

Introduction to the Gibbs formalism

(canonical ensemble)

Gabriel T. Landi

References and notes

Most books introduced the canonical ensemble using a different route than the one we will use here. They start with a theory called the microcanonical ensemble. We will study this theory later on in the course.

The derivation of the canonical ensemble, as is done here, may also be found in:

- Feynman, Statistical Mechanics: a set of lectures. chap 1
- Landau, Statistical Physics, chap 1.

As for the applications (page 10 onwards), you may consult

- Salinas, chapter 6
- Reif, chapter 7

Formulation of the problem

In order to completely specify the "state" of a physical system, one must know all its coordinates and momenta. This is what we call a microstate of the system.

$g = \text{microstate} = \text{specification of all coordinate and momenta}$

(1)

If your system is composed of a single particle, then

Single particle: $g = (x_1, x_2, x_3, p_1, p_2, p_3)$

(2)

If you have N particles then

N particles: $g = (x_1, \dots, x_{3N}, p_1, \dots, p_{3N})$

(3)

The microstate of a system of N particles is specified by giving $3N$ coordinates and $3N$ momenta.

In the context of classical mechanics, it is possible to associate to each microstate g a corresponding energy (or Hamiltonian)

$H(g) = E(g) = \text{energy of microstate } g$

(4)

[The extension of these ideas to quantum systems will be done later on].

Example: a harmonic oscillator

The microstate of a harmonic oscillator in 1D is $g = (x, p)$.
The Hamiltonian is

$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2 \quad (5)$$

— " —

The Gibbs formalism models the contact of a system with a heat bath

Heat bath: a very large system at a constant temperature T (I always imagine a bucket of water)

(6)

The bath must always be much larger than the system. So if the "system" is a DNA molecule, a glass of water is already very large. But if the system is a ping-pong ball, then perhaps it is better to use a bath tub.

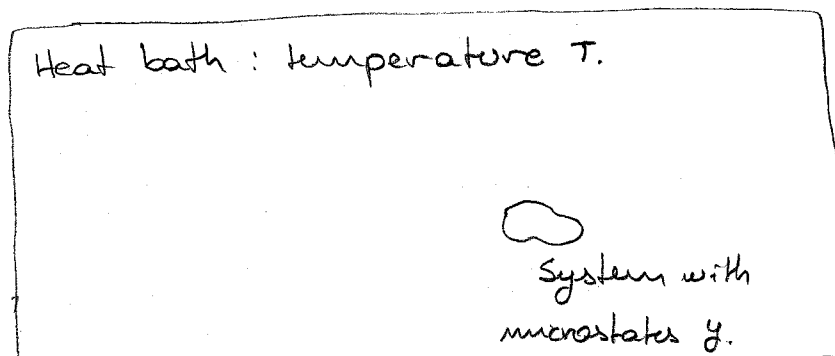
The contact between the system and the bath will cause the system to change randomly. The molecules in the bath will continually collide with the system and exchange energy with it. Consequently, the microstate g of the system will be a random variable.

This process is precisely a Markov chain: the microstate g will keep on jumping from one possible state to the other (In this case both time and g are continuous, but the idea is the same).

As we have learned in our study of Markov chains, after some time has passed, the system will eventually converge to a steady-state (we assume that the chain is irreducible and regular). In the context of a system connected to a heat bath, we call the steady-state "thermal equilibrium".

Thermal equilibrium: the system is said to be in equilibrium with the bath when the Markov chain of y_t reaches a steady-state (7)

The goal of the Gibbs formalism is to give a formula for the probabilities $P(y)$ for finding the system in a given microstate y in thermal equilibrium. You see, $P(y)$ is therefore the steady-state distribution of the Markov chain.



The postulate of equal a priori probabilities

It is a remarkable fact that we can obtain an expression for $P(\gamma)$ using only one basic postulate:

All microstates γ with the same energy $H(\gamma)$ will be equally likely

(8)

This is called the postulate of equal a priori probabilities. We can also state it as follows

$P(\gamma) = f(H(\gamma)) = \text{function only of } H(\gamma)$

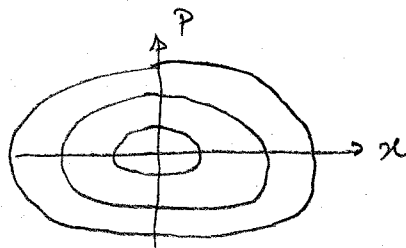
(9)

This is a postulate. It is not something we can demonstrate. It is something we assume in order to construct our theory. Notwithstanding, you must admit it is quite intuitive

Example: harmonic oscillator again

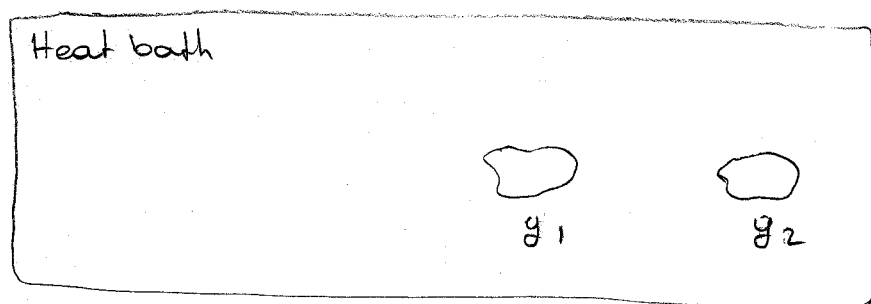
$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2$$

the contours of constant energy form ellipses on the (x, p) plane



what Eq (8) says is that the probabilities depend only on which ellipse you are on, and not on which part of the ellipse you are.

