Krzysztof A. Meissner

Statistical Physics

not for distribution

FACULTY OF
PHYSICS
UNIVERSITY
OF WARSAW

Warsaw, September 2025

These notes were written for the lectures of Statistical Physics at the Master course at the Faculty of Physics at the University of Warsaw. The notes do not refer to the original sources but they are based mostly on: L.D. Landau, J. Lifszyc ,,Statistical physics", G.H. Wannier ,,Statistical Physics" and K. Huang ,,Statistical Mechanics" but with inclusion of many other scattered sources.

Contents

1 Thermodynamics 5
- 1.1 Introduction . . . 5
- 1.2 State functions . . . 6
- 1.3 Temperature . . . 7
- 1.4 Internal energy . . . 8
- 1.5 Second Law of Thermodynamics (SLT) . . . 9

2 Carnot cycle 11
- 2.1 Carnot cycle and absolute temperature . . . 11
- 2.2 Entropy . . . 13
- 2.3 Equations of state (EoS) . . . 14

3 Thermodynamical potentials 17
- 3.1 Thermodynamical potentials . . . 17
- 3.2 Thermodynamical inequalities . . . 19
- 3.3 Maxwell relations . . . 20

4 Thermodynamical processes 23
- 4.1 Thermodynamical identities . . . 23
- 4.2 Joule-Thomson effect . . . 25
- 4.3 Entropy of an Ideal Gas . . . 26
- 4.4 Entropy of the VdW gas . . . 27
- 4.5 Entropy of the gas of photons . . . 27

5 Phase transitions 29
- 5.1 Chemical potential . . . 29
- 5.2 Clausius-Clapeyron equation . . . 30
- 5.3 Phase transition in VdW equation of state . . . 31
- 5.4 Critical point in VdW equation of state . . . 32
- 5.5 Critical exponents . . . 34

6 Classical statistical physics 37
- 6.1 Liouville theorem . . . 37
- 6.2 Poincaré recurrence theorem . . . 38
- 6.3 Liouville's equation . . . 38

6.4 Entropy 39
6.5 Derivation of the Gibbs-Boltzmann factor 40

7 Microcanonical and canonical ensembles 43
7.1 Microcanonical ensemble 43
7.2 Canonical ensemble 44

8 Grand canonical ensemble 47
8.1 Classical grand canonical ensemble 47
8.2 Quantum grand canonical ensemble 47
8.3 Planck distribution 49
8.4 Distributions at low temperatures 50

9 Imperfect gases 51
9.1 Partition function with interactions 51
9.2 Approximations of the interacting partition function 51
9.3 Cluster expansion 52
9.4 Hard disks 54

10 Debye theory of solids 57
10.1 Equipartition of energy 57
10.2 Debye theory of specific heat of solids 58
10.3 Electronic specific heat in metals 62

11 Thermodynamics of magnetic systems 65
11.1 Thermodynamics of the magnetic field 65
11.2 Ferromagnetism 68
11.3 Ising model in 1D 69
11.4 Ising model in 2D 70

1 Thermodynamics

1.1 Introduction

The enormous success of the Newton's equations with the subsequent development of electrodynamics did not touch large part of everyday's life experience like boiling of water or burning wood. Wherever there was a large number of particles (in the modern language) these equations of Newton or Coulomb had no chance of explaining any phenomena involved. In the XIXth century the effort to fill this gap was undertaken by people not connected to physics or mathematics – the beer brewer and a doctor. They noticed that the work done on any object is somehow related to heat, although the last notion was very vague. It has taken quite some time to start to define objects and find relations between work, heat transfer, temperature etc.

The first observation done millennia ago is that a body left alone in some favorable conditions attains a state that is stable (or metastable as we now know). It was a crucial observation that even if we start with a complicated state of a large macroscopic body requiring a lot of data to describe it once the stable state is attained it is independent of the path leading to it and of initial conditions and requires only a couple of numbers to describe it in the final equilibrium. The quest to identify those numbers led finally to the formulation of thermodynamics that was purely phenomenological but very successful. The insight provided by thermodynamical considerations is surprisingly rich and deep taking into account that in XIXth century even the existence of atoms was unclear, the first direct proof being provided by Smoluchowski and Einstein in 1905. We now now the microscopic content of matter and we derive the thermodynamic properties from statistical physics but the ideas borne out in the context of thermodynamics are still extremely important.

As already said in thermodynamics we assume that a given body is in an equilibrium state – no macroscopic properties of the body change with time. Of course we cannot assume that it is strictly kept since we would like to change the state in some processes. To use the thermodynamic concepts we assume that any process is so slow that the system is (approximately) always in some equilibrium or extremely close to it.

It is also (now) clear that the equilibrium state is very dynamical and interpolates between extremely large number of microstates approximately indistinguishable among themselves macroscopically. What fluctuations are sufficiently close to each other and can be treated as indistinguishable as a macrostate and how probable they are in comparison to the large fluctuations is the question that can be posed only within statistical physics and will be discussed at length later.

The general conservation of energy did not belong to the Newton's theory – in the latter only the concept of the conservation of mechanical energy for potential forces was introduced. The first quantitative observation of the possible relation of such a general nature is now called the first law of thermodynamics. It identifies some, as yet not well specified, internal energy U assumed to be the same for a given body in the same external conditions. It therefore belongs to the so called state functions, i.e. the numbers characterizing a given body in these conditions. In some instances it is possible to identify what the U is composed of – for example for a very weakly interacting gas it should be the sum of the kinetic energies of the particles of the gas.

The fundamental statement (the first law of thermodynamics) reads

$$\mathrm{d}U = \text{đ}W + \text{đ}Q \tag{1.1}$$

where U is the internal energy of the body, W is the work done **on** the body and Q is the energy in the form of heat **added to** the body. The first and second notions are (more or less) well defined, it is the third one that goes beyond the mechanical notions and definitions present in the Newton's theory. Within thermodynamics the best approach is to treat this equation as a definition of the energy transfer in the form of heat i.e. $\text{đ}Q = \mathrm{d}U - \text{đ}W$.

It is important to emphasize the usage of đ on the right-hand side of this equation. The d symbol usually used un the theory of forms means a differential and acting on an n-form produces an $(n+1)$-form. On the LHS we have this symbol since we assume that the internal energy U is a well defined function. However, neither the work nor the heat transfer are well defined state functions so we cannot treat their change as differentials and that's why we use a different symbol to denote their contribution to the change of internal symmetry.

1.2 State functions

We will now discuss possible candidates for the state functions characterizing any equilibrium. To be more concrete we discuss a gas in a container. Once the system 'settles down' there is a number of properties that are rather obvious as state functions. This 'settling down' when no further macroscopic changes are observed is called *thermodynamical equilibrium*. Although the system is composed of extremely large number of particles the thermodynamical equilibrium is characterized by only a few numbers (state functions).

We distinguish extensive and intensive state functions. The first ones are proportional to the size of the system; if we add two identical systems and neglect their mutual interactions an extensive function would double its value. As an example one could quote the internal energy U or the volume of the gas V.

On the other hand an intensive state function for such a doubled system remains the same as in each of the individual systems. As an example one can quote here the pressure or the temperature (although the definition of the latter remains to be provided).

We have to always compare intensive with intensive and extensive with extensive state functions.

If we have a state function, say V, it is legitimate to use $\mathrm{d}V$ as a differential of the state function. Since pressure p is another state function for the extremely slow, reversible processes (we will discuss these conditions later on) we can write

$$đW = -p\,\mathrm{d}V \tag{1.2}$$

where the minus sign comes from our definition that $đW$ is the work performed on the body (if the volume grows it is the body that performs work on the environment). In the language of differential forms p is an integrating factor since it brings $đW$ to a bona fide differential

$$-\frac{đW}{p} = \mathrm{d}V \tag{1.3}$$

We will now discuss the 'thermal' part of (1.1).

1.3 Temperature

Although the notion of being 'hotter' or 'colder' is known for millennia the more precise definition had to wait until mid XIX century. It is the core of thermodynamics so we have to discuss it in great detail. As we emphasized the heat transfer $đQ$ goes beyond the Newton's theory so we have to resort to a different reasoning.

The notion of 'temperature' as state function is based on the following phenomenological observation. A body left alone homogenizes 'the temperature' (whatever it is) in such a way that there are no internal macroscopic heat transfers. If two bodies are in a diabatic (i.e. allowing for heat transfers) contact then, after some time, the common 'temperature' for both bodies is settled and there are no further heat transfers. Two bodies have the same 'temperature' if after bringing them to a diabatic contact there are no macroscopic heat transfers – then we say that they are in *thermal equilibrium*. On the basis of these observations a 0th Law of Thermodynamics was proposed:

if two bodies A and B are in thermal equilibrium and also B and C are in thermal equilibrium then also A and C are in thermal equilibrium

This simple and intuitive fact leads to the notion of 'temperature' as a state function.Imagine that the body A is in thermal contact with a chosen fixed body K. Then we change the state functions of the body A (like its volume or pressure) in such a way that the thermal contact with K does not spoil the thermal equilibrium with K – we call such a hypersurface *isotherm*. Denoting these state functions by x_i^A we can write

$$\phi^A(x_i^A) = t \tag{1.4}$$

where the function ϕ_A is such that in spite of changing x_i along the isotherm its value does not change. The condition can be written as

$$\mathrm{d}\phi^A = \sum_i \partial_i \phi^A \mathrm{d}x_i^A = 0 \tag{1.5}$$

when $\mathrm{d}x_i^A$ are tangent to the isotherm but is non vanishing when they are not. Then we can call ϕ^A an empirical temperature in the vicinity of t. This procedure does not specify which choice of ϕ is the most 'physical' or the most convenient. It however suggests the procedure how to define empirical temperature for $t+\Delta t$: we choose $\mathrm{d}x_i^A$ orthogonally to the isotherm t, calculate new $t+\Delta t$, then find a new body K' that is in thermal equilibrium with new $\phi^A\prime$ and repeat the procedure with K' and $t+\Delta t$ extending the definition to $t+2\delta t$ and so on. We have the freedom to reparametrize the empirical temperature $t \to f(t)$ as long as the derivative is everywhere positive. Since this procedure is neither unique nor the most useful, below we will introduce the notion of 'absolute temperature' based on the Carnot cycles.

1.4 Internal energy

We will now discuss the notion of internal energy U used already in the First Law. We imagine a gas in a container that is adiabatic i.e. to a good approximation it does not allow for any heat transfers so that $\not{d}Q = 0$. However, there is a possibility to provide or extract energy to or from the container by means of a piston. Then the First Law reads

$$\mathrm{d}U = \not{d}W \tag{1.6}$$

and it seems straightforward to measure changes in the internal energy U since we can measure very precisely the work done or extracted to or from the system.

However, we must be careful. Imagine two situations: the first one when we move the piston extremely slowly going through intermediate equilibrium states. Then indeed the pressure is well defined all along and we can use the formula $\not{d}W = -p\mathrm{d}V$. The second situation is when we violently jerk the piston. Then there will be sound waves excited in the gas and to bring the equilibrium to the system these waves should be absorbed by the walls of the container. Therefore even if we measure precisely the external work done in the process we don't know exactly how much energy was absorbed by gas as an internal energy and how much by the walls. We usually assume that the container is so big that the bulk energy absorbed is much bigger than the energy absorbed by the walls so this absorption is negligible. It may indeed be so in most situations but we have to remember about subtleties connected to eventual non-equlibrium intermediate states and the exact energy bilance.

Barring these difficulties we can write for a situation with the walls non-permeable to heat transfer

$$\Delta U = \Delta W \tag{1.7}$$

what allows us to determine the differences of the internal energy of a given system. If we additionally add the condition that $U \to 0$ when the temperature goes to 0 then we can introduce the notion of 'absolute' internal energy. It is a state function so it does not depend how we arrived at given conditions but for a given body only on the actual pressure and volume or volume and temperature.

1.5 Second Law of Thermodynamics (SLT)

We will now introduce the notion of entropy as a core notion both in thermodynamics and in statistical physics. The definition in statistical physics relates the entropy with the number of microstates realizing the same (within some interval) macrostate - we will discuss it in great detail later on since it is of fundamental importance. In thermodynamics, as a phenomenological theory, such an approach is impossible.

It was a great achievement in the middle of XIXth century to find the relevant functions of state and the integrating factor (in the modern language) for $đQ$. As we said this notion is based on the First Law (in some specific, controlled situations)

$$đQ = \mathrm{d}U - đW \tag{1.8}$$

and it defines the heat transfer.

We confine ourselves again to extremely slow (quasi-static) processes. Rudolf Clausius in 1850 formulated a law, the celebrated Second Law of Thermodynamics SLT(C)
It is impossible to devise an engine which, working in a cycle, shall produce no effect other than the transfer of heat from a colder body to a hotter body
usually quoted in the equivalent later formulation in 1851 by Kelvin SLT(K):
It is impossible to devise an engine which, working in a cycle, would produce no effect other than the extraction of heat from a reservoir and performance of an equivalent amount of mechanical work

We will see in a moment how these formulations are related to the most famous statement

Entropy never decreases

It is important to introduce now an idea much earlier than the SLT which was actually a motivation to introduce SLT.

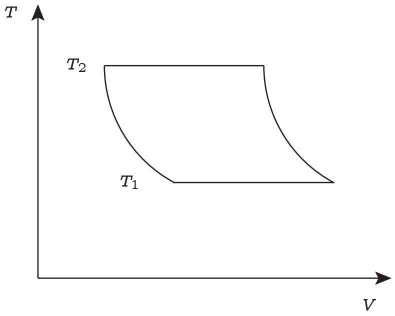
2 Carnot cycle

2.1 Carnot cycle and absolute temperature

In 1824, much earlier than the introduction of the notion of entropy and the Second Law of Thermodynamics, Sadi Carnot introduced a cycle (engine) if we have at our disposal two reservoirs of different temperatures T_1 and T_2 ($T_2 > T_1$).

The cycle is composed of 4 quasi-static processes:

- contact with reservoir with temperature T_2 (isothermic expansion)
- adiabatic expansion from T_2 to T_1
- contact with reservoir with temperature T_1 (isothermic contraction)
- adiabatic contraction from T_1 to T_2



The work done by the container after the cycle (since the internal energy returns to its original value)

$$W = Q_2 - Q_1 \tag{2.1}$$

where Q_2 is the heat taken from reservoir 2 and Q_1 is the heat given to reservoir 1. Therefore the efficiency i.e the ratio of the work to the absorbed heat is equal to

$$\eta = \frac{W}{Q_2} = 1 - \frac{Q_1}{Q_2} \tag{2.2}$$

There are two facts that follow from SLT.

It was experimentally obvious for Carnot but was proven only later with SLT that if $W > 0$ then also $Q_2, Q_1 > 0$ (and, respectively, if $W < 0$ then also $Q_2, Q_1 < 0$). Indeed, assume that $W, Q_2 > 0$ and $Q_1 < 0$ with $T_2 > T_1$. Then we take Q_2 from reservoir 2 and return $W = Q_2 - Q_1 > Q_2$ in the form of heat to the same reservoir – the only net effect is to transfer heat $-Q1$ from the colder reservoir 1 to the hotter reservoir 2 and this is impossible according to SLT(C).

The second fact is that the efficiency (2.2) is the maximal one for any engine absorbing heat Q_2 and giving away heat Q_1. Imagine that we have two such engines, one Carnot and the other arbitrary, both reversible and both working between temperatures T_2 and T_1, $T_2 > T_1$. Then

$$W = Q_2 - Q_1, \qquad \tilde{W} = \tilde{Q}_2 - \tilde{Q}_1 \tag{2.3}$$

If we find such N and $\tilde{N}$ that

$$\frac{Q_2}{\tilde{Q}_2} = \frac{\tilde{N}}{N} \tag{2.4}$$

then running both engines backwards, respectively N times and $\tilde{N}$ times, we get

$$W_{tot} = \tilde{N}\tilde{W} - NW = (\tilde{N}\tilde{Q}_2 - NQ_2) - (\tilde{N}\tilde{Q}_1 - NQ_1) = -(\tilde{N}\tilde{Q}_1 - NQ_1) \tag{2.5}$$

But the work cannot be positive according to SLT(K) so that

$$\tilde{N}\tilde{Q}_1 - NQ_1 \geqslant 0 \Rightarrow Q_2\tilde{Q}_1 - \tilde{Q}_2 Q_1 \geqslant 0 \Rightarrow \frac{\tilde{Q}_1}{\tilde{Q}_2} \geqslant \frac{Q_1}{Q_2} \tag{2.6}$$

hence

$$1 - \frac{Q_1}{Q_2} \geqslant 1 - \frac{\tilde{Q}_1}{\tilde{Q}_2} \tag{2.7}$$

what was to be proven. The proof also shows that all Carnot engines working between the same temperatures have the same efficiency.

Based on these facts about the Carnot cycle one introduces the absolute temperature T by the relation

$$\frac{Q_1}{T_1} = \frac{Q_2}{T_2} \tag{2.8}$$

that connects any two reservoirs by the Carnot cycle. Then we have the efficiency of the Carnot cycle

$$\eta = 1 - \frac{T_1}{T_2} \tag{2.9}$$

Defining the temperature of some reservoir, connecting it to any other reservoir by a Carnot cycle and measuring the heat transfers we can theoretically measure the temperature of any reservoir.

