^3\underline{r} \ \int d^3\underline{r}' \ \phi_{\alpha'}^*(\underline{r}) \ \phi_{\beta'}^*(\underline{r}') \ V_{int}(\underline{r}, \underline{r}') \ \phi_{\alpha}(\underline{r}) \ \phi_{\beta}(\underline{r}')$$

$$\qquad (930)$$

are the matrix elements of the two-body interaction terms.

Second Quantization for Fermions

For fermi particles, one can formally introduce a set of operators $\hat{c}_{\alpha}^{\dagger}$ corresponding to the set of single-particle quantum numbers α. The operators are defined by

$$\hat{c}_{\alpha}^{\dagger} \ \Phi_{n_{\alpha_1}, \ldots, n_{\alpha}, \ldots}(\underline{r}_1, \ldots, \underline{r}_N) \ = \ \sqrt{n_{\alpha} \ + \ 1} \ \Phi_{n_{\alpha_1}, \ldots, n_{\alpha}+1, \ldots}(\underline{r}_1, \ldots, \underline{r}_{N+1})$$

$$\qquad (931)$$

for anti-symmetric states Φ. The Hermitean conjugate operator $\hat{c}_{\alpha}$ is then found to satisfy

$$\hat{c}_{\alpha} \ \Phi_{n_{\alpha_1}, \ldots, n_{\alpha}, \ldots}(\underline{r}_1, \ldots, \underline{r}_N) \ = \ \sqrt{n_{\alpha}} \ \Phi_{n_{\alpha_1}, \ldots, n_{\alpha}-1, \ldots}(\underline{r}_1, \ldots, \underline{r}_{N-1}) \qquad (932)$$

Due to the antisymmetric nature of the states Φ, the creation and annihilation operators must satisfy the anti-commutation relations

$$\{ \ \hat{c}_{\alpha} \ , \ \hat{c}_{\alpha'}^{\dagger} \ \} \ = \ \delta_{\alpha,\alpha'} \qquad (933)$$

and

$$\{ \ \hat{c}_{\alpha}^{\dagger} \ , \ \hat{c}_{\alpha'}^{\dagger} \ \} \ = \ 0$$
$$\{ \ \hat{c}_{\alpha} \ , \ \hat{c}_{\alpha'} \ \} \ = \ 0 \qquad (934)$$

where the anti-commutator of two operators $\hat{A}$ and $\hat{B}$ are defined by

$$\{ \hat{A} , \hat{B} \} = \hat{A} \hat{B} + \hat{B} \hat{A} \tag{935}$$

These anti-commutation relations restricts the occupation numbers n_α for a single-particle state α to zero or unity, as required by the Pauli exclusion principle. This can be seen since the anti-commutation relation yields

$$\hat{c}_\alpha \, \hat{c}_\alpha \, \Phi_{n_{\alpha_1},\ldots}(\underline{r}_1,\ldots,\underline{r}_N) = - \hat{c}_\alpha \, \hat{c}_\alpha \, \Phi_{n_{\alpha_1},\ldots}(\underline{r}_1,\ldots,\underline{r}_N)$$
$$= 0 \tag{936}$$

which results in the identification

$$\sqrt{ (n_\alpha - 1) \, n_\alpha } \; \Phi_{n_{\alpha_1},\ldots,n_\alpha-2,\ldots}(\underline{r}_1,\ldots,\underline{r}_{N-2}) = 0 \tag{937}$$

so either $n_\alpha = 1$ or $n_\alpha = 0$. Incorporating this restriction into the definition of the creation operator yields

$$\hat{c}_\alpha^\dagger \, \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) = \Phi_{n_{\alpha_1},\ldots,n_\alpha+1,\ldots}(\underline{r}_1,\ldots,\underline{r}_{N+1}) \quad \text{for } n_\alpha = 0$$
$$= 0 \quad \text{for } n_\alpha = 1 \tag{938}$$

and similarly for the annihilation operator

$$\hat{c}_\alpha \, \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) = \Phi_{n_{\alpha_1},\ldots,n_\alpha-1,\ldots}(\underline{r}_1,\ldots,\underline{r}_{N-1}) \quad \text{for } n_\alpha = 1$$
$$= 0 \quad \text{for } n_\alpha = 0 \tag{939}$$

The occupation number operator can be defined as

$$\hat{n}_\alpha = \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \tag{940}$$

which is seen to have the eigenvalue equation

$$\hat{n}_\alpha \, \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) = n_\alpha \, \Phi_{n_{\alpha_1},\ldots,n_\alpha,\ldots}(\underline{r}_1,\ldots,\underline{r}_N) \tag{941}$$

similar to the case for bosons. The total number operator, $\hat{N}$, is then defined as

$$\hat{N} = \sum_\alpha \hat{n}_\alpha$$
$$= \sum_\alpha \hat{c}_\alpha^\dagger \, \hat{c}_\alpha \tag{942}$$

On defining the vacuum state as $| \, 0 \, >$ one can create any arbitrary basis state by the repeated action of the creation operators

$$| \, \Phi_{n_{\alpha_1},n_{\alpha_2}\ldots} \, > = \prod_\alpha \left\{ (\, c_\alpha^\dagger \,)^{n_\alpha} \right\} | \, 0 \, > \tag{943}$$

where the product runs over all the single-particle quantum numbers α.

Any operator can also be expressed in terms of the creation and annihilation operators. For example, the Hamiltonian for a system of interacting particles can be written as

$$\hat{H} = \sum_{\alpha;\alpha'} < \alpha' \mid \hat{H}_0 \mid \alpha > \hat{c}^{\dagger}_{\alpha'} \hat{c}_{\alpha} + \frac{1}{2!} \sum_{\alpha,\beta;\alpha',\beta'} < \alpha' \beta' \mid \hat{V}_{int} \mid \alpha \beta > \hat{c}^{\dagger}_{\alpha'} \hat{c}^{\dagger}_{\beta'} \hat{c}_{\beta} \hat{c}_{\alpha}$$

(944)

where

$$< \alpha' \mid \hat{H}_0 \mid \alpha > = \int d^3\underline{r} \, \phi^*_{\alpha'}(\underline{r}) \left[\frac{\hat{p}^2}{2\,m} + V_0(\underline{r}) \right] \phi_{\alpha}(\underline{r})$$

(945)

is the matrix elements of the single-particle energy and $< \alpha' \beta' \mid \hat{V}_{int} \mid \alpha \beta >$ represents the matrix elements of the two-body interaction between properly anti-symmetrized two-electron states. Since ordering is important for fermions, it is important to note that the order of the annihilation operators is in the reverse order of the corresponding single-particle quantum numbers in the matrix elements.

Coherent States

Since the occupation number is unrestricted, an unusual type of state is allowed for bosons. We shall focus our attention on one single-particle quantum level, and shall drop the index α which labels the level. A coherent state $\mid a_{\varphi} >$ is defined as an eigenstate of the annihilation operator[27]

$$\hat{a} \mid a_{\varphi} > = a_{\varphi} \mid a_{\varphi} >$$

(946)

where a_{φ} is a complex number. For example, the vacuum state or ground state is an eigenstate of the annihilation operator, in which case $a_{\varphi} = 0$.

The coherent state[28] can be found as a linear superposition of eigenstates of the number operator with eigenvalues n

$$\mid a_{\varphi} > = \sum_{n=0}^{\infty} C_n \mid n >$$

(947)

On substituting this form in the definition of the coherent state

$$\hat{a} \mid a_{\varphi} > = \sum_{n} C_n \, \hat{a} \mid n >$$

$$= a_{\varphi} \sum_{n} C_n \mid n >$$

(948)

[27] E. Schrödinger, "Der stetige Übergang von der Mikro- zur Makromechanik", Naturwissenschaften 14, 664 (1926).

[28] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field", Phys. Rev. **131**, 2766 (1963).

and using the property of the annihilation operator, one has

$$\sum_n C_n \sqrt{n} \mid n - 1 > \quad = \quad a_\varphi \sum_n C_n \mid n > \tag{949}$$

On taking the matrix elements of this equation with the state $< m \mid$, and using the orthonormality of the eigenstates of the number operator, one finds

$$C_{m+1} \sqrt{m + 1} \quad = \quad a_\varphi \, C_m \tag{950}$$

Hence, on iterating downwards, one finds

$$C_m \quad = \quad \left(\frac{a_\varphi^m}{\sqrt{m!}} \right) C_0 \tag{951}$$

and the coherent state can be expressed as

$$\mid a_\varphi > \quad = \quad C_0 \sum_{n=0}^{\infty} \left(\frac{a_\varphi^n}{\sqrt{n!}} \right) \mid n > \tag{952}$$

The normalization constant C_0 can be found from

$$1 \quad = \quad C_0^* \, C_0 \sum_{n=0}^{\infty} \left(\frac{a_\varphi^{n*} \, a_\varphi^n}{n!} \right) \tag{953}$$

by noting that the sum exponentiates to yield

$$1 \quad = \quad C_0^* \, C_0 \, \exp \left[\, a_\varphi^* \, a_\varphi \, \right] \tag{954}$$

so, on choosing the phase of C_0, one has

$$C_0 \quad = \quad \exp \left[\, - \, \frac{1}{2} \, a_\varphi^* \, a_\varphi \, \right] \tag{955}$$

From this, it can be shown that if the number of bosons in a coherent state are measured, the result n will occur with a probability given by

$$P(n) \quad = \quad \frac{(\, a_\varphi^* \, a_\varphi \,)^n}{n!} \, \exp \left[\, - \, a_\varphi^* \, a_\varphi \, \right] \tag{956}$$

Thus, the boson statistics are governed by a Poisson distribution. Furthermore, the quantity $a_\varphi^* \, a_\varphi$ is the average number of bosons $\bar{n}$ present in the coherent state.

The coherent states can be written in a more compact form. Since the state with occupation number n can be written as

$$\mid n > \quad = \quad \frac{(\, \hat{a}^\dagger \,)^n}{\sqrt{n!}} \mid 0 > \tag{957}$$

179

the coherent state can also be expressed as

$$
\mid a_\varphi \; > \; = \; \exp \left[\; - \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \; \sum_{n=0}^{\infty} \frac{(\, a_\varphi \, \hat{a}^\dagger \,)^n}{n!} \mid 0 \; > \tag{958}
$$

or on summing the series as an exponential

$$
\mid a_\varphi \; > \; = \; \exp \left[\; - \; \frac{1}{2} \, a_\varphi^* \, a_\varphi \; \right] \; \exp \left[\; a_\varphi \, \hat{a}^\dagger \; \right] \mid 0 \; > \tag{959}
$$

Thus the coherent state is an infinite linear superposition of states with different occupation numbers, each coefficient in the linear superposition has a specific phase relation with every other coefficient.

The above equation represents a transformation between number operator states and the coherent states. The inverse transformation can be found by expressing a_φ as a magnitude a and a phase φ

$$
a_\varphi \; = \; a \, \exp \left[\; i \, \varphi \; \right] \tag{960}
$$

The number states can be expressed in terms of the coherent states via the inverse transformation

$$
\mid n \; > \quad = \quad \frac{\sqrt{n!}}{a^n} \, \exp \left[\; + \; \frac{1}{2} \, a^2 \; \right] \; \int_0^{2\pi} \frac{d\varphi}{2\pi} \, \exp \left[\; - \; i \, n \, \varphi \; \right] \mid a_\varphi \; >
$$
$$
\tag{961}
$$

by integrating over the phase φ of the coherent state. Since the set of occupation number states is complete, the set of coherent states must also span Hilbert space. In fact, the set of coherent states is over-complete.

A number of systems do have states which have properties that closely resemble coherent states, such as the photon states in a laser or such as the superfluid condensate of He^4 at low temperatures. Although they do not have the precise mathematical form of the coherent states investigated by Glauber, the approximate states are characterized by having sufficiently large fluctuations in their occupation numbers so that the expectation value of the annihilation operator is well-defined.

8 Fermi-Dirac Statistics

8.1 Non-Interacting Fermions

An ideal gas of non-interacting fermions is described by the Hamiltonian $\hat{H}$ which is given by the sum

$$\hat{H} \; = \; \sum_\alpha \; \epsilon_\alpha \; \hat{n}_\alpha \tag{962}$$

where $\hat{n}_\alpha$ represents the occupation number of the α-th single-particle energy level. This is just the sum of the contributions from each particle, grouped according to the energy levels that they occupy. Likewise, the operator $\hat{N}$ representing the total number of particles is given by

$$\hat{N} \; = \; \sum_\alpha \; \hat{n}_\alpha \tag{963}$$

where the sum is over the single-particle energy levels. Hence, for non-interacting fermions, the density operator is diagonal in the occupation number representation.

The Grand-Canonical Partition function Ξ is given by

$$\Xi \; = \; \text{Trace} \; \exp\left[\; - \beta \; (\; \hat{H} \; - \; \mu \; \hat{N} \;) \; \right] \tag{964}$$

where the trace is over a complete set of microscopic states with variable N for the entire system. A convenient basis is given by the N-particle states $\mid \Phi_{n_{\alpha_1}, n_{\alpha_2}, \dots} >$, since both $\hat{H}$ and $\hat{N}$ are diagonal in this basis. The trace reduces to the sum of all configurations $\{ \; n_\alpha \; \}$, and since the total number of particles N is also being summed over, the trace is unrestricted. Therefore, the trace can be evaluated by summing over all possible values of the eigenvalues n_α for each consecutive value of α. That is

$$\text{Trace} \left\{ \; \dots \; \right\} \equiv \sum_{n_{\alpha_1}=0}^{1} \sum_{n_{\alpha_2}=0}^{1} \sum_{n_{\alpha_3}=0}^{1} \dots \left\{ \; \dots \; \right\} \tag{965}$$

Since the exponential term in Ξ reduces to the form of the exponential of a sum of independent terms, it can be written as the product of exponential factors

$$\exp\left[\; - \beta \; (\; \hat{H} \; - \; \mu \; \hat{N} \;) \; \right] \; = \; \exp\left[\; - \beta \; \sum_\alpha (\; \epsilon_\alpha \; - \; \mu \;) \; \hat{n}_\alpha \; \right]$$

$$= \; \prod_\alpha \exp\left[\; - \beta \; (\; \epsilon_\alpha \; - \; \mu \;) \; \hat{n}_\alpha \; \right] \tag{966}$$

Therefore, the trace in the expression for Ξ can be reduced to

$$\Xi \; = \; \prod_\alpha \left\{ \; \sum_{n_\alpha=0}^{1} \exp\left[\; - \beta \; (\; \epsilon_\alpha \; - \; \mu \;) \; n_\alpha \; \right] \; \right\} \tag{967}$$

where the sum is over all the occupation numbers n_α of either zero or unity as is allowed by Fermi-Dirac statistics. On performing the sum of the geometric series, one obtains

$$\sum_{n_\alpha=0}^{1} \exp\left[- \beta(\epsilon_\alpha - \mu)n_\alpha \right] = 1 + \exp[-\beta(\epsilon_\alpha - \mu)] \tag{968}$$

Therefore, the Grand-Canonical Partition Function is given by

$$\Xi = \prod_\alpha \left\{ 1 + \exp[-\beta(\epsilon_\alpha - \mu)] \right\} \tag{969}$$

The Grand-Canonical Potential Ω is found from

$$\Xi = \exp\left[- \beta \, \Omega \right] \tag{970}$$

which, on taking the logarithm, yields

$$- \beta \, \Omega = \sum_\alpha \ln \left[1 + \exp[-\beta(\epsilon_\alpha - \mu)] \right] \tag{971}$$

or

$$\Omega = - k_B \, T \sum_\alpha \ln \left[1 + \exp[-\beta(\epsilon_\alpha - \mu)] \right] \tag{972}$$

On introducing the density of single-particle states $\rho(\epsilon)$ defined as

$$\rho(\epsilon) = \sum_\alpha \delta(\epsilon - \epsilon_\alpha) \tag{973}$$

the summation can be transformed into an integral over ϵ

$$\Omega = - k_B \, T \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \ln \left[1 + \exp[-\beta(\epsilon - \mu)] \right] \tag{974}$$

where the density of states is bounded from below by ϵ_0. After evaluating the integral, one may find all the thermodynamic properties of the system from Ω.

8.2   The Fermi-Dirac Distribution Function

The average number of particles in the system $\overline{N}$ can be determined from Ω via the relation

$$\overline{N} = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} \tag{975}$$

which is evaluated as

$$\begin{aligned} \overline{N} &= k_B \, T \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\beta \, \exp[-\beta(\epsilon - \mu)]}{1 + \exp[-\beta(\epsilon - \mu)]} \\ &= \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{1}{\exp[\beta(\epsilon - \mu)] + 1} \end{aligned} \tag{976}$$

182

or equivalently

$$\overline{N} \;=\; \sum_\alpha \frac{1}{\exp[\beta(\epsilon_\alpha - \mu)] \;+\; 1} \tag{977}$$

but, by definition, one has

$$\overline{N} \;=\; \sum_\alpha \overline{n}_\alpha \tag{978}$$

Therefore, the function $f(\epsilon_\alpha)$ defined by

$$f(\epsilon_\alpha) \;=\; \frac{1}{\exp[\beta(\epsilon_\alpha - \mu)] \;+\; 1} \tag{979}$$

represents the average number of particles $\overline{n}_\alpha$ in a quantum level with a single-particle energy ϵ_α. The function $f(\epsilon)$ is the Fermi-Dirac distribution function. The Fermi-Dirac distribution vanishes as

$$f(\epsilon) \;\approx\; \exp\left[\,-\, \beta \,(\, \epsilon \,-\, \mu \,) \right] \tag{980}$$

for $\epsilon - \mu > k_B \, T$. However, for $\mu - \epsilon > k_B \, T$, the Fermi-Dirac distribution function tends to unity

$$f(\epsilon) \;\approx\; 1 \;-\; \exp\left[\beta \,(\, \epsilon \,-\, \mu \,) \right] \tag{981}$$

and falls off rapidly from 1 to 0 at $\epsilon \;=\; \mu$ where it takes on the value

$$f(\mu) \;=\; \frac{1}{2} \tag{982}$$

The range of ϵ over which the function differs from either 1 or 0 is governed by
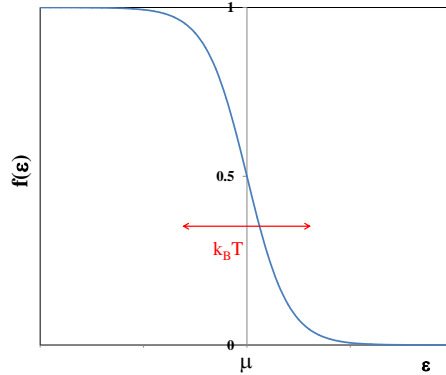
Figure 39: The Fermi-Dirac Distribution function.

$k_B \, T$.
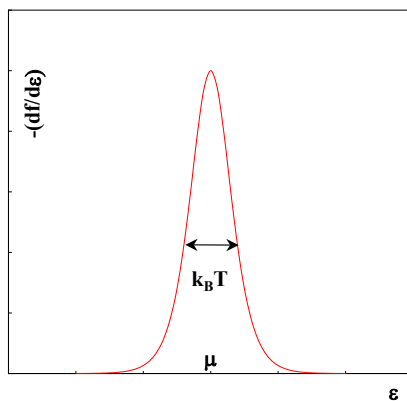
Figure 40: The negative energy derivative of Fermi-Dirac distribution function.

At zero temperature, all the states with single-particle energies less than μ are occupied and all states with energies greater than μ are empty. The value of $\mu(T = 0)$ is called the Fermi-energy and is denoted by ϵ_F. Note that since

$$\int_{-\infty}^{\infty} d\epsilon \; \frac{\partial f(\epsilon)}{\partial \epsilon} \; = \; -1 \tag{983}$$

and as $T \to 0$ one has

$$\frac{\partial f(\epsilon)}{\partial \epsilon} \; \approx \; 0 \qquad \text{if } |\, \epsilon \, - \, \mu \,| \; > \; k_B \, T \tag{984}$$

then

$$-\frac{\partial f(\epsilon)}{\partial \epsilon} \; \sim \; \delta(\, \epsilon \, - \, \mu \,) \tag{985}$$

as it resembles the Dirac delta function.

Thermodynamic Properties

The thermodynamic energy U coincides with the average energy $\overline{E}$. The thermodynamic energy can be obtained from the Grand-Canonical Potential Ω via its definition

$$\Omega \; = \; U \; - \; T \, S \; - \; \mu \, \overline{N} \tag{986}$$

which, together with

$$S \; = \; -\left(\frac{\partial \Omega}{\partial T}\right)_{\mu, V}$$

$$\overline{N} \; = \; -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T, V} \tag{987}$$

184

can be inverted to yield

$$
\begin{aligned}
U &= \Omega + T\,S + \mu\,\overline{N} \\
&= \Omega - T\left(\frac{\partial \Omega}{\partial T}\right)_{\mu,V} - \mu\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V}
\end{aligned}
\tag{988}
$$

On substituting

$$
\Omega = -\,k_B\,T\int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\ln\Big[\,1 + \exp[-\beta(\epsilon-\mu)]\,\Big]
\tag{989}
$$

one obtains

$$
\begin{aligned}
U &= k_B\,T^2\int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\frac{(\epsilon-\mu)\left(-\frac{\partial\beta}{\partial T}\right)\exp[-\beta(\epsilon-\mu)]}{1 + \exp[-\beta(\epsilon-\mu)]} \\
&\quad + \mu\,k_B\,T\int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\frac{\beta\,\exp[-\beta(\epsilon-\mu)]}{1 + \exp[-\beta(\epsilon-\mu)]} \\
&= \int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\frac{\epsilon}{\exp[\beta(\epsilon-\mu)] + 1} \\
&= \int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\epsilon\,f(\epsilon) \\
&= \sum_{\alpha} \epsilon_{\alpha}\,f(\epsilon_{\alpha}) \\
&= \sum_{\alpha} \epsilon_{\alpha}\,\overline{n}_{\alpha}
\end{aligned}
\tag{990}
$$

which shows that the thermodynamic energy for a system of particles is just the average energy, since the average energy of a system of non-interacting particles is just the sum over the average energies for each particle. This reinforces the interpretation of $f(\epsilon)$ as the average number of fermions in a single-particle state with energy ϵ.

The entropy S is determined from the equation

$$
S = -\left(\frac{\partial \Omega}{\partial T}\right)_{\mu,V}
\tag{991}
$$

which yields

$$
\begin{aligned}
S &= k_B\sum_{\alpha}\ln\Big[\,1 + \exp[\beta(\mu-\epsilon_{\alpha})]\,\Big] \\
&\quad - \sum_{\alpha}\frac{(\mu-\epsilon_{\alpha})}{T}\,\frac{\exp[\beta(\mu-\epsilon_{\alpha})]}{1 + \exp[\beta(\mu-\epsilon_{\alpha})]} \\
&= -\,k_B\sum_{\alpha}\ln\Big[\,1 - f(\epsilon_{\alpha})\,\Big]
\end{aligned}
$$

185

$$+ \; k_B \; \sum_\alpha \; \frac{(\mu - \epsilon_\alpha)}{T} \; f(\epsilon_\alpha) \tag{992}$$

However, one may rewrite the factor

$$\frac{(\mu - \epsilon_\alpha)}{k_B \; T} \tag{993}$$

as

$$
\begin{aligned}
\frac{(\mu - \epsilon_\alpha)}{k_B \; T} \;&=\; \ln \left[\; \exp[\beta(\mu - \epsilon_\alpha)] \; \right] \\
&=\; - \; \ln \left[\; 1 \; - \; f(\epsilon_\alpha) \; \right] \; + \; \ln \left[\; f(\epsilon_\alpha) \; \right]
\end{aligned}
\tag{994}
$$

Therefore, on combining the above expressions one finds that the entropy of the non-interacting fermion gas can be expressed as

$$S \;=\; - \; k_B \; \sum_\alpha \; \left(\; (\, 1 \, - \, f(\epsilon_\alpha) \,) \; \ln \left[\; 1 \, - \, f(\epsilon_\alpha) \; \right] \; + \; f(\epsilon_\alpha) \; \ln \left[\; f(\epsilon_\alpha) \; \right] \; \right) \tag{995}$$

The entropy has the form of

$$S \;=\; - \; k_B \; \sum_\gamma \; p_\gamma \; \ln \; p_\gamma \tag{996}$$

where $p_\gamma \;=\; f(\epsilon_\alpha)$ is the probability that the α-th level is occupied and $p_\gamma \;=\; 1 \; - \; f(\epsilon_\alpha)$ is the probability that the level is empty. This form of the entropy follows naturally from the assumption that the non-interacting particles are statistically independent together with the definition of the entropy.

8.3   The Equation of State

The equation of state for a gas of non-interacting fermions can be found from Ω by noting that

$$\Omega \;=\; - \; P \; V \tag{997}$$

The equation of state can be obtained directly when the single-particle density of states $\rho(\epsilon)$ has the form of a simple power law

$$
\begin{aligned}
\rho(\epsilon) \;&=\; C \; \epsilon^\alpha \qquad \text{for } \epsilon \; \geq \; 0 \\
&=\; 0 \qquad \text{otherwise}
\end{aligned}
\tag{998}
$$

where α is a constant. The Grand-Canonical Potential Ω is found as

$$
\begin{aligned}
\Omega \;&=\; - \; P \; V \\
&=\; - \; k_B \; T \; \int_{-\infty}^{\infty} d\epsilon \; \rho(\epsilon) \; \ln \left[\; 1 \; + \; \exp[-\beta(\epsilon - \mu)] \; \right]
\end{aligned}
\tag{999}
$$

186

Hence, one has

$$
\begin{aligned}
\frac{P\,V}{k_B\,T} \;&=\; C\,\int_0^\infty d\epsilon\;\epsilon^\alpha\;\ln\left[\,1\;+\;\exp[-\beta(\epsilon-\mu)]\,\right] \\
&=\; \frac{C}{\alpha\,+\,1}\,\int_0^\infty d\epsilon\;\frac{d\epsilon^{(\alpha+1)}}{d\epsilon}\;\ln\left[\,1\;+\;\exp[-\beta(\epsilon-\mu)]\,\right] \quad (1000)
\end{aligned}
$$

On integrating by parts, one obtains

$$
\begin{aligned}
\frac{P\,V}{k_B\,T} \;&=\; \frac{C}{\alpha\,+\,1}\left[\,\epsilon^{\alpha+1}\;\ln\left(\,1\;+\;\exp[-\beta(\epsilon-\mu)]\,\right)\,\right]\Big|_0^\infty \\
&\quad +\; \frac{C}{\alpha\,+\,1}\,\int_0^\infty d\epsilon\;\epsilon^{(\alpha+1)}\;\frac{\beta\;\exp[-\beta(\epsilon-\mu)]}{1\,+\,\exp[-\beta(\epsilon-\mu)]} \quad (1001)
\end{aligned}
$$

The boundary terms vanish, since the density of states vanishes at the lower limit of integration $\epsilon = 0$ and the logarithmic factor vanishes exponentially when $\epsilon \to \infty$. Thus, on canceling a factor of β, one finds

$$
\begin{aligned}
P\,V \;&=\; \frac{1}{\alpha\,+\,1}\,\int_0^\infty d\epsilon\;C\;\epsilon^{(\alpha+1)}\;\frac{1}{\exp[\beta(\epsilon-\mu)]\,+\,1} \\
&=\; \frac{1}{\alpha\,+\,1}\,\int_{-\infty}^\infty d\epsilon\;\rho(\epsilon)\;\epsilon\;\frac{1}{\exp[\beta(\epsilon-\mu)]\,+\,1} \\
&=\; \frac{U}{\alpha\,+\,1} \quad\quad\quad (1002)
\end{aligned}
$$

That is, the equation of state for the system of non-interacting fermions is found as

$$
P\,V \;=\; \frac{U}{\alpha\,+\,1} \quad\quad\quad (1003)
$$

For $\alpha = \frac{1}{2}$, the relation is identical to that found for the classical ideal gas.

In fact, the high-temperature limit of the equation of state for the system of non-interacting bosons can be evaluated as

$$
\begin{aligned}
\Omega \;&=\; -\,P\,V \\
&=\; -\,k_B\,T\,\int_0^\infty d\epsilon\;\rho(\epsilon)\;\ln\left[\,1\;+\;\exp[-\beta(\epsilon-\mu)]\,\right] \\
&\approx\; -\,k_B\,T\,\int_0^\infty d\epsilon\;\rho(\epsilon)\;\exp[-\beta(\epsilon-\mu)] \\
&\approx\; -\,k_B\,T\,\int_0^\infty d\epsilon\;\rho(\epsilon)\;\frac{1}{\exp[\beta(\epsilon-\mu)]\,+\,1} \\
&\approx\; -\,\overline{N}\;k_B\,T \quad\quad\quad (1004)
\end{aligned}
$$

since at high-temperatures $\mu < 0$ and since we have assumed that the single-particle density of states is zero below $\epsilon = 0$. Therefore, we have re-derived the ideal gas law from the high temperature limit of a set of non-interacting particles which obey Fermi-Dirac statistics.

8.4 The Chemical Potential

We have seen that, at high temperatures, the equation of state of a gas of particles obeying Fermi-Dirac statistics reduces to the equation of state for particles obeying Classical statistics if

$$\exp\left[\beta\,\mu\right] \ll 1 \tag{1005}$$

(where we have restricted $\epsilon \geq 0$) since under these conditions the Fermi-Dirac distribution function reduces to the Boltzmann distribution function. The above restriction is also consistent with our previous discussion of the ideal gas, where

$$\exp\left[\beta\,\mu(T)\right] = \left(\frac{V}{\lambda^3}\right) \tag{1006}$$

We shall examine the temperature dependence of the chemical potential and show that Fermi-Dirac statistics is similar to classical Maxwell-Boltzmann statistics at sufficiently high temperatures.

For a system with a fixed number of fermions, $\overline{N}$, governed by a condition like electrical neutrality, then the chemical potential is temperature dependent and is found as the solution for $\mu(T)$ found from the equation

$$\overline{N} = \int_0^\infty d\epsilon\,\rho(\epsilon)\,f(\epsilon) \tag{1007}$$

For large and negative values of $\mu(T)$, one can expand the Fermi-Dirac distribution function

$$
\begin{aligned}
\overline{N} &= \int_0^\infty d\epsilon\,\rho(\epsilon)\,\frac{\exp[-\beta(\epsilon-\mu)]}{1 + \exp[-\beta(\epsilon-\mu)]} \\
&= \int_0^\infty d\epsilon\,\rho(\epsilon)\,\exp[-\beta(\epsilon-\mu)]\sum_{n=0}^\infty (-1)^n\,\exp[-n\beta(\epsilon-\mu)] \\
&= \int_0^\infty d\epsilon\,\rho(\epsilon)\sum_{n=1}^\infty (-1)^{n+1}\,\exp[n\beta(\mu-\epsilon)] \\
&= \sum_{n=1}^\infty (-1)^{n+1}\,\exp[n\beta\mu]\int_0^\infty d\epsilon\,\rho(\epsilon)\,\exp[-n\beta\epsilon] \tag{1008}
\end{aligned}
$$

On substituting for the chemical potential in terms of the fugacity z

$$z = \exp[\beta\mu] \tag{1009}$$

and on introducing an expression for the density of states

$$\rho(\epsilon) = C\,\epsilon^\alpha \tag{1010}$$

one has

$$\overline{N} \;=\; C \sum_{n=1}^{\infty} (-1)^{n+1} \; z^n \; \int_0^{\infty} d\epsilon \; \epsilon^\alpha \; \exp[-n\beta\epsilon] \qquad (1011)$$

On changing the variable of integration from ϵ to the dimensionless variable x where x is defined as

$$x \;=\; n\beta\epsilon \qquad (1012)$$

the expression for the average number of particles has the form

$$
\begin{aligned}
\overline{N} \;&=\; C \; (\; k_B \; T \;)^{\alpha+1} \int_0^{\infty} dx \; x^\alpha \; \exp[-x] \sum_{n=1}^{\infty} (-1)^{n+1} \; \frac{z^n}{n^{\alpha+1}} \\
&=\; C \; (\; k_B \; T \;)^{\alpha+1} \; \Gamma(\alpha+1) \sum_{n=1}^{\infty} (-1)^{n+1} \; \frac{z^n}{n^{\alpha+1}} \qquad (1013)
\end{aligned}
$$

where $\Gamma(x)$ is the Gamma function. Since one can re-write the above equation as

$$\left(\frac{\overline{N}}{\Gamma(\alpha+1) \; C \; (\; k_B \; T \;)^{\alpha+1}} \right) \;=\; \sum_{n=1}^{\infty} (-1)^{n+1} \; \frac{z^n}{n^{\alpha+1}} \qquad (1014)$$

one can see that, for fixed $\overline{N}$ and high T, z must be small. In the case where $z \ll 1$, which occurs for sufficiently high temperatures, one may only retain the first term in the power series for z. This leads to the solution for z and, hence, the chemical potential

$$z \;\approx\; \left(\frac{\overline{N}}{\Gamma(\alpha+1) \; C \; (\; k_B \; T \;)^{\alpha+1}} \right) \qquad (1015)$$

or alternatively

$$\mu(T) \;\approx\; - \; k_B \; T \; \ln \left[\frac{\Gamma(\alpha+1) \; C \; (\; k_B \; T \;)^{\alpha+1}}{\overline{N}} \right] \qquad (1016)$$

which illustrates that $\mu(T)$ must be temperature dependent. Furthermore, we see that at sufficiently high temperatures, only the first term in the expansion of the Fermi-Dirac distribution contributes. In this limit, Fermi-Dirac statistics reduces to classical Maxwell-Boltzmann statistics. We shall see later that something similar happens in the high temperature limit with Bose-Einstein statistics.

One also sees that on decreasing the temperature downwards, starting from the high temperature limit, then z increases. Therefore, one must retain an increasing number of terms in the expansion

$$\left(\frac{\overline{N}}{\Gamma(\alpha+1) \; C \; (\; k_B \; T \;)^{\alpha+1}} \right) \;=\; \sum_{n=1}^{\infty} (-1)^{n+1} \; \frac{z^n}{n^{\alpha+1}} \qquad (1017)$$

if one wants to determine the chemical potential accurately for lower temperatures. The reversion of the series is only practical if $z \leq 1$, above which the Fermi-Dirac gas is said to be non-degenerate. For temperatures below the degeneracy temperature, at which $\mu = 0$ and therefore $z = 1$, the chemical potential must be found by other methods.

8.5   The Sommerfeld Expansion

Many thermodynamic properties of a system of electrons can be written in the form

$$A(T) = \int_{-\infty}^{\infty} d\epsilon \; \Phi(\epsilon) \; f(\epsilon) \tag{1018}$$

where $\Phi(\epsilon)$ is some function of the single-particle energy ϵ multiplied by the single-particle density of states $\rho(\epsilon)$ and $f(\epsilon)$ is the Fermi-function.

Integrals of this form can be evaluated very accurately, if the temperature T is of the order of room temperature $T \sim 300$ ÅK, which corresponds to an energy scale of

$$k_B \; T \sim \frac{1}{40} \quad \text{eV} \tag{1019}$$

and if the typical energy scale for $\Phi(\epsilon)$ is between 1 and 10 eV. Typical electronic scales are given by the binding energy of an electron in a Hydrogen atom ~ 13.6 eV or the total band width in a transition metal which is about 10 eV. In Na, the chemical potential μ measured from the bottom of the valence band density of states is about 3 eV and is about 12 eV for Al. Clearly, under ambient conditions in a metal one has

$$\mu \gg k_B \; T \tag{1020}$$

so a metal can usually be thought as being below its degeneracy temperature.

The Sommerfeld expansion[29] expresses integrals of the form

$$A(T) = \int_{-\infty}^{\infty} d\epsilon \; \Phi(\epsilon) \; f(\epsilon) \tag{1021}$$

in terms of a sum of the $T = 0$ limit of the integral and a power series of $k_B \; T/\mu$.

