STATISTICAL THERMAL PHYSICS: BASIC IDEAS

Andrew M. Steane
Exeter College and Department of Atomic and Laser Physics, University of Oxford.

March 5, 2025

Abstract

This note presents basic ideas in statistical thermal physics, especially the calculation of entropy. The aim is to be clear, especially about any approximations employed, and to avoid some widespread misconceptions. The starting point is some preliminary understanding of the idea of a quantum state. The text includes derivations of basic formulae for internal energy, entropy, temperature and free energy.

1 Background: lost among the microstates

Once upon a time a man woke up somewhere in a huge city which was unfamiliar to him. Not knowing where he was, he started wandering around. With no other plan to guide him, he just took random turns at the junctions. Occasionally he came across an unusual building or a park, but mostly the streets were quite like one another. The man did not know it, but this city was built according to a plan where a few areas had shops or public buildings, but they were small areas. Most of the city was taken up by a huge suburb full of residential housing.

It was not long before the wandering man found himself in the residential area, because that area took up so much of the city. And once in it, he was very unlikely to find his way out of it because it was so large in comparison to the other areas.

This parable of the wandering man is an illustration of the internal dynamics of any physical system with a large number of physical states but little internal structure. Left to its own devices (i.e. isolated, but not at zero temperature) the system will *wander among the microstates*: that is, its physical configuration will evolve randomly among all those internal states which have the same energy as the one the system started out in, and therefore, if those states are in groups of different sizes, then subsequently the system is most likely to be found in one of the states in the largest group. This idea will be made more precise in the rest of the discussion. The main thing to note at the outset is that we make the following claim (and c.f. Fig. 1):

> An isolated system at non-zero temperature wanders around all the mutually orthogonal quantum states at some given energy in such a way as to spend equal amounts of time, on average, in each of them.

This claim can be justified by study of the dynamics as described by Schrödinger's equation. For the present discussion we shall not present that justification at this stage, but simply make the claim and discover what follows from it. It turns out that a huge amount will follow: almost all of thermal physics! For, by making the claim we come to understand the *entropy* of a thermal system. And if we also understand the *energy* then these two together suffice to open the way to understanding all the thermal behaviour.

The above statement about spending equal times per state can be used to justify the following, closely related, claim:

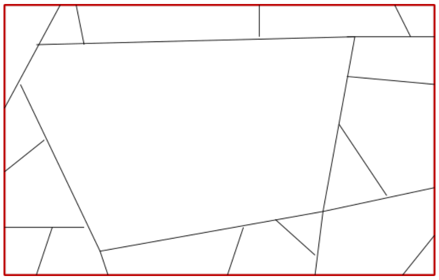


Figure 1: Lost among the microstates. The large rectangle represents a collection of quantum states (often called microstates) all at the same energy. Each tiny square (not shown) inside the rectangle represents a single microstate. The regions inside the rectangle represent collections of microstates based on some further property in addition to energy. For example, for a gas in a box, we could define the quantity $f = N_{\text{left}}/N$ as the fraction of particles located in the left half of the box, and then group the microstates based on values of f. The group with $0.05 \le f < 0.15$ will be small; the group with $0.15 \le f < 0.25$ will be larger; the group with $0.45 \le f \le 0.55$ will be the largest, etc. A system which is equally likely to be in any of the states of the whole rectangle will be most likely to be found in the largest such region.

> *For an isolated system in equilibrium, all accessible microstates are equally likely.*

This is sometimes said to be 'the fundamental assumption of statistical mechanics'.

Terminology around states.

1. A *microstate* is a single quantum state of the entire system. All different microstates are mutually orthogonal (in the quantum mechanical sense). It is convenient to adopt energy eigenstates as microstates.
2. A *macrostate* is a (typically large) set of microstates, all of which are consistent with macroscopic constraints applied to the system.

For example, an ensemble of 10 harmonic oscillators and total energy 20 units has 9900605 microstates, 1 macrostate and 21 states accessible to each system (these states are the low-lying ones, having energy 0 to 20 units). If one had further information about the system, one may choose to say that the single macrostate defined by total energy is made of two or more macrostates specified by their energy and some other property. In that case the 9900605 microstates of given energy will be divided among the macrostates.

1.1 Justifying the assumption: Liouville's theorem

This section can be omitted on first reading, but I think it is better if you do read it.

If we assert the fundamental assumption as just that—an assumption—then it does not need to be justified. It will prove its worth, after the fact, if it turns out to lead to telling insights, and predictions that are in agreement with empirical observation. And it does turn out that way. But the assumption (of equal times or, equivalently, equal probabilities per state) was not originally suggested by mere guess-work. It can be strongly motivated by *Liouville's theorem*. This is a mathematical result, in classical physics, concerning motion of sets of particles under conservative forces (i.e. forces that conserve energy).

We describe the motion by furnishing, for each particle, its position and momentum, thus for N particles moving in 3 dimensions there are $6N$ variables, and each state of motion of the complete system can be specified by a single point in an abstract $6N$-dimensional space called *phase space*. Let us gather all the $3N$ position variables into a $3N$-component vector called $\mathbf{r}$ and let us gather all the $3N$ momentum variables into a $3N$-component vector called $\mathbf{p}$.

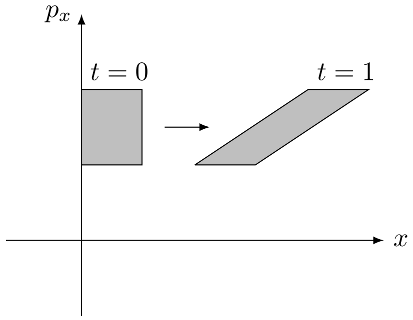


Figure 2: A simple example of Liouville's theorem: free motion in one dimension. The diagram represents states of motion a *single* particle (which is indicated by the dimensionality of the phase space). The rectangular shaded region is a set of initial conditions; the parallelogram shows the associated states of motion after some elapsed time. The two shapes have the same area.

For motion under conservative forces the dynamics can be described by furnishing a Hamiltonian H, and the evolution of the system is then governed by Hamilton's canonical equations of motion:[1]

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}}, \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}}. \tag{1}$$

Liouville's theorem concerns not individual points in phase space, but regions of phase space. We define a region (often called a cell) Γ in phase space as some continuous collection of points $\{\mathbf{r}, \mathbf{p}\}$. The phase-space volume of the cell is then

$$V = \iint_\Gamma \mathrm{d}^{3N}\mathbf{r}\, \mathrm{d}^{3N}\mathbf{p} \tag{2}$$

This is simply the integral over both position and momentum for the given cell. (Suitable SI units for such a 'volume' would be $\mathrm{J}^{3N}\,\mathrm{s}^{3N}$). The theorem can be derived from Hamilton's equations, and it asserts:

> **Liouville's theorem:** Under conservative dynamics, the phase space volume of a cell specified by a given collection of states stays constant as the states evolve.

The idea is that as the system evolves, a given set of points $\{\mathbf{r},\ \mathbf{p}\}$ will move around phase space, such that the cell specified by them will itself move and will typically change shape in complicated ways. But, according to the theorem, this shape-change preserves the volume of the cell. If it gets longer in one direction, it must get shorter in another. A simple example is shown in Fig. 2. For systems of many interacting particles what typically happens is that a cell which started out as some simple shape, such as a round blob, will soon be stretched out very thin and contorted into spirals and the like, all while preserving its volume.

The importance of Liouville's theorem for us is that it suggests that the dynamics of a system of many particles will not tend to send the system into any one region of phase space in preference to another. The particles will not, for example, tend to gather in one place or at one momentum. If the state was uncertain at the outset, say falling into some given phase-space region but it is not known where in that region, then later on it will be somewhere in a region of possibly a different shape but of the same size as the starting region.

The 'fundamental assumption' we make for statistical mechanics is only concerned with one type of initial condition: the one called *equilibrium*. This is the condition the system will be in after it has been left undisturbed for long enough. In such an equilibrium the particles are still moving (just as fast as ever on average) but not in any particular pattern. Their motion will not correspond to sound waves or vortices, for example. In consequence we may assert,

---

[1] The notation $(\partial H/\partial \mathbf{p})$ indicates that vector whose i'th component is $(\partial H/\partial p_i)$.

in the case of an isolated system, that the cell in phase space in which the system state is to be found includes the whole phase-space region at one particular energy, and the system state thereafter stays somewhere in that cell, but not in any particular part of it in preference to another. This does not prove, but motivates, the assumption of equal probabilities which we adopt for statistical mechanics.

The quantum treatment of the motion reproduces the classical one in the relevant limit, and adds some further details. The main thing to note about the quantum result is that, at any given energy, *uniform probability-density in phase space* corresponds to *uniform probability per mutually orthogonal quantum state* at the given energy.

2 Entropy

There are two basic statements about entropy in statistical mechanics. The first is the one due to Boltzmann:

$$S = k_{\rm B} \ln W. \tag{3}$$

The second is associated with the name of Gibbs:

$$S = -k_{\rm B} \sum_i p_i \ln p_i. \tag{4}$$

In Boltzmann's formula, W is the number of **microstates** accessible to an isolated system.

In Gibbs's formula, p_i is the probability that the system will be found to be in its i'th microstate in more general conditions, including, for example, when the system can exchange energy with other systems. On these two formulae hang all the law of statistical mechanics. There are no other formulae greater than these.

