

2. Microstates & Macrostates

Goal: To define microstates in a macroscopic system, and learn how to count them; and formulate our basic probabilistic assumptions.

We will for the most part deal here with systems as described by quantum mechanics, where it is more obvious how to count states. The basic idea is pretty simple – we imagine a macroscopic system to be built up from a set of microscopic systems, and we wish to find out how many individual quantum states of the system (the ‘microstates’) are associated with a given macrostate. As we saw in the last chapter, the macrostate will be specified by fixing a few extensive and intensive thermodynamics variables such as density, volume, pressure, magnetization etc. The question then is to count all of the different microstates which give the same macrostate.

At first glance this problem seems pretty straightforward, provided the individual constituents of the quantum system are simple enough. Suppose, eg., we have a set of N elementary constituents (particles, oscillators, spins, etc.); then if we now know how to count states for one of these, then it seems a fairly simple problem to do this for N of them. Let’s see how we might do this.

Examples of individual sub-systems: In QM a microstate is specified by the giving the *quantum numbers* of each relevant system. Let’s see how to do this for a few simple systems.

(i) Non-relativistic particle in a 1d box:

$$\begin{aligned} \text{The allowed energies are: } E_n &= \frac{(\hbar k_n)^2}{2m} = \frac{\pi^2 \hbar^2}{2mL^2} n^2 & (1) \\ \text{with quantum numbers: } n &= 1, 2, 3, \dots \end{aligned}$$

(ii) Relativistic particle in a 1d box:

$$\begin{aligned} \text{The allowed energies are: } E_n &= \hbar k_n c = \frac{hcn}{2L} & (2) \\ \text{with quantum numbers: } n &= 1, 2, 3, \dots \end{aligned}$$

(iii) Spin S particle in magnetic field $\mathbf{B} = zB$:

$$\begin{aligned} \text{The allowed energies are: } E_m &= -\gamma \hbar m B & (3) \\ \text{with quantum numbers: } m &= -S, -S+1, \dots, +S \end{aligned}$$

where $\mu = \gamma \hbar s$ is the magnetic moment, and m is the z component of spin (in units of \hbar). Thus there are $2S+1$ possible values for m . Notice the simplest example of a spin S system – this is the two-level system (TLS), or spin- $1/2$ particle (often known these days

as a ‘qubit’). There are then two possible states, \uparrow or \downarrow (which then give a magnetization $M = \mu$ or $-\mu$ for this TLS).

(iv) 1-d simple harmonic oscillator (SHO):

$$\begin{aligned} \text{The allowed energies are: } E_n &= \left(n + \frac{1}{2}\right)h\nu, & (4) \\ \text{with quantum numbers: } n &= 0, 1, 2, \dots \end{aligned}$$

where $\nu = \frac{1}{2\pi} \left(\frac{a}{m}\right)^{1/2}$ and a is the restoring force constant: i.e. $F = -ax$.

(v) quantum rotator e.g. diatomic molecule:

$$\begin{aligned} \text{The allowed energies are: } E_j &= j(j+1)E_0 & (5) \\ \text{With quantum numbers: } j &= 0, 1, 2, \dots \end{aligned}$$

where $E_0 = \frac{\hbar^2}{2I}$ is the basic energy unit, where j is the angular momentum quantum number, and where I is the moment of inertia about the center of mass – if we have 2 masses m_1 and m_2 at distances r_1 and r_2 from the com, then $I = m_1r_1^2 + m_2r_2^2$. Note that E_0 is around 10^{-4} eV for typical small molecules.

We observe here that the important inputs from quantum mechanics are the energy spectrum – the eigenvalues - and their associated degeneracies. The particle wave functions are of lesser importance in SM. This is good because it is often much easier to determine the energy eigenvalues of a quantum system than the eigenstates.

For more details of some of these Hamiltonians, see the Appendix at the end of this chapter.

2(a) An N-Particle System: the Spin $\frac{1}{2}$ Magnet (N qubits)

Probably the simplest example of a thermodynamic system that can be described microscopically is a set of N qubits, or spin-1/2 systems. These may or may not be coupled together, and/or in an external magnetic field – the exact form of the Hamiltonian does not change how we count the states. .

Let's start off by letting $N=2$, so that we have a pair of qubits. Let's also put them in a field so as to classify the states according to their projection along the z-axis. Then for this pair of spins there are 4 states: one state ($\downarrow\downarrow$) with $M = -2\mu$, two states ($\downarrow\uparrow$ and $\uparrow\downarrow$) with $M = 0$ and one state ($\uparrow\uparrow$) with $M = 2\mu$.

In the same way we can count 8 states for 3 spins: there is one state ($\downarrow\downarrow\downarrow$) with $M = -3\mu$, 3 states ($\downarrow\downarrow\uparrow, \uparrow\downarrow\downarrow, \downarrow\uparrow\downarrow$) with $M = -\mu$, 3 states ($\uparrow\uparrow\downarrow, \downarrow\uparrow\uparrow, \uparrow\downarrow\uparrow$) with $M = \mu$ and one state ($\uparrow\uparrow\uparrow$) with $M = 3\mu$.