The defined temperature is usually taken as the triple point of water (point of equilibrium of vapor, liquid and ice) that occurs at the pressure 611.73 Pa – then the corresponding temperature is defined as 273.16 K. Such a definition gives the so called absolute temperature measured in Kelvin. It corresponds approximately to more frequently used 273.15 K at 1 atm (101 325 Pa).

There are many types of 'thermometers', easier to use than the Carnot cycle – from the point of view of thermodynamics the best are gaseous ones since they approximately are proportional to the absolute temperature.

It is important to note that such a temperature scale has a point '0' where the efficiency of the Carnot cycle is equal to 1 – the necessity of the existence of such an 'absolute zero' in temperature was discovered by William Thomson (lord Kelvin) in 1848, hence the name of the unit of absolute temperature. Such a point cannot be attained since extracting the heat is more and more difficult when we approach this point. The physical reason for the existence of such a point can be justified only within statistical physics.

2.2 Entropy

Using the relation (2.8) and SLT(K) Clausius has shown that for an arbitrary cycle between a given reservoir of temperature T_0 and many others with temperatures T_i we have along the cycle

$$\sum_i \frac{Q_i}{T_i} \leqslant 0 \tag{2.10}$$

since along the cycle the absorbed heat from the given reservoir is equal to

$$Q = \sum_i Q_i = T_0 \sum_i \frac{Q_i}{T_i} \tag{2.11}$$

and by SLT(K) it has to be negative.

If we additionally assume that the process is reversible then we can run the cycle in the opposite way getting the result

$$\sum_i \frac{Q_i}{T_i} = 0 \tag{2.12}$$

Therefore in a sequence of quasistatic, reversible processes from state A to state B the quantity

$$\sum_i \frac{đQ_i}{T_i} \tag{2.13}$$

is independent of the path form A to B. It suggested that there exists a state function that Clausius in 1865 called 'entropy' and denoted by S. The name was coined from Greek *en-* 'in' and *tropos* 'change,transformation' from Praindoeuropean *trep-* 'turn'; hence for example 'troposphere' the lowest part of the atmosphere rotating with the Earth, 'psychotropic' 'changing the mind', 'tropic' 'where the Sun changes direction'. It is interesting to note the origin of 'energy' - it comes from Greek *en-* and *ergon* 'work' from PIE 'werg'– hence for example allergy ('allos' strange), argon ('not working'), lethargy ('lethe' forgetfulness'); organ; and work itself.

Therefore for quasistatic, reversible processes we can write

$$\frac{\text{đ}Q}{T} = \mathrm{d}S \tag{2.14}$$

and the first Law reads

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V \tag{2.15}$$

It is important to emphasize that the First Law in this form is valid for all quasistatic processes but for not reversible ones $T\mathrm{d}S$ does not have the meaning of the transferred heat and $-p\mathrm{d}V$ does correspond to the performed work.

The fundamental role of entropy as a state function will be discussed in detail in quantum statistical physics later on since only there it has a deep physical meaning – in thermodynamics its meaning is rather vague and unclear.

2.3 Equations of state (EoS)

The First Law for a fixed number of particles

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V \tag{2.16}$$

suggests that there are two independent state functions and all the other depend on them. It means that on top of U and S there should be a third relation involving p, V and T. This relation is called the equation of state. We will derive later on the equation of state for the ideal gas of noninteracting particles (we assume that the number of particles N is kept fixed)

$$pV = NkT \tag{2.17}$$

where N is the number of particles and $k := 1.380649 \cdot 10^{-23}$ J/K is the defined value of the Boltzmann's constant. Introducing the more practical measure of the number of particles by defining a mole N_A

$$N_A := 6.02214076 \cdot 10^{23} \tag{2.18}$$

we rewrite (2.17) as

$$pV = nRT \tag{2.19}$$

where n is expressed in moles and the gas constant $R := N_A k := 8.31446261815324$ J/K/mol.

Including the other real factors like the interactions leads to more complicated equations. The most famous is the Van der Waals equation that takes into account both the interactions and the volume of the particles themselves – it reads

$$\left(p + \frac{an^2}{V^2}\right)(V - nb) = nRT \tag{2.20}$$

b represents the volume of one mole of particles. It was proposed by Johannes Van der Waals in 1873 in his PhD thesis at the University of Leiden – he postulated the existence of molecules and the forces between them. At that time the idea of molecules was generally rejected but the VdW equation and the explanation of liquid-gas phase transition raised enthusiastic reaction for example from James C. Maxwell. The first actual proof of the existence of molecules was given 30 years later by Smoluchowski and Einstein from the theoretical explanation of the Brownian motion and the dipole-dipole long range forces between molecules were given the name Van der Waals forces. To justify the introduction of a as a result of intermolecular interactions we rewrite (2.20) as

$$p = \frac{nRT}{V - nb} - \frac{an^2}{V^2} = \frac{RT}{v - b} - \frac{a}{v^2} \tag{2.21}$$

where v is the molar volume. We see that positive a decreases the pressure by a factor proportional to the number of pairwise attractive interactions between particles.

We will discuss this equation in detail in the context of liquid-gas phase transitions although VdW equation of state is rather inaccurate in the liquid part (for small v). For water in the gaseous state

$$a \sim 0.55 \ \mathrm{Pa} \cdot \mathrm{m}^6/\mathrm{mol}^2, \quad b \sim 3 \cdot 10^{-5} \ \mathrm{m}^3/\mathrm{mol} \tag{2.22}$$

so it is clear that the VdW EoS cannot be used in the liquid part since $v \sim 1.8 \cdot 10^{-5}$ m^3/mol is less than b.

3 Thermodynamical potentials

3.1 Thermodynamical potentials

The first Law of Thermodynamics written in the form

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V \tag{3.1}$$

allows to use the powerful language of differential forms to derive many, sometimes very nontrivial, identities among physically measured quantities. We start from using Legendre transform to introduce new state functions than the ones used in FLT. At this point we assume that the number of particles in the system is fixed – a possibility of the exchange of particles with the reservoir will be discussed separately later.

1. **Free energy F**
   We define
   $$F := U - TS \tag{3.2}$$
   Then
   $$\mathrm{d}F = \mathrm{d}U - \mathrm{d}TS - T\mathrm{d}S = -S\mathrm{d}T - p\mathrm{d}V \tag{3.3}$$
   so that the independent variables for F are temperature and volume. Temperature is usually much more easily controlled than entropy so F is usually more useful than the internal energy U. We have
   $$\left(\frac{\partial F}{\partial T}\right)_V = -S, \qquad \left(\frac{\partial F}{\partial V}\right)_T = -p \tag{3.4}$$
   For constant temperature we can write SLT as
   $$\frac{\Delta Q}{T} \leqslant \Delta S \tag{3.5}$$
   Hence
   $$\Delta F = \Delta U - T\Delta S \leqslant \Delta U - \Delta Q = \Delta W = -W \Rightarrow W \leqslant -\Delta F \tag{3.6}$$
   where W is the amount of work that the system can perform and the equality is for a reversible process. For $W = 0$ we have the statement that F *for an isolated*

system at constant temperature does not grow and therefore the equilibrium at constant temperature is attained for minimal F. If two systems are in equilibrium and in contact at a fixed temperature and their volumes are left free (but $\mathrm{d}V_1 = -\mathrm{d}V_2$) then

$$0 = \mathrm{d}F = -p_1\mathrm{d}V_1 - p_2\mathrm{d}V_2 = -(p_1 - p_2)\mathrm{d}V_1 \Rightarrow p_1 = p_2 \tag{3.7}$$

leading to an (obvious) conclusion that their pressures have to be equal.

2. **Enthalpy H**
We define

$$H := U + pV \tag{3.8}$$

Then

$$\mathrm{d}H = \mathrm{d}U + \mathrm{d}pV + p\mathrm{d}V = T\mathrm{d}S + V\mathrm{d}p \tag{3.9}$$

so that the independent variables for H are entropy and pressure. Enthalpy is very often used in chemistry since we often interested in the released heat of a given reaction at a fixed pressure. From the 0LT we have under constant pressure p

$$Q = U_f - U_i + p(V_f - V_i) = H_f - H_i \tag{3.10}$$

i.e. the absorbed or released heat is a difference of the final and initial enthalpies.

We have

$$\left(\frac{\partial H}{\partial S}\right)_p = T, \qquad \left(\frac{\partial H}{\partial p}\right)_S = V \tag{3.11}$$

3. **Free enthalpy (Gibbs potential) G**
We define

$$G := U - TS + pV \tag{3.12}$$

Then

$$\mathrm{d}G = \mathrm{d}U - \mathrm{d}TS - T\mathrm{d}S + \mathrm{d}pV + p\mathrm{d}V = -S\mathrm{d}T + V\mathrm{d}p \tag{3.13}$$

so that the independent variables for G are temperature and pressure. It is important to emphasize that G is an extensive quantity so it cannot depend only on intensive quantities – it still depends on the number of particles N not written up to now explicitly.

For constant temperature and pressure we can write SLT as

$$\frac{\Delta Q}{T} \leqslant \Delta S \tag{3.14}$$

Hence

$$\Delta G = \Delta U - T\Delta S + p\Delta V \leqslant \Delta U - \Delta Q + p\Delta V = \Delta W + p\Delta V = 0 \tag{3.15}$$

where the equality is for a reversible process. Then we have the statement that *G for an isolated system at constant temperature and pressure does not grow* and therefore the equilibrium in these conditions is attained for minimal G. We have

$$\left(\frac{\partial G}{\partial T}\right)_p = -S, \qquad \left(\frac{\partial G}{\partial p}\right)_T = V \tag{3.16}$$

3.2 Thermodynamical inequalities

We assume that the system has fixed temperature T and pressure p. Then Gibbs potential G has a minimum in equilibrium. We therefore have the condition that all deviations from the equilibrium should lead to a growth of G (δ denotes small deviations but not infinitesimally small)

$$\delta G = \delta U - T\delta S + p\delta V > 0 \tag{3.17}$$

The two last terms are exact and we expand δU up to quadratic deviations in δS and δV

$$\delta U = \frac{\partial U}{\partial S}\delta S + \frac{\partial U}{\partial V}\delta V + \frac{1}{2}\left(\frac{\partial^2 U}{\partial S^2}\delta S^2 + 2\frac{\partial^2 U}{\partial S \partial V}\delta S \delta V + \frac{\partial^2 U}{\partial V^2}\delta V^2\right) + \dots \tag{3.18}$$

The linear parts cancel in (3.17) (since G is in the minimum) and the matrix of quadratic parts should be a positive matrix what gives two conditions.

The first one reads

$$\frac{\partial^2 U}{\partial S^2} = \left(\frac{\partial T}{\partial S}\right)_V = \frac{T}{T\left(\frac{\partial S}{\partial T}\right)_V} = \frac{T}{C_V} > 0 \tag{3.19}$$

i.e. specific heat at constant volume has to be positive.

The second condition of a positive determinant reads

$$\frac{\partial^2 U}{\partial S^2}\frac{\partial^2 U}{\partial V^2} - \left(\frac{\partial^2 U}{\partial S \partial V}\right)^2 > 0 \tag{3.20}$$

It can be written as a jacobian

$$\frac{\partial\left(\frac{\partial U}{\partial S}, \frac{\partial U}{\partial V}\right)}{\partial(S,V)} = \frac{\partial(T,-p)}{\partial(S,V)} > 0 \tag{3.21}$$

We transform this inequality

$$-\frac{\partial(T,p)}{\partial(S,V)} = -\frac{\frac{\partial(T,p)}{\partial(T,V)}}{\frac{\partial(S,V)}{\partial(T,V)}} > 0 \tag{3.22}$$

and hence

$$\frac{\left(\frac{\partial p}{\partial V}\right)_T}{\left(\frac{\partial S}{\partial T}\right)_V} = \frac{T}{C_V}\left(\frac{\partial p}{\partial V}\right)_T < 0 \tag{3.23}$$

i.e.

$$\left(\frac{\partial p}{\partial V}\right)_T < 0 \tag{3.24}$$

so pressure has to decrease with the growth of volume.

3.3 Maxwell relations

Assume that we have 3-dim manifold with a hypersurface defined by $f(x,y,z) = 0$ and we would like to derive some identities between the partial derivatives wrt to different pairs of variables (since only 2 are independent) – they are extensively used in thermodynamics.

We start with

$$\begin{aligned} \mathrm{d}x &= \left(\frac{\partial x}{\partial y}\right)_z \mathrm{d}y + \left(\frac{\partial x}{\partial z}\right)_y \mathrm{d}z \\ \mathrm{d}y &= \left(\frac{\partial y}{\partial x}\right)_z \mathrm{d}x + \left(\frac{\partial y}{\partial z}\right)_x \mathrm{d}z \end{aligned} \tag{3.25}$$

Plugging dy from the second equation into the first we get

$$\left(\frac{\partial x}{\partial y}\right)_z = \frac{1}{\left(\frac{\partial y}{\partial x}\right)_z} \tag{3.26}$$

and the triple product formula

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1 \tag{3.27}$$

The minus sign may seem a little surprising - we may illustrate it with an example of a sphere (we assume below $x, y, z > 0$)

$$x^2 + y^2 + z^2 = R^2 \Rightarrow \left(\frac{\partial x}{\partial y}\right)_z = -\frac{y}{x}, \left(\frac{\partial y}{\partial z}\right)_x = -\frac{z}{y}, \left(\frac{\partial z}{\partial x}\right)_y = -\frac{x}{z} \tag{3.28}$$

Similarly we can write

$$\begin{aligned} \mathrm{d}x &= \left(\frac{\partial x}{\partial y}\right)_z \mathrm{d}y + \left(\frac{\partial x}{\partial z}\right)_y \mathrm{d}z \\ \mathrm{d}x &= \left(\frac{\partial x}{\partial y}\right)_w \mathrm{d}y + \left(\frac{\partial x}{\partial w}\right)_y \mathrm{d}w \end{aligned} \tag{3.29}$$

Writing

$$\mathrm{d}w = \left(\frac{\partial w}{\partial y}\right)_z \mathrm{d}y + \left(\frac{\partial w}{\partial z}\right)_y \mathrm{d}z \tag{3.30}$$

and plugging into the previous equation we get

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial x}{\partial y}\right)_w + \left(\frac{\partial x}{\partial w}\right)_y \left(\frac{\partial w}{\partial y}\right)_z \tag{3.31}$$

and

$$\left(\frac{\partial x}{\partial z}\right)_y = \left(\frac{\partial x}{\partial w}\right)_y \left(\frac{\partial w}{\partial z}\right)_y \tag{3.32}$$

We now introduce the so called Maxwell relations.
We start (abstractly) with the relation

$$\mathrm{d}U = T\,\mathrm{d}S - p\,\mathrm{d}V \tag{3.33}$$

Differentiating with d we get

$$\mathrm{d}T \wedge \mathrm{d}S = \mathrm{d}p \wedge \mathrm{d}V \tag{3.34}$$

If all these quantities depend on two variables x and y then

$$\left(\frac{\partial T}{\partial x}\mathrm{d}x + \frac{\partial T}{\partial y}\mathrm{d}y\right) \wedge \left(\frac{\partial S}{\partial x}\mathrm{d}x + \frac{\partial S}{\partial y}\mathrm{d}y\right) = \left(\frac{\partial p}{\partial x}\mathrm{d}x + \frac{\partial p}{\partial y}\mathrm{d}y\right) \wedge \left(\frac{\partial V}{\partial x}\mathrm{d}x + \frac{\partial V}{\partial y}\mathrm{d}y\right) \tag{3.35}$$

Gathering the coefficients we get

$$\left(\frac{\partial T}{\partial x}\right)_y \left(\frac{\partial S}{\partial y}\right)_x - \left(\frac{\partial p}{\partial x}\right)_y \left(\frac{\partial V}{\partial y}\right)_x = \left(\frac{\partial T}{\partial y}\right)_x \left(\frac{\partial S}{\partial x}\right)_y - \left(\frac{\partial p}{\partial y}\right)_x \left(\frac{\partial V}{\partial x}\right)_y \tag{3.36}$$

Choosing x and y as pairs out of (T, S, p, V) we get 6 Maxwell relations. For example for the pair (T, V) we get

$$\left(\frac{\partial S}{\partial V}\right)_T - \left(\frac{\partial p}{\partial T}\right)_V = 0 \tag{3.37}$$

while for the pair (T, p) we get

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p \tag{3.38}$$

Another example for the pair p, V:

$$\left(\frac{\partial T}{\partial p}\right)_V \left(\frac{\partial S}{\partial V}\right)_p - 1 = \left(\frac{\partial T}{\partial V}\right)_p \left(\frac{\partial S}{\partial p}\right)_V \tag{3.39}$$

4 Thermodynamical processes

4.1 Thermodynamical identities

It is important to connect the thermodynamical notions like for example specific heats to the directly measurable quantities like isothermic compressibility.

