

Chapter 2. Entropy and Temperature

1 The Fundamental Assumption

The *fundamental assumption* of statistical mechanics is that a closed system in equilibrium is equally likely to be in any one of the (quantum) states accessible to it. A *closed system* has no contact with any other system and so has fixed total energy, number of particles, volume and constant values for any external parameters which affect the system.

The meaning of *accessible* is a little more vague. First, any accessible state of the system must be consistent with the fixed parameters of the system: energy; number of particles *etc* (within the constraints of quantum mechanics). The problem with this definition is that sometimes there are states of the system satisfying the full specification, but which are only reached on timescales greatly exceeding those of interest (the duration of a measurement). Despite the ambiguity, it is usually not difficult to decide which states of the system we regard as really accessible.

In principle it is possible to precisely specify the (unique) quantum state of a system, but for macroscopic systems this is (almost) never a practical possibility. Usually the specification is limited to a small number of parameters.

2 Probabilities

If there are g accessible states to a closed physical system, then the probability of finding the system in any particular state is

$$P(s) = \frac{1}{g}.$$

Of course $P(s) = 0$ for any inaccessible state of the system. The P 's are normalized. Later will deal with situations where the $P(s)$ is not uniform over all states of the system — non-isolated systems.

As above, the mean value of any physical property X of the system is

$$\langle X(s) \rangle = \sum_s X(s)P(s),$$

where the sum is taken over the states of the system. For an isolated system this is $\langle X \rangle = \sum_s X(s)/g$.

At any instant a real physical system is in a “single” accessible state, so that the meaning of the average is difficult to apply to a single system. We introduce the artificial concept of an *ensemble* of systems in order that the average defined above can be given a well defined instantaneous meaning. An ensemble is a very large number of identical systems (identical to the extent of the specification of the system). The average is then regarded as an *ensemble average*, the instantaneous average value of X taken over the whole ensemble. For *ergodic* systems the time averaged value of any property X is equal to the ensemble average — we will assume that all systems we deal with are ergodic.

(Note that the definition of ensemble given in Kittel & Kroemer is not the standard definition. Their definition is even more artificial than the usual — particularly for systems involving small numbers of states.)

2.1 The most probable state

We now consider an isolated system that is made up of 2 subsystems in *thermal contact* with one another, ie able to exchange energy but nothing else. Symbolically we write $\mathcal{S} = \mathcal{S}_\infty + \mathcal{S}_\epsilon$. The corresponding total energy is fixed, ie $U = U_1 + U_2 = \text{constant}$. We want to know what determines the division of the energy between the 2 systems.

We again take the spin system as our example. Consider $\mathcal{S}_\infty$ to be made up of N_1 spins and $\mathcal{S}_\infty$ to have N_2 spins (N_1 and N_2 fixed). The spin excess of the combined system is $2s = 2s_1 + 2s_2$. The energy of the combined system in a magnetic field B is then

$$U(s) = U_1(s_1) + U_2(s_2) = -2mB(s_1 + s_2) = -2mBs.$$

Since the total energy is fixed, s is fixed.

We assume that there is some means of energy interchange between the 2 systems so that they can come into thermal equilibrium with one another.

The multiplicity function of the combined system is given simply by

$$g(N, s) = \sum_{s_1} g_1(N_1, s_1) g_2(N_2, s - s_1),$$

where we have used the conservation of s (note the range of summation of s_1). This assumes that the possible states of the 2 systems are independent of one another.

Kittel & Kroemer refer to the set of all states of the combined system for a given s_1 and s_2 as a *configuration*. Thus we are interested in determining the *most probable configuration* of the system. We will denote the most probable values for any quantity with a caret accent — hence $\hat{s}_1$, $\hat{s}_2$, etc

If we know the density of states as a function of U_1 , then we can calculate the average value of the energy in each of the systems. However, like our simple model system, for any system consisting of a large number of particles the density of the accessible states considered as a function of U_1 has a very sharp peak into which most of the states are concentrated. As a result there is almost no distinguishing the *most probable* value for the energy U_1 (the mode) from its mean value.

Thus the mean distribution of the energy between the 2 systems is very close to the most probable, and the latter is much easier to determine. In statistical mechanics we always make the assumption that the most probable value of any property of a closed system is indistinguishable from the mean. This approach is always valid when we are dealing with a closed system consisting of a large number of particles. The average (most probable) value of any property of a system determined this way is known as its *thermal equilibrium value*.

