

Reading

K&K, Chapter 3. Also, for a little culture, there is a handout which is a one page article from *Science*, 1997, G. Bertsch, vol. 277, p. 1619. It describes melting in clusters consisting of 139 atoms! So $1/\sqrt{N} \approx 10\%$, quite a bit larger than the one part in a trillion we've been talking about when there is a mole of particles! One might expect things to be less well defined, and sure enough, these clusters seem to melt over an appreciable range in temperature (rather than a single T). This may be a case where statistical mechanics is just on, or just past the edge!

The Boltzmann Factor

Let's imagine we have an object which is isolated from the rest of the world, so its energy, entropy, and so on are constant. Furthermore, it has come to equilibrium, so all parts of the object are at the same temperature, pressure, and so on. We imagine dividing this object into two pieces: a small piece, called the system; and a large piece called the heat bath or heat reservoir (or just bath or reservoir). The system is supposed to be sufficiently small that it's useful to think of it as being in a single (quantum) state. A system might be a single atom or molecule, but it could be a larger entity if such an entity can be described as being in a single state. The remainder of the object, the bath, is supposed to be very large and might consist of a very large number of similar atoms or molecules. (In principle, we should let $N \rightarrow \infty$, where N is the number of molecules in the bath.) The system and the bath interact so that the system is in a particular state only with some probability. We are going to calculate this probability. We want to speak of the system as having energy E . This means that the interaction between the system and the bath must be weak in order that we can ascribe a definite energy to the system.

Consider two states of the system, 1 and 2, with energies E_1 and E_2 . When the system is in state 1, the bath has energy $U - E_1$, where U is the total (fixed) energy of the bath plus system. The number of states that correspond to this energy in the system is

$$g(U - E_1) \times 1 = \exp(\sigma(U - E_1)) \times 1.$$

Where the first factor is the number of states in the bath and the second factor is the number of states (just 1) of the system. To have the factor 1 here is the reason that we insisted the system be in a single state. Similarly, the number of states in the case that the system has energy E_2 is

$$g(U - E_2) \times 1 = \exp(\sigma(U - E_2)) \times 1.$$

Our fundamental assumption is that each state (of the system plus bath together) that is compatible with the constraints is equally probable. So, the ratio of the probability

that the system is in state 2 to the probability that it is in state 1 is

$$\frac{P(\text{state 2})}{P(\text{state 1})} = \frac{g(U - E_2)}{g(U - E_1)} = e^{\sigma(U - E_2) - \sigma(U - E_1)} .$$

Now, E is an energy on the scale of an energy of a single molecule while U is an energy typical of a mole. So we can expand σ in a Taylor series around $E = 0$,

$$\sigma(U - E) = \sigma(U) - E \frac{\partial \sigma}{\partial U} + \dots = \sigma(U) - \frac{E}{\tau} + \dots .$$

Note that The first term on the right is $\sim N$ times larger than the second term, and we expect that the first omitted term will be another factor of $\sim 1/N$ smaller. Inserting this expansion into our expression for the probability ratio, we have

$$\frac{P(\text{state 2})}{P(\text{state 1})} = e^{\sigma(U) - E_2/\tau - \sigma(U) + E_1/\tau} = e^{-(E_2 - E_1)/\tau} .$$

The probability ratio is an exponential in the energy difference divided by the temperature. As noted when we discussed the model paramagnetic system, this is called the Boltzmann factor, and we've just shown that this is a general result.

It may seem we got something for nothing: we made a few definitions, did some mathematical hocus-pocus and voila, out pops the Boltzmann factor. A key ingredient is our postulate that all states which satisfy the constraints are equally probable. The number of states goes up with increasing energy U . The rate of increase (in the logarithm) is measured by the temperature, τ . When the system is in a state with energy E , it necessarily received that energy from the heat bath. The more energy the heat bath gives to the system, the fewer states it “has left.” This is what makes higher energy states of the system less probable.

In principle, we should consider an ensemble of identically prepared heat baths and systems. An ensemble in which the probability follows the Boltzmann form ($\propto \exp(-E/\tau)$) is called a *canonical ensemble*.

Systems with Several Forms of Energy

A single system in a canonical ensemble might have several ways to store energy. For example, if the system is a gas molecule, it has energy associated with translation of the center of mass, rotation about the center of mass, and various forms of internal energy such as vibration or other electronic excitations and interactions among spins. If the energies can be assigned independently, then the probabilities are independent. For example, if we want to know the probability that a molecule is in a particular rotational state, with energy E_{rot} , we can think of the bath as including the translational and internal motions of the molecule under consideration as well as all the motions of all the other molecules. Similarly, if we consider the translational motion, with energy E_{tran} , all other forms of energy in this molecule and all other molecules may constitute the heat bath. The probabilities are

$$\begin{aligned} P(E_{\text{tran}}) &\propto e^{-E_{\text{tran}}/\tau} , \\ P(E_{\text{rot}}) &\propto e^{-E_{\text{rot}}/\tau} , \\ &\dots \end{aligned}$$

