

VOL. 4 – THERMAL AND STATISTICAL PHYSICS

The Undergraduate Companion to Theoretical Physics

Andrea Kouta Dagnino[‡]

[‡]*Open University, Milton Keynes, UK.*

E-mail: k.y.dagnino@gmail.com

Contents

1	<i>Preliminaries</i>	4	4.5	Statistical interpretation of entropy	30
1.1	How much stuff?	4	4.5.1	Gibb's paradox	31
1.2	The thermodynamic limit	4	4.5.2	Maxwell's demon	31
1.3	Approximating large numbers	5	4.6	Gibbs expression for entropy	32
1.4	The ideal gas law	6	5	<i>Thermodynamic potentials</i>	33
1.5	What is heat	6	5.1	Definitions	33
1.6	Probability and statistics	7	5.2	Differentials	33
1.7	What is temperature	7	5.3	Integrals under constant variable	35
1.8	Partial differentiation	7	5.4	Legendre transformations	37
I	Classical Thermodynamics	8	5.5	Equilibrium conditions	38
2	<i>The First Law</i>	9	5.5.1	Constant pressure and temperature	39
2.1	Definitions	9	5.6	Connecting the potentials: Maxwell's relations	40
2.2	The first law of thermodynamics	10	6	<i>Rods, bubbles and gases</i>	45
2.3	Heat capacity	11	6.1	Elastic materials	45
2.4	Reversible processes	13	6.2	Surface tension	46
2.5	Isothermal expansion	14	6.3	Magnetization	48
2.6	Adiabatic expansion	15	II	Kinetic Theory of Gases	51
2.7	Isochoric and Isobaric expansions	16	7	<i>Speed distribution</i>	52
3	<i>The Second Law</i>	18	7.1	Energetic considerations	52
3.1	Perpetual motion machine	18	7.2	Pressure	53
3.2	The second law of thermodynamics	18	7.3	Deriving the Boltzmann velocity distribution	56
3.3	The Carnot Engine	19	7.4	Effusion	59
3.4	Carnot's theorem	20	7.5	Collisions	62
3.5	The equivalence of Clausius' and Kelvin's statements	22	8	<i>Heat transport phenomena</i>	63
3.6	Clausius' theorem	23	8.1	Particle distribution function	63
4	<i>Entropy</i>	26	8.2	Maxwellian equilibrium	64
4.1	Definition of Entropy	26	8.3	Conservation laws	65
4.2	Reformulating the First Law	28			
4.3	Joule Expansion	29			
4.4	Entropy change for ideal gas	30			

III Statistical Physics	67		
9 <i>The microcanonical ensemble</i>	68		
9.1 What is statistical physics? . . .	68		
9.2 The classical Phase Space . . .	68		
9.3 Probability density and expectation values	71		
9.4 Liouville's theorem	72		
9.5 Ergodic hypothesis and Equal a priori probability . .	73		
9.6 The microcanonical ensemble	74		
9.7 Relating the Boltzmann and Gibbs entropy	75		
9.8 Entropy of an ideal gas (classical)	76		
9.9 Gibb's paradox	78		
9.10 Entropy of an ideal gas (quantum)	80		
9.11 General procedure using the microcanonical ensemble . . .	83		
9.12 Ultra-relativistic gas	83		
9.13 Harmonic oscillators	87		
9.14 Two-state system	87		
10 <i>The canonical ensemble</i>	88		
10.1 Canonical phase-space density	88		
10.2 Important examples	91		
10.2.1 The ideal gas	91		
10.2.2 The ultra-relativistic gas	93		
10.2.3 The harmonic oscillator	94		
10.3 The equipartition and virial theorems	95		
11 <i>Canonical examples of the canonical ensemble</i>	99		
11.1 The Quantum Harmonic Oscillator	99		
11.2 The Diatomic gas	99		
11.3 The Paramagnet	99		
12 <i>The Grand Canonical Ensemble</i>	100		
12.1 Adding the chemical potential	100		
12.2 Euler's relation and the Gibbs-Duhem equation	101		
12.3 The Grand canonical phase space density	102		
12.4 The ideal gas in the GCE . . .	103		
		12.5 Particle number fluctuation .	103
		12.6 Mechanical stability condition	105
		IV Quantum statistical physics	107
		13 <i>Density operator methods</i>	108
		13.1 Making the jump to quantum	108
		13.2 Quantum statistical ensembles	110
		13.3 Particle in a box	111
		13.4 Identical particles	112
		13.5 Boson and fermion gases in the GCE	114
		14 <i>Bose-Einstein statistics</i>	116
		14.1 The density of states	116
		14.2 Photon gas - a statistical perspective	117
		14.3 Photon gas - a kinetic perspective	119
		14.4 Blackbodies and Kirchoff's law	121
		14.5 Phonons	121
		14.6 High temperature Bose gases	121
		14.7 Bose-Einstein condensation .	121
		15 <i>Fermi-Dirac statistics</i>	127
		15.1 The Ideal Fermi gas	127
		15.2 Zero temperature behaviour - Fermi energy	129
		15.3 The Sommerfield approximation	130
		15.4 Magnetism	133
		15.4.1 Landau diamagnetism	133
		15.4.2 Pauli paramagnetism .	134
		15.5 White Dwarf stars	134
		16 <i>Real gases and Interacting systems</i>	137
		16.1 The virial expansion	137
		16.2 The Second Virial coefficient .	137
		16.3 Van der Waals equation . . .	139
		16.4 The cluster expansion	140
		17 <i>Phase transitions</i>	146
		17.1 The Van der Waals gas-liquid phase transition	146
		17.2 Phase transitions in the Ising model	150

17.3 Landau-Ginzburg theory: a taster	155	V Statistical field theory	159
		19 <i>Path integrals</i>	160
		20 <i>Landau-Ginzburg theory</i>	161
18 <i>Diffusion and Brownian motion</i>	158	21 <i>Renormalisation group</i>	162
18.1 Langevin equation	158	<i>Bibliography</i>	163

*If you are out to describe the truth, leave
elegance to the tailor.*

— Ludwig Eduard Boltzmann



References

Several textbooks, online courses/resources were referenced heavily (to the extent of making this text completely unoriginal, yet hopefully helpful for revision) throughout the writing of these lecture notes. Using a typical bibliography (research paper style) would be a formidable task. Pinpointing exactly where each reference as been used is quite difficult for such a large and well-referenced subject, and would probably change the writing style to a far too formal one for lecture notes. Therefore we instead list the most relevant below giving a brief comment on which topics they were mostly used for:

- S. Blundell, K. Blundell *Concepts in Thermal Physics*
Fantastic introduction to thermodynamics. All of part I was typeset referencing this textbook heavily. I particularly enjoyed the extra special topics section with discussions of more contemporary advances in physics based on statistical mechanics.
- W. Greiner, L. Neise, H. Stocker *Thermodynamics and Statistical Mechanics*
This is a more advanced version of Blundell, covering the same topics but more in-depth. I used this mostly while writing parts III and IV.
- M. Kardar *Statistical Physics of Particles*
Comprehensive but advanced, good for a second review of the subject. It is supplemented by an equally fantastic lecture series on MIT OCW. It was used mostly in part IV on Quantum statistical physics.
- A. Schekhhovichin course notes on *A3 Statistical Physics*
Discusses kinetic theory and statistical physics from a more mathematical point of view. It was really helpful for part II.
- Course notes by D. Tong on *Part II Statistical Mechanics*
Excellent notes as usual, particularly recommended for statistical mechanics. It skips through thermodynamics at the beginning (indeed part II students taking the course will have already studied thermodynamics beforehand in part IB), but does dedicate an entire chapter to it later on.
- E. Terentjev *Part II Thermal and Statistical Physics*
Similar content to the notes by David Tong, but the lecture slides made the content really succinct and easy to digest while revising.

Preliminaries

1.1 How much stuff?

The branch of thermal physics often involves the statistical study of macroscopic systems, assemblies of many, many atoms. For example, the number of atoms in an average human is about a staggering 7×10^{27} , that's more than the number of sand particles on Earth (10^9 times more in fact). Therefore, we can't use words like dozen, billion, billion trillion, we must instead invent a name for large quantities of matter, the **mole**.

Definition 1: The mole

A **mole** is the quantity of matter equivalent to the number of atoms in 12 g of ^{12}C . It is equivalent to an **Avogadro number** N_A of atoms:

$$N_A = 6.022 \times 10^{23} \quad (1.1.1)$$

For example a mole of ping pong balls would be 6.022×10^{23} ping pong balls.

The **molar mass** of a substance is the mass of 1 mole of the substance. Carbon-12 therefore has a molar mass of exactly 12g ¹

1.2 The thermodynamic limit

Let us consider the following problem. We have 1 kg of nitrogen gas in a container, which is approximately 2×10^{25} molecules, and we want to model the motion of every single one of these. One year contains 3.2×10^7 seconds, so for a 3 GHz computer counting one molecule per cycle, it would count 9.5×10^{16} molecules in one year. So, to count all the molecules in 1kg of N_2 , it would take 0.2 billion years. That is just for counting the molecules, imagine what it would take to model their positions and velocities!

It is clear that we need to approach this problem from a more statistical perspective. Instead of making measurements on individual atoms, we can instead consider their average. For an infinitely large sample (10^{25} molecules is definitely a good approximation), we can then ignore any fluctuations which get smoothed out for a randomized sample. We refer to this limit of an infinitely large sample as the **thermodynamic limit**.

¹this is not the value found in periodic tables, which is around 12.01 g. Indeed, the periodic table provides the **relative atomic mass**, which is the weighted average of the molar masses for all isotopes.

Suppose the container of gas had a volume V , temperature T , pressure p and total kinetic energy U . If we then slice the container, these variables change to V', T', p', U' . How do these relate to their initial values? By definition, slicing the container in half implies its volume is halved:

$$V' = \frac{V}{2} \quad (1.2.1)$$

Since there are half as many particles the total kinetic energy will also be halved (assuming a randomized distribution of particles, which is clearly the case):

$$U' = \frac{U}{2} \quad (1.2.2)$$

The temperature however should not change, nor does pressure as we will see later:

$$p' = p \text{ and } T' = T \quad (1.2.3)$$

Definition 2: Intensive/Extensive variables

An **intensive** variable, such as p, T , does not scale with the size of the system.
An **extensive variable**, such as V, U does scale with the system size.

For example, consider combining 2 volumes of identical gases with state variables N, V, p, T each. The combined system will have state variables $2N, 2V, p, T$. Hence, volume and the number of particles are extensive, whereas pressure and temperature are intensive.

1.3 Approximating large numbers

We will often come across numbers much larger than N_A . Dealing with statistics and probability, we will often come across factorials, such as $10^{23}!$. In such cases, we may not need an exact value, which would be difficult to calculate, and an approximation would suffice.

To do so, we can use **Stirling's formula**:

$$\ln n! \approx n \ln n - n \quad (1.3.1)$$

In the case of $10^{95}!$, Stirling's formula gives:

$$\ln(10^{95}!) \approx 10^{95} \ln 10^{95} - 10^{95} \approx 2.2 \times 10^{97} \quad (1.3.2)$$

Therefore:

$$10^{95}! \approx \exp(2.2 \times 10^{97}) = 10^{9.55 \times 10^{94}} \quad (1.3.3)$$

which is a one followed by more zeros than the number of particles in the universe.

1.4 The ideal gas law

Experiments in the 18th century showed the relation between pressure, temperature and volume. Boyle's law showed that:

$$p \propto \frac{1}{V} \quad (1.4.1)$$

and Charles' law showed that:

$$V \propto T \quad (1.4.2)$$

and finally Gay-Lussac's law showed that:

$$pV \propto T \quad (1.4.3)$$

These three can be combined into the **ideal gas law**.

The Ideal Gas Law

For a system containing N molecules of an **ideal gas**:

$$pV = Nk_B T \quad (1.4.4)$$

where k_B is the **Boltzmann constant**, with value:

$$k_B \approx 1.38 \times 10^{-23} \text{JK}^{-1} \quad (1.4.5)$$

This equation makes several assumptions on the properties of the gas. Firstly, it assumes that a gas can be modelled as an assembly of tiny particles which can bounce off of each other and the walls of the container. Secondly, it assumes that the molecules are point-like, and that there are no intermolecular forces. It also models the gas molecules as non-relativistic.

Definition 3: Ideal Gas

An **ideal gas** is made of molecules of point size obeying Newtonian principles. They do not interact through intermolecular forces, and have no potential energy, and do not condense when cooled.

1.5 What is heat

Definition 4: Heat

Heat is defined as thermal energy in transit.

1.6 Probability and statistics

1.7 What is temperature

1.8 Partial differentiation

Suppose we have a set of variables x, y, z related through the constraint:

$$F(x, y, z) = 0 \iff x = x(y, z) \tag{1.8.1}$$

We may express this in differentials as:

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \tag{1.8.2}$$

Analogously one may also define:

$$z = z(x, y) \implies dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy \tag{1.8.3}$$

and thus substituting (1.8.3) into (1.8.2) we find that:

$$dx = \left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\left(\frac{\partial x}{\partial y}\right)_z + \left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial y}\right)_x\right) dy \tag{1.8.4}$$

If we choose x, y to be independent then $dy = 0$ yields:

$$\boxed{\left(\frac{\partial x}{\partial z}\right)_y \frac{\partial z}{\partial x} \Big|_y = 0 \iff \left(\frac{\partial x}{\partial z}\right)_y = \frac{1}{\left(\frac{\partial z}{\partial x}\right)_y}} \tag{1.8.5}$$

whereas $dx = 0$ yields:

$$\boxed{\left(\frac{\partial x}{\partial z}\right)_y \left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z = -1} \tag{1.8.6}$$

Part I

Classical Thermodynamics

The First Law

2.1 Definitions

A **system** is whatever part of the universe we choose to study. The **surroundings** are the regions near the system. A system is:

- (i) **Isolated**: if it does not interact with its surroundings. Here the energy, volume and particle number are all conserved.
- (ii) **Closed**: if it can only exchange energy with the surroundings. Volume and particle number are still conserved.
- (iii) **Open**: if it can interact with the surroundings
- (iv) **Homogeneous**: if it has the same properties everywhere throughout it
- (v) **Heterogeneous**: if it has discontinuities in its properties

The **phases** of a system are the individual components of it which are homogeneous, and they are separated by **phase boundaries**.

A system is in **thermal equilibrium** when the macroscopic observables have no time dependence, this set of observables form a particular **equilibrium state**. These observables, which have a well-defined value for each equilibrium state of the system are then called **functions of state**.

If we define equilibrium surfaces from equations of state, then every point on these surfaces will represent a state of equilibrium: If we let $\mathbf{x} = (x_1, x_2, \dots)$ specify a state of a system, and let $f(\mathbf{x})$ be a function of state. Allowing the system to change from \mathbf{x}_i to \mathbf{x}_f , then:

$$\Delta f = \int_{\mathbf{x}_i}^{\mathbf{x}_f} df = \int_{\mathbf{x}_i}^{\mathbf{x}_f} \nabla f(\mathbf{x}) \cdot d\mathbf{x} = f(\mathbf{x}_f) - f(\mathbf{x}_i) \quad (2.1.1)$$

independent of how the change occurs, it only depends on the end points. We call df an **exact differential**, and a quantity that does not have an exact differential is not a function of state.

For example, if $\bar{d}g = ydx$, then for a change from $(0, 0)$ to $(1, 1)$ then along a straight line

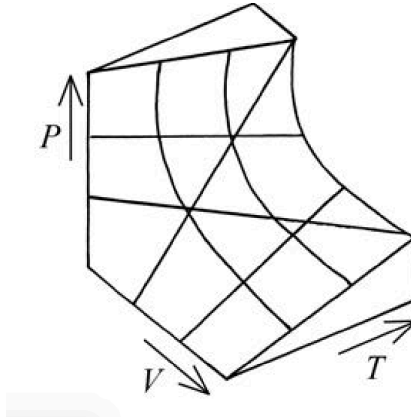


Figure 2.1. pVT surface defines an equilibrium surface for ideal gases which follow the state equation $pV = k_B nT$

path connecting these points:

$$\Delta g = \int_{(0,0)}^{(1,1)} y dx = \int_0^1 x dx = \frac{1}{2} \quad (2.1.2)$$

whereas for a path going to $(1, 0)$ and then $(1, 1)$ we have:

$$\Delta g = \int_{(0,0)}^{(1,0)} y dx + \int_{(1,0)}^{(1,1)} y dx = 0 \quad (2.1.3)$$

So clearly $\bar{d}g$ is an inexact differential, which is symbolized by the small cross through the d . Since Δg depends on the path taken, g is not a function of state.

More generally, to test whether $df = \mathbf{F} \cdot d\mathbf{x}$ it is sufficient to check whether or not \mathbf{F} is conservative

$$\nabla \times \mathbf{F} = 0 \quad (2.1.4)$$

2.2 The first law of thermodynamics

In 1789 Lavoisier proposed a model of heat as a weightless, conserved fluid called **caloric**. It could neither be created nor destroyed, but could be released through combustion. This however did not explain other forms of heating, such as friction which was observed by Rumsford in 1798. So in 1842 Mayer proposed the mechanical equivalent of heat by frictionally generating heat in paper. Joule performed similar experiments but more carefully. He let a mass tied to a string descend a certain height, making a paddle immersed in a mass of water turn. This motion heats the water, and after making the mass fall a certain height Joule measured the temperature rise in the water, deducing correctly the mechanical equivalent of heat. He also measured the heat output of a resistor, and showed that for the same energy input, the same amount of heat was produced, thus suggesting that heat was a form of energy.

The first law of thermodynamics

Energy is conserved, and heat and work are forms of energy, but not functions of state. If U is the **internal energy** of the system, the sum of the energy of all its degrees of freedom, then:

$$dU = dQ + dW \quad (2.2.1)$$

where dQ is the heat supplied to the system, and dW is the work done on the system.

A **thermally isolated system** cannot exchange heat with its surroundings so that:

$$dU = dW \quad (2.2.2)$$

Consider the work done by compressing gas of pressure p and volume V :

$$dW = Fdx = pAdx = -pdV \quad (2.2.3)$$

since decreasing the volume requires positive work. It is important to note that this equation is only true for reversible changes, as we will see in the following chapter.

2.3 Heat capacity

Generally the internal energy will be a function of temperature and volume, so we may set $U = U(T, V)$ and:

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV \quad (2.3.1)$$

Inserting this into $dQ = dU + pdV$ we find:

$$dQ = \left(\frac{\partial U}{\partial T}\right)_V dT + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right] dV \quad (2.3.2)$$

and dividing by dT we find:

$$\frac{dQ}{dT} = \left(\frac{\partial U}{\partial T}\right)_V + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right] \frac{dV}{dT} \quad (2.3.3)$$

Using this we can now calculate the amount of heat we must add to have a change in temperature. If we keep the volume constant:

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V \quad (2.3.4)$$

since the second term in 2.3.3 vanishes at constant volume. If we keep the pressure constant:

$$C_p = \left(\frac{\partial Q}{\partial T}\right)_p = \left(\frac{\partial U}{\partial T}\right)_V + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right] \left(\frac{\partial V}{\partial T}\right)_p \quad (2.3.5)$$

We then define the specific heat capacities as the heat capacity per mass of gas:

$$c_V = \frac{C_V}{M} \quad (2.3.6)$$

$$c_p = \frac{C_p}{M} \quad (2.3.7)$$

where M is the mass of the material. We can also define the ratio of C_p to C_V , called the **adiabatic index** γ :

$$\gamma = \frac{C_p}{C_V} \quad (2.3.8)$$

Example (5.1.1 Sh)

For an ideal monoatomic gas, the internal energy is all due to the kinetic energy, and can be shown (see kinetic theory of gases) to be $U = \frac{3}{2}RT$ per mole, where $R = N_A k_B$ is a constant. Therefore:

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V = \frac{3}{2}R \quad (2.3.9)$$

per mole. Also:

$$C_p = C_V + \left[\left(\frac{\partial U}{\partial V} \right)_T + p \right] \left(\frac{\partial V}{\partial T} \right)_p \quad (2.3.10)$$

but $\left(\frac{\partial U}{\partial V} \right)_T = 0$, and using the ideal gas law for one mole of gas:

$$pV = RT \implies \left(\frac{\partial V}{\partial T} \right)_p = \frac{R}{p} \quad (2.3.11)$$

so that:

$$C_p = C_V + R = \frac{5}{2}R \quad (2.3.12)$$

per mole. Therefore the adiabatic index is given by:

$$\gamma = \frac{5}{3} \quad (2.3.13)$$

It follows that for an ideal gas:

$$dQ = C_V dT + p dV \implies dU = C_V dT \quad (2.3.14)$$

We can also use the result that for an ideal gas $C_p = C_V + R$ to write that:

$$dQ = C_V dT + p dV = (C_p - R) dT + p dV \quad (2.3.15)$$

Hence using the ideal gas law $pV = RT$ $p dV + V dp = R dT$ we find that:

$$dQ = C_p dT - V dp \quad (2.3.16)$$

First Law for Ideal Gases

For an ideal gas undergoing a reversible process, we have that:

$$dU = C_V dT, \quad dW = -pdV \quad (2.3.17)$$

and consequently:

$$dQ = C_V dT + pdV \quad (2.3.18)$$

$$dQ = C_p dT - V dp \quad (2.3.19)$$

2.4 Reversible processes

Pretty much all processes in the universe are governed by time-reversible laws of physics. This is given by the fact that if we looked at a video of the entire universe, we would not be able to realize whether it is in reverse or not. The reason we don't see a broken egg reassemble itself and float on top of a table is due to energy dissipation. When an egg falls off a table and breaks, part of its potential energy is dissipated into its surroundings as heat. This **dissipation** can't be traced back exactly (this has to do with entropy). Similarly, when we delete a file on our computers, we're not truly deleting any information, but converting it into other forms of energy (mostly heat). Tracing back the heat to the original 1s and 0s that encoded the information is impossible, the process is thus irreversible.

For a process to be reversible we require that it be:

1. **Quasistatic**: the system is always in an equilibrium state or infinitesimally close to one (so that any changes in the state variables are infinitesimal). Suppose for example that we push a piston down a column of gas. If we do this slowly enough, the gas will have enough time to adjust to the moving piston, and hence it will always be in an equilibrium state. If instead we introduce a rapid perturbation, this will send shock waves through the gas, exciting its constituent atoms and creating regions of higher temperature. There is no way to extract the sound waves from these regions of high temperature by moving the piston back.
2. **No hysteresis**: hysteresis occurs when a systems time evolution depends on its history. The most common example is the magnetization of iron, where if we remove the magnetizing field slowly, the iron won't go back to its original state, but will instead follow a hysteresis loop as shown below.

Reversibility and dissipation

A process involving dissipation is **irreversible**. A reversible process, called **quasistatic** is a process whose direction can be reversed by an infinitesimal change. A reversible process is therefore **quasistatic** and without **hysteresis** or heat dissipation.

The property that the system is always in an equilibrium state or infinitesimally close to one is analogous to a ball free to roll on a smooth table. If we want to maintain the ball

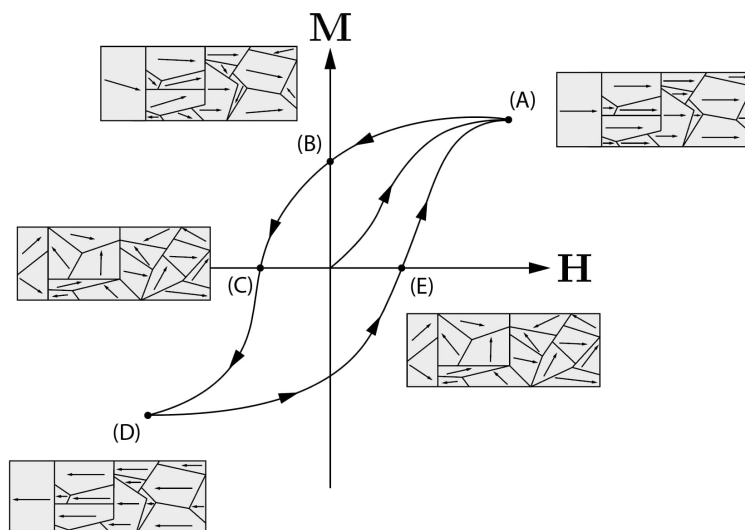


Figure 2.2. Hysteresis loop for the magnetization of iron

in an equilibrium state, so that it does not slip, we must move the table very, very, very slowly. In real life forces such as friction make it so that it doesn't take an infinite amount of time, but in an idealized world this process would take infinitely long.

Since we may define equilibrium as the set of points on an equilibrium surface defined by an equation of state, reversible processes may be defined as processes on this surface. Irreversible processes on the other hand leave this surface.

Hence when systems approach equilibrium, their trajectory is outside of the equilibrium surface and intersects it once equilibrium has been reached. This process is therefore irreversible, as expected from the second law of thermodynamics.

2.5 Isothermal expansion

Consider the heat change in a reversible **isothermal** expansion of an ideal gas, where the temperature remains constant. For an ideal gas we have that $dU = C_V dT = 0$. Inserting this into the first law of thermodynamics:

$$dW = -pdV = -dQ \quad (2.5.1)$$

Therefore for an isothermal expansion from V_1 to V_2 , the heat that must be added is:

$$\Delta Q = \int dQ = \int_{V_1}^{V_2} pdV = \int_{V_1}^{V_2} \frac{RT}{V} dV = RT \ln \frac{V_2}{V_1} \quad (2.5.2)$$

Hence, for an expansion, $V_2 > V_1$, and $\Delta Q > 0$, and heat must therefore be added to the system, whereas for a compression ΔQ heat must be taken from the system.

2.6 Adiabatic expansion

An **adiathermal** process involves no flow of heat, so a system bounded by adiabathermal walls is said to be thermally isolated. A process is further classified as **adiabatic** if it is also reversible. In such a case we have that for an ideal gas:

$$dQ = 0 \text{ and } dU = C_V dT \quad (2.6.1)$$

Hence, using the first law of thermodynamics:

$$C_V dT = -pdV = -\frac{RT}{V} dV \implies \ln \frac{T_2}{T_1} = -\frac{R}{C_V} \ln \frac{V_2}{V_1} \quad (2.6.2)$$

but using the adiabatic index:

$$\gamma = 1 + \frac{R}{C_V} \quad (2.6.3)$$

we can write:

$$TV^{\gamma-1} = \text{const.} \quad (2.6.4)$$

or alternatively:

$$pV^\gamma = \text{const.} \quad (2.6.5)$$

The work done on the system is given by:

$$W = \Delta U = c_v \Delta T = c_v \left(\frac{p_2 V_2}{R} - \frac{p_1 V_1}{R} \right) = \frac{1}{\gamma - 1} (p_2 V_2 - p_1 V_1) \quad (2.6.6)$$

Using $pV^\gamma = \text{const.}$ we then find that:

$$\Delta U = W = \frac{p_1 V_1^\gamma}{\gamma - 1} \left(V_2^{1-\gamma} - V_1^{1-\gamma} \right) \quad (2.6.7)$$

Therefore during an adiabatic process, reducing the internal energy of a system allows for work to be done on the surrounding environment.

Example. Adiabatic atmosphere The pressure due to a thickness dz of atmosphere (which we model as nitrogen gas) with density ρ is given by:

$$dp = -\rho g dz \implies \frac{dp}{dz} = -\rho g \quad (2.6.8)$$

and since $\rho = \frac{mN}{V}$ and $p = \frac{Nk_B T}{V}$ where m is the mass of one molecule, n is the number of nitrogen molecules, we can write:

$$\frac{dp}{dz} = -\frac{mgp}{k_B T} \implies T \frac{dp}{p} = -\frac{mg}{k_B} dz \quad (2.6.9)$$

If we model the atmosphere to be adiabatic, so that each parcel of air does not exchange heat with its surrounding, then we can use:

$$p^{1-\gamma} T^\gamma = \text{const.} \quad (2.6.10)$$

to find differentiating:

$$(1 - \gamma)p^{-\gamma}T^\gamma \frac{dp}{dT} + \gamma p^{1-\gamma}T^{\gamma-1} = 0 \implies (1 - \gamma)\frac{dp}{p} + \gamma\frac{dT}{T} = 0 \quad (2.6.11)$$

Combining this with 2.6.7 we can finally find:

$$\frac{dT}{dz} = -\left(\frac{\gamma - 1}{\gamma}\right)\frac{mg}{k_B} \quad (2.6.12)$$

but using the relation $R = N_A k_B = C_p - C_V$ we quickly find that:

$$\frac{dT}{dz} = -\frac{Mg}{C_p} = -\frac{g}{c_p} \quad (2.6.13)$$

where $c_p = \frac{C_p}{M}$ is the specific heat capacity and M is the molar mass of nitrogen.

2.7 Isochoric and Isobaric expansions

In an **Isochoric process**, the gas has a constant volume throughout. Therefore, the work done on the system is null, because $dW = pdV = 0$. Therefore, we find that the heat released is:

$$Q = C_V \Delta T \quad (2.7.1)$$

and from the first law of thermodynamics:

$$\Delta U = Q = C_V \delta T \quad (2.7.2)$$

So in an isochoric process, heat enters (leaves) the system thus increasing (decreasing) the internal energy. No work is done on the system, and no work is therefore obtainable from the process.

Instead, in an **Isobaric process**, the gas is kept at a constant pressure throughout. Therefore, the work done on the system is:

$$W = -\int_{V_1}^{V_2} pdV = -p\Delta V \quad (2.7.3)$$

The heat supplied is then equal to:

$$Q = C_p \Delta T = (C_V + R)\Delta T \quad (2.7.4)$$

and from the first law of thermodynamics we find that:

$$\Delta U = \Delta Q + \Delta W = (C_V + R)\Delta T - p\Delta V \quad (2.7.5)$$

In an isobaric process, heat enters (leaves) the system. Part of the heat is used by the system to increase (decrease the internal energy) and the rest is used to do work on the surrounding environment.

2.7. ISOCHORIC AND ISOBARIC EXPANSIONS

Thermodynamic quasi-static processes				
Process	Condition	ΔU	ΔQ	ΔW
isothermal	$dT = 0$	$\Delta U = 0$	$\Delta Q = -\Delta W$	$W = RT \ln \frac{V_1}{V_2}$
adiabatic	$dQ = 0$	$\Delta U = \Delta W$	$\Delta Q = 0$	$W = \frac{1}{\gamma-1}(p_2V_2 - p_1V_1)$
isochoric	$dV = 0$	$\Delta U = Q$	$\Delta Q = C_V \Delta T$	$\Delta W = 0$
isobaric	$dp = 0$	$\Delta U = \Delta Q + \Delta W$	$\Delta Q = C_p \Delta T$	$\Delta W = -p \Delta V$

The Second Law

This chapter will introduce the second law of thermodynamics and its applications in the study of engines, machines operating between two reservoirs of different temperatures.

3.1 Perpetual motion machine

Throughout history several attempts have been made to construct a **perpetual motion machine**. Mainly two types of such machines have been thought of:

- (i) machine that produces more energy than it uses
- (ii) machine that produces exactly the same amount of energy as it uses, but continues running indefinitely by converting its waste heat into work

Clearly, the first type violates the first law of thermodynamics. The second type of machine however, is not in violation of this principle, and would theoretically be possible given the thermal physics we have studied.

We are clearly missing a fundamental principle that forbids the "re-usability" of energy, so that no machine is able to convert heat purely into work indefinitely. This principle is the second law of thermodynamics.

3.2 The second law of thermodynamics

Empirical evidence seem to show that heat always flow from hot bodies to cold bodies, and the reverse process never occurs in thermodynamic isolation. This formulation leads to the **Clausius' statement** of the second law of thermodynamics:

Clausius' statement of the second law of thermodynamics

No process can have the sole net result of transferring heat from a colder to a hotter body.

A second, seemingly unrelated statement of the same law can be made, regarding how easy energy can be converted in different forms. This is called **Kelvin's statement of the second law of thermodynamics**

Kelvin's statement of the second law of thermodynamics

No process can have the sole net result of converting heat into work completely.

To see the relation between these two statements, we must first look at **engines**.

3.3 The Carnot Engine

An **engine** is a machine operating between two reservoirs that converts some heat into work by operating through a **Carnot cycle**. The cycle consists of:

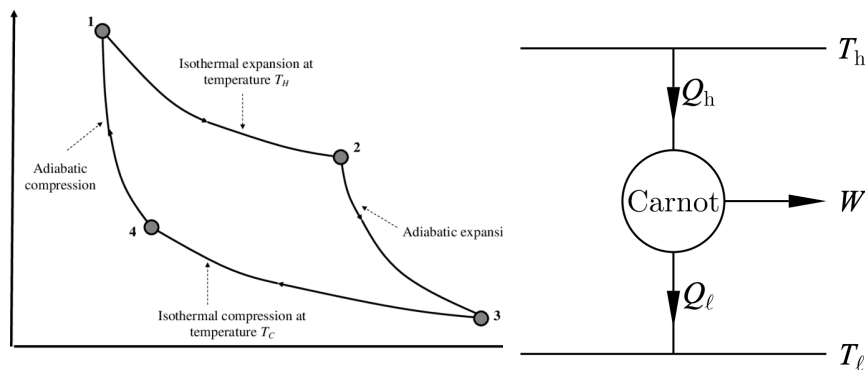


Figure 3.1. Carnot cycle and engine

- 1 → 2 is an isothermal expansion of the gas operating at a higher temperature T_h in which heat Q_h enters from the hot reservoir.
- 2 → 3 is an isentropic (which means adiabatic and reversible) expansion in which the gas is isolated from the reservoirs, but continues expanding by sacrificing its internal energy. The temperature therefore drops to T_l .
- 3 → 4 is an isothermal compression of the gas operating at a lower temperature T_l in which heat Q_l exits into the cold reservoir.
- 4 → 1 is an isentropic (which means adiabatic and reversible) compression in which the gas is isolated from the reservoirs, but continues compressing increasing its internal energy, and thus temperature to T_h .

Note that the process is cyclic, and therefore in one cycle the total internal energy U does not change overall. The first law of thermodynamics then tells us that the work output by the engine is:

$$W = Q_h - Q_l \quad (3.3.1)$$

Using the results from the previous chapter on isothermal and isentropic processes:

$$1 \rightarrow 2 : Q_h = RT_h \ln \frac{V_2}{V_1} \quad (3.3.2)$$

$$2 \rightarrow 3 : \frac{T_h}{T_l} = \left(\frac{V_3}{V_2} \right)^{\gamma-1} \quad (3.3.3)$$

$$3 \rightarrow 4 : -Q_l = RT_l \ln \frac{V_4}{V_3} \quad (3.3.4)$$

$$4 \rightarrow 1 : \frac{T_l}{T_h} = \left(\frac{V_1}{V_4} \right)^{\gamma-1} \quad (3.3.5)$$

where the negative sign in the third equation is due to the fact that the change in energy is negative, whereas the heat transfer Q_l is positive. The first and second equations give:

$$\frac{V_2}{V_1} = \frac{V_3}{V_4} \quad (3.3.6)$$

and dividing the second equation by the fourth, and substituting 3.2.6 we find:

$$\frac{Q_h}{Q_l} = \frac{T_h}{T_l} \quad (3.3.7)$$

The **efficiency** η of an engine can be seen as the ratio of the output energy and the input energy, how much work must be done to obtain a certain energy output. For the Carnot engine, we find that this efficiency is the ratio of the work output W by the machine, and the heat Q_h provided by the reservoirs (which must be kept at the same temperature). Therefore:

$$\eta_C = \frac{W}{Q_h} = 1 - \frac{T_l}{T_h} \quad (3.3.8)$$

Carnot's Engine

A Carnot engine is a machine operating between two reservoirs of temperature T_l and T_h through a Carnot cycle, and has efficiency $\eta_C = 1 - \frac{T_l}{T_h}$.

Since, by equation 3.2.1 $W < Q_h$, the efficiency of the Carnot machine is always lower than 100%, and as we will see soon, no machine can be more efficient than the Carnot engine.

It is now evident how the second law of thermodynamics can explain the impossibility of a *perpetuum mobile* of the second type.

3.4 Carnot's theorem

It turns out that the efficiency of the Carnot engine is maximized, as encapsulated in the **Carnot's theorem**.

Carnot's Theorem

Of all the heat engines working between two temperatures, the most efficient one is the Carnot engine. Thus all reversible engines working between two temperatures have the same efficiency η_C .

Proof. Suppose that there does exist an engine E more efficient than the Carnot engine. Because all the processes in a Carnot cycle are reversible, we run the Carnot engine in reverse and connect it to engine E as shown below.

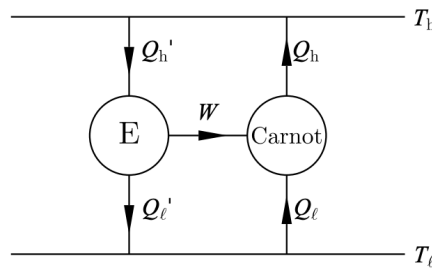


Figure 3.2. Engine E connected to Carnot engine run in reverse

From the first law of thermodynamics we have:

$$W = Q_h' - Q_l' = Q_h - Q_l \implies Q_h - Q_h' = Q_l - Q_l' \quad (3.4.1)$$

Now because $\eta_C > \eta_E$, we must have that $\frac{W}{Q_h'} < \frac{W}{Q_h}$ and thus $Q_h > Q_h'$.

So $Q_h - Q_h' > 0 \implies Q_l - Q_l' > 0$. Since $Q_h - Q_h'$, the amount of heat dumped into the reservoir T_h is positive and equal to $Q_l - Q_l'$, the amount of heat extracted from reservoir T_l , this engine has the sole purpose of transferring heat from a cold body to a hotter body, and therefore violates Clausius's statement of the second law of thermodynamics.

To prove that all reversible engines have the same efficiency, consider a reversible engine R connected to a Carnot engine as shown below:

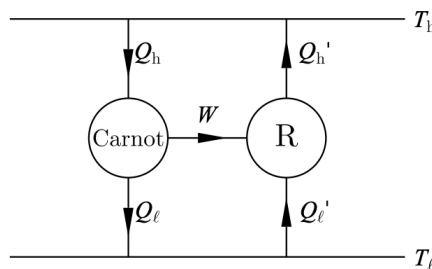


Figure 3.3. Reversible engine R connected to a Carnot engine.

Then its efficiency, by Carnot's theorem, is $\eta_E \leq \eta_C$. Using the same logic as before, we find that the system only dumps heat from the cold reservoir to the hot reservoir for $\eta_E < \eta_C$, thus violating the Clausius' statement. The only possibility is therefore to have $\eta_E = \eta_C = 1 - \frac{T_l}{T_h}$. \square

3.5 The equivalence of Clausius' and Kelvin's statements

To prove that the two are equivalent, it suffices to prove that:

- (i) if a system violates Kelvin's statement then it violates Clausius' statement.
- (ii) if a system violates Clausius' statement violates Kelvin's statement

Proof. (i) If an engine violates Kelvin's statement (so that all the heat is converted to work), then we can connect it to a Carnot engine as shown below: From the first law:

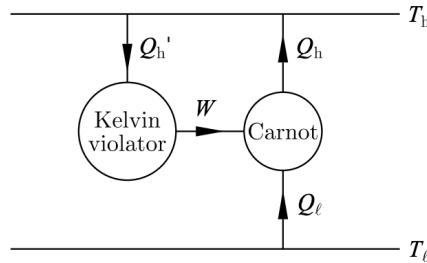


Figure 3.4. Kelvin violator is connected to a Carnot engine

$$Q'_h = W, \quad Q_h = W + Q_l \tag{3.5.1}$$

So the heat dumped into the reservoir T_h is $Q_h - Q'_h = Q_l$ which is the heat extracted from the reservoir T_l . Hence this system is transferring heat from a colder body to a warmer body.

- (ii) If a system violates Clausius' statement, then its only effect is to transfer heat from the reservoir at T_l to the reservoir at T_h . We can then connect it to a Carnot engine as shown:

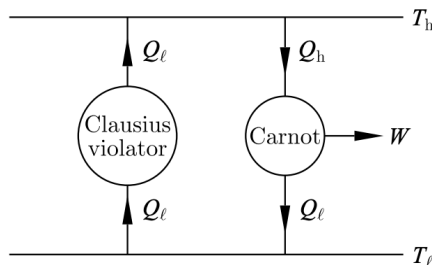


Figure 3.5. Clausius violator is connected to a Carnot engine

From the first law $Q_h - Q_l = W$, so the sole effect of the system is to convert heat into work, thus violating Kelvin's statement. □

Example.

(a) **Refrigerator**

The refrigerator is a heat engine run in reverse. The cold reservoir is the cold food inside the refrigerator, and the hot reservoir is the kitchen. Therefore the efficiency here will be the heat sucked out of the contents (which is what we wish to get) divided by the work required to achieve this:

$$\eta = \frac{Q_l}{W} \tag{3.5.2}$$

For a refrigerator with a Carnot engine, we can show:

$$\eta_C = \frac{T_l}{T_h - T_l} \tag{3.5.3}$$

which is sometimes greater than 100%

(b) **Heat pump**

A heat pump is a refrigerator, used to pump heat from a reservoir to another place. The effect we want is to add heat Q_h to this place, and W is the work that must be done, so that the efficiency is:

$$\eta = \frac{Q_h}{W} \tag{3.5.4}$$

So for a heat pump with a Carnot engine we have that:

$$\eta_C = \frac{T_h}{T_h - T_l} \tag{3.5.5}$$

which is always greater than 100%.

3.6 Clausius' theorem

Consider a general cycle where in an infinitesimal step i heat dQ_i enters at some point connected to a reservoir at temperature T_i . The total work extracted is:

$$\Delta W = \sum_{cycle} dQ_i \tag{3.6.1}$$

We now let each point be supplied with heat via a Carnot engine as shown in (b), operating between a reservoir at temperature T providing heat $dQ_i + dW_i$, and a reservoir T_i . Each

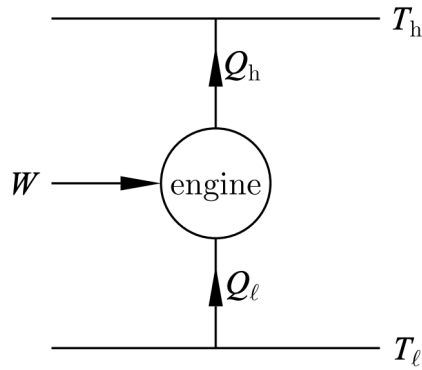


Figure 3.6. Refrigerator/heat pump designs

engine produces work dW_i such that:

$$\frac{\text{heat to } T_i}{T_i} = \frac{\text{heat from } T}{T} \quad (3.6.2)$$

and so:

$$\frac{dQ_i}{T_i} = \frac{dQ_i + dW}{T} \implies dW_i = dQ_i \left(\frac{T}{T_i} - 1 \right) \quad (3.6.3)$$

Because we can't heat completely into work, we must have that the total work in the entire cycle is negative or zero, otherwise we would have converted heat to work.

