1. Introduction

Depending on how you look at it, statistical mechanics is either the least fundamental or most fundamental of all fields of physics. That is because it is not really science at all. It is pure mathematics.

In other subjects, you learn about “natural laws”: Newton’s second law, Maxwell’s equations, Schrödinger’s equation, etc. These laws are not derived from anything else. They were discovered experimentally and then assumed to reflect fundamental aspects of reality. But statistical mechanics does not involve any natural laws of this sort. Instead, it is a set of techniques that can be applied to nearly any physical system, no matter what laws that system obeys. That is why I call it the most fundamental field of physics. New theories may replace old ones, and natural laws may turn out to be merely approximations to deeper laws. But statistical mechanics remains valid through it all, and whatever new laws are discovered, it will almost certainly work just as well with them as it did with the old ones.

Statistical mechanics was developed in the second half of the 19th century. It was primarily the work of Ludwig Boltzmann, who personally published more than a hundred papers on the subject during his lifetime. Other scientists also contributed to it, of course, especially James Clerk Maxwell (the same one Maxwell’s equations are named after) and Josiah Willard Gibbs.

It grew out of thermodynamics, an earlier theory that described the behavior of a mysterious substance called “heat” or “caloric”. Thermodynamics was a physical theory of the more conventional sort. It involved natural laws discovered by experiment, and made no claims about why those laws happened to hold. Caloric was supposed to be a substance much like other forms of matter. But this view turned out to be incorrect. Heat is actually an emergent phenomenon: a mathematical quantity defined in terms of a more detailed theory (the movement of individual atoms). The “laws” of thermodynamics can be derived from that deeper theory by applying statistical techniques. If you ignore the details of how each atom is moving, you find they collectively behave in a way that resembles a continuous fluid. That is what statistical mechanics is all about: deriving high level descriptions by starting from lower level ones and then averaging out lots of details.

Is it possible other theories could be explained in the same way? That is an open question, and a fascinating one. Many physicists suspect gravity is an emergent phenomenon, that it arises from the collective behavior of some deeper degrees of freedom. Statistical interpretations have also been proposed for quantum mechanics. These are all very speculative, of course, and they could easily turn out to be wrong. But they also could easily turn out to be right. Based on what we know today, it is entirely possible that the very structure of spacetime is a consequence of statistical mechanics.

2. The Statistical Description of Physical Systems

Given a physical system, how do you apply statistics to describe it? What does that even mean? There are lots of aspects of the system you might try to compute statistics for, but what approach is the most useful? This chapter will attempt to answer these questions by introducing the basic techniques and concepts of statistical mechanics.

Along the way, we will define quantities with names like “temperature”, “pressure”, and “entropy”. Of course you have heard these words before and you probably have an idea of what you think they mean. For the moment, try to put those ideas aside. In Chapter 4 we will come back to them and try to build up an intuitive understanding of these quantities. But for now, just treat them as mathematical definitions.

2.1. Microstates and Macrostates

To describe an isolated physical system with statistics, we begin by making the following fundamental assumption:

Assumption

The system can exist in a discrete (but possibly infinite) set of *microstates*. A microstate defines the values of all possible *microscopic variables*.

In a classical system of point particles, for example, a microstate defines the position and momentum of every particle. In a quantum mechanical system, it defines the value of the wavefunction at every point in space. To specify what microstate the system is in, you must give the most detailed description you will ever care about.

You may be wondering whether this is a valid assumption. What if the microstates are not discrete? For example, the position and momentum of a classical particle are continuous, not discrete. Whether the universe is actually discrete or continuous at its most fundamental level is still an open question.

Fortunately, this turns out not to matter very much. You can always turn a continuous variable into a discrete one by dividing it into very small bins. For the classical particle, we treat all positions between xx and $x+\delta x$x+δx and all momenta between pp and $p+\delta p$p+δp as a single microstate. As long as we choose δxδx and δpδp sufficiently small, the exact values turn out to have no effect on most of our results.

This is illustrated in Figure 2-1, which shows the space of possible microstates for a single particle in one dimension. Each microstate is defined by its values of xx and pp. This space is known as *phase space*, and we will use it often. More generally, a system of NN particles in dd dimensions has a $2dN$2dN dimensional phase space. Every point (or rather, each tiny volume as shown in Figure 2-1) in this phase space represents a microstate.

Figure 2-1. The phase space of a single particle in one dimension. Each axis is divided into tiny intervals. The volume defined by the intersection of one interval from each axis is a microstate.

For typical systems, the number of microstates is huge and they describe the system in far more detail than we usually care about. Consider a box filled with gas. You have no way to measure the exact position and momentum of every last gas molecule, and you would not care about them even if you could measure them. Instead, you are usually interested in a small number of *macroscopic variables*: the total energy of the system, the total number of gas molecules, the volume of space it takes up, etc. These are things you can measure and that have practical importance.

A *macrostate* is defined by specifying the value of every macroscopic variable. There may be a huge number of microstates all corresponding to the same macrostate. For example, suppose you measure the total energy and volume of a box of gas. There are an enormous number of arrangements of the individual gas molecules that all add up to that energy and volume. You know the gas is in *one* of those states, but you have no idea *which* one.

2.2. The Density of States

The number of microstates corresponding to a macrostate is called the *density of states*. It is written $\Omega(E,V,\ldots)$, where the arguments are the macroscopic variables defining the macrostate.

Since we assumed the microstates are discrete, that means Ω is also a discrete function. Each variable can only take on specific values, and it is only defined for those values. In practice, however, we usually treat Ω as a continuous function. As long as the allowed values are sufficiently close together, the total number of states in any given interval will be huge and we can simply think of Ω as the number of microstates per unit volume in the space defined by the macroscopic variables. That is why we call it the “density” of states.

Example

Compute $\Omega(E)$ for a single free particle in three dimensions.

The energy is given by $E=|p|^2/2m$. The microstates are defined by a six dimensional phase space, but since the energy does not depend on the positions, we can ignore them and just consider the three dimensional space of momentum coordinates.

Each value of E corresponds to a two dimensional surface in this three dimensional space, so Ω is proportional to the area of that surface. In this case, it is just a sphere of radius $|p|$. Since the surface area of a sphere is equal to $4\pi r^2$,

$$\Omega(E) \propto |p|^2 \propto E$$

The proportionality constant in this example depends on what value we choose for δp in discretizing the momenta. In most cases, however, we will only be interested in the ratios of Ω for different values of E, not in its absolute value for single energies, so we can ignore it.

In this example, the density of states increases with increasing energy. That is nearly always true. In this case it grows only linearly with E, but it often grows much faster. Consider another example.

Example

Compute $\Omega(E)$ for a collection of N free particles in three dimensions.

This is nearly identical to the previous example, except we now have $3N$ momentum coordinates. Each value of E corresponds to a $3N-1$ dimensional surface. Combining all the momenta into a single vector, we can write the result as

$$\Omega(E) \propto |p|^{3N-1} \propto E^{(3N-1)/2}$$

For macroscopic systems, N will be on the order of 10^{23}, and even a tiny increase in energy will produce a huge increase in the density of states. For this reason, we usually work with $\log(\Omega)$ instead, which is a much smoother, slower changing function.

2.3. The Postulate of Equal *a priori* Probabilities

Suppose you have measured a set of macroscopic variables for an isolated system. You now know which macrostate it is in, but there may be a huge number of microstates all consistent with that macrostate. What can you say about which one of those it is most likely to be in?

The answer, of course, is that you have no idea. Your measurements do not provide any further information to answer that question. Nonetheless, to calculate any averages or other statistical quantities you must assume something. This leads us to the following assumption:

The Postulate of Equal *a priori* Probabilities

A system has an equal probability of being in any microstate that is consistent with its current macrostate.

Is there any justification for this assumption? That turns out to be a very complicated question. In practice it works very well for many different cases, and there are theoretical arguments for it. On the other hand, there certainly are cases where it is not correct. So instead we will approach the question from a slightly different direction and treat it as a definition:

Definition

An isolated system that satisfies the postulate of equal *a priori* probabilities is said to be in *equilibrium*.

(But note that this definition is usually written in a slightly different way. We will come back to it in the next section.)

For the initial part of this book, we will only consider systems in equilibrium. We will therefore treat this postulate as a given. Of course, there are many interesting situations involving systems that are *not* in equilibrium. We will examine some of those later in the book. We will also look at the process by which systems come to be in equilibrium, and what happens if they are then disturbed from it.

2.4. Time Averages and Ensemble Averages

I was a bit careless with terminology in the last section. I spoke of the “probability” of a system being in a particular microstate, but never defined what that meant. After all, at any given moment the system *is* in a particular microstate and not in any other. There is no probability about it.

There are two approaches one can take to defining probabilities in this context. The first is to realize that the degrees of freedom making up the system are constantly changing. At one instant the atoms have particular positions and momenta, but the next instant they are different. We therefore define the probability of the system being in a microstate as the *fraction of time* it spends in that state. Averages computed using this definition of probability are called *time averages*.

This is the older of the two definitions. Boltzmann used this definition in most of his work. It turns out to have problems, however, which led to the introduction of a new definition of probability.

Instead of looking at just one system, imagine preparing many identical systems by following the exact same procedure many times. All of these systems are in the same macrostate, but each one is in a different microstate. You have to simply imagine doing this, because in practice you have only one system in one microstate—but you do not know which one it is. We define the probability of the system being in a microstate as the *fraction* of these imaginary systems that are in that state. The set of all the systems is called a *statistical ensemble*, and averages computed with this definition of probability are called *ensemble averages*.

You might wonder whether these definitions are equivalent. Do both types of averages give the same results? The answer is an emphatic, “Sometimes.” For some systems they do, and for other system they do not. We therefore turn this into another definition:

Definition

A system for which time averages and ensemble averages are equal is said to be *ergodic*.

(Isn’t that a wonderful word? Ergodic. It comes from the Greek words for “work” and “path”. I recommend working it into your conversation frequently.)

Having hopefully clarified that, we should reconsider the definition from the previous section. I said that a system satisfying the postulate of equal *a priori* probabilities (can I abbreviate that PoEapP?) is in equilibrium. That is not the definition you will find in most books. Here is the more common definition:

Definition (take 2)

An isolated system is in *equilibrium* if the probability distribution of its microstates does not change with time.

It can be proven that if an isolated system ever satisfies the PoEapP, even for a moment, it will then continue to satisfy it forever after. So what I said before was certainly true: an isolated system that satisfies the PoEapP is in equilibrium by either definition. Furthermore, when using ensemble averages, the PoEapP is usually true by definition. To perform any calculation you must first specify what statistical ensemble you are using, and ensembles nearly always stipulate that macroscopically indistinguishable microstates have equal probabilities. (Whether an ensemble is a good description of a particular physical system is a different question, of course, one which must be answered by experiment.)

But if you use time averages, the situation is more complicated. A system is initially in some particular microstate. It will then proceed through a series of other microstates as time passes, but there is no reason it *must* pass through every microstate that is macroscopically indistinguishable from the original one. If the system is not ergodic, it might only pass through a subset of them, never going into others. Its probability distribution would not satisfy the PoEapP, but it would still be constant with time.

In this book we will usually work with ensemble averages. Unless I specifically say otherwise, you should always assume that probabilities are defined by a statistical ensemble, not by an average over time.