It should be obvious by now that to classify and then count the states for N qubits, one is dealing with the same problem as that of a run of N coin tosses, where the states 'heads' and 'tails' correspond to the two spin states for a single spin. Thus if we flip a coin N times and determine the number of heads N_h and tails N_t , we see that this is the same problem, with $n \equiv N_h - N_t$, and where $P(N, n)$ is the probability of observing the difference n .

Let's therefore define $n = N_\uparrow - N_\downarrow$. Since $N = N_\uparrow + N_\downarrow$ one can rewrite $N_\uparrow = (N + n)/2$ and $N_\downarrow = (N - n)/2$. The number of states for which we have net magnetization $M = n\mu$ is then:

$$\Omega(N, n) = \frac{N!}{N_\uparrow! N_\downarrow!} = \frac{N!}{\left(\frac{N+n}{2}\right)! \left(\frac{N-n}{2}\right)!} \quad (6)$$

This number is sometimes called the "multiplicity" for the set of states corresponding to this fixed value of magnetization. Note that fixing the magnetization in this way is essentially fixing the macroscopic state for the system; and the multiplicity then counts the number of microstates associated with this macrostate.

The **probability** for the system to have magnetization $M = n\mu$ can then be found if we make some assumption about how likely each one of the microstates is. Suppose we assume here that all of the 2^N microstates are equally likely. Then the probability of getting this particular macrostate is just

$$P(N, n) = \frac{\Omega(N, n)}{\Omega_{tot}} = \frac{\Omega(N, n)}{2^N} \quad (7)$$

This result may not appear to be very helpful – but if N is large we can use Stirling's approximation (see Appendix for more details), according to which

$$\ln N! \cong N \ln N - N + \frac{1}{2} \ln(2\pi N).$$

Now let us write: $x \equiv \frac{n}{N}$ so that $\frac{(N+n)}{2} = \frac{N}{2}(1+x)$ (8)

So that for small x , ie., for $x \ll 1$, one has: $\ln(1 \pm x) \approx \pm x - \frac{x^2}{2}$.

It is then straightforward to show that the logarithm of the probability $P(N, n)$ given in eqtn (1) above is just

$$\ln P(N, n) \approx +\frac{1}{2} \ln\left[\frac{2}{\pi N}\right] - \frac{1}{2} N x^2 \quad (\text{where we neglect terms of order } x^3).$$

To get the desired result for $P(N, n)$ we then exponentiate this expression for $\ln[P(N, n)]$ to get

$$P(N, n) \cong \left(\frac{2}{\pi N}\right)^{1/2} \exp\left[-\frac{n^2}{2N}\right] \quad (9)$$

and for the multiplicity $\Omega(N, n) \cong 2^N \left(\frac{2}{\pi N}\right)^{1/2} \exp\left[-\frac{n^2}{2N}\right]$ (10)

To get some feeling for these numbers, let's consider a simple example. A human body contains about 10^{26} proton spins, of spin-1/2 each, and each carrying a magnetic moment μ . In zero magnetic field the average moment is zero. The maximum moment is $N\mu$ so the rms deviation from zero is

$$\sqrt{N} \mu = \frac{1}{\sqrt{N}} N \mu = 10^{-13} N \mu.$$

This is a very sharp distribution due to the large number of spins. Note that this means that as a function of time, the magnetic moment will be fluctuating, mostly in this very small range around zero.

Notice that we can think of the result (10) for the multiplicity in a somewhat different way. Suppose we ask how many microstates there are as a function of total energy of the system. According to equation (3), in the absence of any interactions between the spins, the energy of each spin-1/2 system is given by $\varepsilon_m = -m\mu B$, where m is the polarization along the field axis (ie., $m = 1$ or -1).

It is then clear that the energy of N spins in a state with net polarization n must be given by $E_n = -n\mu B$, with $n = N, N-1, N-2, \dots, -N$. These are the only allowed energies. We can then define the " **N -particle density of states**" for the system as the sum

$$\mathcal{N}(E) = \sum_n \Omega(N, n) \delta(E - E_n^{(N)}) \quad (11)$$

where $\delta(x)$ is just the Dirac delta function, and $E_n^{(N)}$ refers to the allowed energies of the N -particle system; and the sum is over the different energy levels of the system, labelled by the quantum number n . This density of states is just a sequence of equally-spaced delta functions, with magnitude given by the Gaussian in (10), peaked around $n = 0$, so that the n -particle density of states $\mathcal{N}(E)$ is peaked around $E = 0$.

Related Systems: Clearly there are going to be a lot of 2-state systems in Nature, so that assemblages of N of any of these will have their states counted in the same way. Thus we can say that the spin $\frac{1}{2}$ magnet ‘maps onto’ or is closely related to many other systems. Here are two examples:

1. **Random walk in 1D:** We will assume a total of N steps, with N_l steps to the left and N_r steps to the right. $n \equiv N_r - N_l$. The probability for taking a net total of n steps to the right is again $P(N, n)$

2. **Binary Alloy:** Consider a bcc lattice of CuZn with 50% Zn and 50% Cu. The ordered state has all the Cu at the body center of each cubic elementary cell, and only Zn at the corners of the cube. Assume there are a total of $2N$ sites – *ie.*, N regular Cu sites and N regular Zn sites. The total multiplicity in the fully ordered state is then just $\Omega = 1$.

