COMMONWEALTH OF AUSTRALIA

Copyright Regulations 1969

WARNING

This material has been reproduced and communicated to you by or on behalf of the University of Queensland pursuant to Part VB of the Copyright Act 1968 (the Act).

The material in this communication may be subject to copyright under the Act. Any further reproduction or communication of this material by you may be the subject of copyright protection under the Act.

Do not remove this notice.
Acknowledgement

These notes were first prepared by Joel Gilmore, Kevin Brake, and Mark Dowling, based on Ross McKenzie’s handwritten notes for PHYS3020 in 2001. They have since been corrected and modified by Matthew Davis (2004–6). Thanks to Amanda Beasley for an extensive list of typos in 2004.

Very Important

These notes are not complete, are not a substitute for the textbook, and are not a substitute for attending lectures.

The material outlined herein will be covered in greater detail in lectures and tutorials.

While care has been taken to ensure the accuracy of these notes, there will undoubtedly be errors, which will be picked up and corrected during class.
What is the goal of statistical mechanics?

Quick answer: to traverse from the microscopic to the macroscopic using statistical methods.

E.g. from describing the physics of individual atoms to that of bulk matter in the form of gases, liquids, solids.

Thermodynamics describes the macroscopic properties of matter without any reference to the microscopic atoms and molecules that make up a particular system. It places restrictions on transformations of these properties that seem to be universal. However, it gives us no insight into why these restrictions exist.

Theoretically, if we know the forces that act between the atoms and molecules that make up a thermodynamic system, we can write down and solve the equations of motion for every particle (either using quantum mechanics or classical mechanics). However, it is obviously impractical to do so, and usually we aren’t interested in the detailed dynamics of every particle. However we are interested in the bulk thermodynamics properties such as pressure and temperature that we are familiar with from everyday life. We can obtain these using the laws of quantum or classical mechanics combined with statistics. For example: later in the course we will show that classical non-interacting particles obey the ideal gas law, \( PV = N k_B T \).

Statistical mechanics has a wide range of applications from condensed matter to astrophysics to quantum optics. Some that we will consider are:

- Cosmic microwave background
- White dwarf stars
- “Unzipping” DNA
- Electronic properties of metals
- Magnetic phase transitions
- Bose-Einstein condensation

Some fundamental questions

Most importantly — the second law of thermodynamics.

- Why does heat disperse as time passes?
- Why does time seem to flow in only one direction?

Explicitly, the microscopic equations of motion are time reversal invariant. E.g. Newton’s equation of motion for all \( N \) atoms in a gas

\[
m rac{d^2 \vec{r}_i}{dt^2} = \vec{F}_i(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)
\]

are invariant under the transformation \( t \to -t \).
Hamilton’s equations of motion

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

are invariant under \( t \to -t, p_i \to -p_i, \dot{p}_i \to -\dot{p}_i, \dot{q}_i \to -\dot{q}_i \). However, the behaviour of macroscopic objects is not time reversal invariant. Have you ever seen

- water go up a waterfall?
- a fresh cup of tea get hotter and not colder with time?
- sugar “undissolve” itself from the same cup of tea?

The fact that playing a video tape backwards makes (some of) us laugh within seconds demonstrates that the macroscopic world is irreversible — time flows in one direction.

**Significance**

Statistical mechanics has played a crucial role in the history of physics. In particular it is at the heart of

- the atomic hypothesis
- the proposal of the quantum by Planck
- even the origin of the universe! (via the observation of the cosmic microwave background)

**Course summary**

You may remember that the areas of classical mechanics, quantum mechanics, general relativity, and electromagnetism can be summarised by just a few equations. It turns out that statistical mechanics can be as well.

Consider a closed system at constant temperature \( T \), volume \( V \). Enumerate all of the possible microscopic configurations of the system by \( i \), with \( E_i \) the energy of the \( i \)th state. Then the Helmholtz free energy \( F(V, T) \) is related to the partition function via

\[ \exp(-F/k_BT) = Z \equiv \sum_i \exp(-E_i/k_BT) \]

where \( k_B = 1.381 \times 10^{-23} \) J/K = \( R/N_A \) is Boltzmann’s constant. This relates the macroscopic potential \( F(V, T) \) to our microscopic knowledge of the system.

This course is concerned with

1. Deriving this result, along with variations on it for e.g. variable particle number \( N \).
2. Applying this result to obtain (and hopefully understand!) the thermodynamic properties of a large number of model systems.
N.B. It turns out that enumerating the microstates, finding their energies and performing the sum is almost always a highly non-trivial exercise!

However, before we get stuck into statistical mechanics, we will use the first couple of lectures of this course reviewing what you should remember from thermodynamics.
Problem solving strategy

It is easy to get lost sometimes when approaching a problem in statistical mechanics — it happens to everyone. The strategy described below applies to just about every problem that we will come across in this course, and should help you get started if you are stuck.

1. Define the system. What are the macrostates?
2. Which ensemble is appropriate: microcanonical, canonical or grand canonical?
3. What are the microstates? This means working out what are the quantum numbers of the system, their associated energies, and the degeneracies of each energy.
4. Evaluate the partition function: i.e. perform the sum.
5. You now have the free energy. From its derivatives you can find thermodynamic properties such as the internal energy, entropy, and pressure (equation of state).
6. What is the relevant microscopic energy scale? Usually it is the energy level spacing. Call it $\epsilon_0$, say.
7. Consider the high temperature limit $k_B T \gg \epsilon_0$. This is the classical regime. Do your results agree with the equipartition theorem?
8. Consider the low temperature limit $k_B T \ll \epsilon_0$. This is the quantum regime. Do your results for the entropy agree with the third law of thermodynamics?
Some useful things to know

Often in statistical mechanics we run into formulae that are somewhat complicated and difficult to work with. Often, this can be circumvented by making a well-chosen approximation. An important skill to develop in PHYS3020 is an understanding of when and where such approximations can be made. It is also very important to check whether your results make physical sense. For example, at high temperature is your result consistent with what you know about classical physics?

Here we describe a few useful tools that will that turn up time and time again.

Taylor series expansion and exponentials

The Taylor expansion or series of the function $f(x)$ about the point $a$ is as follows

$$f(x) = f(a) + (x - a)f'(a) + \frac{1}{2}(x - a)^2 f''(a) + \frac{1}{6}(x - a)^3 f'''(a) + \ldots$$

(1)

$$= \sum_{n=0}^{\infty} \frac{(x - a)^n}{n!} \frac{d^n f(x)}{dx^n} \bigg|_{x=a}$$

(2)

Taylor series are often taken about the origin, which makes them look somewhat simpler. Probably the most important Taylor series for use in statistical mechanics is that for $e^x$ where

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$$

(3)

If $|x| \ll 1$ then we can often truncate this series at the first term

$$e^{\pm x} \approx 1 \pm x.$$  

(4)

Understand how big things are

It is very important to understand the importance of various terms in different physical situations, and make approximations based on this knowledge. If we have an expression like $1 + x$, we can forget the $x$ if it is very small. However, if it is very large then we can probably drop the one.

As an example, exponentials often turn up in the form $(e^{\varepsilon/k_BT} - 1)$. So when we have $\varepsilon \ll k_BT$ then we can approximate this as

$$(e^{\varepsilon/k_BT} - 1) \approx \left(1 + \frac{\varepsilon}{k_BT}\right) - 1,$$

$$\approx \frac{\varepsilon}{k_BT}.$$  

(5)

In the other limit when $\varepsilon \gg k_BT$, then the exponential term is going to be very large compared to one, so we can use

$$(e^{\varepsilon/k_BT} - 1) \approx e^{\varepsilon/k_BT}.$$  

(6)
Changing variables

A lot of integrals turn up in statistical mechanics, and with all the terms involved they can be quite confusing. A strategy that is often useful is to make a change of variables so that you can see the integral in its simplest form. An example is for the energy density of a hollow box

$$\frac{U}{V} = \int_0^\infty \frac{8\pi}{(hc)^3} \frac{\epsilon^3}{e^{\epsilon/k_B T} - 1} d\epsilon$$

(7)

If we change variables to $x = \epsilon/k_B T$ it becomes

$$\frac{U}{V} = \frac{8\pi (k_B T)^4}{(hc)^3} \int_0^\infty \frac{x^3}{e^x - 1} dx$$

(8)

This now has the advantage that all the parameter dependence is outside the integral, which is just some constant. We can look this up in tables — we find that it is $\pi^4/15$, and the final result is

$$\frac{U}{V} = \frac{8\pi^5 (k_B T)^4}{15 (hc)^3}$$

(9)

Changing variables can also be useful when considering the relative importance of terms.

Stirling’s approximation

Another useful approximation is for $N!$ when $N$ is very large. Factorials turn up quite often when working out different arrangements of particles in the microcanonical ensemble. Stirling’s approximation can make such terms easier to deal with. We have

$$N! \approx \left(\frac{N}{e}\right)^N \sqrt{2\pi N}. \quad (10)$$

This result is derived in Appendix B of Schroeder. Usually we want to know the logarithm of factorials: then we can use the slightly less accurate

$$\ln N! \approx N \ln N - N. \quad (11)$$

Some useful integrals

$$\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}, \quad (12)$$

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad \text{(the gaussian integral)} \quad (13)$$
Chapter 1

Review of thermodynamics

Thermodynamics is a phenomenological theory of matter. It should really be called thermostatics. It describes systems in equilibrium, i.e., systems which when left to themselves are such that large scale observations of them do not reveal any changes with time.

Thermodynamics

• makes no assumptions about the microscopic (atomic) nature of matter.
• gives general relations between measurable quantities such as specific heats and compressibilities, e.g.

\[
\left( \frac{\partial V}{\partial T} \right)_P = -\left( \frac{\partial S}{\partial P} \right)_T, \quad \frac{dP}{dT} = \frac{L}{T\Delta V}
\]

• does not say why these quantities have the values that they do.

1.1 Basics

• A thermodynamic system is any macroscopic system that can be defined by a finite number of co-ordinates, or thermodynamic parameters. Some examples are given in Figures 1.1, 1.2, 1.3.
• Thermodynamic parameters are measurable macroscopic quantities associated with they

![Diagram](image)

Figure 1.1: Gas cylinder with moveable piston
A thermodynamic state is determined by a set of sufficient values of parameters in order to completely describe a system.

An equation of state is a functional relationship between the thermodynamic parameters of a system in equilibrium. An example is the equation of state for an ideal gas

\[ PV = Nk_B T. \]

A thermodynamic transformation is a change of state. If a system is initially in thermodynamic equilibrium, then this can only come about by changes in the external condition of the system.

A transformation is quasi-static if the change is so slow that at any point in time the system is approximately in equilibrium. It is reversible if the transformation reverses itself when the transformation retraces its history in time (i.e. the system returns to the initial state). A reversible transformation must be quasi-static, but a quasi-static transformation is not necessarily reversible. Some examples of transformation from state X to X':

Examples of irreversible transitions:

- X=egg
  X'=broken egg!
- Heat any isolated system.
  X=before power turned on
  X'=after power turned on
- X=egg white (it is transparent). X' is state after heating, when egg white is white. (Egg protein molecules become denatured, i.e. they unfold).
Figure 1.4: Mixing of gases $N_2$ and $O_2$ by removing partition

Figure 1.5: Heating of water

Example of a reversible transition:

Figure 1.6: Transformation is reversible if system is thermally isolated and piston is moved slowly enough (i.e. process is quasi-static.)

- The concept of work is taken over from mechanics. For a system with parameters $P$, $V$, and $T$ the work done on the system is
  \[ dW = -PdV. \]

- Heat is absorbed by a system if its temperature increases yet no work is done. The heat capacity $C$ is defined by
  \[ \Delta Q = C\Delta T \]
  where $\Delta Q$ is the heat absorbed and $\Delta T$ is the change in temperature. The heat capacity depends on how the system is heated, e.g. at constant volume or constant pressure. Heat capacities per unit mass or per mole are called specific heats, and are denoted by a lower case $c$. 

11
A heat reservoir (or simply reservoir) is a system so big that any loss or gain in heat does not change its temperature.

A system is thermally isolated if no heat exchange can take place between it and the rest of the world. Any transformations that occur without the transfer of heat are said to be adiabatic.

Extensive thermodynamic quantities are proportional to the amount of substance that makes up a system. Examples are particle number \( N \), volume \( V \). Intensive quantities are independent of the amount of substance, e.g. temperature \( T \), pressure \( P \).

From this point we move on to the laws of thermodynamics. The historical development of these is quite interesting, and if you have the time you should read the popular book by H. C. von Baeyer, *Warmth disperses and time passes: the history of heat*, library QC318.M35 V66 1999. However, here we will make do with a practical summary.

The laws of thermodynamics may be considered as mathematical axioms defining a mathematical model. It is possible to deduce the logical consequences of these, and it remarkable how much can be derived from very simple assumptions. If you are interested in this approach try reading H.A. Buchdahl, *Twenty Lectures on Thermodynamics* (Perma gon, 1975), library QC311.B9155 1975.