If we recall the parable of the wandering man (i.e. if we make the fundamental assumption), we shall realise that in the case of an isolated system one should adopt

$$p_i = 1/W$$

in the Gibbs entropy (by arguing that all accessible microstates are equally likely for an isolated system in internal equilibrium). Therefore one finds that the two formulae agree for that case:

$$S = -k_{\rm B} \sum_i p_i \ln 1/W = k_{\rm B} \sum_i p_i \ln W = k_{\rm B} \ln W. \tag{5}$$

It follows from this that one does not need both formulae. It is sufficient to use the Gibbs formula. It will reproduce Boltzmann's result exactly when the Boltzmann formula applies.[2]

More generally the Gibbs formula can be seen as a kind of average:

$$S = k_{\rm B} \sum_i p_i \ln W_i = \langle k_{\rm B} \ln W \rangle \tag{6}$$

where $W_i = 1/p_i$. However the probabilities p_i do not always have a natural physical interpretation as the inverse of some number of states.

The above is clear and unambiguous if we assume the isolated system to have a precisely-defined total energy. Such an assumption is unproblematic in classical physics, because, by conservation of energy, if the system is isolated its energy cannot change. However in quantum physics an isolated system can be in a quantum superposition of energy eigenstates. In that case the Boltzmann formula (3) does not apply. That is, the Boltzmann formula does apply to a

---

[2] In the literature there is discussion about entropy which sometimes takes the form"the Gibbs entropy is this, but the Boltzmann entropy is that". But, as we have just shown, the formulae agree exactly in the case where they both apply, so no such ambiguity can arise. (When the ambiguity arises it is because some other formula has been introduced and named after Gibbs or Boltzmann.)

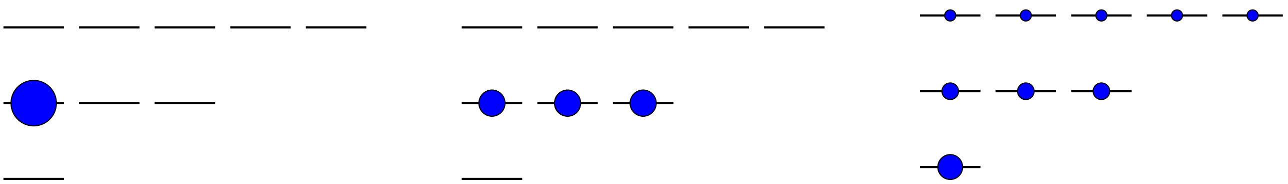


Figure 3: The diagrams show energy levels of a system with 9 quantum states at 3 energy levels (of degeneracies 1,3,5). The filled circles indicate the probability of being in one state or another. Left diagram: system in a single energy eigenstate, entropy = 0. Centre: system in energy equilibrium at energy 1 unit, entropy = $k_{\mathrm{B}} \ln 3 \simeq 1.099\, k_{\mathrm{B}}$. Right: system in thermal equilibrium with mean energy 1 unit, entropy = $2.01\, k_{\mathrm{B}}$.

quantum system with precisely-defined energy (one of its energy eigenvalues), but not otherwise. (We will argue later that the case of a superposition of states of differing energy can be handled by Gibbs's formula.)

Having introduced the entropy, it remains to comment on it. Can Gibbs's formula be derived from other considerations? The reasoning goes as follows. The study of information theory yields a quantity called Shannon entropy which is given by a formula with the same form as the Gibbs entropy (4). The Shannon entropy quantifies information content in a precise way in the context of sending sets of symbols through a communication channel. The associated information theory yields some results which are reminiscent of thermal physics. For example, information processing cannot increase, but can reduce, the Shannon entropy. This is not a proof, but it leads one to propose the Gibbs formula as a plausible candidate for the calculation of entropy in thermal physics. One then studies what follows from this choice, and it turns out that it does make sense. Therefore we accept the plausible candidate as our working definition.

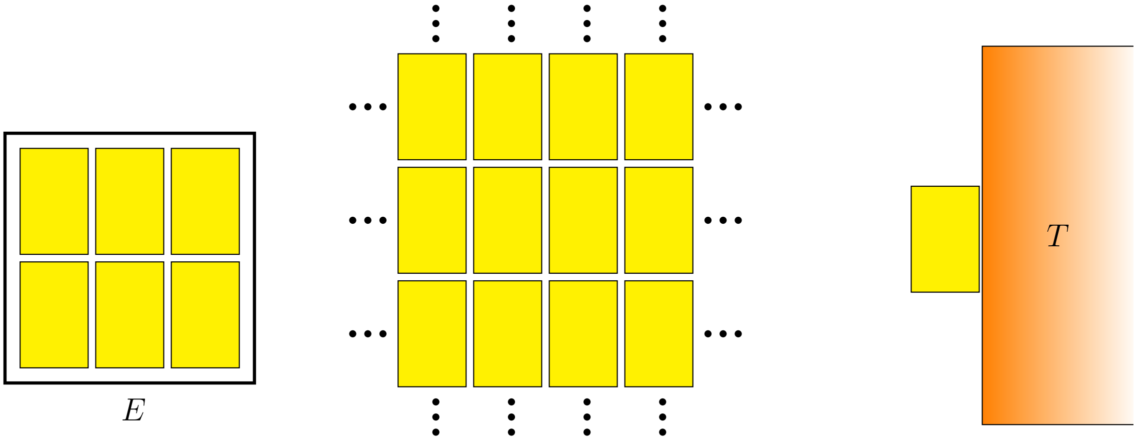
In Section 6 we will derive the Gibbs formula from the Boltzmann formula in the case of a system made of many similar parts, in the limit where the number of parts tends to infinity. This is strong evidence that the Gibbs formula is right. But nowadays we don't usually rely on a derivation like that: we simply assert (4) as an axiom.

3 Ensembles, microstates and macrostates

When a system has a precisely defined energy E there can be more than one accessible microstate because there can be more than one quantum state (energy eigenstate) of energy E. It will be useful to have a name for the case where an isolated system explores those degenerate (i.e. equal-energy) states equally. This is the condition where the Boltzmann entropy formula applies. We shall call it an *energy equilibrium*. Figure 3 gives an example. Another kind of equilibrium is *thermal equilibrium*. This is the condition where the system is in thermal contact with a reservoir at fixed temperature and equilibrium has been reached. We will find later that when a system is composed of a large number of smaller systems and has an amount of energy large compared to its energy level spacing, the energy equilibrium and the thermal equilibrium are very similar.

Another concept we shall need is that of a *Gibbs ensemble*. A Gibbs ensemble is a conceptual ensemble of copies of the system in question, such that each of the copies is described by the same macroscopic quantities (such as volume and mass) but the copies do not need to be all in the same quantum state (microstate). In a Gibbs ensemble we usually also make a further assumption, namely that the collection of systems has relaxed to an equilibrium state in which each system has the same average energy when averaged over time. For a large enough ensemble this equilibrium is the one called thermal equilibrium in thermodynamics. However in the statistical approach we do not need to assume the thermodynamic limit (which is the limit of very large numbers of accessible microstates), and for this reason the equilibrium reached by a finite ensemble may or may not be a thermal equilibrium. For example, an isolated harmonic oscillator prepared in a state of energy E (an energy eigenvalue) will simply stay in that state, which is not normally called thermal equilibrium. An isolated system with a well-defined energy will relax to an energy equilibrium, not a thermal equilibrium.

Among the various Gibbs ensembles which we can imagine, the following are of particular interest (see figure 4):



microcanonical canonical canonical

Figure 4: Ensembles. The microcanonical ensemble has some number N of systems in thermal contact with one another, the whole ensemble being isolated and in energy equilibrium. The canonical ensemble has an infinite number of systems in thermal contact with one another, in equilibrium. The thermal behaviour of any one member of the canonical ensemble is the same as that of a single system in thermal contact with a reservoir at fixed temperature.

1. **Canonical ensemble.** A collection of N copies which are able to exchange energy but not material (e.g. particles) with their neighbours, in the limit $N \to \infty$.
2. **Microcanonical ensemble.** A finite collection of N copies which are able to exchange energy but not material (e.g. particles) with their neighbours, the whole ensemble having a fixed total energy equal to one of the energy eigenvalues of the entire ensemble.

The terminology is slightly misleading because there is nothing especially small about the microcanonical ensemble. Its members can be large or small systems. Equally, the members of the canonical ensemble can be large or small. The main difference between these ensembles is that the microcanonical ensemble has a fixed total energy, whereas if one were to examine the energy of any finite sample from the canonical ensemble, the value can change from one moment in time to the next and from one sample to another (because energy can flow around the ensemble and typically this will result in local fluctuations).

The word ‘canonical’ has the meaning, roughly, of ‘standard’.