Suppose now, just for the sake of argument, that our system is composed of two disconnected parts, both immersed in the bath



The first part has microstate y_1 and energy $H_1(y_1)$, whereas the second part has microstate y_2 and energy $H_2(y_2)$. If we think of both systems as just one single system, then its microstate will be $y = (y_1, y_2)$. Moreover, since the two systems do not interact, the total energy will be simply the sum of the individual energies

$$H(y) = H_1(y_1) + H_2(y_2) \quad (10)$$

But if the two systems do not interact, we expect that they will be statistically independent. This means that we expect that

$$P(y) = P(y_1, y_2) = P(y_1)P(y_2) \quad (11)$$

But from Eq (9) we must also have

$$P(y_1, y_2) = f(H(y_1, y_2)) = f(H_1(y_1) + H_2(y_2))$$

$$P(y_1) = f(H_1(y_1)) \quad (12)$$

$$P(y_2) = f(H_2(y_2))$$

Here f is an unknown function, which we want to determine. From Eq (11) we see that it must satisfy

$$f(H_1 + H_2) = f(H_1) f(H_2) \quad (13)$$

This is a functional equation determining f . You should now have in mind the exponential:

$$e^{x+y} = e^x e^y \quad (14)$$

Let us then try to see what happens if we choose the ansatz

$$P(y_i) = \frac{e^{\alpha_i H_i(y_i)}}{Z_i} \quad (15)$$

where Z_i and α_i are some constants. If we now look at (11) we get

$$P(y_1) P(y_2) = \frac{e^{\alpha_1 H_1(y_1) + \alpha_2 H_2(y_2)}}{Z_1 Z_2} \quad (16)$$

We see that our ansatz (15) will only work if $\alpha_1 = \alpha_2$. Thus

$$P(y_i) = \frac{e^{\alpha H_i(y_i)}}{Z_i} \quad (17)$$

where α is a constant that must be the same for all systems in the same heat bath.

With some reflection it is possible to conclude that (17) is actually the most general function satisfying (11). Any other function that you come up with may be recast in the form of Eq (17) by a suitable redefinition of α and Z_i .

It is also worth noting that (17) is a direct consequence of the postulate of equal a priori probabilities, Eqs (8) or (9).

In order to arrive at Eq (17) we used the idea of having a system composed of two non-interacting parts. But now we may go back into considering just a single system in contact with a bath. Our main result thus far is that

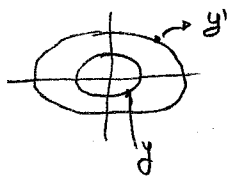
$$P(y) = \frac{e^{\alpha H(y)}}{Z} \quad (18)$$

where Z and α are constants to be determined.

Z is easy: it is just a normalization constant. As for α , here is what we can say about it. Consider two microstates y and y' and let us look at

$$\frac{P(y')}{P(y)} = \frac{e^{\alpha H(y')}}{e^{\alpha H(y)}} = e^{\alpha [H(y') - H(y)]} \quad (19)$$

Suppose $H(y') > H(y)$. For instance, in the harm. oscillator example of page 4, $H(y')$ would be located in a wider ellipse than $H(y)$



looking at (19) we see that

$$\text{If } \alpha > 0 \text{ then } P(y') > P(y) \quad (20)$$

this would mean that it would be more likely to find the system in a high energy state. A system in a bath would always want more. It would capture all the energy of the bath and energize like a power ranger.

This is clearly absurd and we therefore conclude that in order for the theory to make sense, we must have $\alpha < 0$. Thus we write

$$\alpha = -\beta \quad \text{where } \beta > 0 \quad (21)$$

But what can we say about β . Thus far everything has been a theoretical construct. Now experiment must come in. We need to compare our theory with known experimental results. (we will do that along this course). In doing so we will conclude that

$$\beta = \frac{1}{k_B T} \quad (22)$$

where T is the temperature of the bath and k_B is the Boltzmann constant

$$\begin{aligned} k_B &= 1.380648 \times 10^{-23} \text{ J/K} \\ &= 8.617332 \times 10^{-5} \text{ eV/K} \end{aligned} \quad (23)$$

k_B is a constant which converts from Kelvin to energy units such as J or eV. then $[k_B T] = \text{energy}$ and βH is dimensionless.

In summary, the equilibrium probabilities are

$$P(y) = \frac{e^{-\beta H(y)}}{Z}, \quad \beta = \frac{1}{k_B T} \quad (24)$$

The constant Z is called the partition function. It is used to ensure that the probabilities $P(y)$ are correctly normalized:

$$\int P(y) dx_1 \dots dx_{3N} dp_1 \dots dp_{3N} = 1 \quad (25)$$

We must then have

$$\frac{1}{Z} \int e^{-\beta H(y)} dx_1 \dots dx_{3N} dp_1 \dots dp_{3N} = 1$$

so

$$Z = \int e^{-\beta H(y)} dx_1 \dots dx_{3N} dp_1 \dots dp_{3N} \quad (26)$$

If you ever get bored of writing all these differentials, you may instead write

$$Z = \int e^{-\beta H(y)} \mathcal{D}y \quad (27)$$

where $\mathcal{D}y$ is just an abbreviation.