These probabilities are independent, so the probability that our molecule has a total energy $E_{\text{tot}} = E_{\text{tran}} + E_{\text{rot}} + \dots$ is just the product,

$$P(E_{\text{tot}}) \propto e^{-(E_{\text{tran}} + E_{\text{rot}} + \dots)/\tau} = e^{-E_{\text{tot}}/\tau} .$$

Caveat 1: this is the probability for a *single* state of the system having the specified total energy. If the system has several different ways of distributing E_{tot} among its subsystems, then each way has the same probability (our assumption that each state has the same probability) and the probability of being in any state with that E_{tot} is just the sum. Caveat 2: it may not always be true that the energies can be distributed independently. For example, the moment of inertia, and hence the rotational energy levels, may depend on the amplitude of vibration.

In (older) thermo texts, you will often see expressions like

$$\frac{P(\text{energy } 2)}{P(\text{energy } 1)} = \frac{g_2}{g_1} e^{-(E_2 - E_1)/\tau} ,$$

where g_1 and g_2 are called *statistical weights*. The statistical weight of states with energy E is just the number of distinct states with that energy. This expression agrees with the statement above that the probabilities of the independent states add. Note the difference between the probability of a single state with energy E versus the probability of any state with energy E .

Diversion: the Maxwell Velocity Distribution

Consider a low density gas. Generally, the forces between molecules are only large when the molecules are within a few molecular diameters of each other. The mean separation is many molecular diameters. We expect that a molecule in a low density gas satisfies our definition of a system which is weakly interacting with its heat bath (all the other molecules in the gas).

We consider only the translational energy of a molecule. The probability that a molecule is in a particular state with energy $E = mv^2/2$ is just

$$P(E) \propto e^{-E/\tau} = e^{-mv^2/2\tau} ,$$

where m is the mass and v is the velocity of the molecule. Now we have a problem. This expression applies to a particular state. Usually we think of the energy of translation and the velocity as continuous variables. What constitutes a particular state? Let's postpone the counting of states for a bit and see how far we can get with some reasonable assumptions.

First of all, if we want to treat the energy and velocity as continuous variables, we should be talking about a probability density, where the probability that the energy lies in the range $E \rightarrow E + dE$ is given by $p(E) dE$. This will be the probability of getting E no matter how many different states have this E , and this probability must be proportional to the Boltzmann factor and to the number of states that have this E . In other words,

$$p(E) dE = C e^{-E/\tau} n(E) dE ,$$

where C is a constant whose value is adjusted to ensure that $\int p(E) dE = 1$, and $n(E)$, usually called the *density of states*, is the number of distinct states per unit energy with energy E . That is, we've put our ignorance of the number of states into this function!

We note that the Boltzmann factor, written in terms of velocities, is

$$P(E) \propto e^{-mv^2/2\tau} = e^{-m(v_x^2 + v_y^2 + v_z^2)/2\tau} .$$

This strongly suggests that each component of the velocity is normally distributed. We might expect this on other grounds as well. For example, if we pick a molecule and measure the x component of its velocity, we note that v_x is a random variable and its value is determined by the previous history of collisions the molecule has undergone. These are the kinds of conditions that give rise to Gaussian distributions although proving it in this case might be painful. We expect that $\langle v_x \rangle = 0$. Otherwise, our sample of gas would have a non-zero velocity in the x direction and we are considering a sample at rest (of course!). Also, the mean square value of v_x contributes to the energy and we would expect this to be

related to the temperature as the Boltzmann distribution suggests. Finally (among things that we expect), there is no reason to prefer one direction over another, so the probability distributions for v_x , v_y , and v_z should be the same. In fact, the probability of any velocity component should be the same which means that the joint probability should depend only on the magnitude and not on the direction of the velocity.

Putting all this together, we guess that the joint probability that the x , y , and z components of the velocity are in the ranges $v_x \rightarrow v_x + dv_x$, $v_y \rightarrow v_y + dv_y$, and $v_z \rightarrow v_z + dv_z$, is

$$\begin{aligned} p(v_x, v_y, v_z) dv_x dv_y dv_z &= \sqrt{\frac{1}{2\pi\tau/m}} e^{-mv_x^2/2\tau} dv_x \times \sqrt{\frac{1}{2\pi\tau/m}} e^{-mv_y^2/2\tau} dv_y \times \sqrt{\frac{1}{2\pi\tau/m}} e^{-mv_z^2/2\tau} dv_z, \\ &= \left(\frac{1}{2\pi\tau/m}\right)^{\frac{3}{2}} e^{-mv^2/2\tau} dv_x dv_y dv_z. \end{aligned}$$