We now illustrate this for our model system.

We use the approximations for g_1 and g_2 derived above, so that

$$g_1(N_1, s_1) g_2(N_2, s - s_1) = g_1(0) g_2(0) \exp \left(-\frac{2s_1^2}{N_1} - \frac{2(s - s_1)^2}{N_2} \right)$$

(note the shorthand). First we find the most probable configuration. Since the log is a monotonically increasing function, we can maximize $\ln g_1 g_2$, ie maximize

Setting the derivative equal to 0 gives

$$\frac{\hat{s}_1}{N_1} = \frac{s - \hat{s}_1}{N_2} = \frac{\hat{s}_2}{N_2} = \frac{s}{N}.$$

(The 2nd derivative is ≤ 0 , so that this is a maximum.) The equilibrium state of the combined system occurs when the fractional spin excess is the same in either system.

We now want to consider the behaviour of the density of states around the maximum. Set $s_1 = \hat{s}_1 + \delta$ and $s_2 = \hat{s}_2 - \delta$, then

$$\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2} = \frac{(\hat{s}_1 + \delta)^2}{N_1} + \frac{(\hat{s}_2 - \delta)^2}{N_2} = \frac{\hat{s}_1^2 + 2\hat{s}_1\delta + \delta^2}{N_1} + \frac{\hat{s}_2^2 - 2\hat{s}_2\delta + \delta^2}{N_2} = \frac{s^2}{N} + \frac{\delta^2}{N_1} + \frac{\delta^2}{N_2},$$

so that we may write

$$g_1(N_1, s_1) g_2(N_2, s_2) = (g_1 g_2)_{\max} \exp \left(-\frac{2\delta^2}{N_1} - \frac{2\delta^2}{N_2} \right). \quad (1)$$

As long as N_1 and N_2 are both large, this is very sharply peaked.

How often do we expect to find s_1 a given distance from $\hat{s}_1$? The multiplicity function $g_1 g_2$ is Gaussian about the most probable state. We consider a general Gaussian for which the normalized distribution function is

$$\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

The probability that the magnitude of X exceeds x is then given by

$$\begin{aligned} P(|X| > x) &= 2 \int_x^\infty \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \int_0^\infty \exp\left(-\frac{x^2 + 2xz + z^2}{2\sigma^2}\right) dz \quad (y = x + z) \\ &= \sqrt{\frac{2}{\pi}} \frac{1}{\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \int_0^\infty \exp\left(-\frac{xz}{\sigma^2}\right) dz = \sqrt{\frac{2}{\pi}} \frac{\sigma}{x} \exp\left(-\frac{x^2}{2\sigma^2}\right). \end{aligned}$$

Identifying

$$\frac{1}{\sigma^2} = \frac{4}{N_1} + \frac{4}{N_2},$$

we can state that the probability that s_1 deviates from $\hat{s}_1$ by more than δ is less than

$$\frac{1}{\delta} \sqrt{\frac{N_1 N_2}{2\pi N}} \exp\left(-\frac{2\delta^2}{N_1} - \frac{2\delta^2}{N_2}\right).$$

Putting in some numbers, if $N_1 = N_2 = 10^{22}$ and $\delta = 10^{12}$, so that $\delta/N_1 = 10^{-10}$, then the exponential factor is $e^{-400} \simeq 10^{-173.7}$ and the other factors reduce this to $\simeq 10^{-175.3}$. This extremely small.

How small? Consider that the age of the Universe is about 10^{18} s, so that if we sampled the system 10^{10} times per second over the age of the Universe, the chances of one observation exceeding this range is still only $\simeq 10^{-147.3}$ — for a deviation of 1 part in 10^{10} !

In summary, we can safely assumed that we will never see s_1 for this system even as much as one part in 10^{10} away from its equilibrium value.

Of course, we might see much larger fluctuations in a very small system.