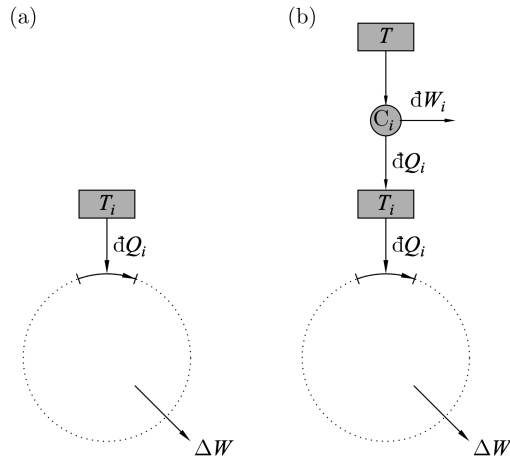


Figure 3.7. (a) A general cycle in which heat dQ_i enters (b) Same cycle, but with heat entering through a carnot engine

Therefore:

$$\Delta W + \sum_{\text{cycle}} dW_i \leq 0 \implies \sum_{\text{cycle}} dQ_i + dW_i \leq 0 \quad (3.6.4)$$

and therefore:

$$T \sum_{\text{cycle}} \frac{dQ_i}{T_i} \leq 0 \implies \frac{dQ_i}{T_i} \leq 0 \quad (3.6.5)$$

since temperature is positive. Turning the sum into an integral:

$$\oint \frac{dQ}{T} \leq 0 \quad (3.6.6)$$

which is the **Clausius inequality**.

If the cycle is reversible, then running it in reverse dQ turns into $-dQ$ and we find:

$$\oint \frac{-dQ}{T} \leq 0 \implies \oint \frac{dQ}{T} = 0 \quad (3.6.7)$$

Clausius' Theorem

For any closed cycle:

$$\oint \frac{dQ}{T} \leq 0 \quad (3.6.8)$$

where equality only holds for reversible cycles.

In essence, Clausius' theorem emphasizes the fact that the heat available to do work in a closed cycle must always decrease or remain constant.

Example. Consider two bodies with heat capacities C_h and C_l used as reservoirs for a Carnot heat engine. We wish to find the total work obtainable. We have that:

$$dQ_h = -C_h dT_h \quad (3.6.9)$$

$$dQ_l = -C_l dT_l \quad (3.6.10)$$

For a Carnot engine we also have:

$$\frac{dQ_l}{T_l} = \frac{dQ_h}{T_h} \quad (3.6.11)$$

Integrating directly we find:

$$C_l \ln \frac{T_f}{T_l} = -C_h \ln \frac{T_f}{T_h} \implies T_f^{C_h+C_l} = T_h^{C_h} T_l^{C_l} \quad (3.6.12)$$

where T_f is the final temperature of the reservoirs. Thus:

$$\Delta Q_h = C_h(T_f - T_h) \quad (3.6.13)$$

$$\Delta Q_l = C_l(T_f - T_l) \quad (3.6.14)$$

so that the work obtainable is:

$$\Delta W = \Delta Q_h - \Delta Q_l = C_h(T_f - T_h) - C_l(T_f - T_l) \quad (3.6.15)$$

Entropy

4.1 Definition of Entropy

Recall that for a reversible process, we have:

$$\oint \frac{dQ_{rev}}{T} = 0 \quad (4.1.1)$$

This implies that $\frac{dQ_{rev}}{T}$ is a function of state defined as **entropy** S as:

Definition: Entropy

The entropy S is a function of state:

$$dS = \frac{dQ_{rev}}{T} \quad (4.1.2)$$

so that the change in entropy between two points A, B is:

$$\Delta S = \int_A^B \frac{dQ_{rev}}{T} \quad (4.1.3)$$

For an irreversible change $A \rightarrow B$, and let us form a closed loop by joining it with a reversible change $B \rightarrow A$. Then the Clausius' inequality reads:

$$\oint \frac{dQ}{T} = \int_A^B \frac{dQ}{T} + \int_B^A \frac{dQ_{rev}}{T} \leq 0 \quad (4.1.4)$$

and rearranging gives:

$$\int_A^B \frac{dQ}{T} \leq \int_A^B \frac{dQ_{rev}}{T} \quad (4.1.5)$$

Therefore:

$$dS \geq \frac{dQ}{T} \quad (4.1.6)$$

For a thermally isolated system $dQ = 0$ always and thus:

$$\boxed{dS \geq 0} \quad (4.1.7)$$

If we assume that the universe is isolated, then we can write that the total entropy tends to a maximum.

Entropic statement of the second law

For a thermally isolated system, the entropy must always increase for an irreversible change and stay the same for a reversible change.

Example.

A large reservoir at constant temperature T_R is placed in thermal contact with a small system at temperature T_S , and both end up with temperature T_R .

The heat transferred to the system is therefore equal to $\Delta Q = C(T_R - T_S)$. The entropy change in the reservoir can then be calculated as:

$$\Delta S_{res} = \int \frac{dQ}{T_R} \quad (4.1.8)$$

$$= -\frac{\Delta Q}{T_R} \quad (4.1.9)$$

$$= \frac{C(T_S - T_R)}{T_R} \quad (4.1.10)$$

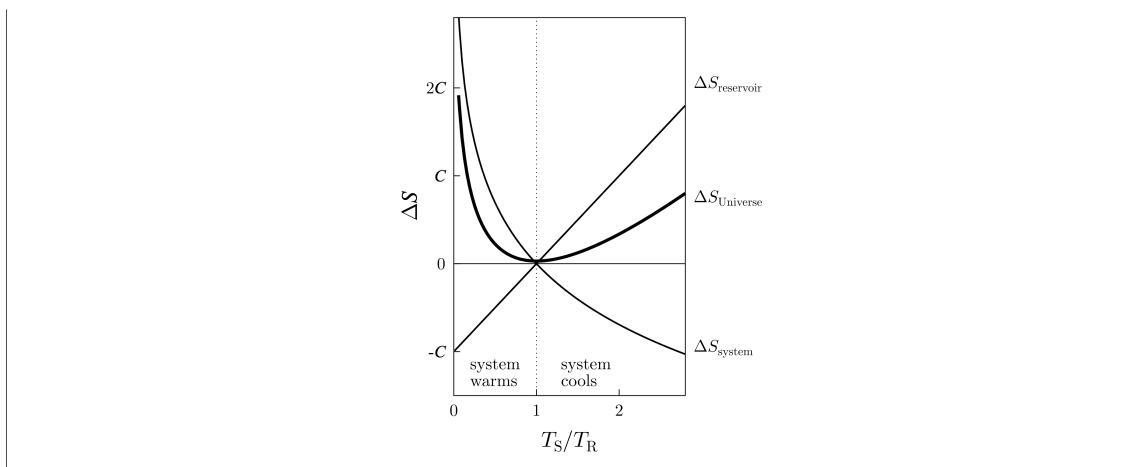
Instead, the entropy change of the system is :

$$\Delta S_{sys} = \int_{T_S}^{T_R} \frac{CdT}{T} = C \ln \frac{T_R}{T_S} \quad (4.1.11)$$

Therefore, the total change in entropy is:

$$\Delta S = C \left(C \ln \frac{T_R}{T_S} + \frac{T_S}{T_R} - 1 \right) \quad (4.1.12)$$

From the plot below, we can see that the entropy is always non-negative, and equal to zero when $T_S = T_R$, which intuitively makes sense since there would be no heat flow.



4.2 Reformulating the First Law

For a reversible change, we can write $dQ = TdS$, $dW = -pdV$. We can then write:

$$dU = TdS - pdV \quad (4.2.1)$$

Luckily, this equation actually holds for irreversible processes as well! This is fundamental equation of thermodynamics:

Fundamental equation of thermodynamics

$$dU = TdS - pdV$$

Note that we may write dU using the chain rule as:

$$dU = \left(\frac{\partial U}{\partial S}\right)_V dS + \left(\frac{\partial U}{\partial V}\right)_S dV \quad (4.2.2)$$

from which it follows that:

$$\begin{cases} T = \left(\frac{\partial U}{\partial S}\right)_V \\ p = \left(\frac{\partial U}{\partial V}\right)_S \end{cases} \implies \frac{p}{T} = -\left(\frac{\partial U}{\partial V}\right)_S \left(\frac{\partial S}{\partial U}\right)_V \quad (4.2.3)$$

Now, from the reciprocity theorem we must have that $\left(\frac{\partial U}{\partial V}\right)_S \left(\frac{\partial S}{\partial U}\right)_V \left(\frac{\partial U}{\partial V}\right)_S \left(\frac{\partial V}{\partial S}\right)_U = 1$ and thus:

$$\frac{p}{T} = -\left(\frac{\partial S}{\partial V}\right)_U \quad (4.2.4)$$

In summary.

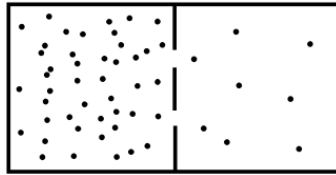
$$\left\{ \begin{array}{ll} dU = \bar{d}Q + \bar{d}W & \text{is always true} \\ \bar{d}Q = TdS & \text{is only true for reversible changes} \\ \bar{d}W = -pdV & \text{is only true for reversible changes} \\ dU = TdS - pdV & \text{is always true} \end{array} \right. \quad (4.2.5)$$

4.3 Joule Expansion

We now go back to the problem of expanding a gas freely and quasistatically in a container, and try to answer the question of why the particles do not reassemble themselves.

Indeed, if this were to happen, then the second law of thermodynamics, $dS \geq 0$, would be violated. We see this by noting that since the system is thermally isolated, $dU = 0$, and if the particles reassemble themselves into their original volume, $dV \leq 0$. Then, using the fundamental equation of thermodynamics we get $dS < 0$, a contradiction.

Physically, this means that the particles reassembling themselves does not satisfy the maximum entropy principle. This explains why, if we let a gas expand freely in a container, this process is irreversible, and we do not observe the gas particles neatly reassemble themselves in the smaller volume they were initially contained within.



The expansion of the gas is therefore irreversible, because $\Delta S > 0$.

Let us calculate the entropy change, we consider one mole of an ideal gas confined to a container of volume V_0 , which freely expands isothermally, upon the opening of a tap, into a container of volume V_1 ¹.

The system is thermally isolated, so $\Delta U = 0$ and since $C_V \Delta T = \Delta U = 0 \implies \Delta T = 0$ so the initial and final temperatures are the same, T . Also, $dU = 0 = TdS - pdV$.

Now using the ideal gas law, $p_i V_0 = RT = p_f V_1 \implies p_f = \frac{V_0}{V_1} p_i$. Now we can write:

$$\Delta S = \int \frac{\bar{d}Q}{T} = \int_{V_0}^{V_1} \frac{p(V)dV}{T} = \int_{V_0}^{V_1} \frac{RdV}{V} = R \ln \frac{V_1}{V_0} \quad (4.3.1)$$

We now write that $\Delta S = \Delta S_{universe} > 0$ since there can be no heat transfer to the surroundings. Hence, the Joule expansion is irreversible.

¹the fact that the process is isothermal is irrelevant, since dS is a function of state and is independent of path

4.4 Entropy change for ideal gas

For an ideal gas we have that:

$$dQ = dU - dW = C_V dT + p dV = C_V dT + \frac{nRT}{V} dV \quad (4.4.1)$$

We cannot integrate this without knowing how the volume changes with temperature, however, if we divide by T we find that:

$$\frac{dQ}{T} = \frac{C_V dT}{T} + \frac{nR}{V} dV \quad (4.4.2)$$

which upon integration gives:

$$\Delta S = C_V \ln \frac{T_2}{T_1} + nR \ln \frac{V_2}{V_1} \quad (4.4.3)$$

4.5 Statistical interpretation of entropy

In the Statistical Mechanics part of these lectures we will see that entropy may alternatively be defined as

$$S = k_B \ln \Omega \quad (4.5.1)$$

called the **Boltzmann-Planck equation**, where Ω is the number of microstates consistent with the given macrostate.

Example.

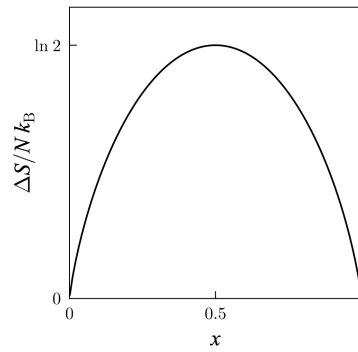
Consider two ideal gases, separated in two containers of volumes xV and $(1-x)V$ at pressure p , temperature T . From the ideal gas law, $p = \frac{Nk_B T}{V}$, so there are xN moles of gas 1 and $(1-x)N$ moles of gas 2.

If we now open the tap between the two containers, the gases will mix resulting in an increasing entropy. If we assume an isothermal expansion, $\Delta U = 0$ so $TdS = pdV \implies dS = \frac{pdV}{T}$ and thus:

$$\Delta S = xNk_B \int_{xV}^V \frac{dV_1}{V_1} + (1-x)Nk_B \int_{(1-x)V}^V \frac{dV_2}{V_2} \quad (4.5.2)$$

$$= -Nk_B (x \ln x + (1-x) \ln(1-x)) \quad (4.5.3)$$

We plot the entropy of mixing below:



We see that entropy is maximum when $x = \frac{1}{2}$ with value $Nk_B \ln 2$, and is lowest when $x = 0, 1$ obviously. The plot is symmetric about $x = \frac{1}{2}$, since the system too is symmetric (we can exchange the two containers, switch $x \leftrightarrow (1 - x)$).

In the case where $x = \frac{1}{2}$, the two containers have equal volume. So, when the tap is opened, Ω , the number of microstates that can exist must be multiplied by 2^N , since every molecule can be in either container 1 or container 2, and not only in its original container. So, the change in entropy is $k_B \ln 2^N = Nk_B \ln 2$ as required.

4.5.1 Gibb's paradox

But what if the two gases were identical? In this previous example we assumed that the gases had some distinctive quality that allowed us to say that each molecule could be in either one container or the other. However, with indistinguishable molecules, it does not make sense anymore in the case where $x = \frac{1}{2}$ that the number of microstates is multiplied by 2^N . The change in entropy would therefore be: $\Delta S = 0$, and not $\Delta S = Nk_B \ln 2$.

We see that the results for distinguishable particles cannot be used for identical particles.

4.5.2 Maxwell's demon

Another paradox that attempts to contradict the second law of thermodynamics is **Maxwell's demon**. A gas is initially in a chamber, and performs a Joule expansion into a second connected chamber. This process, as we explained earlier through the maximum entropy principle, must be irreversible. However, Maxwell came up with the idea of inserting an intelligent valve (the so-called demon), that upon sensing a particle directed towards the initial container, opens only allowing that particle to return to its original vessel. If we repeat this process for long enough, we should get all the particles back to the original container, thus reversing the Joule expansion and reducing the entropy of the system.

The same demon could be used to partition fast gas molecules into one container, and slower gas molecules into the other, thereby transferring heat without doing work.

It turns out that to function properly, the demon must store information to understand where the molecules are going. This leads to an increase in entropy, called the **Shannon entropy**.

4.6 Gibbs expression for entropy

Consider a system with N different microstates. Each i th microstate has n_i contained within it, all identical to each other, and not directly measurable. The probability of observing the system in the i th microstate is $P_i = \frac{n_i}{N}$, with $N = \sum_i n_i$.

The total entropy is, of course, $S_{tot} = k_B \ln N$. However, we may express this total entropy as the sum of the entropy S due to the freedom of choice of each microstate, and the entropy S_{micro} due to the different sub-microstates. So:

$$S_{tot} = S + S_{micro} \quad (4.6.1)$$

We cannot measure S_{micro} directly, we can take the expectation value:

$$S_{micro} = \sum_i P_i S_i = \sum_i k_B \ln n_i \quad (4.6.2)$$

so that:

$$S = k_B \left(\ln N - \sum_i P_i \ln n_i \right) \quad (4.6.3)$$

and since $\ln \frac{n_i}{N} = \ln P_i$ we find that:

$$\boxed{S = -k_B \sum_i P_i \ln P_i} \quad (4.6.4)$$

It follows that the configuration with the most possible microstates will also maximize the entropy, and lead to equilibria. Systems approaching a state of equilibrium will therefore maximize the number of microstates they can occupy.

Thermodynamic potentials

5.1 Definitions

We have seen that the internal energy of a system is a very useful function of state. It turns out that there are several other such quantities with units of energy, called **thermodynamic potentials**. They can all be found by adding to U several combinations of p, V, T, S . We define them below:

Definition: Thermodynamic potentials

We define the following thermodynamic potentials:

$$\left\{ \begin{array}{ll} \text{Internal energy :} & U \\ \text{Enthalpy :} & H = U + pV \\ \text{Hemholtz function :} & F = U - TS \\ \text{Gibbs function :} & G = U + pV - TS \end{array} \right. \quad (5.1.1)$$

5.2 Differentials

Internal energy

The internal energy may be interpreted as the energy needed to create a system, with all its internal degrees of freedom. We find that:

$$dU = TdS - pdV \quad (5.2.1)$$

$$= \left(\frac{\partial U}{\partial S} \right)_V dS + \left(\frac{\partial U}{\partial V} \right)_S dV \quad (5.2.2)$$

which implies:

$$\left\{ \begin{array}{l} T = \left(\frac{\partial U}{\partial S} \right)_V \\ p = - \left(\frac{\partial U}{\partial V} \right)_S \end{array} \right. \quad (5.2.3)$$

This suggests that we use S, V as the pair of independent thermodynamic variables to define the internal energy. Such pairs will be known as **natural variables**.

Definition: Natural variables

The natural variables of a quantity are a set of independent thermodynamic variables that fully specify it through partial differentiation.

Enthalpy

Enthalpy is interpreted as the energy needed to create a system and the work needed to make space for it, at constant pressure. It is therefore the heat released/absorbed by a system undergoing reversible isobaric processes. It is particularly useful in laboratories, where pressure is almost always held constant. We find that:

$$dH = TdS - pdV + pdV + Vdp \quad (5.2.4)$$

$$= TdS + Vdp \quad (5.2.5)$$

$$= \left(\frac{\partial H}{\partial S}\right)_p dS + \left(\frac{\partial H}{\partial p}\right)_S dp \quad (5.2.6)$$

which implies:

$$\begin{cases} T = \left(\frac{\partial H}{\partial S}\right)_p \\ V = -\left(\frac{\partial H}{\partial p}\right)_S \end{cases} \quad (5.2.7)$$

Hemholtz function

The Hemholtz function may be interpreted as the maximum amount of work you can get out of a system at constant temperature and volume. Alternatively, it is the energy needed to create a system, minus the energy that can be supplied by the environment.

We find that:

$$dF = TdS - pdV - TdS - SdT \quad (5.2.8)$$

$$= -SdT - pdV \quad (5.2.9)$$

$$= \left(\frac{\partial F}{\partial T}\right)_V dT + \left(\frac{\partial F}{\partial V}\right)_T dV \quad (5.2.10)$$

which implies:

$$\begin{cases} S = -\left(\frac{\partial F}{\partial T}\right)_V \\ p = -\left(\frac{\partial F}{\partial V}\right)_T \end{cases} \quad (5.2.11)$$

Gibb's function

Finally, we may interpret the Gibb's function as the energy needed to make a system and make room for it, minus the amount of work available from the surroundings. It is alternatively the maximum amount of work that may be performed at constant temperature

and pressure. We find that:

$$dG = TdS + Vdp - TdS - SdT \quad (5.2.12)$$

$$= Vdp - SdT \quad (5.2.13)$$

$$= \left(\frac{\partial G}{\partial p}\right)_T dp + \left(\frac{\partial G}{\partial T}\right)_p dT \quad (5.2.14)$$

which implies:

$$\begin{cases} S = -\left(\frac{\partial G}{\partial T}\right)_p \\ V = \left(\frac{\partial G}{\partial p}\right)_T \end{cases} \quad (5.2.15)$$

In summary:

Potential	First law	Other fundamental variables
$U(S, V) = TS - pV$	$dU = TdS - pdV$	$\begin{cases} T = \left(\frac{\partial U}{\partial S}\right)_V \\ p = -\left(\frac{\partial U}{\partial V}\right)_S \end{cases}$
$H(S, p) = U + pV$	$dH = TdS + Vdp$	$\begin{cases} T = \left(\frac{\partial H}{\partial S}\right)_p \\ V = -\left(\frac{\partial H}{\partial p}\right)_S \end{cases}$
$F(T, V) = U - TS$	$dF = -SdT - pdV$	$\begin{cases} S = -\left(\frac{\partial F}{\partial T}\right)_V \\ p = -\left(\frac{\partial F}{\partial V}\right)_T \end{cases}$
$G(T, p) = U - TS + pV$	$dG = -SdT + Vdp$	$\begin{cases} S = -\left(\frac{\partial G}{\partial T}\right)_p \\ V = \left(\frac{\partial G}{\partial p}\right)_T \end{cases}$

5.3 Integrals under constant variable

Isochoric processes

At constant volume, we have an isochoric process. Here the internal energy U is the most useful, since the expression for U simplifies to:

$$dU = TdS \quad (5.3.1)$$

and

$$TdS = T\left(\frac{\partial S}{\partial T}\right)_V = \left(\frac{\partial Q_{rev}}{\partial T}\right)_V \equiv C_V \quad (5.3.2)$$

. Therefore the internal energy change:

$$\Delta U = \int_{T_1}^{T_2} C_V dT \quad (5.3.3)$$

represents the heat reversibly absorbed by the system.

Isobaric processes

At constant pressure, we have an isobaric process. Here the enthalpy is the most useful, since the expression for H simplifies to:

$$dH = TdS \quad (5.3.4)$$

and

$$TdS = T \left(\frac{\partial S}{\partial T} \right)_p = \left(\frac{\partial Q_{rev}}{\partial T} \right)_p \equiv C_p \quad (5.3.5)$$

so

$$\Delta H = \int_{T_1}^{T_2} C_p dT \quad (5.3.6)$$

is the heat reversibly absorbed by the system.

- (i) If heat enters the system, $\Delta H > 0$ and we have an **endothermic** reaction.
- (ii) If heat leaves the system, $\Delta H < 0$ and we have an **exothermic** reaction.

Isothermal processes

At constant temperature, we have an isothermal process. Here the hemholtz function is the most useful, since the expression for F simplifies to:

$$dF = -pdV \quad (5.3.7)$$

so:

$$\Delta F = \int_{V_1}^{V_2} -pdV \quad (5.3.8)$$

is the work done reversibly on the system (remember $dW = -pdV$ only for reversible processes).

Isothermal and isobaric processes

If we keep both pressure and temperature constant, then $dG = 0$ so that:

$$\Delta G = 0 \quad (5.3.9)$$

so the Gibbs function is a conserved quantity.

Note that if we know any one of the thermodynamic potentials, we get the other free for free. Consider for example the case where the Hemholtz free energy $F(T, V)$ is known. Then we have found that:

$$\begin{cases} S = -\left(\frac{\partial F}{\partial T}\right)_V \\ p = -\left(\frac{\partial F}{\partial V}\right)_T \end{cases} \quad (5.3.10)$$

The internal energy is then:

$$U(S, V) = F + TS = F - T \left(\frac{\partial F}{\partial T} \right)_V = -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T} \right)_V \quad (5.3.11)$$

and similarly the enthalpy will be:

$$H(S, p) = U + pV = F - T \left(\frac{\partial F}{\partial T} \right)_V - \left(\frac{\partial F}{\partial V} \right)_T \quad (5.3.12)$$

Finally the Gibbs free energy will be:

$$G(T, p) = F + pV = F - V \left(\frac{\partial F}{\partial V} \right)_T = -V^2 \frac{\partial}{\partial V} \left(\frac{F}{V} \right)_T \quad (5.3.13)$$

5.4 Legendre transformations

The Legendre Transformations

Consider a function $f(x)$ such that $\frac{df}{dx} = u(x)$ with $u(x)$ invertible. Let us then define:

$$g(u) = f(x(u)) - u \cdot x(u) \implies \frac{dg}{du} = x(u) \quad (5.4.1)$$

One calls f and g **Legendre transforms** of each other, and to transform from one to the other we simply exchange $x \longleftrightarrow u$. Legendre transforms are such that their derivatives are inverses of each other.

Let us take for example $f(x) = x^3$. Then $f'(x) = u(x) = 3x^2 \implies x(u) = \sqrt{\frac{u}{3}}$ and hence $g(u) = \left(\sqrt{\frac{u}{3}} \right)^3 - u \sqrt{\frac{u}{3}} = \left(\frac{u}{3} \right)^{3/2}$. Hence the Legendre transform of f will be $g(x) = \left(\frac{x}{3} \right)^{3/2}$, which is indeed $f(x(u))$.

Legendre transformations are particularly relevant when discussing thermodynamic potentials, which are all Legendre transforms of each other.

Indeed, starting from $dU = TdS - pdV$ suppose we wish to instead find some potential H with natural variables S, p rather than S, V . Then we clearly must interchange p, V , and we may do so by applying the Legendre transformation with $x = -p$ and $u = V$. We find that:

$$dH = dU + d(pV) = TdS - pdV + pdV + Vdp = TdS + VdP \quad (5.4.2)$$

as we found earlier. This process can be repeated to find the other potentials.

Not only, suppose we had a system with more than two degrees of freedom, such as a rod. In the next chapter we will see that the first law reads:

$$dU = TdS + f dL - pdV \quad (5.4.3)$$

Suppose we want a new potential X with natural variables T, L, p rather than S, L, V . Then

we must write that:

$$dX = TdS + fdL - pdV - d(TS) + d(pV) = -SdT + fdL + Vdp \quad (5.4.4)$$

implying that we use:

$$X = U - TS + pV \quad (5.4.5)$$

5.5 Equilibrium conditions

It is still unclear how these quantities may be put to use effectively. We will see however that the thermodynamic potentials provide an effective way to define how equilibrium may be reached in different conditions.

Consider a system with T, p absorbing heat dQ . The entropy of the surroundings changes by $dS_0 = -\frac{dQ}{T}$, where all the heat exchange is reversible since the surroundings are in thermal equilibrium. Since $dS_{tot} = dS_0 + dS \geq 0$, where dS is the change in entropy of the system, we must have that $TdS \geq dQ$. Using the first law:

$$dW \geq dU - TdS \quad (5.5.1)$$

Now notice that $dW = -p_0dV + dW_{mech}$ where dW_{mech} is any mechanical work that can be done on the system, and $-p_0dV$ is the work done by the surroundings on the system as its volume changes.

So, we find that:

$$dW_{mech} \geq dU + p_0dV - T_0dS = d(U + p_0V - T_0S) \quad (5.5.2)$$

since p_0, T_0 are constants (the surroundings are in thermodynamic equilibrium).

Definition: Availability

We define the availability as:

$$A \equiv U + p_0V - T_0S \quad (5.5.3)$$

which is the maximum amount of work which we can extract from a system in surroundings with temperature T_0 and pressure p_0 .

To see why we can interpret the availability as such, suppose we have a system with temperature T and pressure p fitted with a piston, in surroundings with temperature T_0 and pressure p_0 . The maximum amount of work is extracted when the process is reversible, hence when:

$$dU = TdS - pdV \quad (5.5.4)$$

and thus

$$dA = (T - T_0)dS - (p - p_0)dV \quad (5.5.5)$$

The second term is the net mechanical work done on the piston through a change of volume. To see the significance of the first term, suppose we run a reversible Carnot engine

between the system and the surroundings. We must have that the work done by the engine be:

$$dW_e = T_0 dS - T dS = -(T - T_0) dS \quad (5.5.6)$$

Therefore the first term represents the maximum work we may extract from the system through a change of entropy.

We can finally write:

$$\boxed{dW \geq dA} \quad (5.5.7)$$

so the mechanical work done on a system must always be greater than or equal to the available work the system can do. Normally, the system is mechanically isolated, so we may write:

$$dA \leq 0 \quad (5.5.8)$$

Equilibrium is reached when $dA = 0$, that is, when we minimize A , as a direct result of maximizing the global entropy (we started with the inequality $dS_{tot} \geq 0$).

In the following we consider mechanically isolated systems, so that we may assert $dA \leq 0$.

Thermally isolated system at constant volume

No heat can leave, and the volume can't change so that no work can be done by the system. Therefore $dU = 0$ and so $dA = -T_0 dS$, the availability is minimized when the system's **entropy is maximum**.

Thermally isolated system at constant pressure

Since no heat can leave we have that $dS = 0$, as well as $dp = 0$ due to the constant pressure constraint. Hence $dA = dU + p_0 dV = dU + p dV + V dp = dH$. Consequently, the availability is minimized when the system's **Enthalpy is minimum**.

Constant volume and temperature

Since $dV = 0$ and $dT = 0$ we have that $dA = dU - T_0 dS = dU - T dS = dF$. Consequently the availability is minimized when the system's **Helmholtz function is minimum**

5.5.1 Constant pressure and temperature

We have that $dp = dT = 0$ so $dA = dU + p_0 dV - T_0 dS = dU + p dV + T dS = dG$, the availability is minimized when the system's **Gibb's function is minimum**.

Note that these conditions are all equivalent formulations of the maximum *global* entropy principle, but are more or less useful in different situations where different constraints are imposed.

Our set of equilibrium conditions therefore reads:

$$dS = 0, dV = 0, dU = 0 \quad (5.5.9)$$

$$dS = 0, dp = 0, dH = 0 \quad (5.5.10)$$

$$dT = 0, dV = 0, dF = 0 \quad (5.5.11)$$

$$dT = 0, dp = 0, dG = 0 \quad (5.5.12)$$

If any two of the quantities in an above row is specified and kept constant, then the third must be minimized (or maximized for entropy). We therefore get the following conditions (only the most useful) for equilibria:

Variables	Equilibrium condition
T, p	G minimized
T, V	F minimized
U, V	S maximized
S, p	H minimized

5.6 Connecting the potentials: Maxwell's relations

Consider a state function $f(x, y)$, then we may write that:

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy = F_x dx + F_y dy \quad (5.6.1)$$

where $F_x = \left(\frac{\partial f}{\partial x}\right)_y$ and $F_y = \left(\frac{\partial f}{\partial y}\right)_x$. Also, by symmetry of the second derivative:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \quad (5.6.2)$$

so that:

$$\left(\frac{\partial F_x}{\partial y}\right)_x = \left(\frac{\partial F_y}{\partial x}\right)_y \quad (5.6.3)$$

We can apply this idea to the thermodynamic potentials to obtain Maxwell's relations.

For example, the Maxwell relation based on G , the Gibb's function, can be found as follows. Firstly, we write down the exact differential equation:

$$dG = -SdT + Vdp \quad (5.6.4)$$

Hence:

$$-\left(\frac{\partial S}{\partial p}\right)_T = \left(\frac{\partial V}{\partial T}\right)_p \quad (5.6.5)$$

Repeating this for the other three potentials, we find that:

$$-\left(\frac{\partial S}{\partial p}\right)_T = \left(\frac{\partial V}{\partial T}\right)_p \quad (5.6.6a)$$

$$-\left(\frac{\partial p}{\partial S}\right)_V = \left(\frac{\partial T}{\partial V}\right)_S \quad (5.6.6b)$$

$$\left(\frac{\partial V}{\partial S}\right)_p = \left(\frac{\partial T}{\partial p}\right)_S \quad (5.6.6c)$$

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V \quad (5.6.6d)$$

To remember these, the following mnemonic may be used:

- (i) Cross multiplication of the variables gives $(TS) = (pV)$.
- (ii) Opposite pairs of variables (both in numerator or both in denominator) are set as constants.
- (iii) The sign is positive when T appears with p in the derivative.

The Maxwell's equations are supremely powerful, because they relate partial differentials that are easy to measure, shown to the right of the equality, with partial differentials that are very difficult to measure shown to the left of the equality.

Example. Find expressions for $\left(\frac{\partial C_p}{\partial p}\right)_T$ and $\left(\frac{\partial C_V}{\partial V}\right)_T$.

Solution Firstly, note that:

$$C_V = \left(\frac{\partial Q}{\partial T}\right)_V \quad (5.6.7)$$

Also:

$$dU = dQ + dW = TdS - pdV \quad (5.6.8)$$

Taking the volume to be constant and differentiating with respect to temperature

$$\left(\frac{\partial U}{\partial T}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V \quad (5.6.9)$$

Since in a reversible process $dU = dQ$, we find that:

$$\left(\frac{\partial Q}{\partial T}\right)_V = C_V = T\left(\frac{\partial S}{\partial T}\right)_V \quad (5.6.10)$$

and similarly:

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_p \quad (5.6.11)$$

Therefore:

$$\left(\frac{\partial C_p}{\partial V}\right)_T = \left[\frac{\partial}{\partial p} T \left(\frac{\partial S}{\partial T}\right)_p\right]_T \quad (5.6.12)$$

$$= T \left[\frac{\partial}{\partial T} \left(\frac{\partial S}{\partial p}\right)_T\right]_p = -T \left(\frac{\partial^2 V}{\partial T^2}\right)_p \quad (5.6.13)$$

Similarly:

$$\left(\frac{\partial C_V}{\partial V}\right)_T = T \left(\frac{\partial^2 p}{\partial T^2}\right)_V \quad (5.6.14)$$

In general, the following strategies may help in solving problems regarding Maxwell's relations:

(i) write down the differential of a thermodynamic potential:

$$df = \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy \quad (5.6.15)$$

(ii) use the reciprocal theorem:

$$\left(\frac{\partial x}{\partial z}\right)_y = \frac{1}{\left(\frac{\partial z}{\partial x}\right)_y} \quad (5.6.16)$$

(iii) Use the reciprocity theorem:

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \quad (5.6.17)$$

(iv) Use the heat capacities:

$$\frac{C_V}{T} = \left(\frac{\partial S}{\partial T}\right)_V, \quad \frac{C_p}{T} = \left(\frac{\partial S}{\partial T}\right)_p \quad (5.6.18)$$

(v) Define a generalized susceptibility as how a variable changes when a generalized force, a variable which is a partial derivative of the internal energy, such as T or p.

$$\left\{ \begin{array}{ll} \text{Isobaric expansivity :} & \beta_p = \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_p \\ \text{Adiabatic expansivity :} & \beta_V = \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_S \\ \text{Isothermal compressibility :} & \kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_T \\ \text{Adiabatic compressibility :} & \kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_S \end{array} \right. \quad (5.6.19)$$

Note the expansivities express fractional change in volume as temperature varies, whereas compressibilities fractional change in volume as pressure varies.

Example. Show that $C_p - C_V = \frac{VT\beta_p^2}{\kappa_T}$ by considering $S = S(T, V)$.

Solution We begin by writing the differential entropy as:

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV \quad (5.6.20)$$

so that:

$$\left(\frac{\partial S}{\partial T}\right)_p = \frac{C_p}{T} = \left(\frac{\partial S}{\partial T}\right)_V + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p = \frac{C_V}{T} + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p \quad (5.6.21)$$

Using Maxwell's relations and the reciprocity theorem:

$$\left(\frac{\partial S}{\partial V}\right)_T = +\left(\frac{\partial p}{\partial T}\right)_V = -\left(\frac{\partial p}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p \quad (5.6.22)$$

so that:

$$\left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p = \frac{(V\beta_p)^2}{V\kappa_T} = \frac{V\beta_p^2}{\kappa_T} \quad (5.6.23)$$

Finally we find that:

$$C_p - C_V = T \left[\left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p \right] = \frac{VT\beta_p^2}{\kappa_T} \quad (5.6.24)$$

as desired. ◀

Example. Find the entropy of 1 mole of ideal gas.

Solution For this quantity of an ideal gas, the state equation is $pV = RT$. Hence, we may write $S = S(T, V)$, or, in differential form:

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV \quad (5.6.25)$$

$$= \frac{C_V}{T} dT + \left(\frac{\partial p}{\partial T}\right)_V dV \quad (5.6.26)$$

Also:

$$\left(\frac{\partial p}{\partial T}\right)_V = \frac{\partial}{\partial T} \frac{RT}{V} = \frac{R}{V} \quad (5.6.27)$$

so that:

$$dS = \frac{C_V}{T} dT + \frac{R}{V} dV \quad (5.6.28)$$

which upon integration gives:

$$S = C_V \ln T + R \ln V + \text{const} \quad (5.6.29)$$



Rods, bubbles and gases

6.1 Elastic materials

We look at a rod of cross-section A and length L , held at a constant temperature T . If the rod is made of elastic material, then it can be placed under tension df . We then define the isothermal Young's modulus E_T as the ratio of stress $\sigma = df/A$ and strain $\epsilon = dL/L$, so:

$$E_T = \frac{L}{A} \left(\frac{\partial f}{\partial L} \right)_T \quad (6.1.1)$$

We also define the linear expansivity as:

$$\alpha_f = \frac{1}{L} \left(\frac{\partial L}{\partial T} \right)_f \quad (6.1.2)$$

It follows that a wire held at constant length is characterized by a tension-temperature relation:

$$\left(\frac{\partial f}{\partial T} \right)_L = - \left(\frac{\partial f}{\partial L} \right)_T \left(\frac{\partial L}{\partial T} \right)_f = -AE_T \alpha_f \quad (6.1.3)$$

So, for a material with positive linear expansivity, tension decreases as temperature increases. This effect is particularly familiar to stringed-instrument players, where the metal strings slacken when placed in hot environments.

Rubber bands, on the other hands, have negative expansivity, so that as we increase the temperature, the tension too increases. Consequently, if we hang a weight at the end of a rubber band, and subsequently heat it up, the increase in tension will cause the weight to rise.

Further, we may work with the Hemholtz potential and write:

$$dF = -SdT + fdL = - \left(\frac{\partial S}{\partial T} \right)_L dT + \left(\frac{\partial f}{\partial L} \right)_T dL \quad (6.1.4)$$

implying that:

$$S = -\left(\frac{\partial S}{\partial T}\right)_L \quad (6.1.5)$$

$$f = \left(\frac{\partial f}{\partial L}\right)_T \quad (6.1.6)$$

Following the derivation of the Maxwell relations, we may then write:

$$\left(\frac{\partial S}{\partial L}\right)_T = -\left(\frac{\partial f}{\partial T}\right)_L = AE_T\alpha_f \quad (6.1.7)$$

Therefore, for materials such as a metal wire, the act of stretching it isothermally results in an increase in entropy. Indeed, the expansion distorts the crystallites which were previously carefully aligned in a steady state.

More specifically, in the case of an isothermal, reversible extension, the heat absorbed is:

$$\Delta Q = T\Delta S = AE_T T\alpha_f \Delta L \quad (6.1.8)$$

However, in the case of a rubber band, $\alpha_f < 0$, so that an isothermal extension results in entropy decrease. Indeed, unlike the case of metal wires, the rubber molecules are initially coiled up, and stretching it aligns these molecules introducing more order to the system and reducing entropy.

One can indeed show that increasing the link distance between rubber molecules decreases the entropy.

One final consideration that must be made is what happens to the internal energy of the rod when it is extended isothermally?

We can write the first law of thermodynamics as:

$$dU = TdS + fdL \implies \left(\frac{\partial U}{\partial L}\right)_T = T\left(\frac{\partial S}{\partial L}\right)_T + f = AT E_T \alpha_f + f \quad (6.1.9)$$

This term is the sum of the work done on the rod by the tension force, and the heat flowing into the rod due to the change of length.

6.2 Surface tension

Consider the surface of some liquid surface of area A . To change the area of the surface by dA , work must be done against the cohesive forces of the interface. This work can be quantified as:

$$dW = \gamma dA \quad (6.2.1)$$

where γ is known as the surface tension. Now consider the set up shown below, with a piston pushing down and doing work $dW = pdV$ on an incompressible liquid.

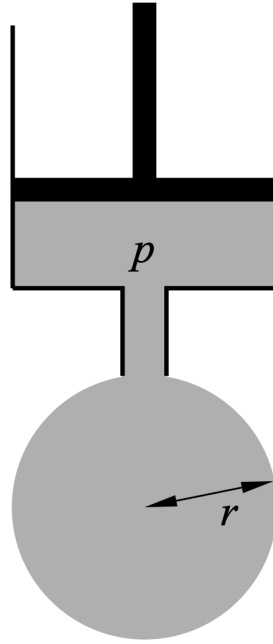


Figure 6.1. Piston pushing on spherical droplet held by surface tension

The liquid forms a spherical droplet suspended from the end of a thin pipe connected to the piston, whose change in volume is $dV = 4\pi r^2 dr$. Then:

$$dA = 4\pi(r + dr)^2 - 4\pi r^2 \approx 8\pi r dr \implies dW = 8\pi\gamma r dr \quad (6.2.2)$$

At the same time, $dW = pdV = 4\pi pr^2 dr$, and equating the two expressions for the work done we find that:

$$8\pi\gamma r dr = 4\pi pr^2 dr \implies p = \frac{2\gamma}{r} \quad (6.2.3)$$

This is the pressure difference between inside and outside the droplet.

For the case of a spherical bubble, then the surface tension acts on two surfaces. Hence, $da = 16\pi\gamma r dr$ and thus the pressure difference is $p = \frac{4\gamma}{r}$.

The first law of thermodynamics for our surface takes the following form:

$$dU = TdS + \gamma dA \quad (6.2.4)$$

so that:

$$\left(\frac{\partial U}{\partial A}\right)_T = T\left(\frac{\partial S}{\partial A}\right)_T + \gamma \quad (6.2.5)$$

The Hemholtz function can be written as:

$$dF = -SdT + \gamma dA \quad (6.2.6)$$

The Maxwell's relation deriving from these expressions is:

$$\left(\frac{\partial S}{\partial A}\right)_T = -\left(\frac{\partial \gamma}{\partial T}\right)_A \quad (6.2.7)$$

Therefore, (6.2.5) can be written as:

$$\left(\frac{\partial U}{\partial A}\right)_T = \gamma - \left(\frac{\partial \gamma}{\partial T}\right)_A \quad (6.2.8)$$

Note that usually, surface tension decreases as we increase temperature, so that both terms in the above expression are positive. The first term describing the energy flowing in the surface due to the external work, and the second expressing the heat flow into the surface.

For an isothermal stretching of a surface, the heat absorbed by the surface is:

$$\Delta Q = T \left(\frac{\partial S}{\partial A}\right)_T \Delta A = -T \Delta A \left(\frac{\partial \gamma}{\partial T}\right)_A > 0 \quad (6.2.9)$$

6.3 Magnetization

We know from the electromagnetism course that the work supplied to a magnetic dipole immersed in a magnetic field $d\mathbf{B}$ is:

$$dW = -\mathbf{m} \cdot d\mathbf{B} \quad (6.3.1)$$

A system of non-interacting magnetic moments is called a paramagnet. When we apply a magnetic field in the vicinity of a paramagnet, these magnetic moments line up exhibiting paramagnetism.

The first law of thermodynamics can then be written as:

$$dU = TdS - mdB \quad (6.3.2)$$

where $m = MV$ is the magnetic moment, M the magnetization and V the volume. B , the magnetic field is expressed as $B = \mu_0(H + M)$. The magnetic susceptibility is defined as:

$$\chi = \lim_{H \rightarrow 0} \frac{M}{H} \quad (6.3.3)$$

and is usually much much smaller than 1. Consequently $B \approx \mu_0 H$, and the susceptibility may be approximated as:

$$\chi \approx \frac{\mu_0 M}{B} \quad (6.3.4)$$

Now the Hemholtz differential may be expressed as:

$$dF = -SdT - mdB \quad (6.3.5)$$

yielding:

$$\left(\frac{\partial S}{\partial B}\right)_T = \left(\frac{\partial m}{\partial T}\right)_B \approx \frac{VB}{\mu_0} \left(\frac{\partial \chi}{\partial T}\right)_B \quad (6.3.6)$$

Therefore for an isothermal change of the magnetic field:

$$\Delta Q = T \left(\frac{\partial S}{\partial B}\right)_T \Delta B = \frac{TVB}{\mu_0} \left(\frac{\partial \chi}{\partial T}\right)_B \quad (6.3.7)$$

Now, paramagnetic systems obey Curie's law $\chi \propto \frac{1}{T}$ as shall be shown later. It follows that $\left(\frac{\partial \chi}{\partial T}\right)_B < 0$ and that consequently $\Delta Q < 0$, heat is emitted by the paramagnet.