It should be remembered that this model is purely empirical, and does not rigorously correspond to the real world. In particular it ignores the atomic structure of matter, so it will inevitably fail in the atomic domain. However, in the macroscopic world thermodynamics is extremely powerful and useful. It allows one to draw both precise and far reaching conclusions from a few seemingly commonplace observations.

### 1.2 The Second Law and Entropy

A good many times I have been present at gatherings of people who, by the standards of the traditional culture, are thought highly educated and who have with considerable gusto been expressing their incredulity at the illiteracy of scientists. Once or twice I have been provoked and have asked the company how many of them could describe the Second Law of Thermodynamics. The response was cold: it was also negative. Yet I was asking something which is about the scientific equivalent of: Have you read a work of Shakespeare’s?


We can order all the states of a system according to whether they are adiabatically accessible to one another.

\[
X \rightarrow X' \rightarrow X''
\]

This ordering defines an empirical entropy function, \( S(X) \). The second law can be written:

The entropy of the final state of an adiabatic transition is never less than that of the initial state i.e. \( dS \geq 0 \).
N.B. $S(X)$ is not uniquely defined. If $f(S)$ is a monotonically increasing function then $f(S(X))$ is equally appropriate as an empirical entropy function. It doesn’t require the definition of heat or temperature.

### 1.2.1 Statements of the second law

**Kelvin:** There exists no thermodynamic transformation whose *sole* effect is to extract a quantity of heat from a given heat reservoir and to convert it entirely into work.

**Claussius:** There exists no thermodynamic transformation whose *sole* effect is to extract a quantity of heat from a colder reservoir and deliver it to a hotter reservoir.

The key word in both of these statements is *sole*!

### 1.3 The First Law, Internal Energy, and Heat

In a mechanical system with conservative forces the energy of the equilibrium state only depends on the co-ordinates $x_1, x_2, x_3, \ldots, x_N$ and not on the history of the system. This allows definition of a potential energy function $V(x_1, x_2, x_3, \ldots, x_N)$

#### The First Law of Thermodynamics

*The amount of work done by a system in an adiabatic transition depends only on the co-ordinates of the final state.*

As a result we can define an *Internal Energy Function*, $U$, which has the property that the work $W_0$ done by a system in an adiabatic transition is equal to the decrease $-\Delta U$ of its internal energy.

$$0 = W_0 + \Delta U$$

Suppose the system $K$ is no longer adiabatically isolated and undergoes a transition from state $X$ to $X'$. In general, we expect $W_0 + \Delta U \neq 0$. The new quantity

$$Q = W + \Delta U$$

is called the *heat* absorbed by $K$.

For an infinitesimal transformation the first law can be written

$$dU = dQ + dW$$  \hspace{1cm} (1.1)$$

It should be remembered that $dQ$ and $dW$ considered individually are inexact differentials — i.e. they depend on the path of the transformation. However, $dU$ itself is an exact differential, which means that it has a unique value for a particular thermodynamic state.
A trivial consequence of the first law is that no heat is exchanged in an adiabatic process.

Example:

Thermos of ice water with heating element, as shown in Figure 1.7
Initial state $X = \text{ice water}$
Final state $X' = \text{boiling water}$
Process 1: lid of thermos is on so system is adiabatically isolated. Electrical work done to get from $X$ to $X'$ does not depend on the time taken. Eg. apply a small current a long time or a large current for a small time or turn off the current every now and then.
Process 2: lid of thermos is off so the system is not adiabatically isolated. Work done does depend on history. Eg. if we wait a few hours before turning on the current the ice melts and less work is required compared to if we quickly turn on a large current to get from $X$ to $X'$.

1.4 The Zeroth Law

Two systems $K_A$ and $K_B$ are said to be in equilibrium with one another if when they are brought into thermal contact the states they are in do not change.

**The Zeroth Law of Thermodynamics**

\[ \sim \]

*If each of two systems are in equilibrium with a third system then they are in equilibrium with one another (Figure 1.8)*

This allows us to define an empirical temperature function, $t$. Like an empirical entropy function is any function that always increases, an empirical temperature function is any value that is the
same for two bodies when they are in thermal equilibrium with each other. This is a way of quantifying the zeroth law.

Let $K_A$ be a system with a single thermodynamic co-ordinate $z$. (eg. a fixed volume of gas with a pressure gauge), i.e. $K_A$ is a “thermometer”.

If $K_C$ (which has co-ordinates $x_1, x_2, ..., x_n$) is in equilibrium with $K_A$ we define the empirical temperature of $K_C$ to be:

$$t(x_1, ..., x_n) = z$$

Note, $t$ is not uniquely defined (hence different temperature scales) since $\theta(t)$ is appropriate where $\theta(x)$ is a monotonical function of $x$.

$U$ must be an increasing or decreasing function of $t$. By convention we take $\delta U/\delta t > 0$. It follows that $\delta S/\delta t > 0$.

### 1.5 Metrical Entropy and Absolute Temperature

Consider a system with 2 co-ordinates, $V$ and $P$, and internal energy $U = U(V, P)$, which undergoes an infinitesimal process. The first law requires that

$$dQ = dU - dW$$

$$= \left( \frac{\partial U}{\partial V} \right)_P dV + \left( \frac{\partial U}{\partial P} \right)_V dP + PdV$$

The entropy is $S = S(V, P)$ and so the entropy change is

$$dS = \left( \frac{\partial S}{\partial V} \right)_P dV + \left( \frac{\partial S}{\partial P} \right)_V dP$$

If the process is adiabatic (reversible) then

$$dQ = 0$$
The second law also requires

\[ dS = 0 \]

It follows that there must exist some function \( \lambda(V, P) \) such that

\[ dQ = \lambda dS. \]

**Proof**

Setting \( dQ = 0 \) in Eq. (1.3) requires

\[ dP = -\frac{1}{\left( \frac{\partial U}{\partial P} \right)_V} \left( P + \left( \frac{\partial U}{\partial V} \right)_P \right) dV \]

substituting this into Eq. (1.3) gives

\[ dS = \left[ \left( \frac{\partial S}{\partial V} \right)_P \left( P + \left( \frac{\partial U}{\partial V} \right)_P \right) \right] dV. \]

We can only have \( dS = 0 \) if

\[ \frac{P + \left( \frac{\partial U}{\partial V} \right)_P}{\left( \frac{\partial U}{\partial P} \right)_V} = \frac{\left( \frac{\partial S}{\partial P} \right)_V}{\left( \frac{\partial S}{\partial V} \right)_P} \]

It follows (from Equation 1.3) that

\[
\frac{dS}{\left( \frac{\partial S}{\partial P} \right)_V} = \left( \frac{\partial S}{\partial V} \right)_P dV + dP = \left( P + \left( \frac{\partial U}{\partial V} \right)_P \right) dV + dP = \frac{1}{\left( \frac{\partial U}{\partial P} \right)_V} dQ
\]

\[ \lambda(V, P) = \left( \frac{\partial U}{\partial P} \right)_V \]

We now show that \( \lambda \) can only depend on the empirical temperature, \( t \). Consider a system, \( C \), composed of two parts, \( A \) and \( B \), as shown in Figure 1.9.

For any infinitesimal process (\( \hat{S}_A \) and \( \hat{S}_B \) are the empirical entropy)

\[
\begin{align*}
    dQ_C &= dQ_A + dQ_B \\
    \lambda_C d\hat{S}_C &= \lambda_A d\hat{S}_A + \lambda_B d\hat{S}_B \\
    d\hat{S}_C &= \frac{\lambda_A}{\lambda_C} d\hat{S}_A + \frac{\lambda_B}{\lambda_C} d\hat{S}_B
\end{align*}
\]
Thus \( \tilde{S}_C \) can only be a function of \( \tilde{S}_A \) and \( \tilde{S}_B \). The same applies to \( \frac{\lambda_A}{\lambda_C} \) and \( \frac{\lambda_B}{\lambda_C} \). But, the only co-ordinate of \( B \) on which \( \lambda_A \) can depend is \( t \). This makes sense if

\[
\lambda_A = T(t) f_A(\tilde{S}_A) \\
\lambda_B = T(t) f_B(\tilde{S}_B) \\
\lambda_C = T(t) f_C(\tilde{S}_A, \tilde{S}_B)
\]

we define the metrical entropy as

\[
S_A = \int f_A(\tilde{S}_A) d\tilde{S}_A
\]

so that

\[
dQ_A = \lambda_A d\tilde{S}_A \\
= T f_A(S_A) d\tilde{S}_A \\
= T dS_A
\]

and

\[
S_C = S_A + S_B
\]

\( T(t) \) is the absolute temperature function. \( S \) is the metrical entropy and is characterised by its additivity (above). It is an extensive property, as are \( U \) and \( V \), i.e., if the system size is doubled

\[
V \rightarrow 2V, \quad S \rightarrow 2S, \quad U \rightarrow 2U
\]

Temperature and pressure are intensive properties, i.e., if the system size is doubled

\[
V \rightarrow 2V, \quad T \rightarrow T, \quad P \rightarrow P
\]

For an infinitesimal process

\[
dQ = T dS = dU + P dV
\]

Hence

\[
\frac{1}{T} = \left( \frac{\partial S}{\partial U} \right)_V
\]

### 1.6 The Third Law


Experimentally, one finds
1. All heat capacities, $C_V(T) = \left( \frac{\partial U}{\partial T} \right)_V$, go to zero as $T \to 0$, at least as fast as $T \to 0$.

2. It is not possible to cool any system to $T = 0$.

Consequences

A) The first condition ensures the convergence of the integral

$$S(T, V) - S(T = 0, V) = \int_0^T \frac{dQ}{T} = \int_0^T \left( \frac{\partial U}{\partial T} \right)_V \frac{dT}{T} = \int_0^T \frac{C_V(T)}{T} dT$$

Hence, we have $(\partial S/\partial T)_V \to 0$ as $T \to 0$.

B) Consider a reversible adiabatic process in which $T$ and $V$ vary.

$$0 = dS = \left( \frac{\partial S}{\partial T} \right)_V dT + \left( \frac{\partial S}{\partial V} \right)_T dV$$

Substituting in that $(\partial S/\partial T)_V = C_V(T)/T$, we see that

$$\Rightarrow \frac{dT}{dV} = \frac{-T}{C_V(T)} \left( \frac{\partial S}{\partial V} \right)_T$$

If $(\partial S/\partial V)_T$ is non-zero as $T \to 0$ then $(dT/dV) > 0$ is non-zero, so by increasing $V$ we can cool the system right the way to $T = 0$. As this is not possible experimentally, we must have $(\partial S/\partial V)_T \to 0$ as $T \to 0$.

These two points show that the gradient of the entropy tends to zero as $T \to 0$. This be summarised in the

**Third Law of Thermodynamics**

\[ \sim \]

*The entropy function has the same finite value for all states which have $T = 0$*

Because entropy is only defined up to an additive constant (like internal energy), we can set this equal to zero.

1.7 Free Energy Tends to Decrease


For an isolated system, the entropy tends to increase, i.e., the equilibrium state is such that $S(U, V, N)$ is a maximum. However, what about a system connected to a heat reservoir?
Consider a system, $K$, at constant temperature, $T$, with fixed $V$ and $N$, and in equilibrium with an environment (reservoir), $R$, which is also at fixed $V$ and $N$.

For an infinitesimal process

\[
\begin{align*}
    dS_{\text{total}} &= dS + dS_R \\
    dU_{\text{total}} &= 0 = dU + dU_R \\
    dS_{\text{total}} &= dS + \frac{1}{T} dU_R \\
    &= dS - \frac{dU}{T} \\
    &= -\frac{1}{T} (dU - T dS) \\
    &= -\frac{dF}{T}
\end{align*}
\]

and since $dS_{\text{total}} \geq 0$ then we have $dF \leq 0$.

i.e., at constant $T$ and $V$, $F$ tends to decrease. Thus the equilibrium state is such that $F(T, V, N)$ is a minimum.

Similarly, at constant $T$ and $P$, the Gibbs free energy

\[G(P, T) \equiv F + PV\]

tends to decrease.

### 1.8 Helmholtz Free Energy

Ref. Schroeder, p. 149-152.

\[
\begin{align*}
    F &= U - TS \\
    dF &= dU - T dS - S dT \\
    &= -S dT - P dV
\end{align*}
\]
Thus, $F = F(T, V)$.

Suppose that in a finite process heat, $Q$, is added and work, $W$, is done on the system at constant temperature

$$
\Delta F = \Delta U - T \Delta S
= Q + W - T \Delta S
$$

If no new entropy is created, i.e., the entropy of the environment does not increase $Q = T \Delta S$ and so $\Delta F = W$.

However, in general $Q \leq T \Delta S$, so

$$
\Delta F \leq W
$$

Thus, $F$ represents the `free` energy, i.e. that available to do work.