2.5. The Maxwell-Boltzmann Distribution

There is only so much to say about isolated systems. Real systems are almost never isolated. They are embedded in some sort of environment, and the interaction with that environment is responsible for much of their complexity.

We can use a simple trick to extend our analysis to non-isolated systems. Begin with an isolated system, then split it into two parts. Call them A and B. A is the part we really care about, the thing we want to do experiments on. B is the environment it is connected to. We only care about B to the extent that it affects A. It is called a *heat bath*.

Here are some examples of the sort of thing I mean:

- Part A is the gas contained in a box. Part B is the box itself, along with the whole room the box is sitting in.
- Part A is a test tube with chemicals in it. Part B is a water bath the test tube is sitting in.
- The whole system is the air in a room. Part A is the carbon dioxide molecules in the air. Part B is everything else (the nitrogen, oxygen, and other trace gasses).

The energy of the system can be decomposed as

$ET=EA+EB+EAB$ET=EA+EB+EAB

The total energy of the system is the sum of three terms: one that depends only on the degrees of freedom that make up A, one that depends only on the degrees of freedom that make up B, and one that depends on both parts of the system.

We now make a series of assumptions.

Assumption

EABEAB is small enough that we can ignore it and write

(1)

$$E_T \approx E_A + E_B$$

This is a somewhat odd assumption. If E_{AB} were really zero, the two subsystems would not interact at all, and we would just have two independent isolated systems. Obviously that is not what we want. But we do want them to be *weakly coupled*. E_{AB} should be nonzero, but still much smaller than either E_A or E_B. Actually, what we really care about is that it is much smaller than the *variations* in the energies of A and B. If the energy of A increases, we assume that energy has primarily come out of B, not just from a reduction in the interaction energy between A and B.

Assumption

The degrees of freedom of A and B are specified independently so the density of states factorizes:

(2)

$$\Omega_T = \Omega_A \Omega_B$$

This is another aspect of requiring the subsystems to be weakly coupled. The first assumption restricted them from being coupled through the energy function. This one restricts them from being coupled through the definitions of their degrees of freedom. We must be free to choose a state for A, and then independently to choose a state for B. One must not restrict the other.

Assumption

The log of the density of states of B can be approximated as linear in energy:

(3)

$$\log(\Omega_B(E_B)) \approx \alpha + \beta E_B$$

Any function can be approximated as linear over sufficiently small intervals (aside from pathological cases like singularities and discontinuities). So essentially we are assuming that we only care about a small range of values for E_B. In practice, what this really means is that B must be much larger than A. However much the energy of A fluctuates, it must only have a very small effect on B. A is a small test tube while B is a large water bath. A is a cup of coffee while B is the surrounding room.

We now want to answer the following question: if the total energy of the system is ETET, what is the probability of A being in *one particular microstate* whose energy is EAEA?

We can reason this out in a series of steps, using each of our assumptions in turn.

1. The complete system can be in ΩT(ET)ΩT(ET) possible microstates. By the PoEapP, every one of them is equally probable. Some of those microstates involve A being in the desired microstate and others do not. We therefore conclude:

   The probability of A being in the desired microstate equals the fraction of microstates of the whole system for which A is in that microstate.

2. By the first assumption above, the energy of B is EB=ET−EAEB=ET−EA.
3. By the second assumption, ΩBΩB is completely independent of what microstate A is in, and depends only on EBEB. Therefore, the number of microstates of the whole system for which A is in the desired microstate is simply equal to ΩB(EB)=ΩB(ET−EA)ΩB(EB)=ΩB(ET−EA).
4. By the third assumption, ΩB(EB)=eα+βEBΩB(EB)=eα+βEB.

Combining these results, we find the probability of A being in the desired microstate is

p(EA)∝ΩB(ET−EA)∝e−βEAp(EA)∝ΩB(ET−EA)∝e−βEA

It is conventional to write this in a slightly different form by defining

(4)

1kT≡β≡∂log(ΩB(E))∂E1kT≡β≡∂log(ΩB(E))∂E

TT is known as the *temperature* and kk is *Boltzmann's constant*, which equals $1.3806488\cdot10^{-23}$ Joules/Kelvin. ββ is called the *inverse temperature*. With this definition, the probability can be written as

p(EA)∝e−EA/kTp(EA)∝e−EA/kT

This last step is quite a strange one. For no obvious reason, we have just introduced a constant with a seemingly arbitrary value, along with a completely new set of units. (Just what are "Kelvins"?) Of course, you probably have already figured out why. The concept of "temperature" was established long before statistical mechanics was developed, and this definition is needed to make the statistical definition match the pre-existing one. We will examine the correspondence in Chapter 4. For the moment, though, just think of it as an arbitrary mathematical definition.

We have almost answered our question. The only thing still missing is the proportionality constant. That is easily found: we just require that the probabilities of all microstates add to 1. (The system is certain to be in *some* state, after all.) The normalization constant is therefore given by

(5)

$$Z=\sum e^{-E_A/kT}$$

where the sum is taken over all microstates of A. Z is called the *partition function*. Aside from being a normalization constant, it turns out to be an interesting function in its own right with some useful properties. We will see more of it later.

We can now give the probability for A to be in the desired microstate:

(6)

$$p(E_A)=\frac{e^{-E_A/kT}}{Z}$$

This is called the *Maxwell-Boltzmann distribution*, and it is probably the single most important equation in this chapter (or possibly even in this entire book). Maxwell originally derived it in 1860 based on a mechanical model of gas molecules. In the following years it was repeatedly re-derived based on a variety of arguments that extended its generality. As you have seen, it is not in any way specific to gas molecules or classical mechanics. It is valid for any system that satisfies a very general set of assumptions.

2.6. Thermodynamic Forces

In the last section we assumed the only macroscopic variable we cared about was energy. Let us now extend this to more general cases. As a concrete example, assume we have two macroscopic variables: energy and volume. Perhaps we are dealing with a balloon filled with helium, so it can stretch and contract, exchanging volume as well as energy with the surrounding air. The density of states is now a function of both variables, $\Omega(E,V)$.

We can repeat the exact same argument as in the last section, simply replacing EE with VV. This time we assume $\log(\Omega B)\log(\Omega B)$ is linear in both variables:

(7)

$$\log(\Omega B(EB,VB))\approx\alpha+\beta EB+\gamma VB\log(\Omega B(EB,VB))\approx\alpha+\beta EB+\gamma VB$$

from which we conclude that the probability of a microstate is given by

$$p(EA,VA)\propto\Omega B(ET-EA,VT-VA)\propto e-\beta EA-\gamma VAp(EA,VA)\propto\Omega B(ET-EA,VT-VA)\propto e-\beta EA-\gamma VA$$

Once again it is conventional to write this in a slightly different form by defining a new quantity:

(8)

$$P\equiv kT\partial\log(\Omega B(E,V))\partial VP\equiv kT\partial\log(\Omega B(E,V))\partial V$$

PP is called the *pressure*. Using this definition, the probability for A to be in a particular microstate is

(9)

$$p(EA,VA)=e-(EA+PVA)/kTZp(EA,VA)=e-(EA+PVA)/kTZ$$

where the partition function now equals

(10)

$$Z=\sum e-(EA+PVA)/kTZ=\sum e-(EA+PVA)/kT$$

There is nothing special about volume. The same calculation can be done for any macroscopic variable, producing an identical result. One other example that is especially important in thermodynamics is NN, the number of particles in the system. Perhaps we are studying a box filled with gas, but the box has a small hole in it allowing molecules to diffuse in and out. Rather than defining subsystem A to be particular set of molecules (whatever volume of space they occupy), we instead define it to be a particular volume of space (whatever molecules it happens to contain at any moment). We then define

(11)

$$\mu \equiv -kT \frac{\partial \log(\Omega_B(E,N))}{\partial N}$$

μ is called the *chemical potential*. The negative sign in front of it is just a matter of convention. The probability of a microstate is

(12)

$$p(E_A,N_A) = \frac{e^{-(E_A-\mu N_A)/kT}}{Z}$$

Quantities like P and μ are called *thermodynamic forces*. Each one is said to be *conjugate* to the macroscopic variable we differentiated with respect to. Together, the macroscopic variable and the thermodynamic force (V and P, or N and μ) form a *conjugate pair*.

“Thermodynamic force” is another very suggestive name. How do they relate to forces of the more conventional sort? Do they act to produce accelerations? Are they derivatives of potential functions? We will examine these questions in Chapter 4. As with everything else in this chapter, just treat them as arbitrary mathematical definitions for now.

Having said that, I now need to indulge in a brief rant. μ is a “thermodynamic force”, but it is also called the “chemical potential”. So is it a force, or is it a potential? They are not the same thing! Just to make matters worse, we will soon encounter another type of quantity called a “thermodynamic potential” (of which μ is *not* an example). Could we at least use consistent terminology? Sadly, the answer is no, we cannot. These names were established long ago, and now it is impossible to change them, even when they clearly do not make sense.

2.8. Thermodynamic Potentials

We now know how to compute the probability of finding a system in lots of different kinds of states: microstates or macrostates, specified by arbitrary sets of macroscopic variables. In every case, the probability takes exactly the same form:

(17)

$$p = \frac{e^{-\Phi/kT}}{Z}$$

where the only difference is the quantity Φ appearing in the exponent. This suggests the idea of *thermodynamic potentials*, energy-like functions that capture the differences between different probability distributions. Several of the most common thermodynamic potentials have special names:

(18)

$$\begin{aligned} H &= E + PV && \text{(Enthalpy)} \\ A &= E - TS && \text{(Helmholtz free energy)} \\ G &= E + PV - TS && \text{(Gibbs free energy)} \\ \Phi_G &= E - \mu N - TS && \text{(Grand potential)} \end{aligned}$$

The term *free energy* can also be used more generally to refer to any thermodynamic potential that describes the probabilities of macrostates (that is, any thermodynamic potential that includes a TS term). The Helmholtz free energy, the Gibbs free energy, and the grand potential are all examples of free energies.

You can think about the differences between thermodynamic potentials in two equivalent ways. First, you can think of starting with E and then adding in terms based on the ensemble you want to use. You want volume to be variable? Then add PV. You want to work with macrostates instead of microstates? Subtract TS. And so on.

Alternatively, you can think of all these potentials as special cases of a single potential that includes all possible terms. If the volume is held fixed, then PV is a constant and can be ignored. It just changes the proportionality constant, which gets normalized away when we require the probabilities to add to 1. If the number of particles is fixed, then μN is similarly a constant and can be ignored. If you are working with microstates then TS is zero. (Think of a microstate as being a tiny macrostate with exactly one microstate, so Ω is 1 and S is 0).

Just as there are special names for common thermodynamic potentials, some of the corresponding ensembles also have special names. These names are purely historical. They do not have any particular meaning, but they are still widely used, so you will need to know them.

The *microcanonical ensemble* refers to an isolated system whose energy is constant. It has equal probability of being in any microstate with the specified energy, and of course no chance at all of being in any microstate with a different energy.

The *canonical ensemble* refers to a system that can exchange energy with a heat bath at a specified temperature. The corresponding thermodynamic potential is E for microstates, or $E-TS$ for macrostates.

The *grand canonical ensemble* refers to a system that can exchange both energy and particles with a heat bath of specified temperature and chemical potential. The corresponding thermodynamic potential is $E-\mu N$ for microstates, or $E-\mu N-TS$ for macrostates.