The change in temperature in an adiabatic change, instead, is:

$$\left(\frac{\partial T}{\partial B}\right)_S = -\left(\frac{\partial T}{\partial S}\right)_B \left(\frac{\partial S}{\partial B}\right)_T \quad (6.3.8)$$

where we used the reciprocity theorem. Since by definition:

$$C_B = T \left(\frac{\partial S}{\partial T}\right)_B \quad (6.3.9)$$

we may finally write:

$$\left(\frac{\partial T}{\partial B}\right)_S = -\frac{TVB}{\mu_0 C_B} \left(\frac{\partial \chi}{\partial T}\right)_B > 0 \quad (6.3.10)$$

It follows that adiabatic demagnetization leads to a drop in temperature. This method may be used in experiments requiring temperatures as low as mK or even μK

Our physical intuition backs up this result. Indeed, consider a sample of paramagnetic material. Before applying a magnetic field, all the magnetic moments have a stochastic distribution (assuming there is no interaction between them). As we apply a magnetic field, the moments will tend to line up, decreasing the entropy of the system. However, as we raise the temperature, the thermal energy $k_B T$ begins to contrast the magnetic energy, until at some point we reach a random distribution again. This state is characterized by a very high entropy.

At a low temperature instead, all the magnetic moments align. The ways we can arrange the system is therefore unique, and hence:

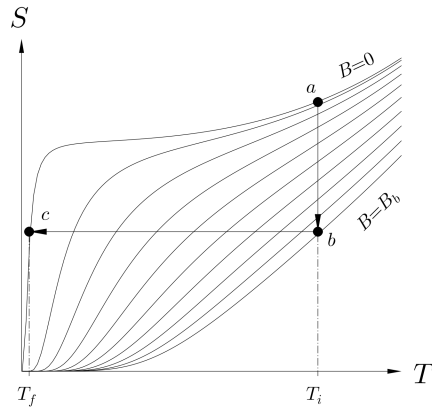
$$S = k_B \ln \Omega = k_B \ln 1 = 0 \quad (6.3.11)$$

as required.

How does this magnetic cooling occur? The typical process is depicted below:

Firstly, ($a \rightarrow b$) isothermal magnetization is initiated. At constant temperature, the paramagnetic material is subjected to an increasing magnetic field. This leads to a decrease in entropy.

6.3. MAGNETIZATION



Then, ($b \rightarrow c$) adiabatic demagnetization occurs. The sample is thermally isolated and the magnetic field is slowly reduced to zero. This lowers the temperature of the paramagnet, while keeping the entropy constant.

How does this entropy remain the same? In truth, there is an entropy exchange between the phonons holding together the lattice and the spins of the magnetic moments. As we lower the magnetic field the phonons, the so-called "spring forces" holding together the molecules in the lattice, lose entropy, and exchange it to the magnetic moments, the spin, of the paramagnet, which randomize their orientation.

Part II

Kinetic Theory of Gases

Speed distribution

7.1 Energetic considerations

Consider a sample of gas, constituted by N molecules which we shall model as particles of mass m moving with velocity $\mathbf{v} = (v_x, v_y, v_z)$. Ignoring rotational and vibrational degrees of freedom, the mean energy of is due only to the mean kinetic energy, thus:

$$\langle E \rangle = N \left\langle \frac{mv^2}{2} \right\rangle \quad (7.1.1)$$

Suppose the gas overall is moving with velocity $\langle \mathbf{v} \rangle = \mathbf{u}$. Then, we define the individual velocity of a molecule in this gas as $\mathbf{w} = \mathbf{v} - \mathbf{u}$, that is, the velocity of the molecule as viewed in the frame moving with the gas. Then, since $\langle \mathbf{w} \rangle = 0$ by definition, we find that:

$$\langle E \rangle = \frac{Nm}{2} \langle |\mathbf{u} + \mathbf{w}|^2 \rangle = \frac{Mu^2}{2} + N \left\langle \frac{mw^2}{2} \right\rangle \quad (7.1.2)$$

where $M = Nm$ is the mass of the entire gas. The first term is clearly the kinetic energy of the entire system, this is the one we commonly associate with energy. However, there is an additional energy due to the particle's random motion. This is the internal energy we studied in classical thermodynamics, which is only visible when studying the gas at a microscopic scale.

If we assume that $\mathbf{u} = 0$ by changing frame of reference, then:

$$\langle E \rangle = U = N \left\langle \frac{mw^2}{2} \right\rangle \quad (7.1.3)$$

How does this compare with the exact energy of the gas? The exact energy is just the sum of the kinetic energy of each molecule, hence:

$$E = \sum_i \frac{mv_i^2}{2} \quad (7.1.4)$$

so that:

$$N \left\langle \frac{mw^2}{2} \right\rangle = \sum_i \left\langle \frac{mv_i^2}{2} \right\rangle \quad (7.1.5)$$

Now, since v_i and v_j are independent for different particles, $\langle f(v_i)g(v_j) \rangle = \langle f(v_i) \rangle \langle g(v_j) \rangle$. Note that:

$$E^2 = \left(\sum_i \frac{mv_i^2}{2} \right)^2 = \sum_{i,j} \frac{m^2 v_i^2 v_j^2}{4} \neq \sum_i \left(\frac{mv_i^2}{2} \right)^2 \quad (7.1.6)$$

Keeping this in mind, we get that:

$$\langle E^2 \rangle - U^2 = \sum_{i,j} \left\langle \frac{m^2 v_i^2 v_j^2}{4} \right\rangle - \left(\sum_i \left\langle \frac{mv_i^2}{2} \right\rangle \right)^2 \quad (7.1.7)$$

$$= \sum_i \left\langle \frac{m^2 v_i^4}{4} \right\rangle + \sum_{i \neq j} \left\langle \frac{mv_i^2}{2} \right\rangle \left\langle \frac{mv_j^2}{2} \right\rangle - \left(\sum_i \left\langle \frac{mv_i^2}{2} \right\rangle \right)^2 \quad (7.1.8)$$

$$= N \left\langle \frac{m^2 v^4}{4} \right\rangle + N(N-1) \left\langle \frac{mv^2}{2} \right\rangle^2 - \left(N \left\langle \frac{mv^2}{2} \right\rangle \right)^2 \quad (7.1.9)$$

$$= \frac{Nm^2}{4} (\langle v^4 \rangle - \langle v^2 \rangle^2) \quad (7.1.10)$$

Hence the standard deviation in the energy is:

$$\sigma_E = \sqrt{\langle E^2 \rangle - \langle E \rangle^2} = \sqrt{\frac{Nm^2}{4} (\langle v^4 \rangle - \langle v^2 \rangle^2)} \quad (7.1.11)$$

The relative error is then:

$$\frac{\sigma_E}{U} = \frac{\sqrt{\frac{Nm^2}{4} (\langle v^4 \rangle - \langle v^2 \rangle^2)}}{N \left\langle \frac{mv^2}{2} \right\rangle} = \frac{1}{\sqrt{N}} \sqrt{\frac{\langle v^4 \rangle}{\langle v^2 \rangle^2} - 1} \quad (7.1.12)$$

For large numbers of particles, this relative error is very very small. Hence, we can indeed regard large distributions of particles as sharply peaked near $\langle E \rangle$. In other words, as the number of particles increases, the relative amount that each individual particle's properties varies from the averaged value of the property decreases. This is exactly the prediction made by taking the thermodynamic limit.

7.2 Pressure

Consider a container immersed in a gas, how can we model the pressure exerted on the walls of the container?

Let's assume that the particles hitting the wall are colliding elastically, and let the z -axis be aligned with the normal to one of the container walls. Then, the velocity components of any given particle before and after the collision are given by:

$$v_z^{\text{after}} = -v_z^{\text{before}} \implies \Delta p_m = 2mv_z \quad (7.2.1)$$

where Δp_m , not to be confused with pressure, is the change in momentum. For the sample of $N(v_z)$ particles whose v_z lie in the infinitesimal interval $[v_z, v_z + dz]$, we define the

differential particle flux as:

$$d\Phi(v_z) = \frac{dN(v_z)}{At} \quad (7.2.2)$$

that is, the number of particles with such speeds hitting the box per unit time per unit area. Their contribution to the pressure is:

$$dp(v_z) = \underbrace{\frac{\Delta p_m N(v_z)}{t}}_{\text{force on wall}} \frac{1}{A} = 2mv_z d\Phi(v_z) \quad (7.2.3)$$

Suppose the ratio of particles with speed in the interval $[v_z, v_z + dv_z]$ is given by $f(v_z)dv_z$, so that:

$$N(v_1, v_2) = \int_{v_1}^{v_2} f(v_z)dv_z \quad (7.2.4)$$

is the number of particles with speed between v_1 and v_2 . It is the probability density function of velocities, also known as the **velocity distribution**. Then, $dN(v_z)$ is the density of particles in the gas, $n = \frac{N}{V}$, times the volume in which the particles must reside for them to hit the container in time t , $Av_z t$, times the ratio of particles with speed in the interval $[v_z, v_z + dv_z]$, $f(v_z)dv_z$:

$$dN(v_z) = Av_z t \cdot n \cdot f(v_z)dv_z \quad (7.2.5)$$

Inserting this into (7.2.2) we find:

$$dp(v_z) = 2mnv_z^2 f(v_z)dv_z \quad (7.2.6)$$

and integrating for all particles (so with v_z ranging from 0, still particles, to ∞):

$$p = \int_0^\infty 2mnv_z^2 f(v_z)dv_z = \int_{-\infty}^\infty mnv_z^2 f(v_z)dv_z = mn \langle v_z^2 \rangle \quad (7.2.7)$$

where we assumed that the velocity distribution is symmetrical, $f(v_z) = -f(-v_z)$. This is justified by the fact that there should be no preferred motion for the particles to move. Alternatively, one could also consider a rotation in frame of reference, observing a particle hitting the container's wall from two opposite points of view makes no change physically.

In three dimensions, we can introduce the velocity distribution $f(\mathbf{v})$. The derivation is exactly the same, and we find that:

$$p = mn \int v_z^2 f(\mathbf{v}) d^3\mathbf{v} = mn \langle v_z^2 \rangle \quad (7.2.8)$$

Furthermore, for isotropic distributions, where, in addition to orientations, no directions are preferred, then:

$$\langle v_z^2 \rangle = \langle v_x^2 \rangle = \langle v_y^2 \rangle = \frac{1}{3} \langle v^2 \rangle \quad (7.2.9)$$

so the pressure is:

$$p = \frac{1}{3} mn \langle v^2 \rangle \quad (7.2.10)$$

The internal energy is:

$$U = \frac{1}{2}mN \langle v^2 \rangle = \frac{3pV}{2} \iff p = \frac{2U}{3V} = \frac{2}{3}u \quad (7.2.11)$$

where u is the energy density.

Moreover, isotropy forces the velocity distribution to be independent of direction. Using spherical coordinates in the (v_x, v_y, v_z) -space:

$$f(\mathbf{v}) \rightarrow f(v, \theta, \phi) \quad (7.2.12)$$

and thus:

$$f(\mathbf{v})dv_xdv_ydv_z = f(\mathbf{v}) \left| \frac{\partial(v_x, v_y, v_z)}{\partial(v, \theta, \phi)} \right| dv d\theta d\phi \quad (7.2.13)$$

$$= f(\mathbf{v})v^2 \sin \theta dv d\theta d\phi = f(v, \theta, \phi) dv d\theta d\phi \quad (7.2.14)$$

implying that:

$$f(v, \theta, \phi) = f(\mathbf{v})v^2 \sin \theta \quad (7.2.15)$$

The particle speed distribution is then:

$$g(v) = \int_0^\pi \int_0^{2\pi} f(v, \theta, \phi) d\phi d\theta = 4\pi v^2 f(\mathbf{v}) \quad (7.2.16)$$

We can use some geometric insight to get to the same result. In the velocity space, the ratio of particles with speed between v and $v + dv$ is given by the volume of the spherical shell of width dv :

$$g(v)dv = 4\pi v^2 dv \cdot f(\mathbf{v}) \quad (7.2.17)$$

coinciding with the results found above.

Example. Let us consider an anisotropic system, where there exists one special direction in space (call it z), which affects the distribution of particle velocities.

- a) What is the expression for pressure p_{\parallel} on the wall perpendicular to the z -axis?
What is the expression for pressure p_{\perp} on a wall parallel to the z -axis?
- b) What is a pressure normal to z at an angle θ ?

Solution a) Since we now have a special direction z , we may use cylindrical coordinates-
The velocity distribution will depend on v_s and v_z . Then:

$$f(\mathbf{v}) = dv_x dv_y dv_z = f(\mathbf{v})v_s d\phi dv_s dv_z \quad (7.2.18)$$

The pressure on a wall with normal along z evaluates to:

$$P = mn \langle v_z^2 \rangle \quad (7.2.19)$$

For a wall whose normal is parallel to z instead, we need to repeat our calculations, but this time with $v_z \mapsto v_s$. We find that:

$$\Delta p = 2mv_s \implies dP(v_s) = 2mv_s d\Phi(v_s) \quad (7.2.20)$$

Since $d\Phi(v_s) = v_s n f(v_s) dv_s$ we get $dP(v_s) = 2mv_s^2 n f(v_s) dv_s$. Finally, using the fact that the distribution in xy plane is isotropic, $f(v_s) = f(-v_s)$:

$$P = mn \int_{-\infty}^{\infty} v_s^2 f(v_s) dv_s = mn \langle v_s^2 \rangle \quad (7.2.21)$$

b) Consequently, for a wall at an angle θ to z .

$$P = P_{\parallel} \cos \theta + P_{\perp} \sin \theta \quad (7.2.22)$$

$$= mn (\langle v_z^2 \rangle \cos \theta + \langle v_s^2 \rangle \sin \theta) \quad (7.2.23)$$

where we decomposed the pressure on the wall into a component along z and perpendicular to z .



7.3 Deriving the Boltzmann velocity distribution

We now face the task of determining what form $f(\mathbf{v})$ takes. We will assume that:

- (i) particle-particle interactions are negligible, except for ideal elastic collisions
- (ii) they are point particles
- (iii) classical, non-relativistic particles

Suppose then that we have taken this gas in a container of volume V , and we have waited for macroscopic equilibrium to occur, so that the velocity distribution is time-independent.

We can argue that in space there are no preferred directions, so we can use isotropy. Moreover, we may also argue that the velocity vector's components must be independent:

$$f(\mathbf{v}) = g(v_x^2)g(v_y^2)g(v_z^2) = h(v^2) \iff \ln(g(v_x^2)g(v_y^2)g(v_z^2)) = \ln h(v^2) \quad (7.3.1)$$

If we let $\phi(v_x^2) = \ln g(v_x^2)$ and $\varphi(v^2) = \ln h(v^2)$:

$$\phi(v_x^2) + \phi(v_y^2) + \phi(v_z^2) = \varphi(v_x^2 + v_y^2 + v_z^2) \quad (7.3.2)$$

If we take the derivative with respect to v_x^2 :

$$\partial v_x^2 \phi(v_x^2) = \partial v_x^2 \varphi(v_x^2 + v_y^2 + v_z^2) \quad (7.3.3)$$

Similarly:

$$\partial v_y^2 \phi(v_x^2) = \partial v_y^2 \varphi(v_x^2 + v_y^2 + v_z^2) = \partial v_x^2 \varphi(v_x^2 + v_y^2 + v_z^2) \quad (7.3.4)$$

and

$$\partial v_z^2 \phi(v_x^2) = \partial v_z^2 \varphi(v_x^2 + v_y^2 + v_z^2) = \partial v_x^2 \varphi(v_x^2 + v_y^2 + v_z^2) \quad (7.3.5)$$

implying that:

$$\partial v_x^2 \phi(v_x^2) = \partial v_y^2 \phi(v_y^2) = \partial v_z^2 \phi(v_z^2) = \quad (7.3.6)$$

The only possible family of functions ϕ satisfying (7.3.6) are functions linear in their argument (we use $-\alpha$ as a convention, it turns out that α is positive):

$$\phi(v_x^2) = -\alpha v_x^2 + \beta, \quad \phi(v_y^2) = -\alpha v_y^2 + \beta, \quad \phi(v_z^2) = -\alpha v_z^2 + \beta, \quad (7.3.7)$$

so that:

$$\varphi(v^2) = -\alpha v^2 + 3\beta \quad (7.3.8)$$

Thus:

$$f(\mathbf{v}) = e^{3\beta} e^{-\alpha v^2} = c e^{-\alpha v^2} \quad (7.3.9)$$

We can now use the normalization condition:

$$1 = \int f(\mathbf{v}) d^3 \mathbf{v} \quad (7.3.10)$$

$$= c \int e^{\alpha v_x^2} dv_x \int e^{\alpha v_y^2} dv_y \int e^{\alpha v_z^2} dv_z \quad (7.3.11)$$

$$= c \left(\frac{\pi}{\alpha} \right)^{3/2} \quad (7.3.12)$$

$$\iff c = \left(\frac{\alpha}{\pi} \right)^{3/2} \quad (7.3.13)$$

Thus, the velocity distribution becomes:

$$f(\mathbf{v}) = \left(\frac{\alpha}{\pi} \right)^{3/2} e^{-\alpha v^2} \quad (7.3.14)$$

Let us analyze the dimensions of α . Clearly, $\frac{1}{\sqrt{\alpha}} = v_{th}$ where v_{th} is some velocity characteristic of our gas sample. Then:

$$\boxed{f(\mathbf{v}) = \frac{1}{(\sqrt{\pi} v_{th})^3} e^{-v^2/v_{th}^2}} \quad (7.3.15)$$

and the related speed distribution is:

$$\boxed{g(v) = \frac{4\pi v^2}{(\sqrt{\pi} v_{th})^3} e^{-v^2/v_{th}^2}} \quad (7.3.16)$$

Note that:

$$\langle v \rangle = \int_0^\infty v g(v) dv = \int_0^\infty \frac{4\pi v^3}{(\sqrt{\pi} v_{th})^3} e^{-v^2/v_{th}^2} dv \quad (7.3.17)$$

$$= \frac{4\pi v_{th}^4}{2(\sqrt{\pi} v_{th})^3} \quad (7.3.18)$$

$$= \frac{2}{\sqrt{\pi}} v_{th} \quad (7.3.19)$$

Moreover:

$$\frac{dg(v)}{dv} = \frac{4}{\sqrt{\pi} v_{th}} \frac{2v e^{-v^2/v_{th}^2} (v_{th}^2 - v^2)}{v_{th}^2} = 0 \iff v = v_{th} \quad (7.3.20)$$

In other words, v_{th} is the most likely speed, since the speed distribution is peaked there.

Now note that:

$$P = \frac{1}{3} mn \int v^2 \frac{1}{(\sqrt{\pi} v_{th})^3} e^{-v^2/v_{th}^2} d^3 \mathbf{v} = \frac{nm v_{th}^2}{2} \quad (7.3.21)$$

so we have that:

$$v_{th} = \sqrt{\frac{2P}{nm}} \quad (7.3.22)$$

For an ideal gas:

$$P = nk_B T \implies \boxed{\frac{mv_{th}^2}{2} = k_B T} \quad (7.3.23)$$

This is a monumental result. The temperature of an ideal gas is the kinetic energy of a particle moving with the most likely speed of the Maxwell speed distribution. Using this expression, the Maxwell distribution takes its final form:

$$\boxed{f(\mathbf{v}) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp\left(-\frac{mv^2}{2k_B T} \right)} \quad (7.3.24)$$

Also, we may rewrite $\langle v \rangle$ as:

$$\langle v \rangle = \sqrt{\frac{8k_B T}{\pi m}} \quad (7.3.25)$$

and similarly:

$$\langle v^2 \rangle = \frac{3k_B T}{m} \quad (7.3.26)$$

Then, we may retrieve the ideal gas law:

$$p = \frac{1}{3} Nm \langle v^2 \rangle = nk_B T \implies pV = Nk_B T \quad (7.3.27)$$

7.4 Effusion

We place an ideal gas in a container with a small hole in it. The diameter d of the hole is much smaller than the average path that particles travel between each collision, known as the particle mean free path λ_{mfp} . This ensures that the particles are removed randomly, without changing the overall velocity distribution.

The differential of the particle flux through the hole (which is the number of particles per unit area per unit time with velocities lying in $[\mathbf{v}, \mathbf{v} + d\mathbf{v}]$) is given by:

$$d\Phi(\mathbf{v}) = nv_z f(\mathbf{v}) d^3\mathbf{v} = nv^3 f(\mathbf{v}) dv \cos\theta \sin\theta d\theta d\phi \quad (7.4.1)$$

since $v_z = v \cos\theta$, and $d^3\mathbf{v} = v^2 \sin\theta dv d\theta d\phi$ for an isotropic distribution inside the container. For an isotropic sample, we may further assume that $f(\mathbf{v}) = f(v)$.

It follows that the distribution of the particles exiting the container is not isotropic, due to the extra $\cos\theta$ factor. This is physically interpreted by the fact that particles with a small θ , that is perpendicular to the wall with the hole, are more likely to be ejected. Similarly, the distribution is not Maxwellian due to the extra v factor, since faster particles are more likely to exit the container.

The flux of the particles with speed within $[v, v + dv]$ can be found by integrating¹:

$$d\hat{\Phi}(v) = nv^3 f(v) dv \int_0^{\pi/2} d\theta \cos\theta \sin\theta \int_0^{2\pi} d\phi = \pi nv^3 f(v) dv = \frac{1}{4} nvg(v) dv \quad (7.4.2)$$

The total flux of effusing particles is then:

$$\Phi = \int_0^{\infty} dv \frac{1}{4} nvg(v) = \frac{1}{4} n \langle v \rangle \quad (7.4.3)$$

Since we assume a Maxwell speed distribution inside the container, we find that:

$$\Phi = \frac{n}{4} \sqrt{\frac{8k_B T}{\pi m}} = \frac{P}{\sqrt{2\pi m k_B T}} \quad (7.4.4)$$

This is the number of particles effusing per unit time, per unit area through the hole. The total effusion rate is given by $\Phi \cdot A$.

Example.

a) A gas effuses into a vacuum through a small hole of area A and collimated by passing through a very small circular hole of radius a , in a screen a distance $d \gg a$ from the first hole. Calculate the rate at which particles emerge from the circular hole. b) Show that if a gas were allowed to leak into an evacuate sphere and the particles condensed where they first hit the surface they would form a uniform coating.

¹note that the upper limit for θ is $\frac{\pi}{2}$, since we only consider the particles moving towards the hole

Solution a) We know that the number of particles passing through the first hole per unit time is the particle flux times the surface of the first hole:

$$T' = \Psi \cdot a = \frac{1}{4}nA \langle v \rangle \quad (7.4.5)$$

The ratio of the particles which then pass through the second collimating hole is:

$$T = An \int_0^\infty v^3 \int_0^\pi \int_0^{\theta'} v f(v) \cdot v^2 \sin \theta dv d\theta d\phi \quad (7.4.6)$$

$$= 2\pi An \int_0^\infty v^3 f(v) \left[\frac{1}{2} \sin^2 \theta \right]_0^{\theta'} dv \quad (7.4.7)$$

$$= \frac{An}{4} \sin^2 \theta' \int_0^\infty v g(v) dv \quad (7.4.8)$$

$$= \frac{An \langle v \rangle}{4} \sin^2 \theta' \quad (7.4.9)$$

Since $d \gg a$, we may use small angle approximation to write $\sin \theta' \approx \frac{a}{d}$ so that:

$$T = \frac{Ana^2 \langle v \rangle}{4d^2} \quad (7.4.10)$$

b) The differential flux through the hole into the evacuated sphere is:

$$d\Phi = nv^3 f(v) \sin \theta \cos \theta dv d\theta d\phi \quad (7.4.11)$$

$$= \frac{1}{2}nv^3 f(v) \sin 2\theta dv d\theta d\phi \quad (7.4.12)$$

$$= \frac{1}{4}nv^3 f(v) \sin 2\theta' dv d\theta' d\phi \quad (7.4.13)$$

where $\theta' = 2\theta$. Taking the origin to be centered at the sphere's center, then we see that θ' becomes the angular measure and replaces θ . We see that this does indeed produce a uniform coating. ◀

Example. A gas is a mixture of hydrogen and deuterium in the proportion 7000 : 1. As the gas effuses through a small hole from a vessel at constant temperature into a vacuum, the composition of the remaining mixture changes. By what factor will the pressure in the vessel have fallen when the remaining mixture consists of hydrogen and deuterium in the proportion 700 : 1.

Solution The flux of particles effusing through the hole is:

$$\Phi_H = \frac{1}{4} n_H \langle v_H \rangle = n_H \sqrt{\frac{k_B T}{2\pi m_H}} = \frac{p_H}{\sqrt{2\pi m_H k_B T}} \quad (7.4.14)$$

$$\Phi_D = \frac{1}{4} n_D \langle v_D \rangle = n_D \sqrt{\frac{k_B T}{2\pi m_D}} = \frac{p_D}{\sqrt{2\pi m_D k_B T}} \quad (7.4.15)$$

Using the definition of particle flux as the number of particles moving through a surface of unit area per unit time, then taking a volume V and recalling $N = n \cdot V$ for a homogeneous system:

$$\frac{1}{A} \frac{dN_H}{dt} = -\frac{N_H}{V} \sqrt{\frac{k_B T}{2\pi m_H}} \quad (7.4.16)$$

$$\frac{1}{A} \frac{dN_D}{dt} = -\frac{N_D}{V} \sqrt{\frac{k_B T}{2\pi m_D}} \quad (7.4.17)$$

Therefore, defining the characteristic time constants $\tau_H = \frac{V}{A} \sqrt{\frac{2\pi m_H}{k_B T}}$ and $\tau_D = \tau_H = \frac{V}{A} \sqrt{\frac{2\pi m_D}{k_B T}}$ then:

$$N_H = N_{0H} e^{-t/\tau_H} \quad (7.4.18)$$

$$N_D = N_{0D} e^{-t/\tau_D} \quad (7.4.19)$$

It follows that if at time $t = 0$ we have $\frac{N_{0H}}{N_{0D}} = 7000$, and at t we have $\frac{N_H}{N_D} = 700$ then:

$$e^{t/\tau_H - t/\tau_D} = 10 \implies t = \ln 10 \frac{\tau_H \tau_D}{\tau_D - \tau_H} \quad (7.4.20)$$

Using the ideal gas law, $\frac{pV}{k_B T} = N$ we find that:

$$p_H = \frac{k_B T N_{0H}}{V} e^{-t/\tau_H} = p_{0H} e^{-t/\tau_H} \quad (7.4.21)$$

$$p_D = \frac{k_B T N_{0D}}{V} e^{-t/\tau_D} = p_{0D} e^{-t/\tau_D} \quad (7.4.22)$$

so that:

$$\frac{p}{p_0} = \frac{p_{0H} e^{-t/\tau_H} + p_{0D} \frac{k_B T N_{0D}}{V} e^{-t/\tau_D}}{p_{0H} + p_{0D}} \quad (7.4.23)$$

$$= \frac{p_{0H} 10^{-\frac{t}{\tau_D - \tau_H}} + p_{0D} 10^{-\frac{t}{\tau_D - \tau_H}}}{p_{0H} + p_{0D}} \quad (7.4.24)$$

$$= \frac{N_{0H}}{N_0} 10^{-\frac{t}{\tau_D - \tau_H}} + \frac{N_{0D}}{N_0} 10^{-\frac{t}{\tau_D - \tau_H}} \quad (7.4.25)$$

$$\approx 0.000412 \quad (7.4.26)$$



7.5 Collisions

In the previous section, we mentioned the particle mean free path, but how would we calculate such a quantity?

Let us model the particles in the gas as rigid spheres of diameter d . For each particle, we can construct a virtual cylinder whose axis is aligned with v , and whose cross section is $\sigma = \pi d^2$. Any particle whose center is within this cylinder must therefore collide with this particle, which is why σ is known as the collisional cross section.

After some time t each particle will have swept a volume σvt , and since the average number of particles in this volume is $n_c = \sigma vtn$, we define the collision time $t = \tau_c$ such that $n_c = 1$, that is:

$$\sigma v \tau_c n = 1 \implies \tau_c = \frac{1}{\sigma n v} \quad (7.5.1)$$

But which velocity to use? It is common notation to use the thermal velocity, thus defining:

$$\tau_c = \frac{1}{\sigma n} \sqrt{\frac{m}{2k_B T}} \quad (7.5.2)$$

It is then straightforward to calculate the mean free path, that is, the average distance travelled by a particle between collisions:

$$\lambda_{mfp} = v_{th} \tau_c = \frac{1}{\sigma n} \quad (7.5.3)$$

This is also the length of the collision cylinder such that at least one particle is included in it.

Heat transport phenomena

8.1 Particle distribution function

Let us define the **particle distribution function in the phase space** as the average number of particles with velocities in the velocity space volume $d^3\mathbf{v}$ and in the spatial volume $d^3\mathbf{r}$, denoted as:

$$F(t, \mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} \quad (8.1.1)$$

Normalization thus requires that for a system of N particles in total:

$$\iint d^3\mathbf{r} d^3\mathbf{v} F(t, \mathbf{r}, \mathbf{v}) = N \quad (8.1.2)$$

In a homogeneous system we obviously require that $F(\mathbf{r}, \mathbf{v}) = F(\mathbf{v}) = n f(\mathbf{v})$ where $n = n(\mathbf{r})$ is the density of particles, which has to be constant.

Indeed, if we integrate F over the velocity space:

$$\int d^3\mathbf{v} F(t, \mathbf{r}, \mathbf{v}) \equiv n(t, \mathbf{r}) = n \quad (8.1.3)$$

The **first moment** of F is defined as the mean momentum density:

$$\int d^3\mathbf{v} m\mathbf{v} F(t, \mathbf{r}, \mathbf{v}) = mn(t, \mathbf{r})\mathbf{u}(t, \mathbf{r}) \quad (8.1.4)$$

where $\mathbf{u}(t, \mathbf{r})$ is the mean velocity of the gas. Indeed F is the number of particles per unit volume with speed in the neighborhood of \mathbf{v} , so multiplying it by the momentum carried by such a particle, $m\mathbf{v}$, and integrating gives the average momentum density.

In a similar fashion, the **second moment** is defined as the mean energy density:

$$\int d^3\mathbf{v} \frac{mv^2}{2} F(t, \mathbf{r}, \mathbf{v}) = \int d^3\mathbf{w} \frac{m|\mathbf{u} + \mathbf{w}|^2}{2} F(t, \mathbf{r}, \mathbf{v}) \quad (8.1.5)$$

$$= \frac{mu^2}{2} \int d^3\mathbf{w} F + \cancel{mu \cdot \int d^3\mathbf{w} \mathbf{w} F} + \int d^3\mathbf{w} \frac{mw^2}{2} F \quad (8.1.6)$$

$$= \frac{mnu^2}{2} + \left\langle \frac{mw^2}{2} \right\rangle n \quad (8.1.7)$$

where \mathbf{u} is defined as in the first moment, and $\mathbf{w} = \mathbf{v} - \mathbf{u}$ is the peculiar velocity of each particle in the gas frame. Indeed:

$$\int d^3\mathbf{w}\mathbf{w}F = \int d^3\mathbf{v}(\mathbf{v} - \mathbf{u})F = \mathbf{u}n(t, \mathbf{r}) - n(t, \mathbf{r})\mathbf{u} = 0 \quad (8.1.8)$$

Now the term $\frac{mnu^2}{2}$ may be interpreted as the kinetic energy density due to the gas' mean motion \mathbf{u} . Instead, the second term $\left\langle \frac{mw^2}{2} \right\rangle n$ is the internal energy density due to the gas' internal degrees of freedom. We may then define:

$$K = \int d^3\mathbf{r} \frac{mnu^2}{2} \quad (8.1.9)$$

as the total kinetic energy of the gas, and:

$$U = \int d^3\mathbf{r} \left\langle \frac{mw^2}{2} \right\rangle n \quad (8.1.10)$$

as the total internal energy.

All this is nice and dandy, but we are assuming that we are given an expression for $F(t, \mathbf{r}, \mathbf{v})$.

8.2 Maxwellian equilibrium

In most classical scenarios the dependence of F on position and time is associated with macroscopic perturbations, such as an oven or an open window. Therefore, the scales of these inhomogeneities are much larger than τ_c and τ_{mfp} .

If we then break up the gas into small elements of size δl in time intervals δt such that:

$$l \gg \delta l \gg \lambda_{mfp}, \quad t \gg \delta t \gg \tau_c \quad (8.2.1)$$

Each one of these small gas elements thus behaves as a homogeneous gas, and may be modelled with the local Maxwellian distribution:

$$F_M(t, \mathbf{r}, \mathbf{v}) = n(t, \mathbf{r}) \left(\frac{m}{2\pi k_B T(t, \mathbf{r})} \right)^3 \exp \left[- \frac{m|\mathbf{w}(t, \mathbf{r})|^2}{2k_B T(t, \mathbf{r})} \right] \quad (8.2.2)$$

where $\mathbf{w}(t, \mathbf{r}) = \mathbf{v} - \mathbf{u}(t, \mathbf{r})$ is the constituent particle's particular velocity in the gas element's center of mass frame.

As before, if we define $v_{th} = \sqrt{\frac{2k_B T}{m}}$, then the maxwellian distribution reduces to:

$$F_M(t, \mathbf{r}, \mathbf{v}) = \frac{n(t, \mathbf{r})}{(\sqrt{\pi}v_{th}(t, \mathbf{r}))^3} e^{-w^2/v_{th}^2} \quad (8.2.3)$$

Local pressure is then

$$P(t, \mathbf{r}) = n(t, \mathbf{r})k_B T(t, \mathbf{r}) = \frac{2}{3}\varepsilon(t, \mathbf{r}) \quad (8.2.4)$$

so that:

$$k_B T(t, \mathbf{r}) = \frac{2}{3} \frac{\epsilon(t, \mathbf{r})}{n(t, \mathbf{r})} = \frac{2}{3} \left\langle \frac{mw^2}{2} \right\rangle \quad (8.2.5)$$

8.3 Conservation laws

Gases can transport momentum, energy of particles through their motion. The properties of momentum, energy and particle transport are called transport properties of gases, and are closely related to conservation laws.

One quantity that must always be conserved is the total number of particles:

$$\int d^3 \mathbf{r} n = N \quad (8.3.1)$$

We must also have conservation of total momentum, which we can set to 0 by moving to an appropriate reference frame:

$$\int d^3 \mathbf{r} m n \mathbf{u} = 0 \quad (8.3.2)$$

Finally, the total energy must also be conserved:

$$\int d^3 \mathbf{r} \frac{m n u^2}{2} + \int d^3 \mathbf{r} \left\langle \frac{m w^2}{2} \right\rangle n = K + U \quad (8.3.3)$$

Now how do we evolve the variables n, \mathbf{u}, T in such a way as to satisfy these conservation laws?

Let us initially consider a gas with $\mathbf{u} = 0$ and constant density. All the energy is clearly in the internal energy:

$$\epsilon = n c_1 T(t, \mathbf{r}) \quad (8.3.4)$$

where c_1 is the heat capacity per particle. Furthermore we also assume that the system is one dimensional with z -dependence.

If we have a hot region and a cold region, then there will be a flux of heat $J(z)$, defined as the internal energy flowing per unit time through unit area perpendicular to z -axis.

So for a small volume $A \times \left[z - \frac{dz}{2}, z + \frac{dz}{2} \right]$ then:

$$\frac{\partial}{\partial t} (n c_1 T \cdot A dz) = A J_z \left(z - \frac{dz}{2} \right) - A J_z \left(z + \frac{dz}{2} \right) \quad (8.3.5)$$

which yields:

$$n c_1 \frac{\partial T}{\partial t} = - \frac{J_z(z + dz/2) - J_z(z - dz/2)}{dz} = - \frac{\partial J_z}{\partial z} \quad (8.3.6)$$

In three dimensions this turns into:

$$\frac{\partial}{\partial t} \int_V d^3 \mathbf{r} n c_1 T = - \int_{\partial V} d\mathbf{A} \cdot \mathbf{J} = - \int_V d^3 \mathbf{r} \nabla \cdot \mathbf{J} \quad (8.3.7)$$

for all volumes V , so that:

$$\boxed{nc_1 \frac{\partial T}{\partial t} + \nabla \cdot \mathbf{J} = 0} \quad (8.3.8)$$

The total flux of heat out of a system is equal to the rate of change of its internal energy, and thus is proportional to the rate of change of its temperature. This is just a restatement of conservation of energy, heat leaving the system must be equal to the energy being lost.

Similarly, the rate of change of a volume V of a gas is:

$$\frac{\partial}{\partial t} \int_V d^3r mn\mathbf{u} \quad (8.3.9)$$

We now define the momentum flux as Π as a tensor with components Π_{ij} as the momentum component j leaving through a surface of unit area perpendicular to i per unit time.

Analogously to energy conservation, momentum conservation then requires:

$$\frac{\partial}{\partial t} \int_V d^3r mn\mathbf{u} = - \int_{\partial V} d\mathbf{A} \cdot \mathbf{\Pi} \quad (8.3.10)$$

We may not apply Gauss' law directly since we have a tensor, instead we may take the j -th component of momentum:

$$\frac{\partial}{\partial t} \int_V d^3r mn u_j = - \int_{\partial V} d\mathbf{A} \cdot \mathbf{\Pi}_j = - \int_V d^3r \nabla \cdot \mathbf{\Pi}_j \quad (8.3.11)$$

Consequently:

$$mn \frac{\partial u_j}{\partial t} + \nabla \cdot \mathbf{\Pi}_j = 0 \quad (8.3.12)$$

But how do we find the momentum flux? Momentum fluxes are usually due to velocity gradients, and in gases one significant contribution is due to viscosity.

Viscosity is a fluid's resistance to deformation through shear stress. That is, if we place a fluid between two straight, parallel plates, and move the top plate, then viscosity will cause layers of gas closer to the top plate to move with greater speed compared to layers farther from the top plate.

Suppose the plates lie in the xy plane and the top one is moved in the x direction, causing shear stress $\tau_{xz} = \frac{F}{A}$. Here A is the area of the plate and F the force applied. Suppose the plate moves at speed u , then this will produce a velocity gradient $\frac{d\langle u_x \rangle}{dz}$ in the gas, because each layer in the gas will try to drag the neighbouring layers with it. Viscosity η is then defined as:

$$\tau_{xz} = \frac{F}{A} = \eta \frac{d\langle u_x \rangle}{dz} \quad (8.3.13)$$

Momentum conservation yields:

$$mn \frac{\partial u_x}{\partial t} + \eta \frac{\partial^2 \langle u_x \rangle}{\partial z^2} \quad (8.3.14)$$

Part III

Statistical Physics

The microcanonical ensemble

9.1 What is statistical physics?

The aim of statistical physics is to derive the laws of physics at the macro-scale, using the governing laws of molecular dynamics. It provides a link between the behaviour of the micro-scale and the resulting effects in bulk properties at the macro-scale we are familiar with.

For example, what makes the specific heat of materials different from each other?

To do so we will need to find an intermediary connection between classical mechanics and thermodynamics, entropy and the phase space.

9.2 The classical Phase Space

Recall that in classical mechanics the coordinates $(\mathbf{q}(t), \mathbf{p}(t))$ are enough to fully specify the state of a system of N particles. These coordinates belong to an alternate space, called the **phase space**. For a three-dimensional “real space” with N particles, it follows that the phase space will be $6N$ -dimensional.

Any point in the phase space is associated to a micro-state of the whole system’s behaviour, that is, an instant in the motion of the system. The way this micro-state behaves through time is given by Hamilton’s equations:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (9.2.1a)$$

The solutions to the above equations are curves $(\mathbf{q}(t), \mathbf{p}(t))$ in the phase space. These curves are known as phase-space trajectories.

For an observable $A(\mathbf{q}(t), \mathbf{p}(t), t)$ its time-evolution is given by:

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial A}{\partial q_i} \dot{q}_i + \frac{\partial A}{\partial p_i} \dot{p}_i \right) \quad (9.2.2)$$

It is now convenient to introduce the Poisson bracket $\{A, H\} = \sum_{i=1}^{3N} \left(\frac{\partial A}{\partial q_i} \dot{q}_i + \frac{\partial A}{\partial p_i} \dot{p}_i \right)$ to simplify

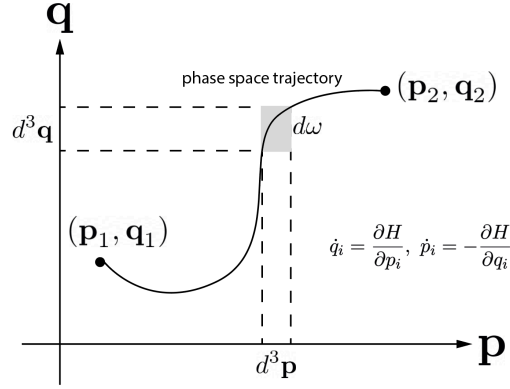


Figure 9.1. The phase space

the sum (9.2.2) to write:

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} + \{A, H\} \quad (9.2.3)$$

Just like normal three dimensional space may be divided into infinitesimal volume elements, one may also partition the phase space into infinitesimal phase space volume elements $d\omega$ defined by:

$$d\omega = d^{3N} \mathbf{q} d^{3N} \mathbf{p} \quad (9.2.4)$$

Now if we consider two energy surfaces in the phase-space, with energies E and $E + \Delta E$, then the corresponding phase volume will be given by:

$$\Delta\omega = \omega(E + \Delta E) - \omega(E) = \left. \frac{\partial\omega}{\partial E} \right|_{V,N} \Delta E \quad (9.2.5)$$

However, we also have that if the area of the energy surface $H(\mathbf{q}, \mathbf{p}) = E$ is defined as $\sigma(E)$, the density of states at energy E , then:

$$\sigma(E) = \int_{H=E} d\sigma \quad (9.2.6)$$

so that:

$$d\omega = \sigma(E)dE \implies \boxed{\sigma(E) = \left. \frac{\partial\omega}{\partial E} \right|_{V,N}} \quad (9.2.7)$$

We could also see this using the fundamental theorem of calculus:

$$\omega(E, V, N) = \int_{E' < E} \sigma(E', V, N) dE' \iff \sigma(E, V, N) = \left. \frac{\partial\omega}{\partial E} \right|_{V,N} \quad (9.2.8)$$

from which it follows that:

$$\int_{E < H < E + \Delta E} d\omega = \sigma(E) \Delta E \quad (9.2.9)$$

Now for a closed, isolated system we have that the variables E (energy), N (number of particles) and V (volume) are enough to specify its macroscopic properties. They specify a **macrostate**.

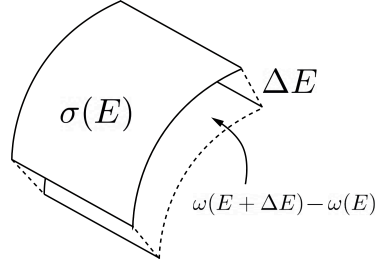


Figure 9.2. Phase space volume between energy surfaces $H(q_i, p_j) = E$ and $H(q_i, p_j) = E + \Delta E$.