### 1.9 Gibbs Free Energy

We define the Gibbs free energy as

$$
G = U - T S + PV
$$

An infinitesimal change in the Gibbs free energy is

$$
dG = dU - T dS - S dT + P dV + V dP
= -S dT + V dP
$$

Thus the Gibbs free energy is a function of the temperature and pressure of the system.

In a finite process where heat $Q$ is added to the system and the total work done on the system is $W$ while the system is at constant temperature and pressure then

$$
\Delta G = \Delta U - T \Delta S + P \Delta V
= Q + W - T \Delta S + P \Delta V
$$

The work term in this equation includes the work done on the system by the environment and any work such as electrical or chemical work done on the system, so we let

$$
W = -P \Delta V + W_{\text{other}}
$$

The $-P \Delta V$ term in the work cancels out with the one in Eq. (1.6) leaving us with

$$
\Delta G = Q + W_{\text{other}} - T \Delta S
$$

As we saw with the Helmholtz free energy, $Q - T \Delta S$ is always less than zero, so we get the relationship

$$
\Delta G \leq W_{\text{other}}
$$

when the temperature and pressure are constant. The Gibbs free energy represents the energy available in the system for non-mechanical work, which includes electrical work, chemical reactions and other forms of work.
1.10 Phase Transitions

Ref. Schroeder, p. 166.

A phase of a system is a form it can take. For example, the phases of water are solid (ice), liquid (water) and gas (steam). Mercury has these three phases and also has a superconducting phase at low very temperatures. A phase diagram shows the equilibrium phase of the system as a function of the relevant parameters of the system. Figure 1.11 is the phase diagram of a ferromagnet. It shows the phase as a function of temperature and pressure. A phase transition occurs when an infinitesimal change in the system’s environment causes a discontinuous change in at least one of the system properties. Examples of phase transitions include melting ice, boiling water, mercury entering a superconducting state, and helium-4 entering a superfluid state. Phase transitions are classified by looking for discontinuities in the derivatives of the Gibbs free energy of the system. A system undergoing a zeroth order phase transition has a discontinuity in the Gibbs free energy. A system which undergoes a first order phase transition has discontinuities in entropy and volume, which are the first derivatives of the Gibbs free energy. Phase transitions which have discontinuities in the second derivatives of the Gibbs free energy such as the specific heat, are termed second order phase transitions. Second order and higher phase transitions are also known as continuous phase transitions.
Chapter 2

Microcanonical Ensemble

Key Concepts

- Microstates and ensembles
- Fundamental assumption
- Multiplicity function $\Omega$
- Stirling’s approximation
- $\Omega$ for a paramagnet
- Sharpness of $\Omega$
- $\Omega$ for an ideal gas
- $\Omega$ and entropy
- $\Omega$ and temperature

Reading

- Schroeder, ch. 2 & 3
- Kittel & Kroemer, ch. 1 & 2
- Callen, ch. 15
2.1 Microstates and Macrostates

Up until now we have only considered thermodynamic (or macroscopic) states of a system. These are defined by a small number of co-ordinates (e.g. $T$, $V$, $N$), and are known as macrostates.

A microstate is a full description of the microscopic (i.e. atomic) composition of a system. There are usually a huge number of microstates that are consistent with a single macrostate.

**Example.**

*A Classical Microstate:* Consider a system of $N$ classical particles inside a fixed volume $V$. The microstate is defined by the $6N$ co-ordinates ($\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, \ldots, \vec{q}_N, \vec{p}_N$) where $\vec{q}_i$ is the position and $\vec{p}_i$ the momentum of the $i$th particle.

*A Quantum Microstate:* Consider a set of $N$ harmonic oscillators, each with frequency $\omega$. The quantum state of each oscillator is defined by the quantum number $n$, corresponding to the energy level $(n + 1/2)\hbar\omega$. The microstate is defined by the set of $N$ integers $(n_1, n_2, \ldots, n_N)$. The total energy of the microstate is

$$U = \frac{N\hbar\omega}{2} + \sum_{i=1}^{N} n_i \hbar\omega.$$

An equilibrium state of a system is a system state in which the macrostate parameters are stationary in time.

In principle, if we could identify the microstate of a system at a point in time then it is possible to solve for the system’s state at any time in the future. However, if the system has a large number of constituents then this becomes an impossible task. It is much easier to proceed from a statistical point of view.

In order to make use of the concepts of probability, it is necessary to consider an ensemble of a very large number $N'$ of systems that are prepared in an identical manner. The probability of a particular event occurring is then defined with respect to this ensemble.

2.2 The Microcanonical Ensemble

Consider a closed system, i.e. one with a fixed internal energy, $U$, fixed volume, $V$, and a constant number, $N$, of particles. (The system is adiabatically isolated). All external parameters such as electric, magnetic, and gravitational fields are also fixed. The *microcanonical ensemble* consists of all possible microstates that are consistent with the constants ($U$, $V$, $N$, etc.).

If you remember anything from this course, you should remember the next statement:
Fundamental assumption of statistical mechanics

All microstates in the microcanonical ensemble are equally probable

Let \( \Omega = \Omega(U, V, N) \) be the number of microstates in the ensemble. \( \Omega(U, V, N) \) is called the multiplicity function. The probability of finding the system in a particular state \( j \) is

\[
P(j) = \frac{1}{\Omega},
\]

as all states are equally likely. Suppose we are interested in some macroscopic property \( X \) of the system such that \( X(j) \) is the value for the microstate \( j \). Then the ensemble average of \( X \) is

\[
\langle X \rangle = \sum_j P(j)X(j) = \sum_j \frac{1}{\Omega}X(j) = \frac{1}{\Omega} \sum_j X(j),
\]

because all microstates are equally.

2.2.1 Transitions Between Microstates and the Ergodic Hypothesis

You may be wondering how a system in the microcanonical ensemble moves between different microstates. For example — with the ideal gas, although the position co-ordinates of the particles change, the momenta do not (apart from collisions with the walls). Also: in a quantum mechanical system: if it is in an eigenstate, then the state is time independent.

The answer is that these are idealised examples, and any realistic system will have some sort of interaction between the constituents (in fact it must to reach thermodynamic equilibrium). These interactions will define a time scale for transitions between the microstates, but will barely affect the results we will derive from treating the systems as idealised.

At this point it is worth mentioning the ergodic hypothesis. In simple terms, it basically says that at equilibrium a time average of a macroscopic parameter can be replaced by an ensemble average:

\[
\langle y \rangle_{\text{ensemble}} = \frac{1}{\theta} \int_0^\theta y(t)dt,
\]

for \( \theta \to \infty \). This says that in a time \( \theta \) sufficiently many of the system microstates will be sampled that the result will be the same as the ensemble average over all microstates. This is not an obvious result, and has only been proven for some specially designed mathematical models. However, it seems that it does apply to many systems across physics.

2.3 Multiplicity Function for a Paramagnet

We now come to our first example. Consider a system of \( N \) spins. Each spin can be in the \( |\uparrow\rangle \) or \( |\downarrow\rangle \) quantum state. The system has \( 2^N \) possible quantum states.
Suppose there are $N_\uparrow$ spins in the $|\uparrow\rangle$ state, $N_\downarrow$ spins in the $|\downarrow\rangle$ state ($N = N_\uparrow + N_\downarrow$).

$$S \equiv N_\uparrow - N_\downarrow = N - 2N_\downarrow$$

$S$ is called the spin excess.

**Example.** $N = 3$

<table>
<thead>
<tr>
<th>Spin Excess ($S$)</th>
<th>Multiplicity ($\Omega$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+3$</td>
<td>$\uparrow\uparrow\uparrow$</td>
</tr>
<tr>
<td>$+1$</td>
<td>$\uparrow\uparrow\downarrow$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$\downarrow\uparrow\uparrow$</td>
</tr>
<tr>
<td>$-3$</td>
<td>$\downarrow\downarrow\downarrow$</td>
</tr>
</tbody>
</table>

In general there are

$$\frac{N!}{N_\uparrow!N_\downarrow!}$$

ways of making a state with particular spin excess $S$. This is because in total there are $N!$ ways of arranging all the particles, but because the spins are indistinguishable, we need to reduce this by the $N_\uparrow!$ identical ways of arranging the up spins and the $N_\downarrow!$ identical ways of arranging the down spins. In combinatorics, this is often written as

$$\binom{N}{N_\uparrow} = \frac{N!}{N_\uparrow!(N - N_\uparrow)!} = \binom{N}{N_\downarrow}$$

and is the number of different ways of choosing $N_\uparrow!$ of all $N$ particles to be spin up: “$N$ choose $N_\uparrow$”.

Writing this in terms of the spin excess

$$\Omega(N, S) = \frac{N!}{\left(\frac{1}{2}N + \frac{1}{2}S\right)!\left(\frac{1}{2}N - \frac{1}{2}S\right)!}$$

(2.1)

$\Omega(N, S)$ is the number of states having the same value of $S$.

**Note:**

1. $N$ and $S$ are the macroscopic thermodynamic variables. We can’t distinguish different microstates with the same $S$ using macroscopic measurements.

2. If the particles are spin half, in the presence of an external magnetic field, $\vec{B}$, the total energy of the system is

$$U = -\vec{M} \cdot \vec{B} = +g\mu_B \frac{1}{2}SB.$$  

where in this equation $g$ is the gyromagnetic ratio.
2.4 Sharpness of the Multiplicity Function for a Paramagnet

In this section we look at the relative probability of finding a particular macrostate with a spin excess $S$ in the absence of a magnetic field. In this situation all arrangements have $U = 0$, and by the fundamental assumption of statistical mechanics, all microstates are equally likely.

A simple analogy to this problem is the tossing of a $N$ two-sided coins, with the spin excess being equal to the number of heads less the number of tails. It is intuitively obvious that we expect the most likely result is $S = 0$. But how likely are we to observe macrostates with a particular spin excess?

The probability of a particular spin excess will be

$$P(N, S) = \frac{\Omega(N, S)}{2^N}.$$  

We want to work out an approximate representation of this function in the region of $S \approx 0$.

2.4.1 Mathematical Aside: Stirling’s Approximation

Ref: Schroeder, Appendix B3.

It is troublesome mathematically to deal with the factorials that appear in the multiplicity function for the paramagnet. However, there is a very useful approximation for $n!$ when $n \gg 1$. In this case we can approximate

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \quad (2.2)$$

$$\log(n!) \approx n \log n - n + \frac{1}{2} \log(2\pi n) \quad (2.3)$$

It is fairly simple to derive the $\left(\frac{n}{e}\right)^n$ factor

$$\log n! = \log \left[n(n-1)(n-2)\ldots1\right]$$
$$= \log (n) + \log (n - 1) + \log (n - 2) + \ldots \log (1)$$
$$\approx \int_0^n \log (x)dx$$
$$= (x \log (x) - x)\bigg|_0^n$$
$$= n \log (n) - n$$
$$= n \left(\log \left(\frac{n}{e}\right)\right)$$

$$n! = \left(\frac{n}{e}\right)^n$$

Generally, in statistical mechanics it is sufficient to take

$$\log (n!) \approx n \log (n) - n.$$
Sharpness of the Multiplicity Function continued

Taking the logarithm of Eq. 2.1 we have

\[ \log (\Omega) = \log (N!) - \log ((\frac{1}{2}N + \frac{1}{2}S)!) - \log ((\frac{1}{2}N - \frac{1}{2}S)!) \]

We now use Stirling’s approximation.

\[
\log(\Omega) = \left( \frac{1}{2}N + \frac{1}{2}S + \frac{1}{2}N - \frac{1}{2}S \right) \log N - N \\
- \left( \frac{1}{2}N + \frac{1}{2}S \right) \log \left( \frac{1}{2}N + \frac{1}{2}S \right) + \left( \frac{1}{2}N + \frac{1}{2}S \right) \\
- \left( \frac{1}{2}N - \frac{1}{2}S \right) \log \left( \frac{1}{2}N - \frac{1}{2}S \right) + \left( \frac{1}{2}N - \frac{1}{2}S \right) \\
= - \left( \frac{1}{2}N + \frac{1}{2}S \right) \log \left( \frac{1}{2}N + \frac{1}{2}S \right) - \left( \frac{1}{2}N - \frac{1}{2}S \right) \log \left( \frac{1}{2}N - \frac{1}{2}S \right)
\]

Now

\[
\log \left( \frac{1}{2}(1 + \frac{S}{N}) \right) \simeq - \log 2 + \frac{S}{N} - \frac{S^2}{2N^2}
\]

(since \( \log (1 + x) = x - \frac{1}{2}x^2 + \ldots \) for \( x \ll 1 \))

Therefore

\[
\log (\Omega) \simeq N \log 2 - \frac{S^2}{2N}
\]

and finally

\[
\Omega(N, S) \simeq \Omega(N, 0) \exp \left( -\frac{S^2}{2N} \right)
\]

This is a Gaussian distribution. A comparison of this approximate form of the multiplicity with the full version for some small values of \( N \) is given in Figure 2.1 — you can see that it is a pretty good fit, at least in the region where \( \Omega(N, S) \) is significant.