2.9. Averages

The average (or *mean*) of a quantity x is defined as

(19)

$$\langle x\rangle \equiv \sum_i x_i p_i$$

where x_i is its value in the i'th state, and p_i is the probability of that state. The sum can be over either microstates or (if x is a macroscopic variable) macrostates. This is an ensemble average. Its value depends on the probabilities of the system being in different states; or to say that another way, it depends on what ensemble we are calculating the average for. If Φ is the thermodynamic potential for the ensemble, it equals

(20)

$$\langle x\rangle = \frac{\sum_i x_i e^{-\Phi_i/kT}}{\sum_i e^{-\Phi_i/kT}}$$

Two important identities follow directly from the above definition:

(21)

$$\langle x+y\rangle = \sum_i (x_i+y_i)p_i = \sum_i x_i p_i + \sum_i y_i p_i = \langle x\rangle + \langle y\rangle$$

and, if C is a constant,

(22)

$$\langle Cx\rangle = \sum_i C x_i p_i = C\sum_i x_i p_i = C\langle x\rangle$$

Just because x has a particular average value, that does not mean it is always exactly equal to that. Sometimes it is more and sometimes it is less. It can be very useful to know how much a quantity tends to vary about its average. Does it stay within a narrow range, or does it vary widely? A useful measure of this is its *variance*, defined as

(23)

$$\mathrm{Var}(x) \equiv \langle (x - \langle x\rangle)^2 \rangle$$

Remembering that $\langle x\rangle$ is a constant, we can derive a useful identity for the variance:

(24)

$$\mathrm{Var}(x) = \langle x^2 - 2x\langle x\rangle + \langle x\rangle^2 \rangle = \langle x^2\rangle - 2\langle x\rangle\langle x\rangle + \langle x\rangle^2 = \langle x^2\rangle - \langle x\rangle^2$$

Another common measure of how much a value tends to vary is its *standard deviation*, which is simply the square root of the variance. It is represented by the symbol σ. A good rule of thumb is that about 2/3 of the time, the value will be between $\langle x\rangle - \sigma$ and $\langle x\rangle + \sigma$. The exact fraction depends on the probability distribution, of course. We will examine this further in the next chapter.

I mentioned before that the partition function has some interesting and useful properties. One of them is that derivatives of $\log(Z)$ tend to give averages. Remember that the partition function is defined as

$$Z = \sum e^{-\beta\Phi} = \sum e^{-\Phi/kT}$$

For example,

(25)

$$-\frac{\partial \log(Z)}{\partial\beta} = -\frac{1}{Z}\frac{\partial Z}{\partial\beta} = -\frac{1}{Z}\sum\frac{\partial e^{-\beta\Phi}}{\partial\beta} = \frac{1}{Z}\sum \Phi e^{-\beta\Phi} = \langle\Phi\rangle$$

Another useful case is to take the derivative with respect to a state variable (either a microscopic or a macroscopic one).

(26)

$$-kT\frac{\partial \log(Z)}{\partial x} = -\frac{kT}{Z}\frac{\partial Z}{\partial x} = -\frac{kT}{Z}\sum\frac{\partial e^{-\Phi/kT}}{\partial x} = \frac{1}{Z}\sum\frac{\partial\Phi}{\partial x}e^{-\Phi/kT} = \left\langle\frac{\partial\Phi}{\partial x}\right\rangle$$

2.10. Quantum Statistical Mechanics

For simplicity, I will mostly rely on classical mechanics in this book. But nearly everything I say applies equally well to quantum mechanics.

For a quantum system, a microstate simply means a value of the wave function (or, if we need to discretize a continuum of states, a tiny volume of Hilbert space). Of course you can describe the wave function using any set of basis functions you want. The “microscopic variables” of the system are just the amplitudes of the basis functions. If they happen to be position eigenstates, then the microscopic variables are the values of the wave function at each point in pace. But you can just as easily use momentum eigenstates, energy eigenstates, or any other basis you choose.

A “macroscopic variable”, on the other hand, is defined as the expectation value of an operator. If Y is the operator corresponding to some measurable quantity y and the system is in microstate $|\Psi\rangle$, then

(27)

$$y \equiv \langle \Psi | Y | \Psi \rangle$$

This always has a well defined value, even if $|\Psi\rangle$ is not an eigenstate of the operator Y. A particularly important case is energy, which is the expectation value of the Hamiltonian:

(28)

$$E \equiv \langle \Psi | H | \Psi \rangle$$

As long as the system remains isolated, its energy is constant. Conservation of energy applies just as well to quantum mechanics as to classical mechanics.

When dealing with quantum systems, we need to be careful to distinguish between different types of probability. In statistical mechanics, probabilities always refer to either ensemble averages or time averages. The “probability” of a variable having a particular value refers to either a fraction of the members of an ensemble, or to a fraction of time. But quantum mechanics also has its own probabilities that apply even when a system is in a single known state. They describe the probability that a measurement will produce a certain result, *given* that the system is in a particular state.

The probabilistic features of quantum mechanics *only* come up when you perform a measurement, which is to say, when you let the system interact with an external measuring device. As long as the system stays isolated, quantum mechanics is fully deterministic. The system is always in a well defined state, and every microscopic and macroscopic variable has a single well defined value at every moment in time.

It is even possible that the probabilisitic features of quantum mechanics are *also* statistical in nature. When you allow a system to interact with an external measuring device, that will necessarily introduce noise into the system. The state of the system is no longer definitely known, because it is subject to unknown forces. Is it surprising, then, that we cannot predict the result with complete certainty?

This is a large subject of its own, and a very controversial one. It can be proven that *if* quantum mechanics is statistical in nature, it must necessarily possess one or more unintuitive properties such as nonlocality or retrocausality. But quantum mechanics is already one of the most unintuitive physical theories ever developed, so that is hardly an argument one way or the other! Some physicists would even argue that we already have good evidence for both nonlocality *and* retrocausality coming from completely unrelated directions. If so, then statistical interpretations of quantum mechanics might well be among the very simplest and most intuitive ones.

In any case, when applying statistical mechanics to quantum systems, be sure to distinguish the “probabilities” due to statistical ensembles from the “probabilities” due to quantum mechanics itself. At any time other than when you are actually in the middle of making a measurement, the former ones are the only kind that apply.

3. Mathematical Interlude: Very Large Numbers

Problems in statistical mechanics tend to involve large numbers. Very *very* large numbers. For example, we often need to average over all the atoms or molecules in a system. For any macroscopic system, that will be on the order of Avogadro’s number: 10^{23}, plus or minus a few orders of magnitude. In other cases we average over microstates, the number of which usually grows *exponentially* in the number of atoms. These are really astronomically large numbers.

This has an interesting result. In many cases, we can simply treat averages as exact numbers. Any variation about the average value is so small, we can completely ignore it. This is a great simplification, and is one of the reasons statistical mechanics is so successful.

It also is the reason that, despite its name, statistical mechanics involves very little real statistics. A large part of the field of statistics deals with the variations about averages: computing the probability of a distribution producing values in a certain range, determining whether a measured value might plausibly have come from a certain distribution, etc. But in statistical mechanics, we can usually ignore these questions. The chance of measuring any value *except* the average one is negligible.

The main goal of this chapter is to demonstrate this fact. We will do that first through a simple example, and then through a very general theorem.

3.1. The Binomial Distribution

Do you ever worry that all the oxygen molecules in the air around you might spontaneously migrate to the other side of the room, leaving you unable to breathe? It could happen. Each molecule diffuses independently, so there is a small but finite probability that just by chance, they will all end up in the same half of the room at the same time.

I hope I am not worrying you. While this is theoretically possible, it is so absurdly improbable that you really can ignore it. We will first compute just how improbable it is, and then study the statistics of this problem in more detail.

Let NN be the number of oxygen molecules in the room. We will describe the position of each one simply by whether it is in the left or right half of the room. That gives 2N2N possible arrangements of molecules, every one of which is equally probable. Of all those arrangements, only one has all the oxygen in the opposite half of the room from you, so the probability is 1/2N1/2N.

Depending on the size of the room you are in, it probably has somewhere in the vicinity of 10^{26} oxygen molecules in it. So the probability of all the oxygen being on the same side of the room at any given moment is

121026≈11010000000000000000000000000121026≈1101000000000000000000000000
0

That is a very, very large denominator. For comparison here are some much smaller numbers:

- The age of the universe is on the order of 10^{17} seconds.
- There are somewhere around 10^{80} atoms in the visible universe.

In short, you really do not need to worry about this happening. You would be much better off worrying about more likely events. For example, that you will be struck by lightning and then, as you try to pick yourself back up, hit by a meteor falling from space. And then immediately be struck by lightning a second time just for good measure. That is the sort of thing that is far more likely to happen to you.

Perhaps you are still not reassured. You may have realized that even if only 90% of the oxygen moved to the other side of the room, that would still leave you in a rather awkward position. We really want to work out the entire probability distribution. Let m be the number of oxygen molecules in the left half of the room and $N-m$ the number in the right half. Assuming the energy of each molecule is independent of which side of the room it is on, and that the oxygen molecules are far enough apart that we can ignore their interactions (both reasonable assumptions), the probability is simply proportional to $\Omega(m)$. Let us consider how that varies with m.

- There is precisely one arrangement of molecules that has all of them on the left side of the room, so $\Omega(0)=1$.
- There are N possible ways to have exactly one molecule on the left side of the room (any of the N molecules could be the one), so $\Omega(1)=N$.
- To get two molecules on the left side of the room, we have N choices for the first molecule and $N-1$ for the second molecule. But now we have counted each state twice: for any pair of molecules, it does not matter which order we choose them in. So $\Omega(2)=N(N-1)/2$.

Continuing like this, we find the general form:

(1)

$$\Omega(m)=\frac{N!}{m!(N-m)!}$$

The probability is then

(2)

$$p(m)=\frac{1}{2^N}\frac{N!}{m!(N-m)!}$$

This is a special case of the *binomial distribution*. (In the general case, the probability for a particle to be on the left side of the room could be different from its probability to be on the right side, but for this example we can ignore that.) It is shown in Figure 3-1 for several values of N. Notice how, as N increases, the probability becomes steadily more concentrated around its midpoint.

Figure 3-1. The binomial distribution for several values of N. To make them easier to compare, all the curves have been normalized to go from 0 to 1 along each axis.

The name comes from the fact that it has the same form as the terms in the *binomial series*:

(3)

$$(a+b)^N=\sum_{m=0}^{N}\frac{N!}{m!(N-m)!}a^N b^{N-m}$$

If we let $a=b=1/2$, this simplifies to

$$\left(\frac{1}{2}+\frac{1}{2}\right)^N=\sum_{m=0}^{N}\frac{N!}{m!(N-m)!}\frac{1}{2^N}$$

so that

(4)

$$\sum_{m=0}^{N}\frac{N!}{m!(N-m)!}=2^N$$

Let us examine this probability distribution a bit further. We can roughly characterize it by calculating its mean and standard deviation. The mean is easy: every particle has an equal chance to be on each side of the room, so $\langle m\rangle=N/2$. Another way to think about this is that the distribution is symmetric: for any value of m, $p(m)=p(N-m)$, from which it directly follows that the mean must be $N/2$.