As a first approximation, one can estimate $A(T)$ as

$$A(T) \sim \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; + \; 0 \tag{1022}$$

[29] A. Sommerfeld, "Zur Elektronentheorie der Metalle auf Grund der Fermischen Statistik". Zeitschrift für Physik 47, 1-3, (1928).

since at $T = 0$, one can write the Fermi-function as

$$
\begin{aligned}
f(\epsilon) &= 1 && \text{for } \epsilon < \mu(T) \\
f(\epsilon) &= 0 && \text{for } \epsilon > \mu(T)
\end{aligned}
\tag{1023}
$$

We would like to obtain a better approximation, which reflects the temperature dependence of the Fermi-function $f(\epsilon)$. A better approximation for $A(T)$ can be obtained by re-writing the exact expression for $A(T)$ as

$$
\begin{aligned}
A(T) &= \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; + \; \int_{\mu(T)}^{\infty} d\epsilon \; 0 \\
&+ \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; (\; f(\epsilon) \; - \; 1 \;) \; + \; \int_{\mu(T)}^{\infty} d\epsilon \; \Phi(\epsilon) \; (\; f(\epsilon) \; - \; 0 \;)
\end{aligned}
\tag{1024}
$$

In this we have included the exact corrections to the $T = 0$ approximation to each region of the integral. This is evaluated as

$$
\begin{aligned}
A(T) &= \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; + \; \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \left(\; \frac{1}{\exp[\beta(\epsilon - \mu)] \; + \; 1} \; - \; 1 \; \right) \\
&+ \int_{\mu(T)}^{\infty} d\epsilon \; \Phi(\epsilon) \; \frac{1}{\exp[\beta(\epsilon - \mu)] \; + \; 1} \\
&= \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; - \; \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; \frac{1}{\exp[-\beta(\epsilon - \mu)] \; + \; 1} \\
&+ \int_{\mu(T)}^{\infty} d\epsilon \; \Phi(\epsilon) \; \frac{1}{\exp[\beta(\epsilon - \mu)] \; + \; 1}
\end{aligned}
\tag{1025}
$$

where we have substituted the identity

$$
\begin{aligned}
\frac{1}{\exp[\beta(\epsilon - \mu)] \; + \; 1} \; - \; 1 &= \; - \; \frac{\exp[\beta(\epsilon - \mu)]}{\exp[\beta(\epsilon - \mu)] \; + \; 1} \\
&= \; - \; \frac{1}{1 \; + \; \exp[-\beta(\epsilon - \mu)]}
\end{aligned}
\tag{1026}
$$

in the second term of the first line. The two temperature-dependent correction terms in $A(T)$ involve a function of the form

$$
\frac{1}{\exp[\; x \;] \; + \; 1}
\tag{1027}
$$

We shall set $x = \beta \; (\; \epsilon \; - \; \mu(T) \;)$ or

$$
\epsilon = \mu(T) \; + \; k_B \; T \; x
\tag{1028}
$$

in the second correction term which yields

$$
\begin{aligned}
A(T) &= \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; - \; \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \; \frac{1}{\exp[-\beta(\epsilon - \mu)] \; + \; 1} \\
&+ k_B \; T \; \int_{0}^{\infty} dx \; \Phi(\mu + k_B T x) \; \frac{1}{\exp[x] \; + \; 1}
\end{aligned}
\tag{1029}
$$

191

The first correction term can be expressed in terms of the variable $y = -\beta (\epsilon - \mu(T))$ or

$$\epsilon = \mu(T) - k_B T y \qquad (1030)$$

and the boundaries of the integration over y run from 0 to ∞.

$$A(T) = \int_{-\infty}^{\mu(T)} d\epsilon \ \Phi(\epsilon) + k_B T \int_{\infty}^{0} dy \ \Phi(\mu - k_B T y) \ \frac{1}{\exp[y] + 1}$$
$$+ k_B T \int_{0}^{\infty} dx \ \Phi(\mu + k_B T x) \ \frac{1}{\exp[x] + 1} \qquad (1031)$$

Except for the terms $+k_B T x$ and $-k_B T y$ in the arguments of the function Φ, the correction terms would cancel and vanish. On changing the integration variable from y to x in the second term, the integrals can be combined as

$$A(T) = \int_{-\infty}^{\mu(T)} d\epsilon \ \Phi(\epsilon)$$
$$+ k_B T \int_{0}^{\infty} dx \left[\Phi(\mu + k_B T x) - \Phi(\mu - k_B T x) \right] \frac{1}{\exp[x] + 1} \qquad (1032)$$

On Taylor expanding the terms in the large square parenthesis, one finds that only the odd terms in $k_B T x$ survive.

$$A(T) = \int_{-\infty}^{\mu(T)} d\epsilon \ \Phi(\epsilon)$$
$$+ 2 k_B T \int_{0}^{\infty} dx \left[\sum_{n=0}^{\infty} \left. \frac{\partial^{2n+1} \Phi(\epsilon)}{\partial \epsilon^{2n+1}} \right|_{\mu(T)} \frac{(k_B T x)^{2n+1}}{(2n+1)!} \right] \frac{1}{\exp[x] + 1} \qquad (1033)$$

On interchanging the order of the summation and the integration, one obtains

$$A(T) = \int_{-\infty}^{\mu(T)} d\epsilon \ \Phi(\epsilon)$$
$$+ 2 \sum_{n=0}^{\infty} \frac{(k_B T)^{2n+2}}{(2n+1)!} \left. \frac{\partial^{2n+1} \Phi(\epsilon)}{\partial \epsilon^{2n+1}} \right|_{\mu(T)} \int_{0}^{\infty} dx \ \frac{x^{2n+1}}{\exp[x] + 1} \qquad (1034)$$

where all the derivatives of Φ are to be evaluated at $\mu(T)$. The integrals over x are convergent. One should note that the power series only contains terms of even powers in $k_B T$. Since the derivatives of Φ such as

$$\frac{\partial^{2n+1} \Phi(\epsilon)}{\partial \epsilon^{2n+1}} \qquad (1035)$$

192

have the dimensions of Φ/μ^{2n+1}, one might think of this expansion as being in powers of the dimensionless quantity

$$\left(\frac{k_B\ T}{\mu} \right)^2 \tag{1036}$$

which is assumed to be much smaller than unity. Therefore, the series could be expected to be rapidly convergent.

The integral

$$\int_0^\infty dx\ \frac{x^{2n+1}}{\exp[x]\ +\ 1} \tag{1037}$$

is convergent since it is just the area under the curve that varies as x^{2n+1} for small x and vanishes exponentially as $\exp[-x]\ x^{2n+1}$ for large x. The integral
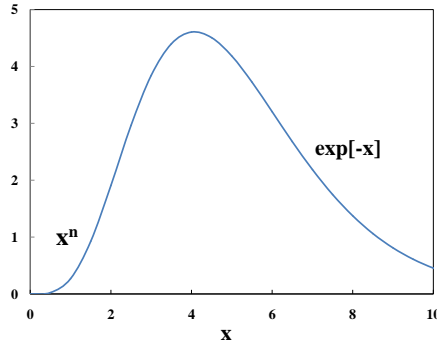
Figure 41: The integrand $\frac{x^n}{\exp[x]+1}$.

can be evaluated by considering I_m given by

$$I_m\ =\ \int_0^\infty dx\ \frac{x^m}{\exp[x]\ +\ 1} \tag{1038}$$

and noting that since $x\ >\ 0$ then $\exp[-x]\ <\ 1$. Therefore, by rewriting the integral as

$$\begin{aligned} I_m\ &=\ \int_0^\infty dx\ \frac{x^m}{\exp[x]\ (\ 1\ +\ \exp[-x]\ )} \\ &=\ \int_0^\infty dx\ \frac{x^m\ \exp[-x]}{(\ 1\ +\ \exp[-x]\ )} \end{aligned} \tag{1039}$$

one can expand the integral in powers of $\exp[-x]$

$$\begin{aligned} I_m\ &=\ \int_0^\infty dx\ x^m\ \exp[-x]\ \sum_{l=0}^\infty (\ -\ 1\ )^l\ \exp[-lx] \\ &=\ \sum_{l=0}^\infty (\ -\ 1\ )^l\ \int_0^\infty dx\ x^m\ \exp[-(l+1)x] \end{aligned} \tag{1040}$$

193

On changing the variable of integration from x to y where

$$y = (l + 1) x \tag{1041}$$

one has

$$I_m = \sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)^{m+1}} \int_0^{\infty} dy \; y^m \; \exp[-y] \tag{1042}$$

The integral $\int_0^{\infty} dy \; y^m \; \exp[-y]$ can be evaluated by successive integration by parts. That is,

$$
\begin{aligned}
\int_0^{\infty} dy \; y^m \; \exp[-y] &= - \int_0^{\infty} dy \; y^m \; \frac{\partial}{\partial y} \exp[-y] \\
&= - y^m \; \exp[-y] \Big|_0^{\infty} + \int_0^{\infty} dy \; m \; y^{m-1} \; \exp[-y]
\end{aligned}
\tag{1043}
$$

The boundary term vanishes like y^m near $y = 0$ and vanishes like $\exp[-y]$ when $y \to \infty$. Therefore,

$$
\begin{aligned}
\int_0^{\infty} dy \; y^m &= m \int_0^{\infty} dy \; y^{m-1} \; \exp[-y] \\
&= m! \int_0^{\infty} dy \; \exp[-y] \\
&= m!
\end{aligned}
\tag{1044}
$$

Thus, we have

$$I_m = m! \sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)^{m+1}} \tag{1045}$$

However, since the Riemann zeta function ζ is defined by

$$\zeta(m+1) = \sum_{l=0}^{\infty} \frac{1}{(l+1)^{m+1}} \tag{1046}$$

one has

$$I_m = m! \left(1 - \frac{2}{2^{m+1}} \right) \zeta(m+1) \tag{1047}$$

Therefore, the Sommerfeld expansion takes the form

$$
\begin{aligned}
A(T) &= \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \\
&+ 2 \sum_{n=0}^{\infty} \left(1 - \frac{1}{2^{(2n+1)}} \right) \zeta(2(n+1)) \; (k_B T)^{2n+2} \left. \frac{\partial^{2n+1} \Phi(\epsilon)}{\partial \epsilon^{2n+1}} \right|_{\mu(T)}
\end{aligned}
\tag{1048}
$$

194

which can be written as

$$
\begin{aligned}
A(T) \;&=\; \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \\[4pt]
&\quad + 2 \sum_{n=1}^{\infty} \left(1 - \frac{1}{2^{(2n-1)}} \right) \zeta(2n) \; (\, k_B \, T \,)^{2n} \; \left. \frac{\partial^{2n-1}\Phi(\epsilon)}{\partial \epsilon^{2n-1}} \right|_{\mu(T)}
\end{aligned}
\tag{1049}
$$

where $\mu(T)$ is the value of the chemical potential at temperature T and the Riemann ζ function has the values

$$
\begin{aligned}
\zeta(2) \;&=\; \frac{\pi^2}{6} \\[4pt]
\zeta(4) \;&=\; \frac{\pi^4}{90} \\[4pt]
\zeta(6) \;&=\; \frac{\pi^6}{945}
\end{aligned}
\tag{1050}
$$

etc. Thus, at sufficiently low temperatures, one expects that one might be able to approximate $A(T)$ by the Sommerfeld expansion

$$
\begin{aligned}
A(T) \;&\approx\; \int_{-\infty}^{\mu(T)} d\epsilon \; \Phi(\epsilon) \\[4pt]
&\quad + \frac{\pi^2}{6} \, (\, k_B \, T \,)^2 \; \left. \frac{\partial \Phi(\epsilon)}{\partial \epsilon} \right|_{\mu(T)} + \; \dots
\end{aligned}
\tag{1051}
$$

8.6 The Low-Temperature Specific Heat of an Electron Gas

The condition of electrical neutrality determines the number of electrons in a metal and keeps the number constant. The number of electrons $\overline{N}$ is given by

$$
\overline{N} \;=\; \int_{-\infty}^{\infty} d\epsilon \; \rho(\epsilon) \; f(\epsilon)
\tag{1052}
$$

which can be approximated by the first few terms in the Sommerfeld expansion

$$
\overline{N} \;=\; \int_{-\infty}^{\mu(T)} d\epsilon \; \rho(\epsilon) \;+\; \frac{\pi^2}{6} \, (\, k_B \, T \,)^2 \; \left. \frac{\partial \rho}{\partial \epsilon} \right|_{\mu} + \; \dots
\tag{1053}
$$

This yields the temperature dependence of $\mu(T)$. Since $\overline{N}$ is independent of temperature

$$
\frac{\partial N}{\partial T} \;=\; 0
$$

195

$$= \frac{\partial}{\partial T} \int_{-\infty}^{\mu(T)} d\epsilon \, \rho(\epsilon) \; + \; \frac{2 \, \pi^2}{6} \, k_B \, (\, k_B \, T \,) \, \frac{\partial \rho}{\partial \epsilon} \bigg|_{\mu} \; + \; \dots$$

$$0 \;=\; \frac{\partial \mu}{\partial T} \, \rho(\mu) \; + \; \frac{2 \, \pi^2}{6} \, k_B \, (\, k_B \, T \,) \, \frac{\partial \rho}{\partial \epsilon} \bigg|_{\mu} \; + \; \dots \tag{1054}$$

Thus, we have found that

$$\frac{\partial \mu}{\partial T} \;=\; - \, \frac{2 \, \pi^2}{6} \, k_B \, (\, k_B \, T \,) \, \frac{\frac{\partial \rho}{\partial \epsilon}}{\rho} \bigg|_{\mu} \tag{1055}$$

which implies that the derivative of μ w.r.t. T has the opposite sign to $\frac{\partial rho}{\partial \epsilon}$. Thus, if $\mu(T = 0)$ is just below a peak in $\rho(\epsilon)$ then the integral expression for N runs over a range of ϵ which avoids the peak. If μ did not decrease with increasing T, then at finite temperatures, the peak when multiplied by the tail of the Fermi-function could give an extra contribution to $\overline{N}$. This increase must
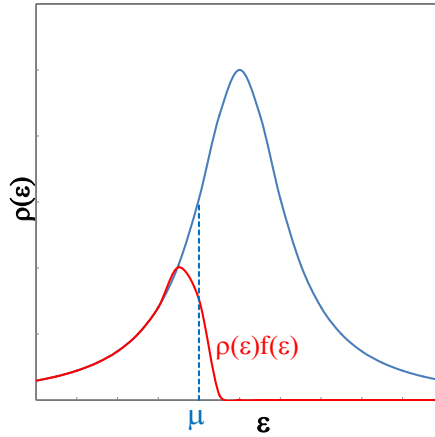
Figure 42: The density of states $\rho(\epsilon)$ and the density of states weighted by the Fermi-Dirac distribution function $\rho(\epsilon) f(\epsilon)$.

be offset by moving $\mu(T)$ down from ϵ_F. so the contribution from the tail at the peak is smaller and is offset by the smaller area under the curve to $\mu(T)$. Similar reasoning applies for the increase in $\mu(T)$ if ϵ_F is located just above a peak in $\rho(\epsilon)$. However, if ϵ_F is located at the top of a symmetric peak in the density of states, then the chemical potential should not depend on temperature.

The internal energy can also be expressed as

$$\overline{E} \;=\; \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \epsilon \, f(\epsilon) \tag{1056}$$

which can be approximated by the first few terms in the Sommerfeld expansion

$$\overline{E} = \int_{-\infty}^{\mu(T)} d\epsilon \, \rho(\epsilon) \, \epsilon \; + \; \frac{\pi^2}{6} \, (\, k_B \, T \,)^2 \left. \frac{\partial(\, \epsilon \, \rho \,)}{\partial \epsilon} \right|_{\mu} \; + \; \ldots \tag{1057}$$

The specific heat is given by the temperature derivative of the internal energy at fixed V

$$
\begin{aligned}
C_V \;=\;& \left(\frac{\partial \overline{E}}{\partial T} \right)_V \\
=\;& \frac{\partial}{\partial T} \int_{-\infty}^{\mu(T)} d\epsilon \, \epsilon \, \rho(\epsilon) \; + \; \frac{2 \, \pi^2}{6} \, k_B \, (\, k_B \, T \,) \left. \frac{\partial(\, \epsilon \, \rho \,)}{\partial \epsilon} \right|_{\mu} \; + \; \ldots \\
=\;& \frac{\partial \mu}{\partial T} \, \mu \, \rho(\mu) \; + \; \frac{2 \, \pi^2}{6} \, k_B \, (\, k_B \, T \,) \left. \frac{\partial(\, \epsilon \, \rho \,)}{\partial \epsilon} \right|_{\mu} \; + \; \ldots \tag{1058}
\end{aligned}
$$

On substituting for $\frac{\partial \mu}{\partial T}$, one finds

$$
\begin{aligned}
C_V \;=\;& - \, \frac{\pi^2}{3} \, k_B \, (\, k_B \, T \,) \, \mu \left. \frac{\partial \rho}{\partial \epsilon} \right|_{\mu} \; + \; \frac{2 \, \pi^2}{6} \, k_B \, (\, k_B \, T \,) \left. \frac{\partial(\, \epsilon \, \rho \,)}{\partial \epsilon} \right|_{\mu} \; + \; \ldots \\
=\;& \frac{\pi^2}{3} \, k_B \, (\, k_B \, T \,) \, \rho(\mu) \; + \; O(T^3) \tag{1059}
\end{aligned}
$$

since on expanding $\frac{\partial \epsilon \rho}{\partial \epsilon}$, one finds the term $\epsilon \frac{\partial \rho}{\partial \epsilon}$ cancels with the term $-\mu \frac{\partial \rho}{\partial \epsilon}$ coming from the temperature dependence of the chemical potential. Hence, the low-temperature electronic specific heat at constant volume is linearly proportional to temperature and the coefficient involves the density of states at the Fermi-energy.

$$C_V \;=\; \frac{\pi^2}{3} \, k_B \, (\, k_B \, T \,) \, \rho(\mu) \tag{1060}$$

The above result is in contrast with the specific heat of a classical gas, which is given by

$$C_v \;=\; \frac{3}{2} \, N \, k_B \tag{1061}$$

The result found using quantum statistical mechanics

$$C_V \;=\; \frac{\pi^2}{3} \, k_B \, (\, k_B \, T \,) \, \rho(\mu) \tag{1062}$$

is consistent with Nernst's law as

$$C_V \;=\; T \left(\frac{\partial S}{\partial T} \right)_T \tag{1063}$$

vanishes as $T \to 0$ since S vanishes as $T \to 0$. This occurs since the quantum ground state is unique so that S/N vanishes as $T \to 0$. The uniqueness occurs since the lowest energy single-particle states are all occupied by one electron, in accordance with the Pauli exclusion principle. Since there is no degeneracy, $S = 0$.

The specific heat of the electron gas is proportional to T. This can be understood by considering the effect of the Pauli exclusion principle. In a classical gas, where there is no exclusion principle, on supplying the thermal energy to the system, on average each particle acquires a kinetic energy of $\frac{3}{2} k_B T$. Hence, the excitation energy of the system is proportional to $k_B T$ and

$$\overline{E} = \frac{3}{2} N k_B T \tag{1064}$$

so

$$\begin{aligned} C_v &= \left(\frac{\partial E}{\partial T} \right)_V \\ &= \frac{3}{2} N k_B \end{aligned} \tag{1065}$$

For fermions, if one supplies the thermal energy to the system, only the electrons within $k_B T$ of the Fermi-energy can be excited. An electron in an energy level far below ϵ_F cannot be excited by $k_B T$ since the final state with higher energy is already occupied. Thus, the Pauli-exclusion principle forbids it to be excited. However, electrons within $k_B T$ of the Fermi-energy can be excited. The initial state is occupied, but the final state is above the Fermi-energy and can accept the excited electron.

Only the electrons within $k_B T$ of the Fermi-energy can be excited. The number of these electrons is approximately given by

$$\rho(\epsilon_F) k_B T \tag{1066}$$

Each of these electrons can be excited by the thermal energy $k_B T$, so the increase in the systems energy is given by

$$\Delta E = \rho(\epsilon_F) (k_B T)^2 \tag{1067}$$

Hence, the specific heat is estimated as

$$\begin{aligned} C_V &= \frac{\Delta E}{T} \\ &= \rho(\epsilon_F) k_B^2 T \end{aligned} \tag{1068}$$

which shows that the linear T dependence is due to the Pauli-exclusion principle.

Similar arguments apply to thermodynamic properties or transport coefficients of the electron gas. The states far from ϵ_F are inert since they cannot be excited by the small energies involved in the process, since their electrons cannot move up in energy because the desired final states are already occupied. The Pauli exclusion principle blocks these states from participating in processes which involve low excitation energies. Thus, they don't participate in electrical conduction, etc. These processes are all dominated by the states near ϵ_F, hence they depend on the density of states evaluated at the Fermi-energy $\rho(\mu)$ or its derivatives $\frac{\partial \rho}{\partial \mu}$, etc. Hence, it may be useful to find other experimental properties which can be used to measure $\rho(\mu)$.

8.7 The Pauli Paramagnetic Susceptibility of an Electron Gas

The Pauli paramagnetic susceptibility provides an alternate measure of the single-particle density of states at the Fermi-energy. The effect of the Pauli exclusion principle limits the number of electrons that are allowed to flip their spins to an energy of the order of $\mu_B \, H$ from the Fermi-energy.

Consider a gas of non-interacting electrons, each of which carries a spin $S = \frac{1}{2}$. These spins couple to an applied magnetic field H^z aligned along the z-axis, via the anomalous Zeeman interaction

$$\hat{H}_{int} \; = \; - \, g \, \mu_B \, S^Z \, H^z \tag{1069}$$

where $g = 2$ is the gyromagnetic ratio. Hence, in the presence of the field, the single-electron energy levels become

$$
\begin{aligned}
\epsilon_\alpha \; - \; \mu_B \, H^z \qquad &\text{for} \qquad S^z \; = \; + \frac{1}{2} \\
\epsilon_\alpha \; + \; \mu_B \, H^z \qquad &\text{for} \qquad S^z \; = \; - \frac{1}{2}
\end{aligned} \tag{1070}
$$

Therefore, the Grand-Canonical potential is given by

$$- \, \beta \, \Omega \; = \; \sum_{\alpha,\pm} \ln \left[\, 1 \; + \; \exp[-\beta(\epsilon_\alpha \; \mp \; \mu_B \, H^z \; - \; \mu) \,] \, \right] \tag{1071}$$

which can be written as an integral over the (zero-field) density of states

$$- \, \beta \, \Omega \; = \; \frac{1}{2} \sum_{\pm} \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \ln \left[\, 1 \; + \; \exp[-\beta(\epsilon \; \mp \; \mu_B \, H^z \; - \; \mu) \,] \, \right] \tag{1072}$$

where the factor $\frac{\rho}{2}$ is the density of states per spin direction (for zero applied fields). The magnetization is defined as

$$M^z \; = \; - \left(\frac{\partial \Omega}{\partial H^z} \right) \tag{1073}$$

which yields

$$
\begin{aligned}
M^z &= \frac{1}{2} \sum_{\pm} (\pm) \, \mu_B \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\exp[-\beta(\epsilon \mp \mu_B H^z - \mu)]}{1 + \exp[-\beta(\epsilon \mp \mu_B H^z - \mu)]} \\
&= \frac{1}{2} \sum_{\pm} (\pm) \, \mu_B \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, f(\epsilon \mp \mu_B H^z) \quad (1074)
\end{aligned}
$$

This can simply be interpreted as μ_B times the excess of spin-up electrons over down-spin electrons

$$
M^z = \sum_{\alpha,\pm} (\pm) \, \overline{n}_{\alpha,\pm} \quad (1075)
$$

which vanishes when $H^z \to 0$. The differential susceptibility is defined as
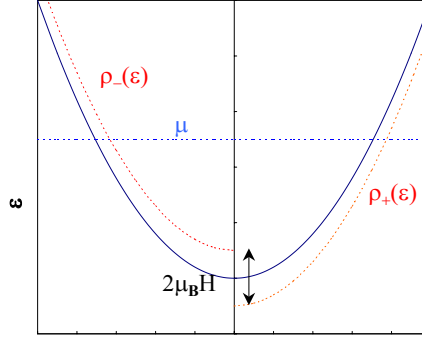
Figure 43: The spin-split single-electron density of states (red), in the presence of a finite magnetic field. Due to the Pauli principle, the field can only realign the spins of electrons which have energies within $\mu_B H^z$ of the Fermi-energy.

$$
\chi^{z,z} = \left(\frac{\partial M^z}{\partial H^z} \right)_{H^z = 0} \quad (1076)
$$

which measure the linear field dependence of M. The susceptibility is evaluated as

$$
\chi^{z,z} = -\frac{1}{2} \sum_{\pm} \mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\partial f}{\partial \epsilon}(\epsilon \mp \mu_B H^z) \quad (1077)
$$

which remains finite in the limit $H^z \to 0$. In the limit of zero field, the susceptibility simplifies to

$$
\begin{aligned}
\chi^{z,z} &= -\frac{1}{2} \sum_{\pm} \mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\partial f}{\partial \epsilon} \\
&= -\mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\partial f}{\partial \epsilon} \quad (1078)
\end{aligned}
$$

In the limit of zero temperature, one has

$$-\frac{\partial f}{\partial \epsilon} = \delta(\epsilon - \mu) \tag{1079}$$

therefore, one finds

$$\chi^{z,z} = \mu_B^2 \, \rho(\mu) \tag{1080}$$

Hence, the ratio of the specific heat and the susceptibility given by

$$\frac{\mu_B^2 \, C_V}{k_B^2 \, T \, \chi^{z,z}} = \frac{\pi^2}{3} \tag{1081}$$

This relation is independent of $\rho(\mu)$ and provided a check on the theory. It is satisfied for all simple metals and for most of the early transition metals. Thus the low temperature limit of $\chi^{z,z}$ is a measure of the density of states at the Fermi-energy.

The leading temperature dependent corrections to χ can be obtained from the Sommerfeld expansion. The susceptibility is given by

$$\chi^{z,z} = -\mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\partial f}{\partial \epsilon} \tag{1082}$$

and on integrating by parts, one obtains

$$\chi^{z,z} = -\mu_B^2 \, \rho(\epsilon) \, f(\epsilon) \Big|_{-\infty}^{\infty} + \mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \frac{\partial \rho}{\partial \epsilon} \, f(\epsilon)$$

$$= \mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \frac{\partial \rho}{\partial \epsilon} \, f(\epsilon) \tag{1083}$$

since the boundary terms vanish. On using the Sommerfeld expansion, one obtains the result

$$\chi^{z,z} = \mu_B^2 \, \rho(\mu) + \mu_B^2 \, \frac{\pi^2}{6} \, (\, k_B \, T \,)^2 \left[\frac{\partial^2 \rho}{\partial \epsilon^2} - \frac{1}{\rho} \left(\frac{\partial \rho}{\partial \epsilon} \right)^2 \right] \Bigg|_{\epsilon_F} + \dots \tag{1084}$$

The corrections aren't important unless $k_B \, T \sim \rho / \frac{\partial \rho}{\partial \epsilon}$, i.e. the temperature is of the order of the energy over which ρ varies. Thus, $\chi^{z,z}$ is approximately temperature independent.

8.8 The High-Temperature Limit of the Susceptibility

The high-temperature limit of the susceptibility

$$\chi^{z,z} = -\mu_B^2 \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \frac{\partial f}{\partial \epsilon} \tag{1085}$$

can be found by using the high-temperature approximation

$$f(\epsilon) \;\approx\; \exp[-\beta(\epsilon - \mu)] \tag{1086}$$

which yields the approximation

$$\chi^{z,z} \;\sim\; \beta\,\mu_B^2 \int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\exp[-\beta(\epsilon - \mu)] \tag{1087}$$

which is evaluated as

$$\chi^{z,z} \;\sim\; \beta\,\mu_B^2\,\overline{N} \tag{1088}$$

on utilizing the expression for the number of electrons

$$\overline{N} \;=\; \int_{-\infty}^{\infty} d\epsilon\,\rho(\epsilon)\,\exp[-\beta(\epsilon - \mu)] \tag{1089}$$

Hence, the Pauli paramagnetic susceptibility turns over into a Curie law at sufficiently high temperatures. Since the high temperature variation first happens
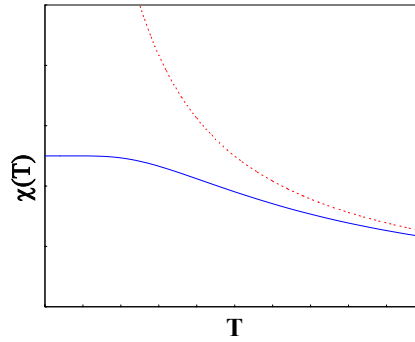
Figure 44: The temperature dependence of the Pauli paramagnetic susceptibility $\chi(T)$ (blue curve). The high temperature Curie law is shown by the dashed red curve.

when $k_B\,T \;\approx\; \mu$, the high temperature limit first applies at temperatures of the order of $T \;\approx\; 12,000$ K.

8.9   The Temperature-dependence of the Pressure of a Gas of Non-Interacting Fermions

The Sommerfeld expansion can also be used to calculate the temperature-dependence of the pressure for a gas of non-interacting fermions at low temperatures. Starting from the expression for the Grand-Canonical potential Ω

and integrating by parts, one finds that

$$
\begin{aligned}
\Omega &= -PV \\
&= -\frac{2}{3} \int_0^\infty d\epsilon \, \rho(\epsilon) \, \epsilon \, \frac{1}{\exp[\,\beta\,(\,\epsilon - \mu\,)\,] + 1} \qquad (1090)
\end{aligned}
$$

where the density of states (including a factor of 2 for both spin directions) is given by

$$
\rho(\epsilon) = \frac{V}{2\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \sqrt{\epsilon} \qquad (1091)
$$

Hence, on substituting for the single-particle density of states and on canceling a factor of V, one finds that the pressure is given by

$$
P = \frac{1}{3\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \int_0^\infty d\epsilon \, \epsilon^{\frac{3}{2}} \, f(\epsilon) \qquad (1092)
$$

where $f(\epsilon)$ is the Fermi-Dirac distribution function. On using the Sommerfeld expansion, one obtains the approximate expression

$$
P \approx \frac{1}{3\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \left[\int_0^{\mu(T)} d\epsilon \, \epsilon^{\frac{3}{2}} + \frac{\pi^2}{6} \, (\,k_B\,T\,)^2 \, \frac{3}{2}\, \mu(T)^{\frac{1}{2}} + \ldots \right] \qquad (1093)
$$

which is evaluated as

$$
P \approx \frac{1}{3\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \left[\frac{2}{5} \, \mu(T)^{\frac{5}{2}} + \frac{\pi^2}{4} \, (\,k_B\,T\,)^2 \, \mu(T)^{\frac{1}{2}} + \ldots \right] \qquad (1094)
$$

The temperature dependence of the chemical potential is approximately given by

$$
\mu(T) \approx \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B\,T}{\epsilon_F} \right)^2 + \ldots \right] \qquad (1095)
$$

Hence, we obtain the final result

$$
\begin{aligned}
P &\approx \frac{1}{3\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \left[\frac{2}{5} \, \epsilon_F^{\frac{5}{2}} + \frac{\pi^2}{6} \, (\,k_B\,T\,)^2 \, \epsilon_F^{\frac{1}{2}} + \ldots \right] \\
&\approx \frac{2}{15\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \epsilon_F^{\frac{5}{2}} \left[1 + \frac{5\,\pi^2}{12} \left(\frac{k_B\,T}{\epsilon_F} \right)^2 + \ldots \right] \qquad (1096)
\end{aligned}
$$

Thus, a gas of non-interacting particles which obey Fermi-Dirac statistics exerts a finite pressure in the limit $T \rightarrow 0$. This can be understood, since at $T = 0$, the particles occupy states with finite momenta up to the Fermi-energy and, therefore, they collide with the container's walls giving rise to pressure. This is in direct contrast with the behavior found for a classical ideal gas and, as we shall see later, is also in contrast with the pressure of a non-interacting gas of bosons at low temperatures.

8.10 Fluctuations in the Occupation Numbers

If one considers a group of energy levels ϵ_α with large enough degeneracy, then it is possible to consider the statistics of the occupation number for α.

The average occupation number of the energy levels ϵ_α, is given by

$$
\begin{aligned}
\overline{n}_\alpha &= \frac{1}{\Xi} \, \mathrm{Trace} \, \hat{n}_\alpha \, \exp\left[\, - \, \beta \, (\, \hat{H} \, - \, \mu \, \hat{N} \,) \, \right] \\
&= \frac{1}{\Xi} \, \mathrm{Trace} \, n_\alpha \, \exp\left[\, - \, \beta \sum_{\alpha'} (\, \epsilon_{\alpha'} \, - \, \mu \,) \, n_{\alpha'} \right] \\
&= \, - \, k_B \, T \, \frac{\partial}{\partial \epsilon_\alpha} \ln \Xi
\end{aligned}
\tag{1097}
$$

where the derivative is w.r.t. the energy level ϵ_α. As expected, this is given by

$$
\begin{aligned}
\overline{n}_\alpha &= \frac{1}{\exp[\beta(\epsilon_\alpha - \mu)] + 1} \\
&= f(\epsilon_\alpha)
\end{aligned}
\tag{1098}
$$

which is just the Fermi-Dirac distribution function.

The mean squared fluctuations around this average is given by

$$
\begin{aligned}
\overline{n_\alpha^2} \, - \, \overline{n_\alpha}^2 &= \left(\, - \, \frac{1}{\beta} \, \frac{\partial}{\partial \epsilon_\alpha} \, \right)^2 \ln \Xi \\
&= \left(\, - \, \frac{1}{\beta} \, \frac{\partial}{\partial \epsilon_\alpha} \, \right) \overline{n}_\alpha \\
&= f(\epsilon_\alpha) \, (\, 1 \, - \, f(\epsilon_\alpha) \,)
\end{aligned}
\tag{1099}
$$

The r.m.s. number fluctuation is reduced from the classical value of $\overline{n}_\alpha$. In fact, due to the Pauli exclusion principle, the fluctuations are only non-zero in an energy width of $k_B \, T$ around the Fermi-energy. The reduction in the fluctuation of fermion occupation numbers is in strong contrast to the fluctuations that are found for particles that obey Bose-Einstein Statistics.

9 Bose-Einstein Statistics

9.1 Non-Interacting Bosons

An ideal gas of non-interacting bosons[30] is described by the Hamiltonian $\hat{H}$ which is given by the sum

$$\hat{H} \; = \; \sum_{\alpha} \epsilon_{\alpha} \; \hat{n}_{\alpha} \tag{1100}$$

where $\hat{n}_{\alpha}$ represents the occupation number of the α-th single-particle energy level. This is just the sum of the contributions from each particle, grouped according to the energy levels that they occupy. Likewise, the operator $\hat{N}$ representing the total number of particles is given by

$$\hat{N} \; = \; \sum_{\alpha} \hat{n}_{\alpha} \tag{1101}$$

where the sum is over the single-particle energy levels. Hence, for non-interacting bosons, the density operator is diagonal in the occupation number representation.

The Grand-Canonical Partition function Ξ is given by

$$\Xi \; = \; \mathrm{Trace} \; \exp\left[\; -\beta \; (\; \hat{H} \; - \; \mu \; \hat{N} \;) \; \right] \tag{1102}$$

where the trace is over a complete set of microscopic states with variable N for the entire system. A convenient basis is given by the N-particle states $| \; \Phi_{n_{\alpha_1}, n_{\alpha_2}, \ldots} > $, since both $\hat{H}$ and $\hat{N}$ are diagonal in this basis. The trace reduces to the sum of all configurations $\{ \; n_{\alpha} \; \}$, and since the total number of particles N is also being summed over, the trace is unrestricted. Therefore, the trace can be evaluated by summing over all possible values of the eigenvalues n_{α} for each consecutive value of α. That is

$$\mathrm{Trace} \left\{ \; \ldots \; \right\} \; \equiv \; \sum_{n_{\alpha_1}=0}^{\infty} \sum_{n_{\alpha_2}=0}^{\infty} \sum_{n_{\alpha_3}=0}^{\infty} \; \ldots \; \left\{ \; \ldots \; \right\} \tag{1103}$$

Since the exponential term in Ξ reduces to the form of the exponential of a sum of independent terms, it can be written as the product of exponential factors

$$\exp\left[\; -\beta \; (\; \hat{H} \; - \; \mu \; \hat{N} \;) \; \right] \; = \; \exp\left[\; -\beta \; \sum_{\alpha} (\; \epsilon_{\alpha} \; - \; \mu \;) \; \hat{n}_{\alpha} \; \right]$$

$$= \; \prod_{\alpha} \exp\left[\; -\beta \; (\; \epsilon_{\alpha} \; - \; \mu \;) \; \hat{n}_{\alpha} \; \right] \tag{1104}$$

---

[30] A. Einstein, "Quantentheorie des einatomigen idealen Gases", Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1, 3-14, (1925).