Width of the distribution

In zero magnetic field (where the energy of an up or down spin is degenerate), the average spin per particle can be calculated as follows:

\[
\frac{\langle S \rangle}{N} = \frac{1}{N} \frac{1}{\Omega} \sum_j S_j,
\]

where \( \Omega = \sum_j = 2^N \) is the total number of microstates. Because for every positive value of \( S_j \) there is a corresponding negative value, we find that

\[
\frac{\langle S \rangle}{N} = 0,
\]
Figure 2.1: Representation of the sharpness of the multiplicity. The bars are the exact values of the multiplicity for the given spin excess, and the solid line is the Gaussian approximation to it. Note the relative narrowing of the distribution for larger $N$. 
as we predicted earlier.

The variance of the distribution is given by

$$\langle \Delta S \rangle^2 = \langle S^2 \rangle - \langle S \rangle^2,$$

and a calculation leads to $\langle S^2 \rangle = N$. Therefore the fractional width of the distribution is

$$\frac{\langle \Delta S \rangle}{N} = \frac{\sqrt{N - 0}}{N} = \frac{1}{\sqrt{N}}.$$

This is reasonably large when $N$ is small. However, the larger $N$ is, the smaller is the fractional width, and by the time $N = 10^{20}$ it is extremely small! (See the example in Figure 2.1). Note also that $\Omega(N, S)$ is reduced to $e^{-1}$ of its maximum value when $S/N = \pm 1/\sqrt{2N}$

There are two important conclusions that can be drawn from this.

Firstly: For a macroscopic system in equilibrium, the chance of seeing a macrostate significantly different from the most probable is exceedingly small. In fact, random fluctuations away from the most likely macrostate are utterly unmeasurable. Once thermal equilibrium has been reached so that all microstates are equally likely, we may as well assume the system is in its most likely macrostate. The limit that measurable fluctuations away from the most likely macrostate never occur is called the thermodynamic limit.

Secondly: we have our first clue about the irreversibility observed in everyday life. Imagine that the system started with all the spins pointing upwards. If we assume the system samples all possible microstates with equal probability, the trend would be overwhelming towards the macrostate that was the most likely. It turns out that the second law of thermodynamics is not really a fundamental law — it is just overwhelmingly probable.

Although the example we have looked at is for one particular system, it turns out that very similar results are obtained for all others.

### 2.5 Multiplicity Function for an Einstein solid

An Einstein solid is a collection of $N$ quantum harmonic oscillators that can share $q$ quanta of energy between them. Each oscillator could potentially have all units of energy — although it turns out that these particular microstates are extremely improbable!

We can graphically represent units of energy as $q$ solid dots in a row, with lines between the dots representing the divisions between the oscillators (it turns out we need $N - 1$ lines). Thus to represent a single microstate, we need $q + N - 1$ symbols, and we need to choose $q$ of them to be dots. It is then relatively easy to see that the multiplicity function is then just

$$\Omega(N, q) = \binom{q + N - 1}{q} = \frac{(q + N - 1)!}{q!(N - 1)!}.$$  \hspace{1cm} (2.4)

You will make use of this result in the tutorial and assignment problems.
2.6 Multiplicity Function for a Classical Ideal Gas

Consider a system of $N$ non-interacting classical particles contained in a volume $V$. The energy of a single particle is its kinetic energy,

$$ U_1 = \frac{1}{2M} (p_x^2 + p_y^2 + p_z^2) \tag{2.5} $$

where $M$ is the mass of the particle. The number of microstates, $\Omega_1$, available to a single particle system is proportional to the number of possible values of $(q_1, p_1)$

$$ \Omega_1 \propto V \cdot V_P $$

where $V_P \equiv \text{Volume}$ of possible $p_1$ values. From the constraint in Equation (2.5) we see that $V_P$ is proportional to the area of the surface of a 3-dimensional sphere of radius $\sqrt{2MU_1}$

$$ V_P \sim 4\pi (\sqrt{2MU_1})^2 $$

For the case of $N$ particles we have the constraint

$$ U = \frac{1}{2M} \sum_{i=1}^{N} \vec{p}_i \cdot \vec{p}_i $$

where $\vec{p}_i = (p_x^i, p_y^i, p_z^i)$. This defines a hypersphere, in $3N$ dimensional space, with radius $r = \sqrt{2MU}$. The multiplicity is

$$ \Omega(U, V, N) = \frac{1}{N!} \left( \frac{V}{h^3} \right)^N A_{3N} $$

where $\frac{1}{N!}$ allows for indistinguishable particles, $h^3$ is inserted to make $g$ dimensionless (almost), and $A_{3N} = \text{area}$ of the hypersphere.

In Appendix B.4, Schroeder shows that the surface 'area' of a d-dimensional 'hypersphere' of radius $r$ is

$$ A_d(r) = \frac{2\pi^{d/2} r^{d-1}}{\Gamma(d/2)} $$

$\Gamma(z)$ is the gamma function, $\Gamma(n+1) = n!$. Since $N \gg 1$ we can write

$$ \Omega(U, V, N) \sim \frac{1}{N! h^{3N}} \left( \frac{V}{\left(3N/2\right)!} \right)^N (2MU)^{3N/2} \equiv f(N) V^N U^{3N/2} $$

2.7 Multiplicity and Entropy

We now show that the multiplicity is an entropy function. Consider a composite system with total energy, $U$ as shown in Figure 2.2
\( \Omega(U) = \) multiplicity of the composite system
\( \Omega_1(U_1) = \) multiplicity of system 1.

From the definition of \( \Omega \) we can write

\[
\Omega(U) = \Omega_1(U_1)\Omega_2(U - U_1)
\]

provided we can neglect the effects of the interface.

Suppose we consider the following process:

Initial state: 2 systems isolated (Figure 2.3)

\[
\Omega_i(U) = \Omega_1(U_{10})\Omega_2(U - U_{10})
\]

Final state: 2 systems brought into thermal contact (Figure 2.4). This is an irreversible transition of an adiabatically isolated system.

\[
\Omega_f(U) = \sum_{U_1} \Omega_1(U_1)\Omega_2(U - U_1)
\]

This sum includes the term \( \Omega_i = \Omega_1(U_{10})\Omega_2(U - U_{10}) \). Thus

\[
\Omega_f(U) \geq \Omega_i(U)
\]
hence we can use $\Omega(U)$ as an empirical entropy function.

Note: $\log(\Omega(U))$ has the additivity property that the absolute (metric) entropy does, i.e. consider a composite system $C$ composed of two components $A$ and $B$, then

$$\Omega_C = \Omega_A \cdot \Omega_B$$

$$\log \Omega_C = \log \Omega_A + \log \Omega_B$$

This led Boltzmann to identify $\log \Omega(U, V)$ with the (metrical) entropy $S(U, V)$.

### 2.8 Multiplicity and Temperature

Suppose systems 1 and 2 are in thermal equilibrium and the system undergoes a reversible adiabatic process

$$\Omega(U) = \sum_{U_1} \Omega_1(U_1) \Omega_2(U - U_1)$$

Reversible $\Rightarrow \Omega = 0$.

Adiabatic $\Rightarrow dU = 0 \Rightarrow dU_1 = -dU_2$.

$$0 = d\Omega = \sum_{U_1} \frac{\partial \Omega_1}{\partial U_1} \Omega_2 dU_1 + \frac{\partial \Omega_2}{\partial U_2} dU_2$$

$$= \sum_{U_1} \Omega_1 \Omega_2 dU_1 \left( \frac{1}{\Omega_1} \frac{\partial \Omega_1}{\partial U_1} - \frac{1}{\Omega_2} \frac{\partial \Omega_2}{\partial U_2} \right)$$

$$\Rightarrow \frac{1}{\Omega_1} \frac{\partial \Omega_1}{\partial U_1} = \frac{1}{\Omega_2} \frac{\partial \Omega_2}{\partial U_2}$$

N.B. The left hand side depends only on system 1, while the right hand side depends only on system 2, therefore both must be equal to a constant. Since 1 and 2 are in thermal equilibrium

$$\frac{1}{\Omega} \frac{\partial \Omega}{\partial U} = \frac{\partial \log \Omega(U)}{\partial U}$$

is an empirical temperature. It is not necessarily the absolute temperature, but just a quantity that is the same for the two systems in equilibrium.

However, remember

$$\frac{1}{T} = \left( \frac{\partial S}{\partial U} \right)_V$$
where $T$ is the absolute temperature, $S$ is the metrical entropy. Hence, we identify $S$ with $\log g$ and write

$$S(U, V) = S = k_B \log \Omega(U, V)$$

This is the Boltzmann entropy. At this stage we cannot say what $k_B$ is, later we will see that it must be Boltzmann’s constant. The microcanonical ensemble turns out to be a very cumbersome way of calculating thermodynamic properties and only works for a very limited number of model systems. We will shortly consider a much more powerful method . . .

The Boltzmann entropy can be defined for both classical and quantum systems. The books by Schroeder and Kittel & Kroemer can be somewhat misleading: they only define it for quantum systems.
Chapter 3

Canonical Ensemble

Key Concepts

- The Boltzmann factor
- Partition function, $Z$
- Relation of $Z$ to $F$
- Classical statistical mechanics
- Classical ideal gas
- Gibbs paradox
- Maxwell speed distribution
- Equipartition of energy
- Quantum ideal gas
- Sackur-Tetrode equation

Reading

- Schroeder, ch. 6
- Kittel & Kroemer, ch. 3
- Callen, ch. 16
3.1 The Boltzmann Factor

In the microcanonical ensemble we have been considering isolated systems that cannot exchange energy with their environment. Now we want to consider systems in contact with a reservoir at a fixed temperature $T$.

To begin with, let’s find the probability, $P(\epsilon)$, that a system, $X$, will be in a state with energy, $\epsilon$. We assume that the number of particles in $R$ and $X$ are both separately fixed (there is no particle exchange between $R$ and $X$).

\[
\begin{array}{|c|c|}
\hline
\text{Total system} & \text{Reservoir} \\
K & R \\
\text{Energy} = U_0 = \text{constant} & \text{Energy} = U_0 - \epsilon \\
T & \\
\hline
\end{array}
\]

We suppose $X$ is in thermal equilibrium with a much larger system, that we refer to as the reservoir $R$. The common temperature is $T$.

Consider two possible microstates, 1 and 2, for the system $X$, with energies, $\epsilon_1$ and $\epsilon_2$, respectively. In state 1 the number of states accessible to $K$ is

\[
\Omega_{R+X_1} = \Omega_R \Omega_{X_1} = \Omega_R(U_0 - \epsilon_1)
\]

since $\Omega_{X_1} = 1$ — we are looking at a single microstate. A similar expression holds for state 2.

By the fundamental assumption of statistical mechanics:

All accessible states are equally probable for a closed system, and so

\[
P(\epsilon_1) = \frac{\Omega_R(U_0 - \epsilon_1)}{\Omega_K(U_0)}, \quad P(\epsilon_2) = \frac{\Omega_R(U_0 - \epsilon_2)}{\Omega_K(U_0)},
\]

and therefore

\[
\frac{P(\epsilon_1)}{P(\epsilon_2)} = \frac{\Omega_R(U_0 - \epsilon_1)}{\Omega_R(U_0 - \epsilon_2)} = \exp \left[ \frac{1}{k_B} (S_R(U_0 - \epsilon_1) - S_R(U_0 - \epsilon_2)) \right]
\]

since $S_R(U) = \text{entropy of } R = k_B \ln \Omega_R(U)$.

We now assume $\epsilon_1, \epsilon_2 \ll U_0$ so we may perform a Taylor expansion. (This is quite reasonable in the limit that the system is very much smaller than the reservoir.)

\[
S_R(U_0 - \epsilon) = S_R(U_0) - \epsilon \left( \frac{\partial S}{\partial U} \right)_{U=U_0} + \ldots
\]

36
But \(1/T = (\partial S/R/\partial U)_V\), and so
\[
\frac{P(\epsilon_1)}{P(\epsilon_2)} = \frac{\exp(-\epsilon_1/k_BT)}{\exp(-\epsilon_2/k_BT)}
\]
The term \(\exp(-\epsilon/k_BT)\) is known as the **Boltzmann factor**. Thus
\[
P(\epsilon) = \frac{\exp(-\beta \epsilon)}{Z}
\]
where \(\beta \equiv 1/k_BT\), and \(\sum_s P(\epsilon_s) = 1\) (the index \(s\) runs over all possible states, \(s\), of \(S\)).

\[
Z = \sum_s \exp(-\beta \epsilon_s)
\]

\(Z\) is called the **Partition Function** (\(Z\) stands for Zustandsmumme, which is German for “state sum”). From \(Z = Z(V,T)\) we can calculate thermodynamic functions!