The standard deviation is a bit more complicated to derive. As a first step, let us first calculate

$$\langle m(m-1)\rangle=\frac{1}{2^N}\sum_{m=0}^{N}m(m-1)\frac{N!}{m!(N-m)!}$$

Notice that the first two terms of the sum are both zero, so we can increase the lower bound to begin from 2. This allows us to then cancel out the factors of $m(m-1)$ that appear in both numerator and denominator:

$$\langle m(m-1)\rangle=\frac{1}{2^N}\sum_{m=2}^{N}\frac{N!}{(m-2)!(N-m)!}=\frac{N(N-1)}{2^N}\sum_{m=2}^{N}\frac{(N-2)!}{(m-2)!(N-m)!}$$

Now make two substitions: define $X=N-2$ and $y=m-2$. This simplifies it to:

$$\langle m(m-1)\rangle=\frac{N(N-1)}{2^N}\sum_{y=0}^{X}\frac{X!}{y!(X-y)!}$$

We immediately recognize the sum as being the same one that appeared in equation (4). Replacing it by 2X=2N−22X=2N−2,

⟨m(m−1)⟩=N(N−1)2N−22N=N(N−1)4⟨m(m−1)⟩=N(N−1)2N−22N=N(N−1)4

We now have all the pieces we need. Recall that

(5)

Var(m)=⟨m2⟩−⟨m⟩2=⟨m(m−1)⟩+⟨m⟩−⟨m⟩2=N(N−1)4+N2−N24=N4Var(m)=⟨m2⟩−⟨m⟩2=⟨m(m−1)⟩+⟨m⟩−⟨m⟩2=N(N−1)4+N2−N24=N4

The standard deviation is then

(6)

σ≡Var(m)−−−−−−√=N−−√2σ≡Var(m)=N2

Let us take a moment to consider these results. The average value is proportional to NN, while the standard deviation is proportional to N−−√N. In most cases, what we really care about is the ratio of the two. You want to know what *fraction* of the oxygen is likely to be on one side of the room, not the specific number of molecules. A number that would be huge in a broom closet would be negligible in an auditorium. Taking the ratio gives

(7)

σ⟨m⟩=1N−−√σ⟨m⟩=1N

For an average sized room with 10^{26} oxygen molecules, that equals 10^{-13}. This is the magnitude of the typical fluctuations, measured as a fraction of the total number of molecules. There are very few physical quantities that can actually be measured to a precision of 13 significant digits, and this is not one of them. The random fluctuations in the number of molecules on each side of the room are simply too small to measure.

3.2. The Central Limit Theorem

Having worked through one example in detail, you might wonder how general our conclusions are. Do other types of statistical problems behave roughly the same way? The answer is yes. There is a very general theorem which guarantees that a wide range of quantities will scale in essentially the same way.

Before presenting that theorem, I first need to introduce one other very important probability distribution, known as the *Gaussian* or *normal* distribution:

(8)

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

It is shown in Figure 3-2. It is not really as complicated as it looks. The basic form is just e^{-x^2}. Replacing x by $x-\mu$ shifts the mean of the distribution to μ, and then the exponent is scaled to make the standard deviation equal σ. The factor in front is just a normalization, to ensure that

$$\int_{-\infty}^{\infty}p(x)dx=1$$

Figure 3-2. A normal distribution with mean of 0 and standard deviation of 1.

The name “normal distribution” gives a sense of just how important this distribution is in statistics. It comes up constantly. We will see the reason for this in just a moment.

In the last section, we considered the number of oxygen molecules m in the left half of the room. You can think of this quantity as a sum of random values:

$$m=\sum_{i=1}^{N}x_i$$

Each value x_i is either 0 or 1 (depending on which half of the room the molecule is in) with equal probability. Now we will generalize this to let the values be drawn from an arbitrary distribution. That distribution could be either discrete or continuous. The only thing we will assume about it is that it has a known mean and standard deviation. There is a remarkable result called the *Central Limit Theorem*:

The Central Limit Theorem

Consider the sum

$$S=\sum_{i=1}^{N}x_i$$

where the values x_i are independently drawn from a distribution with mean μ_x and standard deviation σ_x. In the limit $N\rightarrow\infty$, the sum S is distributed according to a normal distribution with mean $\mu=N\mu_x$ and standard deviation $\sigma=\sqrt{N}\sigma_x$.

This theorem dates back to 1718, when Abraham de Moivre presented a proof of a special case of it in his book *The Doctrine of Chances*. This book was not, as you might guess, an academic treatise for mathematicians. It was a book for gamblers, discussing how to win at various games of chance. Don't ever let anyone tell you that mathematics is not useful!

This is the reason that normal distributions are so important. No matter what distribution you start out with, once you add enough values together the result will always be normally distributed. How large NN needs to be varies, of course, depending on the initial distribution, but it often does not need to be very large. In many cases, summing over 10 values is already enough to give quite a good approximation to a normal distribution.

Notice that our main conclusions from the previous section apply just as well to the general case as they did for our specific example. The standard deviation is proportional to $\sqrt{N}$N, the mean is proportional to NN, and their ratio scales as $1/\sqrt{N}$1/N. No matter what distribution the individual values come from, once you add up a macroscopic number of them the fluctuations will be negligibly small.

4. The Interpretation of Statistical Quantities

In Chapter 2 I defined lots of quantities, but said very little about what they really mean. Now it is time to return to them. We will study them in more detail and try to build up an intuitive understanding of what they represent.

4.1. Temperature

Temperature is a quantity we are all familiar with. You have probably been told that it measures the microscopic jiggling of atoms: the faster they are moving, the hotter an object feels. So when I defined it in terms of a derivative of the density of states, a seemingly unrelated concept, you were probably a little surprised. To resolve this seeming paradox, I am going to prove a very important theorem called the *equipartition theorem*. Once you reach the end, the significance of temperature will become clear.

In general, the energy of a system can depend in an arbitrary way on all of its microscopic variables, which we will call x_1x1, x_2x2, etc. That is, the energy is a function $E(x_1,x_2,\ldots)$E(x1,x2,…). But there are some very important special cases where it takes particular forms. Let us make three assumptions:

1. The energy of the system can be written

   $$E(x_1,x_2,x_3,\ldots)=E_1(x_1)+E_2(x_2,x_3,\ldots)$$

   That is, the energy separates into two terms. The first term depends *only* on x_1, while the second term does not depend on x_1 at all.

2. x_1 can be treated as a continuous variable, and can take on any value from $-\infty$ to ∞.
3. The energy is quadratic in x_1:

   $$E_1(x_1)=Cx_1^2$$

   where C is a constant.

A particularly important example of this situation is the momentum of a classical particle. Each momentum variable p_i contributes $p_i^2/2m$ to the kinetic energy.

Now we want to answer the question, “What is the average value of E_1?” As we learned in section 2.9, that is given by

$$\langle E_1\rangle=\frac{\sum E_1 e^{-\beta E}}{\sum e^{-\beta E}}$$

(For simplicity I am assuming that E is the relevant thermodynamic potential; we are working in the canonical ensemble. This theorem applies just as well to other ensembles, though. Just replace E by Φ in all the equations above.)

Now take advantage of assumption 1. This allows us to factor the above expression:

$$\langle E_1\rangle=\frac{\sum e^{-\beta E_2}\sum E_1 e^{-\beta E_1}}{\sum e^{-\beta E_2}\sum e^{-\beta E_1}}$$

The first sum is identical in both the numerator and denominator, so we can cancel them out:

$$\langle E_1\rangle=\frac{\sum E_1 e^{-\beta E_1}}{\sum e^{-\beta E_1}}=-\frac{\partial \log\left(\sum e^{-\beta E_1}\right)}{\partial \beta}$$

Assumption 2 allows us to replace the sum by an integral:

$$\langle E_1\rangle=-\frac{\partial \log\left(\int_{-\infty}^{\infty} e^{-\beta E_1}dx_1\right)}{\partial \beta}$$

Finally, we make use of assumption 3 and introduce the change of variables $y \equiv \sqrt{\beta}\, x_1$:

(1)

$$\langle E_1\rangle = -\frac{\partial \log\left(\int_{-\infty}^{\infty} e^{-\beta C x_1^2}\, dx_1\right)}{\partial\beta} = -\frac{\partial \log\left(\beta^{-1/2}\int_{-\infty}^{\infty} e^{-Cy^2}\, dy\right)}{\partial\beta} = -\frac{\partial\left(\log(\beta^{-1/2}) + \log\left(\int_{-\infty}^{\infty} e^{-Cy^2}\, dy\right)\right)}{\partial\beta} = \frac{1}{2\beta} = \frac{kT}{2}$$

We now see how the statistical definition of temperature matches up with the informal one you are familiar with. Temperature is a measure of average kinetic energy. To put it precisely:

Interpretation of Temperature

If a system is in equilibrium with a heat bath at temperature T, then its average kinetic energy is $kT/2$ per degree of freedom.

Of course, the equipartition theorem applies equally well to any variable that satisfies the three requirements, not just to momentum variables. In a collection of uncoupled harmonic oscillators, the average *potential* energy will also be $kT/2$ per degree of freedom.

Even more generally, when a system is in equilibrium, its total energy (not just kinetic energy) tends to be evenly distributed among its degrees of freedom. This is just a rule of thumb, not a rigorous result like the equipartition theorem. Still, it is a useful guideline for how to think about systems in equilibrium.

Example

What is the average kinetic energy of a gas of N rigid diatomic molecules at temperature T?

Each molecule is composed of two atoms, so it would normally have 6 degrees of freedom. But because they are rigid (the distance between the two atoms cannot change), that removes one degree of freedom. The average kinetic energy is therefore

$$\langle E_{kinetic}\rangle = 5N \cdot \frac{kT}{2}$$

4.2. Thermodynamic Potentials as a Measure of Probability

Thermodynamic potentials are confusing. Not only are they an unfamiliar concept, but as soon as you start trying to do calculations with them, it is very easy to get all tangled up in a net of subtle mistakes. In the following sections, I will approach thermodynamic potentials from several different directions. I will point out some of the particular issues that make them confusing, and try to help you form a clear understanding of the essential concepts.

The first thing that makes them confusing is that they mix together quantities related to two different systems. Recall that we started with a single isolated system, then split it into two pieces called A and B. A is the system of interest, the thing we really care about. B is the heat bath. Consider the Gibbs free energy as an example. Let me restate the definition, including a subscript on every variable to indicate which one it refers to:

$$GA=EA+PBVA-TBSA$$

It is very easy to get confused about which system each variable relates to. For example, it is easy to see the term TS and remember that T and S are each defined in terms of $\log(\Omega)$, but forget that their definitions refer to completely different Ωs. T is related to the density of states of the heat bath, while S measures the density of states of the system of interest.

A thermodynamic potential is, first and foremost, a measure of how probable it is for a system to be in a particular state. It appears in the exponent of the Maxwell-Boltzmann distribution. Lower values of the potential indicate more likely states, while higher values indicate less likely states.

Let’s go through each term in the Gibbs free energy, and consider why each one should influence the probability.

The first term is simply the energy of the system of interest. All else being equal, states with lower energy are more likely to occur. Why? Because the energy of the complete system A+B is conserved. Less energy in A means more energy in the heat bath, and that means its density of states is larger. (This was one of our assumptions in deriving the Maxwell-Boltzmann distribution. Remember? We assumed $\log(\Omega B)$ was linear in energy.)

The second term is $PBVA$PBVA. This term can be interpeted in exactly the same way, only in terms of volume instead of energy. The total volume of A+B is fixed. Less volume for A means more volume for B, and we assumed the density of states of B increased with volume. Microstates of A with lower volume are therefore consistent with more microstates of B, so they are more likely to occur.