Now define $f \equiv$ fraction of correctly positioned Cu atoms on the N regular Cu sites. Let Ω_1 be the number of ways (multiplicity) of arranging these fN Cu atoms on the N regular Cu sites. Correspondingly let Ω_2 be the number of ways of arranging incorrectly positioned $(1-f)N$ Cu atoms on the N Zn sites. We don’t have to worry about the Zn atoms since once the Cu are positioned that determines where the Zn are. The total multiplicity is the product of Ω_1 and Ω_2 , *ie.*, we have

$$\Omega_1 = \frac{N!}{[fN]![(1-f)N]!} = \Omega_2$$

As expected $\Omega_1 = 1$ for $f = 0$ or 1 . Note in this problem fN corresponds to the number of up spins in the spin $\frac{1}{2}$ magnet and $(1-f)N$ corresponds to the number of down spins. The difference $n = fN - (1-f)N = N(2f - 1) = N\psi$, where $\psi \equiv 2f - 1$, may be considered to be an *order parameter*. Finally, we see that for large N one has

$$\Omega_1 = \frac{2^{1/2}2^N}{(\pi N)^{1/2}} \exp[-n^2 / 2N] = \frac{2^{1/2}2^N}{(\pi N)^{1/2}} \exp[-N\psi^2 / 2] \quad (12)$$

So that the total multiplicity

$$\Omega = \Omega_1 \Omega_2 = \frac{2^{2N+1}}{\pi N} \exp[-N\psi^2] \quad (13)$$

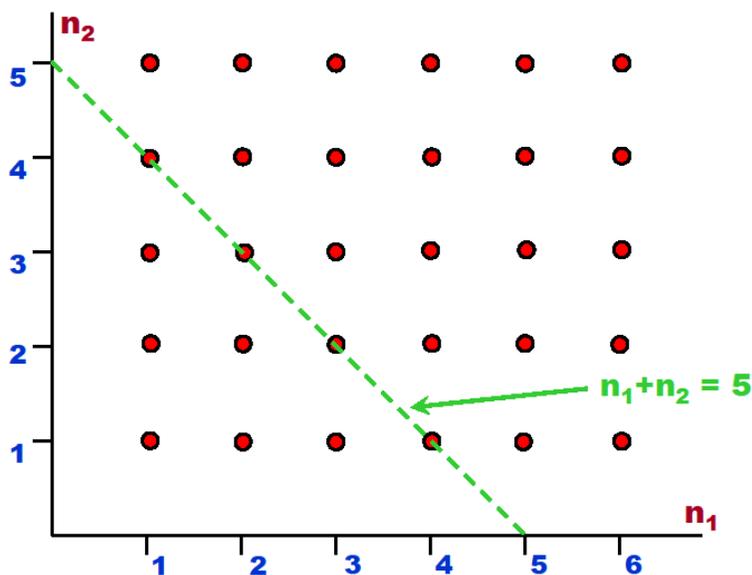
We will come to the idea of an order parameter in more detail later in these notes.

2(b) Other Simple N-Body Systems

We mentioned above that there were various N-body systems that could be described in terms of simple constituents. Here I make a few brief remarks on some of these – we will look in more detail at them in later chapters.

N Particles in a Box: We saw that a 1-d SHO has a spectrum given by eqtn. (1) for a non-relativistic particle, and by (2) for a relativistic particle. There is then a crucial difference between the two, for the energy increases quadratically with quantum number n for the non-relativistic particle, but linearly for the relativistic particle. There is also a key difference between both of these systems and the set of TLS we have just been studying – this is that the energy of each particle is unbounded from above.

Let's look first at the easiest case to analyze, the relativistic particle. With the spectrum given in eqtn. (2), we can start off by plotting a simple graph of the allowed values for a *pair* of particles. Note that a pair of 1-d particles is also equivalent to a single particle moving in a 2-dimensional box of side L ; and we can also generalize to the problem of a single particle moving in a 2-dimensional box with sides L_1 and L_2 .



Let us look first at two 1-d particles, so that in the the 2-D box analogy, we have equal sides, ie., $L_1 = L_2 = L$. The allowed states of the system are now classified by integers n_1 and n_2 , and the allowed energies are

$$E(n_1, n_2) = (hc/2L) [n_1 + n_2] = (hc/2L) p = E_0 p = E_p \quad (14)$$

where $E_0 = hc/2L$, and where $p = (n_1 + n_2) = 2, 3, 4, \dots$ etc.; this last result follows because the sum of any pair of integers is also an integer (see Figure).

We see easily from the figure that the degeneracy of a state with quantum number p and corresponding energy E_p is just $p-1$, and so if we want to write the density of states for this 2-oscillator system, we just have the result

$$\mathcal{N}(E) = \sum_p (p) \Omega_2(p) \delta(E - E_p) \quad (2 \text{ 1d particles}) \quad (15)$$

where $\Omega_2(p) = p-1$. If $p \gg 1$, then we can just write $\Omega_2(p) = p$ for all practical purposes. In any case we see that the density of states is just a series of delta-functions, spaced by energy E_0 , with magnitude increasing linearly with energy.

