Chapter 1

The Methodology of Statistical Mechanics

This chapter provides an overview of statistical mechanics and thermodynamics. Although the discussion is reasonably self-contained, it is assumed that this is not the reader's first exposure to these subjects. More than many other branches of physics, these topics are treated in a variety of different ways in the textbooks in common usage. The aim of the chapter is thus to establish the perspective of the book: the standpoint from which the topics in the following chapters will be viewed. The books by Guénault [1] and by Bowley [2] are recommended to the novice as highly accessible and clearly-argued introductions to the subject; indeed all readers will find many of the examples of this book treated by those books in complementary fashion.

1.1 Terminology and Methodology

1.1.1 Approaches to the subject

Thermodynamics is the study of the relationship between *macroscopic* properties of systems such as temperature, volume, pressure, magnetization, compressibility etc. *Statistical Mechanics* is concerned with understanding how the various macroscopic properties arise as a consequence of the *microscopic* nature of the system. In essence it makes macroscopic deductions from microscopic models.

The power of thermodynamics, as formulated in the traditional manner (Zemansky [3], for example) is that its deductions are quite general; they do not rely, for their validity, on the microscopic nature of the system. Einstein expressed this quite impressively when he wrote [4]:

> *A theory is the more impressive the greater the simplicity of its premises is, the more different kinds of things it relates, and the more extended is its area of applicability. Therefore the deep impression which classical thermodynamics made upon me; it is the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, will never be overthrown.*

On the other hand statistical mechanics, as conventionally presented (Hill [5], for example) is system-specific. One starts from particular microscopic models, say the Debye model for a solid, and derives macroscopic properties such as the thermal capacity. It is true that statistical mechanics will give relationships between the various macroscopic properties of a system, but they will only apply to the system/model under consideration. Results obtained from thermodynamics, on the other hand, are model-independent and general.

Traditionally thermodynamics and statistical mechanics were developed as independent subjects, with “bridge equations” making the links between the two. Alternatively the subjects can be developed together, where the Laws of Thermodynamics are justified by microscopic arguments. This reductionist view was adopted by Landau. In justification he wrote [6]:

> *Statistical physics and thermodynamics together form a unit. All the concepts and quantities of thermodynamics follow most naturally, simply and rigorously from the concepts of statistical physics. Although the general statements of thermodynamics can be formulated non-statistically, their application to specific cases always requires the use of statistical physics.*

The contrast between the views of Einstein and those of Landau is apparent. The paradox, however, is that so much of Einstein’s work was reductionist and microscopic in nature whereas Landau was a master of the macroscopic description of phenomena. This book considers thermodynamics and statistical mechanics in a synthetic manner; in that respect it follows more closely the Landau approach.

1.1.2 Description of states

The state of a system described at the macroscopic level is called a *macrostate*. Macrostates are described by a relatively few variables such as temperature, pressure, volume etc.

The state of a system described at the microscopic level is called a *microstate*. Microstates are described by a very large number of variables.

Classically you would need to specify the position and momentum of each particle in the system. Using a quantum-mechanical description you would have to specify all the quantum numbers of the entire system. So whereas a macrostate is likely to be described by under ten variables, a microstate will be described by over 10^{23} variables.

The fundamental methodology of statistical mechanics involves applying probabilistic arguments about microstates, regarding macrostates as statistical averages. It is because of the very large number of particles involved that the mean behaviour corresponds so well to the observed behaviour – that fluctuations are negligible. Recall (or see Section 1.4.6) that for N particles the likely fractional deviation from the mean will be $1/\sqrt{N}$. So if you calculate the mean pressure of one mole of air at one atmosphere, it is likely to be correct to within $\sim 10^{-12}$ of an atmosphere. It is also because of the large number of particles involved that the mean, observed behaviour corresponds to the most probable behaviour – the *mean* and the *mode* of the distribution differ negligibly.

Statistical Mechanics is slightly unusual in that its formalism is easier to understand using quantum mechanics for descriptions at the microscopic level rather than classical mechanics. This is because the idea of a quantum state is familiar, and often the quantum states are discrete. It is more difficult to enumerate the (micro)states of a classical system; it is best done by analogy with the quantum case. The classical case will be taken up in Section 1.6.

1.1.3 Extensivity and the Thermodynamic Limit

The thermodynamic variables describing macrostates fall into two classes. Quantities such as energy E, number of particles N and volume V, which add when systems are combined, are known as *extensive* variables. And quantities such as pressure p and temperature T, which remain independent when similar systems are combined, are known as *intensive* variables.

As we have argued, thermodynamics concerns itself with the behaviour of "large" systems and usually, for these, finite-size effects are not of interest. Thus, for example, unless one is concerned specifically with surface phenomena, it will be sensible to focus on systems that are sufficiently large that the surface contribution to the energy will be negligible when compared to the volume contribution. This is possible because the energy of interaction between atoms is usually sufficiently short-ranged. In general one will consider properties in the limit $N \to \infty$, $V \to \infty$ while N/V remains constant. This is called the *thermodynamic limit.*

It should be clear that true extensivity of a quantity, such as energy, emerges only in the thermodynamic limit.

Gravity is a long-range force. It is apparent that gravitational energy is not truly extensive; this is examined in Problem 1.2. This non-extensivity makes for serious difficulties when trying to treat the statistical thermodynamics of gravitating systems.

1.2 The Fundamental Principles

1.2.1 The Laws of Thermodynamics

Thermodynamics, as a logical structure, is built on its four assumptions or laws (including the zeroth).

Zeroth Law: **If system A is in equilibrium with system B and with system C then system B is in equilibrium with system C**. Equilibrium here is understood in the sense that when two systems are brought into contact then there is no change. This law was formalized after the first three laws were formulated and numbered. Because it was believed to be more fundamental it was thus called the *Zeroth* Law. The Zeroth Law recognizes the existence of states of equilibrium and it points us to the concept of temperature, a non-mechanical quantity that can label (and order) equilibrium states.

First Law: **The internal energy of a body can change by the flow of heat or by doing work.**

$$\Delta E = \Delta Q + \Delta W. \tag{1.2.1}$$

Here ΔQ is the energy increase as a consequence of heat flow and ΔW is the energy increase resulting from work done. We usually regard this as a statement about the conservation of energy. But in its historical context the law asserted that as well as the familiar mechanical form of energy, heat also was a form of energy. Today we understand this as the kinetic energy of the constituent particles of a system; in earlier times the nature of heat was unclear.

Note that some older books adopt the *opposite sign* for ΔW; they consider ΔW to be the work done *by* the system rather than the work done *on* the system.

Second Law (this law has many formulations): **Heat flows from hot to cold**, or **It is not possible to convert *all* heat energy to work**. These statements have the great merit of being reflections of common experience. There are other formulations such as the Carathéodory statement (see the books by Adkins [7] or Pippard [8]): **In the neighbourhood of any equilibrium state of a thermally isolated system there are states which are inaccessible**, and the entropy statement (see Callen's book [9]): **There is an extensive quantity, which we call entropy, which never decreases in a physical process**. The claimed virtue of the Carathéodory statement is that it leads more rapidly to the important thermodynamic concepts of temperature and entropy: this at the expense of common experience. But if that is believed to be a virtue then one may as well go the "whole hog" and adopt the Callen statement. A major exercise in classical thermodynamics is proving the equivalence of the various statements of the Second Law. In whatever form, the Second Law leads to the concept of entropy and the *quantification* of temperature (the Zeroth Law just gives an *ordering*: A is hotter than B). And it tells us there is an absolute zero of temperature.

Third Law: **The entropy of a body tends to zero as the temperature tends to absolute zero**. The Third Law will be discussed in Section 1.7. We shall see that it arises as a consequence of the quantum behaviour of matter at the microscopic level. However we see immediately that the Third Law is telling us there is an absolute zero of *entropy*.

An even more fundamental aspect of the Zeroth Law is the fact of the *existence* of equilibrium states. If systems did not exist in states of equilibrium then there would be no macrostates and no hope of description in terms of small numbers of variables. Then there would be no discipline of thermodynamics and phenomena would have to be discussed solely in terms of their intractable microscopic description. Fortunately this is not the case; the existence of states of equilibrium allows our simple minds to make some sense of a complex world.

1.2.2 Probabilistic interpretation of the First Law

The First Law discusses the way the energy of a system can change. From the statistical standpoint we understand the energy of a macroscopic system as the *mean* value since the system can exist in a large number of different microstates. If the energy of the jth microstate is E_j and the probability of occurrence of this microstate is P_j then the (mean) energy of the system is given by

$$E = \sum_j P_j E_j. \tag{1.2.2}$$

The differential of this expression is

$$\mathrm{d}E = \sum_j P_j \, \mathrm{d}E_j + \sum_j E_j \, \mathrm{d}P_j; \tag{1.2.3}$$

this indicates the energy of the system can change in two different ways: a) the energy levels E_j may change or b) the probabilities P_j may change.

The first term $\sum_j P_j \, \mathrm{d}E_j$ relates to the change in the energy levels $\mathrm{d}E_j$. This term is the mean energy change of the microstates and we shall show below that this corresponds to the familiar "mechanical" energy: the work done on the system. The second term $\sum_j E_j \, \mathrm{d}P_j$ is a consequence of the change in probabilities or occupation of the energy states. This is fundamentally probabilistic and we shall see that it corresponds to the heat flow into the system.

In order to understand that the first term corresponds to the work done, let us consider a pV system. We shall see (Section 2.1.3) that the energy levels depend on the size (volume) of the system:

$$E_j = E_j(V) \tag{1.2.4}$$

so that the change in the energy levels when the volume changes is

$$\mathrm{d}E_j = \frac{\partial E_j}{\partial V} \mathrm{d}V. \tag{1.2.5}$$

Then

$$\sum_j P_j \mathrm{d}E_j = \sum_j P_j \frac{\partial E_j}{\partial V} \mathrm{d}V. \tag{1.2.6}$$

We are assuming that the change in volume occurs at constant P_j, then

$$\begin{aligned} \sum_j P_j \mathrm{d}E_j &= \sum_j \frac{\partial}{\partial V} P_j E_j \, \mathrm{d}V \\ &= \frac{\partial}{\partial V} E \, \mathrm{d}V. \end{aligned} \tag{1.2.7}$$

But we identify $\partial E / \partial V = -p$, so that

$$\sum_j P_j \mathrm{d}E_j = -p \, \mathrm{d}V. \tag{1.2.8}$$

And thus we see that the term $\sum_j P_j \mathrm{d}E_j$ corresponds to the work done on the system. Then the term $\sum_j E_j \mathrm{d}P_j$ corresponds to the energy increase of the system that occurs when no work is done; this is what we understand as heat flow.

We have learned, in this section, that the idea of heat arises quite logically from the probabilistic point of view.

1.2.3 Microscopic Basis for Entropy

By contrast to macroscopic thermodynamics, statistical mechanics is built on a single assumption, which we will call the *Fundamental Postulate* of statistical mechanics. We shall see how the Laws of Thermodynamics may be understood in terms of this Fundamental Postulate. These ideas date back to Boltzmann. The Fundamental Postulate states: **All microstates of an isolated system are equally likely**. Note, in particular, that an isolated system will have fixed energy E, volume V and number of particles N (fixed extensive quantities). Conventionally we denote by $\Omega(E, V, N)$ the number of microstates corresponding to a given macrostate (E, V, N). Then from the Fundamental Postulate it follows that the probability of a given macrostate is proportional to the number of microstates corresponding to it: $\Omega(E, V, N)$

$$P \propto \Omega(E, V, N). \tag{1.2.9}$$

If we understand the observed equilibrium state of a system as the most probable macrostate, then it follows from the Fundamental Postulate that the equilibrium state corresponds to the macrostate with the largest number of microstates. We are saying that Ω is maximum for an equilibrium state.

Since Ω for two isolated systems is multiplicative, it follows that the *logarithm* of Ω is additive. In other words $\ln \Omega$ is an *extensive* quantity.

Following Boltzmann we define *entropy* S as

$$S = k \ln \Omega. \tag{1.2.10}$$

At this stage k is simply a constant; later we will identify it as Boltzmann's constant. We should note that $\ln \Omega$ is dimensionless, so S will have the same dimensions as k.

Since the logarithm is a monotonic function it follows that the equilibrium state will have maximal entropy. So we immediately obtain the Second Law. And we now understand the Second Law from the microscopic point of view; it is hardly more than the tautology "we are most likely to observe the most probable state"!

1.3 Interactions – The Conditions for Equilibrium

When systems interact their states will often change, the composite system evolving to a state of equilibrium. We shall investigate what determines the final state. We will see that quantities such as temperature emerge in the characterization of these equilibrium states.

1.3.1 Thermal Interaction – Temperature

Let us allow two, otherwise isolated, systems to exchange energy without changing volume or numbers of particles. In other words we allow thermal interaction only; the systems are separated by a *diathermal* wall.