Therefore, the trace in the expression for Ξ can be reduced to

$$\Xi = \prod_{\alpha} \left\{ \sum_{n_\alpha=0}^{\infty} \exp\left[-\beta(\epsilon_\alpha - \mu)n_\alpha \right] \right\} \qquad (1105)$$

where the sum is over all the occupation numbers n_α allowed by Bose-Einstein statistics. The summation is of the form of a geometric series, as can be seen by introducing the variable x_α defined by

$$x_\alpha = \exp\left[-\beta(\epsilon_\alpha - \mu) \right] \qquad (1106)$$

so

$$\sum_{n_\alpha=0}^{\infty} \exp\left[-\beta(\epsilon_\alpha - \mu)n_\alpha \right] = \sum_{n_\alpha=0}^{\infty} x_\alpha^{n_\alpha} \qquad (1107)$$

This geometric series converges if $x_\alpha \leq 1$, i.e.

$$\exp\left[-\beta(\epsilon_\alpha - \mu) \right] \leq 1 \qquad (1108)$$

which requires that $\epsilon_\alpha > \mu$. This condition has to hold for all ϵ_α, so μ must be smaller than the lowest single-particle energy level ϵ_0. Therefore, we require that

$$\epsilon_0 > \mu \qquad (1109)$$

On performing the sum of the geometric series, one obtains

$$\sum_{n_\alpha=0}^{\infty} \exp\left[-\beta(\epsilon_\alpha - \mu)n_\alpha \right] = \frac{1}{1 - \exp[-\beta(\epsilon_\alpha - \mu)]} \qquad \text{where } \epsilon_0 > \mu \quad (1110)$$

Therefore, the Grand-Canonical Partition Function is given by

$$\Xi = \prod_{\alpha} \left\{ \frac{1}{1 - \exp[-\beta(\epsilon_\alpha - \mu)]} \right\} \qquad (1111)$$

The Grand-Canonical Potential Ω is found from

$$\Xi = \exp\left[-\beta\Omega \right] \qquad (1112)$$

which, on taking the logarithm, yields

$$-\beta\Omega = \sum_{\alpha} \ln\left[\frac{1}{1 - \exp[-\beta(\epsilon_\alpha - \mu)]} \right] \qquad (1113)$$

or

$$\Omega = k_B T \sum_{\alpha} \ln\left[1 - \exp[-\beta(\epsilon_\alpha - \mu)] \right] \qquad (1114)$$

206

On introducing the density of single-particle states $\rho(\epsilon)$ defined as

$$\rho(\epsilon) \;=\; \sum_{\alpha} \delta(\,\epsilon \,-\, \epsilon_{\alpha}\,) \tag{1115}$$

the summation can be transformed into an integral over ϵ

$$\Omega \;=\; k_B\,T \int_{-\infty}^{\infty} d\epsilon \; \rho(\epsilon) \; \ln \left[\, 1 \,-\, \exp[-\beta(\epsilon-\mu)]\,\right] \tag{1116}$$

where the density of states goes to zero below ϵ_0. After evaluating the integral, one may find all the thermodynamic properties of the system from Ω.

9.2 The Bose-Einstein Distribution Function

The average number of particles in the system $\overline{N}$ can be determined from Ω via the relation

$$\overline{N} \;=\; -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} \tag{1117}$$

which is evaluated as

$$\begin{aligned}
\overline{N} \;&=\; k_B\,T \int_{-\infty}^{\infty} d\epsilon\; \rho(\epsilon)\; \frac{\beta\;\exp[-\beta(\epsilon-\mu)]}{1\;-\;\exp[-\beta(\epsilon-\mu)]} \\
&=\; \int_{-\infty}^{\infty} d\epsilon\; \rho(\epsilon)\; \frac{1}{\exp[\beta(\epsilon-\mu)]\;-\;1}
\end{aligned} \tag{1118}$$

or equivalently

$$\overline{N} \;=\; \sum_{\alpha} \frac{1}{\exp[\beta(\epsilon_{\alpha}-\mu)]\;-\;1} \tag{1119}$$

but, by definition, one has

$$\overline{N} \;=\; \sum_{\alpha} \overline{n}_{\alpha} \tag{1120}$$

Therefore, the function $N(\epsilon)$ defined by

$$N(\epsilon) \;=\; \frac{1}{\exp[\beta(\epsilon-\mu)]\;-\;1} \tag{1121}$$

represents the average number of particles $\overline{n}$ in a quantum level with a single-particle energy ϵ. The function $N(\epsilon)$ is the Bose-Einstein distribution function. The Bose-Einstein distribution vanishes as

$$N(\epsilon) \;\approx\; \exp\left[\, -\, \beta\,(\,\epsilon\,-\,\mu\,)\,\right] \tag{1122}$$

for $\epsilon \,-\, \mu \,>\, k_B\,T$. For low energies $\mu \,<\, \epsilon \,<\, \mu \,+\, k_B\,T$, the Bose-Einstein
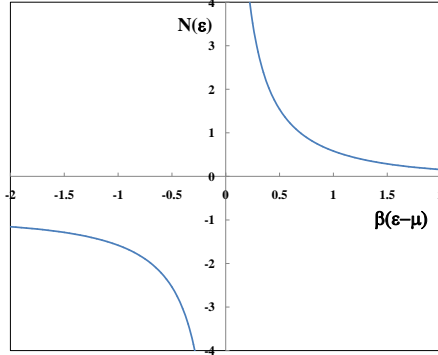
Figure 45: The Bose-Einstein Distribution function.

distribution varies as

$$N(\epsilon) \ \approx \ \frac{k_B \ T}{\epsilon \ - \ \mu} \ - \ \frac{1}{2} \tag{1123}$$

which can become arbitrarily large.

The Bose-Einstein distribution function enters the expressions for other thermodynamic quantities, such as the average energy. The energy can be found from the expression for the entropy S. The entropy can be found from the infinitesimal relation

$$d\Omega \ = \ - \ S \ dT \ - \ P \ dV \ - \ N \ d\mu \tag{1124}$$

which yields

$$
\begin{aligned}
S \ &= \ - \left(\frac{\partial \Omega}{\partial T} \right)_{V,\mu} \\
&= \ - \frac{\Omega}{T} \ + \ k_B \ \beta \int_{-\infty}^{\infty} d\epsilon \ \rho(\epsilon) \ \frac{(\ \epsilon \ - \ \mu \ )}{\exp[\beta(\epsilon - \mu)] \ - \ 1}
\end{aligned}
\tag{1125}
$$

On recalling the expression for $\overline{N}$, one has

$$S \ = \ - \frac{(\ \Omega \ + \ \mu \ \overline{N} \ )}{T} \ + \ \frac{1}{T} \int_{-\infty}^{\infty} d\epsilon \ \rho(\epsilon) \ \epsilon \ \frac{1}{\exp[\beta(\epsilon - \mu)] \ - \ 1} \tag{1126}$$

On using the definition

$$\Omega \ = \ U \ - \ T \ S \ - \ \mu \ \overline{N} \tag{1127}$$

one finds that the thermodynamic energy is given by

$$U \ = \ \int_{-\infty}^{\infty} d\epsilon \ \rho(\epsilon) \ \epsilon \ \frac{1}{\exp[\beta(\epsilon - \mu)] \ - \ 1} \tag{1128}$$

208

or alternately as

$$U = \sum_\alpha \epsilon_\alpha \frac{1}{\exp[\beta(\epsilon_\alpha - \mu)] - 1} \tag{1129}$$

The thermodynamic energy U should be compared with the expression for the average energy of the non-interacting particles

$$\overline{E} = \sum_\alpha \epsilon_\alpha \, \overline{n}_\alpha \tag{1130}$$

This comparison reconfirms our identification of the average number of particles in the α-th energy level as

$$\overline{n}_\alpha = \frac{1}{\exp[\beta(\epsilon_\alpha - \mu)] - 1} \tag{1131}$$

In general, the Bose-Einstein distribution function is defined as

$$N(\epsilon_\alpha) = \frac{1}{\exp[\beta(\epsilon_\alpha - \mu)] - 1} \tag{1132}$$

where $\epsilon \geq \mu$ and corresponds to the thermal average of the occupation number of a quantum level α with single-particle energy ϵ. Since the occupation numbers can have the values $n_\alpha = 0$, 1, 2, $\ldots \infty$, one has

$$\infty \geq n_\alpha \geq 0 \tag{1133}$$

so the average value must also satisfy the same inequalities

$$\infty \geq N(\epsilon_\alpha) \geq 0 \tag{1134}$$

The positivity of $N(\epsilon_\alpha)$ requires that $\epsilon_\alpha \geq \mu$. However, there does exist a possibility that, in thermal equilibrium, a level can be occupied by an average number of particles that tends to infinity. This possibility requires that the energy of this level is sufficiently close to μ, i.e. $\epsilon_0 \approx \mu$.

For a system with a fixed average number of particles $\overline{N}$, the equation

$$\overline{N} = \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, N(\epsilon) \tag{1135}$$

has to be regarded as an implicit equation for $\mu(T)$. Once $\mu(T)$ has been found, one may then calculate other thermodynamic averages. For example, the average energy of our system of non0interacting particles is given by

$$U = \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \epsilon \, N(\epsilon) \tag{1136}$$

etc.

9.3 The Equation of State for Non-Interacting Bosons

The equation of state for a gas of non-interacting bosons can be found from Ω by noting that

$$\Omega = - P V \tag{1137}$$

The equation of state can be obtained directly when the single-particle density of states $\rho(\epsilon)$ has the form of a simple power law

$$\begin{aligned}
\rho(\epsilon) &= C \, \epsilon^\alpha \quad \text{for } \epsilon \geq 0 \\
&= 0 \quad \text{otherwise}
\end{aligned} \tag{1138}$$

where α is a constant. The Grand-Canonical Potential Ω is found as

$$\begin{aligned}
\Omega &= - P V \\
&= k_B T \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \ln \left[1 - \exp[-\beta(\epsilon - \mu)] \right]
\end{aligned} \tag{1139}$$

Hence, one has

$$\begin{aligned}
- \frac{P V}{k_B T} &= C \int_0^{\infty} d\epsilon \, \epsilon^\alpha \, \ln \left[1 - \exp[-\beta(\epsilon - \mu)] \right] \\
&= \frac{C}{\alpha + 1} \int_0^{\infty} d\epsilon \, \frac{d\epsilon^{(\alpha+1)}}{d\epsilon} \, \ln \left[1 - \exp[-\beta(\epsilon - \mu)] \right]
\end{aligned} \tag{1140}$$

On integrating by parts, one obtains

$$\begin{aligned}
- \frac{P V}{k_B T} &= \frac{C}{\alpha + 1} \left[\epsilon^{\alpha+1} \, \ln \left(1 - \exp[-\beta(\epsilon - \mu)] \right) \right] \Big|_0^{\infty} \\
&\quad - \frac{C}{\alpha + 1} \int_0^{\infty} d\epsilon \, \epsilon^{(\alpha+1)} \, \frac{\beta \, \exp[-\beta(\epsilon - \mu)]}{1 - \exp[-\beta(\epsilon - \mu)]}
\end{aligned} \tag{1141}$$

The boundary terms vanish, since the density of states vanishes at the lower limit of integration $\epsilon = 0$ and the logarithmic factor vanishes exponentially when $\epsilon \to \infty$. Thus, on canceling a factor of β, one finds

$$\begin{aligned}
P V &= \frac{1}{\alpha + 1} \int_0^{\infty} d\epsilon \, C \, \epsilon^{(\alpha+1)} \, \frac{1}{\exp[\beta(\epsilon - \mu)] - 1} \\
&= \frac{1}{\alpha + 1} \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \epsilon \, \frac{1}{\exp[\beta(\epsilon - \mu)] - 1} \\
&= \frac{U}{\alpha + 1}
\end{aligned} \tag{1142}$$

That is, the equation of state for an ideal gas of bosons is found as

$$P V = \frac{U}{\alpha + 1} \tag{1143}$$

The same method was used to find the equation of state for an ideal gas of fermions. The result is the same. Therefore, the equation of state holds true, independent of the quantum statistics used. The equation of state must also apply to the classical ideal gas, as the ideal gas can be considered as the high-temperature limiting form of the ideal quantum gasses.

9.4 The Fugacity at High Temperatures

For a system with a fixed number of particles $\overline{N}$, the equation

$$\overline{N} = \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, N(\epsilon) \tag{1144}$$

implicitly determines $\mu(T)$. For non-relativistic bosons with the dispersion relation

$$\epsilon_{\underline{k}} = \frac{\hbar^2 \, \underline{k}^2}{2 \, m} \tag{1145}$$

in three-dimensions, the density of states can be calculated by replacing the sum over discrete values of $\underline{k}$ by an integral over a density if points in phase space

$$
\begin{aligned}
\rho(\epsilon) &= \sum_{\underline{k}} \delta(\epsilon - \epsilon_{\underline{k}}) \\
&= \frac{V}{(2\,\pi)^3} \int d^3\underline{k} \, \delta(\epsilon - \epsilon_{\underline{k}}) \\
&= \frac{V}{2\,\pi^2} \int_0^{\infty} dk \, k^2 \, \frac{2\,m}{\hbar^2} \, \delta\left(\frac{2\,m\,\epsilon}{\hbar^2} - k^2 \right) \\
&= \frac{V}{2\,\pi^2} \frac{m}{\hbar^2} \sqrt{\frac{2\,m\,\epsilon}{\hbar^2}}
\end{aligned}
\tag{1146}
$$

Thus, the average number of particles is expressed as

$$\overline{N} = \frac{V}{4\,\pi^2} \left(\frac{2\,m}{\hbar^2} \right)^{\frac{3}{2}} \int_0^{\infty} d\epsilon \, \epsilon^{\frac{1}{2}} \, \frac{1}{z^{-1} \, \exp[\,\beta\,\epsilon\,] \, - \, 1} \tag{1147}$$

where z is the fugacity is defined as

$$z = \exp[\,\beta\,\mu\,] \tag{1148}$$

The equation for $\overline{N}$ can be re-written in terms of the dimensionless variable $x = \beta\,\epsilon$ as

$$\overline{N} = \frac{V}{4\,\pi^2} \left(\frac{2\,m\,k_B\,T}{\hbar^2} \right)^{\frac{3}{2}} \int_0^{\infty} dx \, x^{\frac{1}{2}} \, \frac{1}{z^{-1} \, \exp[\,x\,] \, - \, 1} \tag{1149}$$

which determines the fugacity z as a function of temperature. This equation can be expressed as

$$\frac{2}{\sqrt{\pi}} \int_0^{\infty} dx \, x^{\frac{1}{2}} \, \frac{1}{z^{-1} \, \exp[\,x\,] \, - \, 1} = \frac{N}{V} \left(\frac{2\,\pi\,\hbar^2}{m\,k_B\,T} \right)^{\frac{3}{2}} \tag{1150}$$

211

or as

$$\frac{1}{\Gamma(\frac{3}{2})} \int_0^\infty dx \; x^{\frac{1}{2}} \; \frac{1}{z^{-1} \; \exp[\; x\;] \; - \; 1} \; = \; \frac{N}{V} \left(\frac{2 \; \pi \; \hbar^2}{m \; k_B \; T} \right)^{\frac{3}{2}} \tag{1151}$$

where

$$\Gamma(\alpha + 1) \; = \; \int_0^\infty dx \; x^\alpha \; \exp[\; - \; x\;] \tag{1152}$$

This type of integral appears frequently in the evaluation of other quantities[31]. We shall denote the integral $I_{\alpha+1}(z)$ as

$$I_{\alpha+1}(z) \; = \; \frac{1}{\Gamma(\alpha + 1)} \int_0^\infty dx \; x^\alpha \; \frac{1}{z^{-1} \; \exp[x] \; - \; 1} \tag{1153}$$

where $z \; < \; 1$ and $x \; > \; 1$. The integrand can be expanded as

$$
\begin{aligned}
I_{\alpha+1}(z) \; &= \; \frac{1}{\Gamma(\alpha + 1)} \int_0^\infty dx \; x^\alpha \; \frac{z \; \exp[-x]}{1 \; - \; z \; \exp[-x]} \\
&= \; \frac{1}{\Gamma(\alpha + 1)} \int_0^\infty dx \; x^\alpha \; \sum_{m=1}^\infty \; z^m \; \exp[-mx]
\end{aligned}
\tag{1154}
$$

On transforming the variable of integration to $y = mx$ one obtains

$$
\begin{aligned}
I_{\alpha+1}(z) \; &= \; \frac{1}{\Gamma(\alpha + 1)} \int_0^\infty dy \; y^\alpha \; \exp[-y] \; \sum_{m=1}^\infty \; \frac{z^m}{m^{\alpha+1}} \\
&= \; \sum_{m=1}^\infty \; \frac{z^m}{m^{\alpha+1}}
\end{aligned}
\tag{1155}
$$

For small z, $z \ll 1$, one only needs to retain the first terms so

$$I_{\alpha+1}(z) \; \approx \; z \tag{1156}$$

while for z equal to unity, $z = 1$, the integral has the value

$$
\begin{aligned}
I_{\alpha+1}(1) \; &= \; \sum_{m=1}^\infty \; m^{-(\alpha+1)} \\
&= \; \xi(\alpha + 1)
\end{aligned}
\tag{1157}
$$

The equation determining $\overline{N}$ can be expressed in terms of the above set of functions as

$$I_{\frac{3}{2}}(z) \; = \; \frac{N}{V} \left(\frac{2 \; \pi \; \hbar^2}{m \; k_B \; T} \right)^{\frac{3}{2}} \tag{1158}$$

---

[31] J.E. Robinson, Phys. Rev. 83, 678 (1951).

where $\alpha = \frac{1}{2}$. This equation can be solved graphically if the function $I_{\frac{3}{2}}(z)$ is plotted vs z since the intersection with the line representing the righthand side yields the solution for z. At high temperatures, one finds that the solution for z is much less than unity so that the chemical potential μ is negative. As the temperature decreases, the value of z increases towards its maximum value of one, which corresponds to μ increasing towards zero. The temperature for which $z = 1$ is given by

$$\xi\left(\frac{3}{2}\right) = \frac{N}{V}\left(\frac{2\,\pi\,\hbar^2}{m\,k_B\,T_c}\right)^{\frac{3}{2}} \tag{1159}$$

For temperatures below T_c, the equation for $\overline{N}$ cannot be satisfied and the lowest energy level has to have a macroscopic occupation number. That is for $T < T_c$, the bosons must condense into the lowest energy state, as first predicted by Einstein. For this low temperature range, it is no longer sufficient to use a continuum expression for the density of single-particle state $\rho(\epsilon)$ which fails to give the proper weight for the lowest energy state. That is, the method used for calculating the density of states approximates the sum over states by an integration over the density of points in phase space states. Therefore, it only calculates the average number of points on the constant energy surface. This approximation fails miserably at very low energies, when the number of points on the constant energy surface is low. A better approximation has the form

$$\rho(\epsilon) = \delta(\epsilon) + C\,\epsilon^{\frac{1}{2}} \tag{1160}$$

which explicitly includes a delta function of weight of unity for the lowest energy state and C is an extensive constant.

9.5 Fluctuations in the Occupation Numbers

9.6 Bose-Einstein Condensation

The expression for the Bose-Einstein distribution function only makes sense when the chemical potential μ is less than the energy of the single-particle quantum state ϵ. Since this is true for all ϵ and since we have defined the lowest single-particle energy as $\epsilon = 0$, we must require that $\mu < 0$. As was first pointed out by Einstein, if μ approaches zero, then there can be a macroscopic occupation of a single-particle quantum level.

When there is a macroscopic occupation of the lowest energy level, the single particle density of states must explicitly include the lowest energy state. In this case, we need to use the better approximation to the density of states given by

$$\rho(\epsilon) = \delta(\epsilon) + C\,\epsilon^{\alpha}\,\Theta(\epsilon) \tag{1161}$$

where the function $\Theta(\epsilon)$ is the Heaviside step function. The delta function represents the lowest energy state. The second term represents the approximation

for the extensive part of the density of states. This expression can also be used at temperatures above T_c since the contribution from the delta function is not extensive and can be ignored.

On using the above expression for the density of states to evaluate the average number of particles, one finds

$$
\begin{aligned}
\overline{N} &= \int_{-\infty}^{\infty} d\epsilon \; \rho(\epsilon) \; N(\epsilon) \\
&= \int_{-\infty}^{\infty} d\epsilon \; \left(\; \delta(\epsilon) \; + \; C \; \epsilon^\alpha \; \Theta(\epsilon) \; \right) \; N(\epsilon) \\
&= N(0) \; + \; \int_0^{\infty} d\epsilon \; C \; \epsilon^\alpha \; N(\epsilon)
\end{aligned}
\tag{1162}
$$

where the first term represents the number of particles in the quantum level with zero energy. On changing the variable of integration to $x \; = \; \beta \; \epsilon$, the expression can be re-written as

$$
\begin{aligned}
\widetilde{N} &= \frac{1}{z^{-1} \; - \; 1} \; + \; C \; (\; k_B \; T \;)^{\alpha+1} \int_0^{\infty} dx \; \frac{x^\alpha}{z^{-1} \; \exp[x] \; - \; 1} \\
&= \frac{z}{1 \; - \; z} \; + \; \Gamma(\alpha+1) \; C \; (\; k_B \; T \;)^{\alpha+1} \; I_{\alpha+1}(z)
\end{aligned}
\tag{1163}
$$

This equation determines z when the average number of particles $\overline{N}$ is fixed. The above equation can be interpreted as the sum of the particles in the lowest
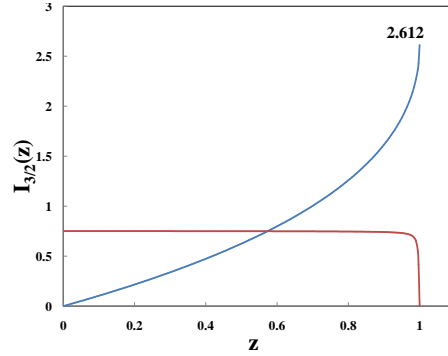
Figure 46: The graphical solution for the fugacity found from plotting both $I_{\frac{3}{2}}(z)$ (blue) and $A[\overline{N} - z/(1 - z)]$ (red) versus z. The point of intersection of the curves yields the value of the fugacity. Note that although $I_{\frac{3}{2}}(z)$ has a divergent derivative at $z = 1$, its value there is finite and is given by 2.612.

energy state N_0 and the number of particles in the excited states N_{exc}

$$
\overline{N} \; = \; N_0 \; + \; N_{exc}
\tag{1164}
$$

where the number of excited particles is given by

$$N_{exc} = \Gamma(\alpha+1) \, C \, (\, k_B \, T \,)^{\alpha+1} \, I_{\alpha+1}(z) \tag{1165}$$

The number of particles in the condensate N_0 is given by

$$N_0 = \frac{z}{1 - z} \tag{1166}$$

Above the Condensation Temperature

For $T > T_c$, one has

$$N_{exc} = \overline{N} \tag{1167}$$

as the explicit equation for $\overline{N}$ determines the value of z to be $z < 1$. Thus, the number of particles in the condensate is not extensive and is negligible, so all the particles can be considered as being in the excited states.

The Condensation Temperature

The Bose-Einstein condensation temperature T_c is evaluated from the condition that $z = 1 - \eta$ where η is a positive small quantity

$$\overline{N} = \Gamma(\alpha+1) \, C \, (\, k_B \, T_c \,)^{\alpha+1} \, \xi(\alpha+1) \tag{1168}$$

which determines the lowest temperature at which the number of the particles in the condensate is still negligibly small ($N_0 \sim \eta^{-1}$). Note that, since $C \propto V$, the condensation temperature T_c depends on the density, or if the number of particles is fixed it depends on the volume $T_c(V)$.

Below the Condensation Temperature

For temperatures below T_c, $T < T_c$, the number of particles in the excited states is temperature dependent and decreases towards zero as T is reduced to zero according to a simple power law.

$$N_{exc} = \Gamma(\alpha+1) \, C \, (\, k_B \, T \,)^{\alpha+1} \, \xi(\alpha+1) \tag{1169}$$

or equivalently as

$$N_{exc} = \overline{N} \left(\frac{T}{T_c} \right)^{\alpha+1} \tag{1170}$$

where we have used the equation for the Bose-Einstein condensation temperature to eliminate the constant C. The number of particles in the condensate is defined as

$$N_0 = \overline{N} - N_{exc} \tag{1171}$$

which is evaluated as

$$N_0 = \overline{N} \left[1 - \left(\frac{T}{T_c} \right)^{\alpha+1} \right] \tag{1172}$$

which tends to N in the limit of $T \to 0$. In this case, one also has
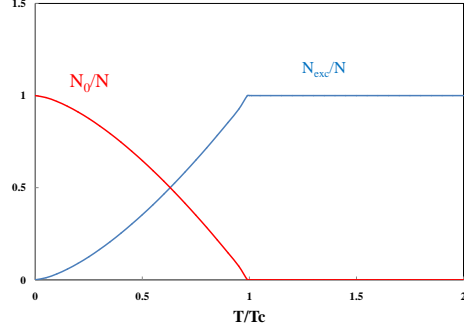
Figure 47: The temperature-dependence of the relative number of condensate particles N_0 and the relative number of excited particles N_{exc}.

$$\frac{z}{1 - z} = \overline{N} \left[1 - \left(\frac{T}{T_c} \right)^{\alpha+1} \right] \tag{1173}$$

which determines z as

$$z = \frac{1}{1 - \frac{1}{N} \left[1 - \left(\frac{T}{T_c} \right)^{\alpha+1} \right]^{-1}} \tag{1174}$$

which confirms that $z \approx 1$.

Properties of the Idealized Condensed Phase

The average energy U of the condensed phase of an ideal gas of bosons is given by

$$U = \int_{-\infty}^{\infty} d\epsilon \, \rho(\epsilon) \, \epsilon \, N(\epsilon) \tag{1175}$$

which together with expression for the single-particle density of states

$$\rho(\epsilon) = \delta(\epsilon) + C \, \epsilon^\alpha \, \Theta(\epsilon) \tag{1176}$$

yields the expression

$$U = \int_0^\infty d\epsilon \, C \, \epsilon^\alpha \, \epsilon \, N(\epsilon) \tag{1177}$$

216

Figure 48: The velocity distribution of a gas of Rb atoms at three different temperatures, showing the evolution of the macroscopic occupation of the lowest energy single-particle state as the temperature is lowered. [M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman, and E.A. Cornell, "Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor", Science 269, 198-201 (1995).]
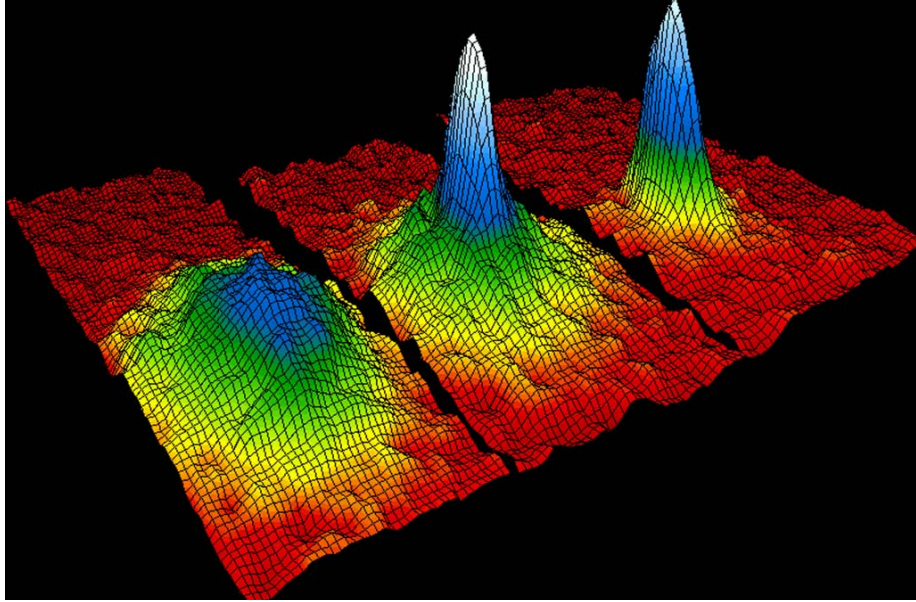
since the $N(0)$ particles with $\epsilon = 0$ don't contribute to the total energy. Then for $z \approx 1$, the integral reduces to

$$
\begin{aligned}
U &= C \ (\ k_B \ T \ )^{\alpha+2} \ \Gamma(\alpha + 2) \ I_{\alpha+2}(1) \\
&= C \ (\ k_B \ T \ )^{\alpha+2} \ \Gamma(\alpha + 2) \ \xi(\alpha + 2)
\end{aligned}
\tag{1178}
$$

Then, the specific heat for $T \leq T_c$ can be found from

$$
\begin{aligned}
C_{V,N} &= \left(\frac{\partial U}{\partial T} \right)_{V,N} \\
&= k_B \ C \ (\ k_B \ T \ )^{\alpha+1} \ \Gamma(\alpha + 3) \ \xi(\alpha + 2)
\end{aligned}
\tag{1179}
$$

which follows a simple power law in T

$$
C_{V,N} = \overline{N} \ k_B \ (\ \alpha \ + \ 2 \ ) \ (\ \alpha \ + \ 1 \ ) \left(\frac{\xi(\alpha + 2)}{\xi(\alpha + 1)} \right) \left(\frac{T}{T_c} \right)^{\alpha+1}
\tag{1180}
$$

The power law variation is simply understood. At $T = 0$ all the particles occupy the state with $\epsilon = 0$. At finite temperatures, the states with energy less than $k_B \ T$ are occupied. For a density of states proportional to ϵ^α, there are

217

approximately $(\,k_B\,T\,)^{\alpha+1}$ occupied states, each of which carries an energy of approximately $k_B\,T$. Therefore, the total energy is proportional to $(\,k_B\,T\,)^{\alpha+2}$. Hence, the specific heat is proportional to $(\,k_B\,T\,)^{\alpha+1}$.
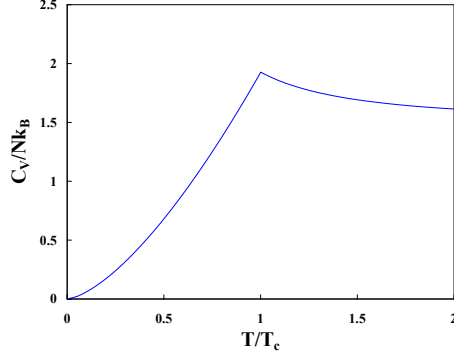
Figure 49: The temperature dependence of the heat capacity for an ideal Bose Gas.

The Cusp in the Heat Capacity at T_c

At the condensation temperature T_c the specific heat has the value of

$$C_{V,N}(T_c) \;=\; \overline{N}\,k_B\,(\,\alpha\,+\,2\,)\,(\,\alpha\,+\,1\,)\left(\frac{\xi(\alpha+2)}{\xi(\alpha+1)}\right) \tag{1181}$$

which for $\alpha = \frac{1}{2}$ takes the value

$$
\begin{aligned}
C_{V,N}(T_c) &= \overline{N}\,k_B\,\frac{15}{4}\left(\frac{\xi(\frac{5}{2})}{\xi(\frac{3}{2})}\right)\\
&= 1.925\,\overline{N}\,k_B \tag{1182}
\end{aligned}
$$

Thus, the specific heat at T_c exceeds the high temperature classical value of $1.5\,\overline{N}\,k_B$. In fact there is a cusp in $C_{V,N}$ at T_c.

The existence of a cusp can be seen by examining the general expression for the heat capacity, valid in the normal liquid and the condensed phase.

$$
\begin{aligned}
C_{V,N} &= \left(\frac{\partial U}{\partial T}\right)_{V,N}\\
&= k_B\,C\,(\,k_B\,T\,)^{\alpha+1}\,\Gamma(\alpha+2)\left[\,(\,\alpha\,+\,2\,)\,I_{\alpha+2}(z)\,+\,T\,\frac{\partial z}{\partial T}\,\frac{\partial}{\partial z}I_{\alpha+2}(z)\,\right]
\end{aligned}
$$
$$\tag{1183}$$

However, since

$$I_{\alpha+1}(z) \; = \; \sum_{m=1}^{\infty} \frac{z^m}{m^{\alpha+1}} \tag{1184}$$

then the derivative w.r.t. z is simply given by

$$\frac{\partial}{\partial z} I_{\alpha+1}(z) \; = \; \frac{1}{z} \, I_\alpha(z) \tag{1185}$$

Thus, the specific heat in the normal phase can be expressed as

$$C_{V,N} \; = \; C \; k_B \; (\, k_B \; T \,)^{\alpha+1} \; \Gamma(\alpha+2) \left[\, (\, \alpha \,+\, 2 \,) \, I_{\alpha+2}(z) \,+\, T \; \frac{\partial z}{\partial T} \; \frac{1}{z} \; I_{\alpha+1}(z) \right] \tag{1186}$$

which requires knowledge of $\frac{\partial z}{\partial T}$ in the normal phase. In the Bose Condensed phase, $z = 1$ so the derivative vanishes. The value of $\frac{\partial z}{\partial T}$ in the normal phase can be found from the condition

$$\left(\frac{\partial \overline{N}}{\partial T} \right)_V \; = \; 0 \tag{1187}$$

which yields

$$\begin{aligned} 0 \; &= \; \left(\frac{\partial \overline{N}}{\partial T} \right)_V \\ &= \; C \, k_B \; (\, k_B \; T \,)^{\alpha} \; \Gamma(\alpha+1) \left[\, (\, \alpha \,+\, 1 \,) \, I_{\alpha+1}(z) \,+\, T \; \frac{\partial z}{\partial T} \; \frac{1}{z} \; I_\alpha(z) \right] \end{aligned} \tag{1188}$$

This has the solution

$$\frac{\partial z}{\partial T} \; = \; -\, (\, \alpha \,+\, 1 \,) \; \frac{z}{T} \; \frac{I_{\alpha+1}(z)}{I_\alpha(z)} \tag{1189}$$

Therefore, the specific heat for the normal phase is given by the expressions

$$\begin{aligned} C_{V,N} \; &= \; C \; k_B \; (\, k_B \; T \,)^{\alpha+1} \; \Gamma(\alpha+2) \left[\, (\, \alpha \,+\, 2 \,) \, I_{\alpha+2}(z) \,-\, (\, \alpha \,+\, 1 \,) \, \frac{I_{\alpha+1}^2(z)}{I_\alpha(z)} \right] \\ &= \; (\, \alpha \,+\, 1 \,) \, \overline{N} \, k_B \left[\, (\, \alpha \,+\, 2 \,) \, \frac{I_{\alpha+2}(z)}{I_{\alpha+1}(z)} \,-\, (\, \alpha \,+\, 1 \,) \, \frac{I_{\alpha+1}(z)}{I_\alpha(z)} \right] \end{aligned} \tag{1190}$$

In the high temperature limit, this can be expanded as

$$C_{V,N} \; \approx \; (\, \alpha \,+\, 1 \,) \, \overline{N} \, k_B \left[\, 1 \,+\, \frac{\alpha}{2^{\alpha+2}} \; z \,+\, \dots \right] \tag{1191}$$

which reaches the classical limit as z approaches zero and increases when z increases. If $\alpha = \frac{1}{2}$, the denominator of the last term in the exact expression diverges at the Bose-Einstein condensation temperature where $z \; = \; 1$. This

occurs as the sum defining $\xi(\frac{1}{2})$ is divergent. The divergence of the $I_{\frac{3}{2}}(z)$ at $z = 1$ causes the last term to vanish and makes the specific heat continuous at T_c. However, the specific heat does have a discontinuity in its slope which is given by

$$\left(\frac{\partial C_{V,N}}{\partial T} \right)\Big|_{T_c-\eta}^{T_c+\eta} = 3.66 \; \frac{\overline{N} \; k_B}{T_c} \tag{1192}$$

A cusp is also seen in the temperature dependence of the specific heat of liquid
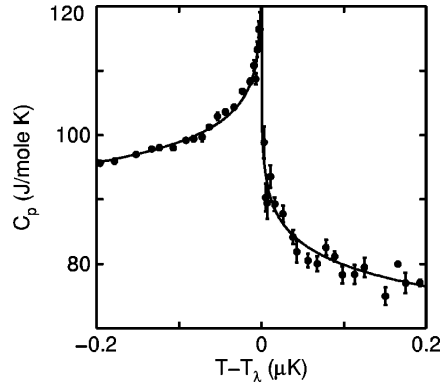
Figure 50: The temperature dependence of the experimentally determined heat capacity of He^4 near the λ point. [J.A. Lipa, J.A. Nissen, D.A. Stricker, D.R. Swanson and T.C.P. Chui, "Specific Heat of Liquid Helium in zero gravity near the Lambda Point.", Phys. Rev B 68, 174518 (2003).]