However, this macrostate is highly degenerate. There are several ways one could arrange a system to give the corresponding values of E, V, N , which we denote by $\Omega(E, V, N)$. Each one of these states is called a **microstate**. We expect that $\Omega(E, V, N)$ will be proportional to $\sigma(E, V, N)$

$$\Omega(E, V, N) = \frac{\sigma(E, V, N)}{\sigma_0} = \frac{1}{\sigma_0} \left. \frac{\partial \omega}{\partial E} \right|_{V, N} \quad (9.2.10)$$

It turns out these microstates all have equal probabilities. Indeed, consider a isolated system which we subdivide into two subsystems with macroscopic variables E_i, V_i, N_i . Because the original system is closed, we require that:

$$dE = 0 \implies dE_1 = -dE_2 \quad (9.2.11)$$

$$dV = 0 \implies dV_1 = -dV_2 \quad (9.2.12)$$

$$dN = 0 \implies dN_1 = -dN_2 \quad (9.2.13)$$

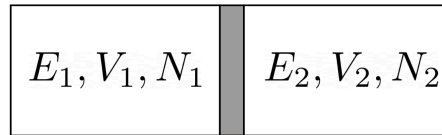


Figure 9.3. Composite system state counting

We must also have from elementary combinatorics that:

$$\Omega(E, V, N) = \Omega_1(E_1, V_1, N_1)\Omega_2(E_2, V_2, N_2) \quad (9.2.14)$$

$$\implies d\Omega = \Omega_1 d\Omega_2 + \Omega_2 d\Omega_1 \quad (9.2.15)$$

$$\implies d \ln \Omega = d \ln \Omega_2 + d \ln \Omega_1 \quad (9.2.16)$$

Now equilibrium occurs whenever we have a maximum number of possible microstates, so when $d\Omega = 0$. Therefore $d \ln \Omega = 0$. In classical thermodynamics, one may recall that equilibrium results in maximal entropy so $dS = dS_1 + dS_2 = 0$. Comparison with (9.2.16)

suggests the relation:

$$S(E, V, N) = k_B \ln \Omega(E, V, N) \quad (9.2.17)$$

where k_B is Boltzmann's constant. This is a remarkable important law because it allows us to find all other thermodynamic variables which we define as:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V, N}, \quad \frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E, N} \quad (9.2.18)$$

9.3 Probability density and expectation values

For isolated systems we have assumed that all microstates on the same energy surface have equal probability of being occupied. This is no longer true for open systems, where we must introduce a probability density $\rho(\mathbf{q}, \mathbf{p})$ which corresponds to the probability density of the system to occupy the phase state (\mathbf{q}, \mathbf{p}) . Normalization is required as always:

$$\int \rho(\mathbf{q}, \mathbf{p}) d^{3N} \mathbf{q} d^{3N} \mathbf{p} = 1 \quad (9.3.1)$$

Suppose we have some observable A of the system, then its expectation value is given by:

$$\langle f \rangle = \int \rho(\mathbf{q}, \mathbf{p}) f(\mathbf{q}, \mathbf{p}) d^{3N} \mathbf{q} d^{3N} \mathbf{p} \quad (9.3.2)$$

For an isolated system for example, we have that:

$$\rho(\mathbf{q}, \mathbf{p}) = \frac{1}{\sigma(E)} \delta(E - H(\mathbf{q}, \mathbf{p})) \quad (9.3.3)$$

This however is a very complicated expression due to the Dirac delta function. We can simplify it by allowing for a small energy deviation ΔE , for which:

$$\rho(\mathbf{q}, \mathbf{p}) = \begin{cases} \rho_0, & E \leq H(\mathbf{q}, \mathbf{p}) \leq E + \Delta E \\ 0, & \text{otherwise} \end{cases} \quad (9.3.4)$$

Then normalization requires:

$$\int_{E \leq H(\mathbf{q}, \mathbf{p}) \leq E + \Delta E} \rho_0 d^{3N} \mathbf{q} d^{3N} \mathbf{p} = 1 \quad (9.3.5)$$

$$\iff \rho_0 \sigma(E, V, N) \Delta E = 1 \quad (9.3.6)$$

$$\iff \rho_0 \Omega(E, V, N) h^{3N} \Delta E = 1 \quad (9.3.7)$$

$$\iff \rho = \frac{1}{\Omega(E, V, N) h^{3N} \Delta E} \quad (9.3.8)$$

since the energy shell volume in the phase space is $\sigma(E, V, N)\Delta E = h^{3N}\Omega(E, V, N)\Delta E$. We may take $\frac{1}{h^{3N}}$ out of the integral by defining:

$$\frac{1}{h^{3N}} \int \rho(\mathbf{q}, \mathbf{p}) d^{3N} \mathbf{q} d^{3N} \mathbf{p} = 1 \quad (9.3.9)$$

An ensemble is then a very very large collection of identical copies of a system with distribution function $\rho(\mathbf{q}, \mathbf{p})$. There are special types of ensembles which we will study:

- (i) micro-canonical ensemble
- (ii) canonical ensemble
- (iii) grand canonical ensemble

In the case of the micro-canonical ensemble it is defined by:

$$\rho_{mc}(\mathbf{q}, \mathbf{p}) = \begin{cases} \frac{1}{\Omega\Delta E}, & E \leq H(\mathbf{q}, \mathbf{p}) \leq E + \Delta E \\ 0, & \text{otherwise} \end{cases} \quad (9.3.10)$$

9.4 Louiville's theorem

Hamilton's equations demand that:

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \{\rho, H\} \quad (9.4.1)$$

Now consider the set of systems in the volume element ω at time t . As they evolve through time each of the phase space points will get mapped to another volume element ω' at time t' . However the initial and final number of phase space points must remain constant. We thus require that if we have a change in phase space volumes then a net amount of phase space points flowing out of ω :

$$\frac{\partial}{\partial t} \int d\omega\rho + \int \rho(\mathbf{v} \cdot \mathbf{n}) \cdot d\sigma \quad (9.4.2)$$

Here \mathbf{v} is the "velocity" in the phase space of the trajectories, given by $(\dot{\mathbf{q}}, \dot{\mathbf{p}})$. Using the divergence theorem we find:

$$\int_{\omega} \left(\frac{\partial\rho}{\partial t} + \nabla(\rho\mathbf{v}) \right) d\omega = 0 \quad (9.4.3)$$

where divergence is taken with respect to the phase space coordinates $\{q_1, q_2, \dots, q_{3N}, p_1, p_2, \dots, p_{3N}\}$. Since (9.4.3) holds for any arbitrary volume ω we find the continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla(\rho\mathbf{v}) = 0 \quad (9.4.4)$$

Let us expand the divergence:

$$\nabla(\rho\mathbf{v}) = \sum_{i=1}^{3N} \left[\frac{\partial}{\partial q_i}(\rho\dot{q}_i) + \frac{\partial}{\partial p_i}(\rho\dot{p}_i) \right] \quad (9.4.5)$$

$$= \sum_{i=1}^{3N} \left[\frac{\partial\rho}{\partial q_i}\dot{q}_i + \rho\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\rho}{\partial p_i}\dot{p}_i + \rho\frac{\partial\dot{p}_i}{\partial p_i} \right] \quad (9.4.6)$$

$$= \sum_{i=1}^{3N} \left[\frac{\partial\rho}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial\rho}{\partial p_i} \frac{\partial H}{\partial q_i} + \rho \left(\frac{\partial^2 H}{\partial q_i \partial p_i} + \frac{\partial^2 H}{\partial p_i \partial q_i} \right) \right] \quad (9.4.7)$$

$$= \{\rho, H\} \quad (9.4.8)$$

Consequently:

$$\boxed{\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \{\rho, H\} = \frac{\partial\rho}{\partial t} + \nabla(\rho\mathbf{v}) = 0} \quad (9.4.9)$$

so the time derivative of the phase-space density is null for a trajectory. Alternatively, a phase space density is constant along the trajectory of a system. This is known as **Louiville's theorem**.

9.5 Ergodic hypothesis and Equal a priori probability

One may wonder why we should even worry about probability distributions in the phase space and taking ensemble averages. After all, when we perform experiments we are really taking a time average of one system, rather than the ensemble average of several systems. If we have some quantity $A(\mathbf{q}, \mathbf{p}, t)$ then its time average will be:

$$\tilde{A} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(\mathbf{q}, \mathbf{p}, t) dt \quad (9.5.1)$$

How can we justify using the ensemble average instead:

$$\langle A \rangle = \int_{\text{ensemble}} A(\mathbf{q}, \mathbf{p}, t) \rho(\mathbf{q}, \mathbf{p}) d\omega \quad (9.5.2)$$

The ergodic hypothesis postulates that these two averages should be equivalent in the thermodynamic limit. This hinges on the condition that the system visits the neighbourhood of all the available phase space points (gets arbitrarily close to them) after a sufficiently long time. If we average over time then under the ergodic hypothesis we are really just averaging over the ensemble.

Also, another fundamental assumption of statistical mechanics is the postulate of equal a priori probability. It states that for a system in equilibrium the micro-states corresponding to a macro-state have the same probability of being occupied.

The link between these two postulates is Louiville's theorem, which states that the a system takes trajectories along constant phase space density surfaces. Since systems in equilibrium have a constant phase space density for each macro-state, it follows that a system's trajectory will be in the neighborhood of all possible microstates.

9.6 The microcanonical ensemble

The microcanonical ensemble is a statistical ensemble used to model isolated systems, where there is no exchange of particles and energy. Therefore we will take E, V, N to be conserved quantities specifying an equilibrium macrostate. For such systems we previously stated without proof that for a given energy they are modelled by a constant phase-space density on the corresponding energy hypersurface, as described in (9.3.10)

Let us prove this result.

Consider an ensemble made up of \mathcal{N} identical copies of an isolated system, each in a different microstate on the energy hypersurface $H(\mathbf{q}, \mathbf{p}) = E$.

Suppose we partition this surface into small elements $\Delta\sigma_i$ with n_i systems each:

$$\mathcal{N} = \sum_i n_i \quad (9.6.1)$$

How many ways are there to distribute the \mathcal{N} systems into the energy surfaces? For \mathcal{N} systems there are $\mathcal{N}!$ different ways to enumerate them. However, since within each surface $\Delta\sigma_i$ there are $n_i!$ internal exchanges which do not alter the overall distribution, the total number of ways to distribute the systems will be:

$$W\{n_i\} = \frac{\mathcal{N}!}{\prod_i n_i!} \quad (9.6.2)$$

In other words, order matters when distributing the n_i systems into the surfaces $\Delta\sigma_i$, but the ordering of each system within $\Delta\sigma_i$ is insignificant.

The most probable distribution will of course maximize $W\{n_i\}$, and hence $\ln W\{n_i\}$ (so entropy will be maximized, corresponding to equilibrium). Therefore we need:

$$\ln W_{tot} = \ln \mathcal{N}! - \sum_i \ln n_i! \quad (9.6.3)$$

$$\approx \mathcal{N} \ln \mathcal{N} - \mathcal{N} - \sum_i (n_i \ln n_i - n_i) \quad (9.6.4)$$

$$\iff d \ln W_{tot} = - \sum_i \ln n_i \, dn_i = 0 \quad (9.6.5)$$

Since we are subject to the constraint (9.6.1) we need to use the method of Lagrange multipliers. We get that:

$$\lambda d\mathcal{N} = \lambda \sum_i dn_i = 0 \quad (9.6.6)$$

which we can sum to (9.6.5) to get:

$$\sum_i (\ln n_i - \lambda) dn_i = 0 \quad (9.6.7)$$

Equating the coefficients to zero one finds that:

$$\boxed{n_i = \text{cnst.}} \quad (9.6.8)$$

We easily identify this distribution as a homogeneous distribution:

$$\rho_{mc}(\mathbf{q}, \mathbf{p}) = \begin{cases} \frac{1}{\Omega\Delta E} & E \leq H(\mathbf{q}, \mathbf{p}) \leq E + \Delta E \\ 0 & \text{otherwise} \end{cases} \quad (9.6.9)$$

as desired. Hence the most likely distribution in a microcanonical ensemble is the ρ_{mc} distribution.

9.7 Relating the Boltzmann and Gibbs entropy

Recall that the Boltzmann entropy was identified to be:

$$S(E, V, N) = k_B \ln \Omega(E, V, N) \quad (9.7.1)$$

In the microcanonical picture however (developed by Gibbs), it is seen as the ensemble average:

$$S(E, V, N) = \frac{1}{h^{3N}} \int_{E \leq H \leq E + \Delta E} \rho(\mathbf{q}, \mathbf{p}) (-k_B \ln \rho(\mathbf{q}, \mathbf{p})) d\omega \quad (9.7.2)$$

which we recognize to be:

$$S(E, V, N) = \langle -k_B \ln \rho(\mathbf{q}, \mathbf{p}) \rangle \quad (9.7.3)$$

Indeed, inserting (9.3.10) into (9.7.2) one finds that:

$$S(E, V, N) = \frac{1}{h^{3N}} \int_{E \leq H \leq E + \Delta E} \frac{1}{\Omega\Delta E} k_B \ln(\Omega\Delta E) d\omega \quad (9.7.4)$$

Since Ω is constant over the energy shell $E \leq H \leq E + \Delta E$ the integrand may be pulled out:

$$S(E, V, N) = \frac{1}{\Omega\Delta E} k_B \ln(\Omega\Delta E) \frac{1}{h^{3N}} \int_{E \leq H \leq E + \Delta E} d\omega \quad (9.7.5)$$

We now recognize:

$$\Omega\Delta E = \frac{1}{h^{3N}} \int_{E \leq H \leq E + \Delta E} d\omega \quad (9.7.6)$$

so that in the microcanonical ensemble:

$$S(E, V, N) = k_B \ln(\Omega\Delta E) = \langle -k \ln \rho \rangle \quad (9.7.7)$$

as desired. Alternatively, in the discrete case, we can insert $p_i = \frac{n_i}{N}$ which is justified by the equal a priori probability postulate into (9.6.4) and find that

$$S = -k_B \left(N \ln N - N \sum_i p_i \ln(N p_i) \right) = \sum_i p_i \ln p_i \quad (9.7.8)$$

as desired. We therefore infer that the Gibbs entropy is a generalization of the Boltzmann entropy, which is only applicable for equilibrium systems in a single macrostate, where the equal a priori probability postulate is justified.

9.8 Entropy of an ideal gas (classical)

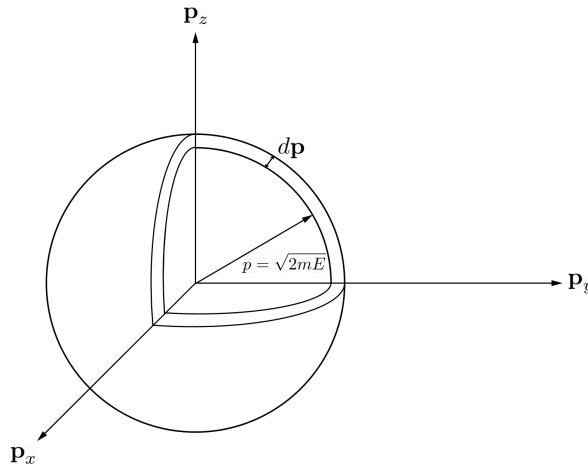
For an ideal gas which is interaction-free, the Hamiltonian takes the form:

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} \quad (9.8.1)$$

The total phase space volume enclosed by the energy surface $E = H(\mathbf{q}, \mathbf{p})$ is

$$\omega(E, V, N) = \int_{H \leq E} d^{3N} \mathbf{q} d^{3N} \mathbf{p} = V^N \int_{H \leq E} d^{3N} \mathbf{p} \quad (9.8.2)$$

since there is no q dependence of the Hamiltonian surface we are integrating over. In the phase-space, we also see that E delineates a spherical surface of radius $p = \sqrt{2mE}$.



So the integral in (9.8.2) is simply the volume of an $3N$ -dimensional sphere:

$$V_{\text{sphere}}^N = \frac{2\pi^{N/2}}{N\Gamma(\frac{N}{2})} R^N \quad (9.8.3)$$

where R is the radius. We then find that:

$$\omega(E, V, N) = V^N \frac{2\pi^{3N/2}}{3N\Gamma(\frac{3N}{2})} (2mE)^{3N/2} \quad (9.8.4)$$

Using (9.2.10) one may then find the number of microstates on the energy surface:

$$\Omega(E, V, N) = \frac{1}{\sigma_0} \frac{\partial}{\partial E} \left(\omega(E, V, N) \right) \quad (9.8.5)$$

$$= \frac{1}{\sigma_0} V^N \frac{\pi^{3N/2}}{\Gamma(\frac{3N}{2})} (2m)^{3N/2} E^{3N/2-1} \quad (9.8.6)$$

Consequently the volume occupied by the micro-canonical ensemble is:

$$\Omega(E, V, N) \Delta E = \frac{1}{\sigma_0} V^N \frac{\pi^{3N/2}}{\Gamma(\frac{3N}{2})} (2mE)^{3N/2} \frac{\Delta E}{E} \quad (9.8.7)$$

One then gets an expression, albeit very complicated, for an ideal gas:

$$S(E, V, N) = k_B \ln \left[\frac{1}{\sigma_0} V^N \frac{\pi^{3N/2}}{\Gamma(\frac{3N}{2})} (2mE)^{3N/2} \frac{\Delta E}{E} \right] \quad (9.8.8)$$

For very large systems where $N \gg 1$ then we may make the approximations $\ln \frac{\Delta E}{E} \approx 0$ (thus ignoring any terms not of order N). We also make use of Stirling's approximation:

$$\ln \Gamma(n) \approx n \ln n - n \quad (9.8.9)$$

to find that:

$$S(E, V, N) \approx Nk_B \left[\ln \left(\frac{V}{\sigma} (2\pi mV)^{3/2} \right) + \frac{3}{2} - \frac{3}{2} \ln \frac{3N}{2} \right] \quad (9.8.10)$$

where $\sigma = \sigma_0^{1/N}$. Further simplification gives:

$$S(E, V, N) \approx Nk_B \left[\frac{3}{2} + \ln \frac{V}{\sigma} \left(\frac{4\pi mE}{3N} \right)^{3/2} \right] \quad (9.8.11)$$

This gives the correct equations of state from classical thermodynamics:

$$\frac{1}{T} = \frac{\partial S}{\partial E} \Big|_{V,N} = \frac{3Nk_B}{2E} \implies E = \frac{3}{2} Nk_B T \quad (9.8.12)$$

$$\frac{p}{T} = \frac{\partial S}{\partial V} \Big|_{E,N} = \frac{Nk_B}{V} \implies pV = Nk_B T \quad (9.8.13)$$

which recast (9.8.11) into:

$$S(E, V, N) \approx Nk_B \left[\frac{3}{2} + \ln \left(\frac{V}{\sigma} (2\pi mk_B T)^{3/2} \right) \right] \quad (9.8.14)$$

However, it turns out that (9.8.11) is incorrect. Indeed, we know that entropy is an extensive quantity, yet the additional logarithmic term produces an extra $\ln \alpha$ term when the system is scaled up by α .

9.9 Gibb's paradox

Let's analyze where this inconsistency comes from. We consider a container containing gas A , and another container containing gas B . Now the total entropy of the two-container system is:

$$S^b = S_A^b(T, V_A, N_A) + S_B^b(T, V_B, N_B) \quad (9.9.1)$$

Now suppose we put the containers in contact so that the gases mix together and occupy



Figure 9.4. Caption

the total volume $V_A + V_B$. Once equilibrium has reached the entropy will be:

$$S^a = S_A^a(T, V_A + V_B, N_A) + S_B^a(T, V_A + V_B, N_B) \quad (9.9.2)$$

The total entropy change is

$$\Delta S = N_A k_B \ln \frac{V_A + V_B}{V_A} + N_B k_B \ln \frac{V_A + V_B}{V_B} > 0 \quad (9.9.3)$$

Throughout this derivation however, we have been assuming that gases A and B are distinguishable. If we use identical gases, then

$$S^b = S_A^b(T, V_A, N_A) + S_B^b(T, V_B, N_B) \quad (9.9.4)$$

$$S^a = S(T, V_A + V_B, N_A + N_B) \quad (9.9.5)$$

We have that:

$$S(T, V_A + V_B, N_A + N_B) = (N_A + N_B) k_B \left[\frac{3}{2} + \ln \frac{V_A + V_B}{\sigma} (2\pi m k_B T)^{3/2} \right] \quad (9.9.6)$$

$$= N_A k_B \left[\frac{3}{2} + \ln \frac{V_A}{\sigma} (2\pi m k_B T)^{3/2} \right] + N_B k_B \left[\frac{3}{2} + \ln \frac{V_B}{\sigma} (3\pi m k_B T)^{3/2} \right] \quad (9.9.7)$$

$$= S^a = S_A^a(T, V_A + V_B, N_A) + S_B^a(T, V_A + V_B, N_B) \quad (9.9.8)$$

so we again find that:

$$\Delta S = N_A k_B \ln \frac{V_A + V_B}{V_A} + N_B k_B \ln \frac{V_A + V_B}{V_B} > 0 \quad (9.9.9)$$

This is quite a problem however, because the mixing of two identical gases is completely reversible, there is no change in the gases when they are put in contact. Therefore, we should have found $\Delta S = 0$.

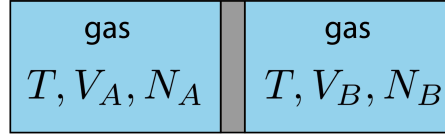


Figure 9.5. Caption

Note that this discrepancy is directly related to the non-extensivity of our (9.8.14), since when we put the two distinguishable gas containers in contact we're essentially scaling the size of the system by two, whereas for identical gas containers we're doing absolutely nothing, yet our expressions for the entropy after mixing are the same for both situations.

To solve this problem, we must employ tools from quantum mechanics. Indeed we know that in quantum mechanics one cannot label particles. Hence different microstates with the same particle distribution but different particle labelling in the classical picture are considered equivalent in quantum mechanics. Since there are $N!$ ways to enumerate N particles it follows that:

$$\Omega(E, V, N) = \frac{1}{N!} \frac{\sigma(E, V, N)}{\sigma_0} \quad (9.9.10)$$

which provides the following correction to entropy:

$$S(E, V, N) = Nk_B \left[\frac{3}{2} + \ln \frac{V}{\sigma} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right] - k_B \ln N! \quad (9.9.11)$$

and using Stirling's formula:

$$S(E, V, N) = Nk_B \left[\frac{5}{2} + \ln \frac{V}{N\sigma} \left(\frac{4\pi m E}{3N} \right)^{3/2} \right] \quad (9.9.12)$$

or alternatively:

$$S(E, V, N) = Nk_B \left[\frac{5}{2} + \ln \frac{V}{N\sigma} (2\pi m k_B T)^{3/2} \right] \quad (9.9.13)$$

We see by inspection that this is indeed an extensive quantity, since the logarithm's argument is intensive overall.

Let's check that this correction does indeed solve Gibbs' paradox. We now get that

$$\Delta S = (N_A + N_B)k_B \left[\frac{5}{2} + \ln \frac{V_A + V_B}{N_A + N_B\sigma} (2\pi m k_B T)^{3/2} \right] \quad (9.9.14)$$

$$- N_A k_B \left[\frac{5}{2} + \ln \frac{V_A}{N_A\sigma} (2\pi m k_B T)^{3/2} \right] - N_B k_B \left[\frac{5}{2} + \ln \frac{V_B}{N_B\sigma} (2\pi m k_B T)^{3/2} \right] \quad (9.9.15)$$

Now we may impose:

$$\frac{V_A}{N_A} = \frac{V_B}{N_B} = \frac{V_A + V_B}{N_A + N_B} \quad (9.9.16)$$

since the density of gases will remain the same. Hence

$$\Delta S = 0 \quad (9.9.17)$$

as desired. It is important to remember that:

The Gibbs' correction factor must not be included for distinguishable microstates.

9.10 Entropy of an ideal gas (quantum)

There is one constant in this entire derivation that may be frustrating the reader, and that is the proportionality constant σ_0 . We can determine its value by counting the number of states from a quantum mechanical point of view, and compare it with our semi-classical results.

Let us try to derive an expression for entropy for an ideal gas made up of distinguishable particles. We will model the container as an infinite quantum well of size L , within which the particles are non-interacting.

The individual particles are in the eigenstates:

$$\psi_{n_x, n_y, n_z} = A \sin\left(\frac{n_x \pi x}{L}\right) \sin\left(\frac{n_y \pi y}{L}\right) \sin\left(\frac{n_z \pi z}{L}\right) \quad (9.10.1)$$

with energy:

$$\epsilon_{n_x, n_y, n_z} = \frac{\hbar^2}{8mL^2} (n_x^2 + n_y^2 + n_z^2) \quad (9.10.2)$$

Therefore, our quantum-mechanical phase space will consist of (n_x, n_y, n_z) . Now the total

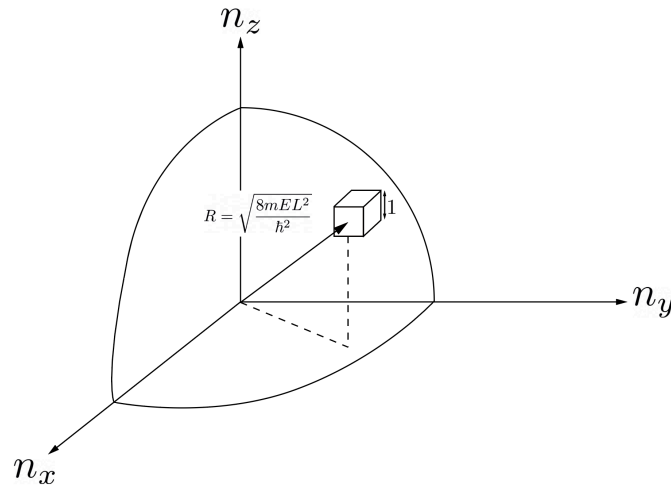


Figure 9.6. Quantum phase space (n_x, n_y, n_z) for an ideal gas

energy of an N -particle system will be given by:

$$E = \frac{h^2}{8mL^2} \sum_{i=1}^{3N} n_i^2 \quad (9.10.3)$$

We can use this to find $\Omega(E, V, N)$, where $V = L^3$, by determining the degeneracy of the energy eigenstate with energy E ? Clearly, we must determine how many ways one can add up a number n in terms of a sum of $3N$ quadratic integer terms $\sum_{i=1}^{3N} n_i^2$.

Unfortunately, this is an impossibly difficult problem to solve. Indeed, if an exact expression does exist, it would be a very irregular function of energy. Luckily however, the average behaviour of $\Omega(E, V, N)$ can be calculated, and we expect it to be increasing with energy. The larger the energy surface, the more likely it is for integer-coordinate points to lie on it.

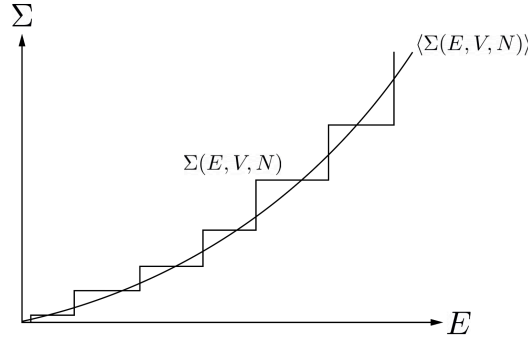


Figure 9.7. Behaviour of Σ and $\langle \Sigma \rangle$

Let us define the number of microstates within the energy sphere:

$$\Sigma(E, V, N) = \sum_{E' \leq E} \Omega(E', V, N) \quad (9.10.4)$$

This is again impossibly complicated to evaluate, but for large enough system, or systems with high enough temperature, then we can approximate it with its average $\langle \Sigma \rangle$. Then, we find that the density of states $g(E, V, N)$ at energy E is approximately:

$$g(E, V, N) = \frac{\partial}{\partial E} (\langle \Sigma(E, V, N) \rangle) \quad (9.10.5)$$

We remark the similarity between $g(E, V, N) \leftrightarrow \sigma(E, V, N)$ and $\langle \Sigma(E, V, N) \rangle \leftrightarrow \omega(E, V, N)$.

We also note that $\langle \Sigma(E, V, N) \rangle$ must be related to the volume enclosed by the energy surface of E . Indeed, suppose we divide this volume into individual unit cubes of unit volume.

Then energy grid point will have an associated unit cube, and the mean number of grid points will then be the total enclosed volume divided by the volume of a unit cube, which by definition is 1.

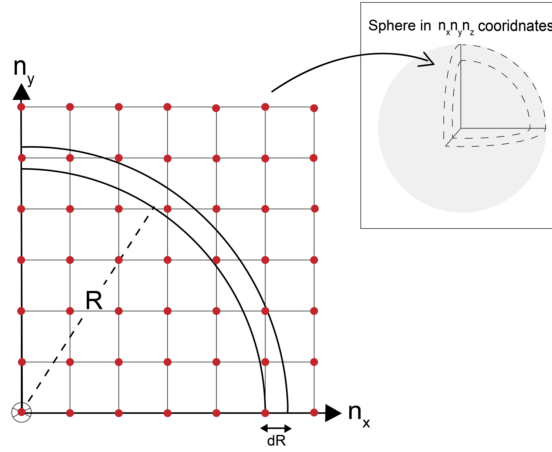


Figure 9.8. Partition of phase space into unit cubes

It is also important to note that since $n_i > 0$ the enclosed volume will be an octant. Therefore, defining:

$$R = \sqrt{\frac{8mEL^2}{h^2}} \quad (9.10.6)$$

we get that $\Sigma(E, V, N)$ is the volume of a $3N$ -hypersphere with this radius:

$$\langle \Sigma(E, V, N) \rangle = V_N(R) = \left(\frac{1}{2}\right)^{3N} \frac{\pi^{3N/2}}{\frac{3N}{2}\Gamma(\frac{3N}{2})} R^{3N} = \left(\frac{V}{h^3}\right)^N \frac{(2\pi mE)^{3N/2}}{\frac{3N}{2}\Gamma(\frac{3N}{2})} \quad (9.10.7)$$

Therefore the mean density of states at energy E is:

$$g(E, V, N) = \left(\frac{V}{h^3}\right)^N \frac{(2\pi m)^{3N/2}}{\frac{3N}{2}\Gamma(\frac{3N}{2})} E^{\frac{3N}{2}-1} \quad (9.10.8)$$

Finally, multiplying this by ΔE and approximating $\frac{3N}{2}N - 1 \approx \frac{3N}{2}$ gives

$$\langle \Omega(E) \rangle = g(E)\Delta E = \left(\frac{V}{h^3}\right)^N \frac{(2\pi m)^{3N/2}}{\frac{3N}{2}\Gamma(\frac{3N}{2})} E^{\frac{3N}{2}} \Delta E \quad (9.10.9)$$

$$= \frac{\pi^{3N/2}}{\Gamma(3N/2)} \frac{1}{2^{3N}} R^{\frac{3N}{2}} \Delta R \quad (9.10.10)$$

Taking the logarithm, we will get an additional term $\ln \Delta R$ which we can set to zero when compared to the $\frac{3N}{2} \ln R$ term. Hence, we get that $\Delta R = 1$, as should be since the subdivision of the phase space was into unit blocks of volume 1. Hence:

$$\langle \Omega(E) \rangle = \frac{\pi^{3N/2}}{\Gamma(3N/2)} \frac{1}{2^{3N}} R^{\frac{3N}{2}} = \left(\frac{V}{h^3}\right)^N \frac{(2\pi mE)^{3N/2}}{\Gamma(\frac{3N}{2})} \quad (9.10.11)$$

Comparison with the classical derivation shows that

$$\sigma = h^3 \iff \sigma_0 = h^{3N} \quad (9.10.12)$$

and hence

$$S(E, V, N) = Nk_B \left[\frac{3}{2} + \ln \frac{V}{h^3} (2\pi m k_B T)^{3/2} \right] \quad (9.10.13)$$

which is known as the **Sackur-Tetrode equation**.

9.11 General procedure using the microcanonical ensemble

This gives us a general procedure to derive classical thermodynamics using the microcanonical ensemble.

- (i) Find the Hamiltonian $H(\mathbf{q}, \mathbf{p})$ of the system.
- (ii) Find the phase space volume enclosed by the energy surface $E = H(\mathbf{q}, \mathbf{p})$.
- (iii) Evaluate the number of microstates on the energy surface $\tilde{\Omega}$:

$$\Omega(E, V, N) = \frac{1}{h^{3N}} \left. \frac{\partial \omega}{\partial E} \right|_{V, N} \quad (9.11.1)$$

- (iv) Evaluate the number of microstates $\tilde{\Omega}$ in the ensemble $\Omega(E, V, N)\Delta E$, which is the number of states in the energy shell $E \leq H(\mathbf{q}, \mathbf{p}) \leq E + \Delta E$:

$$\tilde{\Omega}(E, V, N) = \Omega \Delta E \quad (9.11.2)$$

- (v) Find the entropy:

$$S = k_B \ln \Omega \Delta E = k_B \ln \tilde{\Omega} \quad (9.11.3)$$

and ignoring any terms not order $o(N)$.

- (vi) Derive classical thermodynamic quantities and equations of state:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V, N} \quad (9.11.4)$$

$$\frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E, N} \quad (9.11.5)$$

$$C_V = \left. \frac{\partial E}{\partial T} \right|_V \quad (9.11.6)$$

9.12 Ultra-relativistic gas

Let us consider a gas made of N massless relativistic particles moving at the speed of light c , and thus satisfying the energy-momentum relation:

$$\epsilon = pc \quad (9.12.1)$$

The hamiltonian of this system then reads:

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^N \sqrt{p_{i,x}^2 c^2 + p_{i,y}^2 + p_{i,z}^2} c \quad (9.12.2)$$

Now we find that since the particles are indistinguishable, we must insert the Gibb's correction factor:

$$\Sigma(E, V, N) = \frac{1}{h^{3N}} \omega(E, V, N) = \frac{1}{h^{3N}} N! \int_{H \leq E} d\omega \quad (9.12.3)$$

and since H does not explicitly depend on \mathbf{q} we may integrate over spatial conjugate variables to find:

$$\Sigma(E, V, N) = \frac{V^N}{h^{3N} N!} \int_{H \leq E} d^{3N} \mathbf{p} \quad (9.12.4)$$

We need to find the volume of the energy surface in the momentum space. The square root really complicates things, so we shall instead average the relevant quantities:

$$\langle p^2 \rangle = 3 \langle p_x^2 \rangle = 3 \langle p_y^2 \rangle = 3 \langle p_z^2 \rangle \quad (9.12.5)$$

$$|\langle p^2 \rangle| = \sqrt{\frac{\langle p_x^2 \rangle}{3}} + \sqrt{\frac{\langle p_y^2 \rangle}{3}} + \sqrt{\frac{\langle p_z^2 \rangle}{3}} \quad (9.12.6)$$

$$= \frac{\sqrt{3}}{3} (|p_x| + |p_y| + |p_z|) \quad (9.12.7)$$

Then the volume enclosed by the energy surface $H = E$ is the locus of points in the phase space satisfying:

$$\sum_{i=1}^{3N} \frac{|p_i| c}{\sqrt{3}} \leq E \quad (9.12.8)$$

Hence, substituting $x_i = \frac{p_i c}{\sqrt{3}}$ and $dp_i = \frac{\sqrt{3} E}{c}$ then:

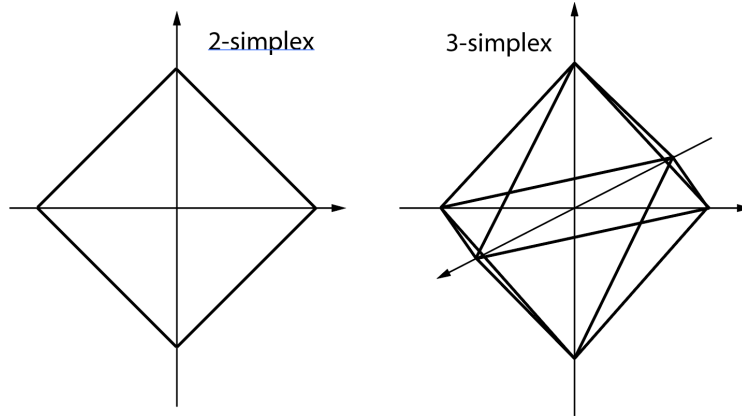
$$\Sigma(E, V, N) = \frac{V^N}{h^{3N} N!} \left(\frac{\sqrt{3} E}{c} \right)^{3N} \int_{\sum_i |x_i| \leq 1} d^{3N} \mathbf{x} \quad (9.12.9)$$

Clearly, if a set of points (x_1, \dots, x_{3N}) satisfies the condition $\sum_i |x_i| \leq 1$ then so will $(-x_1, \dots, -x_{3N})$. The volume over which we are integrating in (9.12.9) is therefore symmetric about the origin. We may therefore find the volume in one octant where $0 \leq x_i$ and then multiply the result by 2^{3N} . We therefore need to evaluate:

$$I_n = \int_{\sum_i x_i \leq 1} d^n x, \quad x_i \in [0, 1] \quad (9.12.10)$$

for $n = 3N$. The geometric region we are integrating over is known as a n -simplex, some are shown below.

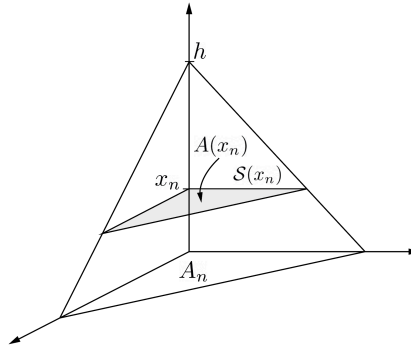
Consider the base of a n -simplex, which will itself be the corresponding $(n-1)$ -simplex in $n-1$ dimensions. Indeed setting $x_n = 0$ then we find $\sum_{i=1}^{n-1} x_i \leq 1$, defining the lower


Figure 9.9. 2-simplex and 3-simplex

dimensional simplex. The “volume” of this hypersurface is

$$V_n = \frac{1}{n} A_n \cdot h = \frac{A_n}{n} \quad (9.12.11)$$

where h is the height (which in our case is taken to be 1 when setting $x_1 = x_2 = \dots = x_{n-1} = 0$) and A_n is the volume of the base.


Figure 9.10. Evaluating volume of a simplex

To see why, let us consider sections of the simplex at constant x_n (height). Using similarity relations we must have that the sides $\mathcal{S}(x_n)$ are all scaled by $\frac{h-x_n}{h}$, so that:

$$A(x_n) = A_n \left(1 - \frac{x_n}{h}\right)^{n-1} \quad (9.12.12)$$

Hence:

$$V_n = \int_0^h A_n \left(1 - \frac{x_n}{h}\right)^{n-1} dx_n = \frac{h}{n} A_n \quad (9.12.13)$$

as desired.

However, note that V_n is really just I_n and $A_n = I_{n-1}$ so that:

$$I_n = \frac{I_{n-1}}{n} \implies I_n = \frac{1}{n!} \quad (9.12.14)$$

implying that:

$$\int_{\sum_i |x_i| \leq 1} d^{3N} \mathbf{x} = \frac{2^{3N}}{(3N)!} \quad (9.12.15)$$

The total number of states enclosed by the energy hypersurface is:

$$\Sigma(E, V, N) = \frac{V^N}{h^{3N} N!} \left(\frac{\sqrt{3}E}{c} \right)^{3N} \frac{2^{3N}}{(3N)!} \quad (9.12.16)$$

The density of states at energy E is:

$$g(E, V, N) = \frac{\partial \Sigma(E, V, N)}{\partial E} = \frac{V^N}{h^{3N} N!} \left(\frac{2\sqrt{3}}{c} \right)^{3N} \frac{E^{3N-1}}{(3N-1)!} \quad (9.12.17)$$

$$\approx \frac{V^N}{h^{3N} N!} \left(\frac{2\sqrt{3}}{c} \right)^{3N} \frac{E^{3N}}{(3N)!} \frac{1}{E} \quad (9.12.18)$$

Then the number of states occupied by the microcanonical ensemble is the volume of the shell $E < H < E + \Delta E$:

$$\Omega \Delta E = \frac{V^N}{N!} \left(\frac{2\sqrt{3}}{hc} \right)^{3N} \frac{E^{3N}}{(3N)!} \frac{\Delta E}{E} \quad (9.12.19)$$

and hence:

$$S(E, V, N) = k_B \ln(\Omega \Delta E) = k_B \ln \left[\frac{V^N}{N!} \left(\frac{2\sqrt{3}}{hc} \right)^{3N} \frac{E^{3N}}{(3N)!} \right] + k_B \ln \frac{\Delta E}{E} \quad (9.12.20)$$

$$= N k_B \ln \left[V \left(\frac{2\sqrt{3}}{hc} \right)^3 \right] - k_B \ln N! - k_B \ln(3N)! + k_B \ln \frac{\Delta E}{E} \quad (9.12.21)$$

Using Stirling's approximation:

$$S(E, V, N) = k_B \left(N \ln \left[V \left(\frac{2\sqrt{3}}{hc} \right)^3 \right] - N \ln N - 3N \ln(3N) + 4N + \ln \frac{\Delta E}{E} \right) \quad (9.12.22)$$

We may ignore all terms not of order N (i.e. $\ln \frac{\Delta E}{E}$), giving upon some simplification

$$\boxed{S(E, V, N) = N k_B \left(4 + \ln \left[\frac{V}{N} \left(\frac{2}{\sqrt{3} N h c} \right)^3 \right] \right)} \quad (9.12.23)$$

We may now derive the equations of state:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N} = \frac{3Nk_B}{E} \iff E = 3Nk_B T \quad (9.12.24)$$

$$\frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E,N} = \frac{Nk_B}{V} \iff pV = Nk_B T \quad (9.12.25)$$

$$C_V = \left. \frac{\partial E}{\partial T} \right|_V = 3Nk_B \quad (9.12.26)$$

and therefore find that:

$$p = \frac{1}{3} \frac{E}{V} \quad (9.12.27)$$

so the pressure is a third of the energy density. Suppose we want the probability distribution of $f(p_1)$. Then we need to integrate the microcanonical ensemble distribution ρ_{mc} over all variables except p_1 :

$$f(p_1) = \int \rho_{mc}(\mathbf{q}, \mathbf{p}) d^{3N} \mathbf{q} d^{3N-1} \mathbf{p} \quad (9.12.28)$$

$$= \int_{E < H < E + \Delta E} \frac{1}{\Omega \Delta E} d^{3N} \mathbf{q} d^{3N-1} \mathbf{p} \quad (9.12.29)$$

9.13 Harmonic oscillators

9.14 Two-state system

The canonical ensemble

10.1 Canonical phase-space density

We have encountered the microcanonical ensemble, which works perfectly for isolated systems, where the natural variables are E, V, N (energy, volume and particle number are conserved). Technically, one could also model an open system as isolated by including its surroundings, however this turns out to be impractical for most cases due to the complex mathematical calculations involved in finding $\omega(E)$. Instead we take an alternative approach, using something known as the partition function \mathcal{Z} to link thermodynamics with statistical mechanics, rather than $\omega(E)$. The advantage is that while the latter requires evaluating hypervolumes, the partition function is delightfully simple to solve in contrast.

Consider a system S at a fixed temperature T and with energy E_S , connected to a heat bath R at the same temperature and with energy E_R . Now the heat bath plus the system form a isolated system so we must have that the total energy:

$$E = E_R + E_S \quad (10.1.1)$$

be a constant of motion. Furthermore, since the heat bath is much larger than the system in the macroscopic limit, we may assume that $\frac{E_S}{E} \ll 1$. Consequently, the microstates will now move on temperature surfaces rather than energy surfaces. All energy surfaces will be available, but the ones with large E_S will be less frequently occupied.