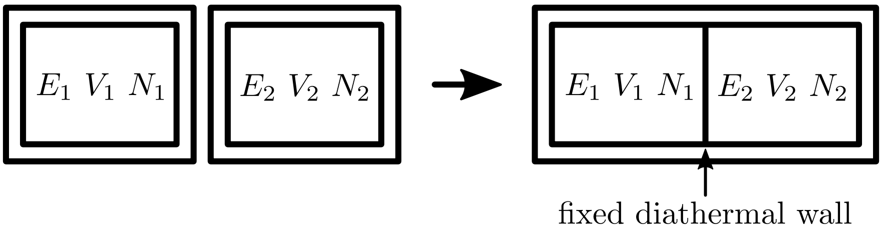


Figure 1.1: Thermal interaction

Now

$$\begin{aligned}\Omega_1 &= \Omega_1(E_1, V_1, N_1)\\ \Omega_2 &= \Omega_2(E_2, V_2, N_2)\end{aligned} \tag{1.3.1}$$

and V_1, N_1, V_2, N_2 are all fixed.

The energies E_1 and E_2 can vary subject to the constraint $E_1 + E_2 = E_0$, a constant.

Our problem is this: after the two systems are brought together what will be the equilibrium state? We know that the systems will exchange energy, and they will do this so as to maximize the total number of microstates for the composite system.

For different systems the Ωs multiply so that we have

$$\Omega = \Omega_1(E_1)\Omega_2(E_2) \tag{1.3.2}$$

— we can ignore V_1, N_1, V_2, N_2 as they don't change.

The systems will exchange energy so as to maximize Ω. Writing

$$\Omega = \Omega_1(E)\Omega_2(E_0 - E) \tag{1.3.3}$$

we allow the systems to vary E so that Ω is a maximum:

$$\frac{\partial\Omega}{\partial E} = \frac{\partial\Omega_1}{\partial E}\Omega_2 - \Omega_1\frac{\partial\Omega_2}{\partial E} = 0 \tag{1.3.4}$$

or

$$\frac{1}{\Omega_1}\frac{\partial\Omega_1}{\partial E} = \frac{1}{\Omega_2}\frac{\partial\Omega_2}{\partial E}. \tag{1.3.5}$$

In this form the left hand side is all about system 1 and the right hand side is all about system 2. In other words the equilibrium state is characterized by a certain property of system 1 being equal to a certain property of system 2.

We note that $\frac{1}{\Omega}\frac{\partial\Omega}{\partial E} = \frac{\partial \ln\Omega}{\partial E}$ so the equilibrium condition may be written

$$\frac{\partial \ln\Omega_1}{\partial E} = \frac{\partial \ln\Omega_2}{\partial E}. \tag{1.3.6}$$

Observe the natural occurrence of the *logarithm* of Ω.

And from the definition of entropy, $S = k\ln\Omega$, we see this means that the equilibrium state is specified by

$$\frac{\partial S_1}{\partial E} = \frac{\partial S_2}{\partial E}. \tag{1.3.7}$$

That is, when the systems have reached equilibrium the quantity $\partial S/\partial E$ of system 1 is equal to $\partial S/\partial E$ of system 2. This is the condition for equilibrium when systems exchange only thermal energy.

Clearly $\partial S/\partial E$ must be related to the temperature of the system. The Second Law requires that $\Delta S \geq 0$; that is

$$\Delta S = \left(\frac{\partial S_1}{\partial E} - \frac{\partial S_2}{\partial E}\right)\Delta E_1 \geq 0. \tag{1.3.8}$$

This means that

$$\begin{aligned} &E_1 \text{ increases (and } E_2 \text{ decreases)} \quad \text{if} \quad \frac{\partial S_1}{\partial E} > \frac{\partial S_2}{\partial E} \\ &E_1 \text{ decreases (and } E_2 \text{ increases)} \quad \text{if} \quad \frac{\partial S_1}{\partial E} < \frac{\partial S_2}{\partial E} \end{aligned} \tag{1.3.9}$$

so energy flows from systems with small $\partial S/\partial E$ to systems with large $\partial S/\partial E$.

Since we know that heat flows from hot systems to cold systems, we therefore identify

$$\begin{aligned} \text{high } T &\equiv \text{low } \frac{\partial S}{\partial E} \\ \text{low } T &\equiv \text{high } \frac{\partial S}{\partial E}. \end{aligned} \tag{1.3.10}$$

There is thus an *inverse* relation between $\partial S/\partial E$ and temperature.

We are led to define *statistical temperature* by

$$\frac{1}{T} = \frac{\partial S}{\partial E}. \tag{1.3.11}$$

When applied to the ideal gas this will give us the result (Section 2.3.3)

$$pV = NkT, \tag{1.3.12}$$

and it is from *this* we conclude that the statistical temperature corresponds to the intuitive concept of temperature as measured by an ideal gas thermometer. Furthermore the *scale* of temperatures will agree with the Kelvin scale (ice point at 273.18 K) when the constant k in the definition of S is identified with Boltzmann's constant.

When the derivative $\partial S/\partial E$ is evaluated, N and V are constant. So the only energy flow is heat flow. Thus the equation defining statistical temperature can also be written as

$$\Delta Q = T\Delta S. \tag{1.3.13}$$

We can now write the energy conservation expression for the First Law:

$$\Delta E = \Delta Q + \Delta W \tag{1.3.14}$$

as

$$\Delta E = T\Delta S - p\Delta V \quad \text{(for } pV \text{ systems).} \tag{1.3.15}$$

1.3.2 Volume change – Pressure

We now allow the volumes of the interacting systems to vary as well, subject to the total volume being fixed. Thus we consider two systems separated by a *movable* diathermal wall.

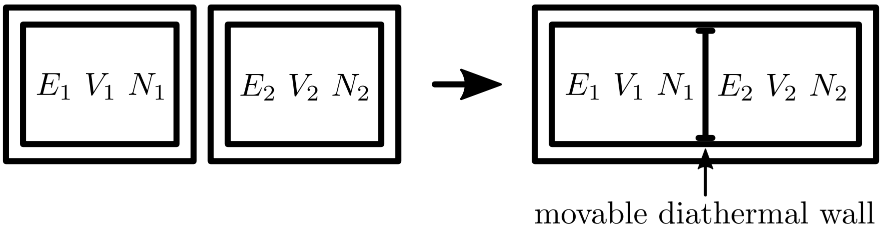


Figure 1.2: Mechanical and thermal interaction

The constraints on this system are

$$\begin{aligned} E_1 + E_2 &= E_0 = \text{const.} \\ V_1 + V_2 &= V_0 \ = \text{const.} \end{aligned} \tag{1.3.16}$$

while N_1 and N_2 are individually fixed.

Maximising the entropy with respect to both energy flow and volume change then gives the two conditions

$$\begin{aligned}\frac{\partial S_1}{\partial E} &= \frac{\partial S_2}{\partial E}\\ \frac{\partial S_1}{\partial V} &= \frac{\partial S_2}{\partial V}.\end{aligned} \tag{1.3.17}$$

The first of these gives, we know, the equality of temperature at equilibrium:

$$T_1 = T_2 \tag{1.3.18}$$

What does the second relation tell us? What is $\partial S/\partial V$? This may be found by rearranging the differential expression for the First Law:

$$\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V. \tag{1.3.19}$$

This may be re-written as

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{p}{T}\mathrm{d}V \tag{1.3.20}$$

so just as we identified

$$\left.\frac{\partial S}{\partial E}\right|_V = \frac{1}{T}, \tag{1.3.21}$$

so we now identify

$$\left.\frac{\partial S}{\partial V}\right|_E = \frac{p}{T}. \tag{1.3.22}$$

Thus the condition that $\partial S/\partial V$ be the same for both systems means that p/T must be the same. But we have already established that T is the same so the new information is that at equilibrium the pressures are equalized:

$$p_1 = p_2. \tag{1.3.23}$$

A paradox arises if the movable wall is not diathermal: that is, if it is thermally isolating. Then one would conclude, from an analysis similar to that above, that while p/T becomes equalized for the two sides, T does not. On the other hand a purely mechanical argument would say that the pressures p should become equal. The paradox is resolved when one appreciates that without a flow of heat, thermodynamic equilibrium is not possible and so the entropy maximum principle is not applicable. Thus p/T will not be equalized. This issue is discussed in greater detail by Callen [9].

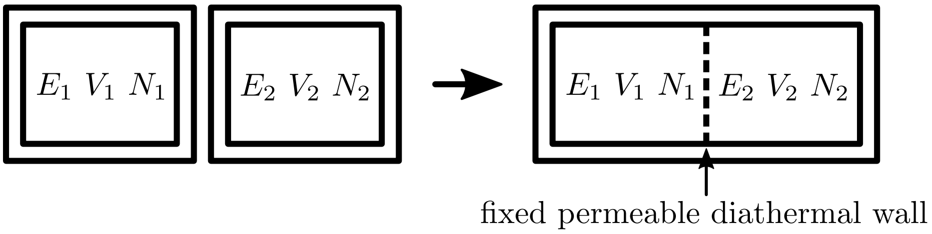


Figure 1.3: Heat and particle exchange

1.3.3 Particle interchange – chemical potential

Let us keep the volumes of the two systems fixed, but allow particles to traverse the immobile diathermal wall.

The constraints on this system are

$$\begin{aligned} E_1 + E_2 &= E_0 = \text{const.} \\ N_1 + N_2 &= N_0 = \text{const.} \end{aligned} \tag{1.3.24}$$

while V_1 and V_2 are individually fixed.

Maximising the entropy with respect to both energy flow and particle flow then gives the two conditions

$$\begin{aligned} \frac{\partial S_1}{\partial E} &= \frac{\partial S_2}{\partial E} \\ \frac{\partial S_1}{\partial N} &= \frac{\partial S_2}{\partial N}. \end{aligned} \tag{1.3.25}$$

The first of these gives, we know, the equality of temperature at equilibrium:

$$T_1 = T_2 \tag{1.3.26}$$

What does the second relation tell us? What is $\partial S / \partial N$? This may be found from the First Law in its extended form:

$$\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N \tag{1.3.27}$$

where μ is the chemical potential. This may be re-written as

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{p}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N \tag{1.3.28}$$

so that we may identify

$$\left.\frac{\partial S}{\partial N}\right|_{E,V} = -\frac{\mu}{T}. \tag{1.3.29}$$

Thus the condition that $\partial S/\partial N$ be the same for both systems means that μ/T must be the same. But we have already established that T is the same so the new information is that at equilibrium the chemical potentials are equalized:

$$\mu_1 = \mu_2. \tag{1.3.30}$$

We see that just as pressure drives volume changes, chemical potential drives particle flow. And arguments similar to those of Section 1.3.1 indicate that particles flow from high values of μ to low values of μ.

1.3.4 Thermal interaction with the rest of the world – the Boltzmann factor

For an isolated system all microstates are equally likely; this is our Fundamental Postulate. It follows that the probability of the occurrence of a given *microstate* is given by

$$P_j = \frac{1}{\Omega}. \tag{1.3.31}$$

But what about a non-isolated system? What can we say about the occurrence of microstates of such a system? Here the probability of a microstate will depend on properties of the surroundings.

In effect we are seeking an extension of our Fundamental Postulate. We shall see how we can use the Fundamental Postulate itself to effect its own extension!

We consider a system interacting with its surroundings through a fixed diathermal wall; this non-isolated system can exchange thermal energy with its surroundings. We ask the question "what is the probability of this non-isolated system being in a given microstate?"

We shall idealize the surroundings by a "large" system, which we will call a *heat bath.* We shall regard the composite system of bath plus our system of interest as isolated – so to *this* we can apply the Fundamental Postulate. In this way we shall be able to find the probability that the system of interest is in a particular microstate. This is the "wine bottle in the swimming pool" model of Reif [10].

The Ωs multiply, thus

$$\begin{array}{ccccc} \Omega_\mathrm{T} & = & \Omega_\mathrm{B} & \times & \Omega \\ \uparrow & & \uparrow & & \uparrow \\ \text{Total} & & \text{Bath} & & \text{System of interest} \end{array} \tag{1.3.32}$$

or

$$\Omega_\mathrm{T} = \Omega_\mathrm{B}(E_\mathrm{T} - E)\Omega(E). \tag{1.3.33}$$

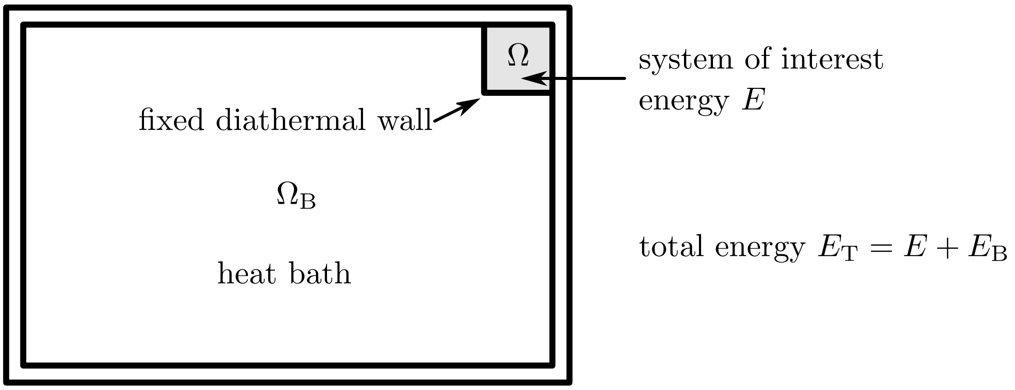


Figure 1.4: Thermal interaction with the rest of the world

Now the Fundamental Postulate tells us that the probability the system of interest has energy E is proportional to the number of microstates of the composite system that correspond to that energy partition

$$P(E) \propto \Omega_B(E_T - E)\Omega(E). \tag{1.3.34}$$

But here $\Omega(E) = 1$ since we are looking at a given *microstate* of energy E; there is one microstate. So

$$P(E) \propto \Omega_B(E_T - E). \tag{1.3.35}$$

It depends solely on the bath. In terms of entropy, since $S = k \ln \Omega$,

$$P(E) \propto e^{S(E_T - E)/k} \tag{1.3.36}$$

where S is the entropy of the *bath.* This type of expression, where probability is expressed in terms of entropy is an inversion of the usual usage where entropy and other thermodynamic properties are found in terms of probabilities. This form was much used by Einstein in his treatment of fluctuations.