Another useful concept is that of a ‘heat bath’ or thermal reservoir. A thermal reservoir is a physical system with a large heat capacity and thermal conductivity, and otherwise it does not matter what kind of system it is. So we can, if we like, model the thermal reservoir as itself a canonical ensemble. From this it follows:

> *After relaxation to thermal equilibrium, a system in a canonical ensemble behaves precisely in the same way as a system (of the same type) which is in thermal contact with a thermal reservoir.*

For this reason we shall speak, in the following, of ‘a canonical ensemble at temperature T’. We have now established the correspondence:

canonical ensemble: each member is in thermal equilibrium (and so is the entire ensemble)

microcanonical ensemble: the entire ensemble is in energy equilibrium (but each member is not, unless there is only one member); in the limit $N \to \infty$ each member is in thermal equilibrium

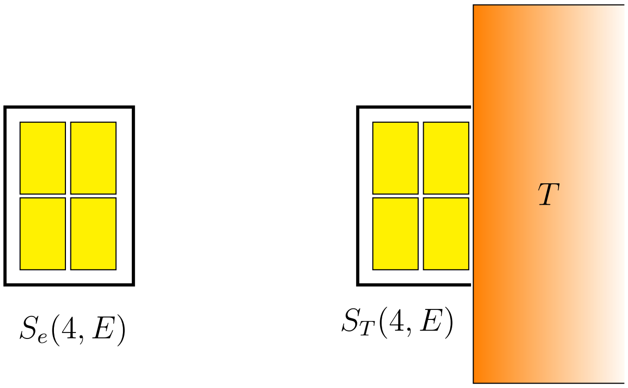


Figure 5: Definitions of $S_e(K, E)$ and $S_T(K, \bar{U})$. The diagrams show an example with $K = 4$. The left diagram shows a microcanonical ensemble with 4 members. The right diagram shows a *single* system taken from a canonical ensemble; we have adopted the strategy of indicating the rest of the (infinite) ensemble by drawing a heat bath. By writing $S_T(K, E)$ in the right hand diagram we indicate the case where $\bar{U}$ in the canonical ensemble is equal to E.

Let us now define, for some given system type (e.g. a harmonic oscillator, or a brick of one mole of particles, or a single particle, or an entire gas):

$S_e(K, E)$ = entropy of a microcanonical ensemble of K systems of the given type, the ensemble having total energy E.

$S_T(K, \bar{U})$ = entropy of a system C which is made up of a K systems of the given type, where now C is a member of a canonical ensemble at temperature T, and T is such that C has average internal energy $\bar{U}$.

Figure 5 may make these definitions clearer. $S_e(K, E)$ may be called the microcanonical entropy and $S_T(K, \bar{U})$ the canonical entropy, but one should not suppose this means they are different types or definitions of entropy; rather they give the entropy of a system in two different physical conditions.

It will emerge that there exists an expression for S_T applicable to any system:

$$S_T(1, \bar{U}) \;=\; -k_{\rm B} \sum_i p_i \ln p_i \qquad \text{with } p_i = e^{-\beta \epsilon_i}/Z \tag{7}$$

where $Z = \sum_i \exp(-\beta \epsilon_i)$ is called the **partition function**, the value of β is such that $\sum_i p_i \epsilon_i = \bar{U}$, and ϵ_i is the energy of quantum state i. This will be derived in Section 4. One may then show, by using the thermodynamic result $T = \partial E / \partial S$, that the temperature T is related to β by $\beta = 1/k_{\rm B}T$; this is shown in Section 9.

We will also derive the following relationships.

$$\begin{aligned}
S_T(K, \bar{U}) &= K S_T(1, \bar{U}/K) && (8)\\
S_T(NK, N\bar{U}) &= N S_T(K, \bar{U}) && (9)\\
S_e(1, E) &= 0 \qquad \text{if the energy level is non-degenerate} && (10)\\
S_e(K, E) &< S_T(K, E) && (11)\\
N S_e(K, E) &< S_e(NK, NE) < N S_T(K, E) && (12)\\
\lim_{K,N\to\infty} \frac{N S_e(K, E)}{S_e(NK, NE)} &= \lim_{N\to\infty} \frac{S_e(NK, NE)}{N S_T(K, E)} = 1 && (13)
\end{aligned}$$

In words, the above can be stated:

- (8), (9): The entropy S_T (which is the thermal equilibrium entropy) is extensive.

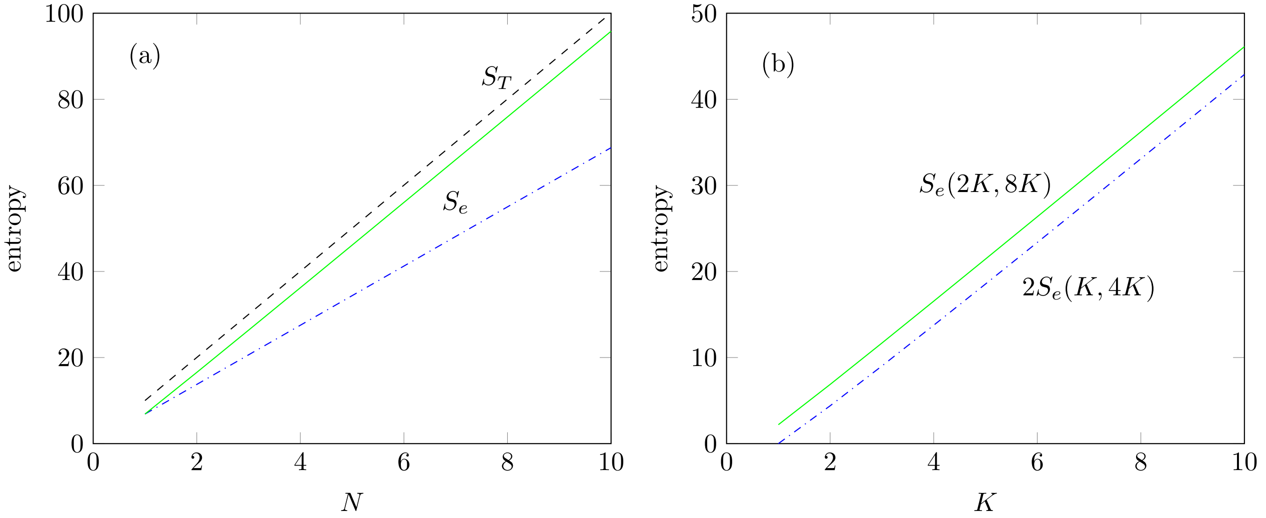


Figure 6: Illustrating (12)–(13). (a) $NS_T(4, 16)$ (black dashed), $S_e(4N, 16N)$ (green full) and $N\, S_e(4, 16)$ (blue dash-dot) for a set of harmonic oscillators. (b) In energy equilibrium the entropy of a pair of systems in thermal contact (and otherwise isolated) is greater than twice the entropy of either system alone. (This results differs from the case of thermal equilibrium, for which $S_T(2K, 2E) = 2S_T(K, E)$).

- (10): A single system with given energy, and having only a single quantum state of that energy, has zero entropy.
- (11): $S_e(K, E)$ is less than the entropy of a system at temperature T having the same size and mean energy.
- (12): $S_e(K, E)$ multiplied by N underestimates the entropy of the corresponding larger system.
- (12): $S_T(K, E)$ multiplied by N overestimates the entropy of the corresponding larger system if the latter has strictly fixed total energy (i.e. having no fluctuation whether for classical or quantum reasons).
- (13): All these entropies agree in the thermodynamic limit.

Figure 6 gives an example of these ideas.

The entropies $S_e(K, E)$ and $S_T(K, \overline{U})$ are both correct; they do not need to agree because they describe different physical conditions. The reason why the entropies do not all agree for finite systems is as follows. When we put N systems next to each other, if they remain entirely isolated from each other then we would expect the entropy simply to go up by a factor N. However if we allow them to be in thermal contact then they gain access to further states in which some have more energy, some less, than the average, at any given time. This explains why $S_e(NK, NE) > NS_e(K, E)$. Similarly, if the energy of the entire ensemble is restricted to a single value then the total entropy is less than that of a collection of the same size and mean energy in thermal equilibrium.

Another way to understand why $NS_T(K, E) > S_e(NK, NE)$ is to start with a set of N systems, each in thermal equilibrium separately, and notice that the internal energy fluctuations of one are uncorrelated with those of another. But if we now impose a constraint such that the total energy is fixed, then the energy fluctuations are correlated: if one system's energy goes down, another's (in the set of N) must go up. Such correlation reduces the entropy. This is related to the concept of *mutual information* in information science.

4 Maximum entropy: Boltzmann distribution

We will prove the following:

For a closed system with fixed average energy $\bar{U}$, the entropy (4) is maximised when the probabilities p_i are given by the **Boltzmann distribution:**

$$p_i \propto e^{-\beta\epsilon_i} \tag{14}$$

Proof. We wish to find a stationary value of $\sum_i p_i \ln p_i$ subject to the constraints

$$\sum_i p_i = 1; \qquad \sum_i p_i \epsilon_i = \bar{U}. \tag{15}$$

The method of Lagrange multipliers converts this problem to that of finding a stationary value of

$$\sum_i (p_i \ln p_i + \alpha p_i + \beta p_i \epsilon_i) \tag{16}$$

with respect to change in the variables p_i, where the Lagrange multipliers α and β have to be set so that the constraints are satisfied. By taking the partial derivative with respect to p_j, with $p_{i\neq j}$ held constant, we obtain

$$\ln p_j + 1 + \alpha + \beta\epsilon_j = 0 \tag{17}$$

hence

$$p_j = \exp(-1 - \alpha - \beta\epsilon_j) = Ae^{-\beta\epsilon_j} \tag{18}$$

where $A = \exp(-1-\alpha)$ is a normalization constant. One can check that this is indeed a maximum not a minimum nor saddle point for the entropy. QED

We can now derive (7). The derivation consists in asserting that *the thermal equilibrium state is the one which maximises the entropy for a given mean energy.* With this definition of thermal equilibrium, (14) implies (7) immediately. The physical reasoning has to be considered carefully in order to justify this step. The *mathematical* steps leading from (15) to (18) are incontrovertible, but to what physical situation does the derivation apply? By allowing the system to have access to many different energies ϵ_i, we imply that the system is not isolated, and yet we are treating a case where its mean energy is fixed. Fixed by what? How fixed? The answer to these questions is that the derivation applies to a system in contact with a heat bath (thermal reservoir) of fixed temperature. The physical claim is that such a system will have an energy which can fluctuate around some fixed value but which does not drift inexorably away from that value. We are also claiming that the statistical definition of entropy agrees with the thermodynamic entropy.