That is all fine as far as microstates are concerned, but the Gibbs free energy describes the probability of macrostates. Lower energy and volume for A correspond to more probable microstates, but there also are likely to be fewer of them. ΩAΩA also increases with energy and volume just like ΩBΩB. That is what the final term $TBSA$TBSA represents. It decreases the potential for macrostates whose entropy (and therefore density of states) is larger, thus making those macrostates more probable. The actual probability of a given macrostate is determined by a balance between all three terms. Increasing the energy and volume of system A increases the first two terms while decreasing the third term. It decreases the probability of each microstate while increasing the total number of microstates.

If we were working with the grand potential instead of the Gibbs free energy, we would have a different term: $-\mu BNA$−μBNA. The interpretation of this term is exactly like the other ones. The total number of particles in A+B is fixed. Fewer particles in A means more in B, which increases its density of states. (Do not be confused by the negative sign in front of this term. That is just to cancel out the extra negative sign that, for purely historical reasons, appears in the definition of μμ.)

4.3. Independent and Dependent Variables

There is a second way that thermodynamic potentials can be interpreted. It assigns a completely different meaning to them, while still being similar enough to be confusing. It has to do with which macroscopic variables we treat as independent and which ones we treat as dependent.

So far, we have always used macroscopic variables such as energy, volume, and particle number. These have the important property that they are also microscopic variables. Every microstate has a well defined energy, a well defined volume, a well defined number of particles. A macrostate simply consists of all the microstates with the specified values for its variables.

Contrast that with variables such as temperature, pressure, and chemical potential. These are *not* microscopic variables. They are defined in terms of the density of states. It makes no sense to speak of the temperature or pressure of a single microstate. They relate to the distribution of microstates, not to the properties of any single state.

On the other hand, from a physical perspective it makes perfect sense to think of these quantities as independent macroscopic variables. Consider these examples:

- You can keep a system isolated so that its energy remains constant, *or* you can put it in contact with a heat bath of fixed temperature. In the former case, energy is the independent variable. It is the quantity you control, and it has a well defined value. In the latter case, it makes more sense to view temperature as the independent variable. That is now the quantity you control. The energy no longer has a single fixed value. It fluctuates continuously as the system exchanges energy with the heat bath. You can, however, calculate an *average* energy $\langle E\rangle$⟨E⟩. Furthermore, as we saw in the last chapter, for macroscopic systems the fluctuations about the average will be completely negligible. In that case, it is reasonable to treat energy as if its value were fixed. But it has now become a *dependent* variable $E(T)$E(T).
- You can fix the volume of a system, *or* you can put it in contact with a heat bath of fixed pressure. In the former case, volume is the independent variable. In the latter case, pressure is the independent variable. The volume now fluctuates as it interacts with the heat bath. For a sufficiently large system, these fluctuations will be negligible. We can therefore treat the volume as a dependent variable with a well defined value $V(P)$V(P).
- Finally, you can fix the number of particles in the system, or you can let it exchange particles with a heat bath of fixed chemical potential. In the former case, NN is an independent variable. In the latter case, it is a dependent variable $N(\mu)$N(μ).

Just to make matters more confusing, the *mathematical* choice of what variables to treat as independent need not match the *physical* choice of what quantities to control. Consider a system in contact with a heat bath of specified temperature and pressure, so that Gibbs free energy is the relevant thermodynamic potential. Physically speaking, TT and PP are the quantities you control, while EE and VV vary in response to them.

Nonetheless, you can still take EE and VV as the macroscopic variables defining a macrostate. In that case, every macrostate has a strictly fixed energy and volume. It is defined to consist of all microstates with those values of EE and VV. On the other hand, the system is no longer in a single, well defined macrostate. It continuously fluctuates through a (usually tiny) range of macrostates as it exchanges energy and volume with the heat bath.

Alternatively, you can choose to treat TT and PP as the macroscopic variables. In that case, the system is in a single, fixed macrostate. On the other hand, that macrostate no longer has fixed values for EE and VV. Only their averages $\langle E\rangle$⟨E⟩ and $\langle V\rangle$⟨V⟩ are well defined. Furthermore, there is no longer a direct mapping from microstates to macrostates! Each macrostate represents a *probability distribution* over microstates, and each microstate has a probability of occurring while the system is in a range of different macrostates.

These two choices of macroscopic variables lead directly to two different interpretations of the thermodynamic potential:

1. If each macrostate has a well defined value for each macroscopic variable and consists of a fixed set of microstates, then the thermodynamic potential also has a well defined value. It describes the probability for the system to be in that macrostate.
2. If we instead treat each macrostate as a probability distribution over microstates, then the thermodynamic potential no longer has a strictly fixed value. Instead we must view it as an average quantity. For example,

   $G=\langle E\rangle+P\langle V\rangle-T\langle S\rangle$G=⟨E⟩+P⟨V⟩−T⟨S⟩

Both of these interpretations are widely used. Furthermore, people often shift fluidly back and forth between them without giving any indication they have just changed their definitions. When in doubt, look carefully to see which macroscopic quantities are being treated as independent variables.

4.4. Thermodynamic Potentials and Thermodynamic Forces

According to classical mechanics, if the potential energy UU of a system depends on a variable xx, the system experiences a *generalized force*

(2)

$Q=-\partial U\partial x$Q=−∂U∂x

When the value of xx changes by an amount ΔxΔx, the system performs *work* equal to

(3)

$$W=Q\Delta x$$

Since we have defined quantities called “thermodynamic potentials” and “thermodynamic forces”, you may be wondering how closely they are related to potentials and forces of the conventional, non-thermodynamic kind. Does the derivative of a thermodynamic potential give a thermodynamic force? Does a thermodynamic force perform work?

I need to be very careful how I answer these questions. All our results so far assume the system is in equilibrium, and if large changes are happening in the macroscopic variables, it clearly is *not* in equilibrium. Until we learn how to deal with situations of this sort, I need to restrict myself to only saying things about systems in equilibrium.

Still, we can at least partly answer these questions now. When a system is in equilibrium, its dependent macroscopic variables take on the values that minimize the thermodynamic potential (and hence maximize the probability). So if we take a derivative of Φ and set it equal to zero, that will provide information about what happens in equilibrium. For example, consider a system in contact with a heat bath of constant temperature and pressure, then take a derivative with respect to volume:

$$\frac{\partial G_A}{\partial V_A}=0=\frac{\partial E_A}{\partial V_A}+P_B-T_B\frac{\partial S_A}{\partial V_A}=\frac{\partial E_A}{\partial V_A}+P_B-kT_B\frac{\partial \log(\Omega_A)}{\partial V_A}=\frac{\partial E_A}{\partial V_A}+P_B-P_A$$

From which we conclude:

(4)

$$-\frac{\partial E_A}{\partial V_A}=P_B-P_A$$

This is a remarkable equation. The left side is an ordinary force of the conventional sort. It is just a derivative of the energy, with nothing statistical about it. Everything we know about forces from classical mechanics can be directly applied to it. The right side, on the other hand, is a thermodynamic force (or rather, the difference between two thermodynamic forces). And this equation says that, when the system is in equilibrium, the two sides must be equal to each other.

If the internal pressure P_A and external pressure P_B are equal, then the equilibrium condition simplifies to just $\frac{\partial E_A}{\partial V_A}=0$: there must be no net mechanical force. If there were, it would cause the volume to change (that being what mechanical forces do). Clearly the pressure is at least *acting* like a mechanical force. For the system to be in equilibrium, the internal pressure, external pressure, and mechanical force must exactly cancel each other out.

So thermodynamic forces *act* like ordinary forces, but does that mean they *are* ordinary forces? Can they produce motion and do work? To answer that, we must examine them more carefully and understand just what is going on when a “thermodynamic force” is applied to a system.

4.5. The Mechanics of Thermodynamic Forces

Consider a balloon filled with gas. It is subject to three different forces: the outward pressure of the gas inside the balloon, the inward pressure of the surrounding air, and the elastic tension of the balloon itself that resists expansion. As seen from equation (4), at equilibrium these three forces exactly balance each other out. If they do not balance, the balloon will expand or contract until they do. But what is actually happening at a microscopic level?

Molecules of gas are constantly striking the surface of the balloon. How do we know that? Because the density of states of the gas increases with increasing volume. If its volume were greater, there would be more microstates available to it. The balloon is restricting it from visiting those microstates, and if the balloon were not there, it would not remain contained in such a restricted volume.

That is what pressure really is: the mechanical force exerted by the gas molecules as they strike against the balloon (or any other object that restricts their motion). So pressure is not merely “like” a force. It *is* a force. It can do all the same things other forces can do, including producing accelerations and performing work.

Chemical potential can be understood in exactly the same way. Imagine a box with a small hole in it, so that air molecules can diffuse in and out. The chemical potential is essentially a measure of the density of air molecules. In equilibrium, we expect the density to be the same inside and outside. If that is not the case, we expect to find a net flow of molecules one way or the other until equilibrium is achieved. On the other hand, if there is another force involved (such as one that repels molecules away from the interior of the box), then we expect to find different densities inside and outside. We can find the equilibrium distribution by looking for the values that minimize the thermodynamic potential.

Here is one more critical observation about thermodynamic forces: as we saw in section 2.6, they are always proportional to the temperature. Given the microscopic description above, this is now easier to understand. For example, pressure is the force of particles randomly striking the walls of a container. The faster they are moving, the harder they strike it. And as we saw in section 4.1, the average velocity of each particle is proportional to the temperature. If the temperature were exactly zero so the particles were not moving at all, there would be no pressure. All thermodynamic forces would disappear, and all thermodynamic potentials would simply become equal to the energy.

4.6. Thermal Equilibrium

In section 4.4 we took the derivative of G with respect to volume, and derived a condition for the system to be in equilibrium. Let’s repeat the same calculation, only instead taking the derivative with respect to energy:

$$\frac{\partial G_A}{\partial E_A} = 0 = 1 - T_B \frac{\partial S_A}{\partial E_A} = 1 - kT_B \frac{\partial \log(\Omega_A)}{\partial E_A} = 1 - \frac{T_B}{T_A}$$

$$\frac{\partial G_A}{\partial E_A} = 0 = 1 - T_B \frac{\partial S_A}{\partial E_A} = 1 - kT_B \frac{\partial \log(\Omega_A)}{\partial E_A} = 1 - \frac{T_B}{T_A}$$

From which we conclude:

(5)

$$T_A = T_B$$

$$T_A = T_B$$

Unsurprisingly, the requirement is that both subsystems must have the same temperature. Two systems whose temperatures are equal are said to be in *thermal equilibrium*. If you bring them into contact with each other, no energy will flow between them. On the other hand, if the systems have different temperatures, there will be a net flow of energy until their temperatures become equal.

This energy transfer is different from the ones seen in the previous sections, in that it does not involve any mechanical work. There is no change to any independent macroscopic variable *other* than energy. It is simply the result of random collisions between molecules that transfer kinetic energy from one subsystem to the other. This type of energy transfer is known as *heat*.

4.7. Intensive and Extensive Variables

Here is a useful bit of terminology.

Macroscopic variables that are independent of the size of the system are called *intensive variables*. Temperature, pressure, and chemical potential are all intensive variables. For example, the temperature of a system has nothing to do with how large that system is.

Macroscopic variables that are proportional to the size of the system are called *extensive variables*. Energy, volume, number of particles, and entropy are all extensive variables.