From FLT we have

$$\mathrm{d}S = \frac{C_V}{T}\mathrm{d}T + \frac{1}{T}\left[\left(\frac{\partial U}{\partial V}\right)_T + p\right]\mathrm{d}V \tag{4.1}$$

where

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V \tag{4.2}$$

Acting with d we get

$$\left(\frac{\partial}{\partial V}\right)_T \left(\frac{C_V}{T}\right)\mathrm{d}V \wedge \mathrm{d}T = \left(\frac{\partial}{\partial T}\right)_V \left[\frac{1}{T}\left(\frac{\partial U}{\partial V}\right)_T + \frac{p}{T}\right]\mathrm{d}V \wedge \mathrm{d}T \tag{4.3}$$

what leads to (using $\frac{\partial^2 U}{\partial V \partial T} = \frac{\partial^2 U}{\partial T \partial V}$)

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p \tag{4.4}$$

If we plug in the ideal gas EoS then we get

$$\left(\frac{\partial U}{\partial V}\right)_T = 0 \tag{4.5}$$

so the internal energy of an ideal gas depends only on temperature (and obviously the number of particles). For the Van der Waals equation (4.30) we get a non zero result

$$\left(\frac{\partial U}{\partial V}\right)_T = \frac{an^2}{V^2} \Rightarrow \frac{U}{n} = B(T) - a\,\frac{n}{V} \tag{4.6}$$

so the specific internal energy for a given temperature decreases with density (because of attractive forces between the particles).

We derive more such identities. From $S = -\left(\frac{\partial F}{\partial T}\right)_V$ we get

$$\left(\frac{\partial C_V}{\partial V}\right)_T = T\frac{\partial^2 S}{\partial V \partial T} = -T\frac{\partial^2}{\partial T^2}\left(\frac{\partial F}{\partial V}\right)_T = T\frac{\partial^2 p}{\partial T^2} \tag{4.7}$$

Returning to the FLT we can write

$$T\mathrm{d}S = C_V \mathrm{d}T + T\left(\frac{\partial p}{\partial T}\right)_V \mathrm{d}V \tag{4.8}$$

On the other hand starting from

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}H - \frac{V}{T}\mathrm{d}p = \frac{1}{T}\left(\left(\frac{\partial H}{\partial T}\right)_p \mathrm{d}T + \left(\frac{\partial H}{\partial p}\right)_T \mathrm{d}p\right) - \frac{V}{T}\mathrm{d}p \tag{4.9}$$

we get from $\mathrm{dd}S = 0$ ($\left(\frac{\partial H}{\partial T}\right)_p = C_p$)

$$\left(\frac{\partial H}{\partial p}\right)_T = V - T\left(\frac{\partial V}{\partial T}\right)_p \tag{4.10}$$

and then

$$T\mathrm{d}S = C_p \mathrm{d}T - T\left(\frac{\partial V}{\partial T}\right)_p \mathrm{d}p \tag{4.11}$$

Introducing

$$\begin{aligned} \alpha &:= \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p \quad \text{(coefficient of thermal expansion)} \\ \kappa_T &:= -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \quad \text{(coefficient of isothermal compressibility)} \end{aligned}$$

and using

$$\left(\frac{\partial p}{\partial T}\right)_V = -\frac{1}{\left(\frac{\partial T}{\partial V}\right)_p \left(\frac{\partial V}{\partial p}\right)_T} = -\frac{\left(\frac{\partial V}{\partial T}\right)_p}{\left(\frac{\partial V}{\partial p}\right)_T} = \frac{\alpha}{\kappa_T} \tag{4.12}$$

we can write

$$\begin{aligned} T\mathrm{d}S &= C_V \mathrm{d}T + \frac{\alpha T}{\kappa_T}\mathrm{d}V \\ T\mathrm{d}S &= C_p \mathrm{d}T - \alpha T V \mathrm{d}p \end{aligned}$$

Subtracting these two equations and treating (p, V) as independent variables we get as a coefficient in front of $\mathrm{d}V$

$$(C_p - C_V)\left(\frac{\partial T}{\partial V}\right)_p - \frac{\alpha T}{\kappa_T} = 0 \tag{4.13}$$

and hence

$$C_p - C_V = \frac{\alpha^2 T V}{\kappa_T} \tag{4.14}$$

so it is positive.

4.2 Joule-Thomson effect

If we decompress the ideal gas into the vacuum without any heat exchange its temperature does not change. However, for real gases the temperature changes and that effect was discovered by Joule and Thomson (lord Kelvin) in 1852. It is an irreversible process with constant enthalpy what can be seen by the following reasoning. We have two parts of the container with higher pressure p_1 on one side and lower pressure p_2 on the other. If we push a small portion of gas V_1 from one part to the other the work performed on the portion in the first part is $p_1 V_1$ while the work performed by the portion in the second part is $p_2 V_2$. The heat is equal to

$$Q = \Delta U - \Delta W = (U_2 - U_1) - (p_1 V_1 - p_2 V_2) = (U_2 + p_2 V_2) - (U_1 + p_1 V_1) = H_2 - H_1 \quad (4.15)$$

Since Q is assumed to vanish we get $H_2 = H_1$ so the enthalpy is constant in the process.

We want to calculate the rate of change of temperature with pressure at constant enthalpy:

$$\nu_{JT} := \left(\frac{\partial T}{\partial p}\right)_H \quad (4.16)$$

Using a triple product identity we rewrite it as

$$\left(\frac{\partial T}{\partial p}\right)_H = -\frac{\left(\frac{\partial H}{\partial p}\right)_T}{\left(\frac{\partial H}{\partial T}\right)_p} \quad (4.17)$$

Since

$$\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}p = T\left(\left(\frac{\partial S}{\partial T}\right)_p \mathrm{d}T + \left(\frac{\partial S}{\partial p}\right)_T \mathrm{d}p\right) + V\mathrm{d}p \quad (4.18)$$

we get

$$\left(\frac{\partial H}{\partial T}\right)_p = C_p \quad (4.19)$$

Therefore (using (4.10) and the coefficient of thermal expansion $\alpha = \left(\frac{\partial V}{\partial T}\right)_p / V$)

$$\nu_{JT} = \frac{V(\alpha T - 1)}{C_p} \quad (4.20)$$

For an ideal gas it is identically 0 but for real gases it is not and even changes sign. At room temperature all gases (except hydrogen, helium, and neon) cool upon expansion ($\nu_{JT} > 0$) but ν_{JT} decreases with growing temperature and at some T_{inv} (called the inversion temperature) it changes sign and above this temperature the gas warms with decreasing pressure. Nitrogen has inversion temperature 621 K (348 °C), oxygen 764 K (491 °C) and hydrogen 202 K (-71 °C).

4.3 Entropy of an Ideal Gas

We know the equation of state and the internal energy of an ideal gas

$$pV = nRT, \qquad U(T,V) = c_v nT \tag{4.21}$$

We have

$$\begin{aligned} \left(\frac{\partial U}{\partial T}\right)_V &= T\left(\frac{\partial S}{\partial T}\right)_V \\ T\left(\frac{\partial S}{\partial V}\right)_T &= p + \left(\frac{\partial U}{\partial V}\right)_T \end{aligned} \tag{4.22}$$

Using $\left(\frac{\partial U}{\partial V}\right)_T = 0$ and solving these equations we get

$$S(T,V) = c_v n \ln T + nR \ln\left(\frac{V}{n}\right) + \text{const} \tag{4.23}$$

where we added n under $\ln V$ to make S extensive. It was an argument of Gibbs that without adding $1/n$ under the logarithm of V dividing (mentally) a container into two parts we would get a different result than for the undivided container what is nonsensical. We will derive this equation (together with the value of the constant) as the so called Sackur-Tetrode equation in the framework of the classical statistical theory and discuss it at length later on since it was a hint on the quantum nature of matter long before quantum mechanics was born. On the other hand it cannot be the full story since the entropy diverges as $T \to 0$ – quantum statistical theory corrects this expression to avoid such a conclusion.

We can also write the formula for the entropy of an ideal gas as a function of T and p

$$S(T,p) = (c_v + R)n \ln T - nR \ln p + \text{const} \tag{4.24}$$

If we have two different ideal gases with identical pressures and temperatures in two containers of volumes V_1 and V_2 and we connect them then the difference of the final entropy (entropy of mixing) reads

$$\delta S(T,V) = n_1 R \ln\left(\frac{V}{V_1}\right) + n_2 R \ln\left(\frac{V}{V_2}\right) = -RV(x_1 \ln x_1 + x_2 \ln x_2) \geqslant 0 \tag{4.25}$$

where

$$x_1 = \frac{n_1}{n_1 + n_2}, \qquad x_2 = \frac{n_2}{n_1 + n_2} \tag{4.26}$$

If we know entropy as a function of U and V we know everything about the system. In the present case we write

$$S(U,V) = c_v n \ln\left(\frac{U}{n}\right) + nR \ln\left(\frac{V}{n}\right) + \text{const} \tag{4.27}$$

Then we use FLT

$$\mathrm{d}S = \frac{\mathrm{d}U}{T} + \frac{p}{T}\mathrm{d}V \tag{4.28}$$

to get

$$U = c_v n T, \qquad pV = nRT \tag{4.29}$$

so we recover both the expression for U and the equation of state.

4.4 Entropy of the VdW gas

Using the EoS for the Van der Waals gas

$$p = \frac{RT}{v-b} - \frac{a}{v^2} \tag{4.30}$$

and the result for the internal energy U (using $\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p$, $\left(\frac{\partial U}{\partial T}\right)_V = c_V$)

$$\frac{U}{n} = c_v T - a\,\frac{1}{v} \Rightarrow c_v T = \left(\frac{U}{n} + \frac{a}{v}\right) \tag{4.31}$$

we can write the FLT as

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T}, \quad \left(\frac{\partial S}{\partial V}\right)_U = \frac{p}{T} = \frac{R}{v-b} - \frac{a}{v^2 T} \tag{4.32}$$

Hence we get for the VdW gas the expression for the entropy

$$S = c_v n \ln\left(\frac{U}{n} + \frac{a}{v}\right) + nR\ln(v-b) + \mathrm{const} \tag{4.33}$$

We rewrite it in a slightly different way

$$S = (c_v + R) n \ln T - nR\ln\left(p + \frac{a}{v^2}\right) + \mathrm{const} \tag{4.34}$$

4.5 Entropy of the gas of photons

The fundamental property of photons is that the internal energy does not depend on the number of photons. If they are closed in a box then the force on a wall is given by

$$F = \frac{2h\nu v_z/c^2}{2L/v_z} = \frac{h\nu}{L}\frac{v_z}{c^2} = \frac{1}{3}\frac{U}{L} \tag{4.35}$$

Hence

$$pV = \frac{1}{3}U \Rightarrow p = \frac{1}{3}\rho \tag{4.36}$$

Assuming that nothing depends on the number of photons we substitute

$$S = \alpha T^m V, \quad p = \beta T^N, \quad \Rightarrow U = 3\beta T^n V \tag{4.37}$$

Using

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V \tag{4.38}$$

we get

$$3\beta n T^{n-1} V \mathrm{d}T + 3\beta T^n \mathrm{d}V = \alpha m T^m V \mathrm{d}T + \alpha T^{m+1} \mathrm{d}V - \beta T^n \mathrm{d}V \tag{4.39}$$

Comparing the expressions we get

$$\alpha = 4\beta, \quad m = 3, \quad n = 4 \tag{4.40}$$

so that

$$U = 3\beta T^4 V, \ \ p = \beta T^4, \quad S = 4\beta T^3 V \tag{4.41}$$

and we recover Stefan-Boltzmann law $U \sim T^4 V$.

It turns out (from the Planck black body distribution) that

$$\beta = \frac{\pi^2 k^4}{45 c^3 \hbar^3} \Rightarrow U = \frac{\pi^2 k^4}{15 c^3 \hbar^3} T^4 V, \ \ S = \frac{4\pi^2 k^3}{45 c^3 \hbar^3} T^3 V \tag{4.42}$$

the number of photons is given by

$$N = \frac{2\zeta(3) k^2}{\pi^2 c^3 \hbar^3} T^3 V = \frac{45\zeta(3)}{2\pi^4} S \tag{4.43}$$

5 Phase transitions

5.1 Chemical potential

In this section we would like to generalize the FLT. The first natural step is to include the number of particles (up to now treated as constant) as a function of state– although it is not a continuous variable even in the classical case, the number is so huge in the generic systems that it can be approximately treated as such. We assume that adding a particle to the system changes its internal energy U by the very fact of its presence. The measure of this change is the so called chemical potential μ. Therefore we write

$$\mathrm{d}U = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N \tag{5.1}$$

Writing

$$\mathrm{d}T \wedge \mathrm{d}S - \mathrm{d}p \wedge \mathrm{d}V + \mathrm{d}\mu \wedge \mathrm{d}N = 0 \tag{5.2}$$

and choosing different 3 variables as independent leads to 20 ($6!/(3! \cdot 3!)$) Maxwell identities with significantly more complicated manipulations than before with only 2 variables.

We would like to emphasize here one important point: if we make a Legendre transform with respect to all extensive variables (S, V, N) and assuming that there no other extensive variables we obtain a function of state that is extensive but formally should depend only on intensive quantities and it is impossible. Therefore we draw the conclusion that such an equation of state should vanish (Euler equation):

$$U - TS + pV - \mu N = 0 \tag{5.3}$$

Acting with d on this equation we get

$$-S\mathrm{d}T + V\mathrm{d}p - N\mathrm{d}\mu = 0 \tag{5.4}$$

what shows that (T, p, μ) are not independent. If the dependence on N is nontrivial (what is not the case for photons) then

$$\left(\frac{\partial \mu}{\partial T}\right)_p = -s \qquad \left(\frac{\partial \mu}{\partial p}\right)_T = v \tag{5.5}$$

where s and v are specific entropy and volume, respectively.

The previously introduced Gibbs potential is therefore equal to

$$G(T,p,N) = U - TS + pV = \mu(T,p)N \tag{5.6}$$

Since G attains a minimum at a given T and p if we have two parts of a system that can exchange particles we have

$$0 = \Delta G = \mu_1 \Delta N_1 + \mu_2 \Delta N_2 = \Delta N_1(\mu_1 - \mu_2) \Rightarrow \mu_1 = \mu_2 \tag{5.7}$$

so that the chemical potentials in thermodynamical equilibrium have to be equal.

If we plug into (5.6) the expressions for photons

$$U = 3\beta T^4 V, \quad p = \beta T^4, \qquad S = 4\beta T^3 V \tag{5.8}$$

we get

$$\mu = 0 \tag{5.9}$$

consistently with the assumption leading to the photon EoS where nothing depends on the number of photons.

If we plug in the expressions for ideal gas $pV = RT$ we get

$$\mu(T,p) = f(T) + RT \ln p = \mu_0(T, p = p_0) + RT \ln \frac{p}{p_0} \tag{5.10}$$

where μ_0 is the value at a given temperature and pressure

If we treat (T,V,N) as independent we get

$$\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \mu(T,V)\mathrm{d}N \tag{5.11}$$

5.2 Clausius-Clapeyron equation

We discuss now the liquid-gas transition. We start from the fact that the mass transfer from the liquid to the gas phase and vice versa is at constant pressure and constant temperature. As we argued the chemical potentials of both phases should be equal but the derivatives can be different for both phases. We recall the equations for these derivatives

$$\mathrm{d}\mu = -s\mathrm{d}T + v\mathrm{d}p \tag{5.12}$$

and therefore

$$\left(\frac{\partial \mu}{\partial T}\right)_p = -s \qquad \left(\frac{\partial \mu}{\partial p}\right)_T = v \tag{5.13}$$

where s and v are specific entropy and volume, respectively. We will use the subscripts 'l' and 'g' for the liquid and gas phases respectively.

We write $(\Delta\mu = \mu_\mathrm{g} - \mu_\mathrm{l})$

$$\begin{aligned} \left(\frac{\partial \Delta\mu}{\partial T}\right)_p &= -(s_\mathrm{g} - s_\mathrm{l}) = -\Delta s < 0, \\ \left(\frac{\partial \Delta\mu}{\partial p}\right)_T &= (v_\mathrm{g} - v_\mathrm{l}) = \Delta v > 0 \end{aligned} \tag{5.14}$$

We now use the identity (valid since $\Delta\mu$ is a function of (T,p))

$$\left(\frac{\partial\Delta\mu}{\partial T}\right)_p \left(\frac{\partial T}{\partial p}\right)_{\Delta\mu} \left(\frac{\partial p}{\partial\Delta\mu}\right)_T = -1 \tag{5.15}$$

to arrive at

$$\left(\frac{\partial p}{\partial T}\right)_{\Delta\mu=0} = \left(\frac{\mathrm{d}p}{\mathrm{d}T}\right) = \frac{\Delta s}{\Delta v} \tag{5.16}$$

Since $T\Delta s = Q_t$ is the specific latent heat of transition we finally get the Clausius-Clapeyron equation

$$\left(\frac{\mathrm{d}p}{\mathrm{d}T}\right) = \frac{Q_t}{T\Delta v} \tag{5.17}$$

that describes the change of the pressure at transition as a function of temperature of transition.

For the liquid-gas transition Δv is always positive (below the critical point) but for the solid-liquid transition it can have either sign – for water it is negative but for vast majority of substances it is positive.

This equation is valid for the so called first order transition when the first derivatives of a state function (in this case it is g) have a jump at transition.