It will emerge, in a study of the dynamics of relaxation to equilibrium, that the statistical definition does reproduce the thermodynamic one in the right limit. A central element of the proof can be seen immediately. Suppose we have two systems in thermal equilibrium with a heat bath and each another, and weakly interacting. Here **weakly interacting** is a much-used concept in statistical physics, it means the systems can exchange energy but such interactions do not significantly change the energy levels of either. The above derivation can be applied to either system, so we shall find

$$p_i = A_1 e^{-\beta_1\epsilon_i}, \qquad p_j = A_2 e^{-\beta_2\epsilon_j}. \tag{19}$$

where i labels the states of system 1 and j labels the states of system 2. Equally, by applying the derivation to the composite system we must find

$$p_{i,j} = Ae^{-\beta(\epsilon_i+\epsilon_j)} \tag{20}$$

since a state (i,j) has energy $\epsilon_i + \epsilon_j$. But $p_{i,j} = p_i p_j$ since when the systems are in thermal equilibrium with the reservoir their states are uncorrelated with one another. Therefore

$$A_1 A_2 e^{-(\beta_1\epsilon_i+\beta_2\epsilon_j)} = Ae^{-\beta(\epsilon_i+\epsilon_j)}. \tag{21}$$

This is not just one equation: it is a set of equations for all the values of i and j. The only solution for all i and j (assuming there are more than 2 of them) is $\beta = \beta_1$ and $\beta = \beta_2$ so we must have $\beta_1 = \beta_2$. The important point is that there exists a solution, and therefore the assumption that the Boltzmann distribution corresponds to thermal equilibrium has not been ruled out by this argument. One may then enquire whether some other probability distribution could possibly also consistently describe the thermal equilibrium of two or more systems. The crucial feature is that a case where the probabilities multiply ($p_{i,j} = p_i p_j$) corresponds to a case where the energies add ($\epsilon_{i,j} = \epsilon_i + \epsilon_j$). It follows that the relationship between probability and energy must be exponential in thermal equilibrium. Hence, the condition of thermal equilibrium and the condition of maximum entropy are one and the same.

4.1 Extensivity of thermal equilibrium entropy

We will now prove (8), which expresses the extensivity of thermal equilibrium entropy. The proof is like the one around (21) just described. We suppose we have a collection of K copies of a system, in thermal equilibrium, and we add one more. The joint probability distribution will be given by (21) where now i labels the states of the whole collection of K copies, and j labels the states of the further system. The entropy of the new collection (of size $K+1$) is

$$\begin{aligned} S_T(K+1,\bar{U}) &= -k_{\rm B} \sum_{i,j} p_{i,j} \ln p_{i,j} \\ &= -k_{\rm B} \sum_{i,j} p_i p_j \ln p_i p_j \\ &= -k_{\rm B} \sum_i \sum_j p_i p_j \left(\ln(p_i) + \ln(p_j)\right) \end{aligned} \tag{22}$$

Now

$$\begin{aligned} \sum_i \sum_j p_i p_j \left(\ln(p_i) + \ln(p_j)\right) &= \left(\sum_i p_i \sum_j p_j \ln p_i\right) + \left(\sum_i p_i \sum_j p_j \ln p_j\right) \\ &= \sum_i p_i \ln(p_i) \left(\sum_j p_j\right) + \sum_j p_j \ln(p_j) \left(\sum_i p_i\right) \\ &= \sum_i p_i \ln p_i + \sum_j p_j \ln p_j \end{aligned} \tag{23}$$

Hence

$$S_T(K+1,\bar{U}) = S_T(K,\bar{U}) + S_T(1,\bar{U}) \tag{24}$$

and (8) follows by induction, and so does (9).

Note that the only assumption we required in the above proof was that the probabilities were uncorrelated. Therefore, more generally we have that the entropy of a system composed of two or more subsystems such that the probability distributions are uncorrelated will be equal to the sum of the entropies of the subsystems.

5 Proving the inequalities

Equation (10) asserts that an isolated system in a single energy eigenstate has zero entropy. This follows immediately from either of the Boltzmann or Gibbs formulae. In the Boltzmann formula one has $W = 1$. In the Gibbs formula one has just a single probability in the sum, equal to one.

Equation (11) asserts that the entropy in energy equilibrium is less than the entropy in thermal equilibrium. This follows from the fact that the thermal equilibrium maximises the entropy at the given mean energy $\bar{U}$ (proved in section

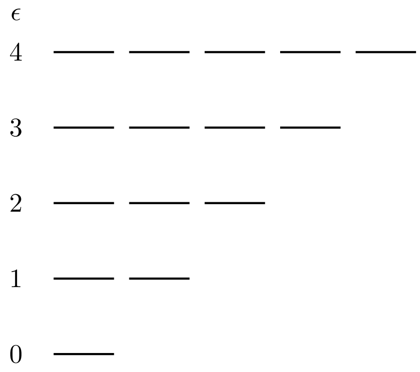


Figure 7: Energy level structure for an example treated in Eqn (27).

4). Therefore the entropy of any state at all (having the given mean energy), other than the thermal equilibrium state, will be less than S_T. It just remains to note that the energy equilibrium state is not the thermal equilibrium state, because the former has all microstate probabilities zero except those at the given energy, whereas the latter has an exponential distribution of probabilities.

By the same argument we can establish the second inequality of (12). For, by extensivity of thermal equilibrium entropy,

$$S_T(NK, NE) = NS_T(K, E) \tag{25}$$

and by maximum entropy for thermal equilibrium,

$$S_T(NK, NE) > S_e(NK, NE). \tag{26}$$

Hence $S_e(NK, NE) < NS_T(K, E)$, which is the second inequality in (12)

We already proved the first inequality of (12) by arguing that a set of N copies of any system has more microstates of given total energy than just the ones in which the energy is equally distributed among the copies. Therefore, whether by the Boltzmann or the Gibbs formula, its entropy is larger.

It is instructive to examine in detail an example of this, as follows. We consider a pair of systems, each of whose energy level structure is as shown in Fig. 7. (Each system could be a 2-dimensional harmonic oscillator, for example). We suppose the total energy available is $E_{\rm tot} = 4$. First we note how many states each system has at energies up to 4. Then we find how many ways there are for the composite system to have total energy 4. This is

$$W = \sum_{E_1=0}^{4} W(E_1) W(4 - E_1) = 1 \times 5 + 2 \times 4 + 3 \times 3 + 4 \times 2 + 5 \times 1 = 35 \tag{27}$$

where $W(E_1)$ is the degeneracy of the energy level at energy E_1 for a single system. In this example $W = 35$ so

$$S = S_e(2, 4) = k_{\rm B} \ln 35. \tag{28}$$

Meanwhile the number of ways for a single system here to have energy 2 is $W = 3$, so

$$S_e(1, 2) = k_{\rm B} \ln 3 \tag{29}$$

Hence

$$2S_e(1, 2) = k_{\rm B} 2 \ln 3 = k_{\rm B} \ln 9 < S_e(2, 4). \tag{30}$$

The fact that the entropy in energy equilibrium, $S_e(N, E)$, is extensive in the limit of large systems is illustrated in Fig. 6, and proved in the next section. The proof turns on the fact that most of the microstates in a microcanonical ensemble of large systems correspond to the case where the energy is equally, or almost equally, divided among the systems.

6 Proving the equality in the limit

In this section we will show that, for a microcanonical ensemble of N systems,

$$S_{\rm tot} = k_{\rm B} \ln W = -k_{\rm B} \sum_i p_i \ln p_i \simeq -N k_{\rm B} \sum_k p_k \ln p_k \tag{31}$$

where i labels the microstates of the entire ensemble and k labels the microstates of one system and the approximation becomes accurate in the limit of large N. In that limit the rest of the ensemble acts as a thermal reservoir for any one of its members, and therefore the energy equilibrium of the ensemble as a whole realises a thermal equilibrium for each member. It follows that, in the limit of large N, the quantity on the right of (31) is equal to S_T, the thermal equilibrium entropy.

The result can be expressed in words thus:

> For an ensemble of N like systems, weakly interacting and in energy equilibrium, in the limit $N \to \infty$ the total entropy (as given by either of (3) and (4)) is equal to N times the entropy of any one system as given by (4).

In (31) the probabilities p_k for any one system are obtained by averaging over the behaviour when the whole ensemble is in energy equilibrium (this will be clarified in the following). The relationship (31) is important because it amounts, almost, to a proof that the Gibbs entropy (4) is the correct generalization of the Boltzmann formula (3). It is not quite a proof, but a strong motivating argument. After receiving the motivation the modern approach is to adopt (4) as a definition and then show afterwards that the behaviour is consistent with the name 'entropy' for this quantity.

The main idea we shall invoke, in order to prove (31), is to regard the entire collection of systems in the microcanonical ensemble as one big isolated system. We can then use Boltzmann's formula for the entropy $S_{\rm tot}$ of the entire ensemble, and having done this, deduce how the entropy of each member of the ensemble is related to its own internal behaviour.