Note that we have made a rather special assumption here, viz., that $L_1 = L_2 = L$. If instead we look at a particle in a 2-D box, and the sides of the box are not equal in size, we will no longer have the degeneracy we just found – the spectrum will be

$$E(n_1, n_2) = (hc/2L_1) n_1 + (hc/2L_2) n_2 \quad (16)$$

and we no longer have all the accidental degeneracy we had when $L_1 = L_2 = L$. However it will be clear that when we go to large energies, the number of states in some finite energy interval will still increase linearly with E . We can then define a ‘smoothed density of states, which will be linear in E .

So much for the 2-dimensional problem. Now suppose we go to the problem of N one-dimensional particles. It is easy to see how the results will generalize; the energy spectrum will now be

$$E(\{n_j\}) = (hc/2L) [n_1 + n_2 + \dots + n_N] = (hc/2L) p = E_0 p = E_p \quad (17)$$

where now we have $p = N-1, N, N+1, \dots$ etc. (so that the minimum energy of the set of N particles is $(N-1) E_0$), and p is now defined as

$$p = (n_1 + n_2 + \dots + n_N) \quad (18)$$

We saw in the case of the pair of particles that the allowed values of the total energy could be defined by diagonal lines in the 2-d plane in the figure above; now we can see that for N particles, we have the allowed energies defined by a set of $(N-1)$ -dimensional hyperplanes, and with a bit of thought you will see that the number of allowed states or

“points” in a hyperplane corresponding to energy E_p will be grow very fast – in fact we have

$$\Omega_N(p) \rightarrow p^{N-1} / (N-1)! \quad (N \text{ 1d particles, for } p \gg 1) \quad (19)$$

so that the N -particle density of states increases incredibly fast with energy. If you want to amuse yourself, try finding the formula for $\Omega_N(p)$ valid for all p .

You can now see that the problem for a non-relativistic particle will look quite different, simply because the allowed energies for the particle are no longer spaced linearly in the quantum number, but quadratically. I will not go through this case - if you are interested, try figuring out what the N -particle density of states looks like for this case.

N spin- S particles: the spectrum for a single spin s in a magnetic field was given in eqtn. (3), and we see that to treat this problem, we must generalize in a straightforward way what was done for the spin-1/2 magnet. We see immediately now that the total number of states for an N -spin system will be $(2S+1)^N$, and that the problem of counting the states and determining the probabilities is isomorphic to that of putting N balls in $2S+1$ different boxes – we get a multinomial distribution instead of the binomial distribution appropriate to spin-1/2 systems.

N Non-Relativistic Oscillators: Just as with the particle in a box, the spectrum here (given in eqtn. (4)) is unbounded from above, and in the same way we can, if we wish, impose an upper cut-off to the spectrum. In fact the problem here is very much like that for a set of N relativistic particles moving in 1 dimension – the only difference is that the lowest energy state for each oscillator is one half the splitting between the higher levels, rather than being equal to it as in the case of the relativistic particle. Otherwise the counting goes through as before.

N quantum rotators: Here the spectrum is given by eqtn. (5), and is again unbounded from above. This case is very similar to the non-relativistic particle in a box, and we will say no more about it here.

2(c) Assignment of Probabilities

It will not have escaped your notice that we have made an assumption in writing equation (7), viz., that the probability of getting any one of the microstates is equal to that for all the others. This assumption of equal weight for all the microstates is one we need to unravel – it involves several components.

A priori Probabilities: The first question we have to deal with is what *a priori* probability we will assign to each microstate. We will see that the answer to this question depends on the ensemble of microstates we choose. In the example chosen above, it was assumed that all the microstates in a given macrostate have the same energy. In this case we assigned equal probability to each of them. In the language of SM this ensemble of equal energy states is called the *Microcanonical Ensemble*. We will deal with it in a more formal and precise way in a subsequent chapter. It corresponds to an ensemble where the total energy is fixed, so we require very strict boundary conditions on the system – no energy is allowed in or out of the system.

This still leaves the question of what reason we have for assigning equal probability to each microstate state in this microcanonical ensemble. We come to this below.

However, the microcanonical ensemble is not the only ensemble we can choose. We can, for example, require that energy is allowed to flow in and out of the system, but that the macroscopic *temperature T* be fixed (such a macrostate will then incorporate microstates whose energies are NOT all the same). The relevant ensemble of microstates here is then called the *Canonical ensemble*. In this case, as we will see, the probabilities for different microstates corresponding to a state of given T will not be the same; in fact they will be proportional to $\exp(-E/kT)$ for a microstate of energy E.

As we will see, one can define many other ensembles, depending on what is held to be constant. Another popular one is the *Grand Canonical ensemble*, for which one even allows particles as well as energy to move in and out of the system – in this case it is the chemical potential that is held constant, as well as, eg., the temperature. In a later chapter we will specify what probabilities are then assigned to the microstates for this ensemble.

At this point you will see that each of these different ensembles of microstates, corresponding to some macrostate, can also be described in terms of the macroscopic states defined by TD (as discussed in Chapter 1). Again, we will discuss this in detail in subsequent chapters.

Notice again that I have not yet given you a reason for believing in these assignments of *a priori* probabilities – I am just telling you what they are assumed to be.