### 3.2 Internal Energy

By definition
\[
U = \sum_s \epsilon_s P(\epsilon_s)
\]
\[
= \frac{1}{Z} \sum_s \epsilon_s \exp(-\beta \epsilon_s)
\]
\[
= \frac{1}{Z} \frac{\partial}{\partial \beta} Z
\]
\[
= -\frac{\partial}{\partial \beta} \ln Z(V,T)
\]
\[
U(V,T) = k_BT^2 \frac{\partial}{\partial T} \ln Z(V,T)
\]

### 3.3 Helmholtz Free Energy

It’s often more convenient to work with the Helmholtz free energy which can be calculated from \(Z\) as follows

\[
F(V,T) = U - TS
\]
\[
dF = -SdT - PdV
\]

so \(S = -(\partial F/\partial T)_V\), \(P = -(\partial F/\partial V)_T\)

\[
U = F + TS
\]
\[
= F - T \left(\frac{\partial F}{\partial T}\right)_V
\]
\[
= -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T}\right)
\]
But \( U = k_B T^2 (\partial \ln Z / \partial T)_V \), therefore

\[
\frac{\partial}{\partial T} \left( k_B \ln Z + \frac{F}{T} \right)_V = 0
\]

Integrate this

\[
\frac{F}{T} = -k_B \ln Z + \alpha(V)
\]

where \( \alpha(V) \) is the integration constant. We now find this constant:

\[
S = -\left( \frac{\partial F}{\partial T} \right)_V = -\frac{\partial}{\partial T}(-k_B T \ln Z + \alpha(V)T) = \frac{\partial}{\partial T}(k_B T \ln Z) - \alpha(V)
\]

Now, as \( T \to 0 \), we can approximate \( Z \to \Omega_0 \exp(-\epsilon_0/k_B T) \). The reason we can do this is that \(-\epsilon_0/k_B T \to -\infty\), meaning \( \exp(-\epsilon_0/k_B T) \to 0 \), so only the term with the smallest \( \epsilon \) is significant. The factor \( \Omega_0 \) is the degeneracy of the ground state term with energy \( \epsilon_0 \). Thus

\[
k_B T \ln Z \to k_B T \left[ \ln \Omega_0 - \frac{\epsilon_0}{k_B T} \right] = k_B T \ln \Omega_0 - \epsilon_0
\]

and \( S(T) \to k_B \ln \Omega_0 - \alpha(V) \)

By our earlier definition of \( S = k_B \ln \Omega \) we must have \( \alpha = 0 \). Therefore, we have derived the dependence of the Helmholtz free energy on the partition function.

\[
F(V, T) = -k_B T \ln Z(V, T)
\]

Using this we can also calculate the pressure directly from the partition function

\[
P = k_B T \frac{\partial}{\partial V} (\ln Z(V, T))
\]

### 3.4 Classical Statistical Mechanics

For a system of \( N \) particles, each of mass, \( M \), the microstate of the system is defined by the positions \((\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_N)\) and momenta \((\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N)\) of the particles. If the particles only interact via conservative forces the dynamics is described by a Hamiltonian

\[
H = \sum_{i=1}^{N} \frac{1}{2M} \vec{p}_i^2 + V(\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_N)
\]

The partition function is

\[
Z(V, T) = \int \prod_{i=1}^{N} \frac{d^3 p_i d^3 q_i}{h^3} \exp \left( -\frac{1}{k_B T} H(\vec{q}_i, \vec{p}_i) \right)
\]
The integral of \( d^3 p_i \) is from \(-\infty\) to \(+\infty\) (in each cartesian direction), the integral of \( d^3 q_i \) is over the volume, \( V \). \( h \) is an undetermined constant (with dimensions of action) required to make \( Z \) dimensionless.

It was introduced by Gibbs in 1875 and turns out to be Planck’s constant!

### 3.5 Classical Ideal Gas

Consider a set of \( N \) non-interacting particles of mass, \( M \), confined to a volume, \( V \). The energy of the state \( \Gamma = (\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N, \vec{q}_1, \vec{q}_2, \ldots, \vec{q}_N) \) is

\[
E_s(\Gamma) = \frac{1}{2M} \sum_{i=1}^{N} p_i^2
\]

The partition function is

\[
Z = \sum_{\Gamma} \exp(-\beta E_s)
= \sum_{\Gamma} \exp(-\beta \sum_{i=1}^{N} p_i^2 / 2M)
= \prod_{i=1}^{N} \exp(-\beta \vec{p}_i^2 / 2M)
\]

Then we convert the sum to an integral of all values of position and momentum, covering all possible states.

\[
Z = \prod_{i=1}^{N} \frac{1}{h^3} \int d^3 p_i \int d^3 q_i \exp \left( -\frac{1}{2Mk_B T} \vec{p}_i^2 \right)
= z^N
\]

where

\[
Z = \frac{1}{h^3} \int d^3 q \int d^3 p \exp \left( -\frac{1}{2Mk_B T} \vec{p}^2 \right)
= V \left[ \int_{-\infty}^{\infty} dp \exp \left( -p^2 / 2Mk_B T \right) \right]^3
= V \left( \frac{2\pi M k_B T}{h^2} \right)^{3/2}
\]

Note:

1. In going from the first line to the second in the above calculation each component, \( p_i \), contributes equally so that it suffices to calculate the integral for one component and then cube it.
2. \( \int_{-\infty}^{\infty} dx \exp(-\alpha x^2) = \sqrt{\pi/\alpha} \)

Continuing from the partition function:

\[
\ln Z = N \ln z \\
= N \left[ \ln V + \frac{3}{2} \left( \ln(k_B T) + \ln \left( \frac{2\pi M}{\hbar^2} \right) \right) \right]
\]

as before

\[
P = -\left( \frac{\partial F}{\partial V} \right)_T = k_B T \frac{\partial}{\partial V} \left( \ln Z(V, T) \right) \\
= \frac{N k_B T}{V} \\
PV = N k_B T
\]

Compare the ideal gas equation: \( PV = nRT \).

This means that the ideal gas constant, \( R = N_A k_B = 8.315 \text{ J mol}^{-1} \text{ K}^{-1} \) where \( N_A \) is the number of atoms in a mole of gas. This was first determined by Perrin from Brownian motion to be \( N_A = 6.02 \times 10^{23} \) (see Section 13.5).

Hence, Boltzmann’s constant is \( k_B = 1.38 \times 10^{-23} \text{ J K}^{-1} \).

The internal energy is given by

\[
U = k_B T^2 \frac{\partial}{\partial T} \ln Z \\
= k_B T^2 \cdot N \cdot \frac{3}{2} \frac{1}{T} \\
= \frac{3}{2} N k_B T = \frac{3}{2} nRT
\]

**Equipartition of energy**: There is an energy of \( k_B T/2 \) associated with each degree of freedom.

\[
C_V = \frac{3}{2} N k_B
\]

**Note**: None of the above results depend on the mass \( M \) of the particles or the type of gas.

### 3.6 Equipartition Theorem

Ref: *Schroeder*, section 6.3

Suppose that there is a degree of freedom \( q \), which has a quadratic term in the classical Hamiltonian,

\[
H = H_0 + \frac{k}{2} q^2
\]
where $H_0$ is the Hamiltonian due to all other degrees of freedom and is independent of $q$. Then, $q$ contributes $k_B T / 2$ to the internal energy.

**Proof:**

\[
Z = Z_0 \int_{-\infty}^{\infty} dq \exp \left( -\frac{kq^2}{2k_B T} \right) = Z_0 \left( \frac{2\pi k_B T}{k} \right)^{1/2}
\]

\[
U(T) = k_B T^2 \frac{\partial}{\partial T} (\ln Z_0 + \frac{1}{2} \ln \left( \frac{2\pi k_B T}{k} \right)) = U_0(T) + \frac{1}{2} k_B T
\]

**Example:**

A classical harmonic oscillator has

\[
H(p, q) = \frac{p^2}{2M} + \frac{M\omega^2}{2} q^2
\]

and so has internal energy $k_B T$. Note the problem of the temperature dependence in Fig. 3.1.

![Figure 1.13](image)

**Figure 1.13.** Heat capacity at constant volume of one mole of hydrogen (H$_2$) gas. Note that the temperature scale is logarithmic. Below about 100 K only the three translational degrees of freedom are active. Around room temperature the two rotational degrees of freedom are active as well. Above 1000 K the two vibrational degrees of freedom also become active. At atmospheric pressure, hydrogen liquefies at 20 K and begins to dissociate at about 2000 K. Data from Woolley et al. (1948).

**Figure 3.1: Heat Capacity at Constant Volume for H$_2$ (Fig. 1.13, Schroeder)**

This was first noted by Maxwell in 1859 and brought up again by Jeans in 1890. It shows that classical physics cannot explain everything!
3.7 Maxwell Speed Distribution

Ref: *Schroeder*, sect. 6.4.

The probability of a gas particle having momentum \( p_x \) in the \( x \)-direction is proportional to

\[
\exp \left( -\frac{1}{2Mk_BT}p_x^2 \right) = \exp \left( -\frac{M}{2k_BT}v_x^2 \right)
\]

By normalising we find

\[
P(v_x) = \left( \frac{M}{2\pi k_BT} \right)^{1/2} \exp \left( -\frac{M}{2k_BT}v_x^2 \right)
\]

so that

\[
\int_{-\infty}^{\infty} P(v_x)dv_x = 1
\]

In spherical polar co-ordinates

\[
dv_xdv_ydv_z = v^2d^2\Omega dv
\]

where \( v = \text{speed} = (v_x^2 + v_y^2 + v_z^2)^{1/2} \)

Normalisation requires

\[
\int_0^{\infty} dv P(v) = 1
\]

The probability density for finding a particle with speed, \( v \), is

\[
P(v) = 4\pi \left( \frac{M}{2\pi k_BT} \right)^{3/2} v^2 \exp \left( -\frac{Mv^2}{2k_BT} \right)
\]

Figure 3.2: The Maxwell speed distribution
Some typical speeds (‘rms’ is “root mean square”):

Most likely speed: Set \( \frac{dP}{dv} = 0 \) \( \Rightarrow \) \( v_{\text{max}} = \sqrt{\frac{2k_BT}{M}} \)

Mean speed \( = \int_{0}^{\infty} dv \, vP(v) \) \( \Rightarrow \) \( \bar{v} = \sqrt{\frac{8k_BT}{\pi M}} \)

Mean square speed \( = \bar{v}^2 = \int_{0}^{\infty} dv \, v^2P(v) \) \( \Rightarrow \) \( v_{\text{rms}} = \sqrt{\frac{3k_BT}{M}} \)

### 3.8 Maxwell’s Demon

References:

H. C. von Baeyer, *Warmth disperses and time passes: a history of heat*


*Feynman lectures on physics*, vol. 1, ch. 46.

![Figure 3.3: Cartoon of Maxwell’s demon in the process of reducing entropy](image)

Proposed by Maxwell in 1867 to demonstrate the statistical nature of the second law of thermodynamics.

The current consensus is that the Demon was not "killed" until 1982. Charles Bennett showed that the erasure of the demons memory is what produces entropy and "saves" the second law from being violated.
3.9 Gibbs Paradox

Let’s calculate the entropy of a classical gas

\[ S = -\left( \frac{\partial F}{\partial T} \right)_V \]
\[ = \frac{\partial}{\partial T}(k_B T \ln Z) \]
\[ = k_B \ln Z + k_B T \cdot \frac{3}{2} \cdot N \cdot \frac{1}{T} \]
\[ \frac{S}{k_B} = N \left( \frac{3}{2} + \ln \left[ V \left( \frac{2\pi M k_B T}{h^2} \right)^{3/2} \right] \right) \]

Difficulties:

1. \( S \) is not extensive. If \( V \to 2V \) and \( N \to 2N \), then we want \( S \to 2S \), but it does not.

2. \( S(T) \) violates the third law of thermodynamics since as \( T \to 0 \), \( S(T) \to -\infty \)

3. Gibbs Paradox

Take the system and insert a partition in the middle.

\[ \Delta S = S_f - S_i \]
\[ = \left( 2 \left( \frac{N}{2} \ln \left( \frac{V}{2} \right) \right) - N \ln V \right) k_B \]
\[ = -Nk_B \ln 2 \]

Hence we have decreased the entropy and violated the second law of thermodynamics!