It will emerge that it will not be necessary for N to be very large in order to get quite good accuracy. The precision will be high even for a finite system because the main approximation is to invoke Stirling's approximation:[3]

$$\ln n! \simeq n \ln n - n + O(\ln n) \tag{32}$$

for any integer n. The relative precision of this approximation is approximately

$$\frac{\ln n}{n(\ln(n) - 1)} \simeq \frac{1}{n} \tag{33}$$

which you can see illustrated by the following table:

n	y $\ln n!$	z $n \ln n - n$	$(y - z)/y$
100	363.739	360.517	0.00886
10^6	1.2815518×10^7	1.2815510×10^7	6×10^{-7}
10^{10}	2.2×10^{11}	2.2×10^{11}	6×10^{-11}
10^{20}	4.5×10^{21}	4.5×10^{21}	6×10^{-21}

(34)

Already at $n = 10^6$ the precision is greater than that of most measurements in science. In practice we will often be dealing with truly vast values of n, of order $\exp(10^{10})$ or more. In that case the precision of Stirling's approximation will be greater than that of all measurements in science.

---

[3] A more full statement of Stirling's approximation is $\ln n! \simeq n \ln n - n + (1/2) \ln(2\pi n)$.

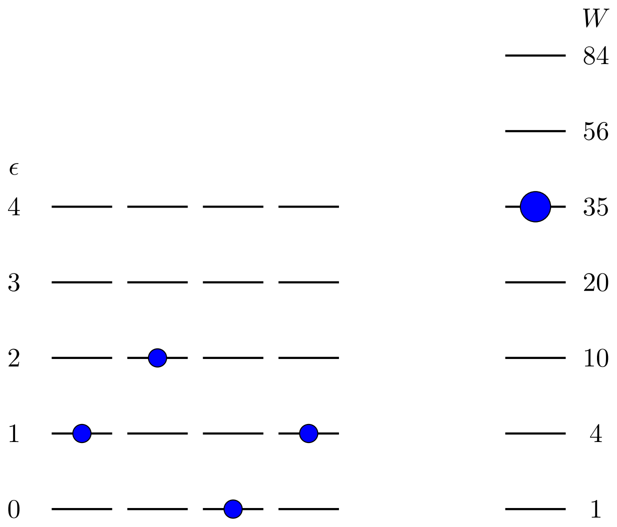


Figure 8: An example case for the calculation of entropy in the microcanonical ensemble (energy equilibrium). We have 4 systems whose energy levels are shown on the left, with an example microstate illustrated by the blue dots, for the case of total energy $E = 4\epsilon$. On the right are shown the energy levels of the entire ensemble, with, indicated, the number of microstates W at that energy (in physics more generally the letter g is often adopted for this quantity). The blue dot indicates the case considered in the example treated in the text, for which there are 35 microstates.

Now we shall treat a microcanonical ensemble and adopt the following notation:

symbol	meaning	example value
N	number of systems	4
E	total energy of ensemble	4
W	number of available microstates for entire ensemble	35
k	quantum state of one system	1–5
i	microstate of entire ensemble	1–35
r	row of occupation number table	1–5
W_r	number of ensemble microstates for r'th group	4, 12, 6, 12, 1

The example values in the table are for a simple case illustrated in Figure 8. W is the number of microstates consistent with some given macrostate, such as a state of given energy, volume and other such properties. The microstates give the details which are left unspecified by any given set of macroscopic measurements. We shall be studying these microstates by noting that they can be gathered into groups (to be shown), with W_r microstates in the r'th group. Therefore

$$W = \sum_r W_r \tag{35}$$

We treat a microcanonical ensemble of N systems of some given type. The individual systems could be single molecules, for example, or they might be themselves quite large, such as a brick made from a mole of molecules. We suppose that there are various states k which any given system might be in, so one way to specify a complete microstate of the whole ensemble would be to write a list of N values of k, specifying which state each system is in. This is not the most convenient way however. A better way is to write down a list saying how many of the systems are in each of the available system states at some given time. Such a list is called a set of **occupation numbers**, denoted thus:

$$\{n_k\}.$$

A set such as $\{n_k\} = \{5, 3, 1, 0, 0, 0, \ldots\}$ means 5 of the systems are in their ground state, 3 are in their first excited state, 1 in the second excited state, and none in the other states. For a fixed number of systems N, we must have

$$\sum_k n_k = N \tag{36}$$

and for a fixed total energy E we must have

$$\sum_k n_k \epsilon_k = E \tag{37}$$

where ϵ_k is the energy of state k of any one system. A complete list of possible sets of occupation numbers is as follows for the example case:

0	1ϵ	2ϵ	3ϵ	4ϵ	5ϵ	6ϵ	W_r
3	0	0	0	1	0	0	4
2	1	0	1	0	0	0	12
2	0	2	0	0	0	0	6
1	2	1	0	0	0	0	12
0	4	0	0	0	0	0	1
							35

The first row in this table shows the energies of the various system states (in this example they are equispaced because we choose to treat a harmonic oscillator). Each row after that gives a possible set of occupation numbers.

There can be many microstates of the ensemble corresponding to any given set $\{n_k\}$. Since the systems are all alike we can make the following assertion:

> The macrostate of the ensemble will be the same for all microstates corresponding to a given set of occupation numbers $\{n_k\}$.

The idea is that energy may move between the systems, so that some gain and some lose energy, but if the occupation numbers stay the same then the total energy and total size of the ensemble has not changed and therefore its **macro**state has not changed, even though its microstate has.

There is no general formula for how many sets of occupation numbers are consistent with any given macrostate of the entire ensemble: it will depend on the type of system. However, it is easy to calculate how many microstates of the ensemble are consistent with a given set of occupation numbers. It is simply the number of ways one can rearrange the states of excitation of the systems so as to get the same occupation numbers. For the case where the systems are distinguishable from one another, for example because they are located at different places, this is

$$W_{\{n\}} = \frac{N!}{\Pi_k n_k!} \tag{38}$$

This is the number of ways to choose n_1 items out of N and n_2 items out of $N - n_1$ and n_3 items out of $N - n_1 - n_2$ and so on.[4]

---

[4] This can be proved by induction quite easily. Here is the general idea:

$$\text{ways to choose } = \frac{N!}{n_1!(N-n_1)!} \times \frac{(N-n_1)!}{n_2!(N-n_1-n_2)!} \times \cdots$$

The following result will be important:

$$
\begin{aligned}
\ln W_{\{n\}} &= \ln N! - \sum_k \ln n_k! && (39)\\
&\simeq N\ln N - N - \sum_k (n_k \ln n_k - n_k) && (40)\\
&= -\sum_k n_k(\ln n_k - \ln N) \qquad \text{using (36)} &&\\
&= -N\sum_k \frac{n_k}{N}\ln\frac{n_k}{N} &&\\
&= -N\sum_k p_k \ln p_k && (41)
\end{aligned}
$$

where we used Stirling's approximation in the first step, and we introduced the probabilities

$$p_k = \frac{n_k}{N}. \tag{42}$$

in the last step. By comparing (41) with the Gibbs definition of entropy (4) we begin to see how the latter can be emerge in the present calculation. To make the connection we now need to argue carefully.

We can write the entropy of the entire ensemble like this:

$$S_{\rm tot} \;=\; k_{\rm B}\ln W = k_{\rm B}\ln\left[\sum_r W_r\right] \tag{43}$$

where the sum is over all possible sets of occupation numbers (i.e. those which are consistent with the macroscopic constraints) and, by using equation (38),

$$W_r = \frac{N!}{\Pi_k n_{k,r}!}, \tag{44}$$

in which $n_{k,r}$ is the occupation of system state k in the r'th set of occupation numbers.

We now make the following observation:

> *If the macrostate of the ensemble were described by just a single set of occupation numbers then the entropy of the ensemble would be equal to N times the Gibbs entropy of each system (up to the precision of Stirling's approximation), with p_k given by (42).*

This follows immediately by replacing the sum in (43) by a single term and then using (41) and (42). Note, this is just an observation, not a proof. But it gives us an idea of how to proceed.

