Thermal Physics[1]

"A theory is the more impressive the greater the simplicity of its premises, the more different kinds of things it relates, and the more extended its area of applicability. Therefore the deep impression that classical thermodynamics made upon me. It is the only physical theory of universal content which I am convinced will never be overthrown, within the framework of applicability of its basic concepts." — A. Einstein

"But although, as a matter of history, statistical mechanics owes its origin to investigations in thermodynamics, it seems eminently worthy of an independent development, both on account of the elegance and simplicity of its principles, and because it yields new results and places old truths in a new light in departments quite outside of thermodynamics."

— J.W. Gibbs

We have seen how a few simple laws (in particular NEWTON'S SECOND LAW) can be combined with not-too-sophisticated mathematics to solve a great variety of problems — problems which eventually are perceived to fall into a number of reasonably well-defined categories by virtue of the mathematical manipulations appropriate to each — and that those distinct classes of mathematical manipulations eventually become familiar enough to acquire familiar names of their own, such as "conservation of impulse and momentum" or "conservation of work and energy" or "conservation of torque and angular momentum." This *emergence* of new tacit paradigms was the great conceptual gift of the Newtonian revolution. But the most profound *practical* impact of the new sciences on society came in the form of the Industrial Revolution, which was made possible only when the science of mechanics was combined with an understanding of how to extract usable mechanical **work** from that most mysterious of all forms of energy, **heat**.

Historically, heat was recognized as a form of energy and *temperature* was understood in terms of its qualitative properties long before anyone truly understood what either of these terms actually meant in any rigorous microscopic model of matter. The link between Newton's mechanics and the thermodynamics of Joule and Kelvin was forged by Boltzmann long after steam power had changed the world, and a simple understanding of many of the finer points of Boltzmann's *statistical mechanics* had to wait even longer until Quantum Mechanics provided a natural explanation for the requisite fact that the number of possible states of any system, while huge, is not infinite, and that small, simple systems are in fact restricted to a countable number of discrete "stationary states." In this drama Albert Einstein was to play a rather important role.

The following conceptual outline of Statistical Mechanics is designed to make the subject as clear as possible, not to be historically accurate or even fair. Having made this choice, however, I hope to be able to display the essence of the most astonishing product of human Science without undue rigamarole, and perhaps to convey the wonder that arises from a deeper and more fundamental understanding.

15.1 Random Chance

With so many miracles to choose from, where do I get off declaring Statistical Mechanics to be "the most astonishing product of human Science?" This is of course a personal opinion, but it is one shared by many physicists — perhaps even a majority. The astonishment is a result of the incredible precision with which one can predict the outcome of experiments on very complicated systems (the more complicated, the more precise!) based on the FUNDAMENTAL ASSUMPTION of STATISTICAL MECHANICS:

> *A system in thermal equilibrium is a priori equally likely to be found in any one of the fully-specified states accessible to it.*

This seemingly trivial statement contains a couple of ringers: the word "accessible" means, for instance, that the total "internal" energy of the system — which is always written U — *i.e.* the sum of the kinetic and potential energies of all the little particles and waves that make up the big system — is fixed. There are many ways to divide up that energy, giving more to one particle and less to another, and the FUNDAMENTAL ASSUMPTION says that they are all equally likely; but in every case the energy must add up to the same U. This can obviously be very confusing, but fortunately we rarely attempt to count up the possibilities on our fingers!

It is the assumption itself that is so amazing. How can anything but total ignorance result from the assumption that we know *nothing at all* about the minute biases a

[1] I have "borrowed" the notation, general approach, basic derivations and most of the quotations shown here from the excellent textbook of the same name by Kittel & Kroemer, who therefore deserve all the credit (and none of the blame) for the abbreviated version displayed before you.

real system might have for one state over another? More emphatically, how can such an outrageous assumption lead to anything but wrong predictions? It amounts to a pronouncement that Nature runs a perfectly honest casino, in which every possible combination of the roll of the dice is *actually* equally likely! And yet every prediction derived from this assumption has been demonstrated to be accurate to the best precision our measurements can provide. And the consequences are numerous indeed!

15.2 Counting the Ways

If we accept the FUNDAMENTAL ASSUMPTION at face value, then it is easy to calculate the *probability* of finding the equilibrated system in any given fully specified state: if the state is *not accessible* [*e.g.* if it takes more energy U than we have at our disposal] then the probability is *zero*; if it *is* accessible, then its probability is just $\frac{1}{\Omega}$, where Ω is the *total number of accessible states*. The first step is therefore to calculate Ω. In general this can get difficult, but we can choose a few simple examples to illustrate how the calculation goes.