3.10 Resolution of the Gibbs Paradox

The above derivation of \( Z(V, T) \) for an ideal gas neglected the fact that the particles are indistinguishable. We should write

\[ Z_N = \frac{1}{N!} \gamma_N \]

As a result

\[ F \quad \rightarrow \quad F_{old} + k_B T \ln N! \]
\[ S \quad \rightarrow \quad S_{old} - k_B \ln N! \]
\[ \approx S_{old} - k_B (N \ln N - N) \]
where we have used Stirling’s approximation. We have now derived the Sackur-Tetrode equation (for the entropy of a classical gas).

\[ S = N k_B \left( \frac{5}{2} \ln \left( \frac{V}{N} \left( \frac{2\pi M k_B T}{\hbar^2} \right)^{3/2} \right) \right) \]

**Note:**

1. This is extensive, so that if we make the replacements \( N \rightarrow 2N \), \( V \rightarrow 2V \) then this results in \( S \rightarrow 2S \). If we partition the system the entropy does not change.

2. There is still the problem that the entropy becomes negative at low temperatures - see tutorial problem.

3. Prior to its theoretical derivation the Sackur-Tetrode equation was derived empirically from experimental data.

### 3.11 Creating Entropy

For the ideal gas

\[ S = N k_B \left( \frac{5}{2} + \ln \left( \frac{V}{N} \right) + \frac{3}{2} \ln T + \text{const.} \right) \]

In general we can increase \( S \) by increasing \( N \), \( V \) or \( T \). If we fix \( N \) and \( V \) and heat the system the entropy increases. However heat is not necessary to increase entropy. Consider the following case of free expansion of a gas into a vacuum (see Figure 3.4).

![Figure 3.4: Free expansion of gas into a vacuum](image)

- Remove the partition
- No heat flows into or out of the gas, \( Q = 0 \)
- The gas does not push against anything when it expands, \( W = 0 \)
- Therefore \( \Delta U = Q + W = 0 \)
- But the entropy of the gas increases. Thus heat is not necessary for an entropy increase.
3.12 Quantum Statistical Mechanics

\[ \exp(-F/k_BT) = Z = \sum_i \exp(-E_i/k_BT) \]

where \(E_i\) are the quantum energy levels of the system.

3.13 Quantum Ideal Gas

Consider a particle of mass, \(M\), in a cube of dimensions \(L \times L \times L\). Schrödinger’s equation is:

\[ -\frac{\hbar^2}{2M} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi(x, y, z) = E \Psi(x, y, z) \]

The boundary conditions are

\[
\begin{align*}
    \Psi(0, y, z) &= \Psi(L, y, z) = 0 \\
    \Psi(x, 0, z) &= \Psi(x, L, z) = 0 \\
    \Psi(x, y, 0) &= \Psi(x, y, L) = 0
\end{align*}
\]

The solution is

\[ \Psi(x, y, z) = A \sin\left(\frac{n_x \pi x}{L}\right) \sin\left(\frac{n_y \pi y}{L}\right) \sin\left(\frac{n_z \pi z}{L}\right) \]

Where \(A\) is a normalisation constant, and \(n_x, n_y, n_z\) are positive integers \(\rightarrow\) quantum numbers.

\[ E = \epsilon(n_x, n_y, n_z) = \frac{\hbar^2}{2M} \left( \frac{\pi}{L} \right)^2 (n_x^2 + n_y^2 + n_z^2) \]

The partition function for a single particle is

\[ z = \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \exp\left(-\epsilon(n_x, n_y, n_z)/k_BT\right) \]

The spacing of the energy levels is

\[ \sim \frac{\hbar^2}{2M} \left( \frac{\pi}{L} \right)^2 \ll k_BT. \]

For sufficiently large systems we can take the continuum limit (i.e., treat the \(n\)’s as continuous variables,

\[ z = \int_0^\infty dn_x \int_0^\infty dn_y \int_0^\infty dn_z \exp\left(-\alpha(n_x^2 + n_y^2 + n_z^2)\right) \]

where

\[ \alpha \equiv \frac{(\pi\hbar)^2}{2ML^2k_BT} \]

We may write this as

\[ z = z^3 \]
where
\[ z = \int_{0}^{\infty} dn \exp (-\alpha n^2) \]
\[ = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}} \]
(c.f. derivation of the partition function for a classical ideal gas)
\[ z = n_Q V \]
\[ V \equiv L^3, \text{ and} \]
\[ n_Q \equiv \left( \frac{Mk_B T}{2\pi \hbar^2} \right)^{3/2} = \frac{1}{\lambda_T^3} \]

Since, using the de Broglie relation, \( p = \hbar/\lambda \)
\[ \langle KE \rangle \sim k_B T = \frac{p^2}{2M} = \frac{\hbar^2}{2M \lambda_T^2} \]
\[ \Rightarrow \lambda \sim \sqrt{\frac{\hbar^2}{Mk_B T}} \]
\( \lambda_T \) is called the (thermal) de Broglie wavelength of a particle with energy \( k_B T \)

### 3.14 Identical Particles

Consider a system with two particles, each of which occupies a separate energy level. Up until now, if we interchange the two particles then this leads to two different terms in the partition function. This is correct as long as the particles are distinguishable: e.g. a helium atom and a neon atom.

However, if the particles are identical they should only be counted once. If there are distinct single particle states
\[ Z_i^N = \left( \sum_i e^{-\beta \epsilon_i} \right)^N \]
then each entry occurs \( N! \) times. This is overcounting! We should write
\[ Z_N = \frac{1}{N!} Z_i^N \]

**Note:**
We are implicitly assuming that there are many more energy levels than particles, i.e. we are not considering the possibility of more than one particle in a level. Later, we will relax this assumption when we consider fermions and bosons.
3.15 Quantum Ideal Gas — continued

\[ Z_N = \frac{1}{N!} (n_Q V)^N \]
\[ P = k_B T \frac{\partial}{\partial V} \ln Z_N = \frac{N k_B T}{V} \]

To calculate the entropy:

\[ S = - \left( \frac{\partial F}{\partial T} \right) \]
\[ = \frac{\partial}{\partial T} (k_B T \ln Z) = k_B \ln Z + k_B T \frac{\partial}{\partial T} \ln Z_N \]

now using \( Z_N = (n_Q V)^N / N! \) and Stirling’s approximation

\[ \ln Z_N = N \ln (n_Q V) - \ln N! \]
\[ \frac{\partial}{\partial T} \ln Z_N = N \frac{\partial}{\partial T} \ln n_Q = \frac{3 N}{2 T} \]

Putting this into the above formula for the entropy, we again have the Sackur-Tetrode equation

\[ S = N k_B \left( \frac{5}{2} + \ln \left( \frac{n_Q V}{N} \right) \right) \]

Comparing the classical and quantum expressions we now see that they agree if the \( h \) in the classical expression is Planck’s constant!
Chapter 4

Paramagnetism

Key Concepts

- Magnetisation of magnetic materials
- Two state paramagnet
  - microcanonical treatment
  - canonical treatment
- $2J + 1$ state paramagnet
- Brillouin function
- Cooling by adiabatic demagnetisation

Reading

D.V. Schroeder, Section 3.3 and p.145-146


Ashcroft and Mermin, Solid State Physics, Ch. 31

Demonstrations

19-22 Paramagnetism and Diamagnetism
19-23 Dysprosium in Liquid Nitrogen
4.1 Magnetisation

Electrons have a magnetic moment \( m \) due to their spin and due to their orbital angular momentum.

For a system of \( N \) electrons or ions the magnetisation is defined as

\[
M = \left\langle \sum_{i=1}^{N} m_i \right\rangle
\]

where \( \langle a \rangle \) denotes a thermal average.

4.1.1 Magnetic Materials

1. **Paramagnets**: \( M = \chi B \) where \( \chi > 0 \)
   i.e. the moments are parallel to an external magnetic field \( B \)
   *Examples*: electrons in a metal, ions in an insulator where the electron shell is partially full.

2. **Diamagnets**: \( M = \chi B \) where \( \chi < 0 \)
   i.e. \( m \) and \( B \) are anti-parallel.
   *Examples*: ions with full electronic shells in an insulator, superconductors.

3. **Ferromagnet**: \( M \neq 0 \) when \( B = 0 \), for \( T < T_c \)
   \[\uparrow\uparrow\uparrow\uparrow\uparrow\]
   Moments form domains, all aligned in the same direction below the critical temperature \( T_c \)
   (also known as the Curie temperature). Above \( T_c \) there is no spontaneous magnetisation.
   *Examples*: Fe, Co, Ni.

4. **Antiferromagnets** \( M = 0 \) but \( \langle m_i \rangle \neq 0 \) for \( B = 0 \)
   \[\uparrow\downarrow\uparrow\downarrow\]
   Adjacent moments are anti-aligned, leading to no overall magnetisation.
   *Examples*: MnO, FeO.

4.2 Non-interacting two-state paramagnet

Consider a system of \( N \) spin-half particles (e.g. electrons) that are localised in the presence of a constant magnetic field \( B \), pointing in the \( z \)-direction. Because they are localised the particles are fixed in position and are hence distinguishable.

The effect of the field is to tend to align the magnetic dipole moment \( m \) of each particle parallel to the field.
The Hamiltonian for a single moment is

\[ H = -\mathbf{m} \cdot \mathbf{B} \]
\[ = -m_z B_z \]
\[ = -mB \]

since we take \( B \) to point in the \( z \)-direction. The energy is lower when the magnetic moment is parallel to the magnetic field (\(-mB\) as opposed to \(mB\)).

For spin-half particles, \( m_z \) can have eigenvalues \( \pm \mu \) so the allowed energy levels are

\[ E = \pm \mu B \]

For electrons,

\[ \mathbf{m} = -g_0 \mu_B \mathbf{S} \]

where \( \mathbf{S} \) is the spin operator, and \( g_0 \) is the \textit{g-factor} or \textit{gyromagnetic ratio}. Its value has been calculated very precisely in QED, and is \( g_0 = 2.002319 \). The Bohr magneton \( \mu_B \) is given by

\[ \mu_B \equiv \frac{e \hbar}{4\pi m_e}, \]
\[ = 9.27 \times 10^{-24} \text{J/T}, \]
\[ = 5.79 \times 10^{-5} \text{eV/T}. \]

The total energy of the system is

\[ U = \mu B(N_\uparrow - N_\downarrow) = \mu B(N - 2N_\uparrow) \quad (4.1) \]

where \( N = N_\uparrow + N_\downarrow \), \( N_\uparrow \) is the number of particles with spin up, and \( \mu = g_0 \mu_B / 2 \). The magnetisation \( M \) is the total magnetic moment of the whole system

\[ M = \mu(N_\uparrow - N_\downarrow) = -\frac{U}{B} \quad (4.2) \]

### 4.2.1 Microcanonical Ensemble

We now find \( M \) as a function of \( B \) and the temperature \( T \) using the microcanonical ensemble.

The multiplicity of the system is

\[ \Omega(U, N) = \frac{N!}{N_\uparrow! N_\downarrow!} \]

as stated in Section 2.3. The entropy is

\[ \frac{S(U, N)}{k_B} = N \ln N - N_\uparrow \ln N_\uparrow - (N - N_\uparrow) \ln(N - N_\uparrow) \quad (4.3) \]

51
where we have used Stirling’s approximation. The temperature is given by

\[
\frac{1}{T} = \left( \frac{\partial S}{\partial U} \right)_{N,B} = \frac{\partial N_\uparrow}{\partial U} \frac{\partial S}{\partial N_\uparrow} = -\frac{1}{2\mu B} \left( \frac{\partial S}{\partial N_\uparrow} \right).
\] (4.4)

Now using (4.3) gives

\[
\frac{1}{k_B} \left( \frac{\partial S}{\partial N_\uparrow} \right) = -\ln N_\uparrow - 1 + \ln(N - N_\uparrow) + 1
= \ln \left( \frac{N - N_\uparrow}{N_\uparrow} \right)
= \ln \left( \frac{N_\downarrow}{N_\uparrow} \right)
\] (4.5)

So that from Equations 4.4 and 4.5

\[
\frac{N_\downarrow}{N_\uparrow} = \exp \left( -\frac{2\mu B}{k_B T} \right).
\]

To find the mean magnetic moment \(M/N\), we use Equation 4.2 and \(N = N_\uparrow + N_\downarrow\),

\[
\frac{M}{N} = \frac{\mu(N_\uparrow - N_\downarrow)}{N_\uparrow + N_\downarrow},
= \frac{\mu(1 - N_\downarrow/N_\uparrow)}{1 + N_\downarrow/N_\uparrow},
= \mu \left[ \frac{1 - \exp(-2\mu B/k_B T)}{1 + \exp(-2\mu B/k_B T)} \right].
\]

Hence,

\[
M = \mu N \tanh \left( \frac{\mu B}{k_B T} \right).
\] (4.6)

As you should be coming familiar with in this course, the first thing that we do is check the high and low temperature limits.

**High temperature / low field:** We have \(\mu B \ll k_B T\). Since \(\tanh x \simeq x\) for \(x \ll 1\) then,

\[
\frac{M}{N B} = \frac{\mu^2}{k_B T}
\]

which is **Curie’s law** that we came across in the tutorials.