The letter Z comes from the German word Zustandssumme, which means "sum over states". This is a good way of remembering what Z is: it is a sum (integral) over all states (the microstates y).

Later on I will discuss the limits of validity of this formalism. But first let us apply it to a few simple examples.

Example: a particle in a box

Consider a particle trapped to move inside a box of volume V . The Hamiltonian will only contain the kinetic energy

$$H = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m} \quad (28)$$

The partition function will then be

$$\begin{aligned} Z_1 &= \int e^{-\beta H} dx dy dz dp_x dp_y dp_z \\ &= \underbrace{\int dx dy dz}_{V = \text{volume of the box}} \int e^{-\beta(p_x^2 + p_y^2 + p_z^2)/2m} dp_x dp_y dp_z \\ &= V \int_{-\infty}^{\infty} e^{-\beta p_x^2/2m} dp_x \int_{-\infty}^{\infty} e^{-\beta p_y^2/2m} dp_y \int_{-\infty}^{\infty} e^{-\beta p_z^2/2m} dp_z \end{aligned}$$

The 3 integrals are identical and they are Gaussian integrals. Using

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad (29)$$

we get

$$\int_{-\infty}^{\infty} e^{-\beta p_x^2/2m} dp_x = \sqrt{\frac{2\pi m}{\beta}} = \sqrt{2\pi m k_B T} \quad (30)$$

Thus the partition function will be

$$Z_1 = V (2\pi m k_B T)^{3/2} \quad (31)$$

The Gibbs probability distribution will then be

$$P(x, y, z, p_x, p_y, p_z) = \frac{e^{-\beta(p_x^2 + p_y^2 + p_z^2)/2m}}{\sqrt{(2\pi m k_B T)^{3/2}}} \quad (32)$$

Notice that the probabilities for the position and momenta factor as a product:

$$P(x, y, z, p_x, p_y, p_z) = P(x, y, z) P(p_x, p_y, p_z) \quad (33)$$

where

$$P(x, y, z) = \frac{1}{V} \quad (34)$$

$$P(p_x, p_y, p_z) = \frac{e^{-\beta(p_x^2 + p_y^2 + p_z^2)/2m}}{(2\pi m k_B T)^{3/2}} \quad (35)$$

The positions are uniformly distributed throughout the box. You are equally likely to find the particle anywhere. As for the momenta, we see that $P(p_x, p_y, p_z)$ itself factors as a product of 3 terms

$$\begin{aligned} P(p_x, p_y, p_z) &= \frac{e^{-\beta p_x^2/2m}}{\sqrt{2\pi m k_B T}} \frac{e^{-\beta p_y^2/2m}}{\sqrt{2\pi m k_B T}} \frac{e^{-\beta p_z^2/2m}}{\sqrt{2\pi m k_B T}} \\ &= P(p_x) P(p_y) P(p_z) \end{aligned} \quad (36)$$

where

$$P(p_i) = \frac{e^{-\beta p_i^2/2m}}{\sqrt{2\pi m k_B T}} \quad (37)$$

We see that the momenta are normally distributed

$$p_i \sim N(0, \sigma^2) \quad \text{with} \quad \sigma^2 = \frac{m}{\beta} = m k_B T \quad (38)$$

This makes sense: the higher the temperature, the higher is the momentum dispersion. We know that if $X \sim N(0, \sigma^2)$ then $\text{Var}(X) = \langle X^2 \rangle = \sigma^2$. Thus we conclude that

$$\langle p_i^2 \rangle = m k_B T \quad (39)$$

It is also customary to write this as

$$\boxed{\frac{\langle p_i^2 \rangle}{2m} = \frac{k_B T}{2}} \quad (40)$$

The average energy of the system will then be, from (28)

$$\boxed{\langle H \rangle = \frac{\langle p_x^2 \rangle}{2m} + \frac{\langle p_y^2 \rangle}{2m} + \frac{\langle p_z^2 \rangle}{2m} = \frac{3}{2} k_B T} \quad (41)$$

An ideal gas: N particles in a box

Suppose now that we have N particles in a box and suppose that the box is very large that they do not interact. The Hamiltonian of the system will be

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} = \sum_{i=1}^N \frac{(p_{x_i}^2 + p_{y_i}^2 + p_{z_i}^2)}{2m} \quad (42)$$

Let us now compute the partition function. To practice, suppose first that $N=2$ (only 2 particles in the box). Then

$$\begin{aligned} Z_2 &= \int e^{-\beta H} d^3r_1 d^3r_2 d^3p_1 d^3p_2 \\ &= \int e^{-\beta H_1} d^3r_1 d^3p_1 \int e^{-\beta H_2} d^3r_2 d^3p_2 \end{aligned} \quad (43)$$