5.3 Phase transition in VdW equation of state

If we have a maximum and a minimum (i.e. $T < T_c$) part of the isotherm is unphysical since it leads to a negative compressibility what points to instability. We can correct it by assuming that there are two coexisting phases and we use only leftmost and rightmost part of the curve joining them by a horizontal line. To the left we have only liquid phase then for some volume we start to have two phases at constant pressure that gradually change the relative abundance until at some volume only the gaseous phase remains and we can again use the VdW equation.

If $T < T_c$ the medium part of the isotherm is non-physical. One usually introduces the so called Maxwell construction that will not be desrcribed here since it is non physical as well. If we assume that the RHS of the curve describing the gas phase is physical (what is well justified) and the LHS of the curve is physical (what is qualitatively justified but quantitatively not so) then we can connect the LHS with the RHS by a horizontal line describing the coexistence of the liquid and the gas phases. The question where the line should be drawn can be answered in the following way.

We recall the expression for the entropy of the VdW fluid:

$$S = (c_v + R)n \ln T - nR \ln\left(p + \frac{a}{v^2}\right) + \text{const} \tag{5.18}$$

Since the phase transition is at constant pressure and temperature the difference of entropies between the RHS and the LHS should be equal to $\Delta S = \Delta Q/T = nQ_t/T$

where Q_t is the latent heat of vaporization and this condition gives the position of the horizontal line. For one mole we get

$$\Delta S = -R\ln\left(p + \frac{a}{v_g^2}\right) + R\ln\left(p + \frac{a}{v_l^2}\right) = \frac{Q_t}{T} \tag{5.19}$$

Hence

$$\frac{p + \frac{a}{v_l^2}}{p + \frac{a}{v_g^2}} = \exp\left(\frac{Q_t}{RT}\right) \tag{5.20}$$

Knowing Q_t and a and assuming that VdW EoS is valid on both sides of the curve we could fix the position of the horizontal line connecting the gas and the liquid phase. Applying this equation to water and the liquid-vapour phase transition

$$a \sim 0.55 \text{ Pa}\cdot\text{m}^6/\text{mol}^2, \quad p \sim 10^5 \text{ Pa}, \quad T \sim 373 \text{ K}, \quad v_l \sim 1.8\cdot 10^{-5} \text{ m}^3 \quad v_g \sim 2.2\cdot 10^{-2} \text{ m}^3 \tag{5.21}$$

we get

$$Q_t(373 \text{ K}) \sim 30000 \text{ J/mol} \tag{5.22}$$

to be compared with the actual 40000 J/mol with the difference mostly due to inapplicability of the VdW EoS to the liquid phase of water. It is, however, important to emphasize that this approach gives automatically $Q(T_c) = 0$ when $v_g(T_c) = v_l(T_c)$.

Another way to approach this problem is via the chemical potential. If we start from the left part of the curve at some pressure and volume (p_l, v_l) and we go to the right along the isotherm then the chemical potential changes as

$$\mu(v) = \mu(v_l) + \int_{p_l}^{p} \mathrm{d}p' v(p') \tag{5.23}$$

where we used the equation $\partial\mu/\partial p = v$. Since at the end on the right hand side we have to end up with the same chemical in the gaseous phase. It gives the so called Maxwell construction giving us the pressure of liquid-gas equilibrium for a given isotherm: since

$$\mu(v_g) = \mu(v_l) + \int_{v_l}^{v_g} \mathrm{d}v(p - p'(v)) = \mu(v_l) \tag{5.24}$$

the areas below the horizontal p line and above this line have to be equal. This approach has a drawback since it uses the unphysical part of the VdW curve to find the pressure of equilibrium at a given temperature but it is easy to visualize.

5.4 Critical point in VdW equation of state

We will now discuss the EoS for the Van der Waals gas

$$p = \frac{RT}{v - b} - \frac{a}{v^2} \tag{5.25}$$

Depending on T the isotherms can have purely negative $\left(\frac{\partial p}{\partial V}\right)_T$ (for $T > T_c$), one saddle point (for $T = T_c$), or one minimum and one maximum (for $T < T_c$). We can calculate T_c

$$\begin{aligned}\left(\frac{\partial p}{\partial v}\right)_T &= 0 = -\frac{RT_c}{(v_c-b)^2} + \frac{2a}{v_c^3} \\ \left(\frac{\partial^2 p}{\partial v^2}\right)_T &= 0 = \frac{2RT_c}{(v_c-b)^3} - \frac{6a}{v_c^4}\end{aligned} \tag{5.26}$$

Hence

$$v_c = 3b, \quad RT_c = \frac{8a}{27b}, \quad p_c = \frac{a}{27b^2} \tag{5.27}$$

If we measure all quantities relative to the critical values we have

$$\tilde{p} = \frac{8\tilde{T}}{3\tilde{v}-1} - \frac{3}{\tilde{v}^2} \tag{5.28}$$

We can compare the parameters of the critical point of the actual gases with the VdW equation. We introduce

$$Z_c = \frac{p_c v_c}{RT_c} \tag{5.29}$$

We read off Z_c for the VdW equation

$$Z_c^{VdW} = \frac{3}{8} \tag{5.30}$$

For water we have

$$p_c \sim 2.2 \cdot 10^7 \text{ Pa}, \quad v_c \sim 56 \text{ cm}^3/\text{mol}, \quad T_c \sim 647\text{K} \tag{5.31}$$

hence

$$Z_c^{H_2O} \sim 0.23 \tag{5.32}$$

For the carbon dioxide we have

$$p_c \sim 7.4 \cdot 10^6 \text{ Pa}, \quad v_c \sim 94 \text{ cm}^3/\text{mol}, \quad T_c \sim 304\text{K} \tag{5.33}$$

hence

$$Z_c^{CO_2} \sim 0.27 \tag{5.34}$$

For the oxygen we have

$$p_c \sim 5 \cdot 10^6 \text{ Pa}, \quad v_c \sim 73 \text{ cm}^3/\text{mol}, \quad T_c \sim 155\text{K} \tag{5.35}$$

hence

$$Z_c^{O_2} \sim 0.28 \tag{5.36}$$

To be closer to the observed values we would have to assume some modified equation of state especially on the liquid branch.

5.5 Critical exponents

We notice that several quantities go to 0 when we approach the critical point. It is important to ask about the rate of this approach – it is encoded in the so called critical exponents. It is a very important area of research in statistical physics connected with the conformal invariance at the critical point, conformal field theory and the renormalization group introduced by Kenneth Wilson. We will show the idea on the example of VdW equation of state.

First we ask about the difference $v_g - v_l$ as we approach $T \to T_c$ from below. An argument would be to write

$$\tilde{p} = \frac{8\tilde{T}}{3\tilde{v}_l - 1} - \frac{3}{\tilde{v}_l^2} = \frac{8\tilde{T}}{3\tilde{v}_g - 1} - \frac{3}{\tilde{v}_g^2} \tag{5.37}$$

Calculating $\tilde{T}$ we get

$$8\tilde{T} = \frac{(\tilde{v}_l + \tilde{v}_g)}{\tilde{v}_l^2 \tilde{v}_g^2}(9\tilde{v}_l\tilde{v}_g - 3\tilde{v}_l - 3\tilde{v}_g + 1) \tag{5.38}$$

The second equation is connected to the equality of chemical potentials

$$0 = \int_{\tilde{v}_l}^{\tilde{v}_g} \mathrm{d}v(\tilde{p} - \tilde{p}'(v)) = (\tilde{v}_g - \tilde{v}_l)\tilde{p} - \frac{8\tilde{T}}{3}\ln\left(\frac{3\tilde{v}_g - 1}{3\tilde{v}_l - 1}\right) - \frac{3}{\tilde{v}_g} + \frac{3}{\tilde{v}_l} \tag{5.39}$$

Solving these two equations in the vicinity of the critical point it turns out that

$$\frac{1}{2}(\tilde{v}_g + \tilde{v}_l) = 1 + \frac{18}{5}(1 - \tilde{T}) + O((1 - \tilde{T})^2), \quad \tilde{v}_g - \tilde{v}_l = 4(1 - \tilde{T})^{\frac{1}{2}} + O((1 - \tilde{T})^{3/2}) \tag{5.40}$$

and writing in general

$$\tilde{v}_g - \tilde{v}_l \sim (1 - \tilde{T})^{\beta} \tag{5.41}$$

we have $\beta = 1/2$.

The next exponent is related to the question how the pressure changes when we approach the volume to v_c. Since $\partial p/\partial v = \partial^2 p/\partial v^2 = 0$ at the critical temperature and close to the critical pressure and density we immediately write

$$\tilde{p} - 1 \sim (\tilde{v} - 1)^{\delta} \tag{5.42}$$

with $\delta = 3$.

The third question concerns the compressibility of the gas. As we know it goes to infinity at the critical point and we observe that

$$\frac{\partial \tilde{p}}{\partial \tilde{v}} \sim -6(\tilde{T} - 1) \tag{5.43}$$

so that

$$\kappa = -\frac{\partial \tilde{v}}{\partial \tilde{p}} \sim (\tilde{T} - 1)^{-\gamma} \tag{5.44}$$

with $\gamma = 1$

For real gases and liquids the critical exponents are slightly different:

$$\beta \sim 0.32, \quad \delta \sim 4.8, \quad \gamma \sim 1.2 \tag{5.45}$$

and it is still an unsolved problem to calculate them analytically in any 3-dimensional model.

6 Classical statistical physics

6.1 Liouville theorem

Imagine the flow of (q_a, p_a) i.e. a tube of close trajectories (in the phase space). Its volume is

$$\delta\Gamma = \mathrm{d}q_1 \dots \mathrm{d}q_n \mathrm{d}p_1 \dots \mathrm{d}p_n \tag{6.1}$$

We ask what will be this infinitesimal volume after time $\mathrm{d}t$. Then

$$q_a \to \tilde{q}_a = q_a + \frac{\partial H}{\partial p_a}\mathrm{d}t, \qquad p_a \to \tilde{p}_a = p_a - \frac{\partial H}{\partial q_a}\mathrm{d}t, \tag{6.2}$$

The jacobian from Γ to $\tilde{\Gamma}$ reads

$$J = \begin{pmatrix} \frac{\partial \tilde{q}_a}{\partial q_b} & \frac{\partial \tilde{q}_a}{\partial p_b} \\ \frac{\partial \tilde{p}_a}{\partial q_b} & \frac{\partial \tilde{p}_a}{\partial p_b} \end{pmatrix} \tag{6.3}$$

We now use the formula

$$\exp(\mathrm{Tr} \ln M) = \det M \tag{6.4}$$

for an arbitrary matrix M with positive eigenvalues. It can be proven using the fact that any matrix can be brought to the diagonal (or Jordan) form by some (complex) matrix A. Indeed, writing

$$M = \mathbf{1} + \delta \tag{6.5}$$

we have (M' is in the diagonal or Jordan form)

$$M' = AMA^{-1} \Rightarrow \mathrm{Tr} \ln M = \mathrm{Tr}\left(\delta + \frac{1}{2}\delta^2 + \dots\right) = \mathrm{Tr}\left(\delta' + \frac{1}{2}\delta'^2 + \dots\right) = \sum \ln \lambda_i \tag{6.6}$$

and we see that both sides of the equation (6.4) are equal to the product of the eigenvalues. In our case

$$M = \mathbf{1} + \delta \qquad \Rightarrow \qquad \det M = 1 + \mathrm{Tr}\delta + O(\delta^2) \tag{6.7}$$

but

$$\mathrm{Tr}\delta = \sum_a \left(\frac{\partial^2 H}{\partial q_a \partial p_a} - \frac{\partial^2 H}{\partial p_a \partial q_a}\right)\mathrm{d}t = 0 \tag{6.8}$$

so that

$$\delta\Gamma = \delta\tilde{\Gamma} \tag{6.9}$$

It says that 'squeezing' the trajectories in q requires 'expanding' them in p – it resembles quantum uncertainty relation but it is very different being purely classical.

6.2 Poincaré recurrence theorem

We now prove one of the most striking theorems in classical mechanics.

We assume that the phase space is of finite phase volume (for example of finite energy and in finite spatial volume). We consider finite time steps $0, T, 2T, \ldots$. The theorem says that for any point P_0 and for any neighborhood D_0 of P_0 in the phase space there exists such n that

$$D_n \cap D_0 \neq \emptyset \tag{6.10}$$

where D_n is D_0 transformed by H after time nT.

The proof consists in showing that since for all n regions D_n have the same volume then there must exist such n' and n'' (different from each other) for which

$$D_{n'} \cap D_{n''} \neq \emptyset \tag{6.11}$$

since otherwise the volume of the phase space would be infinite. Taking for example $n' < n''$ then acting with H backwards n times (action of the hamiltonian is reversible) we get

$$D_0 \cap D_{n''-n'} \neq \emptyset \tag{6.12}$$

what finishes the proof.

6.3 Liouville's equation

For a system of N bodies we can introduce a density on the phase space $\rho(q,p)$. Since the volume of the phase space is constant we get

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial \rho}{\partial t} + \sum_a \left(\frac{\partial \rho}{\partial q_a}\dot{q}_a + \frac{\partial \rho}{\partial p_a}\dot{p}_a \right) = 0 \tag{6.13}$$

what gives the Liouville equation

$$\frac{\partial \rho}{\partial t} = -\sum_a \left(\frac{\partial \rho}{\partial q_a}\frac{\partial H}{\partial p_a} - \frac{\partial \rho}{\partial p_a}\frac{\partial H}{\partial q_a} \right) = -\{\rho, H\}_{PB} \tag{6.14}$$

where we used here the notion of a Poisson Bracket defined as

$$\{f,g\}_{PB} := \sum_a \frac{\partial f}{\partial q_a}\frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q_a} \tag{6.15}$$

An important role is played by time independent densities for which $\frac{\partial \rho}{\partial t} = 0$. The most important example of such stationary distributions is given by

$$\rho = \rho(H(q,p)) \tag{6.16}$$

where H does not depend on time. Then indeed

$$\frac{\partial \rho}{\partial t} = \frac{\partial \rho}{\partial H}\left(-\frac{\partial H}{\partial q_a}\frac{\partial H}{\partial p_a} + \frac{\partial H}{\partial p_a}\frac{\partial H}{\partial q_a} \right) = 0 \tag{6.17}$$

There are two most famous examples of such a distribution. The first one is the microcanonical ensemble where

$$\rho(H(q,p)) = \theta\left(U - H(q,p)\right) \tag{6.18}$$

The second example of such a distribution is the Gibbs-Boltzmann factor in the canonical ensemble

$$\rho(H(q,p)) = \exp\left(-\frac{H(q,p)}{kT}\right) \tag{6.19}$$

which in classical statistical physics for free particles is proven to describe the most probable distribution (if H describes free particles the distribution is usually called Maxwell-Boltzmann distribution).

6.4 Entropy

We have introduced already the concept of entropy in thermodynamics but it is of such a fundamental importance in many different aspects that we will now discuss it in more detail.

As we discussed the concept and the name was introduced by Clausius in 1865. It was soon used by Gibbs and Helmholtz in the definition of other functions of state. In 1877 Ludwig Boltzmann proposed his famous formula

$$S = k \ln \Omega \tag{6.20}$$

where k is a constant (later renamed k_B, the Boltzmann constant) and Ω is a number of microstates realizing a given macrostate. It was loosely treated and understood as a 'measure of disorder' but in fact the equation remained totally mysterious and started to have a meaning only 50 years later in the framework of quantum mechanics. About the same time Erwin Schrödinger introduced a concept of 'negative entropy' when some states are distinguished and have higher probability than the others as in the example of gene replication.

John von Neumann in 1932 introduced the entropy in quantum mechanics using the density matrix ρ reflecting our knowledge about probabilities coming from the measurements (and not the unitary evolution as in Heisenberg or Schrödinger equations). He defined

$$S = -\mathrm{Tr}\rho \ln \rho \tag{6.21}$$

For the pure state we have $\rho^2 = \rho$ and the entropy vanishes, for the maximally mixed state the entropy is $\ln N_H$ where N_H is the size of the Hilbert space (for example for a system of N spins $\frac{1}{2}$ the maximal entropy is $N \ln 2$). In QM it has some strange properties like the entropy of an entangled state can be lower than the entropy of individual components but we leave it aside here. In 1948 Claude Shannon introduced the same concept in information theory and proved that there are intrinsic limits to the lossless transmission of signals (both without noise and with noise).