3 Thermal Equilibrium

We may generalize the argument that we have just used for the 2 spin systems in thermal contact. For any 2 (extensive) systems in thermal contact, the multiplicity function of the combined system is

$$g(N, U) = \sum_{U_1} g_1(N_1, U_1) g_2(N_2, U - U_1).$$

Again defining a configuration of the combined system to be all of the accessible states for a given U_1 (U , N_1 and N_2), the most probable configuration is that for which the product $g_1 g_2$ is the maximum, *ie* for which

$$\frac{1}{g_1} \left(\frac{\partial g_1}{\partial U_1} \right)_{N_1} = \frac{1}{g_2} \left(\frac{\partial g_2}{\partial U_2} \right)_{N_2}.$$

(Note the partial derivative notation.) This may be rewritten

$$\left(\frac{\partial \ln g_1}{\partial U_1} \right)_{N_1} = \left(\frac{\partial \ln g_2}{\partial U_2} \right)_{N_2},$$

so that defining the *entropy* as

$$\sigma(N, U) = \ln g(N, U),$$

the equilibrium condition becomes

$$\left(\frac{\partial\sigma_1}{\partial U_1}\right)_{N_1} = \left(\frac{\partial\sigma_2}{\partial U_2}\right)_{N_2}.$$

This is the general condition for thermal equilibrium.

In more complex physical systems there may be other constraints used to specify the system. The fixed quantity N used above symbolizes all of these possible constraints in the more general case.

4 Temperature

We have just proved an extremely general property of thermal systems, that for any two systems in thermal equilibrium with one another $(\partial\sigma/\partial U)_N$ must be the same in either system. Introducing another system, if each of 2 systems is in thermal equilibrium with a 3rd, then they are in thermal equilibrium with one another. This is just the zeroth law of thermodynamics — the one which tells us that temperature is a well defined quantity. It seems reasonable to assume that this quantity is related to the temperature. In fact, if T is the absolute temperature

$$\frac{1}{k_B T} = \left(\frac{\partial\sigma}{\partial U}\right)_N.$$

The constant of proportionality is the Boltzmann constant

$$k_B = 1.38066 \times 10^{-23} \text{ J K}^{-1} = 1.38066 \times 10^{-16} \text{ erg K}^{-1}.$$

We will use the *fundamental temperature* defined by

$$\frac{1}{\tau} = \left(\frac{\partial\sigma}{\partial U}\right)_N,$$

so that $\tau = k_B T$. (We will not discuss how this relationship is determined. However, we will be deriving results which depend on the Boltzmann constant and we will find our results agree with experiment when using this definition. If we were being rigorous we would not try to fix the relationship between τ and the physical temperature at this stage. All we have shown is that τ is a function of T and we would determine the relationship later when we start calculating how an ideal gas behaves, for example.)

You will commonly see $\beta = \frac{1}{\tau}$ used in other statistical mechanics texts.

The idea of temperature predates the understanding of the physical basis of temperature. If this were not the case we might already be measuring temperatures in Joules or ergs. However, it is far more difficult to determine the constant of proportionality k_B than it is to determine absolute temperatures. As a result temperature still has a place in physics.

Note that we may also write

$$\tau = \left(\frac{\partial U}{\partial\sigma}\right)_N,$$

but beware of the subtle difference in the meaning of this partial derivative and the other one. This form regards U as a function of σ and N .

5 Entropy

We have defined the entropy as the natural log of the number of states accessible to the system (for the given energy, N , etc). In classical thermodynamics the entropy is defined by

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_N,$$

so that in fact

$$S = k_B \sigma = k_B \ln g.$$

We will call S the conventional entropy.

In words, the more states that are accessible to a system, the greater is its entropy. In general the number of states accessible to a system will depend on all of the specified properties of the system, so that the entropy is a function of all of the specified properties of the system. These will always include the energy, and frequently the number of particles and the volume of the system.

We now have an absolute physical definition of the entropy for a physical system.

Suppose that we move a small amount of energy from $\mathcal{S}_\infty$ to $\mathcal{S}_\epsilon$, then the entropy change is

$$\Delta\sigma = \left(\frac{\partial\sigma_1}{\partial U_1}\right)_{N_1} (-\Delta U) + \left(\frac{\partial\sigma_2}{\partial U_2}\right)_{N_2} (\Delta U) = \left(-\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) \Delta U.$$

Thus if $1/\tau_2 > 1/\tau_1$ the entropy increases when energy flows from $\mathcal{S}_\infty$ to $\mathcal{S}_\epsilon$. If the temperatures are both positive, then the entropy increases when energy flows from the system of higher temperature to the system of lower temperature. (Note that temperatures can be negative in some types of system.)