I know this looks ugly. But I wanted you to see that, since the particles do not interact, the partition function will simply be a product of the individual partition functions. Please remember this. It may save your life one day

Energy is a sum $\Rightarrow Z$ and z are products

(44)

thus, if we have N particles we may simply use (31) and write

$$Z_N = z_1^N = V^N (2\pi m k_B T)^{3N/2}$$

(45)

what about the energy. Well, according to (42) the energy is a sum and the $\langle \rangle$ -operation is linear so

$$\langle H \rangle = \sum_{i=1}^N \langle H_i \rangle = \sum_{i=1}^N \frac{3}{2} k_B T = \frac{3N}{2} k_B T \quad (46)$$

when we have a system of N interacting particles, we can do all calculations for just a single particle, as if the others were not even there. This is the biggest triumph of the Gibbs formalism.

An expression relating U and Z

The average energy is, by definition

$$\langle H \rangle = \int H(y) \mathcal{P}(y) \mathcal{D}(y) = \frac{1}{Z} \int H(y) e^{-\beta H(y)} \mathcal{D}y \quad (47)$$

In the jargon of thermodynamics we call $\langle H \rangle$ the internal energy and we reserve for it the symbol U

$$U = \langle H \rangle = \text{internal energy} = \text{average energy} \quad (48)$$

To compute U using Eq (47) requires the computation of 2 integrals: one for Z and the other for $H e^{-\beta H}$. Integrals are hard! There is a neat trick to obtain U directly from Z and therefore obtain U with just a single integral. Looking at (47) we note that

$$\begin{aligned} \int H(y) e^{-\beta H(y)} \mathcal{D}y &= - \frac{\partial}{\partial \beta} \int e^{-\beta H(y)} \mathcal{D}y \\ &= - \frac{\partial Z}{\partial \beta} \end{aligned}$$

thus

$$U = - \frac{1}{Z} \frac{\partial Z}{\partial \beta} \quad (49)$$

we may also write this as

$$U = - \frac{\partial \ln Z}{\partial \beta} \quad (50)$$

For instance, in (45) we found that

$$Z_N = V^N (2\pi m k_B T)^{3N/2} = V^N \left[\frac{2\pi m}{\beta} \right]^{3N/2}$$

so

$$\begin{aligned} \ln Z_N &= \ln \left[V^N (2\pi m)^{3N/2} \right] - \ln \beta^{3N/2} \\ &= \ln \left[V^N (2\pi m)^{3N/2} \right] - \frac{3N}{2} \ln \beta \end{aligned}$$

the first term does not depend on β so

$$U = -\frac{\partial}{\partial \beta} \ln Z = \frac{3N}{2} \frac{\partial}{\partial \beta} \ln \beta = \frac{3N}{2} \frac{1}{\beta}$$

or

$$U = \frac{3N}{2} k_B T$$

which is precisely Eq (46).

Eq (50) is very useful. I always recommend you use it. Notice that it also matches well with the idea of Eq (44): if

$$Z_N = z_1^N$$

then

$$U = -\frac{\partial}{\partial \beta} \ln Z_N = -N \frac{\partial}{\partial \beta} \ln z_1 = N U_1$$

If the z are products, then U must be a sum.

the heat capacity and the specific heat

Another quantity of great importance in thermodynamics and statistical mechanics is the heat capacity. It is defined as

$$C = \frac{\partial U}{\partial T}$$

(51)

It represents the sensitivity of U to changes in T . For the ideal gas $U = \frac{3}{2} N k_B T$ [Eq (46)] so

$$C = \frac{3 N k_B}{2}$$

(52)

The units of C are clear from either (51) or (52): it has the units of k_B , or J/K .

Both U and C are extensive quantities: they increase with the number of particles N in the system. It is often useful to consider these quantities per particle. We write them as

$$u = \frac{U}{N} \quad \text{and} \quad c = \frac{C}{N} \quad (53)$$

the heat capacity per particle is called the specific heat

$$C = \frac{\partial U}{\partial T} = \text{heat capacity}$$
$$c = \frac{C}{N} = \frac{\partial u}{\partial T} = \text{specific heat}$$

(54)

Example: $3N$ harmonic oscillators

Later on we will study the specific heat of solids. We will learn that, to a certain approximation, the solid may be modeled as a collection of $3N$ independent harmonic oscillators, with Hamiltonian

$$H = \sum_{i=1}^{3N} H_i = \sum_{i=1}^{3N} \left[\frac{p_i^2}{2m} + \frac{1}{2} m \omega^2 x_i^2 \right] \quad (55)$$

Since the oscillators are independent, we may start our investigation by considering just what happens to a single harmonic oscillator, with

$$H_1 = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2 \quad (56)$$

The partition function will be

$$z_1 = \int e^{-\beta H} dx dp = \int e^{-\beta p^2/2m} dp \int e^{-\beta m \omega^2 x^2/2} dx$$

Both are Gaussian integrals. Using (29) we then get

$$z_1 = \sqrt{\frac{2\pi m}{\beta}} \sqrt{\frac{2\pi}{m\omega^2\beta}} \quad (57)$$