Now the system of interest is very small compared with the bath; $E \ll E_T$. So we can perform a Taylor expansion of S:

$$S(E_T - E) = S(E_T) - E\frac{\partial S}{\partial E} + \ldots \tag{1.3.37}$$

but

$$\frac{\partial S}{\partial E} = \frac{1}{T}, \tag{1.3.38}$$

the temperature of the bath, so that

$$S(E_T - E) = S(E_T) - \frac{E}{T} \tag{1.3.39}$$

assuming we can ignore the higher terms. Then

$$P(E) \propto e^{S(E_{\mathrm{T}})/k} e^{-E/kT}. \tag{1.3.40}$$

But the first term $e^{S(E_{\mathrm{T}})/k}$ is simply a constant, so we finally obtain the probability

$$P(E) \propto e^{-E/kT}. \tag{1.3.41}$$

This is the probability that a system in equilibrium (with a bath) at a temperature T will be found in a microstate of energy E. The exponential factor $e^{-E/kT}$ is known as the Boltzmann factor, the Boltzmann distribution function or the canonical distribution function.

The Boltzmann factor is a key result. Feynman says [11]:

> *This fundamental law is the summit of statistical mechanics, and the entire subject is either a slide-down from the summit, as the principle is applied to various cases, or the climb-up to where the fundamental law is derived and the concepts of thermal equilibrium and temperature clarified.*

1.3.5 Particle and energy exchange with the rest of the world – the Gibbs factor

We now consider an extension of the Boltzmann factor to account for microstates where the number of particles may vary. Our system here can exchange both energy and particles with the rest of the world. The microstate of our system of interest is now specified by a given energy and a given number of particles. We are asking: what is the probability the system of interest will be found in the microstate with energy E and N particles?

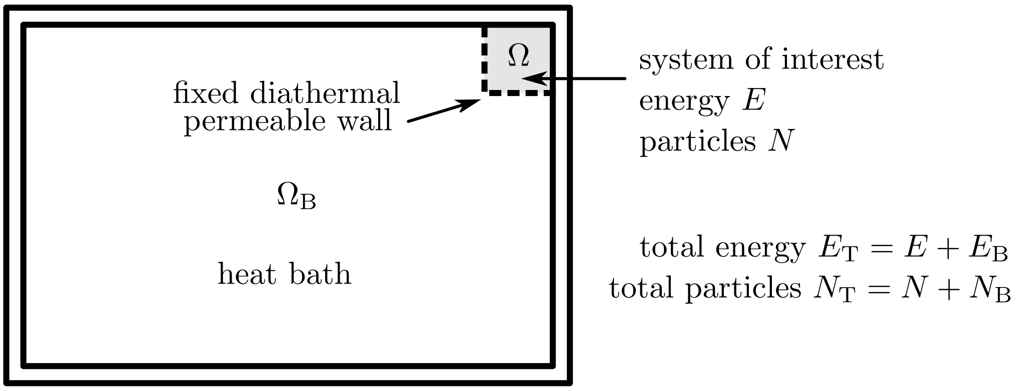


Figure 1.5: Particle and energy exchange with the rest of the world

In this case Ω_{T} is a function of both E and N

$$\Omega_{\mathrm{T}} = \Omega_{\mathrm{B}}(E_{\mathrm{T}} - E, N_{\mathrm{T}} - N)\Omega(E, N). \tag{1.3.42}$$

Now the Fundamental Postulate tells us that the probability the system of interest has energy E and N particles is proportional to the number of microstates of the composite system that correspond to that energy and particle number partition

$$P(E, N) \propto \Omega_{\mathrm{B}}(E_{\mathrm{T}} - E, N_{\mathrm{T}} - N)\Omega(E, N). \tag{1.3.43}$$

But as before, $\Omega(E, N) = 1$ since we are looking at a single microstate. So

$$P(E, N) \propto \Omega_{\mathrm{B}}(E_{\mathrm{T}} - E, N_{\mathrm{T}} - N). \tag{1.3.44}$$

It depends solely on the bath. In terms of entropy, since $S = k \ln \Omega$,

$$P(E, N) \propto e^{S(E_{\mathrm{T}} - E, N_{\mathrm{T}} - N)/k} \tag{1.3.45}$$

where S is the entropy of the *bath*.

Now the system of interest is very small compared with the bath; $E \ll E_{\mathrm{T}}$ and $N \ll N_{\mathrm{T}}$. So, as before, we can perform a Taylor expansion of S:

$$S(E_{\mathrm{T}} - E, N_{\mathrm{T}} - N) = S(E_{\mathrm{T}}, N_{\mathrm{T}}) - E\frac{\partial S}{\partial E} - N\frac{\partial S}{\partial N} + \ldots \tag{1.3.46}$$

but

$$\begin{aligned} \frac{\partial S}{\partial E} &= \frac{1}{T} \\ \frac{\partial S}{\partial N} &= -\frac{\mu}{T} \end{aligned} \tag{1.3.47}$$

so that

$$S(E_{\mathrm{T}} - E, N_{\mathrm{T}} - N) = S(E_{\mathrm{T}}, N_{\mathrm{T}}) - \frac{E}{T} + \frac{\mu N}{T} \tag{1.3.48}$$

assuming we can ignore the higher terms. Then

$$P(E, N) \propto e^{S(E_{\mathrm{T}}, N_{\mathrm{T}})/k} e^{-(E - \mu N)/kT}. \tag{1.3.49}$$

But the first term $e^{S(E_{\mathrm{T}}, N_{\mathrm{T}})/k}$ is simply a constant, so we finally obtain the probability

$$P(E, N) \propto e^{-(E - \mu N)/kT}. \tag{1.3.50}$$

This is the probability that a system in equilibrium (with a bath) at a temperature T and chemical potential μ will be found in a microstate of energy E, with N particles. The exponential factor $e^{-(E - \mu N)/kT}$ is sometimes known as the Gibbs factor, the Gibbs distribution function or the grand canonical distribution function.

1.4 Thermodynamic Averages

The importance of the previously derived probability distribution functions is that they may be used in calculating average (observed) values of various macroscopic properties of systems. In this way the aims of Statistical Mechanics, as outlined in Section 1.1.1 are achieved.

1.4.1 The Partition Function

The probability that a system is in the jth microstate, of energy $E_j(N, V)$, is given by the Boltzmann factor, which we write as:

$$P_j(N, V, T) = \frac{e^{-E_j(N,V)/kT}}{Z(N, V, T)} \tag{1.4.1}$$

where the normalization quotient Z is given by

$$Z(N, V, T) = \sum_i e^{-E_i(N,V)/kT}. \tag{1.4.2}$$

Here we have been particular to indicate the functional dependencies. Energy eigenstates depend on the size of the system (standing waves) and the number of particles. And we are considering our system to be in thermal contact with a heat bath; thus the temperature dependence. We do not, however, allow particle interchange.

The quantity Z is called the (canonical) *partition function*. The letter Z stands for the German word *Zustandssumme*, meaning “sum over states”. Although Z it has been introduced simply as a normalization factor, we shall see that it is a very useful quantity indeed.

1.4.2 Gibbs Expression for Entropy

For an isolated system the micro–macro connection is given by the Boltzmann formula $S = k \ln \Omega$, where Ω is a function of the extensive variables of the system

$$\Omega = \Omega(E, V, N). \tag{1.4.3}$$

But now, at a specified temperature, the energy E is not fixed; rather it fluctuates about a mean value $\langle E \rangle$.

To make the micro–macro connection when E is not fixed we must generalize the Boltzmann expression for entropy. We shall consider a collection of (macroscopically) identical systems in thermal contact, Fig. 1.6[1]. The composite system may be regarded as being isolated, so to *that* we may apply the rule $S = k \ln \Omega$, and from that the mean entropy of a representative single system may be found.

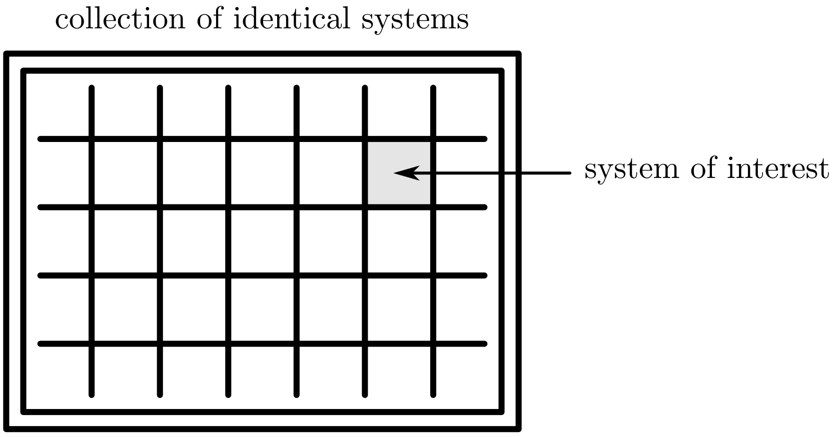


Figure 1.6: Gibbs ensemble for evaluating generalized entropy

Let us consider M identical systems, and let there be n_j of these systems in the jth microstate. We assume that this is a collection of a very large number of systems. Then the systems other than our one of particular interest may be regarded as a heat bath.

The number of possible microstates of the composite system corresponds to the number of ways of rearranging the subsystems:

$$\Omega = \frac{M!}{n_1!\, n_2!\, n_3! \ldots} \tag{1.4.4}$$

and the total entropy of the composite system is then

$$S_{\mathrm{tot}} = k \ln \left(\frac{M!}{n_1!\, n_2!\, n_3! \ldots} \right). \tag{1.4.5}$$

Since all the numbers here are large, we may make use of Stirling's approximation for the logarithm of a factorial, $\ln n! \approx n \ln n - n$, so that

$$S_{\mathrm{tot}} = k \left(M \ln M - \cancel{M} - \sum_j n_j \ln n_j + \cancel{\sum_j n_j} \right) \tag{1.4.6}$$

---

[1] Such a collection is called a Gibbs ensemble; the designation will become clear when you have studied Section 1.6.2.

the cancellation occurring because $M = \sum_j n_j$. Then if we express the first M as $\sum_j n_j$ we have

$$\begin{aligned} S_{\text{tot}} &= k \left(\sum_j n_j \ln M - \sum_j n_j \ln n_j \right) \\ &= -k \sum_j n_j \ln \left(\frac{n_j}{M} \right). \end{aligned} \tag{1.4.7}$$

We are interested in the mean entropy of our particular system. We have been considering a composite of M (macroscopically) similar systems, so the mean entropy of our system is simply the total entropy divided by M. Thus

$$S = -k \sum_j \frac{n_j}{M} \ln \left(\frac{n_j}{M} \right). \tag{1.4.8}$$

But n_j/M is the fraction of systems in the jth state, or the *probability* of finding our representative system in the jth state:

$$P_j = \frac{n_j}{M}. \tag{1.4.9}$$

So we can now express the mean entropy of a non-isolated system in terms of the state probabilities as

$$S = -k \sum_j P_j \ln P_j, \tag{1.4.10}$$

or

$$S = -k \left\langle \ln P \right\rangle, \tag{1.4.11}$$

the average value of the logarithm of the microstate probabilities. Eq. (1.4.10) is known as the Gibbs expression for the entropy. For an isolated system this reduces to the original Boltzmann expression.

1.4.3 Free Energy

In the Gibbs expression for entropy we actually know the values for the probabilities – they are given by the Boltzmann factor:

$$P_j(N, V, T) = \frac{e^{-E_j(N,V)/kT}}{Z(N, V, T)} \tag{1.4.12}$$

where we recall that the normalisation factor is given by the sum over states, the partition function Z

$$Z(N, V, T) = \sum_i e^{-E_i(N,V)/kT} \tag{1.4.13}$$

We then have

$$\ln P_j = -\left(\frac{E_j}{kT} + \ln Z\right). \tag{1.4.14}$$

Thus

$$S = k\left\langle \frac{E_j}{kT} + \ln Z \right\rangle \tag{1.4.15}$$

and since Z is independent of j we have

$$S = \frac{\langle E\rangle}{T} + k\ln Z. \tag{1.4.16}$$

Now in the spirit of thermodynamics we don't distinguish between mean and actual values – since fluctuations will be of order $1/\sqrt{N}$. Thus we write

$$E - TS = -kT\ln Z. \tag{1.4.17}$$

The quantity $E - TS$ is rather important and it is given a special name: *Helmholtz free energy*, or simply *free energy*. The symbol F is used[2]:

$$F = E - TS \tag{1.4.18}$$

so that we can write

$$F = -kT\ln Z. \tag{1.4.19}$$

It might be helpful to recall that $Z(N, V, T)$ is a *dimensionless* quantity.