Ergodicity – Partial or Complete: There is another assumption we must discuss which is also important in the foundations of SM. It concerns the assumption that a given system, which has available to it some set of microstates, will actually sample all of those microstates according to the probabilities assigned to them.

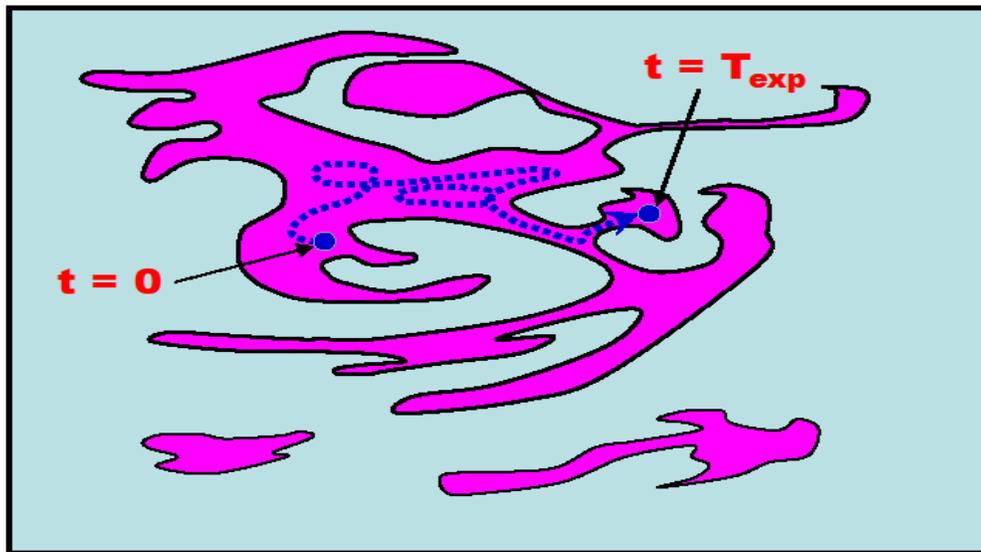
Now it is crucial to understand that this is a question about the *dynamics* of the macrosystem. Suppose that in some experiment, one looks at the system for some period of time T_{exp} , starting at time $t = 0$, and continuing to time $t = T_{\text{exp}}$. At $t = 0$, the macrosystem will start in some microstate. Will, it during the time period T_{exp} , then visit all the other states with equal probability? And, will it visit a reasonable fraction of these states within the time T_{exp} ? Old-style SM theory argued that it would, so that *time*

averages for the macroscopic state would be equivalent to *ensemble averages* over the available microstates. This is known as the “**ergodic hypothesis**” (which can actually come in various forms).

Note how convenient the ergodic hypothesis would be if it were true! For if true, it means that we can assume that not only are the states all a priori equally likely, but that the system will visit them all with equal likelihood over a period of time which one hopes is quite short, i.e., less than T_{exp} . In this case, if we measure any quantity for the system, it will accurately reflect the probabilities for the microstates, in that it will average over all of them according to these probabilities.

The reason why old-style SM theory assumed the ergodic hypothesis, and why quite a few theorists even tried to turn it into a fundamental hypothesis underlying all of SM, is that these theorists spent too much time looking at simple models or at very simple many-body systems for which it was more or less true. Thus they looked at the dynamics of weakly-interacting gases, or weakly-interacting oscillators, or spin systems with simple interactions. These systems were chosen because they were simple, not because they were necessarily realistic, or representative of the real world.

However it was already well understood in the 1950s, at least by some, that the ergodic hypothesis is highly unlikely to be true for most physical systems – and that in fact most realistic physical system depart very strongly from ergodic behavior. Let’s see what is actually involved here.



T

The figure shows a schematic description of the space of possible microstates for a physical system (in reality this is the phase space for the system, and phase space language can be used to describe both classical and quantum systems). Suppose the allowed microstates, corresponding to a given macrostate, are those shown in the pink region. We imagine that at time $t = 0$ the system finds itself in some particular microstate, and evolves in time following the hatched trajectory shown, to finish at a later time T_{exp}

at a rather different point. Now if the ergodic hypothesis is correct, the system will move rapidly all around the pink regions, so that it covers them pretty uniformly, and all memory of where it began is lost after a time less than T_{exp} .

How likely is this? Well, for a simple system like a gas in a box, it is pretty accurate – if we start the gas in some microstate, it will very rapidly cycle through many other states, and spend an amount of time in any one region which is proportional to the ‘size’ of that region (ie., to the number of microstates in that region). Even if we start the gas in some highly organized (and very special) initial state (eg., with all particles moving in the same direction), collisions between the gas particles and with the walls will rapidly cause it to lose memory of this initial state (actually this is not entirely true – 3-body interactions, even in a dilute weakly-interacting gas, cause long-time memory effects and a weak breakdown of ergodicity).

However, the picture also makes clear that the example of a gas is way too special. More often it is not so easy – or for all practical purposes impossible – for the system to make the transition from one pink region to another (notice that some regions in the figure are disconnected from others). Alternatively, the required trajectories to get from one region to another may be so convoluted that they just take a really long time – how this might happen is also obvious from the figure.