Let p_i be the probability of finding S in a microstate i with energy E_i , implying that the reservoir has energy $E - E_i$. The total number of microstates of the combined system is:

$$\Omega(E) = \sum_i \Omega_R(E - E_i) \quad (10.1.2)$$

where we sum over microstates. The entropy of the reservoir is given by the Taylor expansion:

$$S_R = k_B \ln \Omega_R(E - E_i) \approx k \ln \Omega_R(E) - \frac{\partial}{\partial E} (k \ln \Omega_R(E)) E_i + o(E_i^2) \quad (10.1.3)$$

and since $E \approx E_R$:

$$\frac{\partial S_R}{\partial E} = k_B \frac{\partial}{\partial E} (\ln \Omega_R(E)) = \frac{1}{T} \quad (10.1.4)$$

we find that:

$$\Omega_R(E - E_i) \approx \Omega_R(E) \exp \left[-\frac{E_i}{k_B T} \right] \quad (10.1.5)$$

Hence the total number of microstates is:

$$\Omega(E) = \Omega_R(E) \sum_j \exp \left[-\frac{E_j}{k_B T} \right] \quad (10.1.6)$$

We now invoke the postulate of equal a priori probabilities, states of the combined system with the same energy must be equally likely. It follows that out of the $\Omega(E)$ states we have calculated, only $\Omega_R(E) \exp \left(-\frac{E_i}{k_B T} \right)$ correspond to a microstate i . Hence the probability of finding the system in microstate i is therefore normalized to:

$$p_i = \frac{\exp \left(-\frac{E_i}{k_B T} \right)}{\sum_i \exp \left(-\frac{E_i}{k_B T} \right)} \quad (10.1.7)$$

or in more common notation, where $\beta = \frac{1}{k_B T}$:

$$p_i = \frac{\exp(-\beta E_i)}{\sum_j \exp(-\beta E_j)} \quad (10.1.8)$$

It is important to note that we are summing over all states j , rather than energies E_j , since there are $\Omega_S(E_j)$ possible microstates for a given energy E_j . If we wanted to sum over energy then we would have that the number of microstates of the combined system where the system is in microstate i becomes:

$$\Omega(E) = \Omega_S(E_i) \Omega_R(E - E_i) \quad (10.1.9)$$

and thus:

$$p_i = \frac{\exp(-\beta E_i)}{\sum_{E_j} \Omega_S(E_j) \exp(-\beta E_j)} \quad (10.1.10)$$

Furthermore, if we want the probability of the energy E_i rather than the microstate i then:

$$p(E_i) = \frac{\Omega_S(E_i) \exp(-\beta E_i)}{\sum_{E_j} \Omega_S(E_j) \exp(-\beta E_j)} \quad (10.1.11)$$

For continuous microstates we instead have:

$$\rho_c(\mathbf{q}, \mathbf{p}) = \frac{\exp(-\beta H(\mathbf{q}, \mathbf{p}))}{h^{3N} \int \exp(-\beta H(\mathbf{q}, \mathbf{p})) d^3 \mathbf{q} d^3 \mathbf{p}} \quad (10.1.12)$$

known as the canonical phase space density.

Two important remarks must be made about this distribution. Firstly, as expected, microstates with $E_i \geq k_B T$ are much less likely to be occupied. Furthermore, as $T \rightarrow 0$ the

system is forced to occupy the ground state. Indeed, we may rearrange (10.1.8) as:

$$p_{GS} = \frac{1}{1 + \sum_{j \neq GS} e^{-(E_j - E_{GS})/k_B T}} \quad (10.1.13)$$

and since $E_j - E_{GS} > 0$:

$$\lim_{T \rightarrow 0} p_{GS} = \lim_{T \rightarrow 0} \frac{1}{1 + \sum_{j \neq GS} e^{-(E_j - E_{GS})/k_B T}} = 1 \quad (10.1.14)$$

The denominator in (10.1.8) and (10.1.12) is known as the canonical partition function:

$$\mathcal{Z} = \frac{1}{h^{3N}} \int \exp(-\beta H(\mathbf{q}, \mathbf{p})) d^3 \mathbf{q} d^3 \mathbf{p} \quad (10.1.15)$$

Using the ensemble average definition of entropy:

$$S = -k_B \langle \ln \rho_c \rangle = h^{3N} \int \rho_c (-k_B \ln \rho_c) d^3 \mathbf{q} d^3 \mathbf{p} \quad (10.1.16)$$

$$= \hbar^{3N} k_B \int \rho_c (\beta H + \ln \mathcal{Z}) d^3 \mathbf{q} d^3 \mathbf{p} \quad (10.1.17)$$

$$= \frac{1}{T} \langle H \rangle + k_B \ln \mathcal{Z} \quad (10.1.18)$$

where $k_B \ln \mathcal{Z}$ is brought out of the integral since it is independent of the phase space variables. We now note that $\langle H \rangle$ may be identified as the internal energy U of the system, and that thus:

$$\langle U \rangle = -\frac{\partial}{\partial \beta} \ln \mathcal{Z} = k_B T^2 \frac{\partial}{\partial T} (\ln \mathcal{Z}) \quad (10.1.19)$$

and that consequently:

$$S = k_B \left(\ln \mathcal{Z} + T \frac{\partial}{\partial T} (\ln \mathcal{Z}) \right) \quad (10.1.20)$$

which may be elegantly put into the form:

$$S = k_B \frac{\partial}{\partial T} (T \ln \mathcal{Z}) \quad (10.1.21)$$

It follows that both the average energy and the entropy may be expressed in terms of the partition function and temperature alone.

Introducing the Hemholtz free energy:

$$F = U - TS \implies dF = dU - d(TS) \quad (10.1.22)$$

then we may use the first law of thermodynamics to write $dU = TdS - pdV$ and hence:

$$dF = -SdT - pdV = \left. \frac{\partial F}{\partial T} \right|_V dT + \left. \frac{\partial F}{\partial V} \right|_T dV \quad (10.1.23)$$

We then recognize that:

$$S = - \left. \frac{\partial F}{\partial T} \right|_V \quad (10.1.24)$$

$$p = - \left. \frac{\partial F}{\partial V} \right|_T \quad (10.1.25)$$

which should be familiar from the first part of the lecture notes. This new definition of entropy allows us to write:

$$k_B \frac{\partial}{\partial T} (T \ln \mathcal{Z}) = - \left. \frac{\partial F}{\partial T} \right|_V \implies \boxed{F = -k_B T \ln \mathcal{Z} = -\frac{1}{\beta} \ln \mathcal{Z}} \quad (10.1.26)$$

We could have also seen this *a priori* from:

$$F(T, V, N) = E - TS = k_B T^2 \frac{\partial}{\partial T} (\ln \mathcal{Z}) - k_B T \frac{\partial}{\partial T} (T \ln \mathcal{Z}) = -k_B T \ln \mathcal{Z} \quad (10.1.27)$$

Note also that systems will tend to equilibrium by minimizing their Helmholtz free energy, and thus by maximizing the partition function as can be seen from (10.1.26). In other words, systems will tend to macrostates with the largest number of microstates associated with them, thus maximizing entropy as desired.

We see that the partition function in the canonical ensemble assumes the role of Ω in the microcanonical ensemble, where we could find $S(E, V, N)$ from Ω , and then all other macroscopic quantities using Maxwell's relations. Similarly in the canonical ensemble we firstly find the Helmholtz free energy from the partition function \mathcal{Z} , and then derive all other microscopic quantities using Maxwell's relations.

10.2 Important examples

10.2.1 The ideal gas

Let's consider the ideal gas with Hamiltonian:

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{3N} \frac{p_i^2}{2m} \quad (10.2.1)$$

The partition function is then (taking into account the Gibbs factor due to the particle's indistinguishability):

$$\mathcal{Z} = \frac{1}{N!h^{3N}} \int \exp(-\beta H(\mathbf{q}, \mathbf{p})) d^3\mathbf{q} d^3\mathbf{p} \quad (10.2.2)$$

$$= \frac{1}{N!h^{3N}} \int \exp\left(-\beta \sum_{i=1}^{3N} \frac{p_i^2}{2m}\right) d^3\mathbf{q} d^3\mathbf{p} \quad (10.2.3)$$

$$= \frac{V^N}{N!h^{3N}} \prod_{i=1}^{3N} \int \exp\left(-\beta \frac{p_i^2}{2m}\right) dp_i \quad (10.2.4)$$

Substituting $x = \sqrt{\frac{\beta}{2m}} p_i$ then we find that:

$$\int e^{-x^2} dx = \sqrt{\pi} \implies \int \exp\left(-\beta \frac{p_i^2}{2m}\right) dp_i = \sqrt{\frac{2\pi m}{\beta}} \quad (10.2.5)$$

and thus:

$$\mathcal{Z} = \frac{V^N}{N!} \left(\frac{2\pi m}{\beta h^2}\right)^{3N/2} = \frac{V^N}{N!} \left(\frac{2\pi m k_B T}{h^2}\right)^{3N/2} \quad (10.2.6)$$

We may now introduce the thermal wavelength:

$$\lambda = \sqrt{\frac{h^2}{2\pi m k_B T}} \implies \mathcal{Z} = \frac{V^N}{N! \lambda^{3N}} \quad (10.2.7)$$

We may thus construct the Helmholtz free energy as:

$$F(T, V, N) = -k_B T \ln \frac{V^N}{N! \lambda^{3N}} = -k_B T \left[N \ln \frac{V}{\lambda^3} - \ln N! \right] \quad (10.2.8)$$

$$= -k_B T \left[N \ln \frac{V}{\lambda^3} - N \ln N + N \right] \quad (10.2.9)$$

$$= -N k_B T \left[1 + \ln \frac{V}{N \lambda^3} \right] \quad (10.2.10)$$

$$= -N k_B T \left\{ 1 + \ln \left[\frac{V}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] \right\} \quad (10.2.11)$$

$$(10.2.12)$$

Now that we have found the Hemholtz free energy, we can find all other thermodynamic quantities:

$$p = -\left.\frac{\partial F}{\partial V}\right|_{T,N} = \frac{Nk_B T}{V} \implies pV = Nk_B T \quad (10.2.13)$$

$$S = -\left.\frac{\partial F}{\partial T}\right|_{V,N} = Nk_B \left\{ \frac{5}{2} + \ln \left[\frac{V}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] \right\} \quad (10.2.14)$$

$$\mu = \left.\frac{\partial F}{\partial N}\right|_{T,V} = -k_B T \ln \left[\frac{V}{N} \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2} \right] \quad (10.2.15)$$

Also:

$$U = F + TS = \frac{3}{2} Nk_B T \quad (10.2.16)$$

10.2.2 The ultra-relativistic gas

The Hamiltonian for a non-interacting photon gas is given by:

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^N |\mathbf{p}_i| c \quad (10.2.17)$$

Since the particles are indistinguishable, we must include the Gibb's correction factor to the partition function:

$$Z(T, V, N) = \frac{1}{N! h^{3N}} \int \exp\{-\beta H(\mathbf{q}, \mathbf{p})\} d\mathbf{q} d\mathbf{p} \quad (10.2.18)$$

$$= \frac{V^N}{N! h^{3N}} \int \exp\left\{-\beta \sum_{i=1}^N |\mathbf{p}_i| c\right\} d\mathbf{q} d\mathbf{p} \quad (10.2.19)$$

$$= \frac{V^N}{N! h^{3N}} \prod_{i=1}^N \int \exp\{-\beta |\mathbf{p}_i| c\} d\mathbf{p}_i \quad (10.2.20)$$

It is now convenient to adopt spherical coordinates so that:

$$Z(T, V, N) = \frac{V^N}{N! h^{3N}} \prod_{i=1}^N \int \exp\{-\beta r_i c\} r_i^2 \sin \theta_i dr_i d\theta_i d\phi_i \quad (10.2.21)$$

$$= \frac{V^N}{N! h^{3N}} \prod_{i=1}^N \frac{8\pi}{(\beta c)^3} \quad (10.2.22)$$

$$= \frac{(8\pi V)^N}{N! (h\beta c)^{3N}} = \frac{(8\pi V)^N}{N! \lambda^{3N}} \quad (10.2.23)$$

where we define $\lambda = \frac{hc}{k_B T}$. Now that we have found the partition function it is straightforward to calculate the free energy:

$$F(T, V, N) = -k_B T \ln Z = -k_B T \ln \frac{(8\pi V)^N}{N! \lambda^{3N}} \quad (10.2.24)$$

$$= -k_B T \ln \left(\frac{(8\pi V)^N}{N! \lambda^{3N}} \right) \quad (10.2.25)$$

$$= -k_B T \left[N \ln \left(\frac{8\pi V}{\lambda^3} \right) - N \ln N + N \right] \quad (10.2.26)$$

$$= -N k_B T \left[\ln \left(\frac{8\pi V}{N \lambda^3} \right) + 1 \right] \quad (10.2.27)$$

$$= -N k_B T \left\{ \ln \left[\frac{8\pi V}{N} \left(\frac{k_B T}{hc} \right)^3 \right] + 1 \right\} \quad (10.2.28)$$

The other thermodynamic quantities can be found very easily from $F(T, V, N)$ as follows:

$$p = - \left. \frac{\partial F}{\partial V} \right|_{T, N} = \frac{N k_B T}{V} \implies pV = N k_B T \quad (10.2.29)$$

$$S = - \left. \frac{\partial F}{\partial T} \right|_{V, N} = N k_B \left\{ 4 + \ln \left[\frac{8\pi V}{N} \left(\frac{k_B T}{hc} \right)^3 \right] \right\} \quad (10.2.30)$$

$$\mu = \left. \frac{\partial F}{\partial N} \right|_{T, V} = -k_B T \ln \left[\frac{8\pi V}{N} \left(\frac{k_B T}{hc} \right)^3 \right] \quad (10.2.31)$$

giving us an expression for the internal energy:

$$U = F + TS = 3N k_B T \quad (10.2.32)$$

10.2.3 The harmonic oscillator

Consider a collection of N indistinguishable harmonic oscillators in three dimensions. The Hamiltonian for this system reads:

$$H(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{3N} \left(\frac{p_i^2}{2m} + \frac{1}{2} m \omega^2 q_i^2 \right) \quad (10.2.33)$$

The partition function reads:

$$\mathcal{Z}(T, V, N) = \frac{1}{N!h^{3N}} \int \exp\left\{-\beta \sum_{i=1}^{3N} \left(\frac{p_i^2}{2m} + \frac{1}{2}m\omega^2 q_i^2\right)\right\} d\mathbf{p}d\mathbf{q} \quad (10.2.34)$$

$$= \frac{1}{N!h^{3N}} \prod_{i=1}^{3N} \int \exp\left\{-\beta \frac{p_i^2}{2m}\right\} dp_i \int \exp\left\{-\frac{1}{2}\beta m\omega^2 q_i^2\right\} dq_i \quad (10.2.35)$$

$$= \frac{1}{N!h^{3N}} \prod_{i=1}^{3N} \sqrt{\frac{2m\pi}{\beta}} \sqrt{\frac{2\pi}{\beta m\omega^2}} \quad (10.2.36)$$

$$= \frac{1}{N!} \left(\frac{2\pi}{h\beta\omega}\right)^{3N} = \frac{1}{N!} \left(\frac{1}{\hbar\beta\omega}\right)^{3N} \quad (10.2.37)$$

Consequently, the free energy is found to be:

$$F(T, V, N) = -k_B T \ln \mathcal{Z} = -k_B T \ln \left[\frac{1}{N!} \left(\frac{1}{\hbar\beta\omega}\right)^{3N} \right] \quad (10.2.38)$$

$$= -Nk_B T \left\{ \ln \left[\frac{1}{N(\hbar\beta\omega)^3} \right] + 1 \right\} \quad (10.2.39)$$

Therefore the pressure, entropy and chemical potential are found to be:

$$p = -\left. \frac{\partial F}{\partial V} \right|_{T, N} = 0 \quad (10.2.40)$$

$$S = -\left. \frac{\partial F}{\partial T} \right|_{V, N} = Nk_B \left\{ \ln \left[\frac{1}{N(\hbar\beta\omega)^3} \right] + 4 \right\} \quad (10.2.41)$$

$$\mu = \left. \frac{\partial F}{\partial N} \right|_{T, V} = -k_B T \ln \left[\frac{1}{N(\hbar\beta\omega)^3} \right] \quad (10.2.42)$$

and the internal potential thus takes the form of $U = F + TS = 3Nk_B T$.

10.3 The equipartition and virial theorems

Derivation in the canonical ensemble

Let $H(q_i, p_i)$ be the Hamiltonian of a system and let x_i denote all phase coordinates $\{q_i, p_i\}$. Therefore:

$$\left\langle x_i \frac{\partial H}{\partial x_k} \right\rangle = \frac{1}{h^{3N}} \int \rho(\mathbf{x}) x_i \frac{\partial H}{\partial x_k} d\mathbf{x} \quad (10.3.1)$$

$$= \frac{1}{h^{3N} \mathcal{Z}} \int x_i \exp(-\beta H(\mathbf{x})) \frac{\partial H}{\partial x_k} d\mathbf{x} \quad (10.3.2)$$

$$= -\frac{1}{\beta h^{3N} \mathcal{Z}} \int x_i \frac{\partial}{\partial x_k} (x_i \exp(-\beta H(\mathbf{x}))) d\mathbf{x} \quad (10.3.3)$$

We can now integrate by parts. First, we define $d\mathbf{x}'$ to be the infinitesimal volume element excluding dx_k . Moreover, we also assume that $x_i \exp\{-\beta H\}$ vanishes at the boundaries

$\min x_k$ and $\max x_k$ (usually because $H \rightarrow \infty$).

$$\left\langle x_i \frac{\partial H}{\partial x_k} \right\rangle = -\frac{1}{\beta h^{3N} \mathcal{Z}} \left[\int [x_i \exp(-\beta H(\mathbf{x}))]_{\min x_k}^{\max x_k} d\mathbf{x}' - \int \delta_{ik} \exp(-\beta H(\mathbf{x})) d\mathbf{x} \right] \quad (10.3.4)$$

$$= \frac{1}{h^{3N} \beta} \delta_{ik} \int \frac{\exp(-\beta H(\mathbf{x}))}{\mathcal{Z}} d\mathbf{x} \quad (10.3.5)$$

$$= \delta_{ik} k_B T \quad (10.3.6)$$

It follows from that if we choose canonically conjugate variables, the expectation value of their product will be equal to $k_B T$. For example:

$$\left\langle q_i \frac{\partial H}{\partial q_i} \right\rangle = -\langle q_i \dot{p}_i \rangle = k_B T \quad (10.3.7)$$

$$\left\langle p_i \frac{\partial H}{\partial p_i} \right\rangle = \langle p_i \dot{q}_i \rangle = k_B T \implies \left\langle \frac{p_i^2}{2m} \right\rangle = \frac{1}{2} k_B T \quad (10.3.8)$$

The latter can be used to find the mean kinetic energy of the i th particle in three dimensions:

$$\langle T_i \rangle = \frac{3}{2} k_B T \quad (10.3.9)$$

so for a collection of N particles the mean kinetic energy is:

$$\langle T \rangle = \frac{3}{2} N k_B T \quad (10.3.10)$$

Suppose the i th particle is acted on by a conservative force $\mathbf{F}_i = -\nabla V_i$. Then, we see that:

$$-\langle \mathbf{r}_i \cdot \mathbf{F}_i \rangle = \left\langle x_i \frac{\partial V_i}{\partial x_i} \right\rangle + \left\langle y_i \frac{\partial V_i}{\partial y_i} \right\rangle + \left\langle z_i \frac{\partial V_i}{\partial z_i} \right\rangle = 3k_B T \quad (10.3.11)$$

so for a central potential of the form $V_i \propto r_i^\alpha$ then $\mathbf{r}_i \cdot \mathbf{F}_i = -\alpha V_i$ giving:

$$\langle \alpha V_i \rangle = 3k_B T = 2 \langle T_i \rangle \implies \langle T \rangle_i = \frac{\alpha}{2} \langle V_i \rangle \quad (10.3.12)$$

which is known as the Virial theorem. Note that in this derivation we assumed that integration in x could be done. While in classical mechanics there is no ambiguity in this assumption, in quantum mechanics this implies that the temperature $k_B T$ must be larger than the quanta of energy in the system. In the special case where $\alpha = 2$ we get that $\langle T_i \rangle = \langle V_i \rangle$, the potential and kinetic energies contribute equally to the total energy.

More specifically, we have that for a hamiltonian with f degrees of freedom containing only quadratic terms:

$$H = \sum_{i=1}^f (a_i q_i^2 + b_i p_i^2) \quad (10.3.13)$$

then clearly:

$$H = \frac{1}{2} \sum_{i=1}^f \left(q_i \frac{\partial H}{\partial q_i} + p_i \frac{\partial H}{\partial p_i} \right) \quad (10.3.14)$$

$$\iff U = \langle H \rangle = \frac{1}{2} \sum_{i=1}^f \left(\langle q_i \frac{\partial H}{\partial q_i} \rangle + \langle p_i \frac{\partial H}{\partial p_i} \rangle \right) = f k_B T \quad (10.3.15)$$

In other words, since there are $2f$ quadratic terms, each quadratic term contributes $\frac{1}{2}k_B T$ to the internal energy. This result is known as the **equipartition theorem**. For $f = 3N$ then we get the familiar:

$$U = \langle H \rangle = 3N k_B T \quad (10.3.16)$$

which we have found several times already in the previous sections.

Derivation in the microcanonical ensemble

Let $H(q_i, p_i)$ be the Hamiltonian of a system and let x_i denote all phase coordinates $\{q_i, p_i\}$. In the microcanonical ensemble, we have that:

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = \frac{1}{h^{3N}} \int \rho(\mathbf{x}) x_i \frac{\partial H}{\partial x_j} d\omega \quad (10.3.17)$$

$$= \frac{1}{h^{3N}} \frac{1}{\Omega \Delta E} \int_{\Delta E} x_i \frac{\partial H}{\partial x_j} d\omega \quad (10.3.18)$$

$$= \frac{1}{h^{3N}} \frac{1}{\Omega} \frac{\partial}{\partial E} \int_{H < E} x_i \frac{\partial H}{\partial x_j} d\omega \quad (10.3.19)$$

$$= \frac{1}{h^{3N}} \frac{1}{\Omega} \frac{\partial}{\partial E} \int_{H < E} x_i \frac{\partial(H - E)}{\partial x_j} d\omega \quad (10.3.20)$$

where in the last line we used $\frac{\partial E}{\partial x_j} = 0$ where $x = q, p$. This allows us to integrate by parts:

$$\frac{1}{h^{3N}} \frac{1}{\Omega} \frac{\partial}{\partial E} \int_{H < E} x_i \frac{\partial(H - E)}{\partial x_j} d\omega = \frac{1}{h^{3N}} \frac{1}{\Omega} \frac{\partial}{\partial E} \left[\int_{H < E} x_i \left[\frac{\partial(H - E)}{\partial x_j} \right]_{\min x_j}^{\max x_j} d\omega' - \int_{H < E} (H - E) \frac{\partial x_i}{\partial x_j} d\omega \right] \quad (10.3.21)$$

where $d\omega'$ integrates over all variables except x_j . If x_j has an extremal value, then this means that it must lie on the energy surface $H = E$, and consequently the integrand vanishes.

The integral thus simplifies to:

$$-\frac{1}{h^{3N}} \frac{1}{\Omega} \frac{\partial}{\partial E} \int_{H < E} (H - E) \delta_{ij} d\omega = \frac{1}{h^{3N}} \frac{1}{\Omega} \delta_{ij} \int_{H < E} d\omega \quad (10.3.22)$$

$$= \frac{1}{h^{3N}} \frac{1}{\Omega} \delta_{ij} \omega \quad (10.3.23)$$

$$= \frac{\omega'}{\Omega} \delta_{ij} \quad (10.3.24)$$

where we defined $\omega' = \frac{1}{h^{3N}}\omega$. Using the definition of $\Omega = \frac{\partial\omega'}{\partial E}$ we retrieve the remarkable result:

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = \frac{\omega'}{\frac{\partial\omega'}{\partial E}} \delta_{ij} = \frac{\omega'}{1} \frac{\partial}{\partial E} (\ln \omega') \delta_{ij} \quad (10.3.25)$$

$$= \frac{\omega'}{k_B} \frac{\partial S}{\partial E} \delta_{ij} = k_B T \delta_{ij} \quad (10.3.26)$$

as found previously.

Canonical examples of the canonical ensemble

11

- 11.1 The Quantum Harmonic Oscillator
- 11.2 The Diatomic gas
- 11.3 The Paramagnet

The Grand Canonical Ensemble

12.1 Adding the chemical potential

We have seen how to describe isolated systems which conserve both energy and particle number using the microcanonical ensemble, as well as closed systems which conserve only particle number, but allow variable energies, using the canonical ensemble. It is thus fitting to have a third ensemble which can handle open systems where neither energy nor particle number are constant. Such an ensemble is of particular use when describing systems where heat and particles can be exchanged through contact with a heat and particle bath.

The natural variables to be used are T , V and μ , the chemical potential.

The chemical potential is defined as the amount by which the internal energy will change when adding a particle to the system isentropically and isochorically (constant entropy and volume). We must therefore revisit our First Law of Thermodynamics and add a new term:

$$\boxed{dU = TdS - pdV + \mu dN} \quad (12.1.1)$$

giving us the following definition of μ :

$$\mu = \left(\frac{\partial U}{\partial N} \right)_{S,V} \quad (12.1.2)$$

Recalling the definitions of the other thermodynamic potentials:

$$F = U - TS \implies dF = -pdV - SdT + \mu dN \quad (12.1.3)$$

$$G = F + pV \implies dG = Vdp - SdT + \mu dN \quad (12.1.4)$$

and equivalently:

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{V,T}, \quad \mu = \left(\frac{\partial G}{\partial N} \right)_{p,T} \quad (12.1.5)$$

The chemical potential can also be related to entropy and the second law of thermodynamics. Using U , V , N as natural variables then:

$$dS = \left(\frac{\partial S}{\partial U} \right)_{V,N} + \left(\frac{\partial S}{\partial V} \right)_{U,N} + \left(\frac{\partial S}{\partial N} \right)_{V,U} \quad (12.1.6)$$

which when compared with

$$dS = \frac{dU}{T} + \frac{p}{T}dV - \frac{\mu}{T}dN \quad (12.1.7)$$

yields:

$$\boxed{\left(\frac{\partial S}{\partial U}\right)_{V,N} = \frac{1}{T}, \quad \left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{p}{T}, \quad \left(\frac{\partial S}{\partial N}\right)_{V,U} = -\frac{\mu}{T}} \quad (12.1.8)$$

This can be used to investigate how and when a system reaches equilibrium. Indeed, suppose we have two systems decoupled to their surroundings which exchange heat and particles. It is clear that $dU_1 = -dU_2 = dU, dV_1 = -dV_2 = dV, dN_1 = -dN_2 = dN$ since the two systems combined form an isolated system. Then:

$$dS_{tot} = \frac{dU_1}{T_1} + \frac{p_1}{T_1}dV_1 - \frac{\mu_1}{T_1}dN_1 + \frac{dU_2}{T_2} + \frac{p_2}{T_2}dV_2 - \frac{\mu_2}{T_2}dN_2 \quad (12.1.9)$$

$$= \left(\frac{1}{T_1} - \frac{1}{T_2}\right)dU + \left(\frac{p_1}{T_1} - \frac{p_2}{T_2}\right)dV + \left(\frac{\mu_1}{T_1} - \frac{\mu_2}{T_2}\right)dN \quad (12.1.10)$$

For equilibrium to be established we need $dS_{tot} = 0$ and hence:

$$\boxed{T_1 = T_2, \quad p_1 = p_2, \quad \mu_1 = \mu_2} \quad (12.1.11)$$

The first imposes **thermal equilibrium**, the second imposes **mechanical equilibrium** and the third imposes **chemical equilibrium**. Now suppose we have just one system which can exchange U, V, N with a reservoir. We then see that:

$$dS_{tot} = dS - \frac{dU}{T_R} - \frac{p_R}{T_R}dV + \frac{\mu_R}{T_R}dN \quad (12.1.12)$$

Define the availability as we have previously as $dA = -T_R dS_{tot}$, then:

$$dA = -T_R dS + dU + p_R dV - \mu_R dN \quad (12.1.13)$$

and inserting $dU = TdS - pdV + \mu dN$ we finally find:

$$\boxed{dA = (T - T_R)dS - (p - p_R)dV + (\mu - \mu_R)dN} \quad (12.1.14)$$

From our definition of the availability, equilibrium is established when availability is minimized (entropy maximized).

12.2 Euler's relation and the Gibbs-Duhem equation

After having re-introduced a myriad of new notation it may be interesting to know how the intensive variables μ, T and p are related. Note that the internal energy is an extensive quantity and thus satisfies

$$U(\lambda S, \lambda V, \lambda N) = \lambda U(S, V, N) \quad (12.2.1)$$

for any scaling factor λ . Differentiating with respect to λ we find that

$$U(S, V, N) = \left. \frac{\partial U(\lambda S)}{\partial(\lambda S)} \right|_{V, N} S + \left. \frac{\partial U(\lambda V)}{\partial(\lambda V)} \right|_{S, N} V + \left. \frac{\partial U(\lambda N)}{\partial(\lambda N)} \right|_{V, S} N \quad (12.2.2)$$

$$= TS - pV + \mu N \quad (12.2.3)$$

so we have found that

$$\boxed{U = TS - pV + \mu N} \implies dU = d(TS) - d(pV) + d(\mu N) \quad (12.2.4)$$

which is known as Euler's relation. The same could have been derived using any other extensive state function (e.g. entropy). Note however that from the first law, $dU = TdS - pdV + \mu dN$ so it follows that

$$\boxed{SdT - Vdp + Nd\mu = 0} \quad (12.2.5)$$

which is known as the Gibbs-Duhem equation. It shows that T, p, μ are not independent and any one of them can be expressed as a function of the other two!

12.3 The Grand canonical phase space density

We consider an open system S with energy E_S and number of particles N_S at temperature T and chemical potential μ surrounded by a reservoir E_R and number of particles N_R at the same temperature and chemical potential. The system plus the reservoir form an isolated system which we refer to as the total system. Thus we define $E = E_S + E_R$ and $N = N_S + N_R$ to be the energy and number of particles of the total system, and assume that:

$$\frac{N_S}{N} \ll 1, \quad \frac{E_S}{E} \ll 1 \quad (12.3.1)$$

The number of microstates of the system consistent with the macrostate E, N is:

$$\Omega(E, N) = \sum_{ij} \Omega_R(E - E_i, N - N_j) \quad (12.3.2)$$

where we sum over microstates with energy E_i with particle number N_j . Defining the Boltzmann entropy of the reservoir as $S_R = k_B \ln \Omega_R$ we find that:

$$S_R = k_B \ln \Omega_R(E - E_i, N - N_j) \quad (12.3.3)$$

$$\approx k_B \ln \Omega_R(E, N) - E_i \frac{\partial S}{\partial E} - N_j \frac{\partial S}{\partial N} \quad (12.3.4)$$

$$\approx k_B \ln \Omega_R(E, N) - \frac{E_i}{T} + \frac{\mu N_j}{T} \quad (12.3.5)$$

where we used (12.1.8). Consequently we find that:

$$\Omega_R(E - E_i, N - N_j) = \Omega_R(E, N) \exp(-\beta(E_i - \mu N_j)) \quad (12.3.6)$$

giving the following probability distribution:

$$p_{ij} = \frac{1}{\mathcal{Z}} \exp(-\beta(E_i - \mu N_j)), \quad \mathcal{Z} = \sum_{ml} \exp(-\beta(E_m - \mu N_l)) \quad (12.3.7)$$

We define \mathcal{Z} as the grand partition function.

An alternative way to derive (12.3.7) is by maximizing the Gibbs entropy.

We can extend (12.3.7) to the continuous case, letting $\alpha = \beta\mu$ it is given by:

$$\rho(\mathbf{p}, \mathbf{q}, N) = \frac{1}{\mathcal{Z}} \exp(-\beta H + \alpha N), \quad \mathcal{Z} = \sum_N \frac{1}{h^{3N}} \int \exp(-\beta H + \alpha N) d^{3N} \mathbf{p} d^{3N} \mathbf{q} \quad (12.3.8)$$

Using the ensemble average definition of entropy we get that:

$$S = -k_B \langle \ln \rho \rangle = -\frac{k_B}{\mathcal{Z}} \int \exp(-\beta H + \alpha N) [-\ln \mathcal{Z} - \beta H + \alpha N] d\omega \quad (12.3.9)$$

$$= k_B [\ln \mathcal{Z} + \beta U - \alpha \langle N \rangle] \quad (12.3.10)$$

Let $N = \langle N \rangle$ to be the average particle number, then we can define the grand canonical potential Φ_G as:

$$\Phi_G = -\frac{1}{\beta} \ln \mathcal{Z} = -TS + U - \mu N = F - \mu N \quad (12.3.11)$$

Alternatively, using Euler's relation we can write

$$\Phi_G = -pV \quad (12.3.12)$$

which tells us that the Gibbs free energy takes the simple form:

$$G = U + pV - TS = \mu N \quad (12.3.13)$$

Finally, the thermodynamic variables can be expressed in terms of the grand potential:

$$d\Phi_G = dF - \mu dN - Nd\mu = -SdT - pdV - Nd\mu \quad (12.3.14)$$

$$\Rightarrow \mu = -\left. \frac{\partial \Phi_G}{\partial N} \right|_{V,S}, \quad p = -\left. \frac{\partial \Phi_G}{\partial V} \right|_{S,N}, \quad S = -\left. \frac{\partial \Phi_G}{\partial T} \right|_{N,V} \quad (12.3.15)$$

12.4 The ideal gas in the GCE

12.5 Particle number fluctuation

The canonical partition function Z_N for an N particle system can be related to the grand partition function of the same system with variable particle number via a discrete Laplace

transform:

$$\mathcal{Z} = \sum_N z^N Z_N, \quad z = e^{\mu/\beta} \quad (12.5.1)$$

where z is known as the fugacity. We see that, much like the Boltzmann weight is used to average over microcanonical ensembles at different E to get the canonical ensemble, the fugacity is the weight used to average over the canonical ensembles at different N to get the grand canonical ensemble.

The fugacity may be used to express N in a different way. Indeed (we use $\langle N \rangle$ instead of N to avoid confusion):

$$\langle N \rangle = \sum_N N p_N = \sum_N N \sum_i \frac{\exp(-\beta E_i)}{\mathcal{Z}} z^N = \frac{1}{\mathcal{Z}} \sum_N N z^N Z_N \quad (12.5.2)$$

Note also that:

$$\left. \frac{\partial \mathcal{Z}}{\partial z} \right|_{T,V} = \sum_N N z^{N-1} Z_N(T, V) \quad (12.5.3)$$

which allows us to write the particle number as a derivative with respect to fugacity rather than chemical potential:

$$\langle N \rangle = z \left. \frac{\partial}{\partial z} (\ln \mathcal{Z}) \right|_{T,V} = - \frac{z}{\beta} \left. \frac{\partial \Phi_G}{\partial z} \right|_{T,V} \quad (12.5.4)$$

We would also like to find the fluctuation in $\langle N \rangle$ so we need to calculate $\langle N^2 \rangle$:

$$\langle N^2 \rangle = \frac{1}{\mathcal{Z}} \sum_N N^2 z^N Z_N \quad (12.5.5)$$

To do so we investigate the second order derivative of $\ln \mathcal{Z}$:

$$\left. \frac{\partial^2 \mathcal{Z}}{\partial z^2} \right|_{T,V} = \frac{\partial}{\partial z} \left(\frac{1}{\mathcal{Z}} \sum_N N z^{N-1} Z_N \right) \Big|_{T,V} \quad (12.5.6)$$

$$= - \left(\frac{1}{\mathcal{Z}} \left. \frac{\partial \mathcal{Z}}{\partial z} \right|_{T,V} \right)^2 + \frac{1}{\mathcal{Z}} \sum_N N(N-1) z^{N-2} Z_N \quad (12.5.7)$$

$$= - \left(\frac{\langle N \rangle}{z} \right)^2 + \frac{1}{\mathcal{Z}} \sum_N N^2 z^{N-2} Z_N - \frac{1}{\mathcal{Z}} \sum_N z^{N-2} Z_N \quad (12.5.8)$$

$$\implies z^2 \frac{\partial^2 \mathcal{Z}}{\partial z^2} = \langle N^2 \rangle - \langle N \rangle^2 - \langle N \rangle \quad (12.5.9)$$

Substituting (12.5.4) into the above we find that

$$(\Delta N)^2 = z^2 \left. \frac{\partial^2 \mathcal{Z}}{\partial z^2} \right|_{T,V} - z \left. \frac{\partial}{\partial z} (\ln \mathcal{Z}) \right|_{T,v} = \left(z \frac{\partial}{\partial z} \right)^2 (\ln \mathcal{Z}) \Big|_{T,V} \quad (12.5.10)$$

We can also use $\langle N \rangle = z \frac{\partial}{\partial z} = z \frac{\partial}{\partial z} (\ln \mathcal{Z})|_{T,V} = -\frac{z}{\beta} \frac{\partial \Phi_G}{\partial z}|_{T,V}$ to write

$$\boxed{(\Delta N)^2 = z \frac{\partial \langle N \rangle}{\partial z} \Big|_{T,V} = \frac{1}{\beta} \frac{\partial \langle N \rangle}{\partial \mu} \Big|_{T,V}} \quad (12.5.11)$$

since $\frac{\partial \mu}{\partial z} = \frac{1}{\beta}$. Also, note that

$$\langle N \rangle = -\frac{\partial \Phi_G}{\partial \mu} \Big|_{T,V} = V \frac{\partial p}{\partial \mu} \Big|_{T,V} \implies \frac{\partial^2 p}{\partial \mu^2} \Big|_{T,V} = \frac{\beta}{V} (\Delta N)^2 \quad (12.5.12)$$

Note that the LHS is an intensive quantity, so for the RHS to also be intensive we should require

$$\frac{\Delta N}{\langle N \rangle} \sim \frac{\sqrt{V}}{\langle N \rangle} \sim \frac{1}{\sqrt{\langle N \rangle}} \quad (12.5.13)$$

so the fluctuations in particle number are of the order of magnitude of the inverse root of the average particle number. In the thermodynamic limit the number fluctuations thus vanishes thus showing the the canonical and grand canonical ensembles are indeed equivalent.

12.6 Mechanical stability condition

In the canonical ensemble we showed that the heat capacity was proportional to the energy fluctuation squared, and thus had to be non-negative yielding a thermal stability condition. Similarly, we can derive a mechanical stability condition, namely that the thermal compressibility must be non-negative.

Let us define $f = F(T, V, N)/N$ to be the free energy per particle. Note that

$$\mu = \frac{\partial F}{\partial N} \Big|_{V,T} = f + N \frac{\partial f}{\partial N} \Big|_{V,T} \quad (12.6.1)$$

and letting $v = \frac{V}{N}$ then $\frac{\partial f}{\partial N} \Big|_{V,T} = -\frac{V}{N^2} \frac{\partial f}{\partial v} \Big|_{V,T}$ which implies

$$\mu = f - v \frac{\partial f}{\partial v} \Big|_{V,T} \implies \frac{\partial \mu}{\partial v} \Big|_{V,T} = -v \frac{\partial^2 f}{\partial v^2} \Big|_{V,T} \quad (12.6.2)$$

We can do the same for pressure

$$p = -\frac{\partial F}{\partial V} \Big|_{N,T} = -N \frac{\partial f}{\partial V} \Big|_{N,T} = -\frac{\partial f}{\partial v} \Big|_{N,T} \quad (12.6.3)$$

implying that

$$\frac{\partial p}{\partial v} \Big|_{N,T} = -\frac{\partial^2 f}{\partial v^2} \Big|_{N,T} \implies \frac{\partial \mu}{\partial v} \Big|_{V,T} = v \frac{\partial p}{\partial v} \Big|_{N,T} \quad (12.6.4)$$

These relations can be used to express the particle fluctuation as

$$\frac{\beta}{V}(\Delta N)^2 = \frac{\partial^2 p}{\partial \mu^2} \Big|_{T,V} = \frac{\partial v}{\partial \mu} \Big|_{T,V} \frac{\partial}{\partial v} \left(\frac{\partial p}{\partial \mu} \right) \Big|_{T,V} \quad (12.6.5)$$

$$= \frac{1}{v} \frac{\partial v}{\partial p} \Big|_{N,T} \frac{\partial}{\partial v} \frac{1}{v} \Big|_{T,V} = \frac{1}{v^2} \kappa_T \quad (12.6.6)$$

where we defined the thermal compressibility

$$\kappa_T = -\frac{1}{v} \frac{\partial v}{\partial p} \Big|_{N,T} = -\frac{1}{V} \frac{\partial V}{\partial p} \Big|_{N,T} \quad (12.6.7)$$

since the partial derivative is at constant particle number. Thus we have found that the compressibility is related to the particle number fluctuation

$$\boxed{\kappa_T = \beta V \left(\frac{\Delta N}{N} \right)^2 > 0} \quad (12.6.8)$$

The quantities C_V and κ_T are both response functions, they give the system's response to a temperature and pressure perturbation respectively, in the form of the heat and volume changes. What the mechanical stability condition says is that $\kappa_T \delta V > 0$, namely the system's response to a volume increase δV , must be to increase the pressure by $\delta p > 0$ so as to counteract the volume change. If the compressibility were negative then an increase in volume $\delta V > 0$ would lead to a decrease in pressure $\delta p < 0$ making V spiral out of control! Such a macrostate cannot possibly exist as it is extremely unstable to perturbations, an example of this will be seen when discussing the liquid-gas phase coexistence in the Van der Waals model.

The fact that these response functions are related to fluctuations is not a coincidence at all, and goes by the name of the **fluctuation-dissipation theorem**. In general, a system where a quantity is not fluctuating will not respond to a perturbation in this quantity.

Part IV

Quantum statistical physics

Density operator methods

13.1 Making the jump to quantum

In classical statistical mechanics we assigned probabilities to microstates (\mathbf{q}, \mathbf{p}) , which were points in the phase space. In quantum statistical mechanics our microstates will be quantum states $|\psi\rangle$ in the Hilbert space.

We also measure observables $A(\mathbf{q}, \mathbf{p})$ defined as functions in the phase space. In quantum mechanics observables are replaced by linear hermitian operators $\hat{A} \in \mathcal{L}(\mathcal{H})$ acting on the Hilbert space. The expectation value of some operator in a given state gives the expectation value, the average value measured in an experiment.