If you take two independent systems and then view them as a single combined system, all intensive variables will have values that are averages of those for the independent systems, whereas all extensive variables will have values that are the sums of the values for the independent systems. If you combine two identical systems, the resulting system will have twice the volume, twice the energy, and twice the entropy of either of the component systems on its own. But its temperature, pressure, and chemical potential will be identical to those for the component systems.

If you multiply an intensive variable by an extensive one, the result is an extensive variable. Thermodynamic potentials contain many such products. PVPV, TSTS, and μNμN each multiply an intensive variable by an extensive one to produce an extensive variable. Thermodynamic forces are always intensive. Thermodynamic potentials are always extensive.

The ratio of two extensive variables is an intensive variable. For example, dividing the number of particles by the volume produces the *particle density*: $\rho=N/V$ρ=N/V.

Nearly all macroscopic variables fall into one of these two categories. There is nothing especially profound about this. It is just a piece of terminology you will need to know.

4.8. A Brief Rant: Internal Energy

If you read most other books on statistical mechanics, you will find that everywhere I have written EE for energy, they instead write UU for “internal energy”. I have intentionally avoided doing that. At best, internal energy is a useless distinction, and at worst it can be actively misleading. I hesitate even to mention it now. But since you are likely to encounter it sooner or later, I should at least introduce the concept.

The idea is that the total energy of a system can be divided into three parts: the kinetic energy of the system as a whole, the potential energy of the system as a whole, and the "internal energy". Only the last of these affects the internal dynamics of the system. For example, if you put your experimental apparatus on an airplane flying at 1000 km/hour, that greatly increases its kinetic energy, but has no effect on your results. Likewise if you take the apparatus to the top of Mt. Everest, that greatly increases its potential energy but again does not affect your results.

The problem with this idea is that it is simply wrong. Every time the airplane encounters turbulence, the bouncing will add heat to the system, thus affecting your results. The same thing happens to a lesser extent every time it speeds up, slows down, turns, or accelerates in any other way. The only case where the internal energy is fully decoupled from the overall motion of the system is when it moves at a constant speed in a straight line with no acceleration at all. But in that case, the only difference between internal energy and total energy is what reference frame you calculate it in. Relativity tells us that all your results must be independent of what reference frame you use, so the choice of whether to use internal energy or total energy is irrelevant.

Similarly, the gravity at the top of Mt. Everest is slightly weaker than at sea level, and that does have the potential to change your results. In many (but not all!) cases, the difference is negligible, and in that event the only difference between internal energy and total energy is that they are offset by a constant. But the zero point of energy is always arbitrary; adding a constant never affects behavior. So once again, it is irrelevant which one you use.

In summary, the choice to use internal energy instead of total energy *by definition* is only correct if it does not affect any of your results. If it does affect them, that proves you have defined the internal energy incorrectly. But it is very easy to define it incorrectly, such as by neglecting a contribution that actually does matter. And the only way to make sure is to repeat your calculation using the total energy and verify that the results do not change! So why make the distinction in the first place?

5. Thermodynamics

Statistical mechanics grew out of an earlier field called thermodynamics, which was concerned with the thermal properties of liquids and gasses. It grew up around it, and then subsumed it. What we now call “classical thermodynamics” was developed over a period of several hundred years, but much of the most important work was done in just a few decades from the 1820s through the 1850s. It is not at all a coincidence, of course, that this burst of activity coincided with the industrial revolution and the development of the locomotive. Classical thermodynamics was largely developed by people who wanted to learn how to make better steam engines.

Statistical mechanics has come a long way from these humble beginnings, but thermodynamics is still an important field in its own right. In this chapter, I will discuss some of the most important results of classical thermodynamics as seen from a modern statistical viewpoint.

5.1. Thermodynamic Processes

In everything that has come so far, we have considered only systems that were in equilibrium. We will now take a first step toward removing this restriction.

Consider some physical system, and assume it is initially in equilibrium. We now change the system in some way. There are many types of changes we might make. Here are some examples.

- We could directly add energy to the system, for example by shaking it or by placing it over a fire.
- We could bring it into contact with a new heat bath. This might be done by moving it, or by opening a door that previously insulated it and prevented heat from flowing between it and the heat bath.
- We could change the volume of the system. For example, it might be contained in a chamber, one wall of which is attached to a piston allowing it to be moved in or out.
- We could change the potential energy function for the system. For example, we could apply an electric or magnetic field. This will change the energy of each microstate.

Once we make any of these changes, the system will no longer be in equilibrium. If we let it sit long enough, however, it will eventually come back into equilibrium. How long that takes will depend entirely on the system. It could easily be anything from microseconds to millions of years. Assume we do that, so we once again are dealing with a system in equilibrium. This whole process of changing the system and allowing it to equilibrate is known as a *thermodynamic process*. We can now compare the two different states of the system, one before the process was performed and one after.

Notice that we have fundamentally changed our definition of what can constitute a physical system. Up to now we have dealt with two types of systems:

1. An isolated system whose macroscopic properties, like energy and volume, are fixed and unchanging.
2. A system in contact with a heat bath. The properties of the system are determined by the properties of the heat bath. They are free to change, but they do so at random and mostly only within a narrow range. Also, if we view the system of interest plus the heat bath as a single composite system, that composite system is still isolated.

We now have a third option: a system whose properties we can directly control. They change when we choose to change them in exactly the way we choose. We are no longer in any way dealing with an isolated system. We are explicitly making an external change, reaching in from outside to alter the system. On the other hand, we only do that during a restricted time period. Before we make the change, the system (possibly including a heat bath) is isolated. After we are done making the change it is again isolated. But while we make the change, it is not isolated.

5.2. The Laws of Thermodynamics

Classical thermodynamics is based on four laws that are very sensibly numbered starting at zero. When viewed from a modern perspective, three of them are trivial. The remaining one is extremely profound, and even today is often misunderstood.

The Zeroth Law of Thermodynamics

If two bodies are each in thermal equilibrium with a third body, then they are also in thermal equilibrium with each other.

This law was needed to justify that one can define temperature in terms of values measured on a thermometer. Remember that when thermodynamics was first being developed, no one was certain exactly what “temperature” really was. It was strictly an empirical measurement. This law asserts that if you let a thermometer come to equilibrium with two different bodies, and they each produce the same measurement, then bringing those bodies together will produce no net flow of heat.

Once we associate temperature with the average kinetic energy per degree of freedom, this law becomes trivial. It is merely the statement that if two numbers each equal a third number, then they also equal each other.

The First Law of Thermodynamics

The change in a system’s energy during a thermodynamic process equals the heat added to the system, minus the work it performs on its environment. In differential form, this is written

(1)

$$dE=dQ-dW$$

$dE=dQ-dW$

where Q represents heat and W represents work.

This is just a statement of conservation of energy. But once again, remember that the founders of classical thermodynamics did not know exactly what “heat” was. The fact that it was a form of energy, and that it participated in the conservation of energy, had to be discovered through experiment.

Also notice how the form in which it is stated reflects the sorts of problems they were trying to study: mechanical engines that one adds heat to and that do work. In many cases, it is hard to cleanly divide all energy conversions into “heat” and “work”, for example in processes that involve chemical reactions or in which matter is transferred between the system and a heat bath. But of course, energy conservation is a completely general law and applies to all situations equally well.

The next law can be stated in several forms which appear different, but are actually equivalent. Here are two of the most common versions.

The Second Law of Thermodynamics

In any thermodynamic process, the total entropy either increases or remains constant, but never decreases.

or

Heat will never spontaneously flow from a colder body to a warmer one.

This law reflects a deep truth about the nature of the universe we live in. It deserves a whole section of its own, so before I get to that let me first finish off the list of laws.

The Third Law of Thermodynamics

As the temperature of any system approaches zero, its entropy approaches a minimum value.

This is another law that was needed to allow empirical definitions of temperature and entropy. From a statistical viewpoint, it is a statement of the fact that for nearly all physical systems, the density of states increases with energy.

5.3. The Second Law of Thermodynamics

The second law is not a trivial consequence of statistics or mechanics, but at first glance it is easy to *think* it is. The entropy of a system is simply the logarithm of its density of states. For an isolated system, that is a measure of how probable each macrostate is. The second law could therefore be summarized as, “Systems tend to move from less probable states to more probable ones.” That sounds obvious. If a system is initially in a less probable state, we should hardly be surprised to later find it has moved to a more probable one. If it went the other way and moved to a less probable state, *that* would be surprising.

Suppose you have a box filled with gas molecules. Now you remove the cover, so the molecules are free to leave the box. Clearly it is far more likely for the molecules to spread out through the room than to all remain inside the box. There are far more possible arrangements of them when they are spread out through the room. That is a higher entropy state. Once they have diffused out into the room, it is incredibly unlikely you will ever find that, simply by chance, they are all inside the box at the same time again. That is what the second law tells us. Simple, right?

But there is a problem with this logic. So far as we know, the laws of physics are symmetric with respect to time. (Or more accurately they obey CPT invariance, but that distinction is irrelevant for this discussion.) If you reverse the velocity of every particle, the entire system will exactly retrace its trajectory in reverse. There is nothing in the laws of physics that distinguishes between “forward” and “backward” in time. Those are just arbitrary labels we assign to two opposite directions. If we reversed the labels, nothing about the laws of physics would change. (For CPT invariance, you also need to reverse the labels “left” and “right”, as well as the labels “positive” and “negative” charge.)

So we can repeat the same argument in reverse. Consider a room containing gas molecules and an open box. Now close the box. Assume that after you close it, all the molecules are inside the box. At any earlier point in time when the box was still open, it was incredibly unlikely that all the molecules just happened to be inside the box. So this “proves” that entropy tends to decrease.

Stated this way, the argument sounds like nonsense. How can we assume that all the gas just happened to be inside the box at the moment we closed it? But in that case, why is it any more reasonable to assume the gas was initially inside the box in the original version of the argument? Just what is going on here?

Boltzmann himself spent years trying to prove the second law, and several times thought he had succeeded. The most famous of these attempts is known as the “H-theorem”, published in 1872. It appeared to be a completely rigorous proof of the second law for a gas of classical particles. Of course it contained an assumption that broke time symmetry, but that assumption came into it in such a subtle way that for a few years neither Boltzmann nor any of his contemporaries noticed it. Then in 1876, Josef Loschmidt published his “reversibility objection” to the H-theorem. He reasoned as follows.

Start with a system in a low entropy (non-equilibrium) state. Given enough time, it will eventually move to a higher entropy, equilibrium state. Now reverse the velocity of every particle. Since the equations of motion are time symmetric, the system will retrace its trajectory and return to the original, low entropy state (except with velocities reversed). Clearly the second law cannot be a universal law that applies in all cases, because we have just presented a counter-example: a case where the entropy spontaneously decreases.

In response to this objection, Boltzmann slightly modified his position, arguing the second law was still true in a probabilistic sense. He acknowledged that one could carefully design states whose entropy would decrease. But any state that was not specifically designed in this way would almost certainly be one whose entropy increased rather than decreasing.

Unfortunately, this claim is simply wrong. The original low entropy state, the high entropy state it evolves into, and the time reversed version of that high entropy state (which will spontaneously return to a lower entropy one) are all equally probable. There is no basis for accepting some while ignoring others. For every trajectory whose entropy increases, there is a reversed trajectory whose entropy decreases. Both directions are equally probable.