He^4, which is a signature of the so-called λ transition.

The Pressure in the Condensed Phase

The pressure can be found directly from the Grand-Canonical Potential Ω, as

$$\Omega \; = \; - \, P \, V \tag{1193}$$

which is evaluated as

$$\begin{aligned} \Omega \; &= \; - \, P \, V \\ &= \; k_B \, T \, \int_{-\infty}^{\infty} d\epsilon \; \rho(\epsilon) \; \ln\left[1 \; - \; \exp[-\beta(\epsilon - \mu)] \right] \end{aligned} \tag{1194}$$

with

$$\rho(\epsilon) \; = \; \delta(\epsilon) \; + \; C \; \epsilon^{\alpha} \; \Theta(\epsilon) \tag{1195}$$

This yields

$$- \, P \, V \; = \; k_B \, T \, \ln(\, 1 - z \,) + C \, k_B \, T \, \int_0^{\infty} d\epsilon \; \epsilon^{\alpha} \; \ln\left[1 \; - \; z \; \exp[-\beta\epsilon] \right] \tag{1196}$$

220

On changing variable to $x = \beta \epsilon$ one obtains

$$PV = k_B T \ln N(0) - k_B T \ln z - C (k_B T)^{\alpha+2} \int_0^\infty dx\, x^\alpha\, \ln \left[1 - z\, \exp[-x] \right] \tag{1197}$$

and integrating by parts in the last term, yields

$$PV = k_B T \ln N(0) - k_B T \ln z + \frac{C}{\alpha + 1} (k_B T)^{\alpha+2} \int_0^\infty dx\, \frac{x^{\alpha+1}}{z^{-1} \exp[x] - 1} \tag{1198}$$

where the boundary terms have vanished. Hence, we find the equation of state has the form

$$PV = k_B T \ln N(0) - k_B T \ln z + \frac{C}{\alpha + 1} (k_B T)^{\alpha+2} \Gamma(\alpha+2)\, I_{\alpha+2}(z) \tag{1199}$$

For $T < T_c$, one has $z = 1$, therefore the expression reduces to

$$PV = k_B T \ln N(0) + \frac{C}{\alpha + 1} (k_B T)^{\alpha+2} \Gamma(\alpha+2)\, \xi(\alpha+2) \tag{1200}$$

Since C is proportional to the volume, the first and last terms are extensive, while the logarithmic term is not and can be neglected. This further reduces the equation of state to the form

$$PV = C (k_B T)^{\alpha+2} \Gamma(\alpha+1)\, \xi(\alpha+2) \tag{1201}$$

or equivalently

$$P = \frac{C}{V} (k_B T)^{\alpha+2} \Gamma(\alpha+1)\, \xi(\alpha+2) \tag{1202}$$

which as the volume dependence of C cancels with V, one finds that the pressure is independent of volume. The pressure only depends on temperature, for $T < T_c$, Thus, the isotherms become flat on entering the condensed phase.

Furthermore, since

$$N_{exc} = \Gamma(\alpha+1)\, C (k_B T)^{\alpha+1}\, \xi(\alpha+1) \tag{1203}$$

one can write

$$PV = \left(\frac{\xi(\alpha+2)}{\xi(\alpha+1)} \right) N_{exc}\, k_B T \tag{1204}$$

This makes sense since only the excited particles carry momentum and collide with the walls.

The Entropy of the Condensed Phase

The entropy of the condensate can be found from

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V,\mu} \tag{1205}$$

221
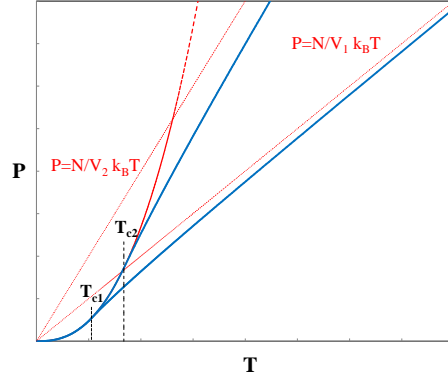
Figure 51: The $P-T$ relations for ideal boson gasses with two different densities (blue) and different T_c's (vertical dashed lines). In the Bose-Einstein condensed phase, the $P - T$ curves collapse onto one curve. The classical asymptotic limiting form for the $P - T$ relations at these two densities are shown by the dashed red lines.
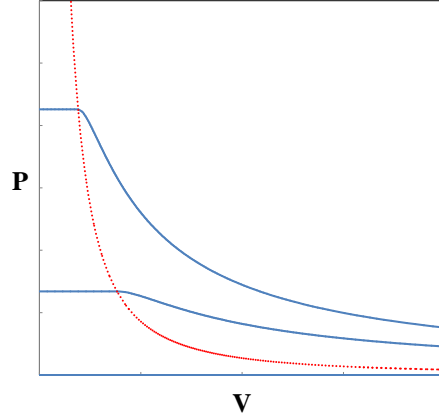
Figure 52: The $P - V$ relations for an ideal boson gas at two different temperatures. Since the number of particles is fixed, the condensation temperature depends on volume $T_c(V) \propto V^{-\frac{3}{2}}$. Thus, the critical pressure P_c varies as $P_c \propto V^{-\frac{5}{3}}$.

which leads to

$$
\begin{aligned}
T\,S &= E \;-\; \Omega \;-\; \mu\,N \\
&= \left(\frac{\alpha+2}{\alpha+1}\right)\,U \;-\; \mu\,\overline{N}
\end{aligned}
\tag{1206}
$$

222

In the Bose condensed phase, $\mu = 0$. Hence, with

$$U = (\alpha + 1) \left(\frac{\xi(\alpha + 2)}{\xi(\alpha + 1)} \right) N_{exc} \, k_B \, T \qquad (1207)$$

one finds that the entropy is given by

$$S = (\alpha + 2) \left(\frac{\xi(\alpha + 2)}{\xi(\alpha + 1)} \right) k_B \, N_{exc} \qquad (1208)$$

This implies that only the excited particles are disordered and contribute to the entropy.

The above discussions relate to non-interacting boson gasses. It is expected that Bose-Einstein condensation may be hindered by interactions, since the density should resemble the squared modulus of the lowest energy single particle wave function and exhibit the same non-uniformity in space. Local interactions are expected to make the fluid's density uniform.

Homework:

Rb^{87} was reported to Bose Condense[32]. The gas was trapped in a three-dimensional harmonic potential with frequency $\omega_0 \approx 750$ Hz.

(i) Determine an approximate form for the density of states.

(ii) For a number density of 2.5×10^{12} per cm^3, estimate the Bose-Einstein condensation temperature and compare it to the reported value of $T_c \approx 170$ nK.

9.7 Superfluidity

He has two electrons located in the $1s$ orbitals, so it has a closed atomic shell and is relatively chemically inert. There does exist a van der Waals interaction between the He atoms. The interatomic potential consists of the sum of the short-ranged exponential repulsion and the attractive van der Waals interaction[33]. The potential has a weak minimum of depth 9 K, at a distance of about 3 Å. Since the atom is relatively light, it doesn't solidify easily. If the atoms did solidify, with a lattice spacing d, then the uncertainty in the momentum Δp of any atom would be of the order of $\hbar/d$. The kinetic energy would be of the order of

$$\frac{\Delta p^2}{2 \, m} \approx \frac{\hbar^2}{2 \, m \, d^2} \qquad (1209)$$

[32]M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman and E.A. Cornell, "Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor", Science, 269, 198-201, (1995).
[33]J.C. Slater and J.G. Kirkwood, "The van der waals forces in gases", Phys. Rev. 37, 682-697 (1931).
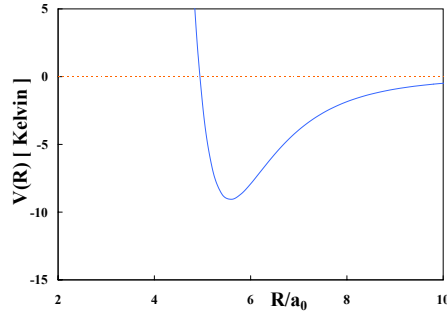
Figure 53: The interaction potential of atomic Helium as a function of radial distance R/a_0, where a_0 is the Bohr radius.

for $d \sim 3$ Å, this energy is greater than the minimum of the potential. Hence, the lattice would melt. Therefore, He remains a liquid at ambient pressures. However, He interacts quite strongly so at pressures of the order of 25 atmospheres, it can solidify at very low temperatures.

The isotope He^3 is a fermion and form a Fermi-liquid at low temperatures. On the other hand, He^4 is a boson and obeys Bose-Einstein Statistics. He^4 can be cooled by a process of evaporation. When it is cooled, the Helium becomes turbulent, just like boiling water. However, at a temperature of 2.2 K it suddenly becomes clear and the turbulence disappears. This signals a phase change
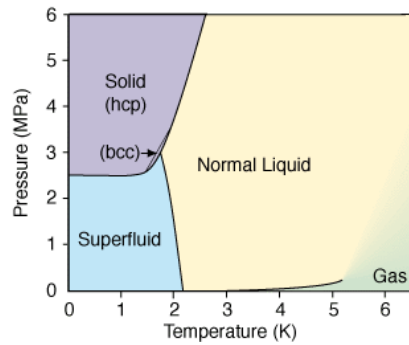
Figure 54: The $P - T$ phase diagram of He^4. He^4 remains a fluid for pressures below 2.5 MPa. The Liquid-Gas critical point is located at a temperature of 5.2 K. The liquid phase undergoes a further transition from He I (the normal liquid) to He II (the superfluid), as the temperature is reduced below 2.18 K.

from the He I phase, to He II.

He II has unusual properties, it flows as if it has a vanishing viscosity, and can flow through narrow capillaries[34]. That is, it exhibits superflow. On the other hand, when the fluid is placed in a cavity contain closely spaced disks that are fee to rotate, and the disks are forced to perform oscillations through some small angles (torsional oscillations), then the effective moment of inertia is temperature dependent. The increased moment of inertia indicates that some fluid is being dragged by the rotating disks. Above the critical temperature T_c it was found by Andronikashvilli[35] that all the fluid was being dragged by the plates, but below T_c the moment of inertia decreased. in fact, the fraction the fluid which is dragged by the rotating disks is found to vanish as T is decreased to zero.

In 1938 Fritz London[36] proposed that the transition to He II is related to Bose-Einstein condensation. The reasons for this proposal was:

(i) The observed critical temperature T_c has the magnitude of 2.18 K, whereas the Bose-Einstein condensation temperature for free bosons is given by

$$k_B \ T_c \ = \ \frac{2 \ \pi \ \hbar^2}{2 \ m} \left[\ \xi \left(\frac{3}{2} \right) \frac{V}{N} \ \right]^{2/3} \tag{1210}$$

which is calculated as 3.14 K.

(ii) The transition occurs in the boson system He^4 but not in the fermionic system He^3.

(iii) For the Bose condensed phase of an ideal gas, the pressure is independent of the density and only a function of temperature. Hence, pressure variations are directly related to temperature variations, as is true for He II. The specific heat of He II also shows a λ anomaly at T_c, whereas the ideal Bose gas is expected to exhibit a cusp at T_c. However, lambda-like anomalies in the specific heat are universally found at second-order phase transitions.

Superfluid He differs from the Bose-Einstein condensate in many details, primarily due to the effects of interactions. For example, an ideal Bose condensate would not be a superfluid. However, interactions can change the form of the dispersion relation for the low-energy excitations so that He II is a superfluid.

The Excitation Spectrum of a Weakly-Interacting Bose Condensate

---

[34]P. Kapitza, "Viscosity of Liquid Helium below the λ-point", Nature, 141, 74-74, (1938), J.F. Allen and A.D. Misener, "Flow of Liquid Helium II", Nature, 141, 75-75, (1938).
[35]E.I. Andronikashvilli, "Direct Observation of two kinds of motion for Helium II", Zh. Eksp. Theor. Fiz., 16, 780 (1946).
[36]F. London, "The Lambda Phenomenon of Liquid Helium and the Bose-Einstein Degeneracy", Nature, 141, 643 (1938).

Consider a gas of weakly-interacting bosons contained in a volume V. The Hamiltonian is described by

$$\hat{H} = \sum_{\alpha;\alpha'} <\alpha'\,|\,\frac{\hat{p}^2}{2\,m}\,|\,\alpha> \hat{a}^\dagger_{\alpha'}\,\hat{a}_\alpha + \frac{1}{2!} \sum_{\alpha,\beta;\alpha',\beta'} <\alpha'\,\beta'\,|\,\hat{V}_{int}\,|\,\alpha\,\beta> \hat{a}^\dagger_{\alpha'}\,\hat{a}^\dagger_{\beta'}\,\hat{a}_\beta\,\hat{a}_\alpha$$

(1211)

where the two-body interaction represents a short-ranged repulsive interaction. The one-body part can be diagonalized by choosing the single-particle wave functions $\phi_\alpha(r)$ to be momentum eigenstates

$$\phi_{\underline{k}}(\underline{r}) = \frac{1}{\sqrt{V}}\,\exp\left[\,i\,\underline{k}\,\cdot\,\underline{r}\,\right]$$

(1212)

The Hamiltonian can be expressed as

$$\hat{H} = \sum_{\underline{k}} \frac{\hbar^2\,\underline{k}^2}{2\,m}\,\hat{a}^\dagger_{\underline{k}}\,\hat{a}_{\underline{k}} + \frac{1}{2!\,V} \sum_{\underline{k},\underline{k}';\underline{q}} V_{int}(\underline{q})\,\hat{a}^\dagger_{\underline{k}-\underline{q}}\,\hat{a}^\dagger_{\underline{k}'+\underline{q}}\,\hat{a}_{\underline{k}'}\,\hat{a}_{\underline{k}}$$

(1213)

where the scattering term conserves momentum and $V_{int}(\underline{q})$ is the Fourier Transform of the interaction potential. Therefore, the two-body scattering does not change the total momentum of the system. We shall assume that the potential is sufficiently short-ranged so that the limit, $\lim_{q\to 0} V_{int}(\underline{q}) = V_{int}(0)$, is well-defined.

At sufficiently low temperatures, the bosons are expected to form a condensate. Let the number of particles in the condensate be N_0 and we shall assume that N_0 is much larger than the number of excited particles N_{exc}

$$\overline{N}_{exc} = \sum_{\underline{k}} \overline{\hat{a}^\dagger_{\underline{k}}\,\hat{a}_{\underline{k}}}$$

(1214)

so

$$\overline{N}_0 = N - \sum_{\underline{k}} \overline{\hat{a}^\dagger_{\underline{k}}\,\hat{a}_{\underline{k}}}$$

(1215)

The condensate, contains a large number of bosons with $\underline{k} = 0$ and, therefore can be considered to be a coherent state. The expectation value of the creation and annihilation $\hat{a}^\dagger_0$ and $\hat{a}_0$, respectively, can be replaced by the complex numbers a_0^* and a_0. Thus, the interaction Hamiltonian can be expanded in powers of a_0 or a_0^* as

$$\begin{aligned} \hat{H} &= \sum_{\underline{k}} \frac{\hbar^2\,\underline{k}^2}{2\,m}\,\hat{a}^\dagger_{\underline{k}}\,\hat{a}_{\underline{k}} + \frac{1}{2!\,V}\,|\,a_0\,|^4\,V_{int}(0) \\[2mm] &+ \frac{1}{2!\,V} \sum_{\underline{k}} V_{int}(\underline{k})\left(\,a_0^2\,\hat{a}^\dagger_{\underline{k}}\,\hat{a}^\dagger_{-\underline{k}} + a_0^{*2}\,\hat{a}_{-\underline{k}}\,\hat{a}_{\underline{k}}\,\right) \\[2mm] &+ \frac{2}{2!\,V} \sum_{\underline{k}} V_{int}(0)\,|\,a_0\,|^2\left(\,\hat{a}^\dagger_{\underline{k}}\,\hat{a}_{\underline{k}} + \hat{a}^\dagger_{-\underline{k}}\,\hat{a}_{-\underline{k}}\,\right) + \ldots \end{aligned}$$

(1216)

226

Terms cubic in a_0 and a_0^* are forbidden due to the requirement of conservation of momentum. In this expression, we have ignored terms involving more than two excited boson creation or annihilation operators. The above form of the Hamiltonian contains terms involving unbalanced creation and annihilation operators. The term with two creation operators represent processes in which two bosons are scattered out of the condensate and the term with two annihilation operators represents the absorbtion of two bosons into the condensate. On replacing $| a_0 |^2$ by $N - \sum_{\underline{k}} a_{\underline{k}}^\dagger a_{\underline{k}}$, one finds

$$
\begin{aligned}
\hat{H} \;=\;& \sum_{\underline{k}} \frac{\hbar^2 \, \underline{k}^2}{2 \, m} \, \hat{a}_{\underline{k}}^\dagger \, \hat{a}_{\underline{k}} \;+\; \frac{N^2}{2! \, V} \; V_{int}(0) \\[2mm]
& + \frac{N}{2! \, V} \, \sum_{\underline{k}} V_{int}(\underline{k}) \, \left(\; \exp[\, + 2 \, i \, \varphi \,] \, \hat{a}_{\underline{k}}^\dagger \, \hat{a}_{-\underline{k}}^\dagger \;+\; \exp[\, - 2 \, i \, \varphi \,] \, \hat{a}_{-\underline{k}} \, \hat{a}_{\underline{k}} \; \right) \\[2mm]
& + \frac{N}{2! \, V} \, \sum_{\underline{k}} V_{int}(0) \, \left(\; \hat{a}_{\underline{k}}^\dagger \, \hat{a}_{\underline{k}} \;+\; \hat{a}_{-\underline{k}}^\dagger \, \hat{a}_{-\underline{k}} \; \right) \;+\; \dots
\end{aligned}
\tag{1217}
$$

where φ is a constant phase. When the condensate adopts a phase, the continuous $U(1)$ phase symmetry of the Hamiltonian has been spontaneously broken.

The Hamiltonian can be put in diagonal form[37] by using a suitably chosen unitary transformation $\hat{U}$, so

$$
\hat{H}' \;=\; \hat{U} \, \hat{H} \, \hat{U}^\dagger
\tag{1218}
$$

The transformed Hamiltonian $\hat{H}'$ has the same spectrum of eigenvalues as the original Hamiltonian $\hat{H}$. We shall choose the transformation to be of the form

$$
\hat{U} \;=\; \exp\left[\; - \sum_{\underline{k}} \Theta_k \left(\, a_{\underline{k}}^\dagger \, a_{-\underline{k}}^\dagger \, \exp[\, + 2 \, i \, \varphi_k \,] \;-\; a_{-\underline{k}} \, a_{\underline{k}} \, \exp[\, - 2 \, i \, \varphi_k \,] \, \right) \right]
\tag{1219}
$$

which is unitary when Θ_k is real. The creation operators transform as

$$
\begin{aligned}
\alpha_{\underline{k}}^\dagger \;=\;& \hat{U} \, a_{\underline{k}}^\dagger \, \hat{U}^\dagger \\[1mm]
=\;& \cosh \Theta_k \, a_{\underline{k}}^\dagger \;+\; \sinh \Theta_k \, \exp[\, - 2 \, i \, \varphi_k \,] \, a_{-\underline{k}}
\end{aligned}
\tag{1220}
$$

and the annihilation operators are found to transform as

$$
\begin{aligned}
\alpha_{\underline{k}} \;=\;& \hat{U} \, a_{\underline{k}} \, \hat{U}^\dagger \\[1mm]
=\;& \cosh \Theta_k \, a_{\underline{k}} \;+\; \sinh \Theta_k \, \exp[\, + 2 \, i \, \varphi_k \,] \, a_{-\underline{k}}^\dagger
\end{aligned}
\tag{1221}
$$

Since the transformation is unitary, it does not affect the canonical conjugate commutation relations

$$
[\, \hat{\alpha}_{\underline{k}} \,,\, \hat{\alpha}_{\underline{k}'}^\dagger \,] \;=\; \delta_{\underline{k},\underline{k}'}
\tag{1222}
$$

---

[37] N.N. Bogoliubov, "On the Theory of Superfluidity", J. Phys. USSR, 11, 23 (1947).

and

$$[\ \hat{\alpha}_{\underline{k}}^\dagger\ ,\ \hat{\alpha}_{\underline{k}'}^\dagger\ ]\ =\ 0$$
$$[\ \hat{a}_{\underline{k}}\ ,\ \hat{\alpha}_{\underline{k}'}\ ]\ =\ 0 \tag{1223}$$

The transformed Hamiltonian takes the form

$$
\begin{aligned}
\hat{H}' \ =\ & \frac{N^2}{2!\,V}\ V_{int}(0) \\[2mm]
& +\ \sum_{\underline{k}}\left(\frac{\hbar^2\,\underline{k}^2}{2\,m}+\frac{N}{V}\,V_{int}(0)\right)\left[\ \cosh^2\Theta_k\ \hat{a}_{\underline{k}}^\dagger\,\hat{a}_{\underline{k}}\ +\ \sinh^2\Theta_k\ \hat{a}_{-\underline{k}}\,\hat{a}_{-\underline{k}}^\dagger\ \right] \\[2mm]
& +\ \frac{N}{V}\sum_{\underline{k}}V_{int}(\underline{k})\ \sinh\Theta_k\ \cosh\Theta_k\ \cos[2(\varphi-\varphi_k)]\left[\ \hat{a}_{\underline{k}}^\dagger\,\hat{a}_{\underline{k}}\ +\ \hat{a}_{-\underline{k}}\,\hat{a}_{-\underline{k}}^\dagger\ \right] \\[2mm]
& +\ \sum_{\underline{k}}\left(\frac{\hbar^2\,\underline{k}^2}{2\,m}+\frac{N}{V}\,V_{int}(0)\right)\sinh\Theta_k\cosh\Theta_k\left[\ \exp[+2i\varphi_k]\,\hat{a}_{\underline{k}}^\dagger\,\hat{a}_{-\underline{k}}^\dagger\ +\ \exp[-2i\varphi_k]\,\hat{a}_{-\underline{k}}\,\hat{a}_{\underline{k}}\ \right] \\[2mm]
& +\ \frac{N}{2!\,V}\sum_{\underline{k}}V_{int}(\underline{k})\left(\cosh^2\Theta_k+\sinh^2\Theta_k\right)\left[\ \exp[+2i\varphi]\,\hat{a}_{\underline{k}}^\dagger\,\hat{a}_{-\underline{k}}^\dagger\ +\ \exp[-2i\varphi]\,\hat{a}_{-\underline{k}}\,\hat{a}_{\underline{k}}\ \right]
\end{aligned}
\tag{1224}
$$

when written in terms of the original creation and annihilation operators. The terms non-diagonal in the particle creation and annihilation operators can be eliminated by the appropriate choice of Θ_k and φ_k. We shall set φ_k equal to the phase of the condensate, $\varphi_k = \varphi$. Then, the off-diagonal terms vanish if one chooses Θ_k to satisfy

$$\tanh 2\Theta_k\ =\ -\ \frac{\frac{N}{V}\,V_{int}(\underline{k})}{\frac{\hbar^2\,\underline{k}^2}{2\,m}+\frac{N}{V}\,V_{int}(0)} \tag{1225}$$

With this choice, the Hamiltonian reduces to

$$
\begin{aligned}
\hat{H}'\ =\ & \frac{N^2}{2!\,V}\,V_{int}(0)\ +\ \frac{1}{2}\sum_{\underline{k}}\left(E(\underline{k})-\frac{\hbar^2\,\underline{k}^2}{2\,m}-\frac{N}{V}\,V_{int}(0)\right) \\[2mm]
& +\ \sum_{\underline{k}}E(k)\,\hat{a}_{\underline{k}}^\dagger\,\hat{a}_{\underline{k}}
\end{aligned}
\tag{1226}
$$

where the first line represents the ground state energy and the second line represents the energy of the elementary excitations. The energy of the elementary excitation $E(\underline{k})$ is given by

$$E(\underline{k})\ =\ \sqrt{\left(\frac{\hbar^2\,\underline{k}^2}{2\,m}+\frac{N}{V}\,V_{int}(0)\right)^2-\left(\frac{N}{V}\,V_{int}(\underline{k})\right)^2} \tag{1227}$$

228

This dispersion relation vanishes identically at $k = 0$ and is approximately linear for small k,

$$E(\underline{k}) \approx k \sqrt{\frac{\hbar^2}{m} \frac{N}{V} V_{int}(0)} \tag{1228}$$

where the excitations have the characteristics of phonons. The excitations are the Goldstone modes associated with the broken gauge symmetry[38]. It should be noted that $V_{int}(0)$ must be positive for this solution to be stable. At higher values of k the dispersion relation reduces to

$$E(\underline{k}) \approx \frac{\hbar^2 \underline{k}^2}{2 m} + \frac{N}{V} V_{int}(0) \tag{1229}$$

which represents the bare particle dispersion relation together with a constant energy shift due to the interaction with the particles in the condensate.

In summary, one observes that due to the interactions with the particles in the condensate, the dispersion of the elementary excitations has changed from quadratic to being linear. This has the important experimental consequence that the specific heat changes from being proportional to $T^{\frac{3}{2}}$ at low temperatures to having a T^3 variation.

The Coherent Nature of The Ground State

The replacement of the condensate creation and annihilation operators may have obscured the physics, specially since the unitary transformation does not conserve the number of particles. An appropriate generalization to the case of conserved particles is given by

$$\hat{U} = \exp\left[-\sum_{\underline{k}} \frac{\Theta_k}{N} \left(a_{\underline{k}}^\dagger a_{-\underline{k}}^\dagger \hat{a}_0 \hat{a}_0 - \hat{a}_0^\dagger \hat{a}_0^\dagger a_{-\underline{k}} a_{\underline{k}} \right) \right] \tag{1230}$$

where, for convenience, we have set the phase of the condensate to zero. The states of the untransformed system can be obtain from those of the transformed system by the inverse transformation

$$| \Psi > = \hat{U}^\dagger | \Psi' > \tag{1231}$$

In the primed frame, the ground state ia an eigenstate of the number operator $\hat{n}_{\underline{k}} = \hat{a}_{\underline{k}}^\dagger \hat{a}_{\underline{k}}$ with eigenvalue zero. Hence, in the primed frame, the ground state simply corresponds to N bosons in the condensate

$$| \Psi' > = \frac{1}{\sqrt{N!}} (\hat{a}_0^\dagger)^N | 0 > \tag{1232}$$

---

[38] Gauge symmetry is broken since each condensate has a particular value for the phase φ, whereas the theory shows that the energy of the Bose-condensed state is independent of φ, where φ is a continuous variable that lies in the range $2\pi > \varphi > 0$. Hence, the Bose-condensate has broken the continuous phase-symmetry of the Hamiltonian.

where $| \, 0 \, >$ is the vacuum. Since the vacuum satisfies

$$\hat{U}^{\dagger} \, | \, 0 \, > \, = \, | \, 0 \, > \qquad (1233)$$

one finds that the ground state in the un-transformed system is given by

$$
\begin{aligned}
| \, \Psi \, > \; &= \; \frac{1}{\sqrt{N!}} \, \hat{U}^{\dagger} \, (\, \hat{a}_0^{\dagger} \,)^N \, \hat{U} \, \hat{U}^{\dagger} \, | \, 0 \, > \\
&= \; \frac{1}{\sqrt{N!}} \, \hat{U}^{\dagger} \, (\, \hat{a}_0^{\dagger} \,)^N \, \hat{U} \, | \, 0 \, > \\
&= \; \frac{1}{\sqrt{N!}} \, \left(\, \hat{U}^{\dagger} \, \hat{a}_0^{\dagger} \, \hat{a}_0^{\dagger} \, \hat{U} \, \right)^{\frac{N}{2}} \, | \, 0 \, > \qquad (1234)
\end{aligned}
$$

Thus, the ground state of the condensate has the form of a product of linear superpositions[39]

$$
\begin{aligned}
| \, \Psi \, > \; &\approx \; \frac{1}{\sqrt{N!}} \, \left(\, \hat{a}_0^{\dagger} \, \hat{a}_0^{\dagger} \, - \, \frac{2}{N} \, \sum_{\underline{k}} \, \Theta_k \, \hat{a}_{\underline{k}}^{\dagger} \, \hat{a}_{-\underline{k}}^{\dagger} \, + \dots \, \right)^{\frac{N}{2}} \, | \, 0 \, > \\
&\approx \; \frac{1}{\sqrt{N!}} \, \left(\, \hat{a}_0^{\dagger} \, \hat{a}_0^{\dagger} \, - \, \sum_{\underline{k}} \, \tanh \Theta_k \, \hat{a}_{\underline{k}}^{\dagger} \, \hat{a}_{-\underline{k}}^{\dagger} \, \right)^{\frac{N}{2}} \, | \, 0 \, > \qquad (1235)
\end{aligned}
$$

in which it is seen that the interaction has scattered pairs of bosons out of the condensate. Conservation of momentum shows that the pairs of particles scattered out of the condensate have zero total momentum. Thus, the number of particles with zero momentum is smaller than the number of particles. The ground state is a form of coherent state, in the sense the number of particles in the condensate is large as is the number fluctuations. It is also seen that the components of the ground state with different numbers of particles in the condensate have definite phase relationships.

The states with a single elementary excitation present are proportional to

$$\hat{U}^{\dagger} \, \hat{a}_{\underline{k}}^{\dagger} \, \hat{U} \, | \, \Psi \, > \, = \, \left(\, \cosh \Theta_k \, \hat{a}_{\underline{k}}^{\dagger} \, \hat{a}_0 \, - \, \sinh \Theta_k \, \hat{a}_0^{\dagger} \, \hat{a}_{-\underline{k}} \, \right) \, | \, \Psi \, > \qquad (1236)$$

Hence, the elementary excitations of the Bose-Einstein condensate are of the form of a linear superposition. The relative weight of the single-particle excitation $a_{\underline{k}}^{\dagger}$ in this excited state is significantly reduced for small k.

Thus, the interaction not only produces a change in the dispersion relations of the excitations of a Bose-Einstein condensate, but also changes the character of the excitations.

---

[39]M. Girardeau and R. Arnowitt, "Theory of Many-Boson Systems: Pair Theory", Phys. Rev. 113, 755-761 (1959).

The Critical Velocity

The change in the character of the dispersion relation at low energies has the consequence that the condensate of a weakly-interacting Bose-Einstein particles exhibits superfluidity. Superfluidity is the property of flowing through narrow capillaries without exhibiting viscosity. Thus, a superfluid will flow in the absence of any driving forces.

We shall consider the flow of a fluid at zero temperature, flowing with velocity $\underline{v}$ through a capillary. In the primed frame of reference moving with the fluid, the walls of the capillary are moving with velocity $-\underline{v}$, In this primed reference frame, consider the fluid to initially be in a Bose-Einstein Condensate and carries no momentum. The total energy of the fluid would be $M\ c^2$, the rest mass energy, and the total momentum is zero. If the viscosity were to be finite, the interaction with the moving capillary walls would cause the fluid to start moving. The change in the state of the initial fluid could only be caused by exciting the internal degrees of freedom of the superfluid. That is, if the viscosity is present, the interaction with the moving capillary walls should produce an excitation in the liquid. In the primed reference frame, the total energy for the condensate in which there is an excitation with momentum $\underline{p}'$ is given by

$$E_T' \ = \ M\ c^2 \ + \ E(\underline{p}') \tag{1237}$$

In the reference frame where the capillary walls are stationary, this energy is given by the Lorentz transformation

$$\begin{aligned} E_T \ &= \ \frac{1}{\sqrt{1 - v/c^2}} \ (\ E_T' \ + \ \underline{v} \cdot \underline{p}'\ ) \\ &\approx \ M\ c^2 \ + \ E(\underline{p}') \ + \ \underline{p}' \cdot \underline{v} \ + \ \frac{M}{2}\ \underline{v}^2 \end{aligned} \tag{1238}$$

and the momentum in the rest frame is given by

$$\begin{aligned} \underline{p} \ &= \ \frac{1}{\sqrt{1 - v/c^2}} \ (\ \underline{p}' \ + \ \underline{v}\ \frac{E'}{c^2}\ ) \\ &\approx \ \underline{p}' \ + \ M\ \underline{v} \end{aligned} \tag{1239}$$

Hence, the energy of the excitation in the stationary reference frame is given by

$$\Delta E \ = \ E(\underline{p}') \ + \ \underline{p}' \cdot \underline{v} \tag{1240}$$

and its momentum is $\underline{p}'$. The excitation energy must be negative, if the excitation is to be allowed. The rationale for this is that since the capillary is at rest and at $T = 0$, it cannot provide the energy necessary to create a positive-energy excitation. On the other hand, since the fluid is moving, it can lose energy by reducing its state of motion and dissipate the excess energy through the creation

of positive-energy excitations in the capillary's walls. Hence, we require that the excitation energy is negative

$$E(\underline{p}') \; + \; \underline{p}' \cdot \underline{v} \; < \; 0 \tag{1241}$$

It is possible to satisfy this criterion if $\underline{p}'$ is anti-parallel to $\underline{v}$. After the excitation has occurred, the liquid is expected to have slowed down.
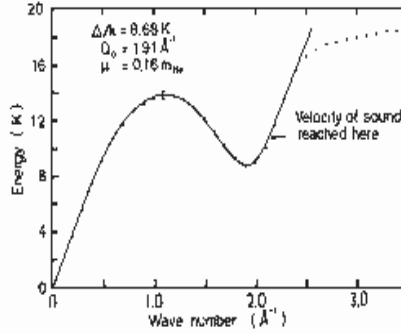
Figure 55: The experimentally determined excitation spectra of He II at $T = 1.1$ K. [R.A. Cowley, and A.D.B. Woods, "Inelastic Scattering of Thermal Neutrons From Liquid Helium", Can. J. Phys. 49, 177 (1971).]

The criterion for dissipation to occur in the fluid is that

$$E(\underline{p}') \; + \; \underline{p}' \cdot \underline{v} \; < \; 0 \tag{1242}$$

On assuming v and p' are oppositely directed, the criterion for viscous flow reduces to

$$v \; > \; \min E(p')/p' \tag{1243}$$

for some value of p'. The critical velocity v_c is defined as

$$v_c \; = \; \min E(p')/p' \tag{1244}$$

The Landau criterion[40] states that superflow occurs when $v_c \; > \; v \; > \; 0$ and viscous flow occurs when $v \; > \; v_c$. Geometrically, the critical velocity is the minimum value of the slope of a line from the origin to a point p_0 on the curve $E(p)$. This point is given by the solution of

$$\frac{E(p_0)}{p_0} \; = \; \left. \frac{dE(p)}{dp} \right|_{p_0} \tag{1245}$$

Any Bose-Einstein condensate with a parabolic dispersion relation $E(\underline{p})$ cannot exhibit superflow, since $p_0 = 0$. Therefore, for a Bose-Einstein condensate with

[40]L.D. Landau, "The Theory of Superfluidity of He II", J. Phys. USSR, 5, 71, (1941).

a parabolic dispersion relation $v_c = 0$ so any flow is viscous. For He^4, the theoretically calculated and experimentally measured dispersion relation $E(p)$ is linear at small p but exhibits a minimum at some finite value of p, which is known as the roton minimum. The roton minima is caused by the strong interactions between the particles which are almost strong enough to result in the formation of a crystal. In fact Feynman has shown that the dispersion relation can be written in terms of the geometric structure factor $S(p)$ via

$$E(p) = \frac{p^2}{2\ m\ S(p)} \tag{1246}$$

For large p, the structure factor has a maximum at a momentum p_0 given by $p_0 \sim \frac{\hbar}{d}$ where d is the interparticle spacing. The maximum in the structure factor gives rise to the observed roton minimum. However, the observed roton minimum gives a critical velocity of 60 m/sec. This velocity is several orders of magnitude greater than the critical velocity observed in He II. Feynman[41] has shown that the critical velocity is very close to that expected for a single vortex ring.