From these considerations it is clear that the ergodic hypothesis is only going to be valid if the dynamics of the system allows it. And the simple fact is that this is **not** the case for almost all solid systems, and for a large variety of other systems where non-linearities are important in the dynamics.

Consider, eg., the systems we mentioned at the beginning of Chapter 1. For these systems we can make the following remarks:

a litre of gas: The dynamics here becomes almost completely ergodic in the limit where interactions between the gas particles are very weak.

a wafer of Si: the phonons relax quickly – depending on the temperature, they behave quasi-ergodically over timescales ranging from seconds to microseconds. However at low T , much of the entropy is in nuclear spins and defects. At low T , in a magnetic field, it may take the nuclear spins days to relax; and the defects may not relax to thermal equilibrium for months or years. If there are lots of defects, interacting with each other, some of them may not relax for astronomically long times. Thus, only a sector of the degrees of freedom behave ergodically (with a slow communication to the other degrees of freedom)

a bar magnet: The same remarks as above apply to the phonons, nuclear spins, and defects. Magnets also have ‘domain wall’ configurations in them. These are in a metastable state but may not relax for astronomical times. They also interact with each other strongly, and this means that there are ‘glassy modes’ connected with their dynamics which may never relax (except over timescales proportional to $\exp[N^\alpha] \tau_0$,

where N is the number of domain walls, α is a number > 1 , and τ_0 is a relaxation times which might be milliseconds. Very far from ergodic.

a melting block of ice: This is not really in equilibrium – however over timescales of minutes or hours (depending on how big the block is) it will behave according to thermodynamics. However only some of the microstates are involved in this – notably the phonons. Again, defects in the ice will relax very inefficiently, even though they are intimately involved in the melting process. Far from ergodic.

a piece of wood: Here the phonons relax quickly, and it is the phonons that define the thermodynamics. There are other modes in the system that are weakly coupled to the phonons, and so should not be included in the definition of the microstates. Far from ergodic.

a piece of glass: Again, only the phonons define the thermodynamics. But here, the behaviour of the glass even up to high temperatures is to a great extent controlled by the coupled dynamics of the defects in the glass (which is highly disordered at the atomic level. This coupled dynamics involves a whole hierarchy of relaxation times, ranging up to $\exp[N^\alpha] \tau_0$, where now N is basically the number of particles in the system. No matter what the value of τ_0 , this timescale is far far larger than the age of the universe. It should be noted that disordered solid systems like glasses are the rule rather than the exception in nature (at least on solid bodies like planets). As far away from ergodic as you can possibly get.

a galaxy: The dynamics of the stars in the galaxy is very far from being ergodic. TD can be applied to certain features of this dynamics – but a proper description, following the KAM theorem, shows that some sectors of the dynamics (ie., some sectors of the phase space) are quasi-ergodic, and others not at all.

a star: Not in equilibrium obviously! And yet TD and SM are extremely useful in analyzing a star's behavior. This is because for long periods it is in a steady state, with energy coming from nuclear fusion being radiated away, mostly as photons. The time it takes the photons to travel from the stellar core to the surface can be millions of years; and the time taken for the macroscopic properties of the star to change ranges from 10^5 up to 10^{14} years, depending on the stellar type. So, astonishingly, if we use TD and SM to treat only the microstates connected with the gaseous degrees of freedom, it works very well (we have to exclude, eg., the nuclear degrees of freedom from our discussion). The behavior is far from ergodic except in small local regions of the star.

the earth: The same remarks apply to the earth, considered as a very large body, this time over timescales of months or years – energy comes from the sun and is reradiated in the form of IR photons. Again, we exclude most of the real microstates from this description. Very far from ergodic.

a lake: Same remarks again, but this time over shorter periods (which depend on the size of the lake). Very far from ergodic.

a fire: This is very far from equilibrium! TD and SM might seem essentially useless for systems like this, undergoing rapid chemical reactions. And yet paradoxically, TD can be applied for very short timescales! This is because even in a fire or a shock wave, there are heat flows which are quasi-adiabatic or quasi-isothermal (depending on the conditions) over short timescales. Very far from ergodic

a person sleeping in a chair: Again, not in equilibrium. But as with a star, if we only look at certain degrees of freedom (here phonons), then for a period of minutes, we can apply TD. Very far from ergodic

a bacterium: Again, not in equilibrium. However, one can apply TD over timescales of milliseconds (to describe the phonons). Very far from ergodic – actually, the bacterial TD is determined largely by the surroundings (ie., the heat bath)

a virus: The same as a bacterium, but now over timescales of microseconds.

So, what have we learned from this, and why does anyone care about the ergodic hypothesis anyway if it so obviously wrong?

There are really three things to take away from this discussion.