There are several properties of the definition

$$S = -\sum_i p_i \ln p_i \tag{6.22}$$

that justify its form. If we have a system consisting of two subsystems A and B then the joined probabilities are products of the individual ones

$$p_{ik} = p_{Ai} p_{Bk}, \quad \sum_{ik} p_{ik} = \sum_i p_{Ai} \sum_k p_{Bk} = 1 \tag{6.23}$$

Then

$$\begin{aligned} S_{AB} &= -\sum_{ik} p_{ik} \ln p_{ik} = -\sum_{ik} p_{Ai} p_{Bk} \ln(p_{Ai} p_{Bk}) = \\ &= -\sum_{ik} p_{Ai} p_{Bk} \ln(p_{Ai}) - \sum_{ik} p_{Ai} p_{Bk} \ln(p_{Bk}) = S_A + S_B \end{aligned} \tag{6.24}$$

6.5 Derivation of the Gibbs-Boltzmann factor

If we have a system that we can divide into two subsystems that can be treated (in first approximation) as independent and in each subsystem we have energy levels E_1^i and E_2^k, $i, k = 1, 2, \ldots$ then we can write

$$E_{i,k} = E_1^i + E_2^k \tag{6.25}$$

The probability to find such a pair (again assuming independence) is given by

$$p_{i,k}(E) = p_1^i \cdot p_2^k \tag{6.26}$$

We assume that the probabilities depend only on the energies i.e.

$$p_{i,k}(E) = p_0(E), \quad p_1^i = p_1(E_i), \quad p_2^k = p_2(E_k) \tag{6.27}$$

Differentiating $p_0(E)$ wrt E_i and E_k and using (6.25) we get

$$\frac{\partial \ln(p_0(E))}{\partial E_i} = \frac{\partial \ln(p_0(E))}{\partial E_k} \Rightarrow \frac{\partial \ln(p_1(E_i))}{\partial E_i} = \frac{\partial \ln(p_2(E_k))}{\partial E_k} \tag{6.28}$$

But LHS depends only on E_i and RHS only on E_k so both sides have to be equal to a constant $-\beta$. Hence for all cases

$$p(E_i) = \frac{\mathrm{e}^{-\beta E_i}}{\sum_i \mathrm{e}^{-\beta E_i}} \tag{6.29}$$

where we imposed that the probabilities have to be normalized to 1.

We can also use the original method of Boltzmann. If we divide a system into subsystems with occupation numbers N_i (N_1 is a number of subsystems in the ground state, N_2 in the first excited state and so on) then the number of ways it can be done is

$$\frac{N!}{N_1! N_2! \ldots} \tag{6.30}$$

where $N = N_1 + N_2 + \ldots$. Then we want to maximize this number (its logarithm) keeping the total number N fixed and the total energy fixed. Therefore we have to introduce Lagrange multipliers and maximize

$$L = \ln\left(\frac{N!}{N_1!N_2!\ldots}\right) + \alpha(N - \sum_i N_i) + \beta(E - \sum_i N_i E_i) \tag{6.31}$$

We use the Stirling approximation for large q

$$\ln q! \sim q \ln q - q + \frac{1}{2}\ln 2\pi q + O(1/(12q)) \tag{6.32}$$

Therefore

$$L \sim N \ln N - N - \sum_i N_i \ln N_i + \sum N_i + \alpha(N - \sum_i N_i) + \beta(E - \sum_i N_i E_i) + O(\ln(N_i)) \tag{6.33}$$

Differentiating wrt N_i we get for each i

$$-\ln N_i - \alpha - \beta E_i = 0 \tag{6.34}$$

which is the desired result.

7 Microcanonical and canonical ensembles

7.1 Microcanonical ensemble

In classical statistical physics we are interested in the classical partition function. One distinguishes different ensembles: microcanonical (isolated, with fixed number of particles and energy), canonical (with fixed number of particles but exchanging energy with a reservoir of a given temperature) and macrocanonical (exchanging both the energy and the number of particles with the reservoir).

Definition of an ensemble is not a mathematically well defined object. We imagine a set of many systems with (almost) the same macroscopic properties or the same system (in equilibrium) seen at many times but both descriptions are rather intuitive. A fundamental assumption that all microstates forming a given macrostate are equally accessible and should be counted with equal probability (ergodic hypothesis) is not rigorously proven until today.

The microcanonical ensemble is described by the number of particles and the volume of the phase space (assuming that it is finite) with the energy U in the small interval ΔU. The number of 'states' in classical physics is formally infinite so to make it well defined we need to appeal to quantum physics where there is a heuristic rule that a new state is possible when $\Delta q \Delta p$ differs by h (the Planck constant). Using this heuristic rule we calculate the number of states in an interval ΔU around

$$Z_N(V,U,\Delta U) = \left(\int^{U} - \int^{U-\Delta U} \right) \frac{\mathrm{d}^{3N}p\,\mathrm{d}^{3N}q}{N!h^{3N}} \sim \Delta U \int \frac{\mathrm{d}^{3N}p\,\mathrm{d}^{3N}q}{N!h^{3N}}\,\delta(U - T_N - V_N) \tag{7.1}$$

if $\Delta U \ll \frac{U}{N}$. $1/N!$ is the Gibbs factor, yet another factor that can be justified only in quantum physics (indistinguishability of identical particles), although Gibbs introduced it by considering the entropy of mixing of two gases and finding a paradox when the gases were identical. $Z_N(V,U)$ is then the number of states around U in the interval ΔU.

According to the famous Boltzmann formula logarithm of Z_N is equal to the entropy (modulo a constant)

$$S = k\ln(W) \tag{7.2}$$

This formula is on Boltzmann's grave in Vienna – it required an incredible ingenuity of Boltzmann to write it down in 1875, 25 years before the Planck's assumption of quantization of photon emissions and absorptions.

It can be justified by the formula (also given by Boltzmann in 1866)

$$S = -\sum P \ln P \tag{7.3}$$

and using equal (maximal) probability $P = 1/Z_N$ for all states ($\sum P = 1$).

In the following we put the Boltzmann's constant k equal to 1 (we measure temperature in units of $8.6173332\ldots \cdot 10^{-5}$ eV). Knowing S we can recover all thermodynamical functions in this ensemble by

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}U + \frac{p}{T}\mathrm{d}V \tag{7.4}$$

i.e.

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V, \quad \frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_U \tag{7.5}$$

As an example let us discuss free non-relativistic particles. Then

$$Z_N(V, U, \Delta U) = \Delta U \frac{V^N \Omega^{(3N-1)}}{N! h^{3N}} (2mU)^{3N/2} \tag{7.6}$$

where $\Omega^{(3N-1)}$ is the volume of $(3N-1)$-dimensional unit sphere

$$\Omega^{(3N-1)} = \frac{2\pi^{3N/2}}{\Gamma(3N/2)} \tag{7.7}$$

Hence (using $\ln(N!) = \ln(\Gamma(N+1)) = N\ln(N) - N + \frac{1}{2}\ln(2\pi N) + O(1/N)$)

$$\frac{S}{N} \to \ln(V/N) + \frac{3}{2}\ln\left(\frac{4\pi mU}{h^2 N}\right) \tag{7.8}$$

We see that without the $N!$ factor in the denominator S would not be proportional to N but there would be logarithmic corrections to S/N growing like $\ln N$. The result is the so called Sackur-Tetrode equation.

Hence we recover the well-known results

$$\frac{1}{T} = \frac{3N}{2U}, \quad \frac{p}{T} = \frac{N}{V} \tag{7.9}$$

7.2 Canonical ensemble

We consider a small subsystem attached to a large one of temperature T. The number of states of the large system when the small one has energy E is approximately equal to

$$\exp(S'(E_0 - E)) = \exp\left(S'(E_0) - E\frac{\mathrm{d}S'(E_0)}{\mathrm{d}E_0}\right) = \exp\left(S'(E_0) - \frac{E}{T}\right) \tag{7.10}$$

and it describes a probability of the small system's energy E when the reservoir has temperature T.

In the canonical ensemble we do not assume that the energy is constant but that the system is immersed in a bath of temperature T and we have a Boltzmann's exponential weight so that the probability of a given state of energy E is given by

$$P = \mathrm{e}^{\beta(F-E)}, \qquad \beta = \frac{1}{T} \tag{7.11}$$

and F is a normalizing factor. Sum of probabilities must be equal to 1 so

$$\mathrm{e}^{-\beta F(V,T)} = \int \frac{\mathrm{d}^{3N}p\mathrm{d}^{3N}q}{N!h^{3N}}\mathrm{e}^{-\beta H(p,q)} \tag{7.12}$$

where $1/N!$ is again the Gibbs factor. We know that the entropy S is given by

$$S = -\sum P\ln P = -\sum \beta(F-E)\mathrm{e}^{\beta(F-E)} = -\beta F + \beta U \tag{7.13}$$

Hence

$$F = U - TS \tag{7.14}$$

and it can be identified with the free energy.

Therefore

$$\begin{aligned} p &= -\left(\frac{\partial F}{\partial V}\right)_T \\ S &= -\left(\frac{\partial F}{\partial T}\right)_V \end{aligned} \tag{7.15}$$

As a first example we consider again free non-relativistic particles. Then

$$\mathrm{e}^{-\beta F(V,T)} = \int \frac{\mathrm{d}^{3N}p\mathrm{d}^{3N}q}{N!h^{3N}}\mathrm{e}^{-\beta p^2/(2m)} \tag{7.16}$$

The density for one particle

$$\int \frac{\mathrm{d}^3p\mathrm{d}^3q}{h^3}\mathrm{e}^{-\beta p^2/(2m)} \tag{7.17}$$

is called Maxwell-Boltzmann distribution and was the first to be introduced in statistical physics.

For N particles we have

$$\mathrm{e}^{-\beta F(V,T)} = \frac{V^N\Omega^{(3N-1)}}{N!h^{3N}}\int p^{3N-1}\mathrm{e}^{-\beta p^2/(2m)}\mathrm{d}p \tag{7.18}$$

The integral is straightforward and we get

$$\mathrm{e}^{-\beta F(V,T)} = \frac{V^N\Omega^{(3N-1)}}{N!h^{3N}}(2m)^{3N/2}\frac{\Gamma(3N/2)}{\beta^{3N/2}} \tag{7.19}$$

Hence

$$-\beta F = NC' + N\ln(V/N) - \frac{3N}{2}\ln(\beta) \tag{7.20}$$

and we recover the known formulae.

8 Grand canonical ensemble

8.1 Classical grand canonical ensemble

In the classical setting the we assume that the system is immersed in a bath of temperature T and chemical potential μ and we have a Boltzmann's exponential weight. We assign a probability for N particles having energy E_N as

$$p_N(E_N) = \mathrm{e}^{\beta(\mu N - E_N)} \tag{8.1}$$

and the normalizing factor Ω defined as

$$\sum \frac{1}{N!}\mathrm{e}^{\beta(\Omega + \mu N - E_N)} = 1 \tag{8.2}$$

We define entropy as

$$S = -\sum p_i \ln p_i = -\sum \frac{1}{N!}\beta(\Omega + \mu N - E_N)\mathrm{e}^{\beta(\Omega + \mu N - E_N)} = -\beta\Omega - \beta\mu < N > + \beta < E > \tag{8.3}$$

and hence

$$\Omega = U - TS + \mu N = -pV \tag{8.4}$$

Therefore

$$\mathrm{e}^{-\beta\Omega} = \sum_{N=0}^{\infty} \frac{1}{N!}\mathrm{e}^{\beta\mu N} \int \mathrm{d}\Gamma_N \mathrm{e}^{-\beta E_N} \tag{8.5}$$

8.2 Quantum grand canonical ensemble

In the quantum setting we assume discrete energy levels and in the grand canonical ensemble we assume that the system is immersed in a bath of temperature T and chemical potential μ and we have a Boltzmann's exponential weight for both so that the probability of a given state of energy E_j occupied by n_i number of particles is given by

$$P_{i,j} = \mathrm{e}^{\beta(\Omega + \mu n_i - n_i E_j)} \tag{8.6}$$

where $\beta = 1/T$ and Ω is a normalizing factor. Sum of probabilities must be equal to 1 so

$$\mathrm{e}^{-\beta\Omega(\mu,V,T)} = \prod_j \sum_{n_i} \mathrm{e}^{\beta(\mu n_i - n_i E_j)} \tag{8.7}$$

We know that the entropy S is given by

$$S = -\sum P_{i,j} \ln P_{i,j} = -\sum \beta(\Omega + \mu n_i - n_i E_j)\mathrm{e}^{\beta(\Omega+\mu n_i - n_i E_j)} = -\beta\Omega - \beta\mu < N > +\beta U \tag{8.8}$$

Hence

$$\Omega = U - \mu N - TS = -pV \tag{8.9}$$

and the average number of particles

$$< N >= -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = V\left(\frac{\partial p}{\partial\mu}\right)_{T,V} \tag{8.10}$$

Therefore

$$\mathrm{d}\Omega = -S\mathrm{d}T - p\mathrm{d}V - N\mathrm{d}\mu \tag{8.11}$$

For non-interacting fermions (occupation numbers $0, 1$)

$$\mathrm{e}^{-\beta\Omega(\mu,T)} = \prod_j \left(1 + \mathrm{e}^{\beta(\mu - E_j)}\right) \tag{8.12}$$

For non-interacting bosons (occupation numbers $0, 1, 2, \ldots$)

$$\mathrm{e}^{-\beta\Omega(\mu,T)} = \prod_j \sum_{n_i} \mathrm{e}^{\beta(\mu n_i - n_i E_j)} = \prod_j \left(1 - \mathrm{e}^{\beta(\mu - E_j)}\right)^{-1} \tag{8.13}$$

giving Fermi-Dirac and Bose-Einstein distributions respectively.

They give for fermions

$$-\frac{\partial\Omega}{\partial\mu} = \sum_j < n_j >= \sum_j \frac{1}{\mathrm{e}^{\beta(E_j-\mu)} + 1} \tag{8.14}$$

while for bosons

$$-\frac{\partial\Omega}{\partial\mu} = \sum_j < n_j >= \sum_j \frac{1}{\mathrm{e}^{\beta(E_j-\mu)} - 1} \tag{8.15}$$

For bosons $\mu \leqslant E_0$ otherwise the expression wouldn't make sense.

Differentiating (8.6) once over μ we get

$$\frac{\partial\Omega}{\partial\mu}\sum \mathrm{e}^{\beta(\Omega+\mu n_i - n_i E_j)} + \sum n_i \mathrm{e}^{\beta(\Omega+\mu n_i - n_i E_j)} = 0 \tag{8.16}$$

and once again

$$\frac{\partial^2\Omega}{\partial\mu^2} + \beta\left(\frac{\partial\Omega}{\partial\mu}\right)^2 + 2\beta\frac{\partial\Omega}{\partial\mu} < N > +\beta < N^2 >= 0 \tag{8.17}$$

Hence

$$< N^2 > - < N >^2 = -T\frac{\partial^2\Omega}{\partial\mu^2} \tag{8.18}$$

so that fluctuations are of the order of $1/\sqrt{N}$ since RHS is of order N.

8.3 Planck distribution

We start from the expression for the average occupational number as a function of energy

$$-\frac{\partial\Omega}{\partial\mu} = \sum_j < n_j > = \sum_j \frac{1}{\mathrm{e}^{\beta(E_j-\mu)}-1} \tag{8.19}$$

For photons we have $\mu = 0$ so we can write

$$\frac{\mathrm{d}E}{\mathrm{d}k} = \frac{2V4\pi k^2\hbar kc}{(2\pi)^3}\frac{1}{\mathrm{e}^{\beta\hbar kc}-1} \tag{8.20}$$

where V is the volume, $E = \hbar\omega = \hbar kc$ is the energy and 2 comes from the number of polarizations. It is more often presented as a function of frequencies

$$\frac{\mathrm{d}E}{\mathrm{d}\omega} = \frac{V\hbar}{\pi^2c^3}\frac{\omega^3}{\mathrm{e}^{\beta\hbar\omega}-1} \tag{8.21}$$

Total energy contained in a box is equal to

$$U = \frac{T^4V}{(\hbar c)^3\pi^2}\int_0^\infty \frac{x^3\mathrm{d}x}{\mathrm{e}^x-1} \tag{8.22}$$

The integral is equal to

$$\int_0^\infty \frac{x^3\mathrm{d}x}{\mathrm{e}^x-1} = \int_0^\infty \mathrm{d}x x^3\left(\mathrm{e}^{-x}+\mathrm{e}^{-2x}+\ldots\right) = 6\zeta(4) = \frac{\pi^4}{15} \tag{8.23}$$

so that finally (introducing the Boltzmann's constant k_B)

$$U = \frac{(k_BT)^4V\pi^2}{15(\hbar c)^3} \tag{8.24}$$

and the energy density

$$\rho = \frac{U}{V} = \frac{(k_BT)^4\pi^2}{15(\hbar c)^3} \tag{8.25}$$

As we know the pressure is equal to

$$p = \frac{1}{3}\rho \tag{8.26}$$

We can now calculate entropy of the photon gas

$$\mathrm{d}S = \frac{\mathrm{d}U}{T} + \frac{p\mathrm{d}V}{T} \Rightarrow S = \frac{4U}{3T} \tag{8.27}$$

and the number density of photons

$$\frac{N}{V} = \frac{(k_BT)^3}{(\hbar c)^3\pi^2}\int_0^\infty \frac{x^2\mathrm{d}x}{\mathrm{e}^x-1} = \frac{(k_BT)^3 2\zeta(3)}{(\hbar c)^3\pi^2} \tag{8.28}$$

If we make small hole of surface S in the container we get the energy emitted

$$\mathrm{d}E=\frac{\int_0^{\pi/2}2\pi\cos(\theta)\sin(\theta)\mathrm{d}\theta}{\int_0^{\pi}2\pi\sin(\theta)\mathrm{d}\theta}\rho Sc\mathrm{d}t=\frac{\rho Sc\mathrm{d}t}{4} \tag{8.29}$$

we get the flux Φ

$$\Phi=\frac{(k_BT)^4\pi^2c}{60(\hbar c)^3}=\sigma T^4 \tag{8.30}$$

where σ is the Stefan-Boltzmann constant.