9.8   The Superfluid Velocity and Vortices

A spatially varying condensate can be characterized by a spatially varying wave function $\psi(\underline{r})$ which is given by the creation amplitude for the particles in the condensate. Then the total number of particles in the condensate N_0 is given by

$$N_0 = \int d^3\underline{r} \mid \psi(\underline{r}) \mid^2 \tag{1247}$$

The spatial variation of the condensate can be described by a phenomenological Landau-Free-Energy

$$F[\psi, \psi^*] = \int d^3\underline{r} \left[\frac{\hbar^2}{2\ m} \mid \underline{\nabla}\ \psi \mid^2 + V(\underline{r}) \mid \psi \mid^2 + V_{int} \mid\psi\mid^4 \right] \tag{1248}$$

where the last term represents a localized interaction between the particles in the condensate. Minimization of the Free-Energy w.r.t ψ^* yields a Schrödinger-like equation with non-linear terms. Conservation of particles requires that the condensate density ρ and current density $\underline{j}$ must satisfy the continuity equation

$$\left(\frac{\partial\rho}{\partial t} \right) + \underline{\nabla} \cdot \underline{j} = 0 \tag{1249}$$

where

$$\rho(\underline{r}) = \mid \psi \mid^2 \tag{1250}$$

---

[41]R.P. Feynman and M. Cohen, "Energy Spectrum of the Excitations in Liquid Helium", Physical Review, 102, 1189-1204, (1956).

and
$$\underline{j}(\underline{r}) \;=\; \frac{\hbar}{2\,m\,i}\left(\psi^{*}\,\underline{\nabla}\psi \;-\; \underline{\nabla}\psi^{*}\,\psi\right) \tag{1251}$$

If the amplitude of the condensate wave function varies slowly compared to the phase, one may write

$$\psi(\underline{r}) \;=\; \sqrt{\frac{N_0}{V}}\;\exp\left[\,i\,\varphi(\underline{r})\,\right] \tag{1252}$$

Hence, one finds that the condensate's current density is given by

$$\underline{j}(\underline{r}) \;=\; \frac{\hbar}{m}\,\frac{N_0}{V}\,\underline{\nabla}\varphi \tag{1253}$$

This allows one to define a superfluid velocity $\underline{v}_s$ via

$$\underline{v}_s \;=\; \frac{\hbar}{m}\,\underline{\nabla}\varphi \tag{1254}$$

which is governed by the spatial variation of the phase of the condensate.

The circulation of the superfluid velocity field is given by the line integral around a closed loop inside the superfluid

$$\oint d\underline{r}\cdot\underline{v}_s \;=\; \frac{\hbar}{m}\,\oint d\underline{r}\cdot\underline{\nabla}\varphi \tag{1255}$$

This equation relates the superfluid circulation to the change of the phase of the condensate wave function at the end point of the loop. Since the condensate wave function must be single-valued, the phase at any point must be defined up to a multiple of $2\,\pi$. Thus, from continuity of the wave function, the circulation must be quantized[42]

$$\oint d\underline{r}\cdot\underline{v}_s \;=\; \frac{\hbar}{m}\,n\,2\,\pi \tag{1256}$$

where n is an integer, $n \;=\; 0,\pm 1,\pm 2,\ldots$. The quantization of circulation is a manifestation of the quantum nature of the superfluid. If one now considers loop in the condensate which is simply connected and shrinks the size of the loop to zero, one finds that the only possible value of the phase quantum number n is zero. Hence, the condensate must be irrotational. On the other hand, if the condensate is multiply connected, then one cannot shrink the loop to zero, and the quantum numbers n can be non-zero. Vortices with low numbers of circulation quanta are preferred energetically. Thus, the circulation of the superfluid can be non-zero when the loop encloses regions in which the condensate density is zero. This analysis has the experimental consequence that if one starts to rotate a cylindrical vessel containing a superfluid, the superfluid liquid will initially remain at rest. However, if the angular velocity is increased

---

[42]L. Onsager, "*Statistical Hydrodynamics*," Nuovo Cim. 6 Suppl. 2 pp. 279-287 (1949).

Ω
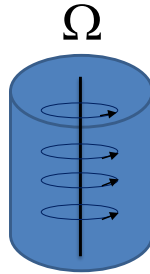
Figure 56: A single vortex excitation in a superfluid contained in a cylinder rotating with angular velocity Ω.
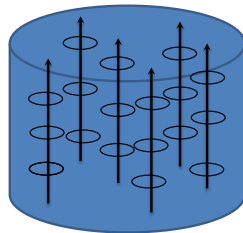
Figure 57: The configuration of vortex excitations in a superfluid contained within a rotating cylinder.

above a critical value, excitations consisting of normal regions called vortices are introduced into the superfluid. A vortex consists of a one-dimensional region of normal fluid (i.e. non-superfluid) that is oriented parallel to the axis of rotation and extends throughout the entire length of the cylinder [See fig(56).]. The vortices are topological excitations. An increase in the angular velocity of the cylinder will result in an increase in the number of vortices that penetrate the superfluid, and the vortex lines will form a two-dimensional array[43]. In a
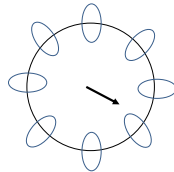
Figure 58: A vortex ring moving through a superfluid.

[43]H.E. Hall and W.F. Vinen, *The Rotation of Liquid He II*", Proc. Roy. Soc. London, Series A, **238**, 204-215 (1956): **238**, 215-234 (1956).

non-rotating superfluid, it is possible to have vortices, if the vortices close up on themselves in loops. Such vortex rings have the form of "smoke rings" [See fig(58).]. In a flowing liquid, large vortex rings can be created before the critical velocity for the creation of rotons is reached. Hence, the critical velocity of a superfluid is usually determined by the vortices. If the arial dimension of the system is reduced below the size of the vortex rings, the critical velocity is bounded from above by the critical velocity inferred from the roton excitations.

10 Phase Transitions

In 1944 Lars Onsager [44] published the exact solution of the two-dimensional Ising Model in zero-field on a square lattice. His exact solution was a *tour de force* of mathematical physics. Onsager used the transfer matrix technique to describe a finite square lattice of size $L \times L$ and then diagonalized the matrix by finding the irreducible representations of a related matrix algebra. The exact solution demonstrated that, in thermodynamic limit, the system exhibited a phase transition which is marked by singularities in the physical properties. The form of the Hamiltonian of the Ising Model is given by

$$\hat{H} \; = \; - \sum_{i,j} J_{i,j} \; S_i^z \; S_j^z \tag{1257}$$

where $J_{i,j} \; = \; J$ if i and j are on neighboring lattice sites and $J_{i,j} \; = \; 0$ otherwise. The spin variables can only have the allowed values $S_i^z \; = \; \pm 1$. We shall assume $J \; > \; 0$ which favors states where the neighboring S^z values have the same sign. Onsager found that the exact partition function is given by the expression

$$\frac{\ln Z}{N} \; = \; \ln 2 + \frac{a^2}{2 \, (\, 2 \, \pi \,)^2} \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} dk_x \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} dk_y \, \ln \left[\, \cosh^2(2\beta J) - \sinh(2\beta J) \, (\, \cos k_x a + \cos k_y a \,) \, \right] \tag{1258}$$

The argument of the logarithm is non-negative and the integral exists for all values of βJ. For $J \; > \; 0$, the minimum value of the argument occurs for $\underline{k} = 0$, and is given by

$$\cosh^2(2\beta J) \; - \; 2 \, \sinh(2\beta J) \; = \; (\, 1 - \, \sinh(2\beta J) \,)^2 \tag{1259}$$

The non-analytic behavior of F occurs at the temperature when this minimum value vanishes. This gives rise to the identification of the critical temperature as the solution of the equation

$$\sinh(2\beta J) \; = \; 1 \tag{1260}$$

or equivalently

$$\tanh \beta_c J \; = \; \sqrt{2} \; - \; 1 \tag{1261}$$

---

[44] L. Onsager, Phys. Rev. **65**, 117 (1944).

For temperatures below the critical temperature, the system is in a ferromagnetic magnetic state. The non-analyticity originates with the long wavelength behavior of the integral, and can be found by approximating the integral by

$$
\frac{\ln Z}{N} \approx \ln 2 + \frac{a^2}{2\,(\,2\,\pi\,)} \int_0^{\frac{\pi}{a}} dk\ k\ \ln\left[\ (\,1\ -\ \sinh(2\beta J)\ )^2\ +\ \frac{1}{2}\ \sinh(2\beta J)\ (\,k\,a\,)^2\ \right]
$$

$$
\approx \ln 2 + \frac{1}{(\,4\,\pi\,)} \int_0^{\frac{\pi^2}{2}} dx\ \ln\left[\ (\,1\ -\ \sinh(2\beta J)\ )^2\ +\ \sinh(2\beta J)\ x\ \right] \qquad (1262)
$$

This yields the expression for the non-analytic part of the Free-Energy

$$
\frac{\ln Z}{N} \sim\ -\ \frac{(\,1\ -\ \sinh(2\beta J)\ )^2}{4\,\pi\ \sinh(2\beta J)} \left[\ \ln\left(\,1\ -\ \sinh(2\beta J)\,\right)^2\ -\ 1\ \right] \qquad (1263)
$$

Hence, the specific heat is found to diverge logarithmically at the transition temperature

$$
\frac{C}{N} \sim\ -\ k_B\ \frac{8}{\pi}\ (\,\beta_c\ J\,)^2\ \ln|\ T\ -\ T_c\ | \qquad (1264)
$$

which is symmetrical around the transition temperature. Onsager stated without proof that, for temperatures below T_c, the zero-field magnetization defined by the average value

$$
M\ =\ \sum_{i=1}^{N}\ <\ S_i^z\ > \qquad (1265)
$$

varies as

$$
M\ =\ N\ \left(\,1\ -\ \sinh^{-4}(2\beta J)\,\right)^{\frac{1}{8}} \qquad (1266)
$$

The proof of this last result was eventually published by C.N. Yang [45]. Onsager's success is of great historical importance and is without parallel, since to date, no exact solutions have been found for models of similar physical importance, such as the three-dimensional Ising model, or such as the two or three-dimensional versions of the Heisenberg model. Onsager's results provided the only rigorous treatment of a phase transition, until three decades later when the renormalization group technique was finally formulated.

10.1 Phase Transitions and Singularities

The non-analytic behavior of the Free-Energy that is seen in the vicinity of phase transition is, at first sight, quite disturbing since this implies that the partition function is also anomalous. This is unexpected since the partition function is defined as

$$
Z(\beta)\ =\ \text{Trace}\ \exp\left[\ -\ \beta\ \hat{H}\ \right] \qquad (1267)
$$

---

[45]C.N. Yang, Phys. Rev. **85**, 808 (1952).

which is understood to be the sum of a finite number of positive terms and, therefore, is an analytic function of β (except possibly when $\beta \to \infty$). Furthermore, the sum of these terms exponentiates to yield a Free-Energy $F_N(\beta)$ which has an extensive part (proportional to N) and other parts which become negligible when $N \to \infty$. This would lead us to expect that the Free-Energy is also an analytic function. However, the sequence of series of analytic functions $F_N(\beta)$ need not be analytic in the limit $N \to \infty$.

Just to make this explicit, consider an Ising Spin Hamiltonian, with long-ranged pair-wise interactions

$$\hat{H} = J \sum_{i>j} S_i^z S_j^z \tag{1268}$$

where $S_i^z = \pm\frac{1}{2}$. Any spin interacts with every other spin with the same interaction strength. The Hamiltonian can be re-written as

$$\hat{H} = \frac{J}{2} \left[\left(\sum_{i=1}^{N} S_i^z \right)^2 - \sum_{i=1}^{N} (S_i^z)^2 \right]$$

$$= \frac{J}{2} \left[\left(\sum_{i=1}^{N} S_i^z \right)^2 - \frac{N}{4} \right] \tag{1269}$$

Strictly speaking, it is necessary to formally define $J = J'/N$, if we were to demand that the Free-Energy be extensive. We shall ignore this simply to avoid carrying around extra factors of N. If J is negative, the ground state corresponds to a ferromagnetic state for which

$$\sum_{i=1}^{N} S_i^z = \pm \frac{N}{2} \qquad \text{for } J < 0 \tag{1270}$$

The ferromagnetic state is two-fold degenerate. On the other hand if $J > 0$ the ground state (when N is even) corresponds to

$$\sum_{i=1}^{N} S_i^z = 0 \qquad \text{for } J > 0 \tag{1271}$$

which is highly-degenerate. It is easy to show that the degeneracy is given by $C_{\frac{N}{2}}^{N}$. The partition function $Z_N(\beta)$ can be expanded in powers of a parameter z defined as

$$z = \exp\left[-\frac{\beta J}{8} \right] \tag{1272}$$

For positive J, z is reduced from 1 to 0 as T is reduced from ∞ to 0. For negative J, z increases from 1 to ∞ as T is reduced from ∞ to 0. The partition

function is found to be given by

$$Z_N(\beta) \;=\; \exp\left[\,-\, N\, \frac{\beta\, J}{8} \,\right] \sum_{m=0}^{N} C_m^N \, z^{(N-2m)^2}$$

$$=\; z^N \sum_{m=0}^{N} C_m^N \, z^{(N-2m)^2} \tag{1273}$$

where C_m^N are the binomial coefficients. This expression contains a factor which is a very high-order polynomial in z. The polynomial has no roots on the positive real axis except, perhaps, at the point $z = 0$. However, it does have pairs of complex conjugate roots in the complex z plane. The roots may be multiple roots. For our model, it is seen that the pairs of roots are located on circles enclosing the origin $z = 0$. As z approaches a point which is a root, the partition
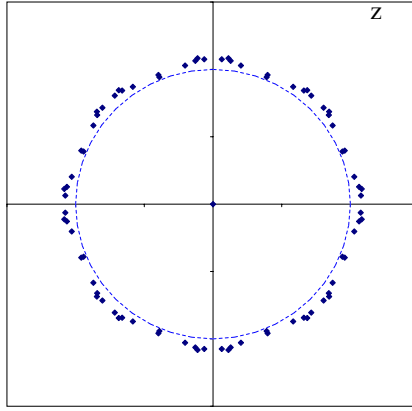
Figure 59: The distribution of zeroes of the partition function $Z_9(z)$ for the Ising Model with long-ranged interactions, in the complex z-plane. The dashed blue circle has a radius of unity.

function approaches zero and the Free-Energy $F_N(z)$ diverges logarithmically.

In general, the partition function is expected to have the form

$$Z_N(z) \;=\; \exp\left[\,-\, N\, A(z) \,\right] \prod_{\alpha} (\, z \,-\, z_\alpha \,)(\, z \,-\, z_\alpha^* \,) \tag{1274}$$

where z_α and z_α^* are the pairs of complex conjugate roots in the complex z plane, and $A(z)$ is a simple function. The Free-Energy $F_N(z)$ is given by

$$F_N(z) \;=\; k_B\, T\, N\, A(z) \,-\, k_B\, T \sum_{\alpha} \ln (\, z \,-\, z_\alpha \,)(\, z \,-\, z_\alpha^* \,) \tag{1275}$$

239

which has singularities in the complex z plane. Lee and Yang[46] proved that the limit

$$\lim_{N\to\infty} \frac{1}{N} F_N(z) = -k_B T \lim_{N\to\infty} \frac{1}{N} \ln Z_N(z) \qquad (1276)$$

exits for all real positive z and is a continuous monotonically increasing function of z. Also, for any region Ω which does not contain any roots of $Z_N(z)$ then $\lim_{N\to\infty} \frac{1}{N} F_N(z)$ is analytic in this region. If these conditions are satisfied for all physical values of z, the system does not exhibit a phase transition.

As the limit $N \to \infty$ is approached, the zeroes of $Z_N(z)$ may approach the real axis and pinch it off at a real value of z, z_c. The conditions of the Lee-Yang theorem do not apply in the immediate vicinity of this point. If the zeroes approach a point z_c on the real axis continuously as N is increased, then the point z_c may be located on a branch cut of $F(z)$ which would yield non-analytic behavior at z_c or equivalently at β_c. In such a case, z_c would defines a critical temperature T_c at which the Free-Energy is singular.

10.2 The Mean-Field Approximation for an Ising Magnet

Consider an Ising Hamiltonian in the presence of an applied magnetic field H^z

$$\hat{H} = -\sum_{i,j} J_{i,j} S_i^z S_j^z - \sum_i \frac{g\,\mu_B}{\hbar} S_i^z H^z \qquad (1277)$$

The operator S^z has $(2S + \hbar)/\hbar$ possible eigenvalues which are $-S$, $-S + \hbar$, ..., $S - \hbar$, S. The interaction $J_{i,j}$ couples the z-components of nearest neighbor spins. We shall assume that the interaction J is short-ranged and takes on the same positive value between each pair of nearest neighboring spins, so that the lowest energy configuration is ferromagnetic in which all the spins are aligned parallel to each other. Although the Hamiltonian has an extremely simple form, the only known exact expressions for the Free-Energy have been found for the special cases where the spins are arranged on one or two-dimensional lattices[47]. Therefore, we shall have to describe this system approximately by using the mean-field approximation, first introduced by Weiss.

We shall define the average magnetization per spin as m and express the Hamiltonian as

$$\hat{H} = -\sum_{i,j} J_{i,j} \left(m + (S_i^z - m) \right) \left(m + (S_j^z - m) \right) - \sum_i \frac{g\,\mu_B}{\hbar} \left(m + (S_i^z - m) \right) H^z$$

$$(1278)$$

---

[46]C.N. Yang and T.D. Lee, "Statistical Theory of Equations of State and Phase Transitions: 1. Theory of Condensation", Phys. Rev. 87, 404-409, (1952).
T.D. Lee and C.N. Yang, 'Statistical Theory of Equations of State and Phase Transitions: 2. Lattice Gas and Ising Model", Phys. Rev. 87, 410-419, (1952).
[47]L. Onsager, "Crystal Statistics I: A two-dimensional model with an order-disorder transition", Phys. Rev. 65, 117-149 (1944).
L. Onsager, unpublished: Nuovo Cimento 6, Suppl. p.261 (1949).

and expand in power of the fluctuations of the spins ($S_i^z - m$) from their average value. To first-order in the fluctuations, one has

$$
\begin{aligned}
\hat{H}_{MF} &= \sum_{i,j} J_{i,j} \, m^2 \\
&- \sum_{i,j} J_{i,j} \, m \, (\, S_i^z + S_j^z \,) - \sum_i \frac{g \, \mu_B}{\hbar} \, S_i^z \, H^z
\end{aligned} \tag{1279}
$$

where we have neglected the terms of second-order in the fluctuations. One should note that the above Hamiltonian resembles that expected for non-interacting spins in an effective magnetic field H_{eff}, given by

$$
H_{eff}^z = H + \frac{\hbar}{g \, \mu_B} \sum_j J_{i,j} \, m \tag{1280}
$$

The mean-field partition function Z can be calculated as

$$
Z_{MF} = \text{Trace} \, \exp \left[- \beta \, H_{MF} \right] \tag{1281}
$$

where the Trace runs over all the possible spin configurations. Thus, the Trace corresponds to the products of sums over the $(2S/\hbar + 1)$ possible configuration of each spin. Since the spins are no longer coupled, the mean-field Hamiltonian factorizes. Hence, the partition function has the form

$$
Z_{MF} = \exp \left[- \beta \sum_{i,j} J_{i,j} \, m^2 \right] \prod_{i=1}^{N} \left\{ \sum_{S_i^z = -S}^{+S} \exp \left[+ \beta \frac{g \, \mu_B}{\hbar} \, S_i^z \, H_{eff}^z \right] \right\} \tag{1282}
$$

The trace can be performed yielding the result

$$
Z_{MF} = \exp \left[- \beta \sum_{i,j} J_{i,j} \, m^2 \right] \left[\frac{\sinh \frac{(2S+\hbar)}{\hbar} \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right)}{\sinh \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right)} \right]^N \tag{1283}
$$

Hence, in the mean-field approximation, the Free-Energy is given by

$$
\begin{aligned}
F_{MF} &= \sum_{i,j} J_{i,j} \, m^2 - N \, k_B \, T \, \ln \left[\sinh \frac{(2S + \hbar)}{\hbar} \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right) \right] \\
&+ N \, k_B \, T \, \ln \left[\sinh \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right) \right]
\end{aligned} \tag{1284}
$$

The magnetization is found from the thermodynamic relation

$$
M^z = - \left(\frac{\partial F}{\partial H^z} \right) \tag{1285}
$$

which yields

$$M^z = N \, g \, \mu_B \left[\frac{2S + \hbar}{2 \, \hbar} \, \coth \, \frac{(2S + \hbar)}{\hbar} \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right) - \frac{1}{2} \, \coth \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right) \right]$$
(1286)

On recognizing that

$$M^z = \frac{g \, \mu_B}{\hbar} \, \sum_i \overline{S^z}_i$$
(1287)

one finds that the average value of S^z is independent of the site and is given by

$$\overline{S_0^z} = \left[\frac{2S + \hbar}{2} \, \coth \, \frac{(2S + \hbar)}{\hbar} \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right) - \frac{\hbar}{2} \, \coth \left(\frac{\beta \, g \, \mu_B \, H_{eff}}{2} \right) \right]$$
(1288)

or, equivalently, on using the definition of m as the average value of the z-component of the spin

$$m = \left[\frac{2S + \hbar}{2} \, \coth \, \frac{(2S + \hbar)}{\hbar} \left(\frac{\beta \, g \, \mu_B \, H^z \, + \, \beta \, \sum_j J_{j,0} \, \hbar \, m}{2} \right) \right.$$
$$\left. - \frac{\hbar}{2} \, \coth \left(\frac{\beta \, g \, \mu_B \, H^z \, + \, \beta \, \sum_j J_{j,0} \, \hbar \, m}{2} \right) \right]$$
(1289)

This non-linear equation determines the value of m. This equation is known as the self-consistency equation, since the equation for m has to be solved self-consistently as m also enters non-linearly in the right-hand side. The equation can be solved graphically.

For $H^z = 0$ the two spin directions are equivalent, and the equation simplifies to

$$m = \left[\frac{2S + \hbar}{2} \, \coth \, \frac{(2S + \hbar)}{\hbar} \left(\frac{\beta \, \sum_j J_{j,0} \, \hbar \, m}{2} \right) \right.$$
$$\left. - \frac{\hbar}{2} \, \coth \left(\frac{\beta \, \sum_j J_{j,0} \, \hbar \, m}{2} \right) \right]$$
(1290)

Both the left and right hand sides are odd functions of m. This symmetry is a consequence of the symmetry under spin inversion $S_i^z \rightarrow - S_i^z$ of the Hamiltonian, when $H^z = 0$. The graphical solution is illustrated in the figure. At high temperatures, the equation has only one solution $m = 0$, whereas at low temperatures, there are three solutions, one solution corresponds to $m = 0$ and the other two solutions corresponds to $m = \pm m_0(T)$ located symmetrically about $m = 0$. The value of $m_0(T)$ increases continuously from 0 and saturates at S as T is decreased towards zero. The critical temperature T_c at which the pair of non-zero solutions first appears can be found by expanding the right-hand side w.r.t. m, since it is expected that $m \approx 0$ just below T_c. This leads to the equation

$$m \approx \beta \, \frac{2}{3} \, \sum_j J_{j,0} \, S \, (\, S \, + \hbar \,) \, m \, + \, O(m^3)$$
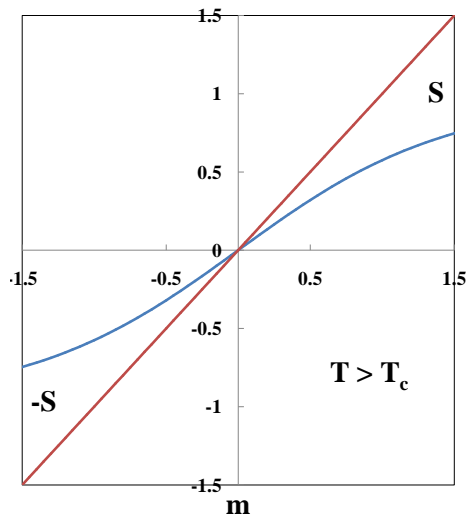(1291)

242

Figure 60: The graphical solution of the mean-field self-consistency equation, for temperatures above T_c.
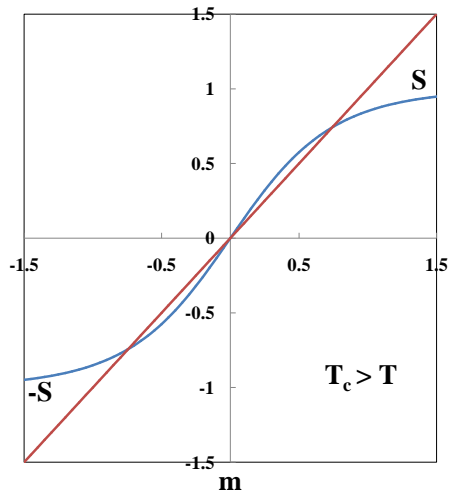

Figure 61: The graphical solution of the mean-field self-consistency equation, for temperatures below T_c.

On assuming that the cubic terms are negligibly small, this has the solution

$m = 0$ unless the temperature is equal to T_c which satisfies

$$k_B \ T_c \ = \ \frac{2}{3} \ \sum_j \ J_{j,0} \ S \ (\ S \ + \hbar \ ) \qquad (1292)$$

At the critical temperature T_c, a non-zero (but still infinitesimal) value of $m_0(T)$ is first allowed. At this temperature, the graphical solution shows that the line $m \ = \ m$ is tangent to the magnetization curve at the origin. For temperatures just below T_c, the cubic terms in m in the self-consistency equation determine the magnitude of the solution m_0 to be a function of $(\ T_c \ - \ T \ )$.

Below T_c, the system must be described by one of the three possible values for m. In equilibrium, the Free-Energy should be minimized. The condition that F_{MF} is an extremum w.r.t. m is equivalent to the above self-consistency condition for m. It is seen that for temperatures for which the non-zero solutions exist, the Free-Energy is minimized for the non-zero solutions for m, and that these minima are degenerate. The system will spontaneously break the
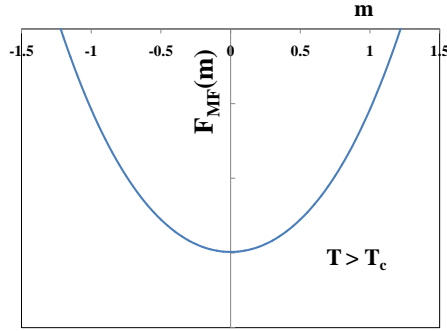
Figure 62: The m-dependence of the mean-field Free-Energy, for temperatures above T_c.

spin inversion symmetry, by settling into one of the two equivalent states. When $H^z \ = \ 0$, the mean-field approximation indicates that there is a second-order phase transition at T_c. The transition occurs at the temperature at which the location of the absolute minimum value of $F_{MF}(m)$ changes from $m \ = \ 0$ for temperatures above T_c to the two equivalent locations of $m \ = \ \pm \ m_0(T)$ for temperatures below T_c.

The mean-field approximation shows the existence of another type of phase transition, which occurs when there is a small symmetry breaking applied magnetic field H^z which is to be varied. In this case, the applied field stabilizes one of the minima (say m_0) w.r.t. to the minima at $- \ m_0$. However, on changing the applied field from a slightly positive to a slightly negative value, the equilibrium value of m jumps discontinuously from $+ \ m_0$ to $- \ m_0$. This discontinuous jump
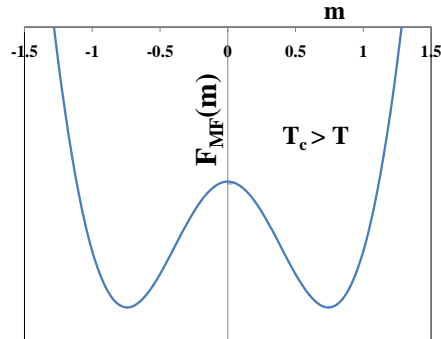
Figure 63: The m-dependence of the mean-field Free-Energy, for temperatures below T_c.

between two pre-existing minima of $F(m)$ is characteristic of a first-order phase transition. The defining characteristic of a first-order transition is the presence
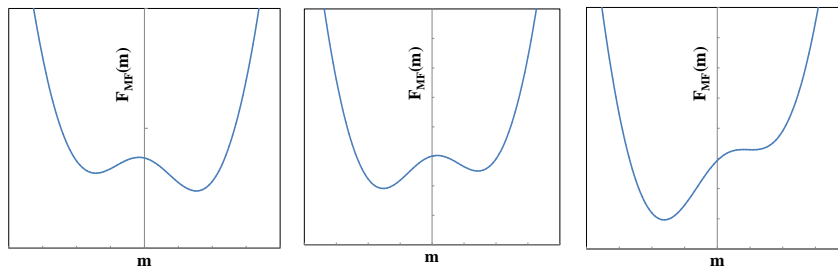
Figure 64: The m-dependence of the mean-field Free-Energy, for temperatures below T_c and various values of the applied field.

of discontinuous first-derivatives of the Free-Energy[48]. Since the entropy is a first-derivative of the Free-Energy with respect to temperature, first-order transitions may involve a latent heat. Since energy is not transferred instantaneously between a system and its environment, first-order transitions are frequently associated with "mixed-phase regimes" in which some parts of the system have completed the transition and others have not. Another characteristic is the presence of hysteresis. For systems where the transformation proceeds slowly, a phase may exist at fields where its Free-Energy is not a global minimum but only a local minimum. Hence, on decreasing the field, the transformation may

---

[48] P. Ehrenfest (1880-1933) proposed a classification of phase transitions based on the discontinuities of the derivatives of the Free-Energy. A transition was defined to be n-th order transition if the lowest-order derivative of the Free-Energy which was discontinuous across the transition was the n-th order derivative. This classification fell into disfavor when it was realized that a derivative could diverge without it being discontinuous.

first occur at a negative value of the field. Furthermore, if the field is subsequently increased, the reverse transformation may occur at a positive value of the field. The point at which the transformation occurs is determined by the rate at which the field is changed and the time-scale required for the system to nucleate the new phase.
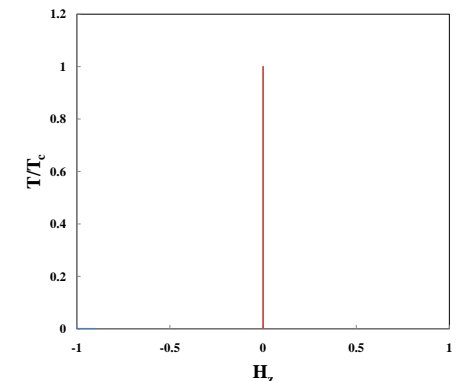
Figure 65: The $H - T$ phase diagram for the mean-field description of a ferromagnet.

The phase diagram in the $H - T$ plane, shows a line of first-order transitions at low-temperatures which ends at the critical point $(H = 0, T = T_c)$. Due to the symmetry of the magnetic system, the line of first-order transitions is vertical. On keeping $H = 0$, the system exhibits a second-order phase transition at the critical point.

Phase transitions are found in many different types of system. Despite the differences between the microscopic descriptions, phase transitions can usually be described in the same manner. For example, the liquid-gas phase transition the role of the magnetization is replaced by the density and the magnetic field is replaced by the pressure. The line of first-order transitions is not vertical but has a finite slope so it can be crossed by changing temperature.

These transitions can be described in similar manners. In this description, the microscopic variables are replaced by coarse grained variables that represent the collective coordinates of a large number of the microscopic degrees of freedom. The resulting description only retains the essential characteristics of the underlying microscopic systems, such as the symmetry, the dimensionality and the character of the ordering.
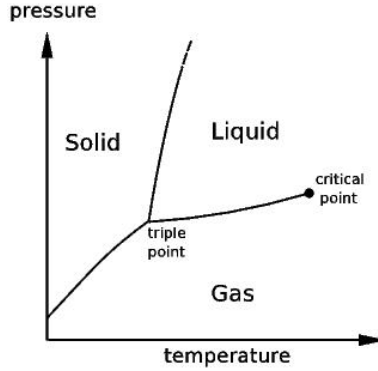
Figure 66: The $P - T$ phase diagram for the liquid-gas transition. The liquid and gas are separated by a line of first-order transitions which ends at a critical point.

10.3   The Landau-Ginzberg Free-Energy Functional

The Landau-Ginzberg formulation of critical phenomena is based on a coarse grained version of statistical mechanics, in which the microscopic degrees of freedom, pertaining to some small region of space, have been replaced by a small set of collective variables that describe the state of the volume element. This leads to a formulation of the statistical mechanics of the system in terms of the collective variables for each small volume element. It is common to introduce an order parameter ϕ defined for the microscopic d-dimensional volume element $d^d\underline{r}$ surrounding the point $\underline{r}$. The order parameter characterizes the change in the system that occurs at the phase transition. For a system which changes symmetry, the order parameter is an extensive quantity (for example it could be a scalar quantity or it could be a vector quantity with a number of components that we will denote by n, etc.) which is non-zero in the state with lower-symmetry and is zero in the state with higher-symmetry. For simplicity, we shall only consider the case when one type of symmetry is broken, in which case, only one order parameter is required. The volume of the system is to be partitioned into a set of identical infinitesimal volumes (or cells) $d^d\underline{r}$ that surrounding a set of points that we shall label by the variable $\underline{r}$. The order parameter $\phi(\underline{r})$ can be expressed as the sum of terms of microscopic quantities ϕ_i that are expressed in terms of the microscopic degrees of freedom contained in the volume element $d^d\underline{r}$ surrounding the point $\underline{r}$

$$\phi(\underline{r}) \; = \; \sum_{i \, \in \, \underline{r}} \phi_i \tag{1293}$$

The partition function Z is expressed as the trace over the Hamiltonian

$$Z \; = \; \text{Trace} \; \exp\left[\; - \; \beta \; \hat{H} \; \right] \tag{1294}$$

247

The Landau-Ginzberg Free-Energy Functional $F[\phi(\underline{r})]$ is a number that depends on the function $\phi(\underline{r})$ and is expressed as in terms of a trace

$$\exp\left[-\beta F[\phi(\underline{r})] \right] = \text{Trace} \prod_{\underline{r}} \left\{ \delta\left(\phi(\underline{r}) - \sum_{i\,\in\,\underline{r}} \phi_i \right) \right\} \exp\left[-\beta \hat{H} \right]$$

(1295)

where the product contains a delta function which constrains the microscopic variables in the volume elements around each point $\underline{r}$ to be consistent with the value of $\phi(\underline{r})$ at that point. Hence, the partition function can be expressed as an integral over the possible values of $\phi(\underline{r})$ for each cell labeled by $\underline{r}$.

$$Z = \prod_{\underline{r}} \left\{ \int d\phi(\underline{r}) \right\} \exp\left[-\beta F[\phi] \right]$$

(1296)

This is recognized as a functional integral, and it should be noted that the set of possible "functions" defined by the values of $\phi(\underline{r})$ at each point of space $\underline{r}$ include many wild functions that change discontinuously from point to point and also includes functions vary smoothly over space. The functional integral over the set of all possible functions $\phi(\underline{r})$ is weighted exponentially by the Landau-Ginzberg Free-Energy Functional. The path integral is conventionally denoted by

$$Z = \int \mathcal{D}\phi(\underline{r}) \exp\left[-\beta F[\phi] \right]$$

(1297)

The Landau-Ginzberg Free-Energy plays the role of a Hamiltonian, which generally depends on T, that describes the physical probabilities in terms of the collective variables $\phi(\underline{r})$ described on the length scale dictated by the choice of the size of the volume elements $d^3\underline{r}$. It contains all the physics that is encoded in the Helmholtz-Free-Energy F. Like the Hamiltonian, the Landau-Ginzberg Free-Energy Functional is a scalar. In principle The Landau-Ginzberg Free-Energy Functional should be calculated from knowledge of the model and its symmetries. In practice, one can understand properties of phase transitions in a quite universal way close to a second-order phase transition or a weakly first-order transition, where the order parameter is quite small. In such cases, one can expand the Landau-Ginzberg Free-Energy Functional in powers of the parameter, keeping only terms of low-order. The constraints imposed by stability and the symmetry on the finite number of terms retained, provides severe restrictions on the form of the Landau-Ginzberg Free-Energy Functional that describes the phase transition of a system. This severe restriction causes all the different phase transitions of physical systems to fall into a small number of universality classes, which are determined only by the dimensionality of the system d and the dimensionality of the order parameter n. Systems which fall into the same universality class have the same types of non-analytic temperature variations.