1.4.4 Thermodynamic Variables

A host of thermodynamic variables can be obtained from the partition function. This is seen from the differential of the free energy. Since

$$\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N \tag{1.4.20}$$

it follows that

$$\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \mu\mathrm{d}N. \tag{1.4.21}$$

We can then immediately identify the various partial derivatives:

$$\left.\begin{aligned} S = -\left.\frac{\partial F}{\partial T}\right|_{V,N} &= \;\;\; kT\left.\frac{\partial \ln Z}{\partial T}\right|_{V,N} + k\ln Z \\ p = -\left.\frac{\partial F}{\partial V}\right|_{T,N} &= \;\;\; kT\left.\frac{\partial \ln Z}{\partial V}\right|_{T,N} \\ \mu = \left.\frac{\partial F}{\partial N}\right|_{T,V} &= -kT\left.\frac{\partial \ln Z}{\partial N}\right|_{T,V}. \end{aligned}\right\} \tag{1.4.22}$$

---

[2] Some texts use A for the Helmholtz free energy.

Since $E = F + TS$ we can then express the internal energy as

$$E = kT^2 \left. \frac{\partial \ln Z}{\partial T} \right|_{V,N}. \tag{1.4.23}$$

Thus we see that once the partition function is evaluated by summing over the states, all relevant thermodynamic variables can be obtained by differentiating Z.

1.4.5 The beta trick

There is a mathematical trick that can be very useful in many applications in Statistical Mechanics. This relies on a change of variables to "inverse temperature" defined by

$$\beta = 1/kT. \tag{1.4.24}$$

Many calculations are more easily done in this way; moreover the physical meaning of such calculations can more apparent.

As an example let's consider a direct calculation of the mean energy of a system $E = \sum_j E_j P_j$ (the internal energy). We now write this mean as

$$E = \frac{1}{Z} \sum_j E_j e^{-\beta E_j}. \tag{1.4.25}$$

We observe that the E_j in the sum may be "brought down" by a differentiation with respect to β:

$$E_j e^{-\beta E_j} = -\frac{\partial}{\partial \beta} e^{-\beta E_j} \tag{1.4.26}$$

So the internal energy is simply

$$E = -\frac{1}{Z} \sum_j \frac{\partial}{\partial \beta} e^{-\beta E_j}. \tag{1.4.27}$$

Then by swapping the order of summation and differentiation this reduces to

$$E = -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_j e^{-\beta E_j} \tag{1.4.28}$$

or

$$E = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}. \tag{1.4.29}$$

Equivalently we may write this in terms of $\ln Z$:

$$E = -\frac{\partial \ln Z}{\partial \beta}. \tag{1.4.30}$$

Upon changing variables back to $T = 1/k\beta$ this is equivalent to

$$E = kT^2 \frac{\partial \ln Z}{\partial T}; \tag{1.4.31}$$

and we have recovered Eq. (1.4.23).

1.4.6 Fluctuations

An isolated system has a well-defined energy. But the energy of a system in contact with a heat bath continually fluctuates about some mean value. The equilibrium with the heat bath is characterized by a temperature T and the system's mean energy is calculated from Eq. (1.4.23) or equivalently, from Eq. (1.4.29). In this section we shall (in contrast to the usual practice in Statistical Mechanics) distinguish between instantaneous and mean values. So here we write this equation as

$$\langle E \rangle = -\frac{1}{Z}\frac{\partial Z}{\partial \beta}. \tag{1.4.32}$$

What is the magnitude of the fluctuations about this mean?

We shall evaluate the RMS (root mean square) of the energy fluctuations σ_E, defined by

$$\sigma_E = \left\langle (E - \langle E \rangle)^2 \right\rangle^{1/2}. \tag{1.4.33}$$

By expanding out (the square) we obtain

$$\begin{aligned} \sigma_E^2 &= \left\langle E^2 \right\rangle - 2\left\langle E \right\rangle^2 + \left\langle E \right\rangle^2 \\ &= \left\langle E^2 \right\rangle - \left\langle E \right\rangle^2 \end{aligned} \tag{1.4.34}$$

In order to evaluate σ_E using this equation we have $\langle E \rangle^2$ by squaring Eq. (1.4.32). We now apply the beta trick to the calculation of $\langle E^2 \rangle$. We require to find

$$\left\langle E^2 \right\rangle = \frac{1}{Z}\sum_j E_j^2 e^{-\beta E_j}. \tag{1.4.35}$$

so we need to bring down *two* lots of E_j. This can be done with *two* differentiations

$$E_j^2 e^{-\beta E_j} = \frac{\partial^2}{\partial \beta^2} e^{-\beta E_j}. \tag{1.4.36}$$

So the mean square energy is simply

$$\left\langle E^2\right\rangle = -\frac{1}{Z}\sum_j \frac{\partial^2}{\partial\beta^2}e^{-\beta E_j}. \tag{1.4.37}$$

Then by swapping the order of summation and differentiation this reduces to

$$\begin{aligned}\left\langle E^2\right\rangle &= \frac{1}{Z}\frac{\partial^2}{\partial\beta^2}\sum_j e^{-\beta E_j}\\ &= \frac{1}{Z}\frac{\partial^2 Z}{\partial\beta^2}.\end{aligned} \tag{1.4.38}$$

We need the second derivative of Z. The first derivative is given, from Eq. (1.4.32), by

$$\frac{\partial Z}{\partial\beta} = -Z\left\langle E\right\rangle. \tag{1.4.39}$$

Differentiating again gives

$$\begin{aligned}\frac{\partial^2 Z}{\partial\beta^2} &= -\frac{\partial Z}{\partial\beta}\left\langle E\right\rangle - Z\frac{\partial\left\langle E\right\rangle}{\partial\beta}\\ &= Z\left\langle E\right\rangle^2 - Z\frac{\partial\left\langle E\right\rangle}{\partial\beta}\end{aligned} \tag{1.4.40}$$

so that

$$\left\langle E^2\right\rangle = \left\langle E\right\rangle^2 - \frac{\partial\left\langle E\right\rangle}{\partial\beta}. \tag{1.4.41}$$

This gives us σ_E^2 as

$$\sigma_E^2 = -\frac{\partial\left\langle E\right\rangle}{\partial\beta}. \tag{1.4.42}$$

We now return to conventional notation, discarding the averaging brackets, so that

$$\begin{aligned}\sigma_E^2 &= -\frac{\partial E}{\partial\beta}\\ &= kT^2\frac{\partial E}{\partial T}\end{aligned} \tag{1.4.43}$$

upon changing back to T as the independent variable. However since we recognize the derivative of energy, $\partial E/\partial T$ as the thermal capacity, the RMS variation in the energy may then be expressed as

$$\sigma_E^2 = kT^2 C_V \tag{1.4.44}$$

(the derivatives, following Eq. (1.4.23) are taken at constant V) or

$$\sigma_E = \sqrt{kT^2 C_V}. \tag{1.4.45}$$

Since C_V and E are both proportional to the number of particles in the system (extensive variables), the *fractional* fluctuations in energy vary as

$$\frac{\sigma_E}{\langle E\rangle} \sim \frac{1}{\sqrt{N}} \tag{1.4.46}$$

which gets smaller and smaller as N increases.

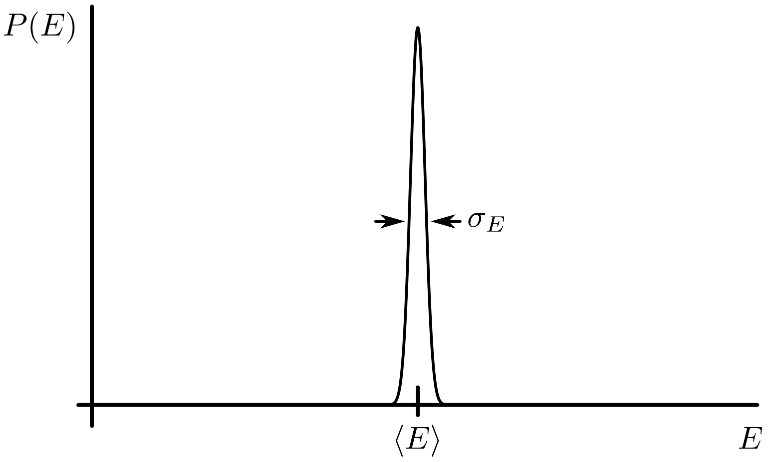


Figure 1.7: Fluctuations in energy for a system at fixed temperature

We thus see that the significance of fluctuations vanishes in the $N \to \infty$ limit, while N/V remains constant: the thermodynamic limit. And it is in this limit that statistical mechanics has its greatest applicability.

We might note that the usual use of Eq. (1.4.44) is the opposite: to show how the heat capacity is related to the energy fluctuations

$$C_V = \frac{1}{kT^2}\sigma_E^2. \tag{1.4.47}$$

1.4.7 The Grand Partition Function

Here we are concerned with systems of variable numbers of particles. The energy of a (many-body) state will depend on the number of particles in the system. As before, we label the (many-body) states of the system by j, but note that the jth state will be different for different N. In other words we need the pair $\{N, j\}$ for specification of a state. The probability that a system has N particles and is in the jth microstate, corresponding to a total energy $E_{N,j}$ is given by the Gibbs factor, which we write as:

$$P_{N,j}(V,T,\mu) = \frac{e^{-[E_{N,j}(N,V)-\mu N]/kT}}{\Xi(V,T,\mu)} \tag{1.4.48}$$

where the normalization constant Ξ is given by

$$\Xi(V,T,\mu) = \sum_{N,j} e^{-[E_{N,j}(N,V)-\mu N]/kT}. \tag{1.4.49}$$

Here both T and μ are properties of the bath. The quantity Ξ is called the *grand canonical partition function.* It may also be written in terms of the canonical partition function Z for different N:

$$\Xi(V,T,\mu) = \sum_{N} Z(N,V,T)e^{\mu N/kT}. \tag{1.4.50}$$

We will see that Ξ also is a useful quantity.

1.4.8 The Grand Potential

The generalized expression for entropy in this case is

$$S = -k\left\langle \ln P_{N,j}\right\rangle . \tag{1.4.51}$$

Here the probabilities are given by the Gibbs factor:

$$P_{N,j}(V,T,\mu) = \frac{e^{-[E_{N,j}(N,V)-\mu N]/kTkT}}{\Xi(V,T,\mu)} \tag{1.4.52}$$

where the normalisation factor, the sum over states, is the grand partition function Ξ. We then have

$$\ln P_{N,j} = -\left(\frac{E_{N,j}}{kT} - \frac{\mu N}{kT} + \ln\Xi\right). \tag{1.4.53}$$

Thus

$$S = k\left\langle \frac{E_{N,j}}{kT} - \frac{\mu N}{kT} + \ln\Xi\right\rangle \tag{1.4.54}$$

which is given by

$$S = \frac{\langle E\rangle}{T} - \frac{\mu\langle N\rangle}{T} + k\ln\Xi. \tag{1.4.55}$$

Now in the spirit of thermodynamics we don't distinguish between mean and actual values – since fluctuations will be of order $1\big/\sqrt{N}$. Thus we write

$$E - TS + \mu N = -kT \ln \Xi. \tag{1.4.56}$$

The quantity $E - TS + \mu N$ is equal to $-pV$ by the Euler relation (see Appendix A) so that we can write

$$pV = kT \ln \Xi. \tag{1.4.57}$$

The quantity pV is referred to as the *grand potential.*

1.4.9 Thermodynamic variables

Just as with the partition function, a host of thermodynamic variables can be obtained from the grand partition function. This is seen from the differential of the grand potential. It's easiest to differentiate pV as $-E + TS - \mu N$; that is

$$\begin{aligned} \mathrm{d}(pV) &= \mathrm{d}(-E + TS - \mu N) \\ &= -\mathrm{d}E + T\mathrm{d}S + S\mathrm{d}T - \mu\mathrm{d}N - N\mathrm{d}\mu. \end{aligned} \tag{1.4.58}$$

But since $\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N$, this gives

$$\mathrm{d}(pV) = S\mathrm{d}T + p\mathrm{d}V + N\mathrm{d}\mu. \tag{1.4.59}$$

We can then identify the various partial derivatives:

$$\left.\begin{aligned} S &= \left.\frac{\partial(pV)}{\partial T}\right|_{V,\mu} = kT \left.\frac{\partial \ln \Xi}{\partial T}\right|_{V,\mu} + k \ln \Xi \\ p &= \left.\frac{\partial(pV)}{\partial V}\right|_{T,\mu} = kT \left.\frac{\partial \ln \Xi}{\partial V}\right|_{T,\mu} = \frac{kT}{V} \ln \Xi \\ \mu &= \left.\frac{\partial(pV)}{\partial N}\right|_{T,V} = kT \left.\frac{\partial \ln \Xi}{\partial N}\right|_{T,V}. \end{aligned}\right\} \tag{1.4.60}$$

Since $E + \mu N = -pV + TS$ we can then express the combination $E + \mu N$ as

$$E + \mu N = kT^2 \left.\frac{\partial \ln \Xi}{\partial T}\right|_{V,\mu}. \tag{1.4.61}$$

Thus we see that once the grand partition function is evaluated by summing over the states, all relevant thermodynamic variables can be obtained by differentiating Ξ.