8.4 Distributions at low temperatures

We start from the expressions for the number of nonrelativistic massive particles

$$\frac{N}{V}=\frac{g8\pi m^{3/2}}{\sqrt{2}h^3}\int_0^\infty\frac{\sqrt{\epsilon}\mathrm{d}\epsilon}{\mathrm{e}^{(\epsilon-\mu)/T}\pm 1}=\frac{g8\pi(mT)^{3/2}}{\sqrt{2}h^3}\int_0^\infty\frac{\sqrt{z}\mathrm{d}z}{\mathrm{e}^{z-\mu/T}\pm 1} \tag{8.31}$$

where $g=2s+1$ is the number of degrees of freedom connected with the spin (for massive particles). This expression gives the chemical potential μ as a function of temperature and density. The energy is given by

$$E=\frac{g8\pi m^{3/2}}{\sqrt{2}h^3}\int_0^\infty\frac{\epsilon^{3/2}\mathrm{d}\epsilon}{\mathrm{e}^{(\epsilon-\mu)/T}\pm 1} \tag{8.32}$$

and the pressure (from $\Omega=-pV$)

$$p=\frac{2E}{3V}=\frac{g8\sqrt{2}\pi m^{3/2}T^{5/2}}{3h^3}\int_0^\infty\frac{z^{3/2}\mathrm{d}z}{\mathrm{e}^{z-\mu/T}\pm 1} \tag{8.33}$$

We now ask when these expressions go to the classical Boltzmann expressions (without ±1 in the denominator). We expand (8.34

$$\frac{N}{V}=\frac{g8\pi(mT)^{3/2}}{\sqrt{2}h^3}\int_0^\infty\frac{\sqrt{z}\mathrm{d}z}{\mathrm{e}^{z-\mu/T}\pm 1}\sim\frac{g8\pi(mT)^{3/2}}{\sqrt{2}h^3}\int_0^\infty\sqrt{z}\mathrm{d}z\left(\mathrm{e}^{-z+\mu/T}\mp\mathrm{e}^{-2z+2\mu/T}+\ldots\right) \tag{8.34}$$

Using

$$\int_0^\infty z^\alpha\mathrm{e}^{-bz}=\frac{\Gamma(\alpha+1)}{b^{\alpha+1}} \tag{8.35}$$

we get the leading expression

$$\frac{N}{V}=\frac{g8\pi(mT)^{3/2}}{\sqrt{2}h^3}\int_0^\infty\sqrt{z}\mathrm{d}z\mathrm{e}^{-z+\mu/T}=\frac{g8\pi^{3/2}(mT)^{3/2}}{\sqrt{2}h^3}\mathrm{e}^{\mu/T} \tag{8.36}$$

Therefore the correction is small when $\mathrm{e}^{\mu/T}\ll 1$:

$$\mathrm{e}^{\mu/T}\ll 1\Rightarrow\frac{N}{V}\ll\left(\frac{mT}{h^2}\right)^{3/2} \tag{8.37}$$

This condition shows when the temperature is high enough for a given density to use the classical expressions $(V/N)^{1/3}\gg h/\sqrt{mT}\sim h/p$ corresponds to the average distance between particles much bigger than the de Broglie wavelength.

9 Imperfect gases

9.1 Partition function with interactions

We start with the partition function in the canonical ensemble including interactions

$$\mathrm{e}^{-\beta F(V,T)} = \int \frac{\mathrm{d}^{3N}p\mathrm{d}^{3N}q}{N!h^{3N}} \mathrm{e}^{-\beta \sum p_i^2/(2m) - \beta U(q_i)} \tag{9.1}$$

The interaction over momenta gives the perfect gas result $\mathrm{e}^{-\beta F_p(V,T)}/V^N$ so we have

$$\mathrm{e}^{-\beta F(V,T)} = \mathrm{e}^{-\beta F_p(V,T)} \int \frac{\mathrm{d}^3q_1 \dots \mathrm{d}^3q_N}{V^N} \mathrm{e}^{-\beta U(q_i)} \tag{9.2}$$

We rewrite it in a slightly different way

$$\mathrm{e}^{-\beta F(V,T)} = \mathrm{e}^{-\beta F_p(V,T)} \left[1 + \int \frac{\mathrm{d}^3q_1 \dots \mathrm{d}^3q_N}{V^N} \left(\mathrm{e}^{-\beta U(q_i)} - 1 \right) \right] \tag{9.3}$$

so that

$$F(V,T) = F_p(V,T) - T \ln \left[1 + \int \frac{\mathrm{d}^3q_1 \dots \mathrm{d}^3q_N}{V^N} \left(\mathrm{e}^{-\beta U(q_i)} - 1 \right) \right] \tag{9.4}$$

Before we present a general approach to this problem we point out some approximations.

9.2 Approximations of the interacting partition function

We start with the approximation that there are only 2-body interactions, they are very short range and the gas is very diluted. Then the vast part of the region of integration does not contribute to the integral. Neglecting the 3-body interactions and expanding $\ln(1+x) \sim x$ we have

$$F(V,T) = F_p(V,T) - \frac{TN^2}{2V^2} \int \mathrm{d}^3q_1\mathrm{d}^3q_2 \left(\mathrm{e}^{-\beta U_{12}} - 1 \right) \tag{9.5}$$

Since U_{12} depends only on the relative positions we can integrate over the center of mass positions to get

$$F(V,T) = F_p(V,T) + \frac{TN^2 B(T)}{V} \tag{9.6}$$

where

$$B(T) = \frac{1}{2}\int \mathrm{d}^3 q_1 \left(1 - \mathrm{e}^{-\beta U_1)}\right) \tag{9.7}$$

If $U_1 = U_{12}(q_2 = 0)$ has finite range this integral depends only on the temperature and the range.

Since

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

we have

$$p = \frac{NT}{V}\left(1 + \frac{NB(T)}{V}\right) \tag{9.8}$$

so the correction is small for large molar volumes.

We now make a different approximation. We divide the region of integration into a 'hard core' of radius R where $U_{12} \to \infty$ and the outside where $U_{12} < 0$ and $|\beta U_{12}| \ll 1$. Therefore the integral (9.7) is equal to

$$B(T) = b - \beta a \tag{9.9}$$

where $b = 4V_R$ and

$$F(V,T) = F_p(V,T) + \frac{TN^2 b}{V} - \frac{N^2 a}{V}. \tag{9.10}$$

Adding the condition that the volume cannot be decreased below certain value we include the first correction

$$p = \frac{NT}{V} + \frac{TN^2 b}{V^2} - \frac{aN^2}{V^2} \sim \frac{NT}{V - Nb} - \frac{aN^2}{V^2} \tag{9.11}$$

and we recover the van der Waals equation of state.

9.3 Cluster expansion

The free energy F is calculated for a fixed number of particles and it is very difficult to get the the systematic expansion of the result for increasing N. A way to obtain the result as an expansion not in N but in powers of density N/V is to calculate the grand canonical ensemble where we use the chemical potential instead of N. We calculate now Ω for the imperfect gas as an expansion in the number of included particles. We write

$$\mathrm{e}^{-\beta\Omega} = \sum_{N=0}^{\infty} \frac{1}{N!} \mathrm{e}^{\beta\mu N} \int \mathrm{d}\Gamma_N \mathrm{e}^{-\beta E_N} \tag{9.12}$$

We have

$$\begin{aligned} E_0 &= 0 && (9.13)\\ E_1 &= \frac{p^2}{2m} && (9.14)\\ E_2 &= \frac{p_1^2}{2m}+\frac{p_2^2}{2m}+U_{12} && (9.15)\\ E_3 &= \frac{p_1^2}{2m}+\frac{p_2^2}{2m}+\frac{p_3^2}{2m}+U_{123} && (9.16)\end{aligned}$$

and so on. Integrals over the momenta can be performed in each case. Hence

$$\Omega = -T\ln\left(1+\xi V+\frac{\xi^2}{2}\int\int \mathrm{e}^{-\beta U_{12}}\mathrm{d}V_1\mathrm{d}V_2+\ldots\ldots\right) \tag{9.17}$$

where

$$\xi = \left(\frac{2\pi mT}{h^2}\right)^{3/2}\mathrm{e}^{\beta\mu} \tag{9.18}$$

We rewrite it as

$$\Omega = -T\xi V - T\ln\left(\mathrm{e}^{-\xi V}\left(1+\xi V+\frac{\xi^2 V}{2}\int \mathrm{e}^{-\beta U_{12}}\mathrm{d}V_2+\ldots\ldots\right)\right) \tag{9.19}$$

where we integrated once over the overall position of pairs. We can expand it in ξ:

$$\begin{aligned}\Omega &= -T\xi V - T\ln\left(\left(1-\xi V+\frac{1}{2}\xi^2V^2-\frac{1}{6}\xi^3V^3+\ldots\right)\cdot\right.\\ &\quad\left.\left(1+\xi V+\frac{\xi^2V}{2}\int\mathrm{e}^{-\beta U_{12}}\mathrm{d}V_2+\frac{\xi^3V}{6}\int\int\mathrm{e}^{-\beta U_{123}}\mathrm{d}V_2\mathrm{d}V_3+\ldots\right)\right) && (9.20)\end{aligned}$$

Since $\Omega=-pV$ and using $\ln(1+x)=x-x^2/2+x^3/3\ldots$ we get the expansion

$$p = T\sum_{n=1}^{\infty}\frac{J_n}{n!}\xi^n \tag{9.21}$$

where we fixed the position of particle 1 and

$$\begin{aligned} J_1 &= 1 && (9.22)\\ J_2 &= \int\left(\mathrm{e}^{-\beta U_{12}}-1)\right)\mathrm{d}V_2 && (9.23)\\ J_3 &= \int\int\mathrm{d}V_2\mathrm{d}V_3\left(\mathrm{e}^{-\beta U_{123}}-1\right)-3V\int\mathrm{d}V_2\left(\mathrm{e}^{-\beta U_{12}}-1\right) && (9.24)\end{aligned}$$

We note that J_3 has a significant contribution only from configurations when all three particles are close to each other. We notice that if the potential U is of finite range (or sufficiently fast decreasing with the distance) then all J_n's are independent of the volume V – they are intensive quantities as they should be.

The average number of particles

$$N = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = V\left(\frac{\partial p}{\partial \mu}\right)_{T,V} \tag{9.25}$$

is equal to ($\frac{\partial \xi}{\partial \mu} = \beta\xi$)

$$N = V\sum_{n=1}^{\infty}\frac{J_n}{(n-1)!}\xi^n \tag{9.26}$$

Solving it with respect to ξ to a given order k and plugging into (9.21) we get the expression for p as a function of T and N/V to the same order k in N/V – the main problem is to actually calculate J_n's...

Writing up to third order

$$\frac{N}{V} = \xi + J_2\xi^2 + \frac{J_3}{2}\xi^3 + \dots \tag{9.27}$$

we invert this relation

$$\xi = \frac{N}{V} - J_2\left(\frac{N}{V}\right)^2 + \left(2J_2^2 - \frac{J_3}{2}\right)\left(\frac{N}{V}\right)^3 + \dots \tag{9.28}$$

Plugging it into the expression

$$\frac{p}{T} = \xi + \frac{J_2}{2}\xi^2 + \frac{J_3}{6}\xi^3 + \dots \tag{9.29}$$

we get

$$\frac{pV}{NT} = 1 - \frac{J_2}{2}\left(\frac{N}{V}\right) + \left(J_2^2 - \frac{J_3}{3}\right)\left(\frac{N}{V}\right)^2 + \dots \tag{9.30}$$

If we write

$$\frac{pV}{NT} = 1 + B_2(T)\left(\frac{N}{V}\right) + B_3(T)\left(\frac{N}{V}\right)^2 + \dots \tag{9.31}$$

then $B_2(T), B_3(T), \dots$ are called virial coefficients.

9.4 Hard disks

The first (unrealistic) example: we assume that the $J_n = (-1)^{n+1}(n-1)!\rho_0^{1-n}$ and then

$$p = T\rho_0 \ln\left(1 + \frac{\xi}{\rho_0}\right) \tag{9.32}$$

Then ($\rho = N/V$)

$$\rho = \frac{\xi}{1 + \frac{\xi}{\rho_0}} \Rightarrow \xi = \frac{\rho}{1 - \frac{\rho}{\rho_0}} \tag{9.33}$$

Hence

$$p = -T\rho_0 \ln\left(1 - \frac{\rho}{\rho_0}\right) \tag{9.34}$$

and the pressure in this example goes to infinity at finite density ρ_0.

The second example is given by a very important model, namely the so called hard spheres model – we will discuss it in two dimensions i.e. hard disks. We have disks of radius R that do not interact except that they are impenetrable so that their centers cannot have smaller distance than $2R$. The value of J_2 is relatively easy to calculate – any position of the second disk that overlaps with the first disk has $U_{12} = \infty$ (and all the other have $U_{12} = 0$) so that

$$J_2 = -\int_0^{2R} 2\pi r \mathrm{d}r = -4\pi R^2 \tag{9.35}$$

Therefore

$$N = V(\xi - 4\pi R^2 \xi^2 + \ldots) \Rightarrow \xi = \frac{N}{V} + \frac{4\pi R^2 N^2}{V^2} + \ldots \tag{9.36}$$

and from (9.30)

$$\frac{pV}{NT} = 1 + \frac{2\pi R^2 N}{V} + \ldots = 1 + 2\eta + \ldots \tag{9.37}$$

where $\eta = \pi R^2 N/V$ is the packing fraction. Since the maximal crystalline packing in 2D is equal to

$$\eta_0 = \frac{\sqrt{3}\pi}{6} \tag{9.38}$$

we expect the pressure grow to infinity when $\eta \to \eta_0$ (when all J_n's should be taken into account) so one can propose

$$\frac{pV}{NT} = \left(1 - \frac{\eta}{\eta_0}\right)^{-1} \left(1 - \frac{2\eta_0 - 1}{\eta_0}\eta\right)^{-1} = \left(1 - 2\eta + \frac{2\eta_0 - 1}{\eta_0^2}\eta^2\right)^{-1} \tag{9.39}$$

to reproduce (9.37) to the order η (when only J_2 is taken into account). The exact equation of state (to all orders in η) for hard disks is unknown.

10 Debye theory of solids

10.1 Equipartition of energy

We start from the statement that every degree of freedom contributes classically $kT/2$ to the energy and discuss the limitations of this statement from quantum mechanics. If the contribution to the hamiltonian from any degree of freedom q is equal to

$$H = \frac{\alpha}{2} q^2 \tag{10.1}$$

then the average energy connected with this degree of freedom is equal to

$$\langle \epsilon \rangle = \frac{\alpha}{2} \frac{\int \mathrm{d}q q^2 \mathrm{e}^{-\beta \alpha q^2/2}}{\int \mathrm{d}q \mathrm{e}^{-\beta \alpha q^2/2}} = \frac{1}{2\beta} = \frac{kT}{2} \tag{10.2}$$

This conclusion is valid even if α depends on other coordinates. Therefore the specific heat equals $k/2$ per degree of freedom

However, as was shown by Einstein in 1907, this conclusion does not hold in quantum mechanics as can be shown on an example of a harmonic oscillator. Classically the specific heat should be equal to k since we have two degrees of freedom. The hamiltonian is given by

$$H = \frac{1}{2}(p^2 + \omega^2 q^2) \tag{10.3}$$

We know that the energies of this system are given by

$$\epsilon_n = \hbar\omega \left(n + \frac{1}{2} \right) \tag{10.4}$$

so that we can calculate average energy:

$$\langle \epsilon \rangle = \hbar\omega \frac{\sum_{n=0}^{\infty} (n + 1/2) \mathrm{e}^{-\beta\hbar\omega(n+1/2)}}{\sum_{n=0}^{\infty} \mathrm{e}^{-\beta\hbar\omega(n+1/2)}} = \frac{\partial}{\partial\beta} \ln \sinh(\beta\hbar\omega/2) = \frac{\hbar\omega}{2} \coth(\beta\hbar\omega/2) \tag{10.5}$$

Calculating the specific heat we get

$$c = -k\beta^2 \frac{\partial \langle \epsilon \rangle}{\partial \beta} = k \frac{(\beta\hbar\omega/2)^2}{\sinh^2(\beta\hbar\omega/2)} \tag{10.6}$$

We see that for large T (small β) we indeed have $c \to k$ but for small T (large β) $c \to 0$. This formula explains the 'freezing' of degrees of freedom: if the first excited state energy is much bigger than kT then this degree of freedom does not contribute to the specific heat. For example for diatomic molecules O_2 and N_2 at room temperature we have only 5 degrees of freedom contributing, 3 translational and 2 rotational, but the oscillational degrees of freedom do not contribute since $\hbar\omega_{osc}$ for these molecules is much bigger than $kT \sim 0.03$ eV for the room temperature and hence the specific heat of air is approximately equal to $5k/2$ per molecule.