	CSM	QSM
Microstate	$(\mathbf{q}, \mathbf{p}) \in \mathcal{P}$	$ \psi\rangle \in \mathcal{H}$
Observable	$A(\mathbf{q}, \mathbf{p})$	$\langle A \rangle, \hat{A} \in \mathcal{L}(\mathcal{H})$
Canonical conjugation	$\{q_i, p_j\} = \delta_{ij}$	$[\hat{q}_i, \hat{p}_i] = i\hbar\delta_{ij}$
Time evolution	$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}$	$\frac{d\langle A \rangle}{dt} = \frac{1}{i\hbar} \langle [\hat{A}, \hat{H}] \rangle + \langle \frac{\partial \hat{A}}{\partial t} \rangle$

Consider a quantum system prepared in an ensemble of states $\{|\psi_i\rangle\}$ each with probability p_i . That is, not only do we have uncertainties inherent to quantum mechanics, but we also have the uncertainty in the wavefunction of the system. The state of a system prepared like this is known as a mixed state. The ensemble averaged expectation value \bar{f} of some operator \hat{f} is given by the weighted average of the expectation value for each state in the ensemble:

$$\langle \bar{f} \rangle = \sum_i p_i \langle \psi_i | \hat{f} | \psi_i \rangle \quad (13.1.1)$$

We can expand this in a given orthonormal basis as:

$$\langle \bar{f} \rangle = \sum_{imn} p_i \langle \psi_i | n \rangle \langle n | \hat{f} | m \rangle \langle m | \psi_i \rangle = \sum_{mn} \langle n | \hat{f} | m \rangle \sum_i \langle m | \psi_i \rangle \langle \psi_i | n \rangle \quad (13.1.2)$$

We define the density matrix related to the state of this system is given by:

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle \psi_i| \quad (13.1.3)$$

which allows us to write:

$$\langle \bar{f} \rangle = \sum_{mn} \langle n | \hat{f} | m \rangle \langle m | \rho | n \rangle = \text{Tr}(\hat{\rho} \hat{f}) \quad (13.1.4)$$

So the expectation value of an observable f for a system in a mixed state defined by a density matrix ρ can be expressed as a trace:

$$\boxed{\langle \bar{f} \rangle = \text{Tr}(\hat{\rho} \hat{f})} \quad (13.1.5)$$

(13.1.4) is reminiscent of the classical formula:

$$\langle f \rangle = \int f(\mathbf{p}, \mathbf{q}) \rho(\mathbf{p}, \mathbf{q}) d\omega \quad (13.1.6)$$

so we should identify the density matrix with the ensemble probability density. In other words $\rho(\mathbf{p}, \mathbf{q})$ got promoted to $\hat{\rho}$.

It is easy to see that:

(i) the density matrix has unit trace:

$$\text{Tr} \hat{\rho} = \sum_j \sum_i p_i \langle j | i \rangle \langle i | j \rangle = \sum_j p_j = 1 \quad (13.1.7)$$

corresponding to the normalization of classical probability densities.

(ii) the density matrix is hermitian:

$$\hat{\rho}^\dagger = \sum_i p_i (|i\rangle \langle i|)^\dagger = \rho \quad (13.1.8)$$

corresponding to the fact that classical probabilities are real numbers.

(iii) The density operator is positive definite:

$$\langle \psi | \hat{\rho} | \psi \rangle = \sum_i p_i |\langle \psi | \psi_i \rangle|^2 > 0, \quad \forall |\psi\rangle \in \mathcal{H} \quad (13.1.9)$$

corresponding to the fact that the classical probability density must be positive.

Recall also that in classical physics the evolution of ρ was given by Liouville's theorem:

$$\frac{\partial \rho}{\partial t} = \{H, \rho\} \quad (13.1.10)$$

Note that in the Schrodinger picture we have a completely analogous result:

$$i\hbar \frac{d\hat{\rho}}{dt} = \hat{H} \sum_i p_i (|\psi_i\rangle \langle \psi_i| - |\psi_i\rangle \langle \psi_i|) \hat{H} = \hat{H} \hat{\rho} - \hat{\rho} \hat{H} \quad (13.1.11)$$

$$\implies \boxed{\frac{d\hat{\rho}}{dt} = \frac{1}{i\hbar} [\hat{H}, \hat{\rho}]} \quad (13.1.12)$$

which is known as the Von Neumann equation. It is important to notice the remarkable duality between Poisson brackets and commutators. Here we assumed that p_i does not change, this is reasonable as long as the system we consider is not interacting with its surroundings.

The equilibrium density operator must be time-independent so we need $[\hat{H}, \hat{\rho}] = 0$. Thus in statistical mechanics we require ρ to be a function of operators commuting with \hat{H} , thus corresponding to conserved quantities. A completely analogous statement applied to classical statistical mechanics.

The time evolution of the density operator is generally non-linear, unlike the Schrodinger evolution which only applies to pure states. It follows that:

$$i\hbar \frac{d\langle f \rangle}{dt} = i\hbar \frac{d}{dt} [\text{Tr}(\hat{\rho}\hat{f})] = \text{Tr} \left([\hat{H}, \hat{\rho}]\hat{f} + i\hbar \hat{\rho} \frac{\partial \hat{f}}{\partial t} \right) = \text{Tr}([\hat{H}, \hat{\rho}]\hat{f}) + i\hbar \left\langle \frac{\partial \hat{f}}{\partial t} \right\rangle \quad (13.1.13)$$

so in equilibrium:

$$\left\langle \frac{\partial \hat{f}}{\partial t} \right\rangle = 0 \quad (13.1.14)$$

13.2 Quantum statistical ensembles

Microcanonical ensemble

We find that:

$$\rho_{mc} = \frac{1}{\Omega(E)} \delta_{H,E} \quad (13.2.1)$$

so the density matrix elements are:

$$\langle m | \rho_{mc} | n \rangle = \frac{1}{\Omega(E)} \begin{cases} 1, & \text{if } E_m = E \text{ and } m = n \\ 0, & \text{if } E_m \neq E \text{ or } m \neq n \end{cases} \quad (13.2.2)$$

The energy condition $E_m = E$ comes from the delta function $\delta_{H,E}$ while the diagonal condition $m = n$ comes from the orthonormality of energy eigenstates. It follows that the density matrix is diagonal whose only non-zero elements are those with energy E . Note that this density operator satisfies the equal-a-priori probability since eigenstates of equal energy have the same probability.

Canonical ensemble

We find that:

$$\hat{\rho} = \frac{1}{Z} e^{-\beta \hat{H}}, \quad Z = \text{Tr}(e^{-\beta \hat{H}}) \quad (13.2.3)$$

or in the energy eigenbasis $\{|n\rangle\}$:

$$\hat{\rho} = \frac{1}{Z} e^{-\beta E_n} |n\rangle \langle n| \quad (13.2.4)$$

13.3 Particle in a box

Consider a free particle in a box of length L with periodic boundary conditions. Then the energy eigenstates in the position basis read:

$$\langle x|k\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \quad (13.3.1)$$

with energy

$$E_k = \frac{\hbar^2 k^2}{2m}, \quad k_{x,y,z} = \frac{2\pi}{L} n_{x,y,z} \quad (13.3.2)$$

We are interested in the matrix elements of the density matrix in the position representation. We look at the infinite L limit where the possible momenta now form a continuum. Then we must replace:

$$\frac{1}{V} \sum_{\mathbf{k}} = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \quad (13.3.3)$$

where $V = L^3$. Therefore:

$$\langle \mathbf{r}|\hat{\rho}|\mathbf{r}'\rangle = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-\beta E(k)} \langle \mathbf{r}|\mathbf{k}\rangle \langle \mathbf{k}|\mathbf{r}'\rangle \quad (13.3.4)$$

$$= \frac{1}{Z} V \int \frac{d^3\mathbf{k}}{(2\pi)^3} \exp \left[-\frac{\beta \hbar^2 \mathbf{k}^2}{2m} + i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}') \right] \quad (13.3.5)$$

$$= \frac{1}{Z} \frac{1}{(2\pi)^3} \left(\sqrt{\frac{2m\pi}{\beta \hbar^2}} \right)^3 \exp \left(-\frac{m}{2\beta \hbar^2} (\mathbf{r} - \mathbf{r}')^2 \right) \quad (13.3.6)$$

We define the thermal wavelength to be:

$$\lambda = \frac{\hbar}{\sqrt{2\pi m k_B T}} \quad (13.3.7)$$

which further simplifies (13.3.6) into:

$$\rho(\mathbf{r}, \mathbf{r}') \equiv \langle \mathbf{r}|\hat{\rho}|\mathbf{r}'\rangle = \frac{1}{Z \lambda^3} \exp \left(-\frac{\pi (\mathbf{r} - \mathbf{r}')^2}{\lambda^2} \right) \quad (13.3.8)$$

Calculating the partition function is also quite easy using the appropriate Gaussian integral:

$$Z = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{-\beta \hbar^2 \mathbf{k}^2 / 2m} = \frac{V}{\lambda^3} \quad (13.3.9)$$

giving:

$$\rho(\mathbf{r}, \mathbf{r}') = \frac{1}{V} \exp \left(-\frac{\pi (\mathbf{r} - \mathbf{r}')^2}{\lambda^2} \right) \quad (13.3.10)$$

Clearly, for N distinguishable particles we would have:

$$\rho(\{\mathbf{r}_i\}, \{\mathbf{r}'_i\}) = \frac{1}{V^N} \exp \left(-\sum_i \frac{\pi (\mathbf{r}_i - \mathbf{r}'_i)^2}{\lambda^2} \right) \quad (13.3.11)$$

13.4 Identical particles

The situation for identical particles gets much more complicated due to the (anti)symmetrization of the (fermionic) bosonic wavefunctions:

$$|k_1, k_2, \dots, k_N\rangle_A = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \text{sgn}(\sigma) P_\sigma |k_1, k_2, \dots, k_N\rangle \quad (13.4.1)$$

$$|k_1, k_2, \dots, k_N\rangle_S = \frac{1}{\sqrt{N! \prod_i (n_i!)}} \sum_{\sigma \in S_N} P_\sigma |k_1, k_2, \dots, k_N\rangle \quad (13.4.2)$$

where n_i are the number of particles in state $|k_i\rangle$. For simplicity let $|\psi\rangle_\eta$ denote the fermionic wavefunction if $\eta = -1$ and the bosonic wavefunction if $\eta = 1$, so that:

$$|k_1, k_2, \dots, k_N\rangle_\eta = \frac{1}{\sqrt{N! \prod_i (n_i!)}} \sum_{\sigma \in S_N} \eta^{\text{sgn}(\sigma)} P_\sigma |k_1, k_2, \dots, k_N\rangle \quad (13.4.3)$$

since $n_k = 0, 1$ for fermions. We then find that:

$$\rho_\eta(\{\mathbf{r}_i\}, \{\mathbf{r}'_i\}) = \frac{1}{Z} \sum_{\{\mathbf{k}_i\}} \exp\left(-\sum_i \frac{\beta \hbar^2 k_i^2}{2m}\right) \langle \mathbf{r}'_1, \dots, \mathbf{r}'_N | \mathbf{k}_1, \dots, \mathbf{k}_N \rangle_\eta \langle \mathbf{k}_1, \dots, \mathbf{k}_N | \mathbf{r}_1, \dots, \mathbf{r}_N \rangle_\eta \quad (13.4.4)$$

Note that the sum over $\{\mathbf{k}_i\}$ is not unrestricted, since different sequences of \mathbf{k}_i can still symmetrize to the same wave-function. It is easy to see that there are $\frac{N!}{\prod_i (n_i!)}$ wavefunctions that symmetrize to the same state. Hence:

$$\rho_\eta(\{\mathbf{r}_i\}, \{\mathbf{r}'_i\}) = \sum_{\{\mathbf{k}_i\}} \frac{\prod_i (n_i!)}{N! Z} \exp\left(-\sum_i \frac{\beta \hbar^2 k_i^2}{2m}\right) \langle \mathbf{r}'_1, \dots, \mathbf{r}'_N | \mathbf{k}_1, \dots, \mathbf{k}_N \rangle_\eta \langle \mathbf{k}_1, \dots, \mathbf{k}_N | \mathbf{r}_1, \dots, \mathbf{r}_N \rangle_\eta \quad (13.4.5)$$

where the sum is now unrestricted. We can now insert (13.4.4) into (13.4.5) to find that:

$$\rho_\eta(\{\mathbf{r}_i\}, \{\mathbf{r}'_i\}) = \sum_{\{\mathbf{k}_i\}} \sum_{\sigma, \sigma' \in S_N} \frac{1}{(N!)^2 Z} \exp\left(-\sum_i \frac{\beta \hbar^2 k_i^2}{2m}\right) \eta^{\text{sgn}(\sigma) + \text{sgn}(\sigma')} \langle \mathbf{r}'_1, \dots, \mathbf{r}'_N | P_{\sigma'} | \mathbf{k}_1, \dots, \mathbf{k}_N \rangle \langle \mathbf{k}_1, \dots, \mathbf{k}_N | P_\sigma | \mathbf{r}_1, \dots, \mathbf{r}_N \rangle \quad (13.4.6)$$

This daunting sum becomes an even more daunting integral as $L \rightarrow \infty$. We do it piece by piece. Firstly:

$$\langle \mathbf{r}_1, \dots, \mathbf{r}_N | P_\sigma | \mathbf{k}_1, \dots, \mathbf{k}_N \rangle = \prod_i e^{i\mathbf{k}_{\sigma(i)} \cdot \mathbf{r}_i} \quad (13.4.7)$$

so that:

$$= \sum_{\sigma, \sigma' \in S_N} \frac{1}{(N!)^2 Z} \eta^{\text{sgn}(\sigma) + \text{sgn}(\sigma')} \int \prod_i \frac{d^3 \mathbf{k}_i}{(2\pi)^3} \exp\left(-\frac{\beta \hbar^2 k_i^2}{2m}\right) e^{i\mathbf{r}'_i \cdot \mathbf{k}_{\sigma'(i)}} e^{-i\mathbf{r}_i \cdot \mathbf{k}_{\sigma(i)}} \quad (13.4.8)$$

$$= \sum_{\sigma, \sigma' \in S_N} \frac{1}{(N!)^2 Z} \eta^{\text{sgn}(\sigma) + \text{sgn}(\sigma')} \int \prod_i \frac{d^3 \mathbf{k}_i}{(2\pi)^3} \exp\left(-\frac{\beta \hbar^2 k_i^2}{2m}\right) e^{i\mathbf{k}_i \cdot (\mathbf{r}_{\sigma^{-1}(i)} - \mathbf{r}'_{\sigma'^{-1}(i)})} \quad (13.4.9)$$

where we rearrange the $e^{i\mathbf{r}'_i \cdot \mathbf{k}_{\sigma(i)}}$ terms so that the exponent contains \mathbf{k}_i . Now we again have:

$$= \sum_{\sigma, \sigma' \in S_N} \frac{1}{(N!)^2 Z} \eta^{\text{sgn}(\sigma) + \text{sgn}(\sigma')} \prod_i \frac{1}{\lambda^3} \exp \left[- \frac{\pi(\mathbf{r}_{\sigma^{-1}(i)} - \mathbf{r}'_{\sigma^{-1}(i)})^2}{\lambda^2} \right] \quad (13.4.10)$$

Finally, note that permuting all of $\sigma(i)$ will give back $\sigma(i)$ in some order. Thus we can write the sum over two permutations as the sum over just one permutation multiplied by $|S_N| = N!$ since the sum will be repeated every time for a different $\sigma' \in S_N$:

$$\rho_\eta(\{\mathbf{r}_i\}, \{\mathbf{r}'_i\}) = \frac{1}{Z_N \lambda^{3N} N!} \sum_{\sigma \in S_N} \eta^{\text{sgn}(\sigma)} \exp \left[- \sum_i \frac{\pi(\mathbf{r}_{(i)} - \mathbf{r}'_{\sigma(i)})^2}{\lambda^2} \right] \quad (13.4.11)$$

Imposing the normalization condition $\text{tr}(\rho) = \int d^{3N} \mathbf{r} \langle \mathbf{r} | \rho | \mathbf{r} \rangle = 1$ we find that

$$Z_N = \frac{1}{\lambda^{3N} N!} \int d^{3N} \mathbf{r} \sum_{\sigma \in S_N} \eta^{\text{sgn}(\sigma)} \exp \left[- \sum_i \frac{\pi(\mathbf{r}_{(i)} - \mathbf{r}'_{\sigma(i)})^2}{\lambda^2} \right] \quad (13.4.12)$$

so the partition function can be found by summing over all permutations of the $N \sim 10^{23}$ particles in our system. The lowest order term comes from the identity permutation, which yields the classical result $Z = \frac{V^N}{\lambda^{3N} N!}$. Note that the Gibbs factor $\frac{1}{h^{3N} N!}$ is already included, it arose naturally from the quantumness of our system without the need of any artificial thought experiments.

The next lowest order term comes from the exchange of just two particles, say particle 1 and 2. There are $\frac{N(N-1)}{2}$ such permutations all yielding the same correction so we find that

$$Z_N = \frac{1}{\lambda^{3N} N!} \int d^{3N} \mathbf{r} \left(1 + \eta \frac{N(N-1)}{2} \exp \left[- \frac{2\pi(\mathbf{r}_{(1)} - \mathbf{r}'_{\sigma(2)})^2}{\lambda^2} \right] + \dots \right) \quad (13.4.13)$$

$$= \frac{V^N}{N! \lambda^{3N}} \left(1 + \frac{N(N-1)}{2V} \eta \int d^3 \mathbf{r}_{12} e^{-2\pi r_{12}^2 / \lambda^2} + \dots \right) \quad (13.4.14)$$

where in the last step we changed to the center of mass relative coordinates $(\mathbf{R}_{12}, \mathbf{r}_{12}) = ((\mathbf{r}_1 + \mathbf{r}_2)/2, \mathbf{r}_1 - \mathbf{r}_2)$. The gaussian integral can be computed as usual and we find that

$$Z_N = Z_{\text{classical}} \left[1 + \frac{N(N-1)}{2V} \left(\frac{\lambda^2}{2} \right)^{3/2} \eta + \dots \right] \quad (13.4.15)$$

The free energy is then

$$F = -k_B T \ln Z_N \approx F_{\text{classical}} + \frac{N^2}{2V} \frac{\lambda^3}{2^{3/2}} \eta + \dots \quad (13.4.16)$$

where we taylor expanded $\ln(1+x)$ in the low density limit $\frac{N}{V} \ll 1$. This finally allows us

to compute the gas pressure as

$$\frac{p}{k_B T} = - \frac{1}{k_B T} \left. \frac{\partial F}{\partial V} \right|_T = \frac{N}{V} - \frac{N^2}{V^2} \frac{\lambda^3}{2^{5/2}} \eta + \dots \quad (13.4.17)$$

It is interesting to note that this expansion is a virial expansion (see (16.1.1)), and the quantum correction we have computed is the second virial coefficient

$$B_2 = - \frac{\eta \lambda^3}{2^{5/2}} \quad (13.4.18)$$

due to the classical potential

$$U(r) = -k_B T \ln(1 + \eta e^{-2\pi r^2/\lambda^2}) \approx -k_B T e^{-2\pi r^2/\lambda^2}, \quad r \gg \lambda \quad (13.4.19)$$

So we see that the effect of quantum statistics at high temperatures (where $\frac{\lambda}{r}$ is small) is that to add an exponential interaction between particles in a classical gas. This interaction is attractive for bosons and repulsive for fermions, just as one would expect!

13.5 Boson and fermion gases in the GCE

Consider a non-interacting quantum gas with single particle dispersion relation $\epsilon_{\mathbf{k}} = \epsilon(\mathbf{k})$. We have seen how to calculate partition function in the free particle case using the momentum basis. We can also use the occupation basis $\{n_{\mathbf{k}}\}$ where:

$$Z = \sum_{\{n_{\mathbf{k}}\}} e^{-\beta \sum_{\mathbf{k}} \epsilon(\mathbf{k}) n_{\mathbf{k}}} \quad (13.5.1)$$

However, since $\sum_{\mathbf{k}} n_{\mathbf{k}} = N$ this sum is restricted. We can relax this condition by moving to the grand canonical ensemble, this will allow us to sum over each $n_{\mathbf{k}}$ independently:

$$\mathcal{Z} = \sum_N e^{\beta \mu N} \sum_{\{n_{\mathbf{k}}\}_N} e^{-\beta \sum_{\mathbf{k}} \epsilon(\mathbf{k}) n_{\mathbf{k}}} \quad (13.5.2)$$

$$= \sum_{\{n_{\mathbf{k}}\}} \exp \left[-\beta \sum_{\mathbf{k}} (\epsilon(\mathbf{k}) - \mu) n_{\mathbf{k}} \right] \quad (13.5.3)$$

$$= \prod_{\mathbf{k}} \sum_{\{n_{\mathbf{k}}\}} \exp \left[-\beta (\epsilon(\mathbf{k}) - \mu) n_{\mathbf{k}} \right] \quad (13.5.4)$$

where $\{n_{\mathbf{k}}\}_N$ is the set of occupation numbers such that $\sum_{\mathbf{k}} n_{\mathbf{k}} = N$. Now for fermions $n_{\mathbf{k}} = 0, 1$ so we find:

$$\mathcal{Z}_- = \prod_{\mathbf{k}} (1 + \exp \left[-\beta (\epsilon(\mathbf{k}) - \mu) \right]) \quad (13.5.5)$$

For bosons instead the occupation numbers are unrestricted $n_{\mathbf{k}} = 0, 1, 2, \dots$ so we will obtain a geometric sum which converges provided that $\epsilon(\mathbf{k}) > \mu$. This must be true for all

\mathbf{k} , so seeing as the ground state energy is $\epsilon_{GS} = 0$ we get that if $\mu < 0$ then:

$$\mathcal{Z}_+ = \prod_{\mathbf{k}} \frac{1}{1 - \exp[-\beta(\epsilon(\mathbf{k}) - \mu)]} \quad (13.5.6)$$

The interpretation of $\mu < 0$ is that since bosons experience an attractive potential, it costs no energy to add a particle to the boson ensemble. The grand canonical potential is given by:

$$\Phi_G = \beta\eta \sum_{\mathbf{k}} \ln(1 - \eta e^{-\beta(\epsilon(\mathbf{k}) - \mu)}) \quad (13.5.7)$$

The average particle number is proportional to derivative of Φ_G with respect to μ :

$$\langle N \rangle_{\eta} = \frac{1}{\beta} \frac{\partial \Phi_G}{\partial \mu} = - \sum_{\mathbf{k}} \frac{e^{-\beta(\epsilon(\mathbf{k}) - \mu)}}{\eta e^{-\beta(\epsilon(\mathbf{k}) - \mu)} - 1} = \sum_{\mathbf{k}} \frac{1}{e^{\beta(\epsilon(\mathbf{k}) - \mu)} - \eta} \quad (13.5.8)$$

The precise importance of μ and the differences between $\eta = 1$ and $\eta = -1$ will be discussed in further detail in the next two chapters.

Bose-Einstein statistics

14.1 The density of states

Consider an ideal gas inside a box of volume $V = L^3$ on which we pose periodic boundary conditions. For a free gas the particles will be in eigenstates:

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}} \quad (14.1.1)$$

where the wave-vector $\mathbf{k} = (k_x, k_y, k_z)$ satisfies the quantization condition:

$$k_{x,y,z} = \frac{2\pi}{L} n_{x,y,z}, \quad n_{x,y,z} \in \mathbb{Z} \quad (14.1.2)$$

Each particle thus has energy:

$$E_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m} = \frac{4\pi^2 \hbar^2}{2mL^2} (n_x^2 + n_y^2 + n_z^2) = \frac{4\pi^2 \hbar^2}{2mL^2} n^2 \quad (14.1.3)$$

We calculated the partition function:

$$\mathcal{Z} = \sum_{\{n_i\}} e^{-\beta E_i} \quad (14.1.4)$$

in the previous section for indistinguishable particles. The sum can also be done by approximating it as an integral. Indeed, note that each term in the above sum is an exponential of (up to some proportion factor) $\lambda^2 n^2 / L^2$. For a sufficiently large box where $\lambda \gg L$ there will be a macroscopic number of states such that $E \lesssim k_B T$ which contribute to the sum. The sum can therefore be transformed into an integral:

$$\sum_n \rightarrow \int d^3 n = \frac{V}{(2\pi)^3} \int d^3 \mathbf{k} = \frac{2V}{(2\pi)^2} \int_0^\infty k^2 dk \quad (14.1.5)$$

where $k = |\mathbf{k}|$. We can use the dispersion relation

$$dE = \frac{\hbar^2 k}{m} dk \implies k^2 dk = \frac{m}{\hbar^2} \sqrt{\frac{2mE}{\hbar^2}} dE \quad (14.1.6)$$

to write (14.1.5) as an integral over energy:

$$\frac{2V}{(2\pi)^2} \int_0^\infty k^2 dk = \int \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E} dE \quad (14.1.7)$$

We define $g(E)dE$ as the number of states with energy in an interval $[E, E + dE]$ divided by dE . Then, we see that:

$$g(E) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E} \quad (14.1.8)$$

This quantity $g(E)$ is known as the density of states. It has a particularly simple form for free gases. For a gas of massless particles, we should instead use the dispersion relation:

$$E = \hbar ck \implies g(E) = \frac{V}{2\pi^2} \frac{E^2}{\hbar^3 c^3} \quad (14.1.9)$$

14.2 Photon gas - a statistical perspective

One might naively think that the calculation in (14.1.9) applies to photons as well. This is not true however, since photons can be polarized in two different states (\mathbf{E} parallel or \mathbf{E} perpendicular) so each \mathbf{k} state is two-fold degenerate. It follows that the density of states for photons is twice that for a normal massless relativistic particle:

$$g(E)dE = \frac{V}{\pi^2} \frac{E^2}{\hbar^3 c^3} dE \quad (14.2.1)$$

Introducing the frequency $\omega = E/\hbar$ we find that:

$$\boxed{g(\omega)d\omega = \frac{V}{\pi^2} \frac{\omega^2}{\hbar c^3} d\omega} \quad (14.2.2)$$

Note that $g(\omega)$ and $g(E)$ are not the same functions. One is the number of states per unit energy while the other is the number of states per unit frequency.

Suppose we have a gas of photons in thermal equilibrium inside a blackbody cavity of volume V at temperature T . Working in the canonical ensemble, the canonical partition for a gas of photons at frequency ω is:

$$Z_\omega = \sum_N e^{-\beta\hbar\omega N} = \frac{1}{1 - e^{-\beta\hbar\omega}} \quad (14.2.3)$$

Instead of summing over all frequencies, we can use the density of states to perform an integral:

$$\ln Z = \int_0^\infty \ln(Z_\omega) g(\omega) d\omega = -\frac{V}{c^3 \pi^2} \int_0^\infty \omega^2 \ln(1 - e^{-\beta\hbar\omega}) \quad (14.2.4)$$

Therefore the Free energy of a photon gas reads

$$F = -\frac{1}{\beta} \ln Z = \frac{V}{\beta c^3 \pi^2} \int_0^\infty \omega^2 \ln(1 - e^{-\beta\hbar\omega}) d\omega \quad (14.2.5)$$

while the internal energy is:

$$U = -\frac{\partial}{\partial \beta} \ln Z = \frac{V}{\pi^2 c^3} \int_0^\infty \frac{\hbar \omega^3 e^{-\beta \hbar \omega}}{1 - e^{-\beta \hbar \omega}} d\omega \quad (14.2.6)$$

$$= \frac{V \hbar}{\pi^2 c^3} \int_0^\infty \frac{\omega^3}{e^{\beta \hbar \omega} - 1} d\omega \quad (14.2.7)$$

The energy density carried by a photon with frequency in the interval $[\omega, \omega + d\omega]$ is known as the spectral energy density and follows the distribution:

$$u(\omega) d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta \hbar \omega} - 1} d\omega \quad (14.2.8)$$

This result is known as Planck's law. The maxima of the energy distribution for a given frequency ω occurs when $\frac{dU}{d\omega} = 0$:

$$3\omega^2(e^{\beta \hbar \omega} - 1) - \omega^3 \beta \hbar e^{\beta \hbar \omega} = 0 \implies \omega_{max} \approx \frac{2.822}{\hbar \beta} \quad (14.2.9)$$

which is known as Wien's displacement law.

Let's now evaluate the integral in (14.2.6) by considering the slightly different integral (which shall be used later):

$$g_n(z) = \frac{1}{\Gamma(n)} \int_0^\infty \frac{x^{n-1} e^{-x}}{z^{-1} - e^{-x}} d\omega \quad (14.2.10)$$

$$= \frac{z}{\Gamma(n)} \int_0^\infty x^{n-1} e^{-x} \sum_{m=0}^\infty z^m e^{-mx} dx \quad (14.2.11)$$

$$= \frac{z}{\Gamma(n)} \sum_{m=0}^\infty z^m \int_0^\infty x^{n-1} e^{-(m+1)x} dx \quad (14.2.12)$$

$$= \frac{1}{\Gamma(n)} \sum_{m=1}^\infty z^m \int_0^\infty x^{n-1} e^{-mx} dx \quad (14.2.13)$$

$$= \frac{1}{\Gamma(n)} \sum_{m=1}^\infty \frac{z^m}{m^n} \int_0^\infty u^{n-1} e^{-u} du \quad (14.2.14)$$

Identifying the Gamma function $\Gamma(n) = \int_0^\infty u^{n-1} e^{-u} du$ we see that $g_n(z)$, known as Bose-Einstein functions, take the form:

$$g_n(z) = \sum_{m=1}^\infty \frac{z^m}{m^n} \implies g_n(1) = \zeta(n) \quad (14.2.15)$$

Consequently:

$$\int_0^\infty \frac{\omega^3}{e^{\beta \hbar \omega} - 1} d\omega = \frac{1}{\beta^3 \hbar^3} \int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{1}{\beta^4 \hbar^4} \zeta(4) \Gamma(4) = \frac{1}{\beta^4 \hbar^4} \frac{\pi^4}{15} \quad (14.2.16)$$

giving an energy density $u = \frac{U}{V}$ of:

$$u = \frac{\pi^2 k_B^4}{15 \hbar^3 c^3} T^4 \quad (14.2.17)$$

The energy per unit area per unit time (radiant emittance) j emitted by an object at temperature T is then given by:

$$j = \frac{uc}{4} = \sigma T^4, \quad \sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2} \quad (14.2.18)$$

To see where $\frac{c}{4}$ comes from, note that the radiant emittance is the integral of $\frac{U}{At} = uc \cos \theta$ where θ is the angle of the emitted photon to the normal. Therefore:

$$j = \int_0^2 \pi \int_0^\pi uc \cos \theta d\theta d\phi = \frac{uc}{4} \quad (14.2.19)$$

as found earlier. Also, the integral in (14.2.5) can be calculated exactly to give an expression for the free energy:

$$F = \frac{V}{\beta c^3 \pi^2} \int_0^\infty \omega^2 \ln(1 - e^{-\beta \hbar \omega}) = -\frac{V}{3\beta c^3 \pi^2} \int_0^\infty \frac{\beta \hbar \omega^3 e^{-\beta \hbar \omega}}{1 - e^{-\beta \hbar \omega}} \quad (14.2.20)$$

$$= -\frac{V \hbar}{3c^3 \pi^2} \frac{1}{(\beta \hbar)^4} \int_0^\infty \frac{x^3}{e^x - 1} dx = -\frac{V \pi^2 k_B^4}{45 \hbar^3 c^3} T^4 \quad (14.2.21)$$

or more succinctly:

$$F = -\frac{4\sigma V}{3c} T^4 \quad (14.2.22)$$

Therefore, the radiation pressure will satisfy the equation of state for a photon gas:

$$p = -\left. \frac{\partial F}{\partial V} \right|_T = \frac{4\sigma}{3c} T^4 \implies \boxed{3pc = 4\sigma T^4} \quad (14.2.23)$$

14.3 Photon gas - a kinetic perspective

Some of the important results in the previous section could have been motivated using kinetic gas theory as well. We consider a gas of photons in thermal equilibrium within a blackbody cavity of volume V and temperature T . For a normal gas with N particles moving randomly, the number of particles per unit volume moving out within a solid angle $d\Omega$ with speed in the interval $[v, v + dv]$ would be given by:

$$\frac{d\Omega}{4\pi} n(v) dv = \frac{1}{4\pi} n(v) \sin \theta dv d\theta d\phi \quad (14.3.1)$$

For photons however $n(v) = \delta(v - c)$ so integrating over all possible speeds and over $\phi \in [0, 2\pi]$ we find that:

$$\frac{1}{2} n \sin \theta d\theta, \quad n = \frac{N}{V} \quad (14.3.2)$$

particles move in the direction $[\theta, \theta + d\theta]$ per unit volume. For a photon in this cavity to hit a wall (oriented so that its normal is along z for simplicity) in a time Δt it must be within a distance $c \cos \theta dt$. Consequently the volume the gas particles must be inside of to collide with the wall is $Ac \cos \theta dt$. The number of gas particles within this volume is:

$$\frac{1}{2}nAc \cos \theta \sin \theta d\theta = \frac{1}{4}nAc \sin(2\theta)d\theta dt \quad (14.3.3)$$

the number of particles hitting the wall per unit area per unit time, or the incident particle flux per unit time is then:

$$\Phi = \int_0^{\pi/2} \frac{1}{4}nc \sin(2\theta)d\theta = \frac{1}{4}nc \quad (14.3.4)$$

where we integrated over $\theta \in [0, \pi/2]$ since particles with $\theta \in [\pi/2, \pi]$ are moving away from the wall.

Each particle has momentum $\frac{E}{c}$ so the momentum density is given by $\frac{u}{c}$. The momentum perpendicular to the wall carried by the gas particles incident on the wall per unit area per unit time is:

$$\int_0^{\pi/2} \frac{1}{4} \frac{u}{c} \cos \theta c \sin(2\theta)d\theta = \frac{1}{6}u \quad (14.3.5)$$

The pressure on the wall is twice this since the particles are also reflected back:

$$p = \frac{1}{3}u \quad (14.3.6)$$

as found earlier. Also, considering that each photon has energy $\hbar\omega$, the radiant emittance (energy per unit time per unit area incident) on the wall is:

$$j = \hbar\omega\Phi = \frac{1}{4}n\hbar\omega c \quad (14.3.7)$$

We may define $u = \langle n\hbar\omega \rangle$ as the energy density to write:

$$j = \frac{1}{4}uc \quad (14.3.8)$$

Finally, using the first law:

$$u = \left. \frac{\partial U}{\partial V} \right|_T = T \left. \frac{\partial S}{\partial V} \right|_T - p = T \left. \frac{\partial p}{\partial T} \right|_V - p = \frac{1}{3}T \left. \frac{\partial u}{\partial T} \right|_V - \frac{u}{3} \quad (14.3.9)$$

which gives after some simplification:

$$4u = T \left. \frac{\partial u}{\partial T} \right|_V \quad (14.3.10)$$

This PDE has a solution $u = AT^4$ for some constant A , yielding:

$$j = \frac{1}{4}AcT^4 \quad (14.3.11)$$

which is the Stefan-Boltzmann law. It is important to note that these calculations do not allow us to calculate the pressure and radiant emittance since we are missing u , the energy density. Only a quantum statistical calculation would have allowed us to calculate this quantity.

14.4 Blackbodies and Kirchoff's law

In this section we try to understand how to quantify the aptitude of a surface in absorbing and emitting electromagnetic radiation at particular frequencies and temperatures. With this in mind, we define the spectral absorptivity $\alpha(\omega)$ as the fraction of the incident radiation absorbed at frequency ω and the spectral emissive power $j^*(\omega)$ for a surface so that $j^*(\omega)d\omega$ gives the power emitted per unit area with frequency in $[\omega, \omega + d\omega]$. In thermal equilibrium the radiation emitted by a surface must be equal to the radiation absorbed. Using our previous definitions we find that:

$$\text{emitted radiation} = j^*(\omega)d\omega \quad (14.4.1)$$

$$\text{absorbed radiation} = \alpha(\omega)\frac{1}{4}cu(\omega)d\omega \quad (14.4.2)$$

so:

$$j^*(\omega)d\omega = \alpha(\omega)\frac{1}{4}cu(\omega)d\omega \implies \boxed{\frac{j^*(\omega)}{\alpha(\omega)} = \frac{1}{4}cu(\omega) = j(\omega) > 0} \quad (14.4.3)$$

This result is known as Kirchoff's law. If we fix ω then the RHS will be a constant, and thus $\alpha \propto j^*$, if a body is a good absorber then it will also be a good emitter. A perfect absorber of radiation, known as a blackbody, will have $\alpha(\omega) = 1$ for all frequencies ω . Thus Kirchoff's law implies that the spectral emissive power $j^*(\omega)$ will be equal to its maximum possible value j , the radiant emittance (all of the absorbed radiation is emitted) which we calculated earlier. If we thus place a gas of photons inside a cavity whose walls are blackbodies, then through perfect absorption and emission of photons the radiation contained within such a black-body is known as black-body radiation. This radiation follows the Planck distribution in (14.2.8) no matter what the absorbed radiation looked like. All that determines the distribution the radiation emitted by a blackbody is the temperature at which it is in equilibrium. It follows that the process of blackbody radiation is not reversible in general.

14.5 Phonons

14.6 High temperature Bose gases

14.7 Bose-Einstein condensation

After having looked at two important examples of Bose-Einstein statistics, let's now give a complete statistical description of the ideal Bose gas at low temperatures. We assume the thermal fluctuations are larger than the energy gaps in the gas, so that we may write the

average number of particles as (we omit the $\langle \rangle$ unless otherwise stated):

$$N = \int \frac{g(E)}{e^{\beta(E-\mu)} - 1} dE = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int \frac{E^{1/2}}{e^{\beta(E-\mu)} - 1} dE \quad (14.7.1)$$

Traditionally, we let $z = e^{\beta\mu}$ be the fugacity. Substituting $x = \beta E$ we find that:

$$N = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \frac{1}{\beta^{3/2}} \int_0^\infty \frac{x^{1/2}}{z^{-1}e^x - 1} dx \quad (14.7.2)$$

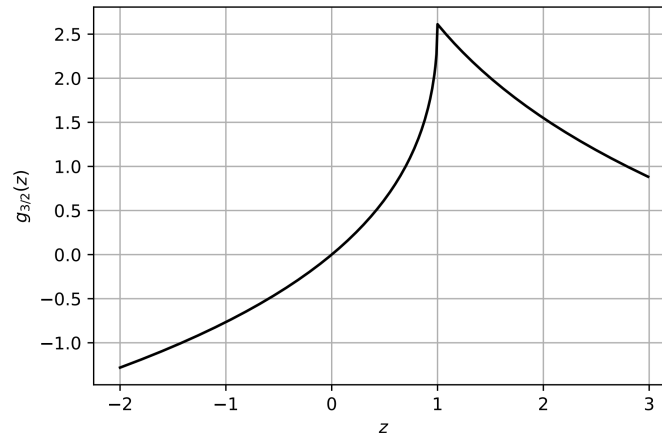
$$= \frac{V}{4\pi^2} \left(\frac{2mk_B T}{\hbar^2} \right)^{3/2} \Gamma(3/2) \sum_{m=1}^\infty \frac{z^m}{m^{3/2}} \quad (14.7.3)$$

$$\Rightarrow n = \frac{1}{\lambda_{th}^3} \sum_{m=1}^\infty \frac{z^m}{m^{3/2}} \equiv \frac{g_{3/2}(z)}{\lambda_{th}^3} \quad (14.7.4)$$

where $\Gamma(3/2) = \frac{\sqrt{\pi}}{2}$ and $n = \frac{N}{V}$ is the particle density. Suppose we slowly reduce the temperature in (14.7.4) keeping n constant. Since $\mu < 0$ the fugacity $z = e^{\beta\mu}$ will also increase¹, but it cannot be larger than 1 since $z \in [0, 1]$ for negative chemical potentials. The critical temperature T_c at which $z = 1$ can be found from (14.7.4):

$$\frac{\zeta(3/2)}{\lambda_{th}^3} = n \Rightarrow T_c = \frac{2\pi\hbar^2}{mk_B} \left(\frac{n}{\zeta(3/2)} \right)^{2/3} \quad (14.7.5)$$

What would happen if we decreased the temperature further? Looking at the plot of $g_{3/2}(z)$:



it seems like $g_{3/2}(z)$ and thus $\frac{N}{V}$ is bounded above by $\zeta(3/2) \approx 2.613$ and $\frac{\zeta(3/2)}{\lambda_{th}^3}$ respectively. So as we continue to decrease the temperature, while z remains fixed at 1 and so $g_{3/2}(z) =$

¹indeed note that the $z \ll 1$ limit is a high temperature limit. To see why, note that as we decrease $T \rightarrow 0$ then we do have that $\beta \rightarrow \infty$, but the chemical potential $\mu(T, N)$ will also change if we keep N fixed. It turns out that $\mu \rightarrow -\infty$ more rapidly than $\beta \rightarrow 0$, since to leading order we need $\frac{N}{V} \sim \frac{z}{\lambda^3}$ to be constant, and thus $z \sim T^{-3/2}$.

$\zeta(3/2)$, λ_{th} will instead decrease. This is not allowed, since the particle number is fixed, so it seems like it is impossible to cool our system below T_c ²

Another way to look at this paradox is by keeping the temperature fixed and increasing the particle density. Again we see that increasing n leads to an compensating increase in z and thus μ . There comes a point however where $\mu = 0 \iff z = 1$, which yields the following critical density

$$n_c = \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} \zeta(3/2) \quad (14.7.6)$$

We cannot further increase n since this would make $\mu > 0$, but this makes no sense!

It turns out that at such low temperatures (or high densities) we have a violation of our initial assumption that a sum over microstates may be approximated by an integral. Indeed the integral in (13.4.5) does not receive a contribution by the ground state $E_{GS} = 0$. Using the Bose-Einstein distribution we see that the number of particles in the ground state is:

$$N_0 = \frac{1}{z^{-1} - 1} \implies N = \frac{V g_{3/2}(z)}{\lambda_{th}^3} + \frac{z}{1 - z} \quad (14.7.7)$$

This resolves our paradox, since as z approaches 1 we see that more and more particles occupy the ground state until N_0 becomes macroscopic when $z = 1$. If we add more particles they will just go to the ground state so the number of excited particles $N_e = \frac{V g_{3/2}(z)}{\lambda_{th}^3}$ will remain constant, making z level off at $z = 1$:

$$z = \begin{cases} 1, & \frac{V}{\lambda_{th}^3} \zeta(3/2) \leq N \\ g_{3/2}^{-1}\left(\frac{N \lambda_{th}^3}{V}\right), & \frac{V}{\lambda_{th}^3} \zeta(3/2) > N \end{cases} \quad (14.7.8)$$

Now as we decrease the temperature below the critical temperature, we have that the ratio of particles occupying the ground state will be:

$$\frac{N_0}{N} = \begin{cases} 1 - \left(\frac{T}{T_c}\right)^{3/2}, & T < T_c \\ 0, & T > T_c \end{cases} \quad (14.7.9)$$

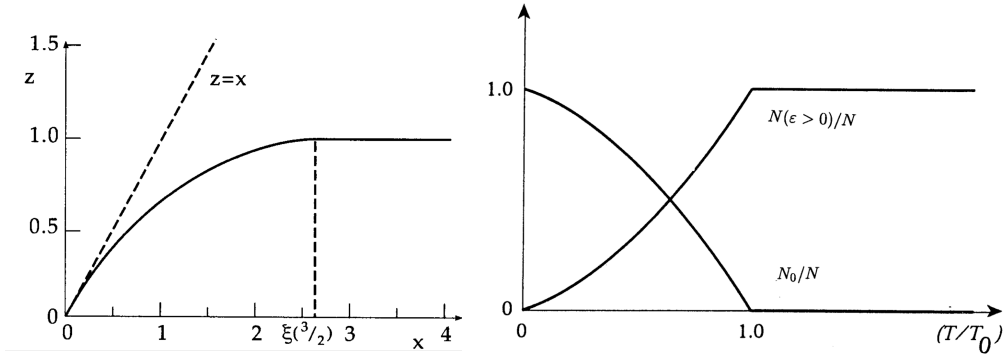
This new state of matter where a macroscopic number of particles occupy the ground state is known as a Bose-Einstein condensate.