1.5 Quantum Distributions

1.5.1 Bosons and Fermions

All particles in nature can be classified into one of two groups according to the behaviour of their wave function under the exchange of identical particles. For simplicity let us consider just two identical particles. The wave function can then be represented as

$$\Psi = \Psi\left(r_1, r_2\right) \tag{1.5.1}$$

where

r_1 is the position of the first particle

and

r_2 is the position of the second particle.

Let us interchange the particles. We denote the operator that effects this by $\mathcal{P}$ (the permutation operator). Then

$$\mathcal{P}\Psi(r_1, r_2) = \Psi(r_2, r_1) \tag{1.5.2}$$

We are interested in the behaviour of the wave function under interchange of the particles. So far we have not drawn much of a conclusion. Let us now perform the swapping operation again. Then we have

$$\begin{aligned} \mathcal{P}^2\Psi(r_1, r_2) &= \mathcal{P}\Psi(r_2, r_1) \\ &= \Psi(r_1, r_2); \end{aligned} \tag{1.5.3}$$

the effect is to return the particles to their original states. Thus the operator $\mathcal{P}$ must obey

$$\mathcal{P}^2 = 1 \tag{1.5.4}$$

And taking the square root of this we find for $\mathcal{P}$.

$$\mathcal{P} = \pm 1. \tag{1.5.5}$$

In other words the effect of swapping two identical particles is either to leave the wave function unchanged or to change the sign of the wave function.

This property continues for all time since the permutation operator commutes with the Hamiltonian. Thus all particles in nature belong to one class or the other. Particles for which

$\mathcal{P} = +1$ are called *bosons*

while those for which

$\mathcal{P} = -1$ are called *fermions.*

Fermions have the important property of not permitting multiple occupancy of quantum states. Consider two particles in the same state, at the same position r. The wave function is then

$$\Psi = \Psi\left(r, r\right). \tag{1.5.6}$$

Swapping over the particles we have

$$\mathcal{P}\Psi = -\Psi. \tag{1.5.7}$$

But $\Psi = \Psi\left(r, r\right)$ so that $\mathcal{P}\Psi = +\Psi$ since both particles are in the same state. The conclusion is that

$$\Psi\left(r, r\right) = -\Psi\left(r, r\right) \tag{1.5.8}$$

and this can only be so if

$$\Psi\left(r, r\right) = 0. \tag{1.5.9}$$

Now since Ψ is related to the *probability* of finding particles in the given state, the result $\Psi = 0$ implies a state of zero probability – an impossible state. We conclude that it is impossible to have more than one fermion in a given quantum state.

This discussion was carried out using r_1 and r_2 to denote *position* states. However that is not an important restriction. In fact they could have designated any sort of quantum state and the same argument would follow.

This is the explanation of the Pauli exclusion principle obeyed by electrons.

We conclude:

- For bosons we can have any number of particles in a quantum state.
- For fermions we can have either 0 or 1 particle in a quantum state.

But what determines whether a given particle is a boson or a fermion? The answer is provided by quantum field theory. And it depends on the *spin* of the particle. Particles whose spin angular momentum is an integral multiple of $\hbar$ are bosons while particles whose spin angular momentum is integer plus a half $\hbar$ are fermions. (In quantum theory $\hbar/2$ is the smallest unit of spin angular momentum.) It is not straightforward to demonstrate this fundamental connection between spin and statistics. Feynman's heroic attempt is contained in his 1986 Dirac memorial lecture [12]. However a slightly more accessible account is contained in Tomonaga's book *The Story of Spin* [13].

For some elementary particles we have:

$$\left.\begin{array}{r}\left.\begin{array}{r}\text{electrons}\\ \text{protons}\\ \text{neutrons}\end{array}\right\} S=\dfrac{1}{2}\end{array}\right\} \rightarrow \text{fermions}$$

$$\left.\begin{array}{rl}\text{photon} & S=1\\ \left.\begin{array}{r}\pi\ \text{meson}\\ K\ \text{meson}\end{array}\right\} & S=0\end{array}\right\} \rightarrow \text{bosons}$$

For composite particles (such as atoms) we simply add the spins of the constituent parts. And since protons, neutrons and electrons are all fermions we can say:

- odd number of fermions → fermion;
- even number of fermions → boson.

The classic example of this is the two isotopes of helium. Thus

- ^{3}He is a fermion;
- ^{4}He is a boson.

Although these are *chemically* equivalent, at low temperatures the isotopes have very different behaviour.

1.5.2 Grand Potential for Identical Particles

The grand potential allows the treatment of systems of variable numbers of particles. We may exploit this in the study of systems of non-interacting (or weakly-interacting) particles in the following way. We focus attention on a single-particle state, which we label by k. The state of the entire system is specified when we know how many particles are in each different (single-particle) quantum state.

$$\begin{aligned}\text{many-particle state} &\equiv \{n_1, n_2, \ldots, n_k, \ldots\}\\ \text{Energy of state} &= \sum_k n_k \varepsilon_k\\ \text{No. of particles} &= \sum_k n_k.\end{aligned} \tag{1.5.10}$$

Here ε_k is the energy of the kth single-particle state. Note that the ε_k are independent of N.

Now since the formalism of the grand potential is appropriate for systems that exchange particles and energy with their surroundings, we may now consider as our "system" the subsystem comprising the particles in a given state k. For this subsystem

$$\begin{aligned} E &= n_k \varepsilon_k \\ N &= n_k \end{aligned} \tag{1.5.11}$$

so that the probability of observing this, i.e. the probability of finding n_k particles in the kth state (provided this is allowed by the statistics) is

$$P_{n_k}(V, T, \mu) = \frac{e^{-(n_k \varepsilon_k - n_k \mu)/kT}}{\Xi_k} \tag{1.5.12}$$

where the grand partition function for the subsystem can be written

$$\Xi_k = \sum_{n_k} \left\{ e^{-(\varepsilon_k - \mu)/kT} \right\}^{n_k}. \tag{1.5.13}$$

Here n_k takes only values 0 and 1 for fermions and 0, 1, 2, $\ldots \infty$ for bosons. The grand potential for the "system" is

$$\begin{aligned} (pV)_k &= kT \ln \Xi_k \\ &= kT \ln \sum_{n_k} \left\{ e^{-(\varepsilon_k - \mu)/kT} \right\}^{n_k}. \end{aligned} \tag{1.5.14}$$

The grand partition function for the entire system is the product

$$\Xi = \prod_k \Xi_k \tag{1.5.15}$$

so that the grand potential (and any other extensive quantity) for the entire system is found by summing over all single particle state contributions:

$$pV = \sum_k (pV)_k. \tag{1.5.16}$$

1.5.3 The Fermi-Dirac Distribution

For fermions the grand potential for the single state is

$$\begin{aligned} (pV)_k &= kT \ln \sum_{n_k = 0,1} \left\{ e^{-(\varepsilon_k - \mu)/kT} \right\}^{n_k} \\ &= kT \ln \left\{ 1 + e^{-(\varepsilon_k - \mu)/kT} \right\}. \end{aligned} \tag{1.5.17}$$

From this we can find the mean number of particles in the state using

$$\begin{aligned}\bar{n}_k &= \left.\frac{\partial (pV)_k}{\partial \mu}\right|_{T,V} = \frac{e^{-(\varepsilon_k-\mu)/kT}}{1+e^{-(\varepsilon_k-\mu)/kT}} \\ &= \frac{1}{e^{(\varepsilon_k-\mu)/kT}+1}.\end{aligned} \tag{1.5.18}$$

This is known as the Fermi-Dirac distribution function.

The grand potential for the entire system of fermions is found by summing the single-state grand potentials

$$pV = kT\sum_k \ln\left\{1+e^{-(\varepsilon_k-\mu)/kT}\right\}. \tag{1.5.19}$$

1.5.4 The Bose-Einstein Distribution

For bosons the grand potential for the single state is

$$\begin{aligned}(pV)_k &= kT\ln\sum_{n_k=0}^{\infty}\left\{e^{-(\varepsilon_k-\mu)/kT}\right\}^{n_k} \\ &= kT\ln\left\{\frac{1}{1-e^{-(\varepsilon_k-\mu)/kT}}\right\} \\ &= -kT\ln\left\{1-e^{-(\varepsilon_k-\mu)/kT}\right\}\end{aligned} \tag{1.5.20}$$

assuming the geometric progression is convergent. From this we can find the mean number of particles in the state using

$$\begin{aligned}\bar{n}_k &= \left.\frac{\partial\,(pV)_k}{\partial \mu}\right|_{T,V} = \frac{e^{-(\varepsilon_k-\mu)/kT}}{1-e^{-(\varepsilon_k-\mu)/kT}} \\ &= \frac{1}{e^{(\varepsilon_k-\mu)/kT}-1}.\end{aligned} \tag{1.5.21}$$

This is known as the Bose-Einstein distribution function.

The grand potential for the entire system of bosons is found by summing the single-state grand potentials

$$pV = -kT\sum_k \ln\left\{1-e^{-(\varepsilon_k-\mu)/kT}\right\}. \tag{1.5.22}$$

(An elegant derivation of the Bose and the Fermi distributions which indicates how the + and − sign in the denominators arises directly from the eigenvalue of the $\mathcal{P}$ operator is given in the Quantum Mechanics text book by Merzbacher [14]. Beware, however – it uses the method of Second Quantization.) See also the simple derivation of the Bose and Fermi distributions given in Appendix E, inspired by Feynman, based on the canonical distribution function.

1.5.5 The Classical Limit – The Maxwell-Boltzmann Distribution

The Bose-Einstein and Fermi-Dirac distributions give the mean numbers of particles in the microstate of energy ε_j as a function of $(\varepsilon_j - \mu)/kT$. When this quantity is large we observe two things. Firstly the denominator of the distributions will be very much larger than one, so the $+1$ or -1 distinguishing fermions from bosons may be neglected. And secondly the large value for the denominator means that the $\bar{n}_j$, the mean occupation of the state, will be very much less than unity.

This condition will apply to all states, down to the ground state of $\varepsilon_j = 0$, if μ/kT is large and negative. This is the classical limit where the issue of multiple state occupancy does not arise and the distinction between fermions and bosons becomes unimportant. We refer to such (hypothetical) particle as maxwellons, obeying Maxwell-Boltzmann statistics. Thus for these particles the mean number of particles in the state is given by

$$\bar{n}_k = e^{-(\varepsilon_k - \mu)/kT}; \tag{1.5.23}$$

This is essentially the Boltzmann distribution function.

The three distribution functions are shown in Fig. 1.8. Observe, in particular, that when $\mu = \varepsilon$ the Fermi occupation is one half, the Maxwell occupation is unity and the Bose occupation is infinite.

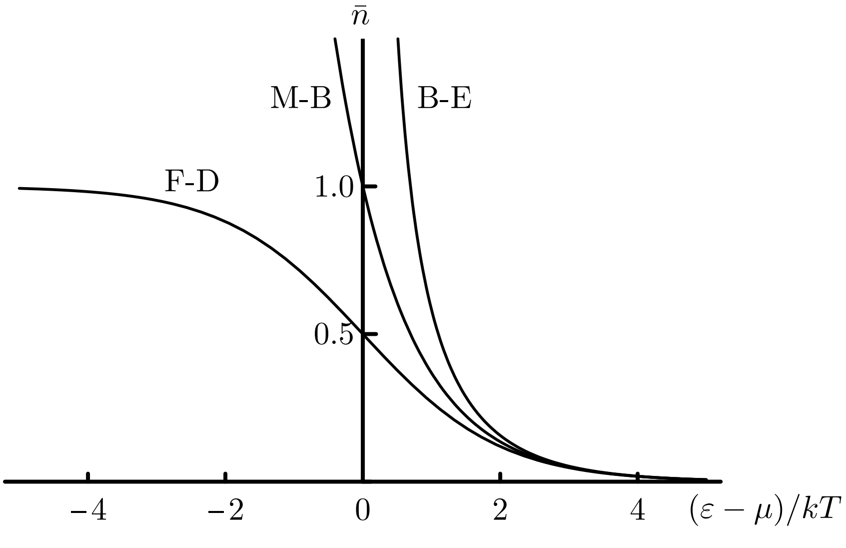


Figure 1.8: Fermi-Dirac, Maxwell-Boltzmann and Bose-Einstein distribution functions

1.6 Classical Statistical Mechanics

1.6.1 Phase Space and Classical States

The formalism of statistical mechanics developed thus far relies very much, at the microscopic level, on the use of (micro)states. We count the number of states, we sum over states etc. etc. This is all very convenient to do within the framework of a quantum-mechanical description of systems where states of a (finite or bound) system are discrete, but what about classical systems. How is the formalism of *classical* statistical mechanics developed – what is a "classical state"?