10.2 Debye theory of specific heat of solids

Applying the rule of the equipartition of energy to solids we would expect that the specific heat should be equal to $3Nk$ where N is the number of atoms – it corresponds to $3N$ translational and $3N$ oscillational degrees of freedom and is called the Dulong-Petit law. Applying the reasoning from quantum mechanics we would still expect the same result even at room temperature since oscillations in solids have frequencies of the order of ω_{osc}/N so much lower than kT for room temperature. However, as we will discuss, the energy levels in solids have much more complicated structure than for the simple harmonic oscillator so the formulae are more complicated as well.

We will discuss the distributions for phonons to derive the formula for the specific heat of solids related to oscillation but not including the electronic heat capacity (dominant at very low temperatures only).

10.2.1 classical computation

If we have a 1D string of atoms with harmonic potential and equilibrium distance a we have

$$m\ddot{x}_n = K(x_{n+1} - x_n) + K(x_{n-1} - x_n) \tag{10.7}$$

To solve this equation we substitute

$$x_n = na + \mathrm{e}^{\mathrm{i}\omega t}\sin(kna), \qquad -\frac{\pi}{a} \le k \le \frac{\pi}{a} \tag{10.8}$$

to get

$$\omega^2 = \frac{4K}{m}\sin^2(ka/2) \tag{10.9}$$

We write

$$\omega = \omega_m \sin(ka/2), \qquad \omega_m = \sqrt{\frac{4K}{m}} \tag{10.10}$$

The group velocity

$$v_g = \frac{\partial\omega}{\partial k} = \frac{\omega_m a}{2}\cos(ka/2) = \frac{Ka^2}{m}\cos(ka/2) = v_{g0}\cos(ka/2) \tag{10.11}$$

Density of states in 3D ($k = \frac{2}{a}\arcsin(\omega/\omega_m)$; there are 2 transverse and 1 longitudinal polarizations)

$$g(\omega)\mathrm{d}\omega = \frac{3V4\pi k^2\mathrm{d}k}{(2\pi)^3} = \frac{V\omega^2\mathrm{d}\omega}{2\pi^2 v_{g0}^3}\cdot\frac{1+\omega^2/(3\omega_m^2)+\ldots}{\sqrt{1-\omega^2/\omega_m^2}} \tag{10.12}$$

Einstein has used the formula for density for one specific frequency $g(\omega) = 3N\delta(\omega - \omega_E)$ – it explained the Dulong-Petit law that the heat capacity tends to $3R$ for large temperatures but was not very good in explaining low temperature behavior of heat capacity.

Debye assumed that all frequencies are present and wrote the formula (without any corrections $\sim \omega^2/\omega_m^2$) to use measured v_{g0}) and defined ω_D by

$$3N = \int_0^{\omega_D} g(\omega)\mathrm{d}\omega = \int_0^{\omega_D} \frac{3V\omega^2\mathrm{d}\omega}{2\pi^2 v_{g0}^3} = \frac{V\omega_D^3}{2\pi^2 v_{g0}^3} \tag{10.13}$$

hence

$$\omega_D = v_{g0}\left(6\rho\pi^2\right)^{\frac{1}{3}}, \qquad \rho = N/V \tag{10.14}$$

and

$$g(\omega)\mathrm{d}\omega = \frac{9N\omega^2\mathrm{d}\omega}{\omega_D^3} \tag{10.15}$$

If transverse and longitudinal speeds are different one may use the averaging

$$\frac{3}{\bar{v}^3} = \frac{2}{v_t^3} + \frac{1}{v_l^3} \tag{10.16}$$

Phonons are bosons so that the energy stored in phonons in temperature T is given by

$$E = \int_0^{\omega_D} \mathrm{d}\omega g(\omega)\hbar\omega \frac{1}{\mathrm{e}^{\frac{\hbar\omega}{kT}} - 1} \tag{10.17}$$

The heat capacity

$$c_p = \frac{\partial E}{\partial T} = \int_0^{\omega_D} \mathrm{d}\omega \frac{9N\hbar^2\omega^4}{\omega_D^3 kT^2} \frac{\mathrm{e}^{\frac{\hbar\omega}{kT}}}{\left(\mathrm{e}^{\frac{\hbar\omega}{kT}} - 1\right)^2} \tag{10.18}$$

It can be rewritten as

$$c_p = 9Nk\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \mathrm{d}x \frac{x^4\mathrm{e}^x}{(\mathrm{e}^x - 1)^2} \tag{10.19}$$

where

$$\theta_D = \frac{\hbar\omega_D}{k} \tag{10.20}$$

For $T \ll \theta_D$ we have

$$c_p \to 9Nk\left(\frac{T}{\theta_D}\right)^3 \cdot \frac{4\pi^4}{15} \tag{10.21}$$

while for $T \gg \theta_D$ we recover the Dulong-Petit law $c_p \to 3R$.

This formula is in much better agreement with experimentally measured values than Einstein's but is not exact either. To have better description one has to take into account the presence of (quantum) characteristic frequencies of a given crystal or dependence of θ_D on temperature. The Debye temperatures of some of the elements (they decrease with the temperature to match the experimental values!): aluminum 433 K, beryllium 1481 K, copper 347 K, lead 105 K, gold 227 K, diamond 2200 K (in room temperature 1840 K).

10.2.2 quantum counterpart of the computation in 1D

The classical hamiltonian for a harmonic oscillator reads:

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} \tag{10.22}$$

In quantum mechanics x and p are operators where Poisson brackets are replaced by commutators (Dirac 1925 equations of motion, Heisenberg 1925 main idea, Born and Jordan 1925 matrix formulation, Born, Heisenberg and Jordan 1926 textbook)

$$\{x, p\}_{PB} = 1 \Rightarrow [\hat{x}, \hat{p}] = \mathrm{i}\hbar \tag{10.23}$$

When acting on functions we can write

$$\hat{x} f(x) := x f(x), \qquad \hat{p} := -\mathrm{i}\hbar \frac{\partial f}{\partial x} \tag{10.24}$$

and the equation for the eigenstate of the hamiltonian

$$\hat{H} f = E f \tag{10.25}$$

is the Schrödinger equation (1926).

Hamiltonian is an operator as well:

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2 \hat{x}^2}{2} \tag{10.26}$$

It is convenient to introduce creation and annihilation operators

$$a := \frac{m\omega\hat{x} + \mathrm{i}\hat{p}}{\sqrt{2m\omega\hbar}}, \qquad a^\dagger := \frac{m\omega\hat{x} - \mathrm{i}\hat{p}}{\sqrt{2m\omega\hbar}}, \tag{10.27}$$

satisfying

$$[a, a^\dagger] = 1 \tag{10.28}$$

and then

$$\hat{H} = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \tag{10.29}$$

Points in a phase space are replaced by states on which the operators act. The crucial role is played by the vacuum state $|0\rangle$– in this case we define it as a state

$$a|0\rangle = 0 \tag{10.30}$$

We can create other states by acting with $a^\dagger$

$$|n\rangle = \frac{1}{\sqrt{n!}} (a^\dagger)^n |0\rangle, \quad a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad a|n\rangle = \sqrt{n} |n-1\rangle \tag{10.31}$$

where the prefactor is needed for $\langle n|n\rangle = 1$.

We can now calculate 'quantum partition function' ($E_n = \hbar\omega(n + \frac{1}{2})$)

$$Z = \sum_{n=0}^{\infty} \mathrm{e}^{-\beta E_n} = \frac{\mathrm{e}^{-\beta\hbar\omega/2}}{1 - \mathrm{e}^{-\beta\hbar\omega}} \tag{10.32}$$

and (using $\sum_{n=0} n\exp(-ny) = \exp(-y)(1 - \exp(-y))^{-2}$)

$$\langle E\rangle = \frac{1}{Z}\sum_{n=0}^{\infty} \hbar\omega(n + \frac{1}{2})\mathrm{e}^{-\beta E_n} = \frac{\hbar\omega}{2}\frac{\mathrm{e}^{\beta\hbar\omega} + 1}{\mathrm{e}^{\beta\hbar\omega} - 1} \tag{10.33}$$

For $\beta \to 0$ we have $\langle E\rangle \to 1/\beta$ (since we have 2 degrees of freedom) and for $\beta \to \infty$ we have $\langle E\rangle \to \hbar\omega/2$ (since it is the ground state energy). Neither Debye nor Einstein included $\frac{1}{2}$ in their calculations but it drops out anyway when we calculate the specific heat what they were interested in.

It is interesting to note that $\frac{1}{2}$ does not drop out if we calculate average displacement. We have

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}(a^\dagger + a), \qquad \hat{p} = \sqrt{\frac{\hbar m\omega}{2}}\mathrm{i}(a^\dagger - a) \tag{10.34}$$

so that

$$\langle n|\hat{x}^2|n\rangle = \frac{\hbar}{m\omega}(n + \frac{1}{2}), \quad \langle n|\hat{p}^2|n\rangle = \hbar m\omega(n + \frac{1}{2}), \tag{10.35}$$

and we recover the Heisenberg uncertainty relation (which can be proven in full generality)

$$\sqrt{\langle\hat{x}^2\rangle\langle\hat{p}^2\rangle} \geq \frac{\hbar}{2} \tag{10.36}$$

and it is saturated for the vacuum state.

In higher dimensions D the calculation would be a little different:

$$E_n = \hbar\omega\left(n + \frac{D}{2}\right), \qquad \text{degeneracy} : \frac{(n + D - 1)!}{n!(D - 1)!} \tag{10.37}$$

For example in 2 dimensions

$$Z = \sum_{n=0}^{\infty}(n + 1)\mathrm{e}^{-\beta E_n} = \frac{\mathrm{e}^{\beta\hbar\omega}}{(\mathrm{e}^{\beta\hbar\omega} - 1)^2} \tag{10.38}$$

and the average energy

$$\langle E\rangle = \frac{1}{Z}\sum_{n=0}^{\infty}(n + 1)\hbar\omega(n + 1)\mathrm{e}^{-\beta E_n} = \hbar\omega\frac{\mathrm{e}^{\beta\hbar\omega} + 1}{\mathrm{e}^{\beta\hbar\omega} - 1} \tag{10.39}$$

For $\beta \to 0$ we have $\langle E\rangle \to 2/\beta$ (since we have 4 degrees of freedom) and for $\beta \to \infty$ we have $\langle E\rangle \to \hbar\omega$ (since it is the ground state energy).

10.3 Electronic specific heat in metals

For electrons we have the contribution to the energy density given by the Fermi-Dirac distribution

$$u = \frac{U}{V} = \int_0^\infty \mathrm{d}\epsilon g(\epsilon)\epsilon \frac{1}{\mathrm{e}^{\frac{\epsilon-\mu_F(T)}{T}}+1} \tag{10.40}$$

where $\mu_F(T)$ is the Fermi level given by the density

$$n = \frac{N}{V} = \int_0^\infty \mathrm{d}\epsilon g(\epsilon) \frac{1}{\mathrm{e}^{\frac{\epsilon-\mu_F(T)}{T}}+1} \tag{10.41}$$

(we will denote $\mu_F(T)$ by $\mu(T)$ later on). If we use the non-relativistic formula (assuming that the temperature is low enough) $\epsilon = \hbar^2 k^2/(2m)$ we have the energy density

$$u(T) = \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^\infty \mathrm{d}\epsilon \frac{\epsilon^{3/2}}{\mathrm{e}^{(\epsilon-\mu(T))/T}+1} \tag{10.42}$$

and the density

$$n = \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^\infty \mathrm{d}\epsilon \frac{\epsilon^{1/2}}{\mathrm{e}^{(\epsilon-\mu(T))/T}+1} \tag{10.43}$$

For $T = 0$ we have ($\mu_0 = \mu_F(0)$)

$$n = \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^{\mu_0} \mathrm{d}\epsilon \epsilon^{1/2} = \frac{32\pi m^{3/2}}{3\sqrt{2}h^3}\mu(0)^{\frac{3}{2}} \tag{10.44}$$

and

$$u(0) = \frac{3}{5}n\mu(0) \tag{10.45}$$

If we have non-zero temperature but very small ($T \ll \mu_0$) we use the method of Sommerfeld. We introduce a new variable $x = (\epsilon - \mu(T))/T$ and we write

$$\begin{aligned}
n &= \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_{-\frac{\mu(T)}{T}}^\infty \mathrm{d}xT\frac{(xT+\mu(T))^{1/2}}{\mathrm{e}^x+1} \\
&= \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^{\frac{\mu(T)}{T}} \mathrm{d}xT(-xT+\mu(T))^{1/2}\left(1-\frac{1}{\mathrm{e}^x+1}\right) \\
&\quad + \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^\infty \mathrm{d}xT(xT+\mu(T))^{1/2}\frac{1}{\mathrm{e}^x+1} \\
&= \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\left(\frac{2}{3}\mu(T)^{\frac{3}{2}} + T^2\mu(T)^{-\frac{1}{2}}\int_0^\infty \mathrm{d}x\frac{x}{\mathrm{e}^x+1} + O(T^4/\mu(T)^{5/2})\right)
\end{aligned} \tag{10.46}$$

up to exponentially small terms. Since the integral is equal to $\zeta(2)(1-2/4) = \pi^2/12$ and the result has to be equal to (10.44) therefore we get

$$\frac{2}{3}\mu(0)^{\frac{3}{2}} = \frac{2}{3}\mu(T)^{\frac{3}{2}} + T^2\mu(T)^{-\frac{1}{2}}\frac{\pi^2}{12} \tag{10.47}$$

and we get

$$\mu(T)=\mu(0)\left(1-\frac{\pi^2T^2}{12\mu(0)^2}+O(T^4/\mu(0)^4)\right) \tag{10.48}$$

Repeating the same steps for the energy (10.42) we get

$$\begin{aligned} u &= \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_{-\frac{\mu(T)}{T}}^{\infty}\mathrm{d}xT\frac{(xT+\mu(T))^{3/2}}{\mathrm{e}^x+1} \\ &= \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^{\frac{\mu(T)}{T}}\mathrm{d}xT(-xT+\mu(T))^{3/2}\left(1-\frac{1}{\mathrm{e}^x+1}\right) \\ &\quad +\frac{16\pi m^{3/2}}{\sqrt{2}h^3}\int_0^{\infty}\mathrm{d}xT(xT+\mu(T))^{3/2}\frac{1}{\mathrm{e}^x+1} \\ &= \frac{16\pi m^{3/2}}{\sqrt{2}h^3}\left(\frac{2}{5}\mu(T)^{\frac{5}{2}}+3T^2\mu(T)^{\frac{1}{2}}\int_0^{\infty}\mathrm{d}x\frac{x}{\mathrm{e}^x+1}+O(T^4/\mu(T)^{5/2}\right) \end{aligned} \tag{10.49}$$

Therefore using (10.47) we get up to T^2 terms

$$u\sim u(0)\left(1-\frac{5\pi^2T^2}{24\mu(0)^2}\right)\left(1+\frac{5\pi^2T^2}{8\mu(0)^2}\right)\sim u(0)\left(1+\frac{5\pi^2T^2}{12\mu(0)^2}\right) \tag{10.50}$$

where $u(0)=\frac{3}{5}n\mu(0)$. Therefore the heat capacity at very low temperatures

$$C_p=\frac{\partial u}{\partial T}=n\frac{\pi^2T}{2\mu(0)} \tag{10.51}$$

Actually this formula has to be corrected by the number of electrons in the conduction band per atom v. Therefore the final formula is

$$C_p=\frac{\partial u}{\partial T}=nv\frac{\pi^2T}{2\mu(0)} \tag{10.52}$$

so it is linear in temperature and much smaller at room temperature than naively expected ($C_p\sim nv$) by the factor $\pi^2T/(2\mu(0))\sim 10^{-2}$. It is however important in comparison with the phonon heat capacity at low temperatures and the total heat capacity of solids at small temperatures is given by

$$c_p=\alpha T+\beta T^3 \tag{10.53}$$

We can obtain experimental α and β by making a figure of measured c_p/T at low temperatures as a function of T^2 which is approximately a straight line with intercept α and slope β.

11 Thermodynamics of magnetic systems

11.1 Thermodynamics of the magnetic field

It is interesting to note that in the case of a magnetic field described by the vector potential $\mathbf{A}$ the Gibbs-Boltzmann factor gives

$$\exp\left(-\frac{(\mathbf{p}-q\mathbf{A})^2}{2mkT}\right)=\exp\left(-\frac{m\dot{\mathbf{r}}^2}{2kT}\right) \tag{11.1}$$

and it is the same distribution in velocities with or without the magnetic field! This is the paradox that in classical physics bodies should not react to a magnetic field while obviously such a reaction exists - this is solved in quantum mechanics where there are quantized levels (Landau levels) and quantized spin degrees of freedom and the classical Gibbs-Boltzmann factor does not describe the real reaction of the bodies to the magnetic field.