For example, a system residing in a d-dimensional Euclidean space which is characterized by an n-dimensional vector order parameter $\underline{\phi} \equiv (\phi_1, \phi_2 \ldots \phi_n)$

and has a Hamiltonian which is symmetric under rotations of the order parameter, can be described by the expanded Landau-Ginzberg Free-Energy Functional

$$
F[\phi] = \int d^d \underline{r} \left[F_0 + F_2 \, \underline{\phi}(\underline{r}) \cdot \underline{\phi}(\underline{r}) + F_4 \, (\, \underline{\phi}(\underline{r}) \cdot \underline{\phi}(\underline{r}) \,)^2 - \underline{\Pi}(\underline{r}) \cdot \underline{\phi}(\underline{r}) \right.
$$
$$
\left. + \sum_{i=1}^{n} c \, (\, \underline{\nabla} \phi_i \cdot \underline{\nabla} \phi_i \,) \right] \tag{1298}
$$

In this expression F_0, F_2, F_4 and c are constants, that might depend on temperature and may also depend on the microscopic length scales of the system. If the above expansion is to describe stable systems that have small values of the order parameter, it is necessary to assume that $F_4 > 0$. The Free-Energy Functional has been expressed in terms of quantities that are invariant under the symmetries of space and the order parameter. The invariant quantities include the identity and the scalar product

$$
\underline{\phi}(\underline{r}) \cdot \underline{\phi}(\underline{r}) = \sum_{i=1}^{n} \phi_i(\underline{r}) \, \phi_i(\underline{r}) \tag{1299}
$$

The first three terms represent the Free-Energy density for the cells, in the absence of an external field. Since the material is assumed to be homogeneous, the coefficients F_0, F_2 and F_4 are independent of $\underline{r}$. The fourth term represents the effect of a spatially varying applied external field $\underline{\Pi}(\underline{r})$ that is conjugate to $\underline{\phi}(\underline{r})$. The application of the field breaks the symmetry under rotations of the order parameter. The final term represents the interaction between neighboring cells, which tends to suppress rapid spatial variations of the order parameter and, hence, gives large weights to the functions $\underline{\phi}(\underline{r})$ which are smoothly varying. The gradient term involves two types of scalar products, one type is associated with the d-dimensional scalar product of the gradients and the other is associated with an n-dimensional scalar product of the vector order parameter. The appearance of the gradient is due to the restriction to large length scales in the Landau-Ginzberg formulation. In this case, expressions such as

$$
\kappa \sum_{\underline{\delta}} \left(\underline{\phi}(\underline{r} + \underline{\delta}) - \underline{\phi}(\underline{r}) \right)^2 \tag{1300}
$$

which tend to keep the value of ϕ in the cell $\underline{r}$ close to the values of ϕ in the neighboring cells at $\underline{r} + \underline{\delta}$ can be expanded, leading to

$$
\kappa \sum_{\underline{\delta}} \sum_{i=1}^{n} \left(\phi_i(\underline{r} + \underline{\delta}) - \phi_i(\underline{r}) \right)^2 \approx \kappa \sum_{i=1}^{n} \sum_{\underline{\delta}} (\, \underline{\delta} \cdot \underline{\nabla} \, \phi_i \,)^2
$$
$$
= c \sum_{i=1}^{n} (\, \underline{\nabla} \, \phi_i \cdot \underline{\nabla} \, \phi_i \,) \tag{1301}
$$

where we have assumed that the higher-order terms in the small length scale δ are negligibly small and that the neighboring cells are distributed isotropically

in space. This assumption of isotropic space and slow variations leads to the Landau-Ginzberg Functional having a form similar to the Lagrangians of continuum Field Theories. Apart from the coefficients F_0, F_2, F_4 and c, the form of the Lagrangian only depends on the values of n and d. However, for systems which undergo more than one type of phase transition, it may be necessary to introduce more than one order parameter, in which case the Landau-Ginzberg Free-Energy functional can have more complicated forms.

Linear Response Theory

For simplicity, we shall consider the case where the order parameter is a scalar. In general, if a system is subject to a uniform applied field with an additional small (perhaps non-uniform) component $\delta\Pi(\underline{r})$, so that

$$\Pi(\underline{r}) \ = \ \Pi_0 \ + \ \delta\Pi(\underline{r}) \tag{1302}$$

then one expects that the additional small component of the field will induce a small additional (non-uniform) component into the expectation value of the local order-parameter $< \ \phi(\underline{r}) \ >$

$$< \ \phi(\underline{r}) \ > \ = \ \phi_0 \ + \ \delta\phi(\underline{r}) \tag{1303}$$

The average of the order-parameter is defined by

$$< \ \phi(\underline{r}) \ > \ = \ \frac{1}{Z} \ \int \mathcal{D}\phi \ \phi(\underline{r}) \ \exp \left[\ - \ \beta \ F[\phi] \ \right] \tag{1304}$$

where the Trace has been replaced by a path integral, and the Hamiltonian H has been replaced by the Landau-Ginzberg Free-Energy Functional $F[\phi]$. The Landau-Ginzberg Free-Energy Functional includes both the uniform applied field and the small (non-uniform) component. On expanding the exponent and denominator in powers of $\delta\Pi(\underline{r})$, one has

$$< \phi(\underline{r}) \ > \ = \ \frac{\int \mathcal{D}\phi \ \phi(\underline{r}) \ \left(\ 1 \ + \ \beta \ \int d^d \underline{r}' \ \phi(\underline{r}') \ \delta\Pi(\underline{r}') \ \right) \ \exp \left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \ \right]}{\int \mathcal{D}\phi \ \left(\ 1 \ + \ \beta \ \int d^d \underline{r}' \ \phi(\underline{r}') \ \delta\Pi(\underline{r}') \ \right) \ \exp \left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}$$
$$\tag{1305}$$

The divisor is expanded to lowest non-trivial order as

$$\frac{1}{\int \mathcal{D}\phi \ \left(\ 1 \ + \ \beta \ \int d^d \underline{r}' \ \phi(\underline{r}') \ \delta\Pi(\underline{r}') \ \right) \ \exp \left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}$$
$$= \ \frac{1}{\int \mathcal{D}\phi \ \exp \left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}$$

$$
- \ \frac{\int \mathcal{D}\phi \ \left(\beta \ \int d^d\underline{r}' \ \phi(\underline{r}') \ \delta\Pi(\underline{r}') \right) \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}{\left[\int \mathcal{D}\phi \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right] \right]^2}
\qquad (1306)
$$

Since the uniform part of the order parameter satisfies

$$
\phi_0 \ = \ \frac{\int \mathcal{D}\phi \ \phi(\underline{r}) \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}{\int \mathcal{D}\phi \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}
\qquad (1307)
$$

then the small additional (non-uniform) component of the order parameter is given by the expression

$$
\delta\phi(\underline{r}) \ = \ \beta \ \int d^d\underline{r}' \ \delta\Pi(\underline{r}') \ \Bigg(\frac{\int \mathcal{D}\phi \ \phi(\underline{r}) \ \phi(\underline{r}') \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}{\int \mathcal{D}\phi \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}
$$

$$
- \ \frac{\int \mathcal{D}\phi \ \phi(\underline{r}) \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}{\int \mathcal{D}\phi \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]} \ \frac{\int \mathcal{D}\phi \ \phi(\underline{r}') \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]}{\int \mathcal{D}\phi \ \exp\left[\ - \ \beta \ F[\phi] \Big|_{\delta\Pi=0} \right]} \Bigg)
$$

$$
(1308)
$$

where the integration over the additional part of the applied field $\delta\Pi(\underline{r}')$ has been taken out of the averages. The above equation can be written in a more compact form as

$$
\delta\phi(\underline{r}) \ = \ \beta \ \int d^d\underline{r}' \ \Big[\ < \ \phi(\underline{r}) \ \phi(\underline{r}') > \ - \ < \ \phi(\underline{r}) \ > \ < \ \phi(\underline{r}') \ > \ \Big] \ \delta\Pi(\underline{r}')
$$

$$
(1309)
$$

in which the averages are calculated with $\delta\Pi = 0$. On defining the two-point correlation function $S(\underline{r}, \underline{r}')$ as

$$
S(\underline{r}, \underline{r}') \ = \ < \ \phi(\underline{r}) \ \phi(\underline{r}') > \ - \ < \ \phi(\underline{r}) \ > \ < \ \phi(\underline{r}') \ >
\qquad (1310)
$$

then the small induced component of the order-parameter is given by

$$
\delta\phi(\underline{r}) \ = \ \beta \ \int d^d\underline{r}' \ S(\underline{r}, \underline{r}') \ \delta\Pi(\underline{r}')
\qquad (1311)
$$

which is a linear response relation which connects the small change in the order-parameter $\delta\phi$ at position $\underline{r}$ to the change in the applied field $\delta\Pi$ at position $\underline{r}'$. For translational invariant systems, the correlation function does not depend separately on $\underline{r}$ and $\underline{r}'$, but only on the relative separation $\underline{r} - \underline{r}'$. For materials

which are translational invariant, one can displace the origin through a distance $\underline{r}'$ leading to the expression

$$S(\underline{r} - \underline{r}') \; = \; < \; \phi(\underline{r} - \underline{r}') \;\; \phi(0) > \; - \; < \; \phi(\underline{r} - \underline{r}') \; > \; < \; \phi(0) \; > \qquad (1312)$$

which only depends on the difference $\underline{r} - \underline{r}'$. In the limit where $\delta\Pi$ becomes uniform one finds that, due to translational invariance, $\delta\phi$ becomes uniform and is given by

$$\delta\phi \; = \; \beta \; \int d^d\underline{r}' \; S(-\underline{r}') \; \delta\Pi \qquad (1313)$$

Due to the isotropy of space, the induced value of $\delta\phi$ can be expressed as

$$\delta\phi \; = \; \beta \; \int d^d\underline{r}' \; S(\underline{r}') \; \delta\Pi \qquad (1314)$$

Hence, the uniform differential susceptibility, χ, defined by

$$\chi \; = \; \left(\frac{\delta\phi}{\delta\Pi} \right) \qquad (1315)$$

is found to be proportional to the spatial integral of the correlation function

$$\chi \; = \; \beta \; \int d^d\underline{r} \; S(\underline{r}) \qquad (1316)$$

where the correlation function $S(r)$ is calculated with $\delta\Pi = 0$.

10.4   Critical Phenomena

The simplest phase diagram exhibiting a phase transition, is a line of first-order transitions which ends in a critical point. If the line of first-order transitions is traversed along its length towards the critical point, the system is approaching a second-order phase transition. The order parameter (which is given by a first-order derivative of the Free-Energy w.r.t. to an applied field Π) is discontinuous as the system is taken on a path which crosses the line of first-order transitions. The difference in the order parameter, from the opposite side of the line, characterizes the difference between the two phases. The magnitude of the discontinuity diminishes for paths that cross the line of first-order transition closer to the critical point. This signifies that there is no discernable difference between the phases at the critical point. Above the critical point, only a unique phase exists and the order parameter is zero.

Phenomena that happen at a critical point are known as critical phenomena. Close to the critical point, the temperature variation in physical quantities may show non-analytic temperature variations. For example, one may consider a uniform order parameter ϕ defined by the derivative of the Free-Energy w.r.t a uniform applied field Π

$$\phi \; = \; - \left(\frac{\partial F}{\partial \Pi} \right) \qquad (1317)$$

a zero-field susceptibility χ defined as

$$\chi = \left(\frac{\partial \phi}{\partial \Pi} \right) \tag{1318}$$

or the heat capacity C

$$
\begin{aligned}
C &= T \left(\frac{\partial S}{\partial T} \right) \\
&= -T \left(\frac{\partial^2 F}{\partial T^2} \right) \tag{1319}
\end{aligned}
$$

In the vicinity of the critical point, these quantities exhibit a non-analytic temperature. A dimensionless parameter t is introduced as

$$t = \frac{T - T_c}{T_c} \tag{1320}$$

as a measure of the "distance" to the critical point. The temperature variation of any quantity is decomposed into a regular part and the non-analytic part. The temperature dependence of the non-analytic part can be expressed in terms of t. It is found that the leading non-analytic part follows temperature variations given by

$$
\begin{aligned}
\phi &\sim |t|^\beta \quad \text{for} \quad T < T_c \quad \text{and} \quad \Pi = 0 \\
\chi &\sim |t|^{-\gamma} \quad \text{for} \quad \Pi = 0 \\
C &\sim |t|^{-\alpha} \quad \text{for} \quad \Pi = 0 \\
\phi &\sim \Pi^{\frac{1}{\delta}} \quad \text{for} \quad T = T_c \tag{1321}
\end{aligned}
$$

where the exponents α, β, γ and δ are known as the critical exponents. Generally, the value of a critical exponent (say λ is the exponent of a quantity A) is determined by taking the limit

$$\lambda = \lim_{t \to 0} \left(\frac{\ln A}{\ln |t|} \right) \tag{1322}$$

The critical exponent describes the leading order temperature variation. However, one expects correction terms, so that a quantity A may vary as

$$A = c\, |t|^\lambda \left(1 + D\, |t|^\mu + \ldots \right) \tag{1323}$$

where $\mu > 0$. For a second-order transition, there should be no latent heat on passing through the transition, thus

$$L = \int_{T_c - \epsilon}^{T_c + \epsilon} dT\, C(T) = 0 \tag{1324}$$

253

so, $1 > \alpha$. The value of α is actually significantly smaller than unity, and for some systems (for example the two-dimensional Ising Model) C varies logarithmically

$$C \sim \ln \left| \frac{T_c}{T - T_c} \right| \tag{1325}$$

when T is close to T_c. Since

$$\ln |t| = \lim_{\alpha \to 0} \frac{1}{\alpha} \left(|t|^{-\alpha} - 1 \right) \tag{1326}$$

the logarithmic variation corresponds to $\alpha = 0$. There are other critical exponents that are introduced to characterize the spatial correlations of the order parameter. Thus, for example, one can introduce a correlation function $S(\underline{r})$ as the an average of the product of the fluctuations of a local order parameter $\Delta\phi(\underline{r})$ defined via

$$\Delta\phi(\underline{r}) = \phi(\underline{r}) - < \phi(\underline{r}) > \tag{1327}$$

The correlation function $S(\underline{r} - \underline{r}')$ is introduced as

$$
\begin{aligned}
S(\underline{r} - \underline{r}') &= < \Delta\phi(\underline{r}) \, \Delta\phi(\underline{r}') > \\
&= < \phi(\underline{r}) \, \phi(\underline{r}') > - < \phi(\underline{r}) > < \phi(\underline{r}') >
\end{aligned} \tag{1328}
$$

The last term has the effect that the correlation function decays to zero at large distances for temperatures above and below T_c. Since we are assuming the system is invariant under translations, one expects that the average value is non-zero below T_c where it satisfies

$$< \phi(\underline{r}) > = < \phi(\underline{r}') > \tag{1329}$$

so that it is independent of the position. Also, due to translational invariance, one has

$$< \phi(\underline{r}) \, \phi(\underline{r}') > = < \phi(\underline{r} - \underline{r}') \, \phi(0) > \tag{1330}$$

so the correlation function only depends on $\underline{r} - \underline{r}$. Furthermore, due to isotropy in space

$$S(\underline{r} - \underline{r}') = S(|\underline{r} - \underline{r}'|) \tag{1331}$$

independent of the orientation of the vector $\underline{r}$. In the limit $|\underline{r} - \underline{r}'| \to \infty$, one expects that the value of ϕ at $\underline{r}$ will be unrelated to the value of ϕ at $\underline{r}'$. Therefore, one expects that in this limit

$$\lim_{|\underline{r}-\underline{r}'|\to\infty} < \phi(\underline{r}) \, \phi(\underline{r}') > \to < \phi(\underline{r}) > < \phi(\underline{r}') > \tag{1332}$$

Hence, one expects that correlation function decays to zero at large distances

$$\lim_{|\underline{r}-\underline{r}'|\to\infty} S(\underline{r} - \underline{r}') \to 0 \tag{1333}$$

One can define a correlation length ξ such that

$$S(\underline{r} - \underline{r}') \sim \left(\frac{1}{|\underline{r} - \underline{r}'|} \right)^{d-2+\eta} \exp \left[- \frac{|\underline{r} - \underline{r}'|}{\xi} \right] \tag{1334}$$

254

where d is the number of spatial dimensions, which holds for $\Pi = 0$ and $T \neq T_c$. The exponent η is the anomalous dimension. The correlation length ξ describes the length scale above which the correlations die out. The correlation length is found to diverge as the critical temperature is approached, so that

$$\xi \sim \mid t \mid^{-\nu} \tag{1335}$$

where the critical exponent is denoted by ν. One can define another exponent η which describes the spatial correlations at $T = T_c$. If one defines the Fourier components $\phi_{\underline{k}}$ of the local order parameter $\phi(\underline{r})$ via

$$\phi_{\underline{k}} \; = \; \frac{1}{\sqrt{V}} \; \int d^d\underline{r} \; \exp\left[\; - \; i \; \underline{k} \, . \, \underline{r} \; \right] \; \phi(\underline{r}) \tag{1336}$$

then one may define a momentum-space correlation function

$$< \phi_{\underline{k}} \, \phi_{-\underline{k}} \; > = \; \frac{1}{V} \; \int d^d\underline{r}_1 \; \int d^d\underline{r}_2 \; \exp\left[\; - \; i \; \underline{k} \, (\; \underline{r}_1 \; - \; \underline{r}_2 \;) \; \right] \; < \phi(\underline{r}_1) \, \phi(\underline{r}_2) \; > \tag{1337}$$

as the Fourier transformation of $S(\underline{r} - \underline{r}')$ for $\underline{k} \neq 0$. For $T = T_c$ one defines the exponent η via

$$< \phi_{\underline{k}} \, \phi_{-\underline{k}} \; > \sim \; \mid \underline{k} \mid^{-2+\eta} \quad \text{as} \quad \underline{k} \; \rightarrow \; 0 \tag{1338}$$

Experiments reveal that to within experimental accuracy the exponents, such as α, γ and ν, have the same value no matter in which direction the limit $t \rightarrow 0$ is taken. That is the exponents have the same values for $T > T_c$ as for $T < T_c$. However, the coefficients of the non-analytic temperature variation usually have different magnitudes for temperatures above and below T_c. That is, the singularities are of the same type no matter whether the transition is approached from above or below.

The same sets of values for the critical exponents are found for many different types of transitions. The same sets of values of the critical exponents are found for systems with transition temperatures that differ by many orders of magnitude, and for transitions that occur in crystalline materials, are independent of the type of crystal structure. Transitions which share the same set of values of the critical exponents form what is known as a universality class.

The values of the critical exponents are not independent of one another. Historically, it was first shown that thermodynamics requires that the six exponents must satisfy (more than) four inequalities. For example, using convexity arguments, Rushbrooke showed that the heat capacity measured at constant ϕ, C_ϕ, must be positive, hence

$$C_\Pi \; \geq \; T \; \frac{\left(\frac{\partial \phi}{\partial T} \right)_\Pi^2}{\chi_T} \tag{1339}$$

Table 2: The experimentally determined values of the critical exponents for a number of three-dimensional systems with order parameters of dimension n.

	Liquid-Gas Xe	Binary Fluid BMethanol-hexane	β-brass Cu-Zn	Norm-Super ^{4}He	Ferromagnet Fe	Antiferromagnet RbMnF$_3$
n	1	1	1	2	3	3
α	0.08 ± 0.02	$0.113 \pm .005$	$0.05 \pm .06$	$-0.014 \pm .016$	$-0.12 \pm .01$	$-0.139 \pm .007$
β	$0.344 \pm .003$	$0.322 \pm .002$	$0.305 \pm .005$	$0.34 \pm .01$	$0.37 \pm .01$	$0.316 \pm .008$
γ	$1.203 \pm .002$	$1.239 \pm .002$	$1.25 \pm .02$	$1.33 \pm .03$	$1.33 \pm .015$	$1.397 \pm .034$
δ	$4.4 \pm .4$	$4.85 \pm .03$	-	$3.95 \pm .15$	$4.3 \pm .1$	-
η	$0.1 \pm .1$	$0.017 \pm .015$	$0.08 \pm .017$	$0.21 \pm .05$	$0.07 \pm .04$	$0.067 \pm .01$
ν	≈ 0.57	$0.625 \pm .006$	$0.65 \pm .02$	$0.672 \pm .001$	$0.69 \pm .02$	-

The above inequality implies that the exponents must also satisfy the inequality

$$\alpha + 2\beta + \gamma \geq 2 \tag{1340}$$

Experimental evidence accumulated that the inequalities were in fact equalities. The exponents of $\alpha = 0$, $\beta = \frac{1}{8}$, $\gamma = \frac{7}{4}$, $\nu = 1$ and $\eta = \frac{1}{4}$, found from Onsager's exact solution of the two-dimensional Ising Model, also satisfy the same equalities. The exponent equalities include the Rushbrooke relation

$$\alpha + 2\beta + \gamma = 2 \tag{1341}$$

the Griffiths relation

$$\alpha + \beta(\delta + 1) = 2 \tag{1342}$$

which describes thermodynamic properties. Widom showed that the exponents would satisfy equalities if the singular part of the Helmholtz Free-Energy F did not depend separately on T and Π but, instead, can be written as

$$F(T, \Pi) = |t|^{\frac{1}{y}} \psi(\Pi / |t|^{\frac{x}{y}}) \tag{1343}$$

where

$$t = \left(\frac{T - T_c}{T_c}\right) \tag{1344}$$

Leo Kadanoff introduced the idea that the exponents expressing spatial correlations are also related. These relations include the Fisher relation

$$(2 - \eta)\nu = \gamma \tag{1345}$$

and the hyper-scaling relation or Josephson relation

$$\nu\, d \;=\; 2 \,-\, \alpha \tag{1346}$$

The Josephson relation is the only relation which involves the dimensionality d. It becomes invalid for sufficiently large d, that is when d exceeds the upper critical dimensionality d_c. For $d > d_c$, all the critical exponents become independent of d.

The scaling that is found in the proximity of a phase transition can be understood as a consequence of the fluctuations of the order parameter that occur as the phase transition is approached. The picture is that as the temperature is decreased towards the critical temperature, the material exhibits islands which are ordered whose spatial extent ξ increases with decreasing t. Furthermore, it is the long ranged large scale fluctuations that dominate the physical divergences. At the transition $t \to 0$ so $\xi \to \infty$, therefore, the system becomes scale invariant. The scaling hypothesis assumes that the correlation length ξ is the only relevant characteristic length scale of the system close to the transition and that all other length scales must be expressible in terms of ξ. Hence, the effects of the microscopic length scale a should be expressible in terms of the ratio $\frac{a}{\xi}$ which vanishes close to the transition. The temperature dependence of static properties can then be inferred from dimensional analysis. Thus, the Free-Energy (measured in units of $k_B T$) per unit volume has dimensions L^{-d} which, on substituting ξ for L, leads to a variation as ξ^{-d}. Since the specific heat has exponent $-\alpha$, and involves the second derivative of F w.r.t. T, F should scale as $|t|^{2-\alpha}$. If the correlation function is normalized to $L^{2-d-\eta}$ then, on noting that $S(r)$ is proportional to $< \phi(r)\,\phi(0) >$, one has

$$< \phi(0) > \; \sim \; L^{\frac{2-d-\eta}{2}} \tag{1347}$$

which then sets

$$< \phi(0) > \; \sim \; \xi^{\frac{2-d-\eta}{2}} \tag{1348}$$

Also since, from linear response theory, the susceptibility χ can be expressed as

$$\chi \; \sim \; \int d^d \underline{r} \; < \phi(\underline{r})\,\phi(0) > \tag{1349}$$

one has $\chi \sim L^{2-\eta}$ or

$$\chi \; \sim \; \xi^{2-\eta} \tag{1350}$$

On using $\xi \sim |t|^{-\nu}$ one obtains the scaling relations

$$
\begin{aligned}
2 \,-\, \alpha \;&=\; d\,\nu \\
\beta \;&=\; \frac{1}{2}\,(\,d\,-\,2\,+\,\eta\,)\,\nu \\
\gamma \;&=\; (\,2\,-\,\eta\,)\,\nu
\end{aligned}
\tag{1351}
$$

The first is recognized as the Josephson hyper-scaling relation and the last is the Fisher relation. The exponent δ can be obtained by first determining the length scale of the conjugate field Π from the definitive relation

$$\phi = -\left(\frac{\partial F}{\partial \Pi}\right) \tag{1352}$$

which leads to the identification that

$$
\begin{aligned}
\Pi \ &\sim \ \frac{F}{\phi} \\
&\sim \ L^{-d} \, L^{-\frac{2-d-\eta}{2}} \\
&\sim \ L^{-\frac{2+d-\eta}{2}}
\end{aligned}
\tag{1353}
$$

Since $\Pi \sim \phi^\delta$, one has

$$
\begin{aligned}
\phi^\delta \ &\sim \ L^{-\frac{2+d-\eta}{2}} \\
&\sim \ \xi^{-\frac{2+d-\eta}{2}}
\end{aligned}
\tag{1354}
$$

Hence, one finds the scaling relation

$$\beta \, \delta = \left(\frac{2 + d - \eta}{2}\right) \nu \tag{1355}$$

The two relations consisting of

$$\beta \, \delta = \left(\frac{2 + d - \eta}{2}\right) \nu \tag{1356}$$

and

$$\beta = \left(\frac{d - 2 + \eta}{2}\right) \nu \tag{1357}$$

can be shown to be equivalent to the Griffiths and Rushbrooke relations. The Griffith relation can be found by simply adding the two relations, yielding

$$\beta \, (\, \delta + 1 \,) = d \, \nu \tag{1358}$$

On using the Josephson relation, the above equation is recognized as yielding the Griffiths relation

$$\beta \, (\, \delta + 1 \,) = (\, 2 - \alpha \,) \tag{1359}$$

Likewise, on subtracting the relations expressed in eqn(1356) and eqn(1357)), one obtains

$$\beta \, (\, \delta - 1 \,) = (\, 2 - \eta \,) \, \nu \tag{1360}$$

which using the Fisher relation, yields

$$\beta \, (\, \delta - 1 \,) = \gamma \tag{1361}$$

Adding the above equation to the Griffiths relation leads to the Rushbrooke relation. Scaling analysis indicates that one may consider there to be only two independent exponents, such as ν and the anomalous dimension η, but does not fix their values.

10.5   Mean-Field Theory

Landau-Ginzburg Mean-Field Theory is an approximation which replaces the exact evaluation of the path integral for the partition function

$$Z = \int \mathcal{D}\phi \, \exp\left[-\beta \, F[\phi] \right] \tag{1362}$$

and, hence, the exact evaluation of the Free-Energy. It can be viewed as an approximate evaluation of the path integral analogous to the evaluation of an ordinary integral by using the method of steepest descents. In the method of steepest descents, one evaluates an integral of the form

$$I = \int_{-\infty}^{\infty} dx \, \exp\left[-\alpha \, f(x) \right] \tag{1363}$$

by finding the value of x, say x_0, for which the exponent $f(x)$ is minimum and then approximating $f(x)$ by a parabola

$$f(x) = f(x_0) + \frac{1}{2} \left(\frac{d^2 f}{dx^2} \right)\Bigg|_{x_0} (x - x_0)^2 + \dots \tag{1364}$$

This approximation is based on the assumption that the value of the integral I has its largest contribution from the region around x_0. This leads to the result

$$I \sim \sqrt{\frac{2\pi}{\alpha \left(\frac{d^2 f}{dx^2} \right)\Big|_{x_0}}} \, \exp\left[-\alpha \, f(x_0) \right] \tag{1365}$$

The dominant contribution to the integral is given by the exponential factor, that is the it is determined by the exponent at the value of x for which the exponent of the integrand is maximized. Mean-field theory approximates the path integral by its extremal value. That is, the partition function z is approximated by

$$Z \sim \exp\left[-\beta \, F[\phi_0] \right] \tag{1366}$$

where ϕ_0 is the ϕ which minimizes the functional $F[\phi]$.

Let us assume that there is a smooth $\phi_0(\underline{r})$ which minimizes $F[\phi]$ and then consider a set of ϕ's with the form

$$\phi(\underline{r}) = \phi_0(\underline{r}) + \lambda \, \delta\phi(\underline{r}) \tag{1367}$$

where $\delta\phi_0(\underline{r})$ is an arbitrary function which vanishes at the boundaries of the system. The parameter λ can be continuously varied from zero to unity, and represents the amplitude of the deviation from ϕ_0. For a fixed $\delta\phi$, the Free-energy functional depends on λ and can be written as $F(\lambda)$. the extremum condition inoplies that

$$\left(\frac{\partial F(\lambda)}{\partial \lambda} \right)\Bigg|_{\lambda=0} = 0 \tag{1368}$$

so the term first-order in λ in the expansion of $F[\phi_0 + \lambda\delta\phi]$ must be zero. That is, if

$$F[\phi_0 + \lambda\delta\phi] \;=\; F[\phi_0] \;+\; \lambda\,\delta F^1 \;+\; \ldots \tag{1369}$$

then for ϕ_0 to be an extremum, one requires that

$$\delta F^1 \;=\; 0 \tag{1370}$$

The explicit form of the above condition is that

$$0 \;=\; \int d^d\underline{r}\left[\, 2\,F_2\,(\,\underline{\phi}_0\cdot\underline{\delta\phi}\,) + 4\,F_4\,(\,\underline{\phi}_0\cdot\underline{\phi}_0\,)\,(\,\underline{\phi}_0\cdot\underline{\delta\phi}\,) - \underline{\mathrm{II}}\cdot\underline{\delta\phi} + 2\,c\,\underline{\nabla}\underline{\phi}_0\cdot\underline{\nabla}\underline{\delta\phi}\,\right] \tag{1371}$$

where the last term must be interpreted as involving two types of scalar product, as previously discussed. On integrating the last term by parts and recalling that $\delta\phi$ vanishes at the boundaries, one can write the integrand as a product of a factor that depends on ϕ_0 and a factor of $\delta\phi$,

$$0 \;=\; \int d^d\underline{r}\,\underline{\delta\phi}\cdot\left[\, 2\,F_2\,\underline{\phi}_0 \;+\; 4\,F_4\,\underline{\phi}_0\,(\,\underline{\phi}_0\cdot\underline{\phi}_0\,) \;-\; \underline{\mathrm{II}} \;-\; 2\,c\,\nabla^2\underline{\phi}_0\,\right] \tag{1372}$$

This must vanish for an arbitrary $\delta\phi$. One can chose $\delta\phi$ such that

$$\delta\phi(\underline{r}) \;=\; \delta^d\!\left(\,\underline{r}\;-\;\underline{r}_0\,\right) \tag{1373}$$

for some arbitrary point $\underline{r}_0$. The integration over $d^d\underline{r}$ can be performed leading to the requirement that ϕ_0 must satisfy the equation

$$0 \;=\; \left[\, 2\,F_2\,\underline{\phi}_0(\underline{r}_0) \;+\; 4\,F_4\,\underline{\phi}_0(\underline{r}_0)\,(\,\underline{\phi}_0(\underline{r}_0)\cdot\underline{\phi}_0(\underline{r}_0)\,) \;-\; \underline{\mathrm{II}}(\underline{r}_0) \;-\; 2\,c\,\nabla^2\underline{\phi}_0(\underline{r}_0)\,\right] \tag{1374}$$

for any arbitrarily chosen point $\underline{r}_0$. The functions $\phi_0(\underline{r})$ which satisfy the above equation extremalizes $F[\phi]$ for any choice of $\delta\phi(\underline{r})$. We shall write eqn(1374) in the form

$$\left[\, 2\,F_2 \;+\; 4\,F_4\,(\,\underline{\phi}_0(\underline{r})\cdot\underline{\phi}_0(\underline{r})\,) \;-\; 2\,c\,\nabla^2\,\right]\underline{\phi}_0(\underline{r}) \;=\; \underline{\mathrm{II}}(\underline{r}) \tag{1375}$$

in which the spatially varying applied field acts as a source. This equation governs all the extrema of $F[\phi]$.

We shall first consider physical properties associated with the extrema for which ϕ is uniform across the system, and then consider the physical properties associated with the spatially varying solutions.

Uniform Solutions

260

The differential equation simplifies for spatially uniform solution and zero applied fields, $\underline{\Pi} = 0$, to

$$\left[2\,F_2 \; + \; 4\,F_4 \; \underline{\phi} \cdot \underline{\phi} \right] \underline{\phi} \; = \; 0 \tag{1376}$$

which has the solutions

$$\underline{\phi} \; = \; 0 \tag{1377}$$

and

$$\underline{\phi} \cdot \underline{\phi} \; = \; -\,\frac{2\,F_2}{4\,F_4} \tag{1378}$$

The second solution only makes sense if

$$\underline{\phi} \cdot \underline{\phi} \; > \; 0 \tag{1379}$$

since the order parameter is assumed to be a real vector quantity. When

$$-\,\frac{2\,F_2}{4\,F_4} \; < \; 0 \tag{1380}$$

the only physically acceptable solution corresponds to $\underline{\phi} = 0$. The magnitude of order parameter may take on a non-zero value if

$$-\,\frac{2\,F_2}{4\,F_4} \; > \; 0 \tag{1381}$$

In this case, in addition to the solution $\underline{\phi} = 0$, there exists the possibility of continuously degenerate solutions. Due to stability considerations, one must have $F_4 > 0$. Hence, a second-order phase transition may occur when F_2 changes sign. This motivates the notation

$$F_2 \; = \; A\,(\,T \; - \; T_c\,) \tag{1382}$$

with $A > 0$. For $T > T_c$, there is only one unique solution which is given by $\underline{\phi} = 0$, so the mean-field value of the Free-Energy is given by

$$F[\phi = 0] \; = \; V\,F_0 \tag{1383}$$

whereas for $T < T_c$, one has the possibility of an additional solution corresponding to

$$\underline{\phi}_0 \cdot \underline{\phi}_0 \; = \; \frac{2\,A\,(\,T_c \; - \; T\,)}{4\,F_4} \tag{1384}$$

which fixes the magnitude of ϕ_0 as

$$\phi_0 \; = \; \sqrt{\frac{2\,A\,(\,T_c \; - \; T\,)}{4\,F_4}} \tag{1385}$$

The non-zero solution is continuously degenerate with respect to the orientation of the vector order parameter. The presence of an infinitesimal applied

magnetic field allows the system to choose a direction of $\underline{\phi}$ which spontaneously breaks the rotational symmetry of the Hamiltonian. The mean-field Free-Energy corresponding the broken symmetry solution is found as

$$
\begin{aligned}
F[\phi_0] & = V \left[F_0 - \frac{2 A^2 (T_c - T)^2}{4 F_4} + \frac{A^2 (T_c - T)^2}{4 F_4} \right] \\
& = V \left[F_0 - \frac{A^2 (T_c - T)^2}{4 F_4} \right]
\end{aligned}
\tag{1386}
$$

which is lower than the Free-Energy of the solution with $\phi = 0$. Hence, in the mean-field approximation, the system will condense into the broken symmetry state. Thus, the mean-field approximation describes a second-order transition at T_c and, for $T < T_c$, the magnitude of the order parameter will have a temperature-dependence given by

$$
| \underline{\phi}_0 | = \sqrt{\frac{2 A (T_c - T)}{4 F_4}}
\tag{1387}
$$

The value of the mean-field critical exponent for the order parameter, β, is, therefore, fixed at $\beta = \frac{1}{2}$.

The order parameter exponent on the critical isotherm, δ, defined via

$$
| \Pi | \sim | \phi |^{\delta}
\tag{1388}
$$

is found from the equation

$$
\left[2 A (T - T_c) + 4 F_4 \underline{\phi}_0 \cdot \underline{\phi}_0 \right] \underline{\phi}_0 = \underline{\Pi}
\tag{1389}
$$

which for $T = T_c$ reduces to

$$
4 F_4 (\underline{\phi}_0 \cdot \underline{\phi}_0) \underline{\phi}_0 = \underline{\Pi}
\tag{1390}
$$

which leads to the identification of the critical exponent $\delta = 3$.

The mean-field specific heat C is calculated from

$$
C = - T \left(\frac{\partial^2 F}{\partial T^2} \right)
\tag{1391}
$$

which, for $T > T_c$, yields

$$
C = 0 \qquad \text{for } T > T_c
\tag{1392}
$$

but, for $T < T_c$ gives

$$
C = V \frac{A^2 T}{2 F_4} \qquad \text{for } T > T_c
\tag{1393}
$$

Hence, in the mean-field approximation, the specific heat exhibits a discontinuous jump at T_c, of magnitude $V \frac{A^2 \, T_c}{2 \, F_4}$. The appropriate mean-field specific heat exponents are $\alpha = \alpha' = 0$.