Thus, the prob. distribution for one harmonic oscillator will be

$$P(x, p) = \frac{e^{-\beta H}}{z} = \frac{e^{-\beta p^2/2m} e^{-\beta m \omega^2 x^2/2}}{\sqrt{\frac{2\pi m}{\beta}} \sqrt{\frac{2\pi}{m\omega^2\beta}}} \quad (58)$$

Again, we see that this prob. splits as a product

$$P(x, p) = P(x)P(p) \quad (59)$$

where

$$P(x) = \sqrt{\frac{m\omega^2\beta}{2\pi}} e^{-\beta m\omega^2 x^2/2} \quad (60)$$

$$P(p) = \sqrt{\frac{\beta}{2\pi m}} e^{-\beta p^2/2m} \quad (61)$$

We now see that both x and p are Gaussian. The dist. of p is also exactly like that for the ideal gas, Eq (37). We therefore have

$$\langle p^2 \rangle = \frac{m}{\beta} \quad (62)$$

$$\langle x^2 \rangle = \frac{1}{m\omega^2\beta}$$

This gives us the average kinetic and potential energies

$$\frac{\langle p^2 \rangle}{2m} = \frac{k_B T}{2} \quad (63)$$

$$\frac{1}{2} m\omega^2 \langle x^2 \rangle = \frac{k_B T}{2}$$

they are the same. This is a manifestation of something called the equipartition of energy.

The internal energy will be, from (56)

$$U_1 = \langle H_1 \rangle = \frac{\langle p^2 \rangle}{2m} + \frac{1}{2} m\omega^2 \langle x^2 \rangle \quad (64)$$

or

$$U_1 = k_B T \quad (65)$$

Now let us return to the 3N oscillators, E_f (55). All oscillators are identical so each will have an internal energy U_i . The total internal energy will then be

$$U = 3N U_i = 3N k_B T \quad (66)$$

The heat capacity will then be

$$C = \frac{\partial U}{\partial T} = 3N k_B \quad (67)$$

and the specific heat will be

$$c = \frac{C}{N} = 3k_B \quad (68)$$

This is a universal result: any solid, independent of the material, should have a specific heat of $3k_B$ per atom.

This result was first noted by Pierre Dulong and Alexis Petit in 1819. Usually, specific heat is measured in units per gram instead of per atom. Here are some results

	Cu	Pb	Ag	Zn	Al
c (J/gK)	0.386	0.128	0.23	0.39	0.90

To really understand what these numbers mean, suppose you take 1 gram of each material at 300K. Now you place each in a different bucket of water, all buckets at 299K. Then Pb will transfer 0.130 J of energy to the water, whereas Al will transfer 0.90 J. This may seem basic, but it is nonetheless quite interesting: two materials with the same mass will heat up water by different amounts. This is why energy and temperature are different concepts. And what connects them is the heat capacity.

The idea of Dulong and Petit was to look not at the specific heat per gram, but the specific heat per atom. To make the conversion, we proceed in 2 steps. First we convert to mols, by multiplying by the atomic mass. For instance

$$M_{Pb} = 207.2 \text{ g/mol}$$

so

$$c \text{ (J/mol K)} = c \text{ (J/g K)} M_{Pb} =$$

$$= 0.128 \times 207.2$$

$$= 26.58 \text{ J/mol K}$$

Our table is then updated to

	Cu	Pb	Ag	Zn	Al
$c \text{ (J/mol K)}$	24.52	26.58	24.81	25.49	24.28

You see now that the numbers are all very similar. To change to c in units of k_B simply divide by the gas constant

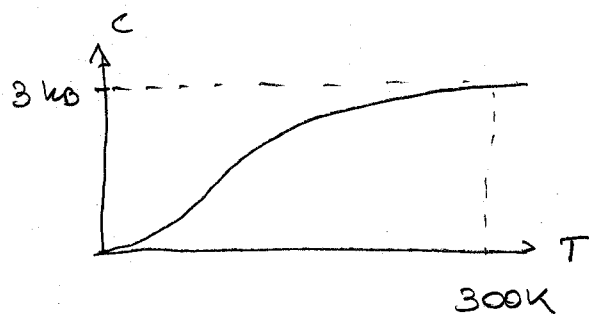
$$R = 8.314 \text{ J/mol K}$$

we then get

	Cu	Pb	Ag	Zn	Al
c/k_B	2.95	3.15	2.98	3.07	2.92

we now see that all results are remarkably similar. This is what led Dulong and Petit to formulate this problem as a law: for a solid, the specific heat per atom should always be $3k_B$.

The law of Dulong and Petit got into some real trouble in the beginning of the 20th century, when cryogenic techniques were invented and allowed scientists to measure the specific heat up to very low temperatures. It was then found that it behaves like this



At 300K we see a roughly constant specific heat (Dulong and Petit). But as we move down the specific heat falls rapidly and tends to zero as $T \rightarrow 0$.

These new effects can only be explained using a quantum theory (we will do just that soon enough).