Notice as \(T \to \infty, M \to 0\), because the electrons are so energetic that the difference between the energies of the spin up and spin down states becomes negligible, and their spins are randomly aligned.
**Low temperature / high field:** We have $\mu B \gg k_B T$. In this case

$$M = \mu N$$

i.e. all the moments are parallel to the field.

The internal energy (again from Equation 4.2) is given by

$$U(T) = -MB,$$

$$= -N \mu B \tanh \left( \frac{\mu B}{k_B T} \right). \quad (4.7)$$

Heat capacity in a constant field is

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

$$= \frac{Nk_B}{\cosh^2\left( \frac{\mu B}{k_B T} \right)} \quad (4.8)$$

These are shown in Figure 4.1. The peak in the $C_B(T)$ versus $T$ curve is known as the **Schottky anomaly**, as we have seen already in tutorials.

### 4.2.2 Canonical ensemble

We now recalculate our result using the canonical ensemble. This calculation will give you an idea about why it is much preferred to the microcanonical!

The partition function

$$Z(B, T) = z(B, T)^N$$

where $z$ is the partition function for a single magnetic moment. Note here that we do **not** have to divide by $N!$ as we did for the ideal gas — this is because the magnetic moments are **distinguishable**.

The partition function for a single moment is particularly simple, as there are only two possible energy states. We have looked at such a system already in tutorials.

$$z(B, T) = \sum_s e^{-\beta E_s},$$

$$= e^{-\beta \mu B} + e^{\beta \mu B},$$

$$= 2 \cosh \left( \frac{\mu B}{k_B T} \right).$$
Figure 4.1: Magnetisation and heat capacity for a two-state paramagnet. Note that you can get negative temperatures for this system. See Appendix E, Kittel and Kroemer for further details. This is also mentioned in Schroeder, pg 101-2.
The internal energy is

\[ U(T) = k_B T^2 \left( \frac{\partial \ln Z}{\partial T} \right)_B, \]

\[ = N k_B T^2 \frac{\partial}{\partial T} \ln \left[ \cosh \left( \frac{\mu B}{k_B T} \right) \right], \]

\[ = -N k_B T^2 \cdot \frac{\mu B}{k_B T^2} \frac{\sinh (\mu B/k_B T)}{\cosh (\mu B/k_B T)}, \]

\[ = -N \mu B \tanh \left( \frac{\mu B}{k_B T} \right). \]

This agrees with the microcanonical ensemble. Notice that this is a lot quicker and easier than the microcanonical formulation.

### 4.3 Insulators of Ions with Partially Full Shells

In solids containing ions such as Fe\(^{3+}\), Fe\(^{2+}\), Cr\(^{2+}\) and Mn\(^{4+}\), the electrons in the outermost atomic orbitals (usually d-orbitals) have spin \( S \) and orbital angular momentum \( L \).

The state is completely specified by \( L, S \), the total angular momentum vector \( J = L + S \) and the magnetic quantum number \( M_J \) which can take \( 2J + 1 \) possible values in the range of \(-J, -J + 1, \ldots, J - 1, J\). (This is very similar to the azimuthal angular momentum \( m \) in a hydrogen atom, only now we consider the total angular momentum.)

Hund’s rules are used to determine the basic electron configuration of an atom or ion. These are described in detail in Chapter 31 of Ashcroft and Mermin, however if you prefer the internet to the library you can find them at [http://hyperphysics.phy-astr.gsu.edu/hbase/atomic/hund.html](http://hyperphysics.phy-astr.gsu.edu/hbase/atomic/hund.html)

In brief, Hund’s rules say that for the ground state, the electrons in the outer shell will arrange themselves to have

1. The maximum value of the total spin \( S \) allowed by the exclusion principle.

2. The maximum value of the total angular momentum \( L \) consistent with this value of \( S \).

3. The value of the total angular momentum \( J \) is equal to \( |L - S| \) when the shell is less than half full, and to \( L + S \) when the shell is more than half full. When the shell is half full the first rule gives \( L = 0 \) so \( J = S \).

The first rule has its origin in the exclusion principle and the Coulomb repulsion between two electrons. The second — is best accepted from model calculations! The third is a consequence of the spin-orbit interaction — for a single electron the energy is lowest when the spin is anti-parallel to the orbital angular momentum.
It can be shown that in a magnetic field \( B = B \hat{z} \) that the magnetic moment of an ion with quantum numbers \( J, L, S, M_J \) is
\[
m = -g(JLS)\mu_B M_J
\]
and that the allowed energies are (similar to the spin-half case)
\[
E(M_J) = -mB = g(JLS)\mu_B M_J B
\]
where the Landé g-factor of the ion is
\[
g(JLS) = \frac{1}{2}(g_0 + 1) - \frac{1}{2}(g_0 - 1) \frac{L(L+1) - S(S+1)}{J(J+1)}
\]
\( g_0 = 2.002319 \) is the electron g-factor.

**Example**

For \( \text{Fe}^{3+} \), one can calculate that \( L = 0 \) and \( S = \frac{5}{2} \), related to the fact that Hund’s rule says that each electron should be placed in a separate state with parallel spins, and \( \text{Fe}^{3+} \) has 5 electrons in its outer orbital, each with spin-\( \frac{1}{2} \). Hence \( J = \frac{5}{2} \), and \( g = g_0 \) since \( J = S \).

### 4.3.1 Partition function

In the same manner as for the two state paramagnet
\[
Z(B) = \sum_s e^{-\beta E_s} = z(B)^N
\]
where \( z(B) \) is the partition function for a single magnetic moment and we assume they do not interact. To save ourselves some writing, we define
\[
a = \frac{g\mu_B B}{k_BT}.
\]
There are \( 2J + 1 \) different energy levels so
\[
z(B) = \sum_{M_J=-J}^{M_J=J} \exp(-aM_J),
\]
\[
= e^{-aJ} + e^{-a(J-1)} + \cdots + e^{a(J-1)} + e^{aJ},
\]
\[
= e^{-aJ} \left( 1 + e^a + e^{2a} + \cdots + e^{2aJ} \right).
\]
We now use the fact that
\[
\sum_{n=0}^{N-1} x^n = \frac{1 - x^N}{1 - x}.
\]
We write the partition function sum as

\[
\begin{align*}
   z(B) &= e^{-aJ} \sum_{n=0}^{2J} (e^a)^n, \\
   &= e^{-aJ} \frac{1 - e^{(2J+1)a}}{1 - e^a}, \\
   &= e^{-aJ} - e^{(J+1)a}, \\
   &= e^{a/2} \left( e^{a/2} - e^{a/2} \right), \\
   &= \frac{\sinh((J + 1/2)a)}{\sinh(a/2)}.
\end{align*}
\]

4.3.2 Magnetisation

If the magnetic field is in the z-direction, the magnetic moment in the state \( |JM_J \rangle \) is \(-g\mu_B M_J\) as stated at the start of this section.

The fractional magnetisation in thermal equilibrium is

\[
\frac{M}{N} = \sum_{M_J} (-g\mu_B M_J) \frac{1}{z} e^{-\beta E(M_J)}
\]

\[
= -g\mu_B \sum_{M_J} \frac{M_J}{z} \exp \left( - \frac{g\mu_B M_J B}{k_B T} \right)
\]

\[
= k_B T \frac{1}{z} \left( \frac{\partial z}{\partial B} \right)_T
\]

\[
M = k_B T N \frac{\partial}{\partial B} (\ln z)
\]

\[
= k_B T N \frac{\partial a}{\partial B} \frac{\partial}{\partial a} \left[ \ln \sinh \left( \frac{(J + 1/2)a}{2} \right) - \ln \sinh \left( \frac{a}{2} \right) \right]
\]

and thus

\[
M = Ng\mu_B \left[ (J + \frac{1}{2}) \coth \left( \frac{(J + 1/2)a}{2} \right) - \frac{1}{2} \coth \left( \frac{a}{2} \right) \right] \quad \text{(4.10)}
\]

We can write

\[
M = Ng\mu_B J B_J(a) \quad \text{(4.11)}
\]

where

\[
B_J(a) = \frac{1}{J} \left[ (J + \frac{1}{2}) \coth \left( \frac{(J + 1/2)a}{2} \right) - \frac{1}{2} \coth \left( \frac{a}{2} \right) \right] \quad \text{(4.12)}
\]

is known as the Brillouin function. Although not immediately obvious, for \( J = \frac{1}{2} \) and \( g = g_0 \) it reduces to the spin-half result. As always, let’s check the limits:

**Low temperature / high field:**

\[
a = \frac{g\mu_B B}{k_B T} \gg 1, \quad \coth \left( \frac{a}{2} \right) \rightarrow 1, \quad \coth \left( \frac{(J + 1/2)a}{2} \right) \rightarrow 1
\]

57
and so we have

\[ M \rightarrow N g \mu_B J \]

i.e. all the magnetic moments are aligned parallel to the field as expected.

**High temperature / low field**

\( a \ll 1 \), and for small \( x \ll 1 \),

\[ \coth x = \frac{1}{x} + \frac{x}{3} + \ldots \]

This means

\[
\frac{M}{N g \mu_B} \approx \left( J + \frac{1}{2} \right) \left( \frac{1}{(J + \frac{1}{2})a} + \frac{1}{3} \left( J + \frac{1}{2} \right) a \right) - \frac{1}{2} \left( \frac{2}{a} + \frac{a}{6} \right)
\]

\[
= \frac{1}{a} - \frac{1}{a} + \frac{a}{3} \left( (J + \frac{1}{2})^2 - \frac{1}{4} \right)
\]

\[
= \frac{1}{3} J(J + 1)a
\]

So

\[
M = N(g \mu_B)^2 \frac{J(J + 1)B}{3k_B T}
\]

which is Curie’s law again.

The Brillouin function can be tested experimentally on crystals such as

\[ \text{Fe}_2(\text{SO}_4)_3\cdot(\text{NH}_4)_2 \text{SO}_4 \cdot 24\text{H}_2\text{O} \]

which have crystal structures such that the magnetic ions (in this case Fe\(^{3+}\)) are very isolated from one another and so the magnetic dipoles interact very weakly with one another. A comparison between experimental results and the Brillouin function is shown in Fig. 4.2.

### 4.3.3 Internal Energy and Entropy

\[ U(T) = -MB \]

We also have

\[ U(T) = k_B T^2 \frac{\partial}{\partial T} \ln Z(B, T) \]

As before, we can show

\[ F \equiv U - TS = -k_B T \ln Z(B, T) \]

Thus,

\[ S = -\left( \frac{\partial F}{\partial T} \right)_B \]
Figure 4.2: Plot of magnetic moment versus $B/T$ for spherical samples of (I) potassium chromium alum, (II) ferric ammonium alum, and (III) gadolinium sulfate octahudrate. Over 99.5% magnetic saturation is achieved at 1.3K and about 5T. The points are experimental results of W.E. Henry (1952) and the solid curves are graphs of the Brillouin function. [Taken from Introduction to Solid State Physics, C. Kittel, 6th ed. 1 kG = $10^3$ G, $10^4$ G = 1 T.]
and so
\[
\frac{S}{k_B} = \ln Z(B, T) + T \frac{\partial}{\partial T} \ln Z(B, T)
\]
\[
= \ln Z + \frac{U}{k_B T}
\]
\[
= \ln Z - MB \frac{\partial}{\partial B} \ln Z(a)
\]
\[
= \ln Z(a) - a \frac{\partial}{\partial a} \ln Z(a)
\]

So the entropy is an increasing function with \(a\):
\[
S(B, T) = k_B f \left( \frac{g\mu_B B}{k_B T} \right).
\] (4.14)
i.e., the entropy is only a function of the ratio \(B/T\). Hence, in an adiabatic process (\(S\) is fixed) and if \(B\) decreases then \(T\) must decrease too.

\[
z(a) = \frac{\sinh[(J + 1/2)a]}{\sinh(a/2)}
\]

\[
\frac{S}{k_B} = \ln (\sinh[(J + 1/2)a]) - (J + 1/2)a \coth[(J + 1/2)a]) - \ln[\sinh(a/2)] + \frac{a}{2} \coth(a/2)
\] (4.15)

As \(a \to 0 (T \to \infty)\):
\[
z(a) \to 2J + 1, \quad S \to k_B \ln(2J + 1).
\]
As \(a \to \infty (T \to 0)\):
\[
z(a) \to e^{Ja}, \quad S \to k_B (Ja - aJ) = 0.
\]

### 4.3.4 Cooling by adiabatic demagnetisation

The paramagnetic crystal to be cooled begins in thermal equilibrium with e.g. liquid helium.

1. The system is held at constant temperature \(T = T_i\). The magnetic field is increased to a field value \(B_i\).
2. The system is adiabatically isolated.
3. The field is slowly reduced to \(B_f\).

The entropy \(S(B/T)\) must remain constant in step 3. Hence, the final temperature is
\[
T_f = T_i \frac{B_f}{B_i} < T_i
\]
Figure 4.3: Cooling by adiabatic demagnetisation process

One might hope that by letting $B_f \rightarrow 0$ we can make $T_f \rightarrow 0$!