5.1 Example: entropy increase on heat flow

Consider two 10 g pieces of copper, one initially at 350 K and the other at 290 K. Assuming that the heat capacity of copper is $0.389 \text{ J g}^{-1} \text{ K}^{-1}$, what is the entropy increase if the 2 blocks are brought into thermal contact and allowed to come into equilibrium with one another?

First, the total heat capacity C of each of the blocks is the same, so that the temperature increase of 1 block is equal to the temperature decrease of the other when heat is exchanged. This means that they will equilibrate at the mean temperature, 320 K.

By definition the heat capacity, and of the temperature,

$$dU = C dT = T dS,$$

which we may integrate to get

$$\Delta S = C \ln \left(\frac{T_f}{T_i} \right).$$

The total increase in the (conventional) entropy of the system is therefore

$$\Delta S = C \ln \left(\frac{T_{f1}}{T_{i1}} \right) + C \ln \left(\frac{T_{f2}}{T_{i2}} \right) = 3.89 \ln \left(\frac{320^2}{350 \times 290} \right) \simeq 0.0343 \text{ J K}^{-1}.$$

In fundamental units this makes the entropy increase

$$\Delta\sigma = \frac{\Delta S}{k_B} = 2.49 \times 10^{21},$$

ie the number of states accessible to the combined system increases by the factor $\exp(2.49 \times 10^{21})$.

5.2 The law of increase of entropy

The entropy always increases when two systems are brought into thermal contact with one another. As we have shown, the multiplicity of the combined system is

$$g(U) = \sum_{U_1} g_1(U_1) g_2(U - U_1).$$

Before the systems were brought into contact, the total number of accessible states was $g_1(U_{i1})g_2(U_{i2})$, which is now one term in the sum for g , all terms in the sum being positive, so that the entropy ($\sigma = \ln g$) must be greater in the combined system.

After the systems have come into thermal equilibrium with one another, if we again separate them the multiplicity will be $(g_1g_2)_{\max}$. By definition this also exceeds the initial multiplicity of the 2 separate systems, so that the entropy will have increased.

The startling thing is that the multiplicity is so enormous that the entropy of the combined system is indistinguishable from $\ln g_{\max}$, *ie* the entropy of the combined system is equal to the sum of the entropies of the two separate systems if they are at the same temperature.

We illustrate this with the spin system again. We have shown for this system that

$$(g_1g_2)_{\max} = g_1(0)g_2(0) \exp\left(\frac{-2s^2}{N}\right),$$

and that

$$g_1(N_1, \hat{s}_1 + \delta)g_2(N_2, \hat{s}_2 - \delta) = (g_1g_2)_{\max} \exp\left(-\frac{2\delta^2}{N_1} - \frac{2\delta^2}{N_2}\right).$$

We can sum (integrate) over s to get the total number of states accessible to the combined system,

$$g(N_1, N_2, s) = \sqrt{\frac{\pi N_1 N_2}{2N}} (g_1g_2)_{\max},$$

so that for the combined system

$$\sigma = \frac{1}{2} \ln\left(\frac{\pi N_1 N_2}{2N}\right) + \ln g_1(0) + \ln g_2(0) - \frac{2s^2}{N}.$$

Now

$$\ln g_i(0) \simeq \frac{1}{2} \ln\left(\frac{2}{\pi N_i}\right) + N_i \ln 2,$$

so that $\ln g_i \sim N$. The range of s is $-\frac{1}{2}N, \dots, \frac{1}{2}N$, so that the term in s can also be $\sim N$. Putting this together,

$$\sigma = \ln(g_1g_2)_{\max} + O(\ln N),$$

and the second term is utterly insignificant compared to the first.

In general, there are terms in the log of the multiplicity which are $O(N)$, while the number of configurations contributing significantly to the sum of all accessible states is $O(N^p)$ for some $p \sim 1$. Thus the total entropy is

$$\sigma = \ln(g_1g_2)_{\max} + p \ln N,$$

and the ratio of the second term to the first is

$$\sim \frac{\ln N}{N}.$$

Even though the sum over all states increases the total number of accessible states by a large factor, the total entropy is negligibly different from the sum of the entropies of the 2 systems.