A gas of interacting particles

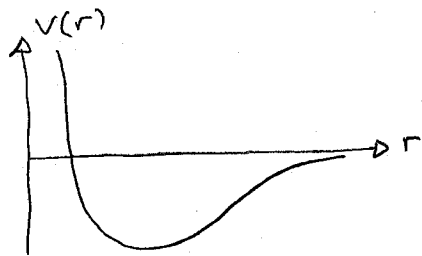
Suppose we have N molecules in a box. The general Hamiltonian describing these particles and their interactions is

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i,j < i} V(r_i, r_j) \quad (69)$$

The second term is the sum of all pair-wise interactions [Experimentally, we find that 3-body interactions are negligible]. The sum is restricted over $j < i$ in order to avoid double counts. We can remove this constraint by dividing by 2:

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \frac{1}{2} \sum_{i,j} V(r_i, r_j) \quad (70)$$

The potentials $V(r_i, r_j)$ could be van der Waals forces between each molecule, or they could describe hard-core repulsion. A typical V will only depend on the distance $r_{ij} = |r_i - r_j|$ between two particles. And it will usually look like this



For large r the interaction is attractive, but if they get too close their electron clouds will touch each other and the interaction becomes repulsive.

The Gibbs distribution corresponding to the Hamiltonian (69) will be

$$P(r_1, \dots, r_N, p_1, \dots, p_N) = \frac{e^{-\beta H}}{Z}$$

Notice how this will factor, as before, in a position term and a momentum term. Thus, we conclude that, even when the particles interact, the momentum p_i of the particles will continue to be statistically independent, with distribution

$$P(p_i) = \frac{e^{-\beta p_i^2/2m}}{(2\pi mk_B T)^{3/2}} \quad (71)$$

[see Eq (35)]. This is quite a remarkable fact and is related to the fact that our theory is classical. In the quantum formalism this is no longer true.

The partition function also factors as

$$Z = \int e^{-\beta H} d^3 r_1 \dots d^3 r_N d^3 p_1 \dots d^3 p_N$$

$$Z = (2\pi mk_B T)^{3N/2} \int d^3 r_1 \dots d^3 r_N \exp\left\{-\frac{\beta}{2} \sum_{ij} V(r_i, r_j)\right\} \quad (72)$$

If you are ever able to compute this integral, let me know. You will likely win a Nobel prize for it. This integral contains all the information about an interacting system of molecules. It therefore contains all the information about phase transitions, from gas to liquid and then from liquid to solid.

The Maxwell - Boltzmann distribution of velocities

we have just seen that, even when the particles interact, the distribution of the momenta will always be statistically independent from the distribution of positions. In particular, if we look at a single particle, its statistical distribution will be given by Eq (35)

$$P(p_x, p_y, p_z) = \frac{1}{(2\pi m k_B T)^{3/2}} \exp\left\{-\frac{p}{2m}(p_x^2 + p_y^2 + p_z^2)\right\} \quad (73)$$

Maxwell and Boltzmann were very interested in this formula because it describes the distribution of momenta in a gas (they derived these formulas before Gibbs' theory, using other methods).

Actually, they preferred to work with velocities: $v_x = \frac{p_x}{m}$. Since this is a 1-to-1 transformation we may use the formula

$$P(v_x, v_y, v_z) = P(p_x, p_y, p_z) \left| \frac{\partial(p_x, p_y, p_z)}{\partial(v_x, v_y, v_z)} \right|$$

the Jacobian in this case is easy

$$\left| \frac{\partial(p_x, p_y, p_z)}{\partial(v_x, v_y, v_z)} \right| = \begin{vmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{vmatrix} = m^3$$

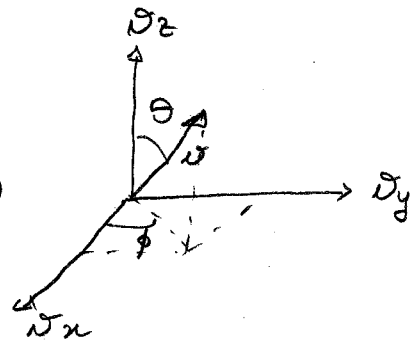
thus

$$P(v_x, v_y, v_z) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp\left\{-\frac{m}{2}(v_x^2 + v_y^2 + v_z^2)\right\} \quad (74)$$

Thus, as expected, the velocities are also independent Gaussians. It is important to notice that this statistical independence is valid only in a non-relativistic theory. If $v_x = 0.9c$, this would restrict the allowed values of v_y and v_z since $|v| = \sqrt{v_x^2 + v_y^2 + v_z^2} < c$. But for sufficiently low velocities we do not need to worry about this.

It is more interesting to look at the distribution of v , the absolute value of the velocity. To do that we change to spherical coordinates

$$\begin{aligned} v_x &= v \sin\theta \cos\phi \\ v_y &= v \sin\theta \sin\phi \\ v_z &= v \cos\theta \end{aligned} \quad (75)$$



the Jacobian of the transformation is $v^2 \sin\theta$ [this is the term that appears when you change variables from $dv_x dv_y dv_z = v^2 \sin\theta dv d\theta d\phi$]. Thus

$$P(v, \theta, \phi) = \left(\frac{m}{2\pi k_B T} \right)^{3/2} v^2 \sin\theta e^{-\frac{m v^2}{2}} \quad (76)$$