Let's also look at the pressure:

$$pV = -\Phi_G = -\frac{1}{\beta} \ln \mathcal{Z} \quad (14.7.10)$$

This time, we take into account the contribution due to the ground state by explicitly adding $\ln(1 - z)$ (note that we can do this since \mathcal{Z} is multiplicative and therefore $\ln(\mathcal{Z}) =$

²even if μ , and thus z were to change below T_c , there is no way $g_{3/2}(z)$ can increase to compensate the increase in λ_{th}^3 since it is bounded above by $\zeta(3/2)$



$\ln(\mathcal{Z}_{GS}) + \ln(\mathcal{Z}_e)$:

$$\ln \mathcal{Z} = -\frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty E^{1/2} \ln(1 - ze^{-\beta E}) dE - \ln(1 - z) \quad (14.7.11)$$

$$= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \frac{2}{3} \beta \int_0^\infty \frac{E^{3/2}}{z^{-1}e^{\beta E} - 1} dE - \ln(1 - z) \quad (14.7.12)$$

$$= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \frac{1}{3} \frac{1}{\beta^{3/2}} \int_0^\infty \frac{x^{3/2}}{z^{-1}e^x - 1} dx - \ln(1 - z) \quad (14.7.13)$$

$$= \frac{V}{\lambda_{th}^3} g_{5/2}(z) = \frac{V}{\lambda_{th}^3} \sum_{m=1}^\infty \frac{z^m}{m^{5/2}} - \ln(1 - z) \quad (14.7.14)$$

Thus

$$p = \frac{k_B T}{\lambda_{th}^3} g_{5/2}(z) - \frac{\ln(1 - z)}{V} \quad (14.7.15)$$

Now note that if $T > T_c$ then $\ln(1 - z)$ is finite and thus the second term is negligible. If instead $T < T_c$ then $z = \frac{1}{N_{gs}+1}$ so we get $\ln(N_{gs})/V$. The numerator diverges more slowly than the denominator so the second term will again vanish. Consequently we find that

$$p = \frac{k_B T}{\lambda_{th}^3} g_{5/2}(z) \quad (14.7.16)$$

regardless of whether we are below or above the critical temperature. The fact that the ground state gives no pressure contribution makes sense since ground state particles have no net momentum ($E_{gs} = 0$) and thus cannot possibly contribute to the pressure. Furthermore, in the condensate regime $T < T_c$ we have that $z = 1$ and thus:

$$p = \frac{k_B T}{\lambda_{th}^3} \zeta(5/2) \quad (14.7.17)$$

Note that this value is independent of the number of particles. This makes sense, since in a BEC any particles that are added will be in the ground state, and such particles will not contribute to the total pressure.

Similarly, the internal energy reads (again the ground state gives a vanishing contribution

to the internal energy in the thermodynamic limit):

$$U = -\frac{\partial}{\partial\beta} \ln Z \Big|_z = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty E^{1/2} \frac{\partial}{\partial\beta} \ln(1 - ze^{-\beta E}) dE \quad (14.7.18)$$

$$= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{E^{3/2}}{z^{-1}e^{\beta E} - 1} dE \quad (14.7.19)$$

$$= \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \frac{1}{\beta^{5/2}} \Gamma(5/2) \int_0^\infty \frac{x^{3/2}}{z^{-1}e^x - 1} dx = \frac{3}{2} \frac{V k_B T}{\lambda_{th}^3} g_{5/2}(z) = \frac{3}{2} pV \quad (14.7.20)$$

The heat capacity is then given by:

$$C_V = \frac{\partial U}{\partial T} \Big|_V = \frac{15}{4} \frac{k_B V}{\lambda_{th}^3} g_{5/2}(z) + \frac{3}{2} \frac{V k_B T}{\lambda_{th}^3} \frac{\partial g_{5/2}(z)}{\partial T} \quad (14.7.21)$$

For $T < T_c$ we have that $z = 1$ is constant, and thus the second term will vanish:

$$c_V = \frac{15}{4} \frac{k_B}{\lambda_{th}^3} \zeta(5/2), \quad T < T_c \quad (14.7.22)$$

Importantly, at the critical temperature:

$$C_V = \frac{15}{4} \frac{\zeta(5/2)}{\zeta(3/2)} N k_B, \quad T = T_c \quad (14.7.23)$$

For $T > T_c$ the second term in (14.7.21) becomes relevant. Since $\frac{V}{\lambda_{th}^3} = \frac{N}{g_{3/2}(z)}$ we have that:

$$U = \frac{3}{2} N k_B T \frac{g_{5/2}(z)}{g_{3/2}(z)} \quad (14.7.24)$$

$$\implies C_V = \frac{3}{2} N k_B \frac{g_{5/2}(z)}{g_{3/2}(z)} + \frac{3}{2} N k_B T \frac{\partial}{\partial T} \left(\frac{g_{5/2}(z)}{g_{3/2}(z)} \right) \quad (14.7.25)$$

$$= \frac{3}{2} N k_B \frac{g_{5/2}(z)}{g_{3/2}(z)} + \frac{3}{2} N k_B T \frac{1}{z} \left(1 - \frac{g_{5/2}(z) g_{1/2}(z)}{g_{3/2}^2(z)} \right) \frac{\partial z}{\partial T} \quad (14.7.26)$$

where we used the fact that $\frac{\partial}{\partial z}(g_n(z)) = \frac{1}{z} g_{n-1}(z)$. We can write $\frac{\partial z}{\partial T}$ in a useful form by differentiating (14.7.4) which applies for $T > T_c$ only:

$$\frac{1}{z} g_{1/2}(z) \frac{\partial z}{\partial T} = -\frac{3}{2} \frac{N}{V} \frac{\lambda_{th}^3}{T} \implies \frac{\partial z}{\partial T} = -\frac{3z}{2T} \frac{g_{3/2}(z)}{g_{1/2}(z)} \quad (14.7.27)$$

This allows us to write:

$$C_V = \frac{3Nk_B}{4} \left(5 \frac{g_{5/2}(z)}{g_{3/2}(z)} - 3 \frac{g_{3/2}(z)}{g_{1/2}(z)} \right), \quad T > T_c \quad (14.7.28)$$

Firstly, note that as $T \rightarrow \infty$ then $z \rightarrow 0$ and since $\lim_{x \rightarrow 0} \frac{g_m(x)}{g_n(x)} = 1$ for all positive integers m, n we have that $C_V \rightarrow \frac{3}{2} N k_B$, which is the correct value for a classic ideal gas. A high

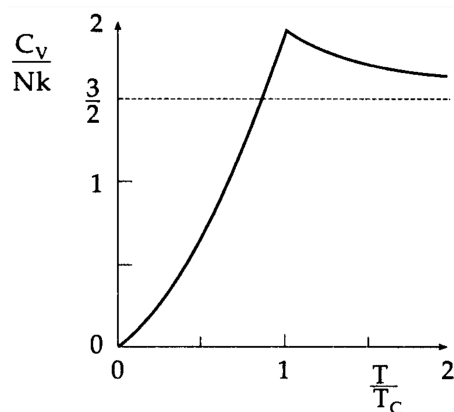
temperature Bose gas behaves as a high temperature classical gas. Secondly, note that as $T \rightarrow T_c$ from above, $z \rightarrow 1$ and since $\zeta(1/2)$ diverges the second term vanishes, giving $C_V \rightarrow \frac{15Nk_B}{4} \frac{\zeta(5/2)}{\zeta(3/2)}$ which is the same value as in the condensate phase. We provide a plot for C_V beside (taken from Greiner).

We see that in the Bose condensation regime $T < T_c$ the specific heat capacity grows as $T^{3/2}$. Instead for $T > T_c$ it decays into the classical ideal gas heat capacity. At $T = T_c$ we see a kink, and while a C_V is continuous its first derivative is not. Since the second derivative of the internal energy is discontinuous, we call this a second order phase transition. In general second order phase transitions occur when the second derivative of a thermodynamic potential suffers from a discontinuity.

It may appear like the only reason we obtained a discontinuity is that we ignored the $\frac{\partial z}{\partial T}$ term in the $T < T_c$ regime:

$$z = \left(1 + \frac{1}{N_0}\right)^{-1} = \left[1 + \frac{1}{N} \left(1 - \frac{T}{T_c}\right)^{-3/2}\right]^{-1} \implies \frac{\partial z}{\partial T} \propto N^{-1} \quad (14.7.29)$$

For finite N this is a nonzero contribution, and indeed the kink would be smooth at some scale. There would be a smooth transition between the two temperature regimes. However, note that as $N \rightarrow \infty$ this contribution vanishes, and we do effectively get a discontinuity. This is true for any phase transition, as we shall soon see.



Fermi-Dirac statistics

15.1 The Ideal Fermi gas

Consider an ideal gas of free, non-interacting fermions with spin s . The number of states will have a degeneracy factor of $g_s = 2s + 1$ (e.g. for spin half each state can be occupied by a spin up and a spin down fermion, giving two-fold degeneracy). The density of states then reads:

$$g(E) = \frac{g_s V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} \quad (15.1.1)$$

Recall that according to Fermi-Dirac statistics the average number of fermions is given by:

$$N = \sum_{\mathbf{k}} \frac{1}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} + 1} \quad (15.1.2)$$

In the high temperature limit the thermal fluctuations are much larger than the energy gap in our gas, so we may approximate this sum by an integral:

$$N = \int \frac{g(E) dE}{z^{-1} e^{\beta\epsilon_{\mathbf{k}}} + 1} \quad (15.1.3)$$

Similarly for the internal energy:

$$U = \sum_{\mathbf{k}} \frac{\epsilon_{\mathbf{k}}}{e^{\beta(\epsilon_{\mathbf{k}} - \mu)} + 1} = \int \frac{g(E) E dE}{z^{-1} e^{\beta\epsilon_{\mathbf{k}}} + 1} \quad (15.1.4)$$

and the grand canonical potential:

$$pV = -\Phi_G = \frac{1}{\beta} \sum_{\mathbf{k}} \ln(1 + e^{\beta(\epsilon_{\mathbf{k}} - \mu)}) = \int g(E) \ln(1 + z e^{-\beta E}) dE \quad (15.1.5)$$

We again encounter a family of integrals $\int_0^\infty \frac{x^{n-1}}{z^{-1}e^x+1} dx$ which are slightly different to $g_n(z)$. They can however be calculated in exactly the same way:

$$f_n(z) = \frac{1}{\Gamma(n)} \int_0^\infty \frac{x^{n-1} e^{-x}}{z^{-1} + e^{-x}} = \frac{1}{\Gamma(n)} \int_0^\infty z e^{-x} x^{n-1} \sum_{m=0}^\infty (-z e^{-x})^m \quad (15.1.6)$$

$$= -\frac{1}{\Gamma(n)} \sum_{m=0}^\infty (-z)^{m+1} \int_0^\infty x^{n-1} e^{-(m+1)x} dx \quad (15.1.7)$$

$$= -\frac{1}{\Gamma(n)} \sum_{m=1}^\infty (-z)^m \int_0^\infty x^{n-1} e^{-mx} dx \quad (15.1.8)$$

$$= -\frac{1}{\Gamma(n)} \sum_{m=1}^\infty \frac{(-z)^m}{m^n} \int_0^\infty s^{n-1} e^{-s} ds = -\sum_{m=1}^\infty \frac{(-z)^m}{m^n} = -g_n(-z) \quad (15.1.9)$$

Consequently the average number of bosons can be expressed as:

$$N = \frac{g_s V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty \frac{E^{1/2}}{z^{-1} e^{\beta E} + 1} dE \quad (15.1.10)$$

$$= \frac{g_s V}{4\pi^2} \left(\frac{2m}{\hbar^2 \beta} \right)^{3/2} \int_0^\infty \frac{x^{1/2}}{z^{-1} e^x + 1} dx \quad (15.1.11)$$

$$= \frac{g_s V}{4\pi^2} \left(\frac{2m k_B T}{\hbar^2} \right)^{3/2} \Gamma(3/2) f_{3/2}(z) \quad (15.1.12)$$

$$= \frac{g_s V}{\lambda_{th}^3} f_{3/2}(z) \quad (15.1.13)$$

In the limit where $z \ll 1$ this can be expanded to second order as:

$$N \approx \frac{g_s V}{\lambda_{th}^3} z \left(1 - \frac{z}{2\sqrt{2}} \right) \quad (15.1.14)$$

We see that $z \ll 1$ is only satisfied if $\lambda_{th}^3 \ll \frac{V}{N}$, so if the thermal wavelength is much smaller than the particle separation. Recalling that the thermal wavelength represents the average de Broglie wavelength of the fermions, this means that we are working in the quasi-classical, high temperature limit. To understand why $z \ll 1$ corresponds to this high temperature limit, note that at fixed particle density $\frac{N}{V}$ we need $\frac{z}{\lambda_{th}^3}$ to be roughly constant, implying that the fugacity scales as $T^{-3/2}$. Indeed one can solve (15.1.14) to get:

$$z \approx \sqrt{2} \left(1 \pm \sqrt{1 - \sqrt{2} \frac{\lambda_{th}^3 N}{g_s V}} \right) \quad (15.1.15)$$

$$\approx \sqrt{2} \left[1 \pm \left(1 - \sqrt{2} \frac{\lambda_{th}^3 N}{g_s V} - \frac{1}{8} \left(\frac{\lambda_{th}^3 N}{g_s V} \right)^2 \right) \right] \quad (15.1.16)$$

$$\approx \frac{\lambda_{th}^3 N}{g_s V} \left(1 + \frac{1}{2\sqrt{2}} \frac{\lambda_{th}^3 N}{g_s V} \right) \quad (15.1.17)$$

where the + solution is discarded since $z \ll 1$. One can similarly find that for the internal energy:

$$U = \frac{3g_s V}{2\beta\lambda_{th}^3} f_{5/2}(z) \approx \frac{3g_s V}{2\beta\lambda_{th}^3} z \left(1 - \frac{z}{4\sqrt{2}}\right) \quad (15.1.18)$$

Inserting (15.1.15) into the above and simplifying, keeping only terms up to second order in $\frac{\lambda_{th}^3 N}{g_s V}$:

$$U \approx \frac{3g_s V}{2\beta\lambda_{th}^3} \frac{\lambda_{th}^3 N}{g_s V} \left(1 + \frac{1}{2\sqrt{2}} \frac{\lambda_{th}^3 N}{g_s V}\right) \left(1 - \frac{1}{4\sqrt{2}} \frac{\lambda_{th}^3 N}{g_s V}\right) \quad (15.1.19)$$

$$\approx \frac{3N}{2\beta} \left(1 + \frac{1}{4\sqrt{2}} \frac{\lambda_{th}^3 N}{g_s V}\right) \quad (15.1.20)$$

For pV , it is convenient to integrate by parts. Letting $G(E) = \int g(E)dE = \frac{2}{3}g(E)E$:

$$pV = \frac{1}{\beta} \left[G(E) \ln(1 + ze^{-\beta E}) \right]_0^\infty + \int_0^\infty \frac{G(E)}{z^{-1}e^{\beta E} + 1} dE \quad (15.1.21)$$

$$= \int_0^\infty \frac{2}{3} \frac{g(E)E}{z^{-1}e^{\beta E} + 1} dE = \frac{2}{3}U \quad (15.1.22)$$

Using our first order approximation for U we get the equation of state:

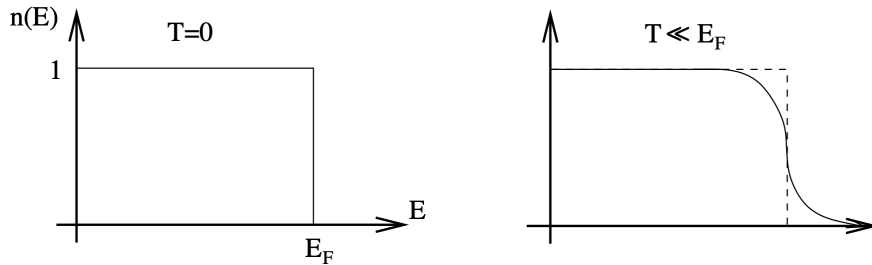
$$pV \approx Nk_B T \left(1 + \frac{1}{4\sqrt{2}} \frac{\lambda_{th}^3 N}{g_s V}\right) \quad (15.1.23)$$

15.2 Zero temperature behaviour - Fermi energy

Let's now look at the extremely low temperature limit $T \rightarrow 0$. It is easy to see that:

$$N_{\mathbf{k}} = \begin{cases} 1, & \text{if } \epsilon_{\mathbf{k}} < \mu \\ 0, & \text{if } \epsilon_{\mathbf{k}} > \mu \end{cases} \quad (15.2.1)$$

This intuitively makes sense. At absolute zero the lowest energy states are favoured, so particles will occupy the energy levels from the ground state up.



The highest occupied energy level E_F is known as the **Fermi energy**, and must be equal

to the zero temperature chemical potential:

$$E_F = \mu(T = 0) \quad (15.2.2)$$

Note that in the GCE the chemical potential and temperature are independent natural variables. However, since we are keeping N constant, if we vary T we must also vary μ giving this temperature dependence of the chemical potential.

For an ideal fermion gas, we can look at the momentum space (k_x, k_y, k_z) . Due to the isotropicity of $\epsilon_{\mathbf{k}} = \frac{\hbar^2 \mathbf{k}^2}{2m}$, when we add particles to the gas they will fill up forming a sphere, a **Fermi sphere** or **Fermi sea**. The radius of that sphere gives the **Fermi momentum** k_F which is related to the Fermi energy by $k_F^2 = \frac{2mE_F}{\hbar^2}$. The states with momentum k_F form the **Fermi surface**.

Let's calculate E_F . Suppose we have $\langle N \rangle = N$ bosons. Then we need:

$$N = \int_0^{E_F} g(E) dE = \frac{g_s V}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{E_F} E^{1/2} dE \quad (15.2.3)$$

$$= \frac{g_s V}{6\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E_F^{3/2} \quad (15.2.4)$$

$$\Rightarrow E_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2 N}{g_s V} \right)^{2/3} \quad (15.2.5)$$

giving the Fermi momentum:

$$k_F = \left(\frac{6\pi^2 N}{g_s V} \right)^{1/3} \quad (15.2.6)$$

Also:

$$U = \int_0^{E_F} g(E) E dE = \frac{g_s V}{10\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E_F^{5/2} = \frac{3}{5} N E_F \quad (15.2.7)$$

$$pV = \frac{2}{3} U = \frac{2}{5} N E_F \quad (15.2.8)$$

so at zero temperature the pressure does not vanish. This is a direct result of the Pauli exclusion principle "repelling" electrons and thus producing an outwards degeneracy pressure.

15.3 The Sommerfield approximation

We have looked at the Fermi-Dirac distribution in the high temperature and absolute zero temperature limits. We are now ready to study the low-temperature limit, and we do so in the Sommerfield approximation.

The first step is to split Fermi-Dirac functions $f_n(z)$ into two integrals:

$$\Gamma(n)f_n(z) = \int_0^{\beta\mu} x^{n-1} \left(1 - \frac{1}{1 + ze^{-x}}\right) dx + \int_{\beta\mu}^{\infty} \frac{x^{n-1}}{z^{-1}e^x + 1} dx \quad (15.3.1)$$

$$= \frac{(\beta\mu)^n}{n} - \int_0^{\beta\mu} \frac{x^{n-1}}{1 + ze^{-x}} dx + \int_{\beta\mu}^{\infty} \frac{x^{n-1}}{z^{-1}e^x + 1} dx \quad (15.3.2)$$

In the two integrals we have two exponentials, $e^{-x+\beta\mu}$ and $e^{x-\beta\mu}$ so it makes sense to use the variables $\eta_1 = -x + \beta\mu$ and $\eta_2 = x - \beta\mu$:

$$\Gamma(n)f_n(z) = \frac{(\beta\mu)^n}{n} - \int_0^{\beta\mu} \frac{(\beta\mu - \eta_1)^{n-1}}{1 + e^{\eta_1}} d\eta_1 + \int_0^{\infty} \frac{(\beta\mu + \eta_2)^{n-1}}{e^{\eta_2} + 1} d\eta_2 \quad (15.3.3)$$

Let's look at the first integral, and expand it as:

$$\int_0^{\beta\mu} \frac{(\beta\mu - \eta_1)^{n-1}}{1 + e^{\eta_1}} d\eta_1 = \int_0^{\infty} \frac{(\beta\mu - \eta_1)^{n-1}}{1 + e^{\eta_1}} d\eta_1 - \int_{\beta\mu}^{\infty} \frac{(\beta\mu - \eta_1)^{n-1}}{1 + e^{\eta_1}} d\eta_1 \quad (15.3.4)$$

In the low temperature limit the integral from 0 to ∞ is small (of order z^{-1}) so it can be safely neglected. Therefore, we get that:

$$\Gamma(n)f_n(z) = \frac{(\beta\mu)^n}{n} + \int_0^{\infty} \frac{(\beta\mu + \eta)^{n-1} - (\beta\mu - \eta)^{n-1}}{e^{\eta} + 1} d\eta \quad (15.3.5)$$

The numerator can be Taylor expanded in $\frac{\eta}{\beta\mu}$:

$$(\beta\mu + \eta_2)^{n-1} - (\beta\mu - \eta_1)^{n-1} = (\beta\mu)^{n-1} \left[\left(1 + \frac{\eta}{\beta\mu}\right)^{n-1} - \left(1 - \frac{\eta}{\beta\mu}\right)^{n-1} \right] \quad (15.3.6)$$

$$\approx (\beta\mu)^{n-1} \left[\left(1 + (n-1)\frac{\eta}{\beta\mu}\right) - \left(1 - (n-1)\frac{\eta}{\beta\mu}\right) \right] \quad (15.3.7)$$

$$= 2(n-1)(\beta\mu)^{n-2}\eta \quad (15.3.8)$$

so that:

$$\Gamma(n)f_n(z) = \frac{(\beta\mu)^n}{n} + 2(n-1)(\beta\mu)^{n-2} \int_0^{\infty} \frac{\eta}{e^{\eta} + 1} d\eta \quad (15.3.9)$$

We recognize the Fermi-Dirac function:

$$\int_0^{\infty} \frac{\eta}{e^{\eta} + 1} d\eta = f_2(1) = \frac{\pi^2}{12} \quad (15.3.10)$$

so we finally find that ($\ln z = \beta\mu$):

$$\Gamma(n)f_n(z) = \frac{(\ln z)^n}{n} + \frac{\pi^2}{6}(n-1)(\ln z)^{n-2} \quad (15.3.11)$$

or in another words:

$$f_n(z) \approx \frac{(\ln z)^n}{\Gamma(n+1)} \left(1 + \frac{\pi^2 n(n-1)}{6 (\ln z)^2} \right), \quad z \ll 1 \quad (15.3.12)$$

This allows us to give a low temperature approximation to N

$$N = \frac{g_s V}{\lambda_{th}^3} f_{3/2}(z) \approx \frac{g_s V}{\lambda_{th}^3} \frac{(\ln z)^{3/2}}{\Gamma(5/2)} \left(1 + \frac{\pi^2}{6} \frac{3}{4(\ln z)^2} \right) \quad (15.3.13)$$

$$= \frac{g_s V}{\lambda_{th}^3} \frac{4(\beta\mu)^{3/2}}{3\sqrt{\pi}} \left(1 + \frac{\pi^2}{8} \frac{1}{(\beta\mu)^2} \right) \quad (15.3.14)$$

$$= \frac{g_s V}{6\pi^2 \hbar^3} (2m\mu)^{3/2} \left(1 + \frac{\pi^2}{8} \frac{1}{(\beta\mu)^2} \right) \quad (15.3.15)$$

and the Fermi energy:

$$E_F = \frac{\hbar^2}{2m} \left[\frac{6\pi^2}{g_s} \frac{g_s V}{\lambda_{th}^3} \frac{4(\beta\mu)^{3/2}}{3\sqrt{\pi}} \left(1 + \frac{\pi^2}{8} \frac{1}{(\beta\mu)^2} \right) \right]^{2/3} \quad (15.3.16)$$

$$= \mu \left(1 + \frac{\pi^2}{8} \frac{1}{(\beta\mu)^2} \right)^{2/3} \approx \mu \left(1 + \frac{\pi^2}{12} \frac{1}{(\beta\mu)^2} \right) \quad (15.3.17)$$

By the same trick as before this can be inverted to give an expression for the chemical potential:

$$\mu \approx E_F \left(1 - \frac{\pi^2}{12} \frac{1}{(\beta E_F)^2} \right) \quad (15.3.18)$$

Finally, the energy density:

$$\frac{U}{V} = \frac{3g_s}{2\beta\lambda_{th}^3} \frac{8(\beta\mu)^{5/2}}{15\sqrt{\pi}} \left(1 + \frac{\pi^2}{6} \frac{15}{4(\beta\mu)^2} \right) \quad (15.3.19)$$

$$= \frac{4g_s}{5\beta\lambda_{th}^3} \frac{(\beta\mu)^{5/2}}{\sqrt{\pi}} \left(1 + \frac{5\pi^2}{8} \frac{1}{(\beta\mu)^2} \right) \quad (15.3.20)$$

To compute C_V we need to take the derivative of the above with respect to temperature. This is problematic since we have μ in our expression, which has gained a temperature dependence at fixed particle number. We fix this by calculating $\frac{U}{N}$:

$$\frac{U}{N} = \frac{4g_s}{5\beta\lambda_{th}^3} \frac{(\beta\mu)^{5/2}}{\sqrt{\pi}} \left(1 + \frac{5\pi^2}{8} \frac{1}{(\beta\mu)^2} \right) \frac{6\pi^2 \hbar^3}{g_s V} \frac{1}{(2m\mu)^{3/2}} \left(1 - \frac{\pi^2}{8} \frac{1}{(\beta\mu)^2} \right) \quad (15.3.21)$$

$$= \frac{3}{5} \mu \left(1 + \frac{\pi^2}{2} \frac{1}{(\beta\mu)^2} \right) \quad (15.3.22)$$

and inserting (15.3.18):

$$\frac{U}{N} \approx \frac{3}{5} E_F \left(1 - \frac{\pi^2}{12} \frac{1}{(\beta E_F)^2} \right) \left[1 + \frac{\pi^2}{2} \frac{1}{(\beta E_F)^2} \left(1 - \frac{\pi^2}{6} \frac{1}{(\beta E_F)^2} \right) \right] \quad (15.3.23)$$

$$= \frac{3}{5} E_F \left(1 - \frac{\pi^2}{12} \frac{1}{(\beta E_F)^2} \right) \left(1 + \frac{\pi^2}{2} \frac{1}{(\beta E_F)^2} \right) \quad (15.3.24)$$

$$= \frac{3}{5} E_F \left(1 + \frac{5\pi^2}{12} \frac{1}{(\beta E_F)^2} \right) \quad (15.3.25)$$

Taking the partial derivative is much simpler now:

$$C_V = \frac{3}{5} E_F N \cdot \frac{5\pi^2}{12} \frac{1}{E_F^2} k_B^2 \cdot 2T = N k_B T \frac{\pi^2}{2T_F} \quad (15.3.26)$$

where $T_F = \frac{E_F}{k_B}$ is the **Fermi temperature**.

15.4 Magnetism

When discussing the motion of electrons in a magnetic field, we must consider the Lorentz force and the coupling of the electron spin to the magnetic field. The first interaction gives rise to “Landau diamagnetism” while the second gives rise to “Pauli paramagnetism”.

15.4.1 Landau diamagnetism

The Hamiltonian of an electron of charge $-e$ in a magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$ is:

$$H = \frac{(\mathbf{p} + e\mathbf{A})^2}{2m} \quad (15.4.1)$$

Interestingly, when calculating the classical partition function, the effect of adding a magnetic field is not noticeable. Indeed:

This result is the **Bohr-van Leeuwen** theorem. Magnetism can therefore appear only in the quantum regime. The Schrodinger equation for the electron in a constant magnetic field $\mathbf{B} = B\hat{z}$ was solved in the Quantum mechanics volume. The energy levels were found to be:

$$E = \frac{\hbar^2 k_z^2}{2m} + \hbar\omega_c \left(n + \frac{1}{2} \right) \quad (15.4.2)$$

where the degeneracy of each oscillator is $\frac{eBL^2}{2\pi\hbar} = \frac{\Phi}{\Phi_0}$ ignoring spin degeneracy. The partition function is then given by:

$$\ln \mathcal{Z} = \frac{L}{2\pi} \int \frac{2\Phi}{\Phi_0} \ln \left[1 + z \exp \left(- \frac{\beta \hbar^2 k_z^2}{2m} - \beta \hbar \omega_c (n + 1/2) \right) \right] dk_z \quad (15.4.3)$$

The Euler sum formula:

$$\sum_{n=0}^{\infty} f\left(n + \frac{1}{2}\right) \approx \int_0^{\infty} f(x) dx + \frac{1}{24} f'(0) \quad (15.4.4)$$

can be applied to:

$$f(x) = \int \ln \left[1 + \exp \left(-\frac{\beta \hbar^2 k_z^2}{2m} + \beta \mu - \beta \hbar \omega x \right) \right] dk_z \quad (15.4.5)$$

We find that:

$$\ln \mathcal{Z} = \frac{VB}{\pi \Phi_0} \sum_{n=0}^{\infty} f(n + 1/2) \quad (15.4.6)$$

$$\approx \frac{VB}{\pi \Phi_0} \int_0^{\infty} f(x) dx - \frac{VB}{\pi \Phi_0} \frac{\hbar \omega_c}{24} \int \frac{1}{z^{-1} \exp(\beta \hbar^2 k_z^2 / 2m) + 1} dk_z \quad (15.4.7)$$

The first term is just the contribution of the partition function of a classical ideal gas, and does not depend on the magnetic field as shown earlier. The second does depend on the magnetic field, and at $T = 0$ the integrand is 1 if $|k| < k_F$ and vanishes otherwise, giving:

$$\ln \mathcal{Z} = \ln \mathcal{Z}_{cl} - \frac{mV}{2\pi^2 \hbar^2} \frac{(\hbar \omega_c)^2}{12} \beta k_F \quad (15.4.8)$$

We now note that:

$$g(E_F) = \frac{mV}{\pi^2 \hbar^2} k_F \quad (15.4.9)$$

so

$$\ln \mathcal{Z} = \ln \mathcal{Z}_{cl} - g(E_F) \frac{(\hbar \omega_c)^2}{24} \beta \quad (15.4.10)$$

The magnetization is then given by:

$$M = \frac{1}{\beta} \frac{\partial}{\partial B} (\ln \mathcal{Z}) = -g(E_F) \frac{(\hbar e)^2}{12m^4} B \quad (15.4.11)$$

Using the Bohr magneton $\mu_B = \frac{|e|\hbar}{2mc}$ we find that:

$$M = -\frac{1}{3} g(E_F) \mu_B^2 B \quad (15.4.12)$$

Note that this magnetization has a negative sign and thus it points away from the direction of the applied magnetic field, it is **diamagnetic**.

15.4.2 Pauli paramagnetism

15.5 White Dwarf stars

Consider a star at temperature $T \rightarrow 0$ that has exhausted all its nuclear fuel, a **white dwarf**. To support themselves and avoid collapsing completely onto itself, these stars need to use the degeneracy pressure calculated earlier.

Note that the fermion energies have both a kinetic contribution and a gravitational contribution, which we have not taken into account in $g(E)$. It can be easily shown that for a

dwarf star of uniform density, mass M and radius R :

$$E_G = -\frac{3}{5} \frac{GM^2}{R} \quad (15.5.1)$$

giving a total energy of:

$$E = -\alpha \frac{M^2}{R} + \gamma \int_0^{E_F} E^{3/2} dE = -\alpha \frac{M^2}{R} + \gamma E_F^{5/2} \quad (15.5.2)$$

Now $V \propto R^3$ and $m = \frac{M}{N}$ so $E_F \propto \frac{M^{2/3}}{mR^2}$. Consequently, for the total energy to be minimized, the equilibrium condition reads that:

$$-\alpha \frac{M^2}{R} + \gamma \frac{M^{5/3}}{R^2} = 0 \implies R \propto M^{-1/3} \quad (15.5.3)$$

This is a perplexing result, the more massive the star gets the smaller its size becomes!

We have not taken into account the fact that as the star shrinks, the Fermi energy grows, until it becomes on the order of magnitude of the electron mass. Here our non-relativistic treatment does not hold anymore. Consequently, we should use the relativistic density of states, and expand it in the ultra-relativistic limit $\frac{m}{E} \ll 1$:

$$g(E) = \frac{V}{\pi^2 \hbar^3 c^3} E^2 \sqrt{1 - \frac{m^2 c^4}{E^2}} \approx \frac{V}{\pi^2 \hbar^3 c^3} \left(E^2 - \frac{m^2 c^4}{2} \right) \quad (15.5.4)$$

giving a total kinetic energy upon integration of:

$$E_{kin} \approx \frac{V}{\pi^2 \hbar^3 c^3} \left(\frac{1}{4} E_F^4 - \frac{m^2 c^4}{4} E_F^2 \right) \quad (15.5.5)$$

To find the Fermi energy, we can express the particle number as:

$$N \approx \frac{V}{\pi^2 \hbar^3 c^3} \left(\frac{1}{3} E_F^3 - \frac{m^2 c^4}{2} E_F \right) \implies E_F \approx \left(\frac{3N}{V} \pi^2 \hbar^3 c^3 \right)^{1/3} \quad (15.5.6)$$

Inserting this into (15.5.5) we find that:

$$E = \left[\frac{3\hbar c}{4} \left(\frac{9\pi M^4}{4m_p^4} \right)^{1/3} - \frac{3}{5} GM \right] \frac{1}{R} - O(R) \quad (15.5.7)$$

where m_p is the proton mass, and $M = Nm_p$ since we may assume that most of the star's mass comes from the nucleus. As long as the term in square brackets is positive, we can always find a minimum energy. However if this is not the case then the star will be unstable, shrinking further and further. The mass at which this occurs is known as the **Chandrasekhar limit**:

$$M_C = \frac{3\sqrt{\pi}}{2} \left(\frac{5\hbar c}{4G} \right)^{3/2} \frac{1}{m^2} \propto \frac{M_P^3}{m_p^2} \quad (15.5.8)$$

where $M_P = \sqrt{\frac{\hbar c}{G}}$ is known as the **Planck mass**. If a star has a mass larger than M_C , its

collapse will either be stopped by the degeneracy pressure of the neutrons, or will form a black hole.

Real gases and Interacting systems

16.1 The virial expansion

Thus far we have only considered non-interacting systems where the partition function could be factorized into the partition function for each particle. When we add interactions to the system this no longer holds and the partition function becomes much more arduous to solve.

In this chapter we will concern ourselves with weakly interacting gases, where the density of particles is low and the temperature is high. We can write a general equation of state as a power expansion of N/V , which in the low density limit reduces to the ideal gas law:

$$\frac{pV}{Nk_B T} = 1 + B_2(T) \frac{N}{V} + B_3(T) \left(\frac{N}{V}\right)^2 + \dots \quad (16.1.1)$$

This is known as the virial expansion. The $B_i(T)$ functions are known as virial coefficients, understanding them will be the goal of this chapter. We begin by calculating $B_2(T)$, the second virial coefficient.

16.2 The Second Virial coefficient

The Hamiltonian for a gas of particles interacting through a central two-body potential $U(r)$ is given by

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i>j} U(r_{ij}), \quad r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \quad (16.2.1)$$

where \mathbf{r}_i is the position of the i th particle. The partition function now becomes:

$$Z = \frac{1}{N! h^{3N}} \int d^{3N} \mathbf{p} d^{3N} \mathbf{r} \exp \left(-\frac{\beta p_i^2}{2m} - \beta \sum_{j>k} U(r_{jk}) \right) \quad (16.2.2)$$

$$= \frac{1}{N! \lambda^{3N}} \int d^{3N} \mathbf{r} e^{-\beta \sum_{j>k} U(r_{jk})} \quad (16.2.3)$$

$$= \frac{1}{N! \lambda^{3N}} \int d^{3N} \mathbf{r} \prod_{j>k} e^{-\beta U(r_{jk})} \quad (16.2.4)$$

Unfortunately, the integral over position looks very intimidating. Let us look more closely at the properties of U .

We want to model neutral atoms which can only interact through Van der Waals interactions. These arise due to temporary dipole-dipole interactions induced by quantum fluctuations in the atom's electronic wave-function, and are of the form $U \sim \frac{1}{r^6}$. Indeed suppose that one atom, due to fluctuations, has a temporary dipole moment p_1 producing an electric field of the form $E \sim \frac{p_1}{r^3}$. As a result the second atom will have an induced dipole moment of $p_2 \sim E \sim \frac{p_1}{r^3}$ meaning that their potential energy scales as $\frac{p_1 p_2}{r^3} \sim \frac{1}{r^6}$. Consequently the potential will be strongly repulsive at small distances, but we should also expect it to be attractive for large distances, vanishing at $r \rightarrow \infty$. One common potential that satisfies these requirements is the Lennard-Jones potential:

$$U(r) \sim \left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^6 \quad (16.2.5)$$

If the gas density is low, or the temperature is high, then we should expect to recover the classical ideal gas. It therefore seems logical to expand the partition function in the ideal gas limit where $e^{-\beta U} \approx 1$. This yields the following expansion parameter

$$f_{jk} = e^{-\beta U(r_{jk})} - 1 \quad (16.2.6)$$

which vanishes in the low density limit $r_{ik} \rightarrow \infty$, and is equal to -1 when $r_{ik} \rightarrow 0$.

$$e^{-\beta \sum_{j>k} U(r_{jk})} = 1 - \beta \sum_{j>k} U(r_{jk}) \quad (16.2.7)$$

We then find that

$$Z = \frac{1}{N! \lambda^{3N}} \int d^{3N} \mathbf{r} \prod_{j>k} (1 + f_{jk}) \quad (16.2.8)$$

$$= \frac{1}{N! \lambda^{3N}} \int d^{3N} \mathbf{r} \left(1 + \sum_{j>k} f_{jk} + \sum_{\substack{j>k \\ l>m}} f_{jk} f_{lm} + \dots \right) \quad (16.2.9)$$

Note note that

$$\int d^{3N} \mathbf{r} f_{jk} = V^{N-2} \int d^3 \mathbf{r}_j d^3 \mathbf{r}_k f_{jk} \quad (16.2.10)$$

We now perform a change of variables from $(\mathbf{r}_j, \mathbf{r}_k)$ to $(\mathbf{R}_{jk}, \mathbf{r}_{jk})$ where

$$\mathbf{R}_{jk} = \frac{1}{2}(\mathbf{r}_j + \mathbf{r}_k), \quad \mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k \quad (16.2.11)$$

whose Jacobian is

$$\frac{\partial(\mathbf{r}_j, \mathbf{r}_k)}{\partial(\mathbf{R}_{jk}, \mathbf{r}_{jk})} = \begin{vmatrix} 1 & \frac{1}{2} \\ 1 & -\frac{1}{2} \end{vmatrix} = -1 \quad (16.2.12)$$

thus yielding

$$\int d^{3N} \mathbf{r} f_{jk} = V^{N-1} \int d^3 \mathbf{r} f(r) \quad (16.2.13)$$

where the limits of integration still span \mathbb{R}^3 (we don't have to worry about what happens at the edges since they provide a negligible contribution).

The sum over $\sum_{j>k} = \frac{1}{2} \sum_j \sum_{k \neq j} = \frac{1}{2} N(N-1)$ implying that the partition function takes the form

$$Z = Z_{\text{ideal}} \left(1 + \frac{N^2}{2V} \int d^3\mathbf{r} f(r) + \dots \right) \quad (16.2.14)$$

It will be helpful (to calculate logarithms more easily) to write the expression in brackets as a power of N . We can do so if the term with the integral is small (that is when we're near the ideal limit), in which case

$$Z = Z_{\text{ideal}} \left(1 + \frac{N}{2V} \int d^3\mathbf{r} f(r) + \dots \right)^N \quad (16.2.15)$$

The free energy then reads

$$F = F_{\text{ideal}} - Nk_B T \ln \left(1 + \frac{N}{2V} \int d^3\mathbf{r} f(r) \right) \quad (16.2.16)$$

For the Lennard-Jones potential the integral is approximately $\sim r_0^3$, where r_0 is the separation minimising the potential. Consequently our expansion holds when

$$\frac{N}{2V} r_0^3 \ll 1 \iff \frac{N}{V} \ll \frac{1}{r_0^3} \quad (16.2.17)$$

so when the gas density is much lower than the atomic density.

16.3 Van der Waals equation

Let us compute the partition function for the Sutherland potential

$$U(r) = \begin{cases} \infty, & r < r_0 \\ -U_0 \left(\frac{r_0}{r} \right)^6, & r \geq r_0 \end{cases} \quad (16.3.1)$$

which models the gas atoms as solid spheres of radii $\frac{r_0}{2}$. Then we find that

$$\int d^3\mathbf{r} f(r) = 4\pi \left[\int_{r_0}^{\infty} dr r^2 (e^{\beta U_0 r_0^6 / r^6} - 1) - \int_0^{r_0} dr r^2 \right] \quad (16.3.2)$$

$$\approx 4\pi \left(-\frac{r_0^3}{3} + \int_{r_0}^{\infty} dr \frac{\beta U_0 r_0^6}{r^4} \right) \quad (16.3.3)$$

$$= \frac{4\pi}{3} (\beta U_0 - 1) r_0^3 \quad (16.3.4)$$

This gives the following approximation to the free energy

$$F \approx F_{\text{ideal}} - Nk_B T \ln \left(1 + \frac{2\pi N}{3V} (\beta U_0 - 1) r_0^3 \right) \quad (16.3.5)$$

Differentiating with respect to volume

$$p = -\left.\frac{\partial F}{\partial V}\right|_T = p_{\text{ideal}} + Nk_B T \frac{-\frac{2\pi N}{3V^2}(\beta U_0 - 1)r_0^3}{1 + \frac{2\pi N}{3V}(\beta U_0 - 1)r_0^3} \quad (16.3.6)$$

$$\approx \frac{Nk_B T}{V} \left(1 - \frac{2\pi N}{3V}(\beta U_0 - 1)r_0^3\right) \quad (16.3.7)$$

$$= \frac{k_B T}{v} \left(1 - \frac{2\pi}{3v} \left(1 - \frac{U_0}{k_B T}\right) r_0^3\right) \quad (16.3.8)$$

where we defined $v = \frac{V}{N}$. We can rearrange this into

$$p + \frac{2\pi r_0^3 U_0}{3v^2} = \frac{k_B T}{v} \left(1 + \frac{2\pi r_0^3}{3v}\right) \quad (16.3.9)$$

Now assuming that $v \gg r_0^3$, that is, the volume per particle is much larger than the atomic volume $4\pi r_0^3/3$ (i.e. low density limit) then we can expand the LHS to give

$$p + \frac{2\pi r_0^3 U_0}{3} \frac{1}{v^2} = \frac{k_B T}{v} \left(1 - \frac{2\pi r_0^3}{3} \frac{1}{v}\right)^{-1} \quad (16.3.10)$$


Letting $a = \frac{2\pi}{3} r_0^3 U_0$ and $b = \frac{2\pi}{3} r_0^3 = \frac{a}{U_0}$ then we find

$$\left(p + \frac{a}{v^2}\right)(v - b) = k_B T \quad (16.3.11)$$

which is the Van der Waals equation of state.

We have calculated the second virial coefficient in its generality and applied it to the Sutherland potential to derive the Van der Waals equation. However, higher order coefficients are harder to solve, and we need a systematic method to evaluate them. The cluster expansion will allow us to do this.