Later in his career, Boltzmann finally came to a clearer understanding of the situation. He recognized that the second law *cannot* be proven. No amount of statistics or probability can turn a symmetric theory into a non-symmetric one. The second law of thermodynamics is not really a “law” at all. It is a *definition*. When we speak of “forward in time,” we really mean “the direction of increasing entropy.” All the phenomena that lead us to view time as asymmetric follow directly from the change in entropy: that it is easy to break a window but hard to fix it; that friction makes things slow down instead of speeding up; even that you can remember the past but not the future. If the arrow of time were reversed so that entropy increased in the opposite direction, you would not notice any difference at all. You would simply perceive the opposite direction as being “forward in time.”

Boltzmann was unable to explain *why* entropy is changing in the first place. If the universe were in equilibrium, entropy should be constant and unchanging in all directions. The best answer he could suggest was an anthropic argument. He observed that even in equilibrium, highly improbable states can still happen. In fact, if you wait long enough, *every* state will eventually be visited. The entire Earth could be a random fluctuation that occurred just by chance. A fluctuation of that sort is incredibly improbable; most of the time, the atoms making up the Earth should be arranged in a generally uniform, featureless way. But in that situation, life could not exist. Given that we *do* exist, we should not be surprised to find the universe in one of the rare states that can support life.

This was a very clever idea, and in many ways a very modern one. It was, however, wrong. It leads to some very specific predictions, and those predictions do not match the universe we actually see around us.

Today we have a better explanation: the change in entropy is an effect of the Big Bang. We know that roughly 13.8 billion years ago, the universe was in a state of incredibly low entropy with all its matter and energy concentrated in a tiny region of space. Entropy has been steadily increasing ever since, and will continue increasing far into the future. That is why we find ourselves in a region of spacetime where there is an entropy gradient. For practical purposes, the phrase “forward in time” really means, “away from the Big Bang.”

We still do not know *why* the Big Bang happened. There are various ideas about that, but we do not yet have enough evidence to say which one is correct. There may be other regions of spacetime where entropy is flat and unchanging. If so, those regions cannot support life, so it is not surprising we are not in one of them. We do not know whether the Big Bang marks the boundary of the universe, or whether there is more universe on the other side of it. If the latter, then entropy in that region presumably increases in the opposite direction. Anyone living there would perceive time as moving in the opposite direction to how we perceive it. Perhaps there are many different Big Bangs and many different universes surrounding them, each with its own arrow of time. Based on our current evidence, we just cannot say.

5.4. Heat and Entropy

Consider a thermodynamic process in which an infinitesimal amount of heat dQ is added to a system, while keeping its volume and other macroscopic variables fixed (so that it does not do any mechanical work). The first law of thermodynamics then simplifies to $dE=dQ$. We can use the chain rule to rewrite this as

(2)

$$dQ=dE=\left(\frac{\partial \log(\Omega)}{\partial E}\right)^{-1} d(\log(\Omega))=TdS$$

This shows a direct connection between heat and entropy. If you add heat to a system, you raise its entropy. If you remove heat, you lower its entropy. This is just our assumption that Ω increases with energy, coming back in yet another form.

Suppose an amount of heat dQ flows from subsystem A to subsystem B. The entropy of A decreases and the entropy of B increases. The total change in entropy of the whole system is the sum of the two:

(3)

$$dS=dS_B+dS_A=\frac{dQ}{T_B}-\frac{dQ}{T_A}$$

If $T_A=T_B$, then dS is zero and the total entropy of the system remains constant. But if the temperatures were equal, no heat would have flowed in the first place. We saw this in section 4.6: if two systems have the same temperature, they are in thermal equilibrium and no heat flows between them.

If $T_A<T_B$, then dS is negative and the total entropy decreases. But this case is forbidden by the second law of thermodynamics. Heat will never

spontaneously flow from a colder body to a warmer one. The total change in entropy cannot be negative. (You see now how these two versions of the second law are equivalent to each other.)

So the only possibility is that $T_A > T_B$. Heat is flowing from a warmer body to a colder one, and the overall entropy of the system increases. This leads us to the following very important conclusion: *whenever heat flows between two bodies, the total entropy of the system increases*. As we will see shortly, this has important consequences for anyone trying to build a steam engine.

5.5. Heat Capacity

If you add heat to a system, you increase its energy and therefore its temperature as well. The proportionality constant is called the *heat capacity*:

(4)

$$C \equiv \frac{dQ}{dT}$$

If the heat capacity is large, that means you can add a lot of heat to the system while having only a small effect on the temperature. Heat capacity is an extensive property; if you double the size of the system, you also double its heat capacity.

As in the previous section, assume the volume is held fixed so the system does no mechanical work. In that case, we can use equation (2) to write the heat capacity at constant volume:

(5)

$$C_V = \frac{(T\,dS)}{dT} = T\left(\frac{\partial S}{\partial T}\right)_V$$

The subscript V refers to the fact that we are keeping the volume fixed. The expression $\left(\frac{\partial S}{\partial T}\right)_V$ is an example of a notation that is common in thermodynamics, and usually used without any explanation of what it really means. In this case, $S = k\log(\Omega(E,V))$. That cannot in any way be considered a function of T! In fact, the temperature is defined by a *derivative* of S. It describes the distribution of microstates with energy. It makes no sense to speak of the "temperature" of a single microstate, or the density of states with a given temperature! So what do we mean by writing a derivative of S with respect to T?

As the system interacts with the heat bath, its energy (and hence entropy) continuously fluctuate. So S is not even well defined in this situation. On the other hand, we can still compute the *average* entropy $\langle S\rangle$. That *is* a well defined function of temperature, so we can reasonably compute its derivative with respect to T. That is what the partial derivative in equation (5) really means: the derivative of $\langle S\rangle$, not of S itself. Of course, the average will vary depending on what ensemble we use to calculate it. We therefore add a subscript to specify the ensemble: in this case, that V is held fixed.

Of course, entropy was first known as an experimental quantity before it was later explained as a statistical quantity. If you think about it in those terms, there is nothing confusing about any of this. Just measure the entropy of a system, and see how it varies with temperature. You find that the result differs depending on whether you fix the volume or the pressure during your experiment, so of course you add a subscript to indicate which one you measured. Simple, right? Just remember that your experiment is actually measuring the *average* entropy over a range of macrostates, not the entropy of any one specific macrostate.

If you choose to fix pressure instead of volume, equation (2) changes to

$$dE=TdS=dQ-PdV$$

and the heat capacity becomes

(6)

$$C_P=\frac{(TdS+PdV)}{dT}=T\left(\frac{\partial S}{\partial T}\right)_P+P\left(\frac{\partial V}{\partial T}\right)_P$$

It can be shown that for nearly all realistic systems, $C_P>C_V$. This is easy to understand. If you let the system expand as you add heat, it performs work on its environment. That reduces the energy of the system. Energy used to perform work is energy that doesn’t go into raising the temperature, so T increases more slowly.

5.6. The Ideal Gas Law

An *ideal gas* is defined as a gas of free particles that do not interact with each other in any way. It makes a very useful model for studying how gasses behave during thermodynamic processes.

The density of states of an ideal gas takes a particularly simple form. Each particle is equally likely to be anywhere in the allowed volume, so for any single particle considered on its own, $\Omega \propto V$. Because the particles do not interact with each other, the density of states for the whole gas is just the product of the ones for the individual particles: $\Omega \propto VN$. The definition of pressure therefore becomes

$$P \equiv kT \partial \log(\Omega) \partial V = NkT \partial \log(V) \partial V = NkTV$$

Rearranging this gives the *ideal gas law*, also known as the *ideal gas equation of state*:

(7)

$$PV = NkT$$

A real gas is not an ideal gas, of course. The particles do interact with each other. Still, this turns out to be a surprisingly good approximation for dilute gasses of non-polar, non-reactive molecules. Air, for example. Furthermore, even when it is not accurate enough to make quantitative predictions, it still gives a good qualitative description of how real gasses behave. So let us take a minute to examine it and see what it can tell us.

Suppose you increase the temperature of a gas, such as by lighting a fire under it. The ideal gas law immediately tells us that the pressure, the volume, or both must increase as well. If you hold the volume fixed, the pressure will increase in direct proportion to the temperature. If instead you hold the pressure fixed while allowing the volume to change, then the gas will expand.

As it expands it does work on its environment: $W = P\Delta V$, where ΔV is the change in volume. That means the energy of the gas must decrease by the same amount to conserve energy, and that in turn means its temperature will decrease. This leads us to another important principle: *when a gas expands, its temperature tends to decrease*.

Of course, it works the other way as well. If you compress the gas by performing work on it, you increase its energy and thus its temperature.

So you see, this simple gas of noninteracting particles actually has an amazing ability: it can convert thermal energy into mechanical work, and vice versa! That certainly sounds like it ought to be a useful ability. Surely *someone* must have come up with something to do with it!

5.7. Heat Engines

A device that transforms thermal energy into mechanical work is called a *heat engine*. From what we saw in the last section, it looks like a gas could be very useful in building one. For example, you might have a chamber filled with gas, and one wall of the chamber is a piston that can move in and out. As the gas expands, the piston moves outward and does work.

But what then? Once it has gone all the way out, it cannot move any further. Before you can extract any more work, you first need to compress the gas again. And to do that, you need to perform work *on* the gas, thus giving back all the useful work you just got out of it. Not so useful after all!

But with a little cleverness we can get around this problem. As the gas expands it does work $P\Delta VP\Delta V$ on its environment, and as we compress it we have to do work $P\Delta VP\Delta V$ on the gas. But what if the pressures in these two expressions were different? We want the pressure to be high while the gas expands and low while we compress it. In that case, we will get more work out of the gas in the expansion stage than we have to put into it in the compression stage.

The ideal gas law tells us how we can do this: by changing the temperature! We need the gas to be hot while it expands and cold while it contracts. That is easily achieved. In addition to the gas (also known as the *working body*), we need two different heat baths: one called the “hot bath” or “source” at temperature $THTH$, and one called the “cold bath” or “sink” at temperature $TCTC$. Here is a specific example of a sequence of steps you might use.

1. Put the working body in contact with the hot bath and let it expand. As it expands it does useful work. Normally this would cause its temperature to drop, but because it is in contact with a heat bath, it steadily absorbs energy and its temperature remains at $THTH$.
2. Now disconnect the working body from the heat bath so it is isolated. Let it continue to expand. It continues to do useful work, and since it is now isolated, its temperature does drop. Let this continue until its temperature reaches $TCTC$.
3. Put it in contact with the cold bath and start compressing it. This would normally cause its temperature to rise, but since it is in contact with the cold bath, it instead expels energy to the bath and remains at $TCTC$.

4. Disconnect the working body from the bath so it is once again isolated. Continue doing work to compress it. This will now cause its temperature to rise. Continue until it reaches T_H.

This is an example of a *thermodynamic cycle*: a sequence of thermodynamic processes that end with the system (other than the heat baths) in exactly the same state it began in, so we can repeat the cycle over and over. In particular, this is known as the *Carnot cycle* after Sadi Carnot, who proposed it in 1824. A heat engine that uses this cycles is called a *Carnot heat engine*.

Let us analyze what happens during this cycle. Let Q_H be the heat absorbed from the hot bath, Q_C the heat expelled to the cold bath, and W the net amount of work performed during the whole cycle (that is, the work performed *by* the working body during steps 1 and 2, minus the work performed *on* the working body in steps 3 and 4). Conservation of energy requires that $W=Q_H-Q_C$.