The above solutions are all uniform, due to the homogeneity of space, and due to the uniformity of the applied fields.

Spatially Varying Solutions

It has been shown that the non-uniform configurations $\underline{\phi}(\underline{r})$ which minimize $F[\phi]$ satisfy the partial differential equation

$$\left[2 \, F_2 \; + \; 4 \, F_4 \; \underline{\phi}(\underline{r}) \cdot \underline{\phi}(\underline{r}) \; - \; 2 \, c \, \nabla^2 \right] \underline{\phi}(\underline{r}) \; = \; \underline{\Pi}(\underline{r}) \tag{1394}$$

We shall consider spatially varying solutions, in which the spatial variations are induced by an applied field which has a small non-uniform part

$$\underline{\Pi}(\underline{r}) \; = \; \underline{\Pi}_0 \; + \; \underline{\delta\Pi}(\underline{r}) \tag{1395}$$

The local order-parameter $\underline{\phi}(\underline{r})$ will be expressed in terms of a uniform component $\underline{\phi}_0$ and a spatially varying part $\underline{\delta\phi}(\underline{r})$

$$\underline{\phi}(\underline{r}) \; = \; \underline{\phi}_0 \; + \; \underline{\delta\phi}(\underline{r}) \tag{1396}$$

The spatially varying part of the order-parameter $\underline{\delta\phi}(\underline{r})$ vanishes in the limit that $\underline{\delta\Pi}(\underline{r})$ vanishes. In this limit, $\underline{\phi}_0$ minimizes $\overline{F[\phi]}$ in the presence of $\underline{\Pi}_0$. The terms of first-order in the small spatially varying components satisfy the equation

$$\left[2 \, F_2 + 4 \, F_4 \, \underline{\phi}_0 \cdot \underline{\phi}_0 - 2 \, c \, \nabla^2 \right] \underline{\delta\phi}(\underline{r}) + 8 \, F_4 \, \underline{\phi}_0 \, (\, \underline{\phi}_0 \cdot \underline{\delta\phi}(\underline{r}) \,) \; = \; \underline{\delta\Pi}(\underline{r}) \tag{1397}$$

This equation indicates that, for temperatures below T_c, the mean field response will be different depending on the relative orientation of the spatially varying field $\underline{\delta\Pi}(\underline{r})$ and the direction of the uniform order parameter $\underline{\phi}_0$. For temperatures above the critical temperature T_c, the vector order parameter vanishes and the equation simplifies to

$$\left[2 \, F_2 \; - \; 2 \, c \, \nabla^2 \right] \underline{\delta\phi}(\underline{r}) \; = \; \underline{\delta\Pi}(\underline{r}) \tag{1398}$$

In this case, the magnitude of the induced vector order parameter is independent of the orientation of the spatially varying applied field. This is expected since, in the absence of the order parameter which has spontaneously broken the symmetry, the system is isotropic.

Longitudinal Response

263

For temperatures below the critical temperature and when the non-uniform part of the applied field $\underline{\delta\Pi}(\underline{r})$ is parallel to $\underline{\phi}_0$, the mean-field response is longitudinal and satisfies the equation

$$\left[2 \, F_2 \; + \; 12 \, F_4 \; \underline{\phi}_0 \cdot \underline{\phi}_0 \; - \; 2 \, c \, \nabla^2 \right] \delta\phi_L(\underline{r}) \; = \; \delta\Pi_L(\underline{r}) \qquad (1399)$$

This equation is also valid for temperatures above T_c, where ϕ_0 vanishes, although longitudinal and transverse is undefined. It should be noted that the equation has a different form in the two temperature regimes.

Transverse Response

The transverse response is only defined for temperatures below the critical temperature. If $\underline{\delta\Pi}(\underline{r})$ is transverse to $\underline{\phi}_0$, the mean-field response is determined from the partial differential equation

$$\left[2 \, F_2 \; + \; 4 \, F_4 \; \underline{\phi}_0 \cdot \underline{\phi}_0 \; - \; 2 \, c \, \nabla^2 \right] \delta\phi_T(\underline{r}) \; = \; \delta\Pi_T(\underline{r}) \qquad (1400)$$

In the limit $\underline{\Pi}_0 \to 0$, the uniform order parameter $\underline{\phi}_0$ satisfies the equation

$$\left[2 \, F_2 \; + \; 4 \, F_4 \; \underline{\phi}_0 \cdot \underline{\phi}_0 \right] \; = \; 0 \qquad (1401)$$

Hence, the partial differential equation for the transverse response in the mean-field approximation simplifies to

$$- \, 2 \, c \, \nabla^2 \, \delta\phi_T(\underline{r}) \; = \; \delta\Pi_T(\underline{r}) \qquad (1402)$$

The solution of this equation determines the order parameter for which the Landau-Ginzberg Free-Energy Functional is extremal.

The Mean-Field Correlation Functions

These difference in the response show that the correlation function $S(\underline{r})$ involved in the linear response theory must be considered as a tensor quantity. The mean-field equations for the order parameter allows one to calculate the (mean-field) tensor correlation function. Linear response theory describes how $\delta\underline{\phi}(\underline{r})$ is related to $\underline{\delta\Pi}(\underline{r})$. In particular, if a tensor correlation function $S_{i,j}(\underline{r})$ is defined via

$$S_{i,j}(\underline{r} - \underline{r}') \; = \; < \, \phi_i(\underline{r}) \; \phi_j(\underline{r}') \, > \; - \; < \, \phi_i(\underline{r}) \, > \, < \, \phi_j(\underline{r}') \, > \qquad (1403)$$

one finds that components satisfy the linear response relations

$$\delta\phi_i(\underline{r}) \; = \; \beta \, \sum_j \, \int d^d\underline{r}' \; S_{i,j}(\underline{r} - \underline{r}') \; \delta\Pi_j(\underline{r}') \qquad (1404)$$

Substitution of the linear response relation into eqn(1397) results in an integral equation which is linear in the components of $\delta\Pi$. The left-hand side involves an integral over $\underline{r}'$ of the components of $\delta\Pi(\underline{r}')$ weighted by a matrix function involving the derivatives of $S_{i,j}(\underline{r} - \underline{r}')$. The right-hand side only depends on a component of $\delta\Pi(r)$. The integral equation has to be satisfied for an arbitrary $\delta\Pi(\underline{r})$. The equation is satisfied if the matrix function is proportional to a d-dimensional delta function $\delta^d(\underline{r} - \underline{r}')$. The integration over $\underline{r}'$ can then be trivially performed. This results in a matrix equation which, after eliminating the arbitrary spatially varying field, determines the tensor response function.

We shall first examine the longitudinal response function S_L which satisfies the linear response relation

$$\delta\phi_L(\underline{r}) \;=\; \beta \int d^d\underline{r}' \; S_L(\underline{r} - \underline{r}') \; \delta\Pi_L(\underline{r}') \tag{1405}$$

On substituting the longitudinal linear response relation into eqn(1399) which describes the mean-field order parameter, one finds

$$\left[2\,F_2 \;+\; 12\,F_4\,\underline{\phi}_0 \cdot \underline{\phi}_0 \;-\; 2\,c\,\nabla_{\underline{r}}^2 \right] \int d^d\underline{r}' \; S_L(\underline{r} - \underline{r}') \; \delta\Pi_L(\underline{r}') \;=\; k_B\,T\,\delta\Pi_L(\underline{r}) \tag{1406}$$

On changing the order of integration and differentiation, one obtains

$$\int d^d\underline{r}' \left[2\,F_2 \;+\; 12\,F_4\,\underline{\phi}_0 \cdot \underline{\phi}_0 \;-\; 2\,c\,\nabla_{\underline{r}}^2 \right] S_L(\underline{r} - \underline{r}') \; \delta\Pi_L(\underline{r}') \;=\; k_B\,T\,\delta\Pi_L(\underline{r}) \tag{1407}$$

where $\delta\Pi_L$ is to be regarded as an arbitrary function. The equation is satisfied for any $\delta\Pi_L(\underline{r})$, if the (mean-field) longitudinal correlation function satisfies

$$\left[2\,F_2 \;+\; 12\,F_4\,\underline{\phi}_0 \cdot \underline{\phi}_0 \;-\; 2\,c\,\nabla_{\underline{r}}^2 \right] S_L(\underline{r} - \underline{r}') \;=\; k_B\,T\,\delta^d\!\left(\underline{r} - \underline{r}' \right) \tag{1408}$$

as can be verified by substitution. Hence, we have determined a partial differential equation which determines the longitudinal correlation function below T_c and also determines the correlation function above T_c,

Likewise, one can determine the transverse correlation function $S_T(\underline{r} - \underline{r}')$ which satisfies the linear response relation

$$\delta\phi_T(\underline{r}) \;=\; \beta \int d^d\underline{r}' \; S_T(\underline{r} - \underline{r}') \; \delta\Pi_T(\underline{r}') \tag{1409}$$

On substituting the transverse linear response relation into eqn(1402), which describes the mean-field order parameter, one finds

$$- 2\,c\,\nabla_{\underline{r}}^2 \int d^d\underline{r}' \; S_T(\underline{r} - \underline{r}') \; \delta\Pi_T(\underline{r}') \;=\; k_B\,T\,\delta\Pi_T(\underline{r}) \tag{1410}$$

On changing the order of integration and differentiation, one obtains

$$\int d^d \underline{r}' \left[\, - \, 2 \, c \, \nabla_{\underline{r}}^2 \, S_T(\underline{r} - \underline{r}') \right] \delta\Pi_T(\underline{r}') \; = \; k_B \, T \, \delta\Pi_T(\underline{r}) \qquad (1411)$$

where $\delta\Pi_T$ is to be regarded as an arbitrary function. The equation is satisfied for any $\delta\Pi_T(\underline{r})$, if the (mean-field) transverse correlation function satisfies

$$- \, 2 \, c \, \nabla_{\underline{r}}^2 \, S_T(\underline{r} - \underline{r}') \; = \; k_B \, T \, \delta^d\!\left(\underline{r} \, - \, \underline{r}' \right) \qquad (1412)$$

The above equation can be used to determine the transverse correlation function in the mean-field approximation.

For $T \; > \; T_c$ the mean-field order parameter is zero, so the correlation function satisfies

$$\left(\, 2 \, A \, (\, T \, - \, T_c \,) \; - \; 2 \, c \, \nabla^2 \, \right) S(\underline{r} - \underline{r}') \; = \; k_B \, T \, \delta^d\!\left(\underline{r} \, - \, \underline{r}' \right) \quad (1413)$$

In the low temperature phase, $T \; < \; T_c$, the order parameter is given by

$$| \, \underline{\phi}_0 \, |^2 \; = \; - \, \frac{A \, (\, T \, - \, T_c \,)}{2 \, F_4} \qquad (1414)$$

therefore, the longitudinal correlation function satisfies the partial differential equation

$$\left(\, - \, 4 \, A \, (\, T \, - \, T_c \,) \; - \; 2 \, c \, \nabla_{\underline{r}}^2 \, \right) S_L(\underline{r} - \underline{r}') \; = \; k_B \, T \, \delta^d\!\left(\underline{r} \, - \, \underline{r}' \right) \quad (1415)$$

Dimensional analysis of the above two equations indicates that the correlation length ξ should be given by

$$\xi \; = \; \sqrt{\frac{c}{A \, (\, T \, - \, T_c \,)}} \qquad (1416)$$

for $T \; > \; T_c$, and

$$\xi \; = \; \sqrt{\frac{c}{2 \, A \, (\, T_c \, - \, T \,)}} \qquad (1417)$$

for $T \; < \; T_c$. Since the correlation length is defined to scale as

$$\xi \; \sim \; | \, T \, - \, T_c \, |^{-\nu} \qquad (1418)$$

one finds that the critical exponents are the same above and below T_c and are given by $\nu \; = \; \nu' \; = \; \frac{1}{2}$.

The above equations can be expressed in terms of the correlation length ξ as

$$\left(\, \xi^{-2} \; - \; \nabla_{\underline{r}}^2 \, \right) S(\underline{r} - \underline{r}') \; = \; \frac{k_B \, T}{2 \, c} \, \delta^d\!\left(\underline{r} \, - \, \underline{r}' \right) \qquad (1419)$$

266

The equation can be solved by Fourier transforming, leading to

$$\left(\xi^{-2} + \underline{k}^2 \right) S(\underline{k}) = \frac{k_B T}{2 c} \tag{1420}$$

Thus,

$$S(\underline{k}) = \left(\frac{k_B T}{2 c} \right) \left(\frac{1}{\xi^{-2} + \underline{k}^2} \right) \tag{1421}$$

The correlation function $S(\underline{r})$ is given by the inverse Fourier Transformation

$$
\begin{aligned}
S(\underline{r}) &= \int \frac{d^d \underline{k}}{(2 \pi)^d} \, \exp\left[i \, \underline{k} \cdot \underline{r} \right] S(\underline{k}) \\
&= \left(\frac{k_B T}{2 c} \right) \int \frac{d^d \underline{k}}{(2 \pi)^d} \, \frac{\exp\left[i \, \underline{k} \cdot \underline{r} \right]}{\xi^{-2} + \underline{k}^2}
\end{aligned}
\tag{1422}
$$

For three dimensions, the integral is evaluated as

$$
\begin{aligned}
S(\underline{r}) &= \left(\frac{k_B T}{2 c} \right) \int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \, \sin\theta \int_0^{\infty} \frac{d^3 k}{(2 \pi)^3} \, \frac{\exp\left[i \, k \, r \, \cos\theta \right]}{\xi^{-2} + k^2} \\
&= \left(\frac{k_B T}{2 c} \right) \int_0^{\pi} d\theta \, \sin\theta \int_0^{\infty} \frac{dk}{(2 \pi)^2} \, k^2 \, \frac{\exp\left[i \, k \, r \, \cos\theta \right]}{\xi^2 + k^2} \\
&= \left(\frac{k_B T}{2 c} \right) \int_0^{\infty} \frac{dk}{(2 \pi)^2} \, \frac{k^2}{i \, k \, r} \left(\frac{\exp\left[+ i \, k \, r \right] - \exp\left[- i \, k \, r \right]}{(\xi^{-2} + k^2)} \right) \\
&= \left(\frac{k_B T}{4 \pi c} \right) \frac{1}{r} \int_{-\infty}^{\infty} \frac{dk}{(2 \pi i)} \, \exp\left[+ i \, k \, r \right] \frac{k}{(\xi^{-2} + k^2)} \tag{1423}
\end{aligned}
$$

where two terms have been combined extending the integration over k to the range $-\infty$ to $+\infty$. The remaining integration can be performed using Cauchy's theorem by completing the contour with a semi-circle in the upper-half complex plane. The contribution of the semi-circular contour at infinity vanishes due to Jordan's lemma. The integral is dominated by the residue at the pole $k = i \, \xi^{-1}$, leading to

$$S(r) = \left(\frac{k_B T}{8 \pi c} \right) \frac{\exp\left[- \frac{r}{\xi} \right]}{r} \tag{1424}$$

which leads to the identification of the (three-dimensional) [49] mean-field value of the critical exponent η as $\eta = 0$.

---

[49] The three-dimensional result is quite special. A more general form of the mean-field form

The susceptibility χ in the high temperature phase and the longitudinal susceptibility χ_L in the low temperature ordered phase can be obtained from linear response theory. The susceptibility is given by

$$\chi = \beta \int d^d \underline{r} \; S(\underline{r}) \tag{1427}$$

which, in three-dimensions leads to the mean-field susceptibility given by

$$
\begin{aligned}
\chi &= \left(\frac{1}{8 \pi c} \right) \int d^3 \underline{r} \; \frac{\exp\left[-\frac{r}{\xi} \right]}{r} \\
&= \left(\frac{1}{2 c} \right) \int_0^\infty dr \; r \; \exp\left[-\frac{r}{\xi} \right] \\
&= \left(\frac{\xi^2}{2 c} \right) \int_0^\infty dx \; x \; \exp\left[-x \right]
\end{aligned} \tag{1428}
$$

where the dimensionless parameter x

$$x = \left(\frac{r}{\xi} \right) \tag{1429}$$

has been introduced in the last line. The integral can be evaluated using integration by parts, leading to

$$\chi = \frac{\xi^2}{2 c} \tag{1430}$$

Hence the susceptibilities are given by

$$\chi = \frac{1}{2 A \left(T - T_c \right)} \tag{1431}$$

for $T > T_c$, whereas, for $T < T_c$, the longitudinal susceptibility is given by

$$\chi_L = \frac{1}{4 A \left(T_c - T \right)} \tag{1432}$$

Both susceptibilities diverge as the transition temperature is approached. Thus, in mean-field theory the susceptibility critical exponent γ is the same in the

of $S(\underline{r})$ in d-dimensions is given by

$$S(r) = \left(\frac{k_B T}{2 c S_d \left(d - 2 \right)} \right) \frac{\exp\left[-\frac{r}{\xi} \right]}{r^{(d-2)}} F\left(\frac{r}{\xi} \right) \tag{1425}$$

where $F(x)$ has the properties that $F(0) = 1$ and for large r

$$F\left(\frac{r}{\xi} \right) \sim \left(\frac{r}{\xi} \right)^{\frac{(d-3)}{2}} \tag{1426}$$

For $d = 3$ the function reduces to a constant, which is unity.

low and high temperature phases and is given by $\gamma = \gamma' = 1$. However, the amplitude of the divergent term is a factor of two greater for the high temperature phase. The transverse susceptibility χ_T can also be calculated from the transverse correlation function $S_T(\underline{r})$. The correlation length for the transverse correlation function diverges for all temperatures below T_c. The divergence of transverse correlation length is connected with Goldstone's theorem. It has the effect that, $S_T(\underline{r})$ satisfies the equation

$$- 2 \, c \, \nabla^2 \, S_T(\underline{r}) \;=\; k_B \, T \, \delta^d(\underline{r}) \tag{1433}$$

One can integrate the equation for S_T over $\underline{r}$, where the integration runs over a volume which contains the origin, and then use Gauss's theorem to express the remaining volume integral as an integral over the surface of the volume

$$- 2 \, c \int d^d\underline{r} \; \nabla^2 \, S_T(\underline{r}) \;=\; k_B \, T \int d^d\underline{r} \; \delta^d(\underline{r})$$

$$- 2 \, c \int d^d\underline{r} \; \nabla^2 \, S_T(\underline{r}) \;=\; k_B \, T$$

$$- 2 \, c \int d\underline{S}^{d-1} \, . \, \underline{\nabla} \, S_T(\underline{r}) \;=\; k_B \, T \tag{1434}$$

where the direction of the $(d-1)$ dimensional surface element, $d\underline{S}^{d-1}$, is defined to be normal to the surface of integration. Since $S_T(\underline{r})$ is spherically symmetric, the integration is easily performed over the surface of a hyper-sphere of radius r, leading to

$$- 2 \, c \, S_d \, r^{d-1} \, \hat{e}_r \, . \, \underline{\nabla} \, S_T(\underline{r}) \;=\; k_B \, T \tag{1435}$$

where S_d is the surface area of a d-dimensional unit sphere. Thus,

$$\left(\frac{\partial S_T}{\partial r} \right) \;=\; - \left(\frac{k_B \, T}{2 \, c \, S_d} \right) \frac{1}{r^{d-1}} \tag{1436}$$

For $d > 2$, the expression can be integrated leading to the transverse correlation function $S_T(r)$ being given by

$$S_T(r) \;=\; \left(\frac{k_B \, T}{2 \, c \, (d-2) \, S_d} \right) \frac{1}{r^{d-2}} \tag{1437}$$

The transverse susceptibility, χ_T, can then be found from S_T via

$$\chi_T \;=\; \beta \int d^d\underline{r} \; S_T(\underline{r})$$

$$\sim \int_0^L dr \; r^{d-1} \, \frac{1}{r^{d-2}}$$

$$\sim \int_0^L dr \; r$$

$$\sim \; L^2 \tag{1438}$$

and diverges in the thermodynamic limit where the linear dimension of the system, L, is sent to infinity. The transverse susceptibility χ_T is infinite since the application of a small transverse field can cause the vector order parameter to re-orient.

The results for the high temperature and the longitudinal susceptibilities could have been determined directly from the equation

$$\left[2\, F_2 + 4\, F_4\, \underline{\phi}_0 \cdot \underline{\phi}_0 - 2\, c\, \nabla^2 \right] \delta\underline{\phi}(\underline{r}) + 8\, F_4\, \underline{\phi}_0\, (\, \underline{\phi}_0 \cdot \delta\underline{\phi}(\underline{r})\,) \; = \; \delta\underline{\Pi}(\underline{r}) \quad (1439)$$

by considering the limit in which $\delta\phi$ and $\delta\Pi$ become independent of $\underline{r}$. In this limit, one obtains

$$\left[2\, F_2 \; + \; 4\, F_4\, \underline{\phi}_0 \cdot \underline{\phi}_0 \right] \delta\underline{\phi} \; + \; 8\, F_4\, \underline{\phi}_0\, (\, \underline{\phi}_0 \cdot \delta\underline{\phi}\,) \; = \; \delta\underline{\Pi} \qquad (1440)$$

The differential longitudinal susceptibility χ_L is defined as

$$\chi_L \;=\; \frac{\delta\phi_L}{\delta\Pi_L} \qquad (1441)$$

is found to be given by

$$\chi_L \;=\; \frac{1}{\left[2\, F_2 \; + \; 12\, F_4\, \underline{\phi}_0 \cdot \underline{\phi}_0 \right]} \qquad (1442)$$

which is evaluated as

$$\chi_L \;=\; \frac{1}{4\, A\, (\, T_c \; - \; T\,)} \qquad (1443)$$

The susceptibility in the disordered phase, where $\phi_0 = 0$, is given by

$$\chi \;=\; \frac{1}{2\, F_2} \qquad (1444)$$

which reduces to

$$\chi \;=\; \frac{1}{2\, A\, (\, T \; - \; T_c\,)} \qquad (1445)$$

as found previously.

10.6 The Gaussian Approximation

The Gaussian approximation is an approximation in which one evaluates the corrections to mean-field theory due to the small amplitude fluctuations about the mean-field ground state. The Gaussian approximation is analogous to the steepest descents evaluation of an integral of the form

$$I \;=\; \int_{-\infty}^{\infty} dx \; \exp\left[\, - \, \alpha\, f(x) \right] \qquad (1446)$$

The mean-field approximation corresponds to approximating the integral by the exponential of the function $f(x_0)$ at the value x_0 which minimizes $f(x)$. The Gaussian approximation includes the corrections due to integrating the lowest-order (quadratic) terms in the Taylor expansion of $f(x)$. This leads to the result

$$
\begin{aligned}
I \quad &\sim \quad \sqrt{\frac{2\,\pi}{\alpha\,\left(\frac{d^2 f}{dx^2}\right)\Big|_{x_0}}} \, \exp\left[\, -\, \alpha\, f(x_0)\,\right] \\[2mm]
&=\quad \exp\left[\, -\, \alpha\, f(x_0)\, -\, \frac{1}{2}\, \ln\left(\frac{\alpha\,\left(\frac{d^2 f}{dx^2}\right)\Big|_{x_0}}{2\,\pi}\right)\right] \qquad (1447)
\end{aligned}
$$

which gives a correction to the exponent which is a factor of α^{-1} smaller than the leading contribution.

The Gaussian approximation goes one step beyond the mean-field approximation. The mean field approximation finds the most probable configuration ϕ_0 and takes into account the small amplitude fluctuations $\underline{\phi}(\underline{r})$. The spatially varying order parameter can be represented as

$$
\underline{\phi}(\underline{r}) \;=\; \underline{\phi}_0 \;+\; \underline{\delta\phi}(\underline{r}) \qquad (1448)
$$

in which the fluctuations $\underline{\delta\phi}(\underline{r})$ are assumed to have small amplitudes

$$
|\,\underline{\delta\phi}(\underline{r})\,| \;\ll\; |\,\underline{\phi}_0\,| \qquad (1449)
$$

The small amplitude fluctuations can be expressed in terms of their Fourier components $\underline{\phi}_{\underline{k}}$ via

$$
\underline{\delta\phi}(\underline{r}) \;=\; \frac{1}{\sqrt{V}} \, \sum_{\underline{k}} \, \exp\left[\, +\, i\,\underline{k}\cdot\underline{r}\,\right] \underline{\phi}_{\underline{k}} \qquad (1450)
$$

It should be noted that since the $\underline{\delta\phi}(\underline{r})$ are real functions, then

$$
\underline{\phi}_{\underline{k}}^{*} \;=\; \underline{\phi}_{-\underline{k}} \qquad (1451)
$$

and $\underline{\phi}_{\underline{k}}$ are complex functions with real and imaginary parts. The $\underline{k}$-space representation can be used to simplify the expression for the Landau-Ginzberg Free-Energy Functional. The path integral can be re-expressed in terms of integrals over the Fourier components $\phi_{i,\underline{k}}$

$$
\begin{aligned}
Z \quad &=\quad \exp\left[\, -\, \beta\, F\,\right] \\[2mm]
&=\quad \prod_{i,\underline{k}} \left\{\, \int d\phi_{i,\underline{k}}\,\right\} \exp\left[\, -\, \beta\, F[\phi]\,\right] \qquad (1452)
\end{aligned}
$$

where we are formally assuming that the Fourier components $\phi_{i,\underline{k}}$ are independent fields. The Gaussian approximation retains terms in $F[\phi]$ up to quadratic order in the $\phi_{i,\underline{k}}$. This allows the functional integral of the resulting approximate integrand to be evaluated exactly.

Like the mean-field approximation, the Gaussian approximation takes on different forms in the ordered and disordered phases.

The Gaussian Approximation for $T > T_c$

For $T > T_c$ the mean-field order parameter is given by $\underline{\phi}_0 = 0$, and the non-trivial part of the Free-Energy Functional can be expressed as

$$\int d^d\underline{r} \left[\, F_2 \, \underline{\delta\phi}^2(\underline{r}) \; + \; c \left(\underline{\nabla} \, \underline{\delta\phi} \right)^2 \; + \; F_4 \left(\underline{\delta\phi}^2(\underline{r}) \right)^2 \right] \tag{1453}$$

which can be written in terms of the Fourier components as

$$\sum_{\underline{k}} \left(F_2 + c \, \underline{k}^2 \right) \underline{\phi}_{\underline{k}} \cdot \underline{\phi}_{-\underline{k}} + \frac{1}{V} \sum_{\underline{k}_1,\underline{k}_2,\underline{k}_3} F_4 \left(\underline{\phi}_{\underline{k}_1} \cdot \underline{\phi}_{\underline{k}_2} \right) \left(\underline{\phi}_{\underline{k}_3} \cdot \underline{\phi}_{-\underline{k}_1-\underline{k}_2-\underline{k}_3} \right) \tag{1454}$$

In obtaining the above expression, we have used the identity

$$\int d^d\underline{r} \, \exp \left[\, i \, (\, \underline{k}_1 \, + \, \underline{k}_2 \, + \, \underline{k}_3 \, + \, \underline{k}_4 \,) \cdot \underline{r} \right] \; = \; V \, \delta_{\underline{k}_1+\underline{k}_2+\underline{k}_3+\underline{k}_4} \tag{1455}$$

where δ is the Kronecker delta function. In the Gaussian approximation, where $\phi_{\underline{k}}$ is assumed to be small, one neglects the fourth-order term proportional to F_4. Hence, one has

$$\begin{aligned} F[\phi] \;\; &\approx \;\; F_0 \, V \; + \; \sum_{\underline{k}} \left(\, F_2 \, + \, c \, \underline{k}^2 \, \right) \underline{\phi}_{\underline{k}} \cdot \underline{\phi}_{-\underline{k}} \\[2mm] &\approx \;\; F_0 \, V \; + \; \sum_{\underline{k}} \sum_{i=1}^{n} \left(\, A \, (\, T \, - \, T_c \,) \, + \, c \, \underline{k}^2 \, \right) \phi_{i,\underline{k}} \, \phi_{i,-\underline{k}} \end{aligned} \tag{1456}$$

The condition

$$\underline{\phi}_{\underline{k}}^* \; = \; \underline{\phi}_{-\underline{k}} \tag{1457}$$

relates the fields at points $\underline{k}$ and $-\underline{k}$. Since the two fields are not independent, it is convenient to partition $\underline{k}$-space into two disjoint regions, one region denoted by $'$ which contains the set of points $\underline{k}$ and a second region that contains all the points $-\underline{k}$ obtained by inversion of the points in the region $'$. The primed region, $'$, is chosen such that all points of k-space are contained in either the region $'$ or its inversion partner. The Gaussian functional integral is evaluated by first re-writing it as

$$Z \;\; = \;\; \exp \left[\, - \, \beta \, F \right]$$

$$
= \prod_{i,\underline{k}}' \left\{ \int d\phi_{i,\underline{k}} \right\} \exp\left[-\beta \, F[\phi] \right]
$$

$$
= \prod_{i,\underline{k}}' \left\{ \int d\phi_{i,\underline{k}} \int d\phi_{i,-\underline{k}} \right\} \exp\left[-\beta \, F[\phi] \right]
$$

$$
= \prod_{i,\underline{k}}' \left\{ \int d\phi_{i,\underline{k}} \int d\phi_{i,\underline{k}}^{*} \right\} \exp\left[-\beta \, F[\phi] \right] \tag{1458}
$$

where the values of $\underline{k}$ in the primed products are restricted to the region $'$. The variable of integration is changed from $\phi_{i,\underline{k}}$ and $\phi_{i,\underline{k}}^{*}$ to the real and imaginary parts of the components of the field, $\Re e \phi_{i,\underline{k}}$ and $\Im m \phi_{i,\underline{k}}$

$$
Z = \prod_{i,\underline{k}}' \left\{ 2 \int d\Re e \phi_{i,\underline{k}} \int d\Im m \phi_{i,\underline{k}} \right\} \exp\left[-\beta \, F[\phi] \right] \tag{1459}
$$

where the Jacobian of the transformation is 2. The approximate Free-Energy functional is also re-written as a summation over $\underline{k}$ where the $\underline{k}$-values in the summation are restricted to the primed region.

$$
\begin{aligned}
F[\phi] &\approx F_0 \, V + 2 \sum_{\underline{k}}' \sum_{i=1}^{n} \left(A \, (T - T_c) + c \, \underline{k}^2 \right) \phi_{i,\underline{k}} \, \phi_{i,-\underline{k}} \\
&\approx F_0 \, V + 2 \sum_{\underline{k}}' \sum_{i=1}^{n} \left(A \, (T - T_c) + c \, \underline{k}^2 \right) \phi_{i,\underline{k}} \, \phi_{i,\underline{k}}^{*} \\
&\approx F_0 \, V + 2 \sum_{\underline{k}}' \sum_{i=1}^{n} \left(A \, (T - T_c) + c \, \underline{k}^2 \right) \left((\Re e \phi_{i,\underline{k}})^2 + (\Im m \phi_{i,\underline{k}})^2 \right)
\end{aligned}
\tag{1460}
$$

On performing the Gaussian integrals, one finds that the Partition function Z is approximated by

$$
\begin{aligned}
Z &\approx \exp\left[-\beta \, F_0 \, V \right] \prod_{i=1}^{n} \prod_{\underline{k}}' \left[\frac{A \, (T - T_c) + c \, \underline{k}^2}{2 \, \pi \, k_B \, T} \right]^{-1} \\
&\approx \exp\left[-\beta \, F_0 \, V \right] \prod_{i=1}^{n} \prod_{\underline{k}} \left[\frac{A \, (T - T_c) + c \, \underline{k}^2}{2 \, \pi \, k_B \, T} \right]^{-\frac{1}{2}} \tag{1461}
\end{aligned}
$$

where in the last line we have restored the product to run over the entire range of $\underline{k}$. In this expression, each of the n components of the order parameter yields an identical factor. Thus, since

$$
Z = \exp\left[-\beta \, F \right] \tag{1462}
$$

273

the Gaussian approximation to the Free-Energy, for $T > T_c$, is given by the expression

$$F \approx F_0 V + \frac{k_B T n}{2} \sum_{\underline{k}} \ln \left[\frac{A (T - T_c) + c \underline{k}^2}{2 \pi k_B T} \right] \qquad (1463)$$

where the summation runs over the full range of $\underline{k}$. The specific heat can be obtained from the expression

$$C = - T \left(\frac{\partial^2 F}{\partial T^2} \right) \qquad (1464)$$

The most divergent term in C is recognized as

$$C \sim \frac{T^2 n k_B}{2} V \int \frac{d^d \underline{k}}{(2 \pi)^d} \frac{A^2}{(A (T - T_c) + c \underline{k}^2)^2} \qquad (1465)$$

As we shall see, this exhibits different types of behavior depending on whether $d > 4$ or $d < 4$. On setting

$$x = k \xi \qquad (1466)$$

where the correlation length ξ is defined by

$$\xi^{-2} = \frac{A (T - T_c)}{c} \qquad (1467)$$

one finds that the leading divergence of the specific heat is given by

$$C \sim \frac{n k_B A^2 T^2}{2 c^2} V \xi^{4-d} \int \frac{d^d \underline{x}}{(2 \pi)^d} \left(\frac{1}{1 + x^2} \right)^2 \qquad (1468)$$

The integral is convergent at large k for $4 > d$, in which case it is independent of the cut-off for k. For larger dimensionalities, d, the integral may exhibit an ultra-violet divergence if the cut-off is ignored. However, the upper cut-off for x does depend on the lattice spacing a and is given by $x_c \sim \pi \xi / a$, so the expression for the part of C displayed becomes independent of ξ when $d > 4$, leading to

$$C \sim \frac{n k_B A^2 T^2}{2 c^2} V a^{4-d} \qquad (1469)$$

Thus, for $d > 4$, we have found that the critical exponent is given by $\alpha = 0$. In general, the only divergences of interest are those which occur at $k \sim 0$ when $\xi \to \infty$. Divergences that occur due to the behavior of the integrand in the region $k \sim 0$ are known as infra-red divergences. For $4 > d$, the integral in eqn(1468) is convergent when ξ is finite, and one obtains

$$\begin{aligned} C & \sim \xi^{4-d} \\ & \sim (T - T_c)^{-\nu(4-d)} \end{aligned} \qquad (1470)$$

274

This expression reflects the infra-red divergence which occurs at $T = T_c$. Thus, in the Gaussian approximation and with $4 > d$, the critical exponent α has been calculated as

$$\alpha = \nu \, (4 - d) \tag{1471}$$

which differs from the value $\alpha = 0$ found from the discontinuous specific heat as calculated in the mean-field approximation. For $d > 4$, the Gaussian approximation was found to yield the specific heat exponent $\alpha = 0$, just like the value of α found in the mean-field approximation. The exponents of the Gaussian and the mean-field approximation first coincide when $d = 4$. The difference in the α the exponents found for $d < 4$ indicates that the fluctuations of the order parameter need to be accounted in dimensions less than four.