- (i) First, that real macroscopic systems are complicated, and that they usually have many different sorts of excitations (excited states, in quantum language), which may or may not communicate with each other or with the external world very effectively; so that they may or may not partake in the thermodynamics and statistical physics. If, over some timescale, some set of internal modes does not couple to those other modes that are determining the thermodynamics behaviour, then these internal modes can be ignored when defining the microstates. The relevant timescales for the coupling between modes can range from microscopic times (less than nanoseconds) to super-astronomical times (immeasurably greater than the lifetime of the universe).
- (ii) Second, that we may assign a priori probabilities to the different microstates, but these probabilities are only meaningful if the system is actually capable of visiting these states – whether they do or not depends on the detailed dynamics of the system.
- (iii) Finally, the description we give of some macroscopic system (using, eg., a Hamiltonian) depends on which modes we are interested in. We can describe an atom without describing the internal modes of the nucleus – these are simply not active or relevant to atomic behavior except in very unusual cases. It is often said that “all Hamiltonians are effective Hamiltonians”. The same is true of microstates for a real system – we select the ones that are relevant to the behavior we are interested in.

2(d) Relation between TD and SM

Let's try and summarize the relationship between thermodynamics (TD) and statistical mechanics (SM). The following table will begin this process – as we go on in the course you will see more of how it goes.

Thermodynamics	Statistical Mechanics
Establishes relationships between macroscopic variables using empirical information	Macroscopic quantities are calculated from assumptions about microscopic constituents (atoms, molecules, etc.)
4 basic laws or axioms – related to the definition of the macroscopic variables	Axioms of probability, and the assumption of microstates for macroscopic systems, with assigned weighting for the microstates. The microstates are described using, e.g. quantum mechanics, with assigned Hamiltonians
0th law : If systems A and B are in equilibrium with C then A is in equilibrium with B. Assumes an operational definition of equilibrium	Equilibrium derived by maximization of probability for collections of microstates, for one or more systems. Temperature defined statistically.
1st law: heat and work are equivalent and total energy is conserved	Total Energy is conserved.
2 nd law: Entropy in an isolated system can only increase. Changes in entropy are well defined. Entropy defined in terms of heat.	Entropy defined for a given collection of microstates – measures the number of such states. System evolve to their most probable states (highest entropy)
3 rd law: At absolute zero ($T=0$) the entropy approaches a constant value.	At $T=0$ the system is in its ground state and entropy is zero (for unique ground state) or finite otherwise.
Only average quantities are calculated – to give the macroscopic variables	Fluctuations and correlations can also be calculated.
Widely applicable, to all macroscopic systems in equilibrium or moving slowly between equilibrium states. Not applicable to systems out of equilibrium.	Even more widely applicable – can treat systems near equilibrium subject to time- and space-dependent perturbations, and their dynamics, correlations, and fluctuations Not applicable to system far from equilibrium.
The underlying microscopic dynamics (quantum, classical, etc.) does not play a role..	Dynamics plays a key role. Quantum mechanics – and indistinguishability - naturally fit in, and play a key role.

As noted above, in TD one deals entirely with quantities that are defined macroscopically, for a system or set of systems, and with the relations between them; we

may also introduce a heat bath as well. The foundations of the subject can be formalized using the 4 laws of TD (and in other ways as well).

In contrast, we shall see that in SM we incorporate in some cases very detailed knowledge of the microscopic constituents. Nevertheless we will be able, from SM, to derive all the thermodynamic state quantities U , T , S , F etc., using probabilistic arguments.

Actually, SM gives us far more than TD. Quite apart from ‘explaining’ TD for systems at equilibrium (in a way that we will see as we go along), it also allows us to look at the microscopic basis of the idea of equilibrium and to look at fluctuations around it, and to look at the properties of a system near to equilibrium which may be subject to time-dependent perturbations. In this latter circumstance it allows us to look at the time-dependent behavior of the system, and at all the correlation functions or ‘response functions’ (written as functions of space and time) which describe how the system reacts to such perturbations.

The only limiting condition on the methods used in SM is that we still require that the system be near equilibrium. Why this is will become clearer as we go along. This is, as far as the real world goes, a significant limiting factor – some key features of what are referred to as ‘far from equilibrium’ behaviour are beyond the reach of SM. Not as far as one might often think, however – it is amazing how much of the dynamics of, eg., shock waves and explosions, can be understood using the methods of SM.

We will in subsequent sections discuss the foundations of SM, and from this delineate its limitations. However, even before we do this, it is interesting and useful to give a brief preview, in which we compare some of the key features of the 2 different approaches to the behavior of macroscopic systems near equilibrium. We do this in tabular form, as shown above.

Appendix: Some Derivations

In the text a number of results are given without explicit demonstration – here I fill in the gaps. We look first at two of the simple Hamiltonians discussed in the beginning of the chapter, and then look at the derivation of Stirling’s approximation.

Simple Model Hamiltonians: At the beginning of the chapter model Hamiltonians are given for a number of simple systems, and we amplify here on two of these, for those who have not met them:

(i) **Relativistic Particle:** in the absence of any charge, a relativistic particle has an energy-momentum dispersion relation given by

$$E^2 = m^2c^4 + p^2c^2 \quad (\text{A.1})$$

Where m is the rest mass of the particle. We do not prove this result here since it relies on detailed knowledge of special relativity, which is beyond the scope of this course. We note two limiting cases of this relation. First, when the momentum p is small, ie., when $p^2 \ll m^2c^2$, then we can expand the energy as

$$\begin{aligned} E &= mc^2 [1 + (p/mc)^2]^{1/2} = mc^2 [1 + (v/c)^2]^{1/2} \\ &= mc^2 [1 + \frac{1}{2}(v/c)^2 + \dots] \\ &= mc^2 + \frac{1}{2}mv^2 + \dots \quad (v \ll c) \end{aligned} \quad (\text{A.2})$$

Second, when the momentum is very high, so that $p^2 \gg m^2c^2$, we have

$$E = pc [1 + (mc/p)^2]^{1/2} = pc + \dots \quad (v \sim c) \quad (\text{A.3})$$

This latter result is in the relativistic limit, and is used in eqn (2) at the beginning of this chapter.