To specify a state in classical mechanics we must know the position and velocity of all particles in the system. (Position *and* velocity, since the equations of motion – Newton's laws – are second-order differential equations). For reasons which become clear in the Lagrangian and the Hamiltonian formulations of mechanics, it proves convenient to use the position and the *momentum* rather than velocity of the particles in the system. This is because it is then possible to work in terms of *generalized* coordinates and momenta – such as angles and angular momenta – in a completely general way; one is not constrained to a particular coordinate system. Thus we will say that a classical state is specified by the coordinates and the momenta of all the constituent particles. A single particle has 3 coordinates x, y, z and three momentum components p_x, p_y, p_z so it needs *six* components to specify its state. The generalized coordinates are conventionally denoted by q and the momenta by p. This p, q space is called *phase space.*

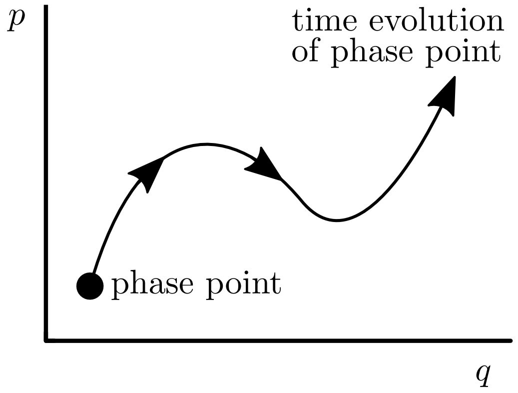


Figure 1.9: Trajectory in phase space

The classical state of a particle is denoted by a *point* in phase space. Its evolution in time is represented by a *curve* in phase space. It should be evident that during its evolution the phase curve of a point cannot intersect itself since there is a unique path proceeding from each location in phase space.

There is a difficulty in counting these states since the p and q vary continuously; there would be an *infinite* number of states in any region of phase space. In classical statistical mechanics it is expedient to erect a *grid* in phase space with cells $\Delta q_x\, \Delta q_y\, \Delta q_z\, \Delta p_x\, \Delta p_y\, \Delta p_z$. Then the classical analogue of a quantum microstate is a cell in phase space. In other words, a system is in a given microstate if it is in a specified cell in phase space.

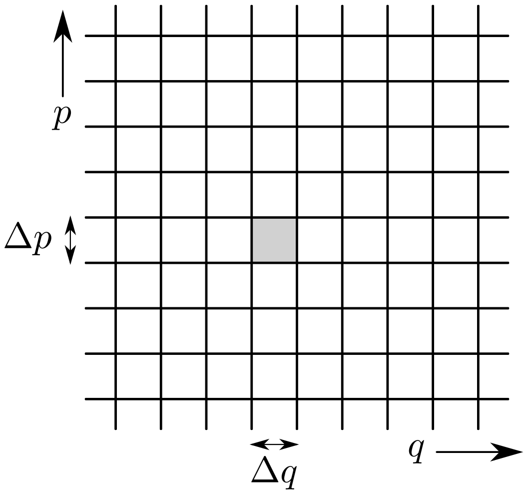


Figure 1.10: Classical microstate in phase space

What should be the size of these cells in phase space? From the historical perspective we would say that the “area” of each product $\Delta p\, \Delta q$ will be a constant h and then we will observe:

- how that constant appears in calculated physical quantities and
- how such calculated physical quantities compare with those calculated quantum-mechanically.

Looking ahead we will see that h is identified with Planck’s constant. This might be surprising. But we note that the Uncertainty Principle of quantum mechanics tells us we cannot discern a state to better than a $\Delta p\, \Delta q$ cell smaller than Planck’s constant.

The general rule is that the sums over states in the quantum case correspond to integrals over phase space, with appropriate normalization, in the classical case:

$$\sum_{\substack{\text{single}\\\text{particle}\\\text{states}}} \to \frac{1}{h^3}\int \mathrm{d}^3p\,\mathrm{d}^3q. \qquad (1.6.1)$$

1.6.2 Boltzmann and Gibbs Phase Spaces

The microstate of a system is represented by a point in phase space. A different microstate will be represented by a different point in the phase space. In developing the statistical approach to mechanics we must talk about different microstates – so we are considering different points in phase space and the probabilities associated with them. Boltzmann and Gibbs looked at this in different ways. Boltzmann's idea was that a gas of N particles would be represented by N points in the 6-dimensional phase space. The evolution of the state of the system with time is then described by the "flow" of the "gas" of points in the phase space. This sort of argument only works for weakly interacting particles, since:

- only then can you talk about the state of an individual particle, and
- later arguments are based on the movement of the points in phase space being independent.

Gibbs adopted a rather more general approach. He regarded the state of a system of N particles as being specified by a single point in a $6N$–dimensional phase space. The 6–dimensional phase space of Boltzmann is sometimes referred to as μ–space and the $6N$–dimensional phase space of Gibbs is sometimes referred to as Γ–space.

In both cases one applies probabilistic arguments to the collection of points in phase space. This collection is called an *ensemble*. So in Boltzmann's view a single particle is the system and the N particles comprise the ensemble while in Gibbs's view the assembly of particles is the system and many imaginary copies of the system comprise the ensemble. In the Boltzmann case one performs averages over the possible states of a single particle, while in the Gibbs case one is considering possible states of the entire system and applying probabilistic arguments to those.

Both views are useful. The Boltzmann approach is easier to picture but it can only be applied to weakly interacting particles. The Gibbs approach is more powerful as it can be applied to strongly interacting systems where the particles cannot be regarded as being even *approximately* independent.

The probability of finding a system in a microstate in the region $\mathrm{d}p\,\mathrm{d}q$ of phase space is given by $\rho\,\mathrm{d}p\,\mathrm{d}q$ where ρ is the density of representative points in the phase space. So the probability density of the microstate p, q is given by $\rho(p, q)$.

1.6.3 The Fundamental Postulate in the Classical Case

If we say that all (quantum) states are equally likely, then the classical analogue will be that all points in phase space are equally likely. The quantum version of the Fundamental Postulate refers to an *isolated system.* This is a system for which E, V and N are fixed. Classically when the energy is fixed, this restricts the accessible region of phase space to a constant-energy *hypersurface.* Similarly, fixing V and N determines which regions of phase space are available. The classical version of the Fundamental Postulate then states that **for an isolated system all available regions of phase space on the constant energy hyper-surface are equally likely**.

The probability of a macrostate corresponding to a region of phase space will then be proportional to the number of phase points in the region. So correctly normalized, it will be given by the *density* of points $\rho(p, q)$ where by p and q we mean the set of all momentum coordinates and all position coordinates.

1.6.4 The classical partition function

The classical analogue of the quantum partition function, Eq. (1.4.2), is given by

$$Z = \frac{1}{h^{3N}} \int e^{-H(p_i, q_i)/kT} \mathrm{d}^{3N} p \, \mathrm{d}^{3N} q. \tag{1.6.2}$$

The function $H(p_i, q_i)$ is the energy of the system expressed as a function of the position and momentum coordinates q_i and p_i (called the Hamiltonian).

1.6.5 The equipartition theorem

The equipartition theorem is concerned with the internal energy associated with individual degrees of freedom of a system. It has important consequences for the behaviour of the thermal capacity of classical systems.

We ask the question “What is the internal energy associated with a given degree of freedom, – say p_i?”. That is easy to write down:

$$\langle E_i \rangle = \frac{1}{Z} \int E_i e^{-E(q_1 \ldots q_i \ldots q_N,\, p_1 \ldots p_i \ldots p_N)/kT} \mathrm{d}^{3N} q \; \mathrm{d}^{3N} p. \tag{1.6.3}$$

Assuming the energy E_i depends only on the p_i and not on the other ps and qs, we can factorize that bit out of the exponential and write the integral as

$$\langle E_i \rangle = \frac{\int E_i e^{-E_i/kT} \mathrm{d}p_i \int \overset{\downarrow \cdots\cdots\cdots \ \text{no}\, p_i \cdots\cdots\cdots \ \downarrow}{e^{-E(q_1 \ldots q_i \ldots q_N, p_1 \ldots p_n)/kT} \mathrm{d}^{3N} q \, \mathrm{d}^{3N-1} p}}{\int e^{-E_i/kT} \mathrm{d}p_i \int [\text{same integral as above} - \text{no}\, p_i]} \tag{1.6.4}$$

So the second integral in numerator and denominator cancel, leaving the simple expression

$$\langle E_i \rangle = \frac{\int E_i e^{-E_i/kT} \mathrm{d}p_i}{\int e^{-E_i/kT} \mathrm{d}p_i}. \tag{1.6.5}$$

This may be simplified by using the beta trick discussed in Section 1.4.5. We write $\beta = 1/kT$ and use

$$E_i e^{-\beta E_i} = -\frac{\partial}{\partial \beta} e^{-\beta E_i} \tag{1.6.6}$$

so that

$$\langle E_i \rangle = -\frac{\frac{\partial}{\partial \beta} \int e^{-\beta E_i} \mathrm{d}p_i}{\int e^{-\beta E_i} \mathrm{d}p_i} \tag{1.6.7}$$

or

$$\langle E_i \rangle = -\frac{\partial}{\partial \beta} \ln \int e^{-\beta E_i} \mathrm{d}p_i. \tag{1.6.8}$$

At this stage we must be more specific about the functional form of $E_i(p_i)$. Since the p_i is a momentum then for a classical particle $E_i(p_i) = p_i^2/2m$: a quadratic dependence. For simplicity let us write simply

$$E_i = b p_i^2 \tag{1.6.9}$$

for some positive constant b. The integral is then

$$\int e^{-\beta E_i} \mathrm{d}p_i = \int e^{-\beta b p_i^2} \mathrm{d}p_i. \tag{1.6.10}$$

We don't actually need to evaluate the integral! Remember that we are going to differentiate the logarithm of the integral with respect to β; so all we want is the β-dependence. Let us make a change of variable and put

$$\beta p_i^2 = y^2. \tag{1.6.11}$$

The integral then becomes

$$\beta^{-1/2} \int e^{-by^2} \mathrm{d}y \tag{1.6.12}$$

so that

$$\begin{aligned}\langle E_i\rangle &= -\frac{\partial}{\partial\beta}\ln\left(\beta^{-1/2}\int e^{-by^2}\mathrm{d}y\right)\\ &= -\frac{\partial}{\partial\beta}\left\{-\frac{1}{2}\ln\beta+\ln\int e^{-by^2}\mathrm{d}y\right\}.\end{aligned} \tag{1.6.13}$$

The second term is independent of β so upon differentiation it vanishes. Thus differentiating we obtain

$$\langle E_i\rangle = \frac{1}{2\beta} \tag{1.6.14}$$

or, in terms of T:

$$\langle E_i\rangle = \frac{1}{2}kT. \tag{1.6.15}$$

So you see we didn't have to evaluate the integral – and the b has vanished; the physics came out of the integral upon change of variables.

The general conclusion here may be stated as the **Equipartition theorem**: For a classical (non-quantum) system each degree of freedom with a quadratic dependence on coordinate or momentum gives a contribution to the internal energy of $kT/2$.

[Incidentally, if $E_i \propto q_i^n$ or p_i^n (for even n) then the corresponding equipartition energy is given by $\langle E_i\rangle = kT/n$.]

1.6.6 Consequences of equipartition

We consider two examples – lattice vibrations, and a gas of particles.

Lattice vibrations

For the case of lattice vibrations each atom is essentially three harmonic oscillators, one in the x, y, and z directions. Thus for N atoms we have $3N$ harmonic oscillators. Now in this case *both* the position and momentum coordinates contribute a quadratic term to the energy. The internal energy is then

$$E = 3NkT \tag{1.6.16}$$

in the non-quantum (high temperature) limit. Differentiation with respect to temperature gives the isochoric thermal capacity

$$\begin{aligned}C_V &= 3Nk\\ &= 3R \text{ per mol.}\end{aligned} \tag{1.6.17}$$

Gas of particles

Considering now a gas of noninteracting particles there is no contribution to the energy from the position coordinates. Only the momentum coordinates contribute a quadratic term to the energy and the internal energy is then

$$E = \frac{3}{2} NkT \tag{1.6.18}$$

in the non-quantum (high temperature) limit. Differentiation with respect to temperature gives the (isochoric) thermal capacity

$$\begin{aligned} C_\mathrm{V} &= \frac{3}{2} Nk \\ &= \frac{3}{2} R \text{ per mol.} \end{aligned} \tag{1.6.19}$$

The thermal capacity of the solid is double that of the fluid because in the solid the position coordinates also contribute to the internal energy.
[In fact the walls of a box of gas can be modelled as an oscillator with a power law potential $V \propto x^n$ where $n \to \infty$.]