16.4 The cluster expansion

We introduce a pictorial notation that will enable us to calculate the cluster expansion terms to any order. Given a term of the type $f_{ij}f_{kl}\dots$ we firstly draw N atoms, and then draw a line between all the pairs appearing as indices, so between i, j , then between k, l etc... For example, we see that the graph associated to $f_{21}f_{23}f_{31}$ with $N = 3$ is . It is clear that in the cluster expansion (16.2.9) we will have all possible graphs G with N nodes

$$Z = \frac{1}{N! \lambda^{3N}} \sum_{G \in \mathcal{G}_N} W[G] \quad (16.4.1)$$

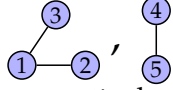
where

$$W[G] = \int d^{3N} \mathbf{r} G \quad (16.4.2)$$

For example we have that if $N = 6$ then

$$W \left[\begin{array}{c} \text{3} \\ \text{1} \text{---} \text{2} \\ \text{5} \end{array} \quad \begin{array}{c} \text{4} \\ \text{5} \end{array} \quad \text{6} \right] = \left(\int d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 d^3 \mathbf{r}_3 f_{13} f_{12} \right) \left(\int d^3 \mathbf{r}_4 d^3 \mathbf{r}_5 f_{45} \right) \left(\int d^3 \mathbf{r}_6 \right) \quad (16.4.3)$$

so we see that the integrals factorize into connected graphs. These connected graphs



are examples of what we will call a 3-cluster, 2-cluster and 1-cluster respectively, since they each have 3, 2 and 1 nodes respectively. Thus a l -cluster will be a connected graph with l -nodes. In general we should find that if we have m_l l -clusters in the cluster expansion then

$$\sum_{l=1}^N m_l l = N \quad (16.4.4)$$

Let us define the following integrals

$$U_l \equiv \int d^{3m_l} \mathbf{r} \sum_{G \in \mathcal{C}_l} G \quad (16.4.5)$$

where \mathcal{C}_l is the set of l -clusters. For example

$$U_3 = \int d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 d^3 \mathbf{r}_3 \left[\begin{array}{c} \text{3} \\ \text{1} \text{---} \text{2} \end{array} + \begin{array}{c} \text{3} \\ \text{1} \text{---} \text{2} \end{array} + \begin{array}{c} \text{3} \\ \text{1} \text{---} \text{2} \end{array} + \begin{array}{c} \text{3} \\ \text{1} \text{---} \text{2} \end{array} \right] \quad (16.4.6)$$

which contains both f^2 and f^3 terms. Then we can write

$$\sum_G W[G] = \sum_{\{m_l\}_N} g \prod_l U_l^{m_l} \quad (16.4.7)$$

where g is a proportionality factor due to any multiple-counting, and $\{m_l\}_N$ is the set of all $\{m_1, \dots, m_N\}$ which satisfy (16.4.4). Our job is to find $g(N)$.

For a given $\{m_l\}_N = \{m_1, m_2, \dots, m_N\}$, we can define the sum of all graphs with m_1 1-clusters etc... as $S\{m_l\}_N$ implying that

$$Q_N \equiv \sum_{G \in \mathcal{G}_N} W[G] = \sum_{\{m_l\}_N} S\{m_l\} \quad (16.4.8)$$

Each term in $S\{m_1, \dots, m_N\}$ is of the form

$$W[(m_1 \text{ 1-clusters}) (m_2 \text{ 2-clusters}) \dots] \quad (16.4.9)$$

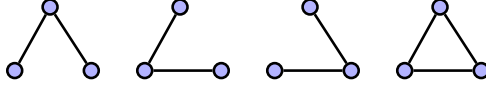
For example if $N = 9$ then a possible partition is $\{m_l\} = \{1, 1, 2\}$, and thus a possible term in $S\{1, 1, 2\}$ is

$$W \left[\begin{array}{c} \text{3} \\ \text{1} \text{---} \text{2} \end{array} \quad \text{4} \quad \begin{array}{c} \text{6} \\ \text{5} \end{array} \quad \begin{array}{c} \text{9} \\ \text{7} \text{---} \text{8} \end{array} \right] \quad (16.4.10)$$

If every l -cluster had the same value upon integration our job would be pretty much done. However, this is clearly not the case e.g. it is easy to see that the two 3-clusters in the above

example will yield different integrals.

If we categorize l -clusters into different types corresponding to different values when integrated, then we see that there is only one type of 1-clusters and 2-clusters, while there are four types of 3-clusters



Note that 3-clusters of a different type don't necessarily have different values, indeed the first three are equal to each other. Nevertheless they are not equivalent as they are distinct terms in the cluster expansion (different indices in $f_{ij}f_{kl}\dots$).

Suppose there are K_l different types of l -clusters. Then the m_l l -clusters can be decomposed into n_1 l -clusters of type 1, n_2 l -clusters of type 2, etc...

$$W[(m_l \text{ } l\text{-clusters})] = W[(n_1 \text{ } l\text{-clusters type 1}) (n_2 \text{ } l\text{-clusters type 2}) \dots] \quad (16.4.11)$$

$$= (W[l\text{-cluster of type 1}])^{n_1} (W[l\text{-cluster of type 2}])^{n_2} \dots \quad (16.4.12)$$

where we require

$$\sum_{i=1}^{K_l} n_i = m_l \quad (16.4.13)$$

In (16.4.10) for example we had two different 3-clusters, each of a different type.

Let us now consider all terms in $S\{m_l\}$ where all but the m_l l -cluster are fixed, the latter ranging over all possible l -clusters. The numbers at the nodes of the m_l l -clusters is therefore fixed, but we further fix the l numbers at the nodes of each of the l -clusters (we'll deal with the permutations later). Therefore two graphs in the set we are considering can only differ in the number of l -cluster types $\{n_1, \dots, n_{K_l}\}$.

The m_l l -cluster contribution is given by summing over all sets of number of l -cluster types $\{n_i\} \equiv \{n_1, \dots, n_{K_l}\}$:

$$W[(m_l \text{ } l\text{-clusters})] = \sum_{\{n_i\} \in A} \prod_{i=1}^{K_l} (W[l\text{-cluster of type } i])^{n_i} \quad (16.4.14)$$

where

$$A = \left\{ \{n_1, \dots, n_{K_l}\} : \sum_i n_i = m_l, 0 \leq n_i \leq m_l \right\} \quad (16.4.15)$$

Repeating this argument for all l we then find that

$$\prod_{l=1}^N W[(m_l \text{ } l\text{-clusters})] = \prod_{l=1}^N \left(\sum_{\{n_i\} \in A} \prod_{i=1}^{K_l} (W[l\text{-cluster of type } i])^{n_i} \right) \quad (16.4.16)$$

is the sum of the integrals of the N -particle graphs contributing to $S\{m_i\}$ with fixed numbers on the nodes on each cluster.

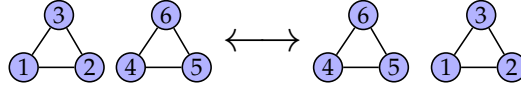
All other graphs can be generated from the ones we have identified in (16.4.16) by permu-

tations of the particle numbers. In other words if we view the calculation we have just done as looking only at the shape of the graphs, now we are tasked with assigning numbers to each node. There are $N!$ such permutations, but not all of these will produce a new graph. Indeed there are two types of permutations that will not produce new graphs

- (i) permutations within each of the clusters: suppose we permute the particle numbers within a given l -cluster. This will either produce the same cluster, or another l -cluster of a different type. Both are accounted for, so we divide by $\prod_{l=1}^N (l!)^{m_l}$:

$$\frac{N!}{\prod_{l=1}^N (l!)^{m_l}} \quad (16.4.17)$$

- (ii) permutations of clusters of same type: suppose we permute the particle numbers so that two clusters of the same type effectively get interchanged



Since there are $n_j!$ permutations of all l -clusters of type j , we should divide (16.4.17) by $n_1!n_2!\dots n_{K_l}!$ so that the correct factor becomes

$$\frac{N!}{\prod_{l=1}^N (l!)^{m_l} \prod_{j=1}^{K_l} n_j!} \quad (16.4.18)$$

We find that

$$S\{m_i\} = N! \prod_{l=1}^N \frac{1}{(l!)^{m_l}} \prod_{i=1}^{K_l} \frac{(W[l\text{-cluster of type } i])^{n_i}}{n_i!} \quad (16.4.19)$$

The multinomial theorem is a generalization of the binomial theorem and tells us that

$$\left(\sum_{i=1}^m x_i \right)^N = N! \sum_{\{n_i\} \in A} \prod_{i=1}^m \frac{x_i^{n_i}}{n_i!} \quad (16.4.20)$$

Consequently

$$S\{m_i\} = N! \prod_{l=1}^N \frac{U_l^{m_l}}{(l!)^{m_l} m_l!} \quad (16.4.21)$$

finally implying that

$$Z = \frac{1}{\lambda^{3N}} \sum_{\{m_l\}_N} \prod_{l=1}^N \frac{U_l^{m_l}}{(l!)^{m_l} m_l!} \quad (16.4.22)$$

Just like when computing quantum statistics, we see that performing the restricted sum over $\{m_l\}_N$ is a messy business. We can simplify matters by moving to the grand canonical

ensemble, where the partition function reads

$$\mathcal{Z} = \sum_{N=0}^{\infty} z^N \left(\frac{1}{\lambda^{3N}} \sum_{\{m_i\}_N} \prod_{l=1}^N \frac{U_l^{m_l}}{(l!)^{m_l} m_l!} \right) \quad (16.4.23)$$

$$= \sum_{m_1, \dots, m_N=0}^{\infty} \left[\prod_{l=1}^{\infty} \frac{U_l^{m_l}}{(l!)^{m_l} m_l!} \left(\frac{z}{\lambda^3} \right)^{l m_l} \right] \quad (16.4.24)$$

$$= \prod_{l=1}^{\infty} \sum_{m_l=0}^{\infty} \frac{1}{m_l!} \left(\frac{U_l z^l}{l! \lambda^{3l}} \right)^{m_l} = \prod_{l=1}^{\infty} \exp \left(\frac{U_l z^l}{\lambda^{3l} l!} \right) \quad (16.4.25)$$

Let us define $b_l = \frac{\lambda^3}{V} \frac{U_l}{l! \lambda^{3l}}$ then

$$\mathcal{Z} = \exp \left(\frac{V}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l \right) \quad (16.4.26)$$

This allows us to calculate the pressure

$$p = \frac{k_B T}{V} \ln \mathcal{Z} = \frac{k_B T}{\lambda^3} \sum_{l=1}^{\infty} b_l z^l \quad (16.4.27)$$

and the particle density

$$\frac{N}{V} = \frac{z}{V} \frac{\partial}{\partial z} (\ln \mathcal{Z}) = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} l b_l z^l \quad (16.4.28)$$

The equation of state for the gas can be expressed as

$$\frac{pV}{Nk_B T} = \frac{\sum_l b_l z^l}{\sum_l l b_l z^l} \quad (16.4.29)$$

Equating this to the virial expansion in (16.1.1) we find that

$$\sum_{l=1}^{\infty} b_l z^l = \sum_{l=1}^{\infty} B_l \left(\frac{N}{V} \right)^{l-1} \sum_{m=1}^{\infty} m b_m z^m \quad (16.4.30)$$

$$= \sum_{l=1}^{\infty} B_l \frac{1}{\lambda^{3(l-1)}} \left(\sum_{n=1}^{\infty} n b_n z^n \right)^{l-1} \sum_{m=1}^{\infty} m b_m z^m \quad (16.4.31)$$

$$= \left[1 + \frac{B_2}{\lambda^3} (z + 2b_2 z^2 + 3b_3 z^3 + \dots) + \frac{B_3}{\lambda^6} (z + 2b_2 z^2 + 3b_3 z^3 + \dots)^2 \right] \times (z + 2b_2 z^2 + 3b_3 z^3 + \dots) \quad (16.4.32)$$

where $b_1 = B_1 = 1$ are the first virial coefficients of course. Keeping only terms up to fourth order

$$z + b_2 z^2 + b_3 z^3 = z + \left(2b_2 + \frac{B_2}{\lambda^3} \right) z^2 + \left(3b_3 + \frac{4b_2 B_2}{\lambda^3} + \frac{B_3}{\lambda^6} \right) z^3 + o(z^4) \quad (16.4.33)$$

Equating coefficients we get

$$B_2 = -\lambda^3 b_2^2 = -\frac{1}{2V} \int d\mathbf{r}_1 d\mathbf{r}_2 f(r_{12}) = -\frac{1}{2} \int d^3r f(r) \quad (16.4.34)$$

just as we found earlier. We can also calculate the third virial coefficient

$$B_3 = -2\lambda^6 b^3 - 4\lambda^3 b_2 B_2 = \lambda^6 (4b_2^2 - 2b^3) \quad (16.4.35)$$

Repeating this process ad infinitum we can systematically find all virial coefficients.

Phase transitions

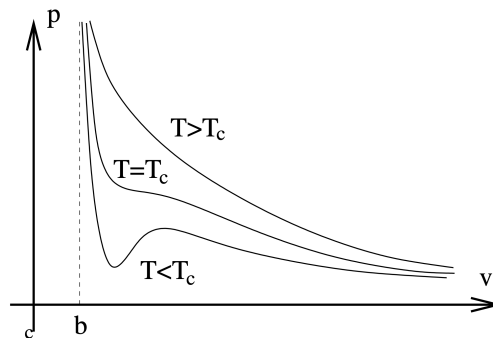
17.1 The Van der Waals gas-liquid phase transition

The critical temperature

We have observed that when a gas of non-interacting bosons is cooled below a certain temperature its properties start to change abruptly. We start with the Van der Waals (VdW) equation we derived earlier:

$$p = \frac{k_B T}{v - b} - \frac{a}{v^2} \quad (17.1.1)$$

where $v = \frac{V}{N}$ is the average volume taken by one particle. At constant temperature we get the following isotherms: where we ignored the $\frac{a}{v^2}$ term since $v > b$. As we can see



for large temperatures the isotherm looks like a smooth, strictly decreasing function. For sufficiently low temperatures however the isotherm seems to have a . To transition from one regime to the other we must have a temperature, a critical temperature, at which the isotherm has an inflection point. A quick calculation yields:

$$\frac{dp}{dv} = \frac{2a}{v^3} - \frac{k_B T_c}{(v - b)^2} = 0, \quad \frac{d^2p}{dv^2} = \frac{2k_B T}{(v - b)^3} - \frac{6a}{v^4} \implies T_c = \frac{8a}{27b} \frac{1}{k_B} \quad (17.1.2)$$

where the inflection point occurs at $v = 3b$.

Looking more closely at the $T < T_c$ isotherm we see that for some pressures three different particle densities are available. The highest particle density point (low v) corresponds to a liquid phase, since $|\frac{dp}{dv}| \gg 1$ implying that this phase is very hard to compress. The lowest particle density point (high v) corresponds to a gas phase, since $|\frac{dp}{dv}| \ll 1$ implying

that this phase is easy to compress. The middle solution is particularly different, if we decrease its pressure it will respond by decreasing its volume, while if we increase its pressure then its volume will also increase. In other words the compressibility is negative

$\kappa = -\left.\frac{\partial v}{\partial p}\right|_T < 0$, thus violating the principle of thermodynamic stability. Indeed either way any perturbation would lead this point to either extreme, it is not a stable phase at all.

Maxwell's construction

To fix the problem of the stable phase, Maxwell proposed a where we replace the unphysical, unstable part of the isotherms with a flat line. This flat line leads to a coexistence between the liquid and gaseous phase. We know that phase equilibrium at constant pres-

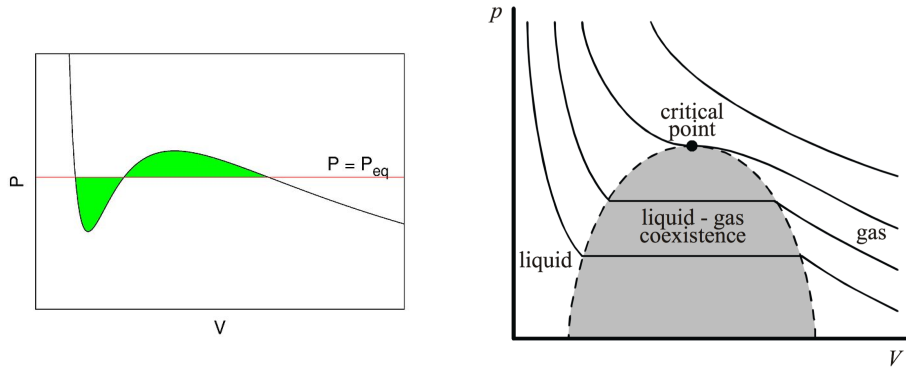


Figure 17.1. Carnot cycle and engine

sure and temperature requires the the chemical potential of the two phases m be equal, so in our case

$$\mu_g = \mu_l \quad (17.1.3)$$

where μ_g and μ_l are the chemical potentials of the gas and liquid phases respectively. Therefore the chemical potential will change along the isotherms by

$$d\mu = \left.\frac{\partial \mu}{\partial p}\right|_T dp \quad (17.1.4)$$

but we know that

$$\left.\frac{\partial G}{\partial p}\right|_{N,T} = \left.\frac{\partial \mu}{\partial p}\right|_T N = V \implies d\mu = \frac{V(p,T)}{N} dp \quad (17.1.5)$$

so that

$$\mu(p,T) = \mu_l + \int_{p_l}^p \frac{V(p',T)}{N} dp' \quad (17.1.6)$$

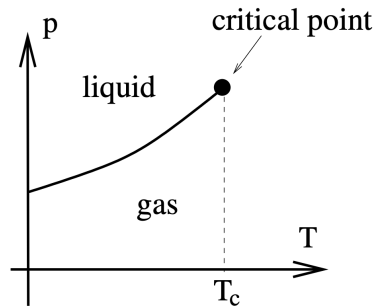
We thus see that the liquid and gas phase coexist at the pressure p where the integral vanishes. Therefore

$$\int_{p_l}^p v(p',T) dp' = \int_A^B d(pv) - \int_A^B p dv = p(v_g - v_l) - \int_A^B p dv$$

However, $p(v_l - v_g)$ is the area of the rectangle under p while the integral is the area under the isotherm. The above vanishes only if the area between the isotherm and p vanishes. If we calculate this pressure at all temperatures then we eventually get a coexistence line, as shown beside. Note that within this coexistence region the isotherms must be flat, reflecting the fact that the gas and liquid phases together, in equilibrium, can have different average densities despite each having a fixed density.

The Clausius-Clapeyron equation

Alternatively we can look at the $p - T$ diagram where the coexistence line now becomes a line, a phase boundary. If we are not on this boundary then the two phases cannot coexist, all particles are either liquid or gaseous, but if we are right on top of it then chemical equilibrium is achieved. At the critical point there is no coexistence as there is the isotherms are monotonic.



Since the phase equilibrium condition is $\frac{G_l}{N_l} = \frac{G_g}{N_g}$ it follows that the Gibbs free energy must be continuous on either side of the phase transition line. Defining $g = \frac{G}{N}$, $v = \frac{V}{N}$ and $s = \frac{S}{N}$ it then follows that

$$-s_l dT + v_l dp = -s_g dT + v_g dp \implies \frac{dp}{dT} = \frac{s_g - s_l}{v_g - v_l} \quad (17.1.7)$$

We further define the specific latent heat L as the energy per particle released as we go through the phase transition

$$L = T(s_g - s_l) \quad (17.1.8)$$

Then we have that

$$\boxed{\frac{dp}{dT} = \frac{L}{T(v_g - v_l)}} \quad (17.1.9)$$

which is known as the Clausius-Clapeyron equation. In the gas-liquid diagram we have a discontinuity in s giving rise to a non-zero latent heat, this is a first order phase transition (we explain the phase transition order in more detail later). As we move along the phase line towards the critical temperature the difference between s_l and s_g decreases, until at T_c it vanishes giving rise to a second order phase transition. Above T_c we no longer have a distinction between, so there is no longer a phase transition.

Let us make some further assumptions, namely that L is a constant, $v_g \gg v_l$ and that the

gas is now ideal. Then

$$\frac{dp}{dT} = \frac{Lp}{k_B T^2} \implies p = p_0 e^{-L/k_B T} \quad (17.1.10)$$

There is another elegant way to calculate the critical point. let's rearrange the VdW equation into

$$pv^3 - (pb + k_B T)v^2 + av - ab = 0 \quad (17.1.11)$$

If we are below the critical temperature then this has three real roots, but if we are above it then there is only one real root. So there must be some temperature where we go from one case to the other, where the three roots must coincide:

$$p_c(v - v_c) = 0 \quad (17.1.12)$$

It is easy to see that this occurs when $v_c = 3b$ and $p_c = \frac{a}{27b^2}$ which gives the desired critical temperature of $k_B T_c = \frac{8a}{27b}$.

Critical exponents

We can rewrite the VdW equation by defining the following dimensionless variables

$$T' = \frac{T}{T_c}, \quad v' = \frac{v}{v_c}, \quad p' = \frac{p}{p_c} \quad (17.1.13)$$

to find that

$$p' = \frac{8T'}{3v' - 1} - \frac{3}{v'^2} \quad (17.1.14)$$

Note also that the following ratio is independent of a and b (and should therefore be the same for all gases)

$$\frac{p_c v_c}{k_B T_c} = \frac{3}{8} \quad (17.1.15)$$

We can solve the VdW equation in this revamped form to yield the two stable solutions

$$T' = \frac{(3v'_l - 1)(3v'_g - 1)(v'_l + v'_g)}{8v_g^2 v_l^2} \quad (17.1.16)$$

There is some sort of universality in the phase transition.

Next let's look at what happens near the critical point as we move along the coexistence curve. We define $\epsilon = v'_g - v'_l$, and noting that v'_l and v'_g are symmetric in (17.1.16) we can set $v'_g = 1 + \epsilon/2$ and $v'_l = 1 - \epsilon/2$. Then

$$T' = \frac{(2 - 3\epsilon/2)(2 + 3\epsilon/2)}{4(1 - \epsilon^2/4)^2} \quad (17.1.17)$$

$$\approx \frac{1}{4} \left(4 - \frac{9}{4}\epsilon^2\right) \left(1 + \frac{\epsilon^2}{2}\right) \quad (17.1.18)$$

$$= \frac{1}{4} \left(4 - \frac{1}{4}\epsilon + o(\epsilon^2)\right) \quad (17.1.19)$$

We therefore find that

$$T' = 1 - \frac{1}{16}(v'_g - v'_l)^2 \implies v_g - v_l \sim (T_c - T)^{1/2} \quad (17.1.20)$$

Next let's look at the pressure as we approach the critical temperature along an isotherm.

Notice that at the critical point $\left. \frac{\partial p}{\partial v} \right|_{T_c} = \left. \frac{\partial^2 p}{\partial v^2} \right|_{T_c} = 0$ so we find that

$$p - p_c \sim (v - v_c)^3 \quad (17.1.21)$$

Finally let's look at the compressibility $\kappa = -\left. \frac{\partial v}{\partial p} \right|_T$. We expect this to diverge at the critical point so to leading order

$$\kappa \sim (T - T_c)^{-1} \quad (17.1.22)$$

Experiments have shown that the correct scaling laws are

$$v_g - v_l \sim (T_c - T)^\beta, \quad \beta \approx 0.32 \quad (17.1.23)$$

$$p - p_c \sim (v - v_c)^\delta, \quad \delta \approx 4.8 \quad (17.1.24)$$

$$\kappa \sim (T - T_c)^{-\gamma}, \quad \gamma \approx 1.2 \quad (17.1.25)$$

The numbers β, γ, δ are known as critical exponents.

17.2 Phase transitions in the Ising model

The Ising model

Consider a d -dimensional spin lattice with N sites, so that the spin on every site either points up ($s=1$) or down ($s=-1$). We insert a magnetic field B along the z -axis and add a (anti)ferromagnetic interaction J between neighbouring spins

$$H = -J \sum_{\langle ij \rangle} s_i s_j - B \sum_i s_i \quad (17.2.1)$$

If $J > 0$ then we have a ferromagnet favouring parallel spins, while if $J < 0$ then we have an antiferromagnet favouring antiparallel spins.

Note that this problem is closely related to that of a lattice gas. Indeed suppose we have N hard-core particles on a d -dimensional lattice. There is an attractive interaction $4J$ between neighbouring sites and a chemical potential μ which yields the following Hamiltonian

$$H = -4J \sum_{\langle ij \rangle} n_i n_j - \mu \sum_i n_i \quad (17.2.2)$$

This is equivalent to the Ising model if we define $s_i = 2n_i - 1$ and remember that $n_i = 0, 1$.

Mean field theory

We consider a d -dimensional spin lattice and write $s_i = \langle s_i \rangle + \delta s_i$ where δs_i is the spin fluctuation at site i . Then the spin interaction term becomes

$$s_i s_j = (\langle s_i \rangle + \delta s_i)(\langle s_j \rangle + \delta s_j) = \langle s_i \rangle \langle s_j \rangle + \delta s_i \langle s_j \rangle + \delta s_j \langle s_i \rangle + \delta s_i \delta s_j \quad (17.2.3)$$

Next we use a mean field (MF) approximation and assume that $(s_i - \langle s_i \rangle)(s_j - \langle s_j \rangle)$ is small, in other words the fluctuations in the spin of neighbouring sites is small. We can therefore ignore the $\delta s_i \delta s_j$ term and

$$s_i s_j = \langle s_i \rangle \langle s_j \rangle + \delta s_i \langle s_j \rangle + \delta s_j \langle s_i \rangle \quad (17.2.4)$$

$$= s_i \langle s_j \rangle + s_j \langle s_i \rangle - \langle s_i \rangle \langle s_j \rangle \quad (17.2.5)$$

Let us now impose translational invariance (and thus periodic boundary conditions) along our chain, meaning that the spin expectation value is the same along all sites: $\langle s_i \rangle = m$ for all i . Then the Ising hamiltonian becomes

$$H = -J \sum_{\langle ij \rangle} (2m s_i - m^2) - B \sum_i s_i \quad (17.2.6)$$

The first sum is over pairs of nearest neighbours, so $\sum_{\langle ij \rangle} s_i = \frac{q}{2} \sum_i s_i$ and $\sum_{\langle ij \rangle} m^2 = \frac{1}{2} N q m^2$ where q is the number of nearest neighbours for a given site. Consequently

$$H = \frac{1}{2} N q m^2 J - (J q m + B) \sum_i s_i \quad (17.2.7)$$

Defining an effective magnetic field $B_{\text{eff}} = B + J q m$ then we get

$$H = \frac{1}{2} N q m^2 J - B_{\text{eff}} \sum_i s_i \quad (17.2.8)$$

We have managed to decouple the spin-interaction and have a non-interacting model which we can easily solve using the tools we have developed thus far. Note that the effective magnetic field is slightly larger (or smaller) since we have absorbed the spin-interaction J into the magnetic field, and the mean field approximation ensured that this would be a constant contribution.

The partition function now reads ¹

$$Z = e^{-\frac{1}{2} \beta N q m^2 J} \sum_{\{s_i\}} \prod_i e^{\beta B_{\text{eff}} s_i} = e^{-\frac{1}{2} \beta N q m^2 J} \prod_i (e^{\beta B_{\text{eff}}} + e^{-\beta B_{\text{eff}}}) \quad (17.2.9)$$

$$= e^{-\frac{1}{2} \beta N q m^2 J} (e^{\beta B_{\text{eff}}} + e^{-\beta B_{\text{eff}}})^N = 2^N e^{-\frac{1}{2} \beta N q m^2 J} \cosh(\beta B_{\text{eff}}) \quad (17.2.10)$$

¹note

$$\sum_{\{s_i\}} \prod_i e^{s_i} = \left(\sum_{s_1=\pm 1} \dots \sum_{s_N=\pm 1} \right) (e^{s_1} \dots e^{s_N}) = \left(\sum_{s_1=\pm 1} e^{s_1} \right) \dots \left(\sum_{s_N=\pm 1} e^{s_N} \right) \dots = \prod_i \sum_{s_i=\pm 1} e^{s_i}$$

The magnetisation is given by

$$m = \frac{1}{N} \sum_{\{s_i\}} p_i \left(\sum_i s_i \right) = \frac{1}{Z_c N} \sum_{\{s_i\}} e^{-\beta E_{\{s_i\}}} \left(\sum_i s_i \right) \quad (17.2.11)$$

$$= \frac{1}{N\beta} \frac{\partial}{\partial B_{\text{eff}}} \ln Z_c \quad (17.2.12)$$

so using (17.2.9) we finally find that

$$m = \tanh(\beta B_{\text{eff}}) = \tanh(\beta B + \beta J q m) \quad (17.2.13)$$

This is a self-consistency equation for m , it ensures that the magnetisation we defined when imposing the mean field approximation yields its own correct value. Unfortunately it has no analytic solution, but we can still view it graphically. For a given T and B we will obtain the magnetisation value predicted by our MFT. Let us first consider the case where $B = 0$.

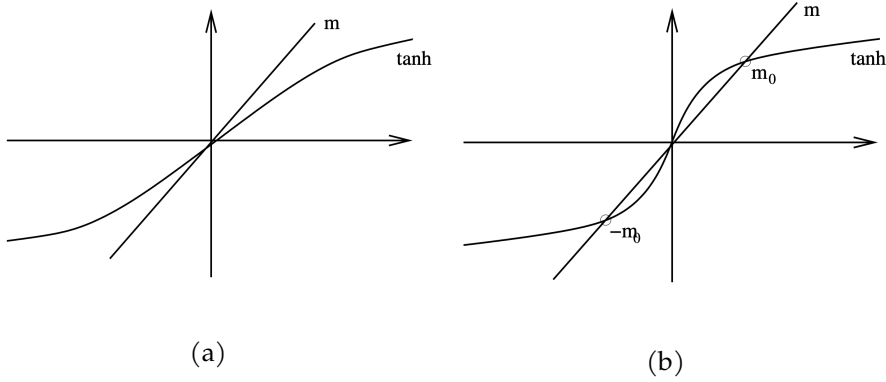


Figure 17.2. Plots of $m(T, B) = m$ and $m(T, B) = \tanh(\beta J q m)$ for (a) $Jq\beta > 1$ and (b) $Jq\beta < 1$.

Then we find that

$$m = \tanh(\beta J q m) \approx \beta J q m + \frac{1}{3}(\beta J q m)^3 \quad (17.2.14)$$

for $Jq \ll k_B T$ (high temperatures), so the slope near the origin is $\beta J q$. It is clear that if $\beta J q > 1$ then there is only one solution, $m = 0$. This implies a paramagnetic phase, the thermal fluctuations overpower the magnetic order preference of aligned spins. If instead $\beta J q < 1$ then there are three solutions, $m = \pm m_0, 0$ for some m_0 . We shall soon see that in this case the $m = 0$ case is unphysical, just like in the VdW equation, so we only have $m = \pm m_0$. This implies a ferromagnetic phase, the magnetic order preferred by the system overcomes the entropic contribution due to thermal fluctuations. As $\beta \rightarrow \infty$ we have $m_0 \rightarrow \pm m$ so we do indeed obtain perfect ferromagnetism.

There is a critical temperature where we transition between these two phases

$$k_B T_c = Jq \quad (17.2.15)$$

Now consider the case where $B \neq 0$. In the there is only one solution $m = m_0$, while in the low temperature limit we again obtain three different roots, but one will be unstable and the other meta-stable leaving only one possible magnetisation. Consequently we will

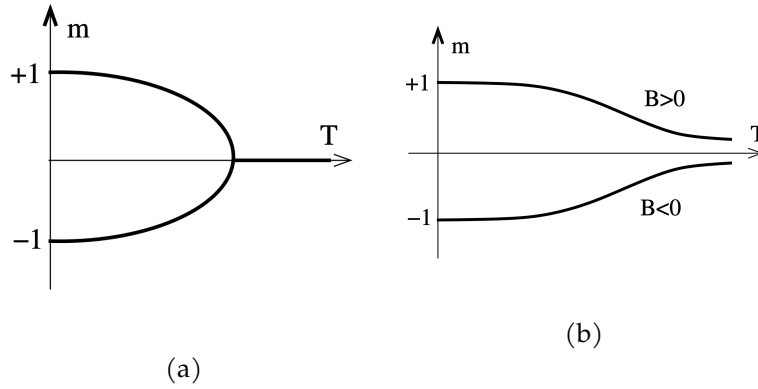


Figure 17.3. Plots of the magnetisation m given by the self-consistency equation for (a) $B = 0$ and (b) $B \neq 0$.

only ever have one solution to the self-consistency equation so there is no phase transition as we vary the temperature with fixed magnetic field. However, if we fix the temperature and vary the magnetic field from negative to positive then we will have a sudden jump from negative to positive magnetisation yielding a phase transition.

Let us try to calculate the critical exponents for the $B = 0$ phase transition. Suppose we go along one of the black lines in ?? . As T approaches T_c from below, the magnetisation m will be small, allowing us to perform a Taylor expansion

$$m \approx \beta J q m - \frac{1}{3}(\beta J q m)^3 = \frac{T_c}{T} m - \frac{1}{3} \left(\frac{T_c}{T} \right)^3 m^3, \quad T \rightarrow T_c^- \quad (17.2.16)$$

This can be rearranged into

$$\left(\frac{T_c}{T} - 1 \right) m + \frac{1}{3} \left(\frac{T_c}{T} \right)^3 m^3 = 0 \quad (17.2.17)$$

which yields the solutions

$$m = \pm \sqrt{3 \left(\frac{T}{T_c} \right)^2 \frac{T_c - T}{T_c}}, \quad T \rightarrow T_c^- \quad (17.2.18)$$

We introduce the dimensionless variable $t = \frac{T - T_c}{T_c} = T' - 1$ which allows us to write

$$m = \pm \sqrt{-3(1+t)^2 t} \approx \pm \sqrt{-3(1+2t)t} \sim |t|^{1/2} \quad (17.2.19)$$

so in other words

$$m \sim \pm (T - T_c)^{1/2} \quad (17.2.20)$$

which is the same critical exponent as in the VdW gas.

We can also sit at the critical temperature so that $\beta J q = 1$ and look at the magnetisation

as $B \rightarrow 0$. Instead of expanding (17.2.13) in m we now expand in B

$$m \approx \frac{B}{Jq} + m - \frac{1}{3}m^3 \quad (17.2.21)$$

so that

$$m \sim B^{1/3} \quad (17.2.22)$$

We can also look at the susceptibility as we approach T_c isothermally:

$$\chi = \left. \frac{\partial m}{\partial B} \right|_T \quad (17.2.23)$$

near the critical point. If we differentiate both sides of the self-consistency equation we get

$$\chi = \beta \operatorname{sech}^2(\beta B + \beta Jqm)(1 + Jq\chi) \quad (17.2.24)$$

implying that

$$\chi = \frac{\beta \operatorname{sech}^2(\beta B + \beta Jqm)}{1 - \beta Jq \operatorname{sech}^2(\beta B + \beta Jqm)} = \frac{\beta}{\cosh^2(\beta B + \beta Jqm) - \beta Jq} \quad (17.2.25)$$

If $B = 0$ this reduces to

$$\chi = \frac{\beta}{\cosh^2(\beta Jqm) - \beta Jq} \quad (17.2.26)$$

and using the fact that $m \sim |t|^{1/2}$ as $T \rightarrow T_c^-$ we can expand $\cosh x \approx 1 + x^2/2$ to find that

$$\chi \approx \frac{\beta}{1 + \beta^2 J^2 q^2 |t| - \beta Jq} \sim |t|^{-1}, \quad T \rightarrow T_c^- \quad (17.2.27)$$

since $\beta Jq \rightarrow 1$ as $T \rightarrow T_c^-$.

So, is our mean field theory approximation right? Not even remotely in $d = 1$, we predicted a phase transition at $k_B T_c = 2Jq$ whereas it turns out that there is no phase transition in one dimension! In 1D the thermal fluctuations are strong enough to forbid ordered phases from forming². In two dimensions we do have a phase transition at $k_B T_c \approx 2.27Jq$ (as opposed to $k_B T_c \approx 4Jq$ according to MFT) and the critical exponents are

$$m \sim (T_c - T)^\beta, \quad \beta = \frac{1}{8} \quad (17.2.28)$$

$$m \sim B^{1/\delta}, \quad \delta = 15 \quad (17.2.29)$$

$$\chi \sim (T - T_c)^{-\gamma}, \quad \gamma = \frac{7}{4} \quad (17.2.30)$$

We see that the mean field approximation is still bad, but at least it correctly predicts the phase transition (but not the critical temperature). However the mean field approximation is justified for large dimensions, where it makes sense for each spin to effectively feel an averaged interaction from its neighbours. In low dimensions this is simply wrong to

²this is not true for quantum phases, where the “quantumness” adds an extra dimension to our system. Bose-Einstein condensation is an example of this.

assume, but in $d \rightarrow \infty$ dimensions (whatever that means) the MFT would provide the exact solutions.

One may wonder how a discontinuity in thermodynamic quantities can occur if they are all derived from the partition function, a sum of smooth, analytic functions. In a finite system this is a finite sum so we cannot possibly get a discontinuity in the derivatives of thermodynamic potentials. In the limit where $N \rightarrow \infty$ however the sum becomes infinite and we can get a discontinuity. Thus phase transitions can only occur in the thermodynamic limit.

Finally, note that in 3D the critical exponents are

$$\beta \approx 0.32, \delta \approx 4.8, \gamma \approx 1.2 \quad (17.2.31)$$

which are the same as those in the VdW gas! This is not merely a cosmic coincidence, it turns out that when systems undergo phase transitions they lose information about their microscopic details. Systems which have the same critical exponents behave mathematically identically near the critical temperature and form universality classes. A single, universal theory can be used to describe systems which superficially look different.

17.3 Landau-Ginzburg theory: a taster

Landau's theory

At the center of Landau's theory of phase transitions is the Free energy, let us compute it for the Ising model

$$F = -\frac{1}{\beta} \ln Z = \frac{1}{2} N q m^2 J - \frac{N}{\beta} \ln(2 \cosh(\beta B_{\text{eff}})) \quad (17.3.1)$$

Suppose we treat F as a function of m , so we magically find a way to go out of equilibrium and keep m fixed and independent of T, V . We can then expand F in m which is small near the critical temperature:

$$F \approx \frac{1}{2} N q m^2 J - \frac{N}{\beta} \ln 2 - \frac{N}{\beta} \ln \left(1 + \frac{(\beta B_{\text{eff}})^2}{2} + \frac{(\beta B_{\text{eff}})^4}{24} \right) \quad (17.3.2)$$

$$\approx \frac{1}{2} N q m^2 J - \frac{N}{\beta} \ln 2 - \frac{N}{\beta} \left[\frac{(\beta B_{\text{eff}})^2}{2} + \frac{(\beta B_{\text{eff}})^4}{24} - \frac{1}{2} \left(\frac{\beta B_{\text{eff}}}{2} \right)^2 \right] \quad (17.3.3)$$

If $B = 0$ then $B_{\text{eff}} = J q m$ so that

$$F \approx \frac{1}{2} N q m^2 J - \frac{N}{\beta} \ln 2 - \frac{N}{\beta} \left[\frac{(\beta J q m)^2}{2} - \frac{(\beta J q m)^4}{12} \right] \quad (17.3.4)$$

$$= -N k_B T \ln 2 + \frac{N J q}{2} \left(1 - \frac{J q}{k_B T} \right) m^2 + \frac{N J^4 q^4}{12 k_B^3 T^3} m^4 \quad (17.3.5)$$

$$= -N k_B T \ln 2 + \frac{N k_B}{2} (T_c - T) m^2 + \frac{N k_B T_c^4}{12 T^3} m^4 \quad (17.3.6)$$

We managed to expand the free energy near the critical temperature T_c as a polynomial in m of the form

$$F = F_0 + a(T - T_c)m^2 + bm^4 \quad (17.3.7)$$

with $F_0 = -Nk_B T \ln 2$, $a = -\frac{Nk_B}{2}$ and $b = \frac{Nk_B T_c^4}{12T^3}$. Here m is what's known as an order parameter, it measures the degree of order across the boundary of a phase transition, with $m = 0$ indicating paramagnetic order while if $m \neq 0$ then we have ferromagnetic order.

Equilibrium occurs at stationary points of F which in our case are given by

$$\frac{\partial F}{\partial m} = 2a(T - T_c)m + 4bm^3 = 0 \implies m = 0 \text{ or } \sqrt{\frac{a(T_c - T)}{2b}} \quad (17.3.8)$$

The number of solutions clearly depends on the signs of a, b and the temperature T . We can assume that $b > 0$ at all temperatures then the free energy looks like

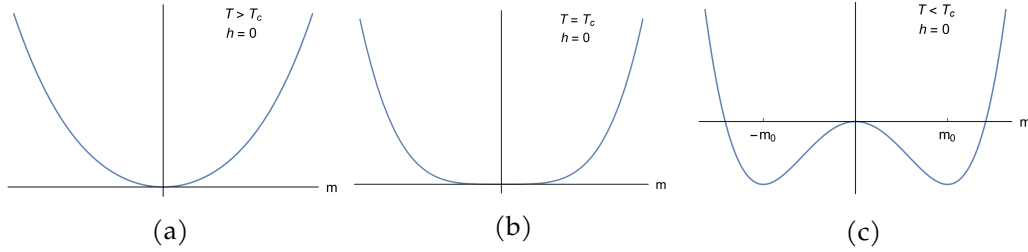


Figure 17.4. Free energy at $B = 0$ and (a) $T < T_c$ (b) $T = T_c$ (c) $T > T_c$.

We see that if $T \geq T_c$ only one solution ($m = 0$) is real and physical, representing the paramagnetic phase. If however $T \leq T_c$ then we get three solutions, the one at the origin is meta-stable and can fall back to the other two under perturbative effects. The other two solutions $m = \pm \sqrt{\frac{a(T_c - T)}{2b}}$ are stable and represents the ferromagnetic phase. This result is in agreement with the critical exponents we calculated earlier.

Now suppose that $B \neq 0$. Then the free energy reads

$$F = -Nk_B T \ln 2 + \frac{JNq}{2}m^2 - \frac{N}{2k_B T}(B + Jqm)^2 + \frac{Nk_B T_c^4}{12T^3}(B + Jqm)^4 \quad (17.3.9)$$

which is not even anymore.

Again for high temperatures we only have one minimum shifted from the origin yielding a non-zero magnetisation. Suppose we instead start at sufficiently low temperatures with $B < 0$, where we have two minima and a local maximum. The plot will not be symmetric so the minimum on the left will be lower than the other. This will be the preferred, physical magnetisation, while the other two will be either unstable (the middle magnetisation) or meta-stable (the other local minima).

As we continue to increase the magnetic field the two local minima get more and more symmetric until at $B = 0$ we get the graph discussed before. We then enter the $B > 0$ regime where the situation is flipped, the minimum on the right will be the physical

magnetisation. Here we encounter a phase transition where we suddenly transition from a negative magnetisation to a positive one, exactly as we had previously predicted.

Note that unlike before the position of the minimum transitions smoothly as we vary the temperature.

Ginzburg's addition

Landau's theory ignores the fluctuations in the system which we can take into account using the revamped Landau theory: the Landau-Ginzburg theory.

Diffusion and Brownian motion

18

18.1 Langevin equation

Part V

Statistical field theory

Path integrals

19

Landau-Ginzburg theory

20

Renormalisation group

Acknowledgments

This is the most common positions for acknowledgments. A macro is available to maintain the same layout and spelling of the heading.

Note added. This is also a good position for notes added after the paper has been written.

Bibliography

- [1] Author, *Title*, *J. Abbrev.* **vol** (year) pg.
- [2] Author, *Title*, arxiv:1234.5678.
- [3] Author, *Title*, Publisher (year).