(ii) **Quantum Rotator:** Here we imagine a pair of masses m_1 and m_2 , separated from each other by a distance $R = r_1 + r_2$, where r_1 and r_2 are the distances of each from the centre of mass. The Hamiltonian of the system when rotating is then given by

$$\begin{aligned} H &= \frac{1}{2\mu R^2} \left[p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} \right] \\ &= -\frac{\hbar^2}{2I} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \end{aligned} \quad (\text{A.4})$$

where the latter expression is just the usual result in spherical coordinates. The moment of inertia of the system is $I = \mu R^2$ (which is the same as $I = m_1 r_1^2 + m_2 r_2^2$), and the reduced mass μ is given in terms of the original masses by

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (\text{A.5})$$

(for which see any classical mechanics text). We can then solve Schrodinger's eqn for the system in the usual way – it is just like that for a particle in a spherically symmetric potential. Once the radial term has been extracted, we are left with an angular equation given by

$$\hat{H}Y_\ell^m(\theta, \varphi) = \frac{\hbar^2}{2I} \ell(\ell+1)Y_\ell^m(\theta, \varphi) \quad (\text{A.6})$$

so that the energy eigenvalues are just those given in eqtn (5) of the main text.

Stirling's Approximation: Stirling's approximation is a typical example of the lowest terms in an "asymptotic expansion" of an integral. The full expansion is

$$\ln n! \sim n \ln n - n + \frac{1}{2} \ln(2\pi n) + \frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \dots \quad (\text{A.7})$$

There are various ways of deriving this systematic expansion. The simplest is to start from the standard expression for the factorial in terms of the Gamma function; for integer n , we have $n! = \Gamma(n+1)$, so that

$$n! = \int_0^\infty x^n e^{-x} dx = \int_0^\infty e^{n \ln x - x} dx. \quad (\text{A.8})$$

which can be verified directly by integrating the right-hand side by parts. We can now treat this integral by "steepest descents", which amounts to expanding the function about its maximum. The maximum of the integrand is when $x = n$, and so we write $x = n + y$, and expand in y , ie., write

$$\begin{aligned} \ln x &= \ln n + \ln \left(1 + \frac{y}{n}\right) \\ &= \ln n + \frac{y}{n} - \frac{y^2}{2n^2} + \frac{y^3}{3n^3} - \dots \end{aligned} \quad (\text{A.9})$$

so that we have

$$n! = \int_{-n}^\infty \exp \left[n \left(\ln n + \frac{y}{n} - \frac{y^2}{2n^2} + \dots \right) - n - y \right] dy \quad (\text{A.10})$$

Notice that the term linear in y makes no contribution here, so we can approximate the result as

$$n! \approx e^{n \ln n - n} \int_{-\infty}^\infty e^{-y^2/(2n)} dy = e^{n \ln n - n} \sqrt{2\pi n} = \sqrt{2\pi n} n^n e^{-n} \quad (\text{A.11})$$

where we use the standard result for a Gaussian integral. This is the desired result. If we continue the expansion in eqtn (A.9) to higher orders in y we can get the higher-order terms appearing in (A.7).

Note that a quick and dirty way to get the leading term in (A.11) is to just write

$$\ln n! = \sum_{j=1}^n \ln j$$

And then to approximate this by an integral, as

$$\sum_{j=1}^n \ln j \approx \int_1^n \ln x \, dx = n \ln n - n + 1.$$

Which gives a correct approximation except for the last term on the right hand-side.

The more general method here is useful. Suppose we wish to integrate the function $\exp [Mf(x)]$ over x , between limits a and b , where M is some constant (typically $M \gg 1$). Let us suppose that $f(x)$ has a maximum for $a < x < b$, at the point x_0 , so that the the exponential is very sharply peaked around x_0 . We write

$$f(x) \approx f(x_0) - \frac{1}{2} |f''(x_0)| (x - x_0)^2 \quad (\text{A.12})$$

so that

$$\int_a^b e^{Mf(x)} \, dx \approx e^{Mf(x_0)} \int_a^b e^{-\frac{1}{2}M|f''(x_0)|(x-x_0)^2} \, dx \quad (\text{A.13})$$

Again we have a Gaussian integral, and the result is

$$\int_a^b e^{Mf(x)} \, dx \approx \sqrt{\frac{2\pi}{M|f''(x_0)|}} e^{Mf(x_0)} \quad (\text{A.14})$$

for $M \gg 1$. Again, we can extend this to higher orders by continuing the expansion in (A.12) to higher orders. Note that the limits a and b do not appear in the final answer, since they are supposed to be sufficiently far from the maximum in the integrand.