What we need to do now is consider the fact that the macroscopic constraints are satisfied by more than one set $\{n_k\}$ of occupation numbers. In this case we need to reconsider what the probabilities p_k refer to. For clarity, let's refer to the set of microstates consistent with a given set of occupation numbers as a *group*. Some groups have very large W_r, making them macroscopic, others have small W_r, making them microscopic. The question is, what do the probabilities p_k used in the Gibbs entropy refer to? They refer to the probabilities that the system in question is in each of its internal (and mutually orthogonal) quantum states. In the present case this is the probability for a system to be in its state k when the whole ensemble can explore all available microstates equally. This means we must replace (42) by

$$p_k = \frac{\langle n_k\rangle}{N} = \frac{1}{N}\sum_r \frac{W_r}{W} n_{k,r} \tag{45}$$

where the second version tells you precisely which average we have in mind when we write $\langle n_k\rangle$. It is an average over the sets of occupation numbers, weighted by the number of microstates in each group. Our example case gives

$$
\begin{aligned}
\langle n_k\rangle &= \frac{60}{35}, \frac{40}{35}, \frac{24}{35}, \frac{12}{35}, \frac{4}{35} && (46)\\
&\simeq 1.71,\ 1.14,\ 0.69,\ 0.34,\ 0.11 && (47)
\end{aligned}
$$

Now consider the total entropy, as given by (43). It has the form of a logarithm of a sum. Notice that for large N the terms in the sum include some very large numbers, of order

$$W_r \sim O\left(\nu e^N\right)$$

(by using (41)), where ν is the number of subsystem states having non-negligible probability.[5] This W_r is truly enormous when N is equal to Avogadro's number, for example. Observe next that *the logarithm of a sum of very large numbers can be accurately estimated by the logarithm of the largest term in the sum:*

$$\ln \sum_i X_i \simeq \ln X^* \qquad \text{where } X^* = \max(X_i) \tag{48}$$

Proof:

$$X^* \le \sum_{i=1}^{m} X_i \le m X^* \tag{49}$$

where m is the number of terms in the sum. Hence

$$\ln X^* \le \ln \sum_i X_i \le \ln m X^* \tag{50}$$

which gives

$$\ln X^* \le \ln \sum_i X_i \le \ln X^* + \ln m. \tag{51}$$

Therefore the fractional error involved in using $\ln(X^*)$ as an estimate of the log of the sum is approximately $\ln(m)/2\ln(X^*)$. We are interested in cases with $X_i \sim \nu \exp(N)$ and $m \simeq \nu N$ so the fractional error is approximately

$$\frac{\ln(\nu N)}{\ln(\nu e^N)} \simeq \frac{\ln \nu + \ln N}{N} \ll 1 \qquad \text{when } N \gg 1, \ln \nu. \tag{52}$$

With this in mind, we now argue that Eqn (43) can be replaced by

$$S_{\rm tot} \quad \simeq \quad k_{\rm B} \ln W^* \tag{53}$$

where W^* is the value of W_r for the largest group (i.e. the set of occupation numbers with the largest microstate count) and the approximation is very good for large N. It follows that if the probabilities p_i are calculated for just that group, then (31) will follow.

This completes the derivation of (31). One should also note, as already remarked, that this derivation provides a strong hint that the Gibbs formula (4) is indeed the correct one to describe entropy in general.

For clarity we will now display the entire derivation again, as a single sequence of steps:

$$\begin{aligned}
S_{\rm tot} &= k_{\rm B} \ln W = k_{\rm B} \ln \left[\sum_r W_r\right] \\
&\simeq k_{\rm B} \ln W^* \\
&= k_{\rm B} \left(\ln N! - \sum_k \ln n_k^*!\right) \\
&\simeq -N k_{\rm B} \sum_k p_k \ln p_k
\end{aligned} \tag{54}$$

---

[5] A rough estimate for this ν in conditions of thermal equilibrium is $\nu \sim k_{\rm B} T/\Delta\epsilon$ where $\Delta\epsilon$ is a typical spacing between subsystem energy levels.

where the first approximate step was explained in equations (51) and (52) and the second approximate step invoked Stirling's approximation for the occupation numbers.

There remain some subtleties which are only important for modest or small values of N. This concerns the slight difference between the probabilities given by (42) (using the most likely occupation numbers) and the average probabilities given by (45) (obtained from the average occupation numbers). If we use the latter we obtain a more accurate statement of the entropy per system, and (54) is still found as long as correlations between systems are negligible. This shows that (54) is more accurate than is implied by (52) if we use the average occupation numbers to give the probabilities, as long as the correlations are small enough.[6]

It follows that the main requirement, in order for the above discussion to be accurate, is that the Stirling approximation be accurate for the most likely sets of occupation numbers. At very low temperatures, or for small isolated systems, this condition may not be met. In that case one must reconsider what exact case we wish to study. For example, a small system in thermal equilibrium with a heat bath can differ substantially from an isolated small system. In modern physics the standard practice is to assert that Gibbs' formula gives the entropy no matter the conditions, but it may require careful study to deduce what are the probabilities p_i in any given case.

7 Thermodynamic entropy

So far we defined a quantity which was given the name 'entropy'. It remains to justify that name, by connecting the Gibbs entropy with the thermodynamic definition of entropy. Our task is to show that the the quantity defined by (4) has the properties associated with thermodynamic entropy in the thermodynamic limit (i.e. for systems with a very large number of accessible states.) These properties are:

1. Entropy is an extensive property in the thermodynamic limit.
2. The internal dynamics of an isolated system can result in increase of entropy, or constant entropy, but not a decrease in the system's entropy.

The first property was treated in section 4.1. The second property requires a more involved study which is not the subject of the present note. The essential idea is that internal dynamics of any system are such that as time goes on, a larger and larger set of microstates can be explored by the system, and the system state is much more likely to be found in a group containing a larger and larger number of microstates as time goes on.

8 Worked example: a set of harmonic oscillators

A particle of mass m undergoing simple harmonic motion at angular frequency ω has the hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2. \tag{55}$$

The quantum theory of this system leads to a set of energy eigenstates whose eigenvalues are $(n + 1/2)\hbar\omega$ where n is a non-negative integer (thus having values $n = 0, 1, 2, \ldots$), with one state at each energy.

In this section we will discuss the energies in units of $\hbar\omega$, and we will choose the zero of potential energy so that the ground state has zero energy. In these units the energies E_1 available to a single harmonic oscillator are simply the non-negative integers: $E_1 = 0, 1, 2, \ldots$.

---

[6] It turns out that the average occupation numbers match the thermal equilibrium distribution very well even for quite small values of N: see Fig. 10(b).

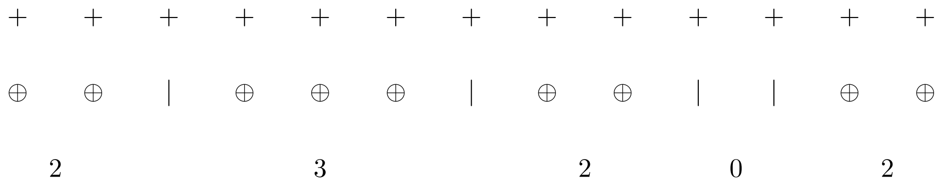


Figure 9: Argument to find the number of microstates of an ensemble of oscillators of total energy E. First we imagine a line of $E+N-1$ locations (top diagram). Next we place a circle on E of these locations and a vertical line on the remaining $N-1$ locations (middle diagram). The number of ways to do this, assuming the circles are all alike and the lines are all alike, is equal to the number of ways to choose $N-1$ things out of $E+N-1$ without respect to ordering. This is given by the binomial coefficient as shown in Eqn (56). Next we interpret the lines as partitions, dividing up the E circles into N groups. Each group corresponds to one oscillator; it receives an energy equal to the number of circles in the group. Thus we derive (56). The argument turns on persuading oneself that this does correctly account for all the ways of dividing up the energy, and without over-counting. Trying out examples with small values of E and N may help to build familiarity with the reasoning.

Because of the simple spectrum of energy levels (an equispaced ladder) it is easy to find how many states are consistent with a given total energy E for an ensemble of N oscillators. It is equal to the number of ways E items can be shared out about N different boxes, with the items all alike. The answer is

$$W(N,E) = \binom{E+N-1}{N-1} = \frac{(E+N-1)!}{(N-1)!E!} \tag{56}$$

See Fig. 9 for an argument to derive this.

Equation (56) solves the problem of calculating entropy for all microcanonical ensembles of harmonic oscillators! The answer is

$$S_e(N,E) = k_\mathrm{B} \ln W(N,E) = k_\mathrm{B} \ln \binom{E+N-1}{N-1}. \tag{57}$$

This was used to plot the examples shown in Fig. 6.

To find the value of S_T (the entropy for the canonical ensemble, i.e. the thermal equilibrium value) we employ the Gibbs' formula (4) with the probabilities given by the Boltzmann distribution (14). We have

$$p_i = \exp(-\beta\epsilon_i)/Z \tag{58}$$

with

$$Z = \sum_{n=0}^{\infty} e^{-\beta\epsilon_i} = \sum_{n=0}^{\infty} e^{-\beta n\hbar\omega} = \frac{1}{1-e^{-\beta\hbar\omega}} \tag{59}$$

where $n\hbar\omega$ is the energy of the state with quantum number n (when the ground state has energy zero) and the final expression comes from summing the geometric series. The quantity Z is called the **partition function** and it is a central concept in statistical physics. Once one has an expression for Z in terms of suitable macroscopic variables such as β and ω, one can derive all the thermodynamic behaviour! One finds, for example, that the entropy is given by

$$S_T = k_\mathrm{B}\left(\ln Z - \beta\frac{\partial \ln Z}{\partial\beta}\right) \tag{60}$$

(the derivation is provided in the next section). Applied to a single harmonic oscillator, this expression gives

$$S_T = k_\mathrm{B}\left[\frac{b}{e^b-1} - \ln\left(1-e^{-b}\right)\right] \tag{61}$$