Breakdown of equipartition

Equipartition breaks down when quantum effects become important. In Section 2.3.3 we shall see that the internal energy of a *single* quantum free particle corresponds to $\frac{3}{2}kT$: the equipartition value for the three spatial degrees of freedom. However once we have a collection of N identical particles comprising a quantum gas, the internal energy is given, in Section 2.7.1, by

$$E = \frac{3}{2} NkT \left\{ 1 + a\sqrt{\frac{2}{\pi}}\frac{1}{3}\left(\frac{\varepsilon_\mathrm{q}}{kT}\right)^{3/2} + a^2 \dots \right\}, \tag{1.6.20}$$

where $a = +1$ for fermions, zero for “classical” particles and -1 for bosons and ε_q is the quantum energy parameter (the Fermi energy in the case of fermions). The equipartition result occurs at high temperatures; as the gas cools, quantum effects become important. For fermions the internal energy increases above the equipartition value, while for bosons the internal energy decreases below the equipartition value.

In Problem 1.16 you will see that the internal energy of a quantum harmonic oscillator may be written as

$$E = kT + \frac{\hbar^2\omega^2}{12kT} + \cdots . \tag{1.6.21}$$

The first term represents the high-temperature equipartition value. The second (and higher) terms indicate the deviations, showing the internal energy increasing above its high-temperature value as the temperature is lowered.

1.6.7 Liouville's theorem

We ask the question "How does a macroscopic system evolve in time?". The answer is that it will develop in accordance with the Second Law of thermodynamics; the system evolves to the state of maximum entropy consistent with the constraints imposed. Can this be understood from microscopic first principles? In other words, can the law of entropy increase be derived from Newton's laws? Both Boltzmann and Gibbs agonized over this.

We need a definition of entropy which will be suitable for use in the classical case. The problem is that there are not *discrete* states now, since the p and the q can vary continuously. By analogy with the Gibbs expression for entropy:

$$S = -k \sum_j P_j \ln P_j, \tag{1.6.22}$$

since the probability (density) is given by the density of points in phase space, we now have

$$S = -k \int \rho \ln \rho \, \mathrm{d}p \, \mathrm{d}q. \tag{1.6.23}$$

It is essentially (minus) the average of the logarithm of the density of points in phase space.

If one calculates the way points move around phase space under the influence of the laws of mechanics one finds that the "flow" is incompressible. Thus the density remains constant. This result is known as Liouville's theorem. We need the machinery of Hamiltonian mechanics to show this. If you are happy with Hamiltonian mechanics the proof is sketched below. But the implication is that since ρ remains constant then the entropy remains constant, so the Second Law of thermodynamics seems to be inconsistent with the laws of mechanics at the microscopic level.

To demonstrate Liouville's theorem we first note that the flow of points in phase space must obey the equation of continuity, since the number of points is conserved:

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{v}\rho = 0. \tag{1.6.24}$$

However in this case ρ depends on the position and momentum coordinates, q and p. Thus the divergence contains all the $\partial/\partial q$ derivatives and all the

$\partial/\partial p$ derivatives. And the "velocity" $\mathbf{v}$ has components $\mathrm{d}p/\mathrm{d}t$ as well as the usual $\mathrm{d}q/\mathrm{d}t$. Thus the divergence term is actually

$$\operatorname{div}\mathbf{v}\rho = \frac{\partial}{\partial p}\left(\frac{\mathrm{d}p}{\mathrm{d}t}\rho\right) + \frac{\partial}{\partial q}\left(\frac{\mathrm{d}q}{\mathrm{d}t}\rho\right) \tag{1.6.25}$$

(these equations really contain all the q and p coordinates; the above, as elsewhere, is a shorthand simplification). We expand the p and q derivatives to give

$$\operatorname{div}\mathbf{v}\rho = \left(\frac{\partial}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t}\right)\rho + \frac{\partial\rho}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial\rho}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} \tag{1.6.26}$$

and then we use Hamilton's equations

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}, \quad \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p} \tag{1.6.27}$$

in the first bracket. Then

$$\frac{\partial}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} = -\frac{\partial^2 H}{\partial p\partial q} + \frac{\partial^2 H}{\partial q\partial p} = 0 \tag{1.6.28}$$

Thus we find that

$$\frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial\rho}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} = 0. \tag{1.6.29}$$

But we recognize this as the *total* derivative of ρ; it is the derivative of the density when moving with the flow in phase space. This is zero. Thus as the representative points evolve and flow in phase space, the local density remains constant. This is the content of Liouville's theorem, expressed in its usual form as

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = 0. \tag{1.6.30}$$

We have a paradox: since ρ remains constant during evolution then the entropy remains constant; the Second Law of thermodynamics seems to be inconsistent with the laws of mechanics at the microscopic level.

1.6.8 Boltzmann's H theorem

The resolution of the paradox of the incompatibility between Liouville's theorem and the Second Law may be understood from the *nature* of the flow of points in phase space. Boltzmann defined a quantity H, which was essentially the integral of $\rho\ln\rho$ over phase space and he obtained an equation of motion for H as a probabilistic differential equation for the flow of points into and out of regions of phase space. We shall adopt a variant of the approach of Gibbs to study the evolution of Boltzmann's H. Please note this H is neither the hamiltonian of the previous section nor the enthalpy function; this is *Boltzmann's* H.

The flow of points in phase space is complicated. Since we have a given number of elements in our ensemble, the number of points in phase space is fixed. And Liouville's theorem is saying that the multi-dimensional volume occupied by the points is constant. But the flow can be "dendritic" with fingers spreading and splitting in all directions, as in Fig. 1.11.

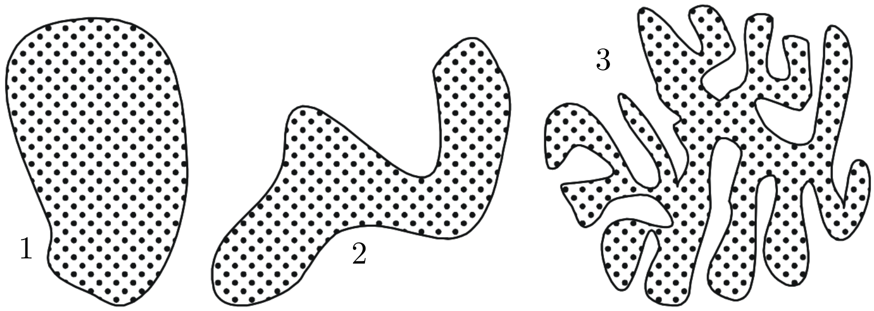


Figure 1.11: Evolution of a region of phase space

And as this happens, there will come a time when it is difficult to distinguish between what is an occupied region and what is an unoccupied region of phase space; they will be continually folded into each other. Gibbs argued that there was a scale in phase space, beyond which it was not possible (or at least reasonable) to discern. If the details of the state "3" in Fig. 1.11 are too fine to discern, then it will simply appear as a region of greater volume and, therefore, lesser density.

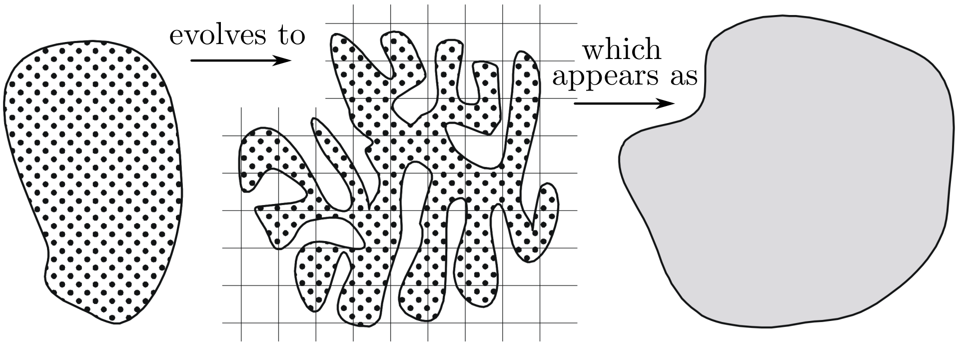


Figure 1.12: Apparent reduction in density in phase space

The procedure of taking an average over small regions of phase space in this way is known as "coarse graining". Gibbs showed rigorously (rather than by just using diagrams as here) that the coarse grained density of points in phase space decreased as time proceeded [15].

Thus the conclusion is that the coarse-grained H decreases; so the coarse-grained entropy increases.

There is a connection with quantum mechanics, which may be invoked in the question of coarse-graining. The volume of a "cell" in phase space is a product of p, q pairs. Now the Uncertainty Principle tells us that we cannot locate a point within a $p\,q$ area to within better than Planck's constant. This gives the ultimate resolution that is achievable in specifying the state of a system – so at the fundamental level there is indeed a firm justification for coarse-graining.

Quantum mechanics has a habit of popping up in the most unexpected areas of statistical thermodynamics. This theme continues into the next section.

1.7 The Third Law of thermodynamics

1.7.1 History of the Third Law

The third law of thermodynamics arose as the result of experimental work in chemistry, principally by the famous physical chemist Walther Nernst. He published what he called his "heat theorem" in 1906. A readable account of the history of the Third Law and the controversies surrounding its acceptance is given by Dugdale [16].

Nernst measured the change in Gibbs free energy and the change in enthalpy for chemical reactions which started and finished at the same temperature. At lower and lower temperatures he found that the changes in G and the changes in H became closer and closer.

Nernst was led to conclude that at $T = 0$ the changes in G and H were the same. And from some elementary thermodynamic arguments he was able to infer the behaviour of the entropy at low temperatures.

Changes in H and G are given by

$$\begin{aligned} \Delta H &= \quad T\Delta S + V\Delta p \\ \Delta G &= -S\Delta T + V\Delta p\,. \end{aligned} \tag{1.7.1}$$

Thus ΔG and ΔH are related by

$$\Delta G = \Delta H - T\Delta S - S\Delta T \tag{1.7.2}$$

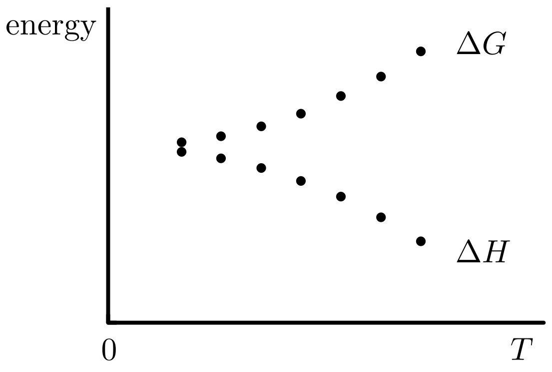


Figure 1.13: Nernst's observations

and if the temperature is the same before and after, $\Delta T = 0$, so then

$$\Delta G = \Delta H - T\Delta S. \tag{1.7.3}$$

This is a very important equation for chemists.

Now Nernst's observation may be stated as

$$\Delta H - \Delta G \to 0 \quad \text{as} \quad T \to 0, \tag{1.7.4}$$

which he realized implied that

$$T\Delta S \to 0 \quad \text{as} \quad T \to 0. \tag{1.7.5}$$

1.7.2 Entropy

On the face of it this result is no surprise since the factor T will ensure the product $T\Delta S$ goes to zero. But Nernst took the result further. He studied *how fast* $\Delta H - \Delta G$ tended to zero. And his observation was that it always went faster than linearly. In other words he concluded that

$$\frac{\Delta H - \Delta G}{T} \to 0 \quad \text{as} \quad T \to 0. \tag{1.7.6}$$

So even though $1/T$ was getting bigger and bigger, the quotient $(\Delta H - \Delta G)/T$ still tended to zero. But we know that

$$\frac{\Delta H - \Delta G}{T} = \Delta S. \tag{1.7.7}$$

So from this Nernst drew the conclusion

$$\Delta S \to 0 \quad \text{as} \quad T \to 0. \tag{1.7.8}$$

The entropy change in a process tends to zero at $T = 0$. The entropy thus remains a constant in any process at absolute zero. We conclude:

- The entropy of a body at zero temperature is a constant, independent of all other external parameters.

This was the conclusion of Nernst; it is sometimes called Nernst's heat theorem. It was subsequently to be developed into the Third Law of thermodynamics.

1.7.3 Quantum viewpoint

From the purely macroscopic perspective the third Law is as stated above: at $T = 0$ the entropy of a body is a constant. And many conclusions can be drawn from this. One might ask the question "what is the constant?". However we do know that thermodynamic conclusions about measurable quantities are not influenced by any such additive constants since one usually differentiates to find observables.

If we want to ask about the constant then we must look into the microscopic model for the system under investigation. Recall the Boltzmann expression for entropy:

$$S = k \ln \Omega$$

where Ω is the number of microstates in the macrostate. Now consider the situation at $T = 0$. Then we know the system will be in its ground state, the lowest energy state. But this is a *unique* quantum state. Thus for the ground state

$$\Omega = 1 \tag{1.7.9}$$

and so

$$S = 0. \tag{1.7.10}$$

Nernst's constant is thus zero and we then have the expression for the Third Law:

- As the absolute zero of temperature is approached the entropy of all bodies tends to zero.