Notice that this again factors as a product, meaning that v , θ and ϕ are statistically independent:

$$P(v, \theta, \phi) = \underbrace{4\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} v^2 e^{-\frac{m v^2}{2}}}_{P(v)} \underbrace{\frac{\sin\theta}{2}}_{P(\theta)} \underbrace{2\pi}_{P(\phi)} \quad (77)$$

I already adjust the coefficients to make sure that the 3 distributions are correctly normalized.

$$\int_0^{2\pi} P(\phi) d\phi = \frac{1}{2\pi} \int_0^{2\pi} d\phi = 1$$

$$\int_0^{\pi} P(\theta) d\theta = \frac{1}{2} \int_0^{\pi} \sin\theta d\theta = \frac{1}{2} (-\cos\theta) \Big|_0^{\pi} = 1$$

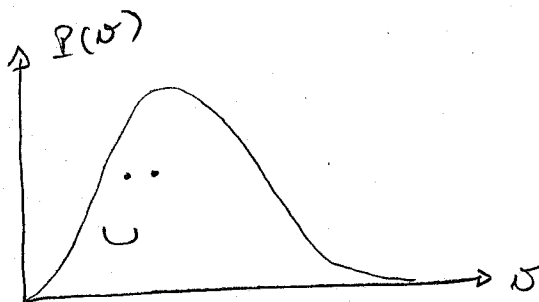
$$\int_0^{\infty} P(v) dv = 1 \quad \leftarrow \text{This is not obvious; need to check.}$$

we see that ϕ is *unif* $(0, 2\pi)$; you can find the velocity in any direction you want. θ , on the other hand, is not uniform. It depends on $\sin\theta$ and is therefore maximum at $\theta = \pi/2$ and zero at $\theta = 0$ or $\theta = \pi$. This is simply due to the fact that the sphere has more points in the equator.

The distribution of velocities is called the Maxwell-Boltzmann distribution

$$P(v) = 4\pi \left(\frac{m}{2\pi k_B T} \right)^{3/2} v^2 \exp \left\{ -\frac{mv^2}{2k_B T} \right\} \quad (78)$$

It looks like this



The most likely speed is the maximum of $P(v)$:

$$v_{\text{most likely}} = \sqrt{\frac{2k_B T}{m}}$$

this is close, although not equal, to $\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} k_B T$.

The Free energy

We have seen that when H is a sum of independent terms, Z will be a product. Consequently $\ln Z$ will also be a sum. It is convenient to define a new quantity, which we call the Helmholtz free energy, as

$$F = -k_B T \ln Z \quad (79)$$

Then, F will be an extensive quantity

$$\text{If } H = H_1 + \dots + H_N \quad \rightsquigarrow \quad F = F_1 + \dots + F_N \quad (80)$$

The term "free energy" comes from thermodynamics. We will understand later on why this energy is "free".

From Eq (50) we see that U and F are related by

$$U = -\frac{\partial}{\partial \beta} \ln Z = -\frac{\partial T}{\partial \beta} \frac{\partial}{\partial T} \ln Z = k_B T^2 \frac{\partial}{\partial T} \ln Z = -T^2 \frac{\partial}{\partial T} (-k_B \ln Z)$$

or

$$U = -T^2 \frac{\partial}{\partial T} \left(\frac{F}{T} \right) \quad (81)$$

This relation is also found from thermodynamics, where it is called the Gibbs - Helmholtz relation.

We may also relate F to the heat capacity (C): with some simplifications, we find

$$C = \frac{\partial U}{\partial T} = -T \frac{\partial^2 F}{\partial T^2} \quad (82)$$

Pressure of an ideal gas

In Eq (45) we found that the partition function of an ideal gas was

$$Z = V^N (2\pi m k_B T)^{3N/2} \quad (83)$$

The Free energy will then be, from Eq (79)

$$F = - N k_B T \ln \left[V (2\pi m k_B T)^{3/2} \right] \quad (84)$$

The student who studied Thermodynamics may recall that the pressure is related to the free energy according to

$$P = - \frac{\partial F}{\partial V} \quad (85)$$

I want to demonstrate this formula from the Gibbs theory. But first let us check it:

$$\begin{aligned} P &= - N k_B T \frac{\partial}{\partial V} (\ln V + \dots) \\ &= - \frac{N k_B T}{V} \end{aligned}$$

Thus we get

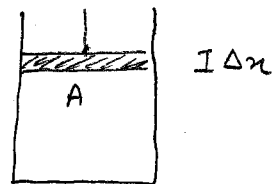
$$pV = N k_B T \quad (86)$$

which is the ideal gas law.

giving us not assume that from Thermodynamics. Indeed, this is all up to now done by (11) formula is microscopic point of view.

Suppose an gas is trapped in a cylinder with a movable piston. If we compress the piston by an amount Δx , the change in energy (mechanical work) will be

$$\Delta E = F \Delta x = p A \Delta x \quad (87)$$



where A = cross section area and p is the pressure. Now, $A \Delta x = -\Delta V$ is the change in volume (the negative sign is because you compressed the cylinder). Thus $\Delta E = -p \Delta V$. Or

$$\text{pressure} = - \frac{\partial H}{\partial V} \quad (88)$$

But this result holds for a given energy $E = H$; i.e., for a given microstate. The actual observed pressure will be an average of p over all possible microstates, viz,

$$p = - \int \frac{\partial H}{\partial V} \frac{e^{-\beta H}}{Z} \mathcal{D}y = \frac{1}{Z} \frac{1}{\beta} \frac{\partial}{\partial V} \int e^{-\beta H} \mathcal{D}y = \frac{k_B T}{Z} \frac{\partial Z}{\partial V}$$

or
$$p = k_B T \frac{\partial \ln Z}{\partial V} = - \frac{\partial}{\partial V} (-k_B T \ln Z)$$

which is exactly Eq (85). qed.