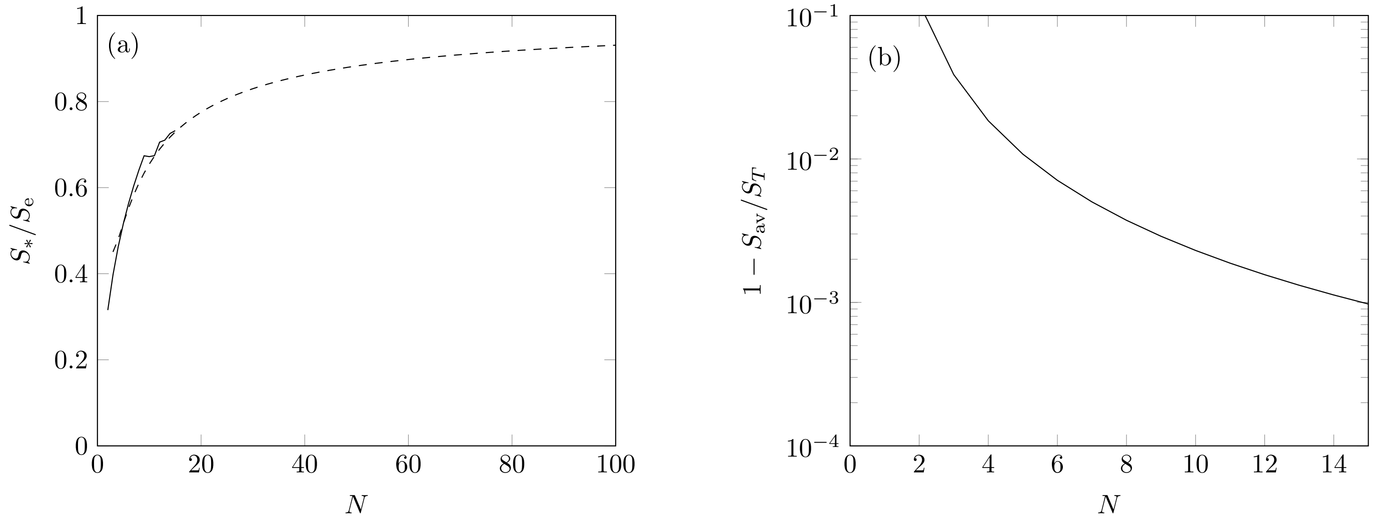


Figure 10: (a) shows the degree to which S_* (66) approximates the microcanonical entropy for small N. The full curve is an exact numerical calculation (for harmonic oscillators); the dashed curve is $1 - 3\ln(N)/2N$. (b) Shows that the entropy (67) calculated using average occupation numbers in a microcanonical ensemble approaches closely to the canonical entropy even for modest values of N.

where $b = \beta\hbar\omega$.

Another important result is the formula for the mean energy:

$$\bar{U} = \sum_i p_i \epsilon_i = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} \tag{62}$$

which will be derived in the next section. Applied to the harmonic oscillator this gives

$$\frac{\bar{U}}{\hbar\omega} = \frac{1}{e^b - 1}. \tag{63}$$

This can be used to find the value of b for a given mean energy, and thus to find $S_T(1, \bar{U})$. One obtains (exercise 3)

$$S_T(1, E) = k_\mathrm{B}\left((E+1)\ln(E+1) - E\ln E\right). \tag{64}$$

where $E = \bar{U}/\hbar\omega$. Hence, by using the extensivity of thermal equilibrium entropy (Eqn (8), proved in Section 4.1),

$$S_T(N, E) = N k_\mathrm{B}\left((E+1)\ln(E+1) - E\ln E\right). \tag{65}$$

This was used to plot the dashed line in Fig. 6.

Fig 6 illustrates how large N needs to be in order that the microcanonical entropy and thermal equilibrium entropy shall approximately agree. One finds, for example, that for the case $E = 10N$ (i.e. 10 units of energy per oscillator), S_e reaches 99% of S_T at $N = 175$, and 99.9% at $N \simeq 2200$. When $E = 40N$, S_e reaches 99% of S_T at $N = 152$.

We will now explore the question, how accurate is the approximation leading to Eqn (54)? This question is largely answered by (52) and by Fig. 10 which shows, for a system of harmonic oscillators, the degree to which the microcanonical and canonical entropies are matched by two approximate entropies defined as:

$$\begin{aligned} S_* &= k_\mathrm{B}\ln W^* && (66) \\ S_\mathrm{av} &= -N k_\mathrm{B}\sum_k p_k \ln p_k \qquad \text{with } p_k = \langle n_k\rangle/N && (67) \end{aligned}$$

where both refer to a microcanonical ensemble. The first of these is the approximation adopted in (53), and Fig. 10(b) illustrates that the second approaches closely to the canonical entropy, even for modest values of N.

9 Some basic formulae

In this section we consider only thermal equilibrium, so that the Boltzmann distribution applies:

$$p_i = \frac{1}{Z} e^{-\beta \epsilon_i} \tag{68}$$

where i labels individual quantum states of the system in question, and Z is the partition function defined by

Partition function

$$Z = \sum_i e^{-\beta \epsilon_i}. \tag{69}$$

This can also be written $Z = \mathrm{Tr}(e^{-\beta H})$ where $\mathrm{Tr}(\cdots)$ signifies a trace operation, and H is the Hamiltonian written in the form of a matrix. In the basis of energy eigenstates, H is diagonal with the eigenvalues ϵ_i on the diagonal, so the trace expression reproduces (69). (The reader unfamiliar with matrix methods can ignore this and simply use (69)).

We will now show how various thermodynamic quantities are related to the partition function.

First consider the average energy which is given by

$$\bar{U} = \sum_i p_i \epsilon_i. \tag{70}$$

To find this, observe that

$$\frac{\partial Z}{\partial \beta} = \sum_i -\epsilon_i e^{-\beta \epsilon_i} = -\sum_i \epsilon_i (Z p_i) \tag{71}$$

where in the partial derivative all the energies ϵ_i are constant. It follows that

Internal energy

$$\bar{U} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}. \tag{72}$$

Next we find the entropy:

$$\begin{aligned} S &= -k_{\mathrm{B}} \sum_i p_i \ln p_i \;=\; k_{\mathrm{B}} \sum_i p_i (\beta \epsilon_i + \ln Z) \\ &= k_{\mathrm{B}} \left(\beta \bar{U} + \ln Z \right) \end{aligned} \tag{73}$$

Hence

Entropy

$$S = k_{\mathrm{B}} \left(-\beta \frac{\partial \ln Z}{\partial \beta} + \ln Z \right). \tag{74}$$

Next we shall find the temperature. The fundamental relation of thermodynamics (1st and 2nd law) is

$$\mathrm{d}U = T\mathrm{d}S + X\mathrm{d}R \tag{75}$$

where X and R are suitable macroscopic variables such that $X\mathrm{d}R$ expresses the work in a reversible change in a simple system.[7] Hence

$$\frac{1}{T} = \left.\frac{\partial S}{\partial U}\right|_R \tag{76}$$

We have

$$\left.\frac{\partial S}{\partial \beta}\right|_\epsilon = k_{\rm B}\left(\left.\frac{\partial \ln Z}{\partial \beta}\right|_\epsilon + \bar{U} + \beta\left.\frac{\partial \bar{U}}{\partial \beta}\right|_\epsilon\right) = k_{\rm B}\beta\left.\frac{\partial \bar{U}}{\partial \beta}\right|_\epsilon \tag{77}$$

where the constant ϵ is a shorthand notation for the condition that all the energy levels are constant. It follows that

$$\left.\frac{\partial S}{\partial \bar{U}}\right|_\epsilon = k_{\rm B}\beta. \tag{78}$$

We now argue that constant energy levels ϵ_i corresponds to constant R in the thermodynamic expression. Therefore

Temperature

$$T = \frac{1}{k_{\rm B}\beta}, \qquad \text{or} \quad \beta = \frac{1}{k_{\rm B}T} \tag{79}$$

where we have taken $\bar{U}$ as the internal energy in the thermodynamic expression.

For a system with internal energy U the *Helmholtz function* F is defined by

$$F = U - TS. \tag{80}$$

Using (72), (73) and (79) we have

Helmholtz function

$$\bar{F} = \bar{U} - k_{\rm B}T(\beta\bar{U} + \ln Z) = -k_{\rm B}T\ln Z. \tag{81}$$

The relationship between Z and $\bar{F}$ can also be expressed $Z = \exp(-\beta\bar{F})$ which is reminiscent of the Boltzmann distribution and can serve as a mnemonic for this important result.

Because (81) is easy to remember, it can serve as a useful alternative way to get the entropy (i.e. an alternative to (72)). We recall the thermodynamic relationship

$$\mathrm{d}F = -S\mathrm{d}T + X\mathrm{d}R \tag{82}$$

(which follows immediately from the definition $F = U - TS$ and the fundamental thermodynamic relation (75)), and thus obtain

$$S = -\left.\frac{\partial F}{\partial T}\right|_R. \tag{83}$$

9.1 ‘Single particle’ partition function

The expressions (68)–(81) all apply to a single system, of whatever size and composition, in thermal equilibrium. In practice for many systems it may be difficult to analyse the energy level structure in full, and difficult to obtain the partition function. An important simplification occurs if our system is itself an ensemble of many smaller, simpler

---

[7] More generally there could be many pairs of such variables giving rise to further terms in the expression for $\mathrm{d}U$; this does not affect the argument here if we allow ‘constant R’ to stand in for ‘constant all such quantities’ so that (76) still holds.

systems which are only weakly interacting. In this situation the partition function of one of the small parts is written Z_1 and if the systems are distinguishable one finds

$$Z = Z_1^N \tag{84}$$

(see exercise 5). Notice that all the main macroscopic quantities mentioned in (72)–(81) depend on $\ln Z$ not Z directly, so they can all be obtained from $N \ln Z_1$ and indeed one finds

$$\bar{U} = N\bar{U}_1, \qquad S = NS_1, \qquad \bar{F} = N\bar{F}_1 \tag{85}$$

in an obvious notation. For systems of this type the quantity Z_1 is often called the 'single particle partition function' even though the small parts do not need to be single particles; they can be anything at all as long as the system is a weakly-interacting set of them.