The Gaussian Approximation for $T_c > T$

For temperatures below T_c, the mean-field order parameter is non-zero and will be parallel to any uniform applied field, Π_L, no matter how small. We shall orient our coordinate system so that the applied field lies along one axis, the longitudinal axis. The longitudinal component of the order parameter, $\phi_L(\underline{r})$, will be written as the sum of the mean-field order parameter ϕ_L and the spatially varying fluctuations

$$\phi_L(\underline{r}) = \phi_L + \delta\phi_L(\underline{r}) \tag{1472}$$

The remaining $(n-1)$ components are transverse components which represent truly spatially varying fluctuations, i.e. they have no uniform components. The transverse components will be denoted as

$$\delta\phi_j(\underline{r}) \tag{1473}$$

for $j = 1, 2, \ldots, n$. On substituting these expressions into the Ginzberg-Landau Functional, one obtains

$$
\begin{aligned}
F[\phi] = \; & V \left[F_0 + F_2 \, \phi_L^2 + F_4 \, \phi_L^4 - \Pi_L \, \phi_L \right] \\
& + \int d^d\underline{r} \left[F_2 \, \delta\phi_L^2(\underline{r}) + 6 \, F_4 \, \phi_L^2 \, \delta\phi_L^2(\underline{r}) + c \, (\underline{\nabla}\delta\phi_L)^2 \right. \\
& + F_2 \sum_{j=1}^{n-1} \delta\phi_j^2(\underline{r}) + 2 \, F_4 \, \phi_L^2 \sum_{j=1}^{n-1} \delta\phi_j^2(\underline{r}) + c \sum_{j=1}^{n-1} (\underline{\nabla}\delta\phi_j)^2 \\
& + 4 \, F_4 \, \phi_L \, \delta\phi_L^3(\underline{r}) + 4 \, F_4 \, \phi_L \, \delta\phi_L(\underline{r}) \sum_{j=1}^{n-1} \delta\phi_j^2(\underline{r}) \\
& \left. + F_4 \, \delta\phi_L^4(\underline{r}) + 2 \, F_4 \, \delta\phi_L^2(\underline{r}) \sum_{j=1}^{n-1} \delta\phi_j^2(\underline{r}) + F_4 \sum_{i,j}^{n-1} \delta\phi_i^2(\underline{r}) \, \delta\phi_j^2(\underline{r}) \right]
\end{aligned}
\tag{1474}
$$

The first line represents the Landau-Ginzberg Free-Energy for a uniform longitudinal order parameter. The Gaussian approximation consists of minimizing the first line, as in mean-field theory, and retains the terms in the second and third lines as they are of quadratic order in the fluctuations. The terms in the last two lines are neglected, since they are of cubic and quartic order in the fluctuations. The fluctuating parts of the fields are expressed in terms of their Fourier components

$$\delta\phi_L(\underline{r}) \;=\; \frac{1}{\sqrt{V}} \; \sum_{\underline{k}} \; \exp\left[\; + \; i \; \underline{k} \cdot \underline{r} \; \right] \; \phi_{L,\underline{k}} \qquad (1475)$$

and

$$\delta\phi_j(\underline{r}) \;=\; \frac{1}{\sqrt{V}} \; \sum_{\underline{k}} \; \exp\left[\; + \; i \; \underline{k} \cdot \underline{r} \; \right] \; \phi_{j,\underline{k}} \qquad (1476)$$

On substituting into the Gaussian approximation for the Free-Energy Functional, one obtains

$$
\begin{aligned}
F[\phi] \;\approx\; V &\left[\; F_0 \;+\; F_2 \; \phi_L^2 \;+\; F_4 \; \phi_L^4 \;-\; \Pi_L \; \phi_L \; \right] \\
&+\; \sum_{\underline{k}} \left[\; F_2 \;+\; 6 \; F_4 \; \phi_L^2 \;+\; c \; \underline{k}^2 \; \right] \phi_{L,\underline{k}} \; \phi_{L,-\underline{k}} \\
&+\; \sum_{j=1}^{n-1} \sum_{\underline{k}} \left[\; F_2 \;+\; 2 \; F_4 \; \phi_L^2 \;+\; c \; \underline{k}^2 \; \right] \phi_{j,\underline{k}} \; \phi_{j,-\underline{k}}
\end{aligned}
$$

$$(1477)$$

Since ϕ_L minimizes the first term in the approximate Free-Energy Functional it satisfies

$$\left[\; 2 \; F_2 \;+\; 4 \; F_4 \; \phi_L^2 \; \right] \phi_L \;=\; \Pi_L \qquad (1478)$$

or

$$F_2 \;+\; 2 \; F_4 \; \phi_L^2 \;=\; \frac{\Pi_L}{2 \; \phi_L} \qquad (1479)$$

On utilizing the expression for ϕ_L, one can express the approximate Free-Energy Functional as

$$
\begin{aligned}
F[\phi] \;\approx\; V &\left[\; F_0 \;-\; F_4 \; \phi_L^4 \;-\; \frac{\Pi_L \; \phi_L}{2} \; \right] \\
&+\; \sum_{\underline{k}} \left[\; 4 \; F_4 \; \phi_L^2 \;+\; \frac{\Pi_L}{2 \; \phi_L} \;+\; c \; \underline{k}^2 \; \right] \phi_{L,\underline{k}} \; \phi_{L,-\underline{k}} \\
&+\; \sum_{j=1}^{n-1} \sum_{\underline{k}} \left[\; \frac{\Pi_L}{2 \; \phi_L} \;+\; c \; \underline{k}^2 \; \right] \phi_{j,\underline{k}} \; \phi_{j,-\underline{k}} \qquad (1480)
\end{aligned}
$$

The longitudinal and transverse fluctuations behave differently. It is seen that the longitudinal fluctuations are primarily stabilized by the non-zero order parameter, whereas only the applied field stabilizes the transverse fluctuations. The Gaussian path integral can be evaluated leading to the Gaussian approximation to the Free-Energy

$$
F \;\approx\; F[\phi_L] \;+\; \frac{k_B\,T}{2} \sum_{\underline{k}}\; \ln\left[\frac{4\,F_4\,\phi_L^2 \;+\; \frac{\Pi_L}{2\,\phi_L} \;+\; c\,\underline{k}^2}{2\,\pi\,k_B\,T}\right]
$$

$$
+\; (\,n \;-\; 1\,)\,\frac{k_B\,T}{2} \sum_{\underline{k}}\; \ln\left[\frac{\frac{\Pi_L}{2\,\phi_L} \;+\; c\,\underline{k}^2}{2\,\pi\,k_B\,T}\right] \tag{1481}
$$

The Free-Energy can be used to calculate the divergent part of the specific heat and its critical exponent α. The specific heat is given by the sum of the contributions the mean-field theory and the longitudinal Gaussian fluctuations. Note that the amplitude of the singular part of the Free-Energy is different above and below T_c, and this leads to an extra factor of n in the specific heat of the high temperature phase, whereas at low temperatures, there is an extra factor of $2^{\frac{(d-4)}{2}}$.

The Ginzberg Criterion

The Ginzberg Criterion provides an estimate of the temperature range in which the results of mean-field theory may be reasonable. Mean-field theory (or the Gaussian approximation) may be considered reasonable whenever the fluctuations in the order parameter are smaller than the average value of the order parameter. The size of the mean-squared fluctuations can be estimated by $S(r)$ evaluated at a length scale given by the correlation length ξ. Hence, the results of mean-field theory may be reasonable when

$$
1 \;>\; \frac{S(\xi)}{\phi_0^2} \tag{1482}
$$

or, equivalently

$$
\phi_0^2 \;>\; \left(\frac{k_B\,T}{8\,\pi\,c}\right)\frac{1}{\xi^{d-2}} \tag{1483}
$$

which leads to

$$
\frac{A\,(\,T_c \;-\; T\,)}{2\,F_4} \;>\; \left(\frac{k_B\,T}{8\,\pi\,c}\right)\left(\frac{2\,A\,(\,T_c \;-\; T\,)}{c}\right)^{\frac{d-2}{2}} \tag{1484}
$$

or

$$
\left(\frac{2\,\pi\,c^2}{k_B\,T\,F_4}\right) \;>\; \left(\frac{2\,A\,(\,T_c \;-\; T\,)}{c}\right)^{\frac{d-4}{2}} \tag{1485}
$$

This suggests that, generally, mean-field theory might be reasonable for temperatures outside the critical region which is a narrow temperature window around

T_c. The fluctuations dominate in the critical region. The Ginzberg criterion also indicates that mean-field theory, or the Gaussian approximation, might also be reasonable for all temperatures in four or higher dimensions. The upper critical dimension d_c^u is the dimension above which the critical point can be treated in the Gaussian approximation, and for an ordinary second-order transition $d_c^u = 4$.

There is also a lower critical dimensionality d_c^l. Mermin and Wagner have shown that a phase with spontaneously broken continuous symmetry is unstable for dimensions less than two, since longwave length transverse fluctuations of the order parameter are divergent. In this case, the lower critical dimensionality d_c^l, below which a phase transition cannot occur, is $d_c^l = 2$. The divergence of the fluctuations for $2 > d$ found in systems with a continuously broken symmetry is related to the presence of Goldstone modes. Due to the divergence of the fluctuations, the average value of the order parameter is not well-defined and, therefore, the fluctuations dynamically restore the broken symmetry. The suppression of ordering can be seen in a different way, by examining how T_c is reduced in the self-consistent Gaussian approximation.

The Self-consistent Gaussian Approximation.

The self-consistent Gaussian approximation starts from the approximate Free-Energy of the Gaussian model in the form

$$F[\phi] \;\approx\; F[\phi_L] \;+\; \frac{k_B\,T}{2}\,\sum_{\underline{k}}\,\ln\left[\frac{F_2 \;+\; 6\,F_4\,\phi_L^2 \;+\; c\,\underline{k}^2}{2\,\pi\,k_B\,T}\right]$$

$$+\,(\,n\,-\,1\,)\,\frac{k_B\,T}{2}\,\sum_{\underline{k}}\,\ln\left[\frac{F_2 \;+\; 2\,F_4\,\phi_L^2 \;+\; c\,\underline{k}^2}{2\,\pi\,k_B\,T}\right] \quad (1486)$$

This expression holds true for both T greater and T smaller than T_c. For temperatures above T_c, one expects ϕ_L will be zero and the two logarithmic terms can be combined since there is no physical distinction between the longitudinal and transverse directions, if $\Pi_L \;=\; 0$. Minimiziation w.r.t. ϕ_L leads to the solutions of either

$$\phi_L \;=\; 0 \quad (1487)$$

or

$$0 \;=\; \left[\,F_2 \;+\; 2\,F_4\,\phi_L^2\,\right] \;+\; \frac{k_B\,T}{2}\,\sum_{\underline{k}}\,\frac{6\,F_4}{F_2 \;+\; 6\,F_4\,\phi_L^2 \;+\; c\,\underline{k}^2}$$

$$+\,(\,n\,-\,1\,)\,\frac{k_B\,T}{2}\,\sum_{\underline{k}}\,\frac{2\,F_4}{F_2 \;+\; 2\,F_4\,\phi_L^2 \;+\; c\,\underline{k}^2}$$

$$=\; \left[\,F_2 \;+\; 2\,F_4\,\phi_L^2\,\right] \;+\; 6\,F_4\,\sum_{\underline{k}}\,<\,\phi_{L,\underline{k}}\,\phi_{L,-\underline{k}}\,>$$

$$+\,(\,n\,-\,1\,)\,2\,F_4\,\sum_{\underline{k}}\,<\,\phi_{T,\underline{k}}\,\phi_{T,-\underline{k}}\,> \quad (1488)$$

where the last two terms have been recognized as involving the fluctuations of the order parameter, as evaluated in the Gaussian approximation. The critical temperature T_c is the temperature at which two infinitesimal but real solutions for ϕ_L first occur. This is to be contrasted with the approximate critical temperature , $T_c^{(0)}$, defined by

$$F_2 \ = \ A \ (\ T \ - \ T_c^{(0)} \ ) \tag{1489}$$

The true critical temperature T_c is determined from the equation

$$0 \ = \ F_2 \ + \ (\ n \ + \ 2 \ ) \ 2 \ F_4 \ \left(\frac{k_B \ T_c}{2} \right) \ \sum_{\underline{k}} \ \frac{1}{F_2 \ + \ c \ \underline{k}^2} \tag{1490}$$

which since the last term is positive, reduces T_c below $T_c^{(0)}$. At $T_c^{(0)}$, the last term can be expressed in terms of an integral

$$(\ n \ + \ 2 \ ) \ 2 \ F_4 \ \left(\frac{k_B \ T_c^{(0)}}{2 \ c} \right) \ V \ \frac{S_d}{(\ 2 \ \pi \ )^d} \ \int dk \ k^{d-3} \tag{1491}$$

For $d \ > \ 3$, the integral is finite and of the order $\frac{a^{2-d}}{(d-2)}$, hence, one expects that the shift of T_c will be reasonably moderate. On the other-hand, for $3 \ > \ d$ the integral representing the order parameter fluctuations is divergent due to the behavior at $k \ \sim \ 0$, thereby suppressing T_c to much lower temperatures. The logarithmic divergence of the correction to T_c that occurs for $d \ = \ 2$ is consistent with the value of the lower critical dimensionality $d_c^l \ = \ 2$ that is inferred from the Mermin-Wagner theorem.

10.7   The Renormalization Group Technique

The scaling behavior shows that there exists a single relevant length scale that describes the large scale, long-ranged, fluctuations that dominate the singular parts of the Free-Energy. The scaling theory and the formulation of the Landau-Ginzberg Free-Energy Functional indicate that the microscopic length scales in the Hamiltonian are irrelevant. The scaling hypothesis describes the change in the fluctuations as this length scale is changed by, for example, changing the temperature. Furthermore, at the critical temperature, the system appears to exhibit the same behavior at all length scales. The renormalization group technique supplements the scaling hypothesis by incorporating the effect of the short scale physics. It shows that if the length scale is changed, then the effective interactions controlling the large scale fluctuations also change. The interactions between the long-ranged fluctuations are re-scaled, when the short-ranged fluctuations are removed by integrating them out. The method involves the following three steps:

(i) Integrating out the short scale fluctuations of the system, thereby increasing the effective short distance cut-off for the system.

(ii) Re-defining all length scales, so that the new cut-off appears indistinguishable from the old cut-off.

(iii) Re-define or renormalize the interactions governing the fluctuations of the order parameter.

The above procedure introduces the idea of an operation that can be compounded, resulting a semi-group rather than a group, since the operations are not uniquely invertible. The operations result in a flow in both the form and the parameters involved in the Landau-Ginzberg Functional as the length scale is changed by successive infinitesimal increments. The set of parameters $\{F_2, F_4, \ldots, c\}$ that describe the most general form of the Landau-Ginzberg Free-Energy Functional describe a point in parameter space. A change in scale by a factor of λ results in a flow between two different points in parameter space

$$\{F_2', F_4', \ldots, c'\} \;=\; \mathcal{R}(\lambda)\{F_2, F_4, \ldots, c\} \tag{1492}$$

At the critical point, the above operations should leave the renormalized Landau-Ginzberg Functional invariant, reflecting the scale-invariance that occurs at the critical point,

$$\{F_2^*, F_4^*, \ldots, c^*\} \;=\; \mathcal{R}(\lambda)\{F_2^*, F_4^*, \ldots, c^*\} \tag{1493}$$

The corresponding invariant point of parameter space $\{F_2^*, F_4^*, \ldots, c^*\}$ is known as a fixed point. Sometimes the properties of a system which is close to a fixed point can be inferred from the flow of the parameters under the renormalization group operations by linearizing the flow around the fixed point. In this case, the procedure results in the recovery of the phenomena described by the scaling hypothesis together with the actual values of the critical exponents.

10.8   Collective Modes and Symmetry Breaking

Goldstone's theorem[50] pertains to phase transition where the Hamiltonian has a continuous symmetry that is spontaneously broken. That is, the ground state is infinitely degenerate and does not have the full symmetry of the Hamiltonian. The theorem states that in the broken symmetry state, the system will have branch of collective boson excitations that has a dispersion relation which reaches $\omega = 0$ at $k = 0$.

Furthermore, in the long wavelength limit, the bosons are non-interacting and a coherent superpositions of these bosons connects the broken symmetry state to the continuum of degenerate states. This is easily understood, since if

---

[50] J. Goldstone, "Field Theories with Superconductor Solutions". Nuovo Cimento, 19, 154-164, (1961).

the system was to be physically transformed from one broken symmetry state to another, no energy would have to be supplied to the system. Thus, the bosons dynamically restore the broken symmetry.

Goldstone bosons in the form of spin waves were already known to exist in ferromagnets and antiferromagnets[51], where the continuous spin rotational symmetry is spontaneously broken at low temperatures. For ferromagnets, the ground state and the spin-wave dispersion relations can be calculated exactly. Ironically, P.W. Anderson had already investigated the dynamic modes associated with a superconductor[52] prior to Goldstone's work. Anderson had found, contrary to the Goldstone theorem, that the bosons in a superconductor had a finite excitation energy similar to the plasmon energy of the metal. A posteriori, this is obvious since metals neither become transparent nor change colour when they start to superconduct. Anderson's idea was subsequently picked up by Peter Higgs[53] and by Tom Kibble and co-workers[54] and also by François Englert and Robert Brout[55] who noted that, if long-ranged interactions were present, the modes would acquire a mass. The massive modes, associated with the breaking of a continuous symmetry in the presence of long-ranged interactions, are known as Kibble-Higgs modes.

Here we shall examine the Goldstone bosons of a Heisenberg ferromagnet, which is a slightly unusual case since the order parameter of a ferromagnet is a conserved quantity.

The Ferromagnetic State

The fully polarized ferromagnetic state $\mid \Phi_0 >$ has all the spins aligned and is an exact eigenstate of the Heisenberg Hamiltonian. The Hamiltonian can be written as a scalar product

$$
\begin{aligned}
\hat{H} &= -\sum_{i,j} J_{i,j} \, \underline{\hat{S}}_i \cdot \underline{\hat{S}}_j \\
&= -\sum_{i,j} J_{i,j} \left[\hat{S}_i^z \, \hat{S}_j^z + \frac{1}{2} \left(\hat{S}_i^+ \, \hat{S}_j^- + \hat{S}_i^- \, \hat{S}_j^+ \right) \right]
\end{aligned}
\qquad (1494)
$$

where the sum runs over pairs of sites. We shall assume that the spontaneous

[51]P.W. Anderson, "An Approximate Quantum Theory of the Antiferromagnetic Ground State", Phys. Rev. 86, 694-701, (1952).

[52]P.W. Anderson, "Random-Phase Approximation in the Theory of Superconductivity", Phys. Rev. 112, 1900-1916, (1958), "Plasmons, Gauge Invariance, and Mass", Phys. Rev. 130, 439 (1963).

[53]P. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", Phys. Rev. Lett. 13. 508-509, (1964).

[54]G. Guralnik, C.R. Hagen and T.W.B. Kibble, "Global Conservation Laws and Massless Particles", Phys. Rev. Lett. 13, 585-587 (1964).

[55]F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", Phys. Rev. Lett. 13, 321-23 (1964).

magnetization is parallel to the z-axis. Then

$$
\begin{aligned}
\hat{H} \mid \Phi_0 > & = -\sum_{i,j} J_{i,j} \left[\hat{S}_i^z \hat{S}_j^z + \frac{1}{2} (\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+) \right] \mid \Phi_0 > \\
& = -\sum_{i,j} J_{i,j} \left[S\, S + \frac{1}{2} (\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+) \right] \mid \Phi_0 > \\
& = -\sum_{i,j} J_{i,j}\, S^2 \mid \Phi_0 > \qquad\qquad (1495)
\end{aligned}
$$

The first line follows since all the spins are aligned with the z-axis and are eigenstates of $\hat{S}_i^z$ with eigenvalue S

$$
\hat{S}_i^z \mid S_i > \; = \; S \mid S_i > \qquad\qquad (1496)
$$

The second line occurs since the spin-flip terms vanish as they all involve the spin-raising operator at a site and

$$
\hat{S}_i^+ \mid S_i > \; = \; 0 \qquad\qquad (1497)
$$

as the spin cannot be raised further. Hence, the fully-polarized ferromagnetic state is an exact eigenstate with eigenvalue $E_0 = -\sum_{i,j} J_{i,j}\, S^2$. This degenerate is infinitely degenerate since any rotation of the total magnetization leads to an equivalent symmetry broken ground state.

Conservation of Magnetization

The z-component of the total magnetization M is defined as

$$
\hat{M}^z = \sum_{i=1}^{N} \hat{S}_i^z \qquad\qquad (1498)
$$

The magnetization commutes with the Hamiltonian and so is conserved. The commutator

$$
[\, \hat{M}^z \,,\, \hat{H} \,] \qquad\qquad (1499)
$$

can be evaluated with the aid of the commutation relations

$$
\begin{aligned}
[\, \hat{S}_i^z \,,\, \hat{S}_j^+ \,] & = + \delta_{i,j}\, \hbar\, \hat{S}_i^+ \\
[\, \hat{S}_i^z \,,\, \hat{S}_j^- \,] & = - \delta_{i,j}\, \hbar\, \hat{S}_i^- \qquad\qquad (1500)
\end{aligned}
$$

From which one finds that

$$
\begin{aligned}
[\, \hat{M}^z \,,\, \hat{H} \,] & = \frac{\hbar}{2} \sum_{i,j} J_{i,j} (\hat{S}_i^+ \hat{S}_j^- - \hat{S}_i^- \hat{S}_j^+) \\
& \quad - \frac{\hbar}{2} \sum_{i,j} J_{i,j} (\hat{S}_i^+ \hat{S}_j^- - \hat{S}_i^- \hat{S}_j^+) \\
& = 0 \qquad\qquad (1501)
\end{aligned}
$$

independent of any choice for the exchange interaction. Hence, the total magnetization is conserved.

The Spin Wave Dispersion Relation

The spin wave state $| \Phi_{\underline{q}} >$ is a linear superpositions of ferromagnetic states with a single flipped spin. The spin wave state can be expressed as

$$
\begin{aligned}
| \Phi_{\underline{q}} > &= \hat{S}_{\underline{q}}^{-} | \Phi_0 > \\
&= \frac{1}{\sqrt{N}} \sum_j \exp \left[i \, \underline{q} \cdot \underline{R}_j \right] S_j^{-} | \Phi_0 >
\end{aligned}
\tag{1502}
$$

which satisfies the energy-eigenvalue equation

$$
\hat{H} \, | \Phi_{\underline{q}} > = E_{\underline{q}} \, | \Phi_{\underline{q}} >
\tag{1503}
$$

where the energy eigenvalue can be expressed in terms of the ground state energy E_0 and the spin wave excitation energy $\hbar \, \omega_{\underline{q}}$

$$
E_{\underline{q}} = E_0 + \hbar \, \omega_{\underline{q}}
\tag{1504}
$$

The dispersion relation can be found from

$$
[\, \hat{H} \, , \, \hat{S}_{\underline{q}}^{-} \,] \, | \Phi_0 > = \hbar \, \omega_{\underline{q}} \, \hat{S}_{\underline{q}}^{-} \, | \Phi_0 >
\tag{1505}
$$

by using the commutation relations

$$
[\, \hat{S}_j^z \, , \, \hat{S}_i^{-} \,] = - \, \delta_{i,j} \, \hbar \, \hat{S}_i^{-}
\tag{1506}
$$

and

$$
[\, \hat{S}_j^{+} \, , \, \hat{S}_i^{-} \,] = 2 \, \delta_{i,j} \, \hbar \, \hat{S}_i^z
\tag{1507}
$$

On using these two commutation relations, one finds that the commutation relation between $S_{\underline{q}}^{-}$ and the Hamiltonian produces

$$
\begin{aligned}
[\, \hat{H} \, , \, \hat{S}_{\underline{q}}^{-} \,] = & \frac{1}{\sqrt{N}} \sum_{i,j} J_{i,j} \, \hbar \left[\hat{S}_i^z \, \hat{S}_j^{-} - \hat{S}_i^{-} \, \hat{S}_j^z \right] \exp \left[i \, \underline{q} \cdot \underline{R}_j \right] \\
& + \frac{1}{\sqrt{N}} \sum_{i,j} J_{i,j} \, \hbar \left[\hat{S}_i^{-} \, \hat{S}_j^z - \hat{S}_i^z \, \hat{S}_j^{-} \right] \exp \left[i \, \underline{q} \cdot \underline{R}_i \right]
\end{aligned}
\tag{1508}
$$

Thus, when acting on the ferromagnetic state the commutation relation reduces to

$$
\begin{aligned}
[\, \hat{H} \, , \, \hat{S}_{\underline{q}}^{-} \,] \, | \Phi_0 > = & \frac{1}{\sqrt{N}} \sum_{i,j} J_{i,j} \, \hbar \, S \left[\hat{S}_j^{-} - \hat{S}_i^{-} \right] \exp \left[i \, \underline{q} \cdot \underline{R}_j \right] | \Phi_0 > \\
& + \frac{1}{\sqrt{N}} \sum_{i,j} J_{i,j} \, \hbar \, S \left[\hat{S}_i^{-} - \hat{S}_j^{-} \right] \exp \left[i \, \underline{q} \cdot \underline{R}_i \right] | \Phi_0 >
\end{aligned}
\tag{1509}
$$

which can be further reduced by noting that since the pairwise interaction only depends on the nearest neighbor separation and not the absolute location in the lattice. Therefore, on expressing this in terms of $\hat{S}_{\underline{a}}^{-}$ and a sum over the nearest neighbors sites, one has

$$[\ \hat{H}\ ,\ \hat{S}_{\underline{q}}^{-}\ ]\ |\ \Phi_0\ >\ =\ \frac{1}{\sqrt{N}}\ \sum_i\ J_{i,0}\ \hbar\ S\ \left(\ 1\ -\ \cos[\ \underline{q}\ .\ \underline{R}_{i,0}\ ]\ \right)\ \hat{S}_{\underline{q}}^{-}\ |\ \Phi_0\ > \tag{1510}$$

Thus, the dispersion relation is evaluated as a sum over nearest neighbor sites

$$\hbar\ \omega_{\underline{q}}\ =\ \sum_i\ J_{0,i}\ \hbar\ S\ \left(\ 1\ -\ \cos[\ \underline{q}\ .\ \underline{R}_{0,i}\ ]\ \right) \tag{1511}$$

which vanishes quadratically as $q \to 0$. Usually, Goldstone modes have linear dispersion relations, however, due to the conserved nature of the order parameter, ferromagnetic spin waves have quadratic dispersion relations.

10.9   Appendix: The One-Dimensional Ising Model

The Transfer Matrix Solution of the One-Dimensional Ising Model in an Applied Field

We shall apply periodic boundary conditions to a one-dimensional chain of length N, so that $S_{N+1}^z\ =\ S_1^z$, so the system of spins becomes equivalent to a ring of spins with N links. The partition function can be written as the trace over N factors each of which depends on the values of two neighboring spins

$$Z\ =\ \text{Trace}\ \left[\ \prod_{i=1}^N\ \left\{\ T(S_i^z, S_{i+1}^z)\ \right\}\ \right] \tag{1512}$$

where the factors of K have been symmetrized

$$T(S_i^z, S_{i+1}^z)\ =\ \exp\left[\ \beta\ J\ S_i^z\ S_{i+1}^z\ +\ \beta\ \frac{B}{2}\ (\ S_i^z\ +\ S_{i+1}^z\ )\ \right] \tag{1513}$$

The factors $T(S_i^z, S_{i+1}^z)$ can be regarded as elements of a matrix $\hat{T}$ (the transfer matrix)

$$T(S_i^z, S_{i+1}^z)\ =\ <\ S_i^z\ |\ \hat{T}\ |\ S_{i+1}^z\ > \tag{1514}$$

where $|\ S_i^z\ >$ is a column vector

$$|\ S_{i+1}^z\ >\ =\ \begin{pmatrix} \left(\frac{1+S_{i+1}^z}{2}\right) \\[2ex] \left(\frac{1-S_{i+1}^z}{2}\right) \end{pmatrix} \tag{1515}$$

and $<\,S_i^z\,|$ is a row vector

$$<\,S_i^z\,|\ =\ \left(\ \left(\frac{1+S_i^z}{2}\right)\qquad \left(\frac{1-S_i^z}{2}\right)\ \right)\tag{1516}$$

Thus, the matrix $\hat{T}$ is given by

$$\hat{T}\ =\ \left(\begin{array}{cc}\exp\left[\ +\ \beta\ (\ J\ +\ B\ )\right] & \exp\left[\ -\ \beta\ J\right] \\ \exp\left[\ -\ \beta\ J\right] & \exp\left[\ +\ \beta\ (\ J\ -\ B\ )\right]\end{array}\right)\tag{1517}$$

Since

$$\sum_{S_{i+1}^z}|\ S_{i+1}^z\ ><\ S_{i+1}^z\ |\ =\ \hat{I}\tag{1518}$$

where $\hat{I}$ is the unit 2×2 matrix

$$\hat{I}\ =\ \left(\begin{array}{cc}1 & 0 \\ 0 & 1\end{array}\right)\tag{1519}$$

the summation over S_{i+1}^z in the expression can be performed as

$$\sum_{S_{i+1}^z}<\ S_i^z\ |\ \hat{T}\ |\ S_{i+1}^z\ ><\ S_{i+1}^z\ |\ \hat{T}\ |\ S_{i+2}^z\ >\ =\ <\ S_i^z\ |\ \hat{T}\ \hat{T}\ |\ S_{i+2}^z\ >$$
$$=\ <\ S_i^z\ |\ \hat{T}^2\ |\ S_{i+2}^z\ >\tag{1520}$$

Using the completeness property iteratively, the successive traces over the variables S_i^z for $i\ =\ 2\ ,\ \dots\ ,N$, in the expression for Z can be replaced by successive multiplication of the matrices $\hat{T}$. Thus

$$Z\ =\ \sum_{S_1^z=\pm1}<\ S_1^z\ |\ \hat{T}^N\ |\ S_1^z\ >\tag{1521}$$

The partition function can be expressed in terms of the eigenvalues λ_i of $\hat{T}$ defined by the eigenvalue equation

$$\hat{T}\ |\ \chi_i\ >\ =\ \lambda_i\ |\ \chi_i\ >\tag{1522}$$

for $i=1$ or $i=2$. We shall assume that $\lambda_1\ >\ \lambda_2$. On defining a 2×2 matrix $\hat{S}$ as a row vector of the two column vectors $|\ \chi_1\ >$ and $|\ \chi_2\ >$

$$\hat{S}\ =\ (\ |\ \chi_1\ >\qquad |\ \chi_2\ >\ )\tag{1523}$$

then, as the similarity transform based on $\hat{S}$ diagonalizes $\hat{T}$, the inverse transformation is given by

$$\hat{T}\ =\ \hat{S}\ \left(\begin{array}{cc}\lambda_1 & 0 \\ 0 & \lambda_2\end{array}\right)\ \hat{S}^{-1}\tag{1524}$$

Then, the partition function can be evaluated as

$$
Z = \sum_{S_1^z = \pm 1} < S_1^z \mid \hat{S} \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} \hat{S}^{-1} \mid S_1^z > \tag{1525}
$$

and on utilizing the cyclic invariance of the trace, one finds the result

$$
\begin{aligned}
Z &= \sum_{S_1^z = \pm 1} < S_1^z \mid \begin{pmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{pmatrix} \mid S_1^z > \\
&= \lambda_1^N + \lambda_2^N
\end{aligned} \tag{1526}
$$

The Free-Energy F is given by

$$
\begin{aligned}
F &= - k_B \, T \, \ln \left(\lambda_1^N + \lambda_2^N \right) \\
&= - k_B \, T \, N \, \ln \lambda_1 - k_B \, T \, \ln \left(1 + \frac{\lambda_2^N}{\lambda_1^N} \right)
\end{aligned} \tag{1527}
$$

The second term can be neglected since it is of order unity. Thus,

$$
F \approx - k_B \, T \, N \, \ln \lambda_1 \tag{1528}
$$

The eigenvalues are determined from the secular equation

$$
\begin{vmatrix} \exp\left[+ \beta \, (\, J \, + \, B \,) \right] - \lambda & \exp\left[- \beta \, J \right] \\ \exp\left[- \beta \, J \right] & \exp\left[+ \beta \, (\, J \, - \, B \,) \right] - \lambda \end{vmatrix} = 0 \tag{1529}
$$

which can be expressed as the quadratic equation

$$
\lambda^2 - 2 \, \lambda \, \exp[\, \beta J \,] \, \cosh \beta B + 4 \, \cosh \beta J \, \sinh \beta J = 0 \tag{1530}
$$

This has the two solutions

$$
\lambda = \exp[\, \beta J \,] \, \cosh \beta B \pm \sqrt{\exp[\, + 2\beta J \,] \, \sinh^2 \beta B + \exp[\, - 2\beta J \,]} \tag{1531}
$$

Thus, the final result for the Free-Energy is given by

$$
F = - N \, k_B \, T \, \ln \left[\exp[\, \beta J \,] \, \cosh \beta B + \sqrt{\exp[\, + 2\beta J \,] \, \sinh^2 \beta B + \exp[\, - 2\beta J \,]} \, \right] \tag{1532}
$$

The magnetization is defined as

$$
M = - \left(\frac{\partial F}{\partial B} \right) \tag{1533}
$$

which yields

$$M \; = \; N \; \frac{\exp[\; \beta J \;] \; \sinh \beta B}{\sqrt{\exp[\; + 2\beta J \;] \; \sinh^2 \beta B \; + \; \exp[\; -2\beta J \;]}} \tag{1534}$$

Since the Free-Energy is an analytic function of B and T, and as $M(B,T)$ is also an analytic function of B for all T, the system does not exhibit a phase transition at any finite T.

Zeroes of the Partition Function

The zeroes of the partition function are determined as

$$Z_N(z) \; = \; \lambda_1^N \; + \; \lambda_2^N \; = \; 0 \tag{1535}$$

we shall express z as

$$z \; = \; \exp[-4\beta J + 2\beta B] \; = \; \exp[-4\beta J + i\theta] \tag{1536}$$

corresponding to $\beta B \; = \; i \, \frac{\theta}{2}$. On defining a parameter r via

$$r \; = \; \exp[-2\beta J] \tag{1537}$$

then

$$z \; = \; r^2 \; \exp[\; i \, \theta \;] \tag{1538}$$

The values of θ for which $Z_N(z) = 0$ are determined from the equation

$$\left[\cos \frac{\theta}{2} + \left(r^2 - \sin^2 \frac{\theta}{2} \right)^{\frac{1}{2}} \right]^N + \left[\cos \frac{\theta}{2} - \left(r^2 - \sin^2 \frac{\theta}{2} \right)^{\frac{1}{2}} \right]^N = 0 \tag{1539}$$

On setting

$$\cos \frac{\theta}{2} \; = \; (\; 1 \; - \; r^2 \;) \; \cos \varphi \tag{1540}$$

the equation for θ can be simplified to

$$(\; 1 \; - \; r^2 \;)^{\frac{N}{2}} \left[(\; \cos \varphi \; + \; i \; \sin \varphi \;)^N \; + \; (\; \cos \varphi \; - \; i \; \sin \varphi \;)^N \right] \; = \; 0 \tag{1541}$$

or, equivalently

$$(\; 1 \; - \; r^2 \;)^{\frac{N}{2}} \; 2 \; \cos N\varphi \; = \; 0 \tag{1542}$$

which has solutions given by

$$\varphi \; = \; \pm \, (\; k \; - \; \frac{1}{2} \;) \; \frac{\pi}{N} \tag{1543}$$

for $k \; = \; 1 \, , \; 2 \, , \dots , \; N$. Since

$$\cos \theta \; = \; 2 \; \cos^2 \frac{\theta}{2} \; - \; 1 \tag{1544}$$

one has the solution

$$\begin{aligned} \cos\theta \;&=\; -\,r^2 \,+\, (\,1 \,-\, r^2\,)\;\cos 2\varphi \\ &=\; -\,r^2 \,+\, (\,1 \,-\, r^2\,)\;\cos\!\left(\frac{(2k-1)\pi}{N}\right) \end{aligned} \tag{1545}$$

for $k \,=\, 1\,,\,2\,,\dots\,,\,N$. This determines the distribution of zeroes of $Z_N(z)$. It is seen that they reside on a circle of radius r^2 in the complex z plane. The zeroes lie on an arc of the circle starting with the angle $\theta_i \,\approx\, \cos^{-1}(\,1 \,-\, 2\,r^2\,) \,=\, 2\,\sin^{-1}r$ and ending with an angle $\theta_f \,=\, 2\,\pi \,-\, \theta_i$. It has a gap of angular width $4\,\sin^{-1}r$ centered on the real z-axis. It only pinches of the real z-axis when $r \to 0$ or, equivalently, when $T \to 0$. Thus, the model can be considered as exhibiting a phase transition at $T = 0$. In this case, the critical exponents are identified as $\alpha \,=\, 1,\, \beta \,=\, 0,\, \gamma \,=\, 1,\, \nu \,=\, 1$ and $\delta \,=\, \infty$. The exponents satisfy the scaling relations. The existence of a phase transition in one-dimension is not inconsistent with the Mermin-Wagner theorem, since the broken symmetry is discrete.

The distribution of zeroes on the circle, $\rho(\theta)$, is defined as

$$\rho(\theta) \;=\; \frac{1}{N}\left(\frac{dk}{d\theta}\right) \tag{1546}$$

and is evaluated in the two regions as

$$\rho(\theta) \;=\; \begin{cases} \dfrac{1}{2\pi}\,\dfrac{\sin\frac{\theta}{2}}{\sqrt{\sin^2\frac{\theta}{2}-r^2}} & \text{if } \theta_f \,>\, \theta \,>\, \theta_i \\[2mm] \qquad 0 & \text{otherwise} \end{cases} \tag{1547}$$

respectively. The distribution is normalized to unity

$$\int_0^{2\pi} d\theta\;\rho(\theta) \;=\; 1 \tag{1548}$$