However, at sufficiently low temperatures, the interactions between the magnetic moments in the magnetic crystal become important and the relation $S = S(B/T)$ no longer holds.

The minimum temperature that can be reached is $T_{\text{min}} \sim T_{\text{int}}$ where $k_B T_{\text{int}}$ is comparable to the typical interaction energy between the magnetic moments.

This cooling method was first proposed in 1926 independently by Gianque and Debye. Using this method in 1933-1935, three different groups reached temperatures in the range 5 mK to 350 mK. The initial temperatures were $T_i \sim 1$ K obtained by rapidly evaporating He$^4$.

In 1956, Simon and Kurti reached $10^{-5}$ K using nuclear magnetic moments.
Chapter 5

Thermal Radiation

Key concepts

- The ultraviolet catastrophe
- The Planck distribution
- Summing over modes
- The Planck spectrum
- Blackbody radiation
- Cosmic background radiation

Reading

Schroeder, Section 7.4
Kittel & Kroemer, Ch 4

Demonstrations


24-18: Radiation spectrum of a hot object
5.1 Ultraviolet catastrophe

Consider classical electromagnetic waves confined to a box of fixed volume, $V = L^3$. There are an infinite number of modes of vibration for the electromagnetic waves, with wavelengths

$$\lambda_n = \frac{2L}{n} \quad n = 1, 2, 3, \ldots$$

By considering the total energy of the electromagnetic field, we can show that each mode behaves like a classical harmonic oscillator. Thus, by the equipartition theorem each mode must have an average energy of $k_B T$ (as there are two degrees of freedom). As there are an infinite number of modes, then the total electromagnetic energy in the box must also be infinite.

![Figure 5.1: Modes of an electromagnetic radiation within a box of width $L$](image)

However, when we open the oven door we don’t get blasted with an infinite amount of energy, so something must be wrong with this model. This is called the ultra-violet catastrophe. Why?

This represents a failure of classical statistical mechanics.

5.2 The Planck distribution

The solution to the ultraviolet catastrophe led to the birth of quantum mechanics (Planck, 1900). Suppose that an oscillator with frequency $\nu$ can only have discrete energies

$$E_n = n\hbar \nu, \quad n = 0, 1, 2, \ldots$$

where $\hbar$ is Planck’s constant. This is ignoring the zero point energy. The partition function for the single oscillator is

$$Z = \sum_{n=0}^{\infty} \exp(-n\hbar \nu/k_B T) = \frac{1}{1 - \exp(-\hbar \nu/k_B T)}.$$

64
The average energy is

\[
U = k_B T \frac{1}{Z} \frac{\partial Z}{\partial T} = \frac{\hbar \nu}{\exp(\hbar \nu / k_B T) - 1}
\]

The average number of quanta of energy \(\hbar \nu\) is then the total energy divided by \(\hbar \nu\)

\[
\bar{n}(\nu) = \frac{U}{\hbar \nu} = \frac{1}{\exp(\hbar \nu / k_B T) - 1}
\]

This is the Planck distribution.

For high frequency modes, \(\hbar \nu \gg k_B T\), \(\bar{n}(\nu) \simeq \exp(-\hbar \nu / k_B T) \ll 1\). These modes are “frozen out” and don’t contribute to the total energy of the system.

### 5.3 Summing over modes

We now find the total energy in the box using the Planck distribution. In one dimension the allowed wavelengths are

\[
\lambda_n = \frac{2L}{n}, \quad n = 1, 2, 3, \ldots
\]

The momenta of these wavelengths are

\[
p_n = \frac{h}{\lambda_n} = \frac{\hbar}{2L} n,
\]

and the energy of each mode is

\[
\varepsilon_n = p_n c = \frac{\hbar c}{2L} n.
\]

In three dimensions we have

\[
\varepsilon = |\mathbf{p}| c = c \left[ p_x^2 + p_y^2 + p_z^2 \right]^{1/2}, \tag{5.1}
\]

where \(p_x, p_y, p_z\) are the \(x, y,\) and \(z\) components of the momentum, and are given by

\[
p_x = \frac{\hbar}{2L} n_x, \quad p_y = \frac{\hbar}{2L} n_y, \quad p_z = \frac{\hbar}{2L} n_z, \quad n_x, n_y, n_z = 1, 2, 3, \ldots
\]

with

\[
\varepsilon_n = \frac{\hbar c}{2L} |\mathbf{n}| \quad \mathbf{n} = (n_x, n_y, n_z)
\]

Each mode (photon) can have two possible polarisations. The total energy is then

\[
U = 2 \sum_n \varepsilon_n \bar{n}(\varepsilon_n),
\]

\[
= \sum_n \frac{\hbar c}{L} |\mathbf{n}| \left[ \frac{1}{\exp(|\mathbf{n}| \hbar c / 2L k_B T) - 1} \right]
\]
Figure 5.2: The Planck distribution compared with the classical Rayleigh-Jeans law (as a function of wavelength, not energy!) Classical physics can account for the long-wavelength, low energy modes where there are many photons per mode. However, it predicts a UV catastrophe for short wavelengths.

For large $L$ we can convert the sum to an integral

$$
\sum_n \rightarrow \frac{1}{8} \times 4\pi \int_0^\infty n^2 dn,
$$

where we multiply by one eighth because all the $n_i$ must be positive, and so we only want one octant of the 3D space. This gives

$$
U = \frac{h c \pi}{2L} \int_0^\infty \frac{n^3 dn}{\exp(hcn/2Lk_B T) - 1}, \quad (5.2)
$$

If we make a change of variables to $\varepsilon = hcn/2L$ then we find the total energy density is

$$
\frac{U}{L^3} = \int_0^\infty \frac{8\pi \varepsilon^3 d\varepsilon}{(hc)^3 \exp(\varepsilon/k_B T) - 1},
$$

where

$$
u(\varepsilon) \equiv \frac{8\pi \varepsilon^3}{(hc)^3 \exp(\varepsilon/k_B T) - 1}, \quad (5.3)
$$

is the spectrum or energy density per unit photon energy, which is plotted in Fig. 5.2. It has a maximum at $\varepsilon = 2.82k_B T$ (Wien’s law).

For $\varepsilon \ll k_B T$, $u(\varepsilon) \sim \varepsilon^2 k_B T$, as expected classically.
For $\varepsilon \gg k_B T$, $u(\varepsilon) \sim \varepsilon^3 \exp(-\varepsilon/k_B T) \ll \varepsilon^2 k_B T$, and so the energy density of high energy radiation is suppressed, preventing the ultra-violet catastrophe.

By making another change of variables $x = \varepsilon/k_B T$, we can change the expression (5.3) into the form

$$\frac{U}{V} = \frac{8\pi (k_B T)^4}{(hc)^3} \int_0^\infty \frac{x^3 dx}{e^x - 1}$$  \hspace{1cm} (5.4)

The integral in Eqn. (5.4) evaluates to $\pi^4/15$ (see the Appendix of Schroeder). Note that $U/V \sim T^4$. Hence, doubling $T$ increases $U/V$ by a factor of 16.

Also the heat capacity $C_V(T) = (\partial U/\partial T)_V \sim T^3$, approaches zero as the temperature goes to zero, which is consistent with the third law of thermodynamics.

### 5.4 Blackbody radiation

Blackbody radiation comprises the photons emitted by any non-reflecting (“black”, i.e. perfectly absorbing) surface at a fixed temperature $T$. A blackbody emits the same photon spectrum as does a hole in a box at the same temperature.

![Figure 5.3: A hollow box and a blackbody in thermal equilibrium.](image)

We can see this by considering a blackbody and a hole in a box with equal area in thermal equilibrium. If the blackbody does not emit the same spectrum as the box, then there will be a net flow of energy from the one emitting more energy to the one emitting less energy. But since the two are in thermal equilibrium, there can be no net energy flow between the two objects.

This does not mean that the spectrums are the same — you could imagine a situation where the spectrum of each object was different, but there was no net energy flow. However, the argument can be extended as we can always insert a filter between the box and the blackbody that blocks a certain frequency range.

Imagine the two spectrums are not the same. This means that the energy being emitted for some frequency range will be different for the two objects. If we place a filter between the two objects
that blocks this range, it means that there will then be a net flow of energy, which violates our assumption of thermal equilibrium.

We can arbitrarily choose any filter, with any frequency range, which means that the radiation emitted by a blackbody must have the same spectrum as a box at the same temperature.

The power per unit area emitted by the hole, and so also the blackbody, is

\[ \frac{P}{A} = \frac{c U}{4V} \]

The factor of 1/4 comes from a somewhat involved geometric calculation that works out the total energy that escapes from the box in a time interval \( dt \) (see Schroeder, pg 300–2). Substituting Eqn. (5.4) into the formula for the power per unit area results in Stefan’s law

\[ \frac{P}{A} = \sigma T^4, \quad \sigma = \frac{2\pi^5 k_B^4}{15h^3 c^2}. \]

The Stefan-Boltzmann constant is \( \sigma = 5.67 \times 10^{-8} \) W m\(^{-2}\) K\(^2\). Stefan’s law was discovered empirically in 1879.

## 5.5 Cosmic background radiation

Cosmic background radiation was predicted in the late 1940’s by Alpher, Hermann and Gamow as being a remnant of the “big bang”. The cosmic background radiation are the photons which were last scattered off electrons 300,000 years after the beginning of the universe, when the temperature was approximately 3000 K. In the meantime the universe has expanded by about a factor of \( 10^3 \) in every direction, and this has caused the wavelengths of this radiation to be stretched — red-shifted, if you like. It was first observed accidentally in 1964 by Penzias and Wilson, who originally thought it was noise due to pigeon droppings in their radio telescope. For this discovery they received the Nobel prize in 1978.

The cosmic background radiation is incredibly uniform and isotropic. In 1989, the Cosmic Background Explorer (COBE), a NASA satellite, found that the cosmic background radiation has the spectrum of a blackbody with a temperature \( T = 2.728 \pm 0.002 \)K and the deviation from the Planck spectrum is less than 0.005%. In 1992 using the differential microwave radiometer on COBE temperature fluctuations of 30 \( \mu \)K on an angular scale of 10 degrees was found. This was important because it provided evidence for spatially inhomogeneous regions needed in the early universe to form galaxies.

Very recent results from the Wilkison Microwave Anisotropy Probe (WMAP) satellite have made a far more accurate measurement of the temperature fluctuations (see Fig. 5.5, 5.6), and these are consistent with the inflationary model of the early universe. The most precise calculation of the age of the universe was made using data from this satellite in 2003 — it gives an age of 13.7 \( \pm \) 0.2 billion years.

A summary of findings from the WMAP measurements (from the WMAP website):
• Universe is 13.7 billion years old, with a margin of error of close to 1%.

• First stars ignited 200 million years after the Big Bang.

• Light in WMAP picture is from 379,000 years after the Big Bang.

• Content of the Universe:
  - 4% Atoms, 23% Cold Dark Matter, 73% Dark Energy.
  - The data places new constraints on the Dark Energy. It seems more like a “cosmological constant” than a negative-pressure energy field called “quintessence”. But quintessence is not ruled out.
  - Fast moving neutrinos do not play any major role in the evolution of structure in the universe. They would have prevented the early clumping of gas in the universe, delaying the emergence of the first stars, in conflict with the new WMAP data.

• Expansion rate (Hubble constant) value: Ho = 71 (km/sec)/Mpc (with a margin of error of about 5%)

• New evidence for Inflation (in polarized signal)

• For the theory that fits our data, the Universe will expand forever. (The nature of the dark energy is still a mystery. If it changes with time, or if other unknown and unexpected things happen in the universe, this conclusion could change.)

Further reading

http://map.gsfc.nasa.gov/
Figure 5.4: Spectrum of the cosmic background radiation, as measured by the COBE satellite. The vertical axis is the energy density. Note that a frequency of $f = 3 \times 10^{11}$ s$^{-1}$ corresponds to a wavelength $\lambda = 1.0$ mm. The point by point uncertainties are too small to show up on this scale; the size of the squares estimates systematic uncertainties. The best fit temperature is $2.728 \pm 0.002$ K. Graph taken from Schroeder.
Figure 5.5: (Colour online) The first detailed, all-sky picture of the infant universe. Colors indicate "warmer" (red) and "cooler" (blue) spots. The oval shape is a projection to display the whole sky; similar to the way the globe of the earth can be projected as an oval. The WMAP image reveals 13 billion year old temperature fluctuations of only a few millionths of a degree that correspond to the seeds that grew to become the galaxies. Encoded in the patterns are the answers to many age-old questions, such as the age and geometry of the Universe. Image courtesy of the NASA/WMAP Science Team.

Figure 5.6: (Colour online) A close up of the temperature fluctuations in the image above. Image courtesy of the NASA/WMAP Science Team.