The evolution of the combined system towards equilibrium is not instantaneous, and we can imagine separating the two systems at some intermediate time before equilibrium is reached. By this means we may define the entropy of the 2 systems as a function of time. The total entropy is the sum of the entropies of the 2 separate systems, and it is an increasing function of time as they evolve towards equilibrium.

In summary, the entropy is *extensive*.

Entropy is not a concept that is easy to grasp intuitively. We will calculate the entropy for various physical systems throughout the course.

Note the seeming conflict between time reversal invariance of the equations of motion and the trend to increasing entropy.

6 Thermodynamic laws

Zeroth law: if two systems are in thermal equilibrium with a third system, then they must be in thermal equilibrium with one another. We have discussed this as a simple consequence of $\tau_1 = \tau_2$ in equilibrium.

First law: this is just the conservation of energy. We have already postulated that “heat” is a form of energy — in fact this is the first mention of heat as the medium of energy exchange between systems in thermal contact.

Second law: the entropy of 2 systems in thermal equilibrium is greater than or equal to the entropy of the 2 systems before being brought into thermal contact. There are other formulations of this law, but they are all equivalent to this statement.

Third law: The entropy of a system approaches a constant as the temperature approaches 0 (not necessarily zero because of ground state degeneracy). This is a consequence of the statistical definition of the entropy. $g(N, U_0)$ is well defined and $g(N, U) \rightarrow g(N, U_0)$ as $U \rightarrow U_0$, the ground state energy. It is not entirely trivial that $\tau \rightarrow 0$, ie

$$\left(\frac{\partial \ln g}{\partial U} \right)_N \rightarrow \infty,$$

as $U \rightarrow U_0$, but this follows because the available volume of configuration space varies as a positive power of $U - U_0$ near to the ground state.

6.1 Some properties of the entropy

We have already shown that it is extensive.

In general the energy of an isolated system is not precisely defined (quantum limits apart). We can define the *density of states* as the number of accessible states per unit energy smoothed over a suitable energy range (to eliminate the effects of quantization). If the energy of an isolated system is not precisely defined, then

$$\sigma = \ln g = \ln[\mathcal{D}(U)\delta U] = \ln \mathcal{D}(U) + \ln \delta U.$$

We can show that the actual value of δU is completely inconsequential to the value of σ , since the second term is always completely overwhelmed by the first. The argument is much the same as the argument we used to show that the entropy is extensive — the term $\ln \mathcal{D}(U)$ contains terms $O(N)$. A reasonable range for the possible values of δU varies with the size of a system as some small power of N . For any value in this range the resulting values of σ are indistinguishable.

The entropy is defined in terms of the density of states [in classical statistical mechanics. We see that the results will usually be identical to those for quantum statistical mechanics.

6.2 Example Problems, Chapter 2

1. Given $g(U) = CU^{3N/2}$, a) show that $U = 3N\tau/2$. We apply the definition of the temperature

$$\frac{1}{\tau} = \left(\frac{\partial \ln g}{\partial U} \right)_N = \frac{3N}{2U},$$

from which the result follows immediately. b) show that

$$\left(\frac{\partial^2 \sigma}{\partial U^2} \right)_N < 0.$$

This is a straight application of the definitions. It is a necessary condition for stability (as we will find later). Note that this is the form of the multiplicity function for the ideal gas.

2. Find the equilibrium value of the fractional magnetization as a function of the temperature for our model spin system. We know that $U = -2mBs$, and that

$$\sigma \simeq \ln g(0) - \frac{2s^2}{N}.$$

First we must eliminate s in terms of U , giving

$$\sigma \simeq \ln g(0) - \frac{U^2}{2m^2B^2N},$$

and then we apply the definition of the temperature,

$$\frac{1}{\tau} = \left(\frac{\partial \sigma}{\partial U} \right)_N = -\frac{U}{m^2B^2N},$$

so that the mean energy is given as a function of the temperature by

$$\langle U \rangle = -\frac{Nm^2B^2}{\tau}.$$

Note that this is physically unreasonable when $\tau \rightarrow 0$ because of the approximations we have used. We now eliminate the energy again in favour of s to get

$$\langle 2s \rangle = \frac{NmB}{\tau},$$

and the fractional magnetization is

$$\frac{\langle M \rangle}{Nm} = \frac{2\langle s \rangle}{N} = \frac{emB}{\tau}.$$

Recalling our approximation that $s \ll N$, we can limit the validity of this result to $\tau \gg mB$.