A small confusion that can arise is that some texts seem to imply that when we treat Z_1 we are treating a microcanonical ensemble. In fact we are not, since the definition of Z_1 implies that the probabilities follow the Boltzmann distribution, so we are treating thermal equilibrium, which is equivalent to a canonical ensemble. The confusion arises because there is a finite number N of subsystems. But the entire collection of N subsystems can itself be in equilibrium with a thermal reservoir, so it is perfectly possible to have precisely thermal equilibrium conditions at a well-defined temperature for a finite system, and this is what we have typically in mind when we employ the single-particle partition function.

9.2 Heat capacity

The heat capacity plays an important role because it is often the most directly measurable physical quantity from which information about energy and entropy can be obtained. A heat capacity always has the form

$$\frac{đQ_x}{\mathrm{d}T} = T\left.\frac{\partial S}{\partial T}\right|_x \tag{86}$$

where Q is heat and T is temperature, but a given system has an infinite number of different heat capacities owing to the fact that $đQ$ depends on the type of change the system undergoes when it is being supplied with heat. For a fluid the two most important heat capacities (called *principle heat capacities*) are the constant-pressure and constant-volume heat capacities, written C_p and C_V. In statistical physics the most important heat capacity, for any system, is the one describing heating or cooling such that the energy eigenstates and their energies do not change. For this heat capacity, and only this one, the change in internal energy is equal to the heat entering, and therefore

$$C = T\left.\frac{\partial S}{\partial T}\right|_\epsilon = \left.\frac{\partial \bar{U}}{\partial T}\right|_\epsilon \tag{87}$$

where the ϵ indicates constant energy levels. This 'constant ϵ' is taken as understood when the expression is written without explicitly indicating what is constant in the partial derivative.

There is an interesting relationship between this heat capacity and the system's energy fluctuations in thermal equilibrium. Starting from the expression for mean energy,

$$\bar{U} = \sum_i p_i \epsilon_i = \sum_i \frac{e^{-\beta\epsilon_i}}{Z}\epsilon_i \tag{88}$$

one obtains

$$\begin{aligned}
\frac{\partial \bar{U}}{\partial \beta} &= \sum_i -\frac{e^{-\beta\epsilon_i}}{Z}\epsilon_i^2 - \frac{1}{Z^2}e^{-\beta\epsilon_i}\frac{\partial Z}{\partial \beta} \\
&= -\langle \epsilon^2 \rangle - \sum_i p_i \epsilon_i \frac{1}{Z}\frac{\partial Z}{\partial \beta} \\
&= -\langle U^2 \rangle + \langle U \rangle^2
\end{aligned} \tag{89}$$

where in the last step we used that $\langle U^2 \rangle$ is simply another name for $\langle \epsilon^2 \rangle$ (the symbols U and ϵ refer to the same quantity), and we used the expression (72) giving $\bar{U}$ in terms of Z. We have, therefore,

$$\frac{\partial \bar{U}}{\partial \beta} = -\Delta U^2 \tag{90}$$

where ΔU^2 is the variance of the energy distribution. But since

$$C = \frac{\partial \bar{U}}{\partial T} = \frac{\partial \bar{U}}{\partial \beta} \frac{\mathrm{d}\beta}{\mathrm{d}T} = -\frac{1}{k_\mathrm{B} T^2} \frac{\partial \bar{U}}{\partial \beta} \tag{91}$$

we have

Relationship between heat capacity and energy fluctuations

$$\Delta U^2 = k_\mathrm{B} T^2 C \tag{92}$$

It makes physical sense that a system with a large heat capacity should also have large variance of its energy distribution when it is in equilibrium with a heat bath. For, if the heat capacity is large, it means a given influx of energy from the reservoir does not change the system's temperature very much, so the system does not easily become hotter and for that reason cause a heat flow back again.

For many systems at 'high' temperatures (i.e. such that $k_\mathrm{B}T$ is large compared to the relevant energy level splittings) one has $\bar{U} \sim CT$ (this is just a rough order of magnitude) and therefore

$$\frac{\Delta U}{\bar{U}} \sim \frac{\sqrt{k_\mathrm{B} T^2 C}}{CT} = \sqrt{\frac{k_\mathrm{B}}{C}} \sim \frac{1}{\sqrt{N}} \tag{93}$$

where in the last step we asserted $C \sim N k_\mathrm{B}$ which is again a rough statement but applicable to many systems at high temperature. From this we deduce that in thermal equilibrium the standard deviation of the system's energy distribution (often called the energy fluctuation) is about a factor $\sqrt{N}$ smaller than the mean energy.

10 The free energy method

This section can be skipped on first reading. It presents an alternative way to derive the Boltzmann distribution.

In Section 4 we derived the Boltzmann distribution as the distribution of probability which maximises entropy under the constraints of fixed average internal energy and fixed total probability. We then considered a pair of (weakly-interacting) systems in thermal contact, and we explored what happens if the apply the Boltzmann distribution to the systems individually (as well as to the composite system overall). We argued that the result makes sense and this justifies the assumption that the constraint associated with mean energy is the right one to describe thermal equilibrium. Arguably, no distribution other than (19) could lead to the maximum entropy of the composite system, but we did not prove this.

We will now present an argument which will underline the conclusion and also provide another way to arrive at it. We begin with the following *thermodynamic* observation: when a system is in thermal equilibrium with a heat bath at a given temperature, the overall entropy of both is maximised when the *free energy* of the system is minimised, where the free energy is given by

$$F = U - TS. \tag{94}$$

The term 'free energy' is a thermodynamic concept with wider application; in the situation of fixed energy levels and number of particles, the free energy is the Helmholtz function.

Using Gibbs's formula for entropy, we have

$$F = \left[\sum_i p_i \epsilon_i\right] + k_{\rm B} T \left[\sum_i p_i \ln p_i\right]. \tag{95}$$

We would like to find the distribution of probability which will minimise this quantity, with just a single constraint:

$$\sum p_i = 1. \tag{96}$$

Note, there is no constraint on the system energy in this argument, because the system can get any energy it may require from the reservoir. Using a Lagrange multiplier α, the minimum free energy occurs at a stationary value of

$$\sum_i p_i \epsilon_i + k_{\rm B} T p_i \ln p_i + \alpha p_i \tag{97}$$

By differentiating with respect to p_j at constant $p_{i \neq j}$ and setting the result to zero we find

$$\epsilon_j + k_{\rm B} T (\ln p_j + 1) + \alpha = 0 \tag{98}$$

and therefore the free energy is minimized when

$$p_j \propto e^{-\epsilon_j / k_{\rm B} T}. \tag{99}$$

This is the Boltzmann distribution once again. Notice that the parameter β never appeared in this argument. Note also that we explicitly considered thermal equilibrium with a reservoir from the outset, so there was a well-defined temperature and no need to assume anything about the system energy.

11 Exercises

1. A system has three energies levels with energies $\{0, 1, 2\}\epsilon$, and degeneracies $1, 3, 5$, as in Fig. 3. Find the value of β for a state in thermal equilibrium with mean energy ϵ. Hence find the probability of occupation of the ground state, and the entropy. [*Ans.* $\beta\epsilon = \ln\sqrt{5}$, $p_0 \simeq 0.29925$, $S \simeq 2.01118\, k_{\rm B}$.]

2. A certain system has equally spaced energy levels, with the following degeneracies (i.e. number of states at that energy):

energy	W
0	1
ϵ	5
2ϵ	15
3ϵ	35
4ϵ	70
5ϵ	126

   (i) What is the entropy when the system in is energy equilibrium with energy 2ϵ?
   (ii) What is the entropy of a pair of such systems in thermal contact and otherwise isolated, with total energy exactly 2ϵ?
   (iii) As part (ii), but with total energy exactly 4ϵ? [*Ans.* $k_{\rm B} \ln 15$, $k_{\rm B} \ln 55$, $k_{\rm B} \ln 715$]

3. Derive (64) from the equations leading up to it.

4. Use the expression (81) relating Helmholtz function to partition function to obtain the thermal equilibrium entropy of a harmonic oscillator of frequency ω as a function of $\beta\hbar\omega$.

5. Show that if systems 1 and 2 are weakly interacting but otherwise independent of one another, then the partition function of the joint system is $Z_1 Z_2$. Hence prove by induction that for a system composed of N weakly interacting subsystems all of the same type but distinguishable from one another, $Z = Z_1^N$.

6. Re-derive all the results of Section 9, relying only on your own wits and memory.

7. Consider the simplest possible system: just two quantum states, with energies $\pm\epsilon$.
   (i) Show that the partition function is $Z = 2\cosh\beta\epsilon$.
   (ii) Find the internal energy, first from $\sum_i p_i \epsilon_i$ and then from (72).
   (iii) Show that the heat capacity is $C = k_{\mathrm{B}}(\beta\epsilon/\cosh\beta\epsilon)^2$.

8. Suppose we shift the zero of potential energy, with the result that all the energy levels change by the same amount V. What happens to the partition function, the internal energy and the entropy?

9. System B is like system A except that for every quantum state of A, there are 3 quantum states of B, all of the same energy. How does this difference affect the quantities $Z, \bar{U}, S$?