Notice how, when you average $\partial H / \partial V$ you get $\partial F / \partial V$ and not $\partial U / \partial V$. This gives an idea of why the free energy is "free"; not all the energy can be used to produce pressure because, as we will see later, part is consumed as heat. The available part of U to produce pressure is the "free" energy.

The heat capacity is the variance of the energy

Now I want to show you another important result. The heat capacity is defined in Eq (51).

$$C = \frac{\partial U}{\partial T} = \frac{\partial \beta}{\partial T} \frac{\partial U}{\partial \beta} = -\frac{1}{k_B T^2} \frac{\partial U}{\partial \beta} \quad (89)$$

But U in form is given by (50), so

$$\checkmark \quad z' = \partial z / \partial \beta$$

$$C = -\frac{1}{k_B T^2} \frac{\partial}{\partial \beta} \left(-\frac{\partial}{\partial \beta} \ln z \right) = \frac{1}{k_B T^2} \frac{\partial}{\partial \beta} \left(\frac{z'}{z} \right)$$
$$= \frac{1}{k_B T^2} \left[\frac{z''}{z} - \left(\frac{z'}{z} \right)^2 \right]$$

We already know that $z'/z = \frac{\partial}{\partial \beta} \ln z = -\langle H \rangle$. As for the first term, we have

$$\frac{z''}{z} = \frac{1}{z} \frac{\partial^2}{\partial \beta^2} \int e^{-\beta H} Dy = \frac{1}{z} \int H^2 e^{-\beta H} Dy = \langle H^2 \rangle$$

thus, we reach the remarkable result:

$$C = \frac{\partial U}{\partial T} = \frac{1}{k_B T^2} \left[\langle H^2 \rangle - \langle H \rangle^2 \right] \quad (90)$$

The heat capacity is actually proportional to the variance of the energy. It is a measure of the fluctuations of H due to the contact with a heat bath.

The variance of any random variable is always non-negative, so we conclude that

$$C = \frac{\partial U}{\partial T} \geq 0$$

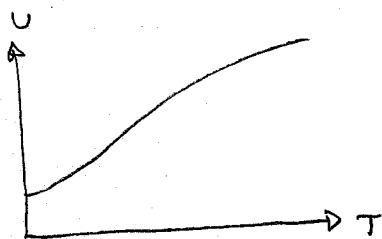
(91)

This means that the slope of $U(T)$ is never negative. Or, we can say that U is a monotonically non-decreasing function of T .

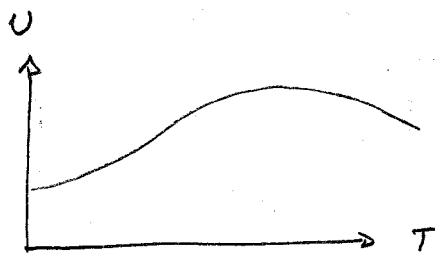
Physically, this means that if you increase T you also increase the energy of the system: hotter systems are always more energetic than colder systems. Eq (91) is therefore a mathematical demonstration of our intuitive idea about temperature.

Of course, all this does not mean that if you double T you will also double U . This would be true when $U \propto T$, which we have seen some examples where this is true, we will soon see many other examples where U will be given by more complicated functions of T . But these functions are always increasing.

Just to be clear, graphically Eq (91) means that



OK



Not OK!

when a system is composed of many independent parts, H will be a sum of terms of the form $H = H_1 + \dots + H_N$. Each H_i is a r.v. and if the subparts do not interact, they will be independent. Thus, according to the central limit theorem, H will be approximately given by a Normal distribution. Its mean is $\langle H \rangle = U$ and, according to (90), its variance will be

$$\text{Var}(H) = \langle H^2 \rangle - \langle H \rangle^2 = Ck_B T \quad (10)$$

Thus,

$$H \sim N(U, Ck_B T) \quad (11)$$

Both U and C are extensive [C is extensive because $C = \partial U / \partial T$ or because of the fact that $\text{Var}(\text{sum}) = \text{sum}(\text{var})$]. Thus we may write $U = Nu$ and $C = Nc$. The mean-squared deviation will then be

$$\frac{\sqrt{\text{Var}(H)}}{\langle H \rangle} = \frac{\sqrt{Nc k_B T}}{Nu} = \frac{1}{\sqrt{N}} \frac{\sqrt{c k_B T}}{u}$$

thus, we see that in the limit $N \rightarrow \infty$ (thermodynamic limit) the relative fluctuations become negligible. This is the law of large numbers: for a macroscopic system the energy tends to a deterministic value.

By the way: we saw that Z played the role of a characteristic function. Now we see that F plays the role of the cumulant generating function (compared to (10)).