This is called the Maxwell velocity distribution. Now let's change variables from Cartesian components to spherical polar components, v , θ , ϕ where

$$\begin{aligned} v_x &= v \sin \theta \cos \phi, \\ v_y &= v \sin \theta \sin \phi, \\ v_z &= v \cos \theta. \end{aligned}$$

This is just like changing coordinates from x , y , z to r , θ , ϕ . The result is

$$p(v, \theta, \phi) dv d\theta d\phi = \left(\frac{1}{2\pi\tau/m}\right)^{\frac{3}{2}} e^{-mv^2/2\tau} v^2 \sin \theta dv d\theta d\phi.$$

Just as we expected, the probability is independent of direction, so we can integrate over direction (θ and ϕ) to obtain the probability density for the velocity and any direction. The integral over directions gives 4π and we have,

$$p(v) dv = 4\pi \left(\frac{1}{2\pi\tau/m}\right)^{\frac{3}{2}} e^{-mv^2/2\tau} v^2 dv.$$

Let's change variables one more time, back to the energy $E = mv^2/2$,

$$p(E) dE = 2\pi \left(\frac{1}{\pi\tau}\right)^{\frac{3}{2}} e^{-E/\tau} E^{\frac{1}{2}} dE.$$

Comparing this with our earlier expression, we conclude that the number of states per unit energy, $n(E)$ must be proportional to \sqrt{E} .

To recap: knowing that the probability of a state with energy E is proportional to $\exp(-E/\tau)$, we argued that it was reasonable to expect the velocity components to be normally distributed. We used symmetry arguments to specify the form of the distribution and we used the probability distribution in energy to set the variance. We made some changes in variables to convert our guessed probability density in velocity to a probability density in energy from which we concluded that the number of kinetic energy states per unit energy is $\propto E^{1/2}$.

Aside—the Gamma Function

I kind of side-stepped the gamma function last week when we were using Stirling's approximation. But I think I should at least introduce the function as it's necessary for some of the integrals you might need to do. For example, to verify that $p(E)$ above is properly normalized, you will have to integrate over the energy and this is most naturally written as a gamma function. Define the gamma function by an integral:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt .$$

Using integration by parts,

$$\begin{aligned} \Gamma(z) &= \int_0^{\infty} t^{z-1} e^{-t} dt , \\ &= -t^{z-1} e^{-t} \Big|_0^{\infty} + (z-1) \int_0^{\infty} t^{z-2} e^{-t} dt \quad (z \geq 1) , \\ &= (z-1) \int_0^{\infty} t^{z-2} e^{-t} dt , \\ &= (z-1) \Gamma(z-1) . \end{aligned}$$

The gamma function satisfies a recursion relation. It is straightforward to show that $\Gamma(1) = 1$ and with the recursion relation, one has

$$n! = \Gamma(n+1) .$$

This relation can be used to extend the definition of factorial to non-integers!

The next most interesting arguments of the gamma function after the integers are half integers. Using the recursion relation, these can be found if one knows

$$\Gamma(1/2) = \sqrt{\pi} .$$

Can you show this?

For completeness, I should mention the incomplete gamma functions. (I couldn't resist, sorry!) These arise when the range of integration does not include the entire interval $[0, \infty)$.

The Partition Function

We've had a diversion and an aside, so maybe it's time to get back to thermal physics. Consider the function

$$Z(\tau) = \sum_s e^{-E_s/\tau},$$

where the sum extends over all states s . Among other things, this function normalizes the probabilities,

$$P(E_s) = \frac{e^{-E_s/\tau}}{Z(\tau)}; \quad \sum_s P(E_s) = 1.$$

Z is called the *partition function*. We've written it as a function of temperature, but it's also a function of the energies of all the states which might be functions of macroscopic parameters of the system.

The average energy of a member of the ensemble is

$$\langle E \rangle = \sum_s E_s P(E_s).$$

Consider the derivative of Z with respect to temperature (the energies of the states, E_s , do not depend on temperature),

$$\begin{aligned} \frac{\partial Z}{\partial \tau} &= \frac{\partial}{\partial \tau} \sum_s e^{-E_s/\tau}, \\ &= \sum_s e^{-E_s/\tau} \frac{\partial}{\partial \tau} \left(\frac{-E_s}{\tau} \right), \\ &= \sum_s e^{-E_s/\tau} \frac{E_s}{\tau^2}, \\ &= \frac{1}{\tau^2} \sum_s E_s e^{-E_s/\tau}, \\ &= \frac{Z}{\tau^2} \frac{1}{Z} \sum_s E_s e^{-E_s/\tau}, \\ &= \frac{Z}{\tau^2} \langle E \rangle, \end{aligned}$$

from which we deduce that

$$\langle E \rangle = \tau^2 \frac{\partial \log Z}{\partial \tau} .$$

Our discussion followed K&K except that we use E rather than ε and K&K appear to change the definition of U in midstream from the energy of heat bath to the average energy of a state of a single system.