As we saw in section 5.4, whenever heat flows into or out of a system, its entropy changes as well. Therefore the entropy of the two heat baths must change over the course of the cycle. The entropy of the hot bath decreases and the entropy of the cold bath increases. From equation (2), $Q_H=T_H\Delta S_H$ and $Q_C=T_C\Delta S_C$.

The *efficiency* of a heat engine, represented by the symbol η, is defined as the ratio of the work done by it to the energy absorbed from the hot bath:

(8)

$$\eta\equiv\frac{W}{Q_H}=\frac{Q_H-Q_C}{Q_H}=1-\frac{Q_C}{Q_H}=1-\frac{T_C\Delta S_C}{T_H\Delta S_H}$$

The second law of thermodynamics requires that the total entropy of the whole system must increase or stay constant over the cycle. The working body ends up in exactly the same state it began in, so its entropy does not change. The engine also does work on its environment, but we cannot assume anything about what happens there. Perhaps all the energy is being stored into a single degree of freedom with no entropy at all. The two heat baths are all we have to work with, so the second law requires that $\Delta S_C\geq\Delta S_H$. We therefore conclude

(9)

$$\eta\leq 1-\frac{T_C}{T_H}$$

Notice how little we assumed in deriving this: merely that the heat engine absorbed heat from one bath, expelled heat to another, and did work. I described the Carnot cycle as an illustration, but no details of the cycle were required for the derivation. Therefore, equation (9) applies equally well to heat engines that use different cycles. I assumed nothing about the nature of the working body. It could be a gas, a liquid, a solid, or even something exotic like a supercritical fluid. I assumed nothing about how the working body performed work. It could involve moving a piston, applying an electric field, or shooting ping-pong balls at a target.

Absolutely any heat engine, no matter what it is made of or how it works, must obey equation (9), a fact known as *Carnot's theorem*. The second law of thermodynamics requires that we *must* have a second heat bath: we decrease the entropy of the hot bath, so we need to make up for that somewhere else. Conservation of energy requires that any heat expelled to the cold bath is not available for doing work. Those two facts place an absolute limit on how efficient any heat engine can ever be.

There is no lower limit, of course. You can make a heat engine as inefficient as you want. You could connect the two heat baths and allow energy to flow directly from one to the other without doing any work at all, thus achieving a spectacular efficiency of zero. The key to making an efficient engine is to minimize all transfers of heat *except* the ones that are absolutely required for the operation of the engine.

Another interesting fact about heat engines is that you can run them backward, which turns them into *heat pumps*. Just perform the same steps in reverse order and changing the direction of movement (so that expanding becomes compressing, for example). In this case, heat flows *out* of the cold bath and *into* the hot bath. Heat is flowing opposite to its normal direction! It does not do this spontaneously, of course. Instead of the engine *producing* work, we now have to do work on it. You are probably already very familiar with this fact: refrigerators and air conditioners have an unfortunate habit of needing to be connected to an external source of energy.

5.8. Free Energy

Consider an even simpler case: just a working body and a single heat bath. Assume the working body has energy EE and entropy SS, and that the heat bath has temperature TT. We now want to answer a simple question: what is the maximum amount of work we can extract from the body?

Ideally, we would like the answer to be E. We want to extract *all* the energy as work. But the second law makes that impossible. As the body's energy decreases, its entropy does too. We need to make up for that by adding heat (and entropy) to the heat bath.

The third law of thermodynamics says that as the body's energy (and hence temperature) goes to zero, its entropy approaches some minimum value. Call that S_0. So its entropy decreases by $S-S_0$, and we need to add heat $T(S-S_0)$ to the heat bath. Therefore, the maximum about of work we can hope to extract from the body is

(10)

$$W_{max}=E-T(S-S_0)$$

If we assume S_0 is zero, this simplifies further to $W_{max}=E-TS$, which we recognize as the Helmholtz free energy.

In general, of course, S_0 is not zero, but that is a result of the statistical definition of entropy. Before statistical mechanics was developed, the zero point of entropy was considered arbitrary. In thermodynamics, only *differences* in entropy are usually important, so it was common to fix the zero point by defining S_0 to be zero.

This is the origin of the term "free energy". It means energy that is "free" in the sense of "available" or "not locked up in an unusable form". It is the maximum amount of energy you can hope to extract as work. Notice that its value depends on the temperature of the heat bath. The colder the heat bath you have access to, the more work you can extract from the working body.

Phase Space

Why are statistics important? In statistical physics, we are in fact dealing with DoFs in the order of Avogadro’s number, most of the time. If we are going to calculate the dynamics of these systems by combining all the dynamics of each particles, it becomes rather difficult. It already takes a huge amount of storage to store one screenshot of the system. Assuming each DoF take only 8 bits, we need 10231023 bytes that is 10171017 GB. It seems to be impossible to store this one snapshot of the system. It is time to steer away from this Newtonian approach.

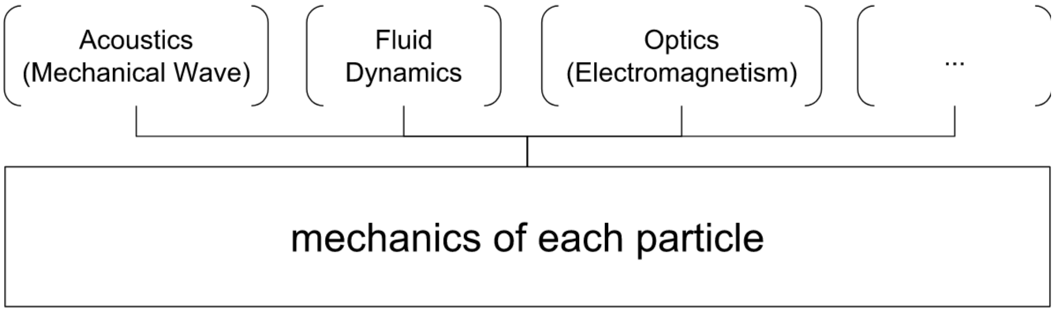


Fig. 8 Newton’s plan of mechanics. Mechanics was in the center of all physics.

What is mechanics? Mechanics is the thinking that deals with dynamics of object in the following way:

- Description of initial state;
- Time evolution of the system;
- Extraction of observables.

The initial state of an object is exactly what the name indicates. In a coordinate system approach, the initial state should include the coordinates and the time derivative of the coordinates since we are interested in dynamics. In fact, the coordinates and time derivative of them form a phase space. In general, a vector in phase space is a complete description of the state of an object. Thus time evolution of the object is simply the motion of state vector in phase space. Finally, we will do whatever is needed to extract observables. For example, we could trivially use the projection of points in phase space to get the position or velocity.

The problem arise when it comes to a system with a large amount of particles. As mentioned previously, it’s in general not possible to record all these DoFs. We need a completely new scheme.

Boltzmann Factor

probability of a point in phase space=exp(−EkBT)probability of a point in phase space=exp⁡(−EkBT)

Boltzmann factor gives us the (not normalized) probability of the system staying on a phase space state with energy EE.

Why Boltzmann Factor

Why does Boltzmann factor appear a lot in equilibrium statistical mechanics? Equilibrium of the system means when we add infinitesimal amount of energy to the

whole thing including system and reservoir, a characteristic quantity C(E)=CSCRC(E)=CSCR won’t change. That is the system and the reservoir will have the same changing rate of the characteristic quantity when energy is changed, i.e.,
∂lnCS∂ES=−∂lnCR∂ER.∂ln⁡CS∂ES=−∂ln⁡CR∂ER.

We have dE1=−dE2dE1=−dE2 in a equilibrium state. They should both be a constant, which we set to ββ. Finally we have something like
∂lnCS∂ES=β∂ln⁡CS∂ES=β

which will give us a Boltzmann factor there.

This is just a very simple procedure to show that Boltzmann factor is kind of a natural factor in equilibrium system.

Partition Function

For a given Hamiltonian H, the (classical) partition function Z is

Z=∫dp∫dxe−βHZ=∫dp∫dxe−βH

A simple example is the Harmonic Oscillator,

H=p22m+12qx2H=p22m+12qx2

The partition function

Z=∫e−βp2/(2m)dq∫e−β12qx2dx=2πm/q−−−−√1βZ=∫e−βp2/(2m)dq∫e−β12qx2dx=2πm/q 1β

Energy

E=1Z∫∫e−βp2/(2m)e−β12qx2Hdpdx=⋯=kBTE=1Z∫∫e−βp2/(2m)e−β12qx2Hdpdx=⋯=kB T

(This result is obvious if we think about equipartition theorem.)

A more clever approach for the energy is to take the derivative of partition function over ββ, which exactly is
⟨E⟩=−∂∂βlnZ⟨E⟩=−∂∂βln⁡Z

In our simple case,

lnZ=−∂∂β(ln(kBT)+SomeConstant)=kBTln⁡Z=−∂∂β(ln⁡(kBT)+SomeConstant)=kBT

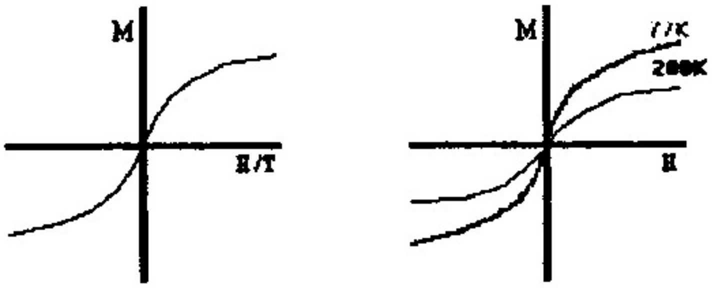
This is the power of partition function. To continue the SHO example, we find the specific heat is

C=kBC=kB

Does This Result Depend on SHO

Magnetization

We have such a result in an experiment of magnetization with N magnetic dipoles in 1D.



How can we describe this with a theory?

It’s not possible to describe the system by writing down the dynamics of each magnetic dipole. So we have to try some macroscpic view of the system. Probability theory is a great tool for this. The probability of a dipole on a energy state $EiEi$ is $P(Ei)=exp(-\beta Ei)\sum ni=1exp(-\beta Ei).P(Ei)=exp(-\beta Ei)\sum i=1nexp(-\beta Ei).$

So the megnetization in this simple case is

$M=(\mu Ne\beta\mu B-\mu Ne-\beta\mu B)/(exp(\beta\mu B)+exp(-\beta\mu B))=\mu Ntanh(\beta\mu B)M=(\mu Ne\beta\mu B-\mu Ne-\beta\mu B)/(exp(\beta\mu B)+exp(-\beta\mu B))=\mu Ntanh(\beta\mu B)$

Heat Capacity

Another category of problems is temperature related. For example, a study of average energy with change temperature.

For the paramagnetic example, the energy of the system is

$E=-(\mu BNe\beta\mu B-\mu Ne-\beta\mu B)/(exp(\beta\mu B)+exp(-\beta\mu B))=-\mu NBtanh(\beta\mu B)E=-(\mu BNe\beta\mu B-\mu Ne-\beta\mu B)/(exp(\beta\mu B)+exp(-\beta\mu B))=-\mu NBtanh(\beta\mu B)$

Obviously, no phase transition would occur. But if we introduce self interactions between dipoles and go to higher dimensions, it’s possible to find phase transitions.