We note that this applies specifically to bodies that are in *thermal equilibrium.* The Third Law can be summarized as

$$\frac{\partial S}{\partial \text{ anything}} \to 0 \quad \text{as} \quad T \to 0. \tag{1.7.11}$$

The above discussion is actually an over-simplification. In reality there may be degeneracy in the ground state of the system; then the above argument appears to break down. However, recall that entropy is an extensive quantity and that the entropy of the system should be considered in the thermodynamic limit. In other words, strictly, we should examine how the intensive quantity S/V or S/N behaves in the limit $V \to \infty, N \to \infty$ while V/N remains constant.

If the degeneracy of the ground state is g then we must look at the behaviour of $\ln(g)/N$. This will tend to zero in the thermodynamic limit so long as g increases with N no faster than exponentially. This is the fundamental quantum-mechanical principle behind the Third Law. The interested reader should consult the paper by Leggett [17] for a deeper discussion of these points.

To complete this discussion it is instructive to see how the Third Law would fail if classical mechanics were to apply down to $T = 0$. We saw, in Section 1.6.7, that the Gibbs expression for entropy:

$$S = -k \sum_j P_j \ln P_j \tag{1.7.12}$$

must be replaced, in the classical case by:

$$S = -k \int \rho \ln \rho \, \mathrm{d}p \, \mathrm{d}q, \tag{1.7.13}$$

where ρ is the density of points in phase space. This is necessary because in the classical case there are no discrete states and the momenta and coordinates, the p and q, can vary continuously.

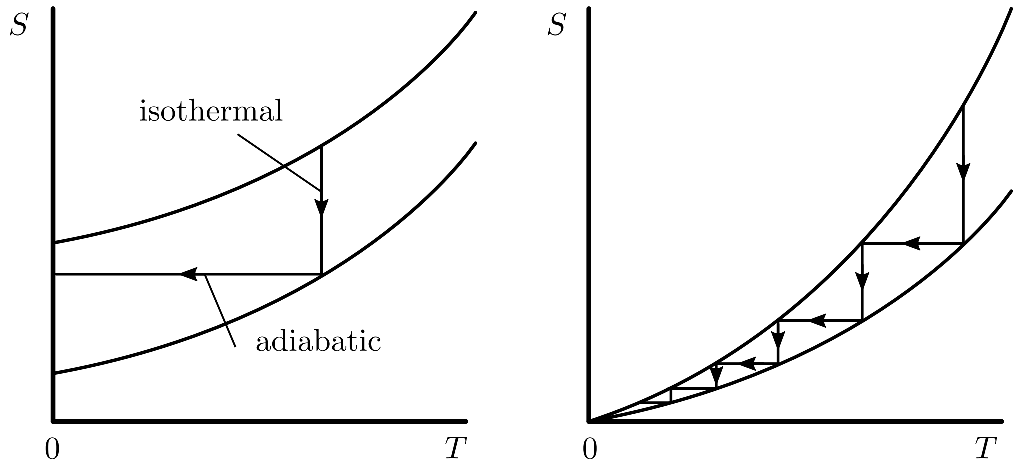
As the temperature is lowered the mean energy of the system will decrease. And corresponding to this, the “volume” of phase space occupied will decrease. In particular the momentum coordinates q will vary over a smaller and smaller range. In the $T \to 0$ limit the momentum range will become localized closer and closer to $q = 0$. The volume of occupied phase space shrinks to zero and the entropy thus tends to $-\infty$. This indeed is the limiting value indicated by the classical treatment of the Ideal Gas, as we shall see in Section 2.3.3.

The Uncertainly Principle of quantum mechanics limits the low temperature position-momentum specification of a system; you cannot localize points in phase space to a volume smaller than the appropriate power of Planck’s constant. This fundamental limitation of the density of phase points recovers the Third Law. Thus again we see the intimate connection between quantum mechanics and the Third law.

The Second Law tells us that there is an absolute zero of temperature. Now we see that the Third Law tells us there is an absolute zero of entropy.

1.7.4 Unattainability of absolute zero

The Third Law has important implications concerning the possibility of cooling a body to absolute zero. Let us consider a sequence of adiabatic and isothermal operations on two systems, one obeying the Third Law and one not, as in Fig. 1.14.



System not obeying Third Law
can get to $T = 0$ in two steps

System obeying Third Law
cannot get to $T = 0$ in a finite number of steps

Figure 1.14: Approaching absolute zero

Taking a sequence of adiabatics and isothermals between two values of some external parameter we see that the existence of the Third Law implies that you cannot get to $T = 0$ in a finite number of steps. This is, in fact, another possible statement of the Third Law.

Although one cannot get all the way to $T = 0$, it is possible to get closer and closer. Fig 1.15, adapted and extended from Pobel's book [18], indicates the success in this venture.

1.7.5 Heat capacity at low temperatures

The Third Law has important consequences for the heat capacity of bodies at low temperatures. Since

$$\begin{aligned} C &= \frac{\partial Q}{\partial T} \\ &= T\frac{\partial S}{\partial T}\,, \end{aligned} \tag{1.7.14}$$

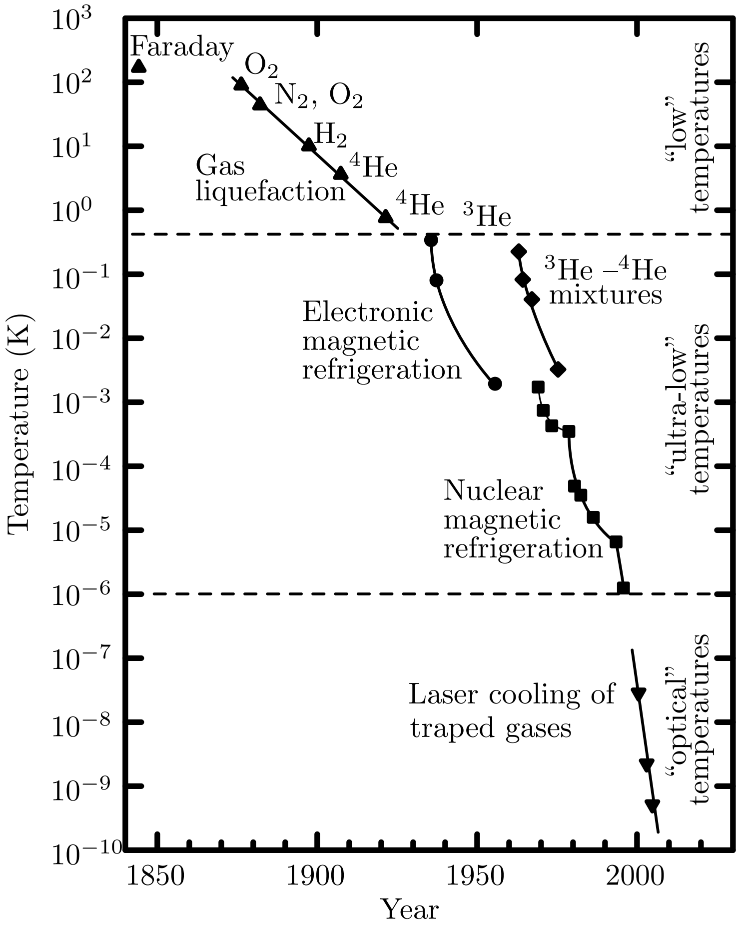


Figure 1.15: The road to absolute zero (adapted from Pobel [18]).

and the Third Law tells us that

$$\frac{\partial S}{\partial T} \to 0 \quad \text{as} \quad T \to 0 \tag{1.7.15}$$

we then have

$$C \to 0 \quad \text{as} \quad T \to 0 \tag{1.7.16}$$

Classical models often give a constant heat capacity – from equipartition. For an ideal gas (Section 2.3.3)

$$C_V = \frac{3}{2}Nk \tag{1.7.17}$$

independent of temperature. The Third Law tells us that this cannot hold at low temperatures. And indeed we shall see that for both Fermi and Bose gases C_V does go to zero as $T \to 0$.

1.7.6 Other consequences of the Third Law

Most "response functions" or susceptibilities – generalized spring constants – go to zero or to a constant as $T \to 0$ as a consequence of the Third Law. This is best seen by examining the relevant Maxwell relation (Appendix B.6). For example consider the thermal expansion coefficient. The Maxwell relation here is

$$\left.\frac{\partial V}{\partial T}\right|_p = -\left.\frac{\partial S}{\partial p}\right|_T . \tag{1.7.18}$$

The right hand side is zero by virtue of the Third Law. Thus we conclude that

$$\left.\frac{\partial V}{\partial T}\right|_p \to 0 \quad \text{as} \quad T \to 0; \tag{1.7.19}$$

the expansion coefficient goes to zero.

An interesting example is the magnetic susceptibility of a paramagnet, which we will treat in Section 2.9. The parallel with the ubiquitous model pV system is made by noting that the magnetic work increment is

$$\Delta W_{\mathrm{m}} = -M\Delta B, \tag{1.7.20}$$

corresponding to the mechanical work increment

$$\Delta W = -p\,\Delta V. \tag{1.7.21}$$

Thus we can take over pV results through the identification

$$\begin{aligned} M &\to p \\ B &\to V \end{aligned} \tag{1.7.22}$$

where M is the total magnetic moment.

The magnetic susceptibility is

$$\chi = \frac{\mu_0}{V}\frac{\partial M}{\partial B} \tag{1.7.23}$$

so that

$$\chi V/\mu_0 \to \frac{\partial p}{\partial V}. \tag{1.7.24}$$

There is no Maxwell relation for this, but consider the variation of the susceptibility with temperature:

$$\begin{aligned} \frac{\partial(\chi V/\mu_0)}{\partial T} &= \frac{\partial^2 M}{\partial T \partial B} \\ &\to \frac{\partial^2 p}{\partial T \partial V}. \end{aligned} \tag{1.7.25}$$

The order of differentiation can be reversed here. In other words

$$\frac{\partial}{\partial B}\frac{\partial M}{\partial T} \to \frac{\partial}{\partial V}\frac{\partial p}{\partial T}. \tag{1.7.26}$$

And now we do have a Maxwell relation:

$$\left.\frac{\partial p}{\partial T}\right|_V = \left.\frac{\partial S}{\partial V}\right|_T \quad \text{gives} \quad \left.\frac{\partial M}{\partial T}\right|_V = \left.\frac{\partial S}{\partial B}\right|_T. \tag{1.7.27}$$

The Third Law tells us that the right hand side of these equations goes to zero as $T \to 0$. We conclude then that

$$\frac{\partial \chi}{\partial T} \to 0 \quad \text{as} \quad T \to 0 \tag{1.7.28}$$

or

$$\chi \to \text{const} \quad \text{as} \quad T \to 0. \tag{1.7.29}$$

The Third Law tells us that the magnetic susceptibility becomes constant as $T \to 0$. But what does Curie's law, Eq. (2.9.12) say? This states

$$\chi = \frac{C}{T} \tag{1.7.30}$$

where C is the Curie constant. From this we conclude

$$\chi \to \infty \quad \text{as} \quad T \to 0\ !! \tag{1.7.31}$$

This is *completely incompatible* with the Third Law.

However Curie's law is a specifically high temperature result (strictly, it applies to the small B/T limit). The general expression for the magnetization of an ideal paramagnet of N spin ½ moments μ is, Eq. (2.9.9):

$$M = N\mu \tanh\left(\frac{\mu B}{kT}\right) \tag{1.7.32}$$

and corresponding to this, the (differential) susceptibility is

$$\chi = \frac{N\mu^2}{VkT}\operatorname{sech}^2\left(\frac{\mu B}{kT}\right). \tag{1.7.33}$$

Now we see that

$$\chi \to 0 \quad \text{as} \quad T \to 0 \tag{1.7.34}$$

in conformity with the Third Law, so long as the magnetic field B is finite. Of course you *have* to use a magnetic field, however small, to measure the susceptibility. Nevertheless, even in the absence of an externally-applied magnetic field, there will be an internal field present: the dipole fields of the magnetic moments themselves. Thus the Third Law is not under threat.

There is a further consideration in the case of fluid magnetic systems. In a fluid, where the particles must be treated as delocalized, the statistics will also have an effect. Consider the behaviour of fermions at low temperatures, to be treated in Section 2.5. Very roughly, only a fraction T/T_{F} of the particles are free and available to participate in "normal" behaviour. We then expect that the Curie law behaviour will be modified to

$$\chi \sim \left(\frac{T}{T_{\mathrm{F}}}\right) \times \frac{C}{T} \qquad (1.7.35)$$

or

$$\chi \sim \frac{C}{T_{\mathrm{F}}} \qquad (1.7.36)$$

which is indeed a constant, in conformity with the Third Law. This result is correct, but a more detailed calculation must be done to determine the numerical constants involved.

1.7.7 Pessimist's statement of the laws of thermodynamics

As we have now covered all the laws of thermodynamics we can present statements of them in terms of what they prohibit in the operation of Nature.

- **First Law**: You cannot convert heat to work at greater than 100% efficiency
- **Second Law**: You cannot even achieve 100% efficiency — except at $T = 0$
- **Third Law**: You cannot get to $T = 0$.

This is a simplification, but it encapsulates the underlying truths, and it is easy to remember.