In magnetic systems we can divide any current into free currents $\mathbf{j}_0$ and bound currents

$$\mathbf{j}(\mathbf{x})=\sum_a \mathbf{j}_0(\mathbf{x})\delta(\mathbf{x}-\mathbf{x}_a)+\sum_b \varepsilon_{ijk}M_j\partial_k\delta(\mathbf{x}-\mathbf{x}_b)+\ldots \tag{11.2}$$

where $\nabla\cdot\mathbf{j}_0=0$ and the second term is written in such a way to satisfy this condition automatically and $\ldots$ stand for higher derivatives of the delta function. Neglecting all higher derivatives and introducing continuous current densities we write

$$\mathbf{j}(\mathbf{x})=\mathbf{j}_0(\mathbf{x})+\nabla\times\mathbf{M}(\mathbf{x}) \tag{11.3}$$

If the changes are slow then we have the equation

$$\nabla\times\mathbf{B}=\mu_0\mathbf{j} \tag{11.4}$$

and rearranging we get

$$\nabla\times(\mathbf{B}-\mu_0\mathbf{M})=\mu_0\mathbf{j}_0 \tag{11.5}$$

where on the RHS we have only free currents. We introduce

$$\mathbf{H}:=\frac{1}{\mu_0}\mathbf{B}-\mathbf{M} \tag{11.6}$$

we have

$$\nabla \times \mathbf{H} = \mathbf{j}_0 \tag{11.7}$$

The field $\mathbf{H}$ is very useful in actual applications since most often we control the external currents (in electrostatics we usually control potentials and therefore the analogous field $\mathbf{D}$ is not so useful) but we have to remember that at the microscopic level we should use only the field $\mathbf{B}$.

We can now write down Maxwell equations in the presence of media

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho_0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{H} &= \mathbf{j}_0 + \frac{\partial \mathbf{D}}{\partial t} \end{aligned} \tag{11.8}$$

where $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$ and it is in this form that Maxwell wrote originally his equations. Only later it became clear that the fields $\mathbf{D}$ and $\mathbf{H}$ are secondary and at the microscopic level everything can (and should) be described by the field $\mathbf{E}$ and $\mathbf{B}$ only.

11.1.1 First Law for magnetic systems

It is not a priori obvious what expression we should use in FLT: $\mu_0 \mathbf{H} \mathrm{d}\mathbf{M}$ or $\mu_0 \mathbf{M} \mathrm{d}\mathbf{H}$ and with which sign. We now derive the appropriate formula.

If we have a large permanent magnet oriented along the x axis and consider a small body moving along the same axis approaching from infinity with magnetization $\mathrm{M}(x)$ then the attractive force on the body is equal to

$$\mathrm{M}(x) \frac{\mathrm{d}\mathbf{H}}{\mathrm{d}x} \tag{11.9}$$

We want this process to be quasi-static so we apply the compensating external force so the work done on the body is negative

$$W = -\int_{-\infty}^{a} \mathrm{M} \frac{\mathrm{d}\mathbf{H}}{\mathrm{d}x} \mathrm{d}x = -\int_0^{H(a)} \mathrm{M}\mathrm{d}\mathbf{H} \tag{11.10}$$

But this result is a sum of magnetization work and the displacement so we subtract the displacement work keeping the final magnetization $\mathrm{M}(a)$ so that the work required for the magnetization alone is equal to

$$W_M = -\int_a^{-\infty} \mathrm{M}(a) \frac{\mathrm{d}\mathbf{H}}{\mathrm{d}x} \mathrm{d}x - \int_0^{H(a)} \mathrm{M}\mathrm{d}\mathbf{H} = \int_0^{M(a)} \mathbf{H}\mathrm{d}\mathrm{M} \tag{11.11}$$

so that

$$\mathrm{d}W_M = \mathbf{H}\mathrm{d}\mathrm{M} \tag{11.12}$$

11.1.2 Paramagnetism

We start from a system of magnetic dipoles $\boldsymbol{\mu}$ in the external magnetic field $\mathbf{B}$ so that the energy is equal to

$$U = -\boldsymbol{\mu}\cdot\mathbf{B} \tag{11.13}$$

Then

$$Z_N = \left(\int 2\pi\sin\theta d\theta e^{-\beta\mu B\cos\theta}\right)^N = (4\pi)^N\left(\frac{\sinh\beta\mu B}{\beta\mu B}\right)^N = e^{-\beta F} \tag{11.14}$$

Hence

$$U = \frac{\partial(\beta F)}{\partial\beta} = -\frac{\partial\ln Z_N}{\partial\beta} = -N\mu B\left(\coth\beta\mu B - \frac{1}{\beta\mu B}\right) \tag{11.15}$$

For large T we get (for $x\to 0$ we have $\coth(x) = x^{-1} + x/3 + \ldots$)

$$U = -\frac{N\mu^2\beta B^2}{3} \tag{11.16}$$

Then the magnetization M

$$M = -\frac{U}{B} = \frac{N\mu^2 B\beta}{3} = \frac{N\mu^2 B}{3kT} \tag{11.17}$$

Therefore the magnetic susceptibility is inversely proportional to the temperature

$$\chi = \frac{M}{NB} = \frac{\mu^2}{3kT} \tag{11.18}$$

what is known as the Curie law (Pierre Curie established it in 1895, the year of the marriage with Maria Skłodowska, discovery of X-rays by Wilhelm Röntgen, and the beginning of the work on radioactivity by Maria that Pierre joined soon after).

In the quantum case

$$\mu = g\mu_B j \tag{11.19}$$

where j is the total angular momentum, g is a gyromagnetic ratio (equal to 2 for pure electron spin $j = 1/2$ and 1 for pure orbital moment $j = l$), μ_B is a Bohr magneton

$$\mu_B = \frac{e\hbar}{2m_e c} \tag{11.20}$$

(in nuclear physics one uses the nuclear magneton where $m_e \to m_p$). Then the magnetization is equal to

$$M_j = N\mu\frac{\sum_{m=-j}^{m=j}\frac{m}{j}\exp(m\beta\mu H/j)}{\sum_{m=-j}^{m=j}\exp(m\beta\mu H/j)} \tag{11.21}$$

The sum in the denominator can be easily calculated and we get

$$M_j = N\mu\left((1+1/(2j))\coth(1+1/(2j))\beta\mu H - 1/(2j)\coth(1/(2j))\beta\mu H\right) \tag{11.22}$$

For $j = \frac{1}{2}$ we get a very simple result

$$M_{1/2} = N\mu \tanh \beta\mu H \tag{11.23}$$

what can be directly seen from (11.21). For large temperatures we again recover the Curie law but with 3 times bigger magnetic susceptibility than in the classical case (which corresponds to $j \to \infty$).

11.2 Ferromagnetism

Ferromagnetism is a phenomenon in some of the transition metals or its alloys where there is a strong intrinsic magnetic field remaining even after the external field is removed. Its microscopic derivation from first principles is still unknown but there are several half-phenomenological descriptions. The origin of this highly surprising, but known from millenia, behavior is the localized partial occupation of the d and f shells in these atoms that produce a very high orbital spin (s and p shell electrons are mostly delocalized and contribute to the conductivity of these elements). An effective hamiltonian that takes into account also the Pauli exclusion principle is the Heisenberg model proposed in 1928:

$$H = -J \sum_{ij} \mathbf{S}_i \cdot \mathbf{S}_j - \mathbf{B} \cdot \sum_i \mathbf{S}_i \tag{11.24}$$

where the spins are usually treated quantum-mechanically (if the coefficients J are different in different directions we get the so called XXY or XYZ models). For $\mathbf{B} = 0$ and vanishing temperature the ground state is either ferromagnetic $J > 0$ or antiferromagnetic $J < 0$ (however for $T > 0$ the system is neither ferro- nor antiferromagnetic in 1 and 2 dimensions, these properties start from 3 dimensions). There are many generalizations of the Heisenberg model (Hubbard model, $t - J$ model and many others) that are used also for other purposes and they are extensively studied both from the physical and from the mathematical perspectives. There exists also a simplification of the model is the so called Ising model that was proposed earlier (1920) where only the z component of spins is used and it is ferromagnetic for $J > 0$ below some critical temperature – it will be discussed below.

Pierre Weiss proposed to use a change in the derivation of the Curie law to describe the ferromagnetic case (we take here the simplest case $j = 1/2$ as an example) in the framework of the so called mean-field theory. The proposal consists of treating H in (11.23) as only part of the field and add the magnetization itself (with some coefficient)

$$m = \frac{M}{V} = \frac{N}{V}\mu \tanh\left(\beta\mu(H + \alpha m)\right) \tag{11.25}$$

Spontaneous magnetization is when there is non-zero m for vanishing external field H:

$$m = \frac{M}{V} = \frac{N}{V}\mu \tanh(\beta\mu\alpha m) \tag{11.26}$$

For $T \to 0$ $m \to N\mu/V$ so all magnets are oriented in the same direction. When we increase the temperature the magnetization m decreases and at some critical temperature T_c the equation doesn't have any nontrivial solutions and the only solution is $m = 0$. Since $\tanh x \sim x$ for small x so that

$$T_c = \frac{N\mu^2\alpha}{V} \tag{11.27}$$

For $T \to T_c$ from below we expand in $T_c - T$ to the next order ($\tanh x = x - x^3/3 + \ldots$)

$$m = \frac{N\mu^2\alpha m}{V(T_c - (T_c - T))} - \frac{N\mu^4\alpha^3 m^3}{3V(T_c - (T_c - T))^3} + \ldots = m\left(1 + \frac{T_c - T}{T_c}\right) - \frac{\mu^2\alpha^2 m^3}{T_c^2} \tag{11.28}$$

so that

$$m \sim \left(1 - \frac{T}{T_c}\right)^{\frac{1}{2}} \tag{11.29}$$

On the other hand, if $T \to T_c$ from above we expand in H and $T - T_c$:

$$\frac{\mu\alpha m}{T_c} = \frac{\mu(H + \alpha m)}{T} \Rightarrow m \sim \frac{HT_c}{\alpha(T - T_c)} \tag{11.30}$$

It replaces the Curie law and is known as the Curie-Weiss law for ferromagnets.

11.3 Ising model in 1D

The Ising model was proposed by Lenz in 1920 and solved in one dimension, therefore without the phase transition, by E. Ising in 1925 in his doctoral dissertation. The main result is due to L. Onsager in 1944 where the exact partition function for the model in 2D in the absence of the magnetic field B was calculated. In 1952 C.N. Yang has proven an exact formula (earlier announced by L. Onsager and B. Kaufman in 1949) for the first derivative of the partition function of the Ising Model in 2D with respect to B at $B = 0$ (magnetization). There are thousands of papers on the subject trying to include the non-vanishing magnetic field and huge body of results, both numerical and analytical, exists for the Ising on finite lattices. The Ising model in 2D has some features of the genuine ferromagnetic materials so its analysis can give us some insight much deeper than the mean-field theory (that turned out to be incorrect in several respect in comparison to the Ising model).

We consider N ($N > 1$) spins on a line with periodic boundary conditions. We introduce a hamiltonian

$$H = -J\sum_i \sigma_i\sigma_{i+1} - B\sum_i \sigma_i. \tag{11.31}$$

where the first sum runs over closest neighbours only ($\sigma_{N+1} = \sigma_1$). We assume that $J > 0$ and $\sigma_i = \pm 1$. The partition function reads

$$Z = \sum \mathrm{e}^{-\beta H} \tag{11.32}$$

where the sum runs over all configurations.

We introduce the so called transfer matrix.

$$M := \begin{pmatrix} \mathrm{e}^{\beta(J+B)} & \mathrm{e}^{-\beta J} \\ \mathrm{e}^{-\beta J} & \mathrm{e}^{\beta(J-B)} \end{pmatrix} \tag{11.33}$$

where the upper line corresponds to $(+,+)$ and $(+,-)$ neighbouring spins and the lower line to $(-,+)$ and $(-,-)$. Then

$$Z = \mathrm{Tr} M^N = \lambda_1^N + \lambda_2^N \tag{11.34}$$

where λ_1, λ_2 are eigenvalues of M. They can easily calculated from the equations

$$\lambda_1 + \lambda_2 = \mathrm{Tr} M = 2\mathrm{e}^{\beta J}\cosh(\beta B), \qquad \lambda_1 \cdot \lambda_2 = \det M = 2\sinh(2\beta J), \tag{11.35}$$

so that

$$\begin{aligned} \lambda_1 &= \mathrm{e}^{\beta J}\cosh(\beta B) + \sqrt{\mathrm{e}^{2\beta J}\sinh^2(\beta B) + \mathrm{e}^{-2\beta J}}, \\ \lambda_2 &= \mathrm{e}^{\beta J}\cosh(\beta B) - \sqrt{\mathrm{e}^{2\beta J}\sinh^2(\beta B) + \mathrm{e}^{-2\beta J}} \end{aligned} \tag{11.36}$$

In the limit $N \to \infty$ the bigger of the eigenvalues is the only relevant one in Z so that in the thermodynamical limit

$$\ln Z = -\beta F = N \ln \lambda_1 \tag{11.37}$$

11.4 Ising model in 2D

On a 2D square lattice with N 'spins' $\sigma = \pm 1$ we introduce a hamiltonian

$$H = -\frac{J}{2} \sum_{i,j,i\neq j} \sigma_i \sigma_j - B \sum_i \sigma_i. \tag{11.38}$$

where the first sum runs over closest neighbours only. We will assume that $J > 0$ (the ferromagnetic Ising model) and the system is on a square $\sim \sqrt{N} \times \sqrt{N}$ with periodic boundary conditions. We will introduce the notation

$$x := \mathrm{e}^{-\beta J}, \qquad z := \mathrm{e}^{\beta B} \tag{11.39}$$

where $\beta = 1/T$ and we assume that $z \geqslant 1$ what corresponds to the choice of the direction of B.

We define the normalized free energy

$$\mathrm{e}^{-\beta N F_N(x,z)} = Z_N(x,z) = \sum \mathrm{e}^{-\beta(E-E_0)} \tag{11.40}$$

where the sum runs over all configurations and E_0 is the lowest energy corresponding to the configuration with all spins pointing in the direction of B for which $\mathrm{e}^{-\beta E_0} = x^{-N^2} z^N$.

The goal is to calculate the free energy per spin F_N in the thermodynamic limit $N \to \infty$. The advantage of calculating $F_N(x,z)$ over $Z_N(x,z)$ is that one has to include only 'connected diagrams'.

A simple case of $J = 0$ gives immediately the result

$$-\beta F_N(1,z) = \ln\left(1 + \frac{1}{z^2}\right) \tag{11.41}$$

The low temperature expansion (which starts from a special configuration with all spins directed in the direction of B) consists of specially ordered contributions from configurations with more and more inverted spins. For N sufficiently large so that the periodic boundary conditions do no not play a role up to a given order we have

$$\begin{aligned} \mathrm{e}^{-\beta N F_N(x,z)} &= 1 + \frac{N}{z^2}x^8 + \frac{N}{2z^4}\left((N-5)x^{16} + 4x^{12}\right) + \ldots \\ &= \left(1 + \frac{x^8}{z^2} + \frac{x^{12}(1-x^4)2}{z^4} + \frac{x^{16}(1-x^4)(-8x^4+6)}{z^6} + \ldots\right)^N \end{aligned} \tag{11.42}$$

where we ordered terms by increasing powers of z^{-2} and the polynomial in parentheses is finite with the last term equal to z^{-2N}.

The task is to calculate the thermodynamic limit of the expression in parentheses

$$\exp(f^\infty(x,z)) = \lim_{N\to\infty} \exp(-\beta F_N(x,z)) = \exp\left(\sum_{m,k} C_{m,k}\, x^{4m} z^{-2k}\right) \tag{11.43}$$

For example

$$\exp(f^\infty(1,z)) = 1 + \frac{1}{z^2} \tag{11.44}$$

The famous result of Onsager gives the full result for the case $z = 1$ ($B = 0$):

$$\exp(f^\infty_{Ons}(x)) = \lim_{z\to 1} \exp(f^\infty(x,z)) = (1+x^4)\exp\left(-\sum_{n=1}^{\infty}\left(\frac{(2n)!}{(n!)^2}\right)^2 \frac{1}{4n}\left(\frac{y}{4(1+y)^2}\right)^n\right) \tag{11.45}$$

where

$$y = \frac{4}{\left(\frac{1}{x^2} - x^2\right)^2} \tag{11.46}$$

In the original article the result was expressed in terms of elliptic functions. The beginning of the expansion:

$$\exp(f^\infty_{Ons}(x)) = 1 + x^8 + 2x^{12} + 5x^{16} + 14x^{20} + 44x^{24} + 152x^{28} + 566x^{32} + \ldots \tag{11.47}$$

The result of Yang for the magnetization at $B = 0$

$$z\frac{\partial f^\infty(x,z)}{\partial z}\bigg|_{z=1} = (1-y^2)^{\frac{1}{8}} - 1 \tag{11.48}$$

although extremely simple was obtained by a very complicated method and it serves as yet another check on the results. The result for susceptibility $z\partial_z(z\partial_z(f^\infty(x,z)))|_{z=1}$ is not known analytically but only as a beginning of an expansion in x^4.

The phase transition for $B = 0$ occurs for a temperature when the expression in parentheses in (11.45) diverges:

$$y = 1 \Rightarrow x_c^4 + x_c^{-4} = 6 \Rightarrow x_c^4 = 3 - 2\sqrt{2} \tag{11.49}$$

The formula for the partition function when $B \neq 0$ is not analytically known.