15.2.1 Conditional Multiplicity

Suppose we have a jar full of pennies, say N pennies, all of which have had unique numbers painted on them so that they can be easily distinguished from each other. Now suppose we shake it thoroughly and dump it out on a nice flat table; each penny falls either "heads" or "tails" with equal *a priori* probability. The probability of penny #1 being "heads" is $\frac{1}{2}$. The probability of penny #1 being "heads" *and* penny #2 being "tails" is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. The probability of penny #1 being "heads" *and* penny #2 being "tails" *and* penny #3 being "tails" is $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$. And so on. If the pennies are all "statistically independent" (*i.e.* how one penny falls has no influence on the other pennies), the probability of *any specific arrangement* of specific pennies falling specific ways [what we call a *fully specified state* of the system] is

$$\left(\frac{1}{2}\right)^N = \frac{1}{2^N}$$

where N is the total number of pennies.

Unfortunately, *this is not what we want to know.* We don't care *which* pennies fall which way,[2] only *how many* of each. This is what we call a *partially specified* or *partially constrained* state of the system. What we really want to know is the *number of ways* we can get n heads and $(N-n)$ tails.[3]

Suppose we *specify* that n pennies are "heads" and the remaining $(N-n)$ are "tails." The *number of ways we can do this* is what we call $\Omega(n, N)$, the *multiplicity function* for the *partially constrained* state specified only by n and N. Here's how we calculate $\Omega(n, N)$: the number of different ways we can rearrange all N coins is

$$N! \equiv N \cdot (N-1) \cdot (N-2) \cdots 3 \cdot 2 \cdot 1$$

because we have N choices of which coin will be first, then we have $(N-1)$ choices of which coin will be second, then we have $(N-2)$ choices of which coin will be third, and so on. The total number of choices is the *product* of the numbers of choices at each step. However, we have *overcounted* by the number of different ways the *heads* can be rearranged *among themselves*, which by the same argument is $n!$, *and* by the number $(N-n)!$ of ways the *tails* can be rearranged among *them*selves. Therefore the total number of *distinguishable* combinations that all give n heads and $(N-n)$ tails is

$$\Omega(n, N) = \frac{N!}{n! \cdot (N-n)!} \tag{1}$$

Another example would be a parking lot with N spaces in which n cars are parked. The number of different ways we can label the spaces is $N!$ but the n occupied spaces can be rearranged amongst themselves $n!$ different ways and the $(N-n)$ empty spaces can be rearranged $(N-n)!$ different ways without altering the partial constraint [namely, that only n of the spaces are filled].[4] Then Eq. (1) describes the number of different ways the cars can be parked without changing the total number of parked cars.

[2] In the present case, we have a *choice* of whether to treat the pennies as "indistinguishable" or not. No two pennies are *really* indistinguishable, of course; even without our painted-on numbers, each one has unique scratches on its surface and was crystallized from the molten state in a unique microscopic pattern. We *could* tell one from another; we just don't *care*, for circumstantial reasons. In QUANTUM MECHANICS, however, you will encounter the concept of *elementary particles* [*e.g.* electrons] which are so uncomplicated that they truly *are* indistinguishable [*i.e. perfectly* identical]; moreover, STATISTICAL MECHANICS provides a means of actually *testing* to see whether they are *really absolutely indistinguishable* or just very similar!

[3] It might be that we get to keep all the pennies that come up heads, but for every penny that comes up tails we have to chip in another penny of our own. In that case our *profit* would be $n - (N-n) = 2n - N$ cents.

[4] If you were the parking lot owner and were charging \$1 per space [cheap!], your profit would be \$$n$. I keep coming back to monetary examples — I guess *cash* is the social analogue of *energy* in this context.

The Binomial Distribution

To generalize, we talk about a *system of* N *particles*,[5] each of which can only be in one of two possible *single-particle states*. A *fully specified* N*-particle state* of the system would have the single-particle state of *each individual particle* specified, and is not very interesting. The *partially specified* N*-particle state* with n of the particles in the first single-particle state and the remaining $(N-n)$ particles in the other single-particle state can be realized in $\Omega(n,N)$ different ways, with $\Omega(n,N)$ given by Eq. (1). Because there are only *two* possible single-particle states, this case of Ω is called the *binomial* distribution. It is plotted[6] in Fig. 15.1 for several values of N.

Note what happens to $\Omega(n,N)$ as N gets bigger: the peak value, which always occurs at $n_{\text{peak}} = \frac{1}{2}N$, gets very large [in the plots it is compensated by dividing by 2^N, which is a big number for large N] and the *width* of the distribution grows steadily *narrower* — *i.e.* values of $\frac{n}{N}$ far away from the peak get less and less likely as N increases. The width is in fact the *standard deviation*[7] of a hypothetical random sample of n, and is proportional to $\sqrt{N}$. The *fractional* width (expressed as a fraction of the total range of n, namely N) is therefore proportional to $\frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$:

$$\text{Fractional Width} \propto \frac{1}{\sqrt{N}} \tag{2}$$

which means that for *really* large N, like $N = 10^{20}$, the binomial distribution will get *really* narrow, like a part in 10^{10}, in terms of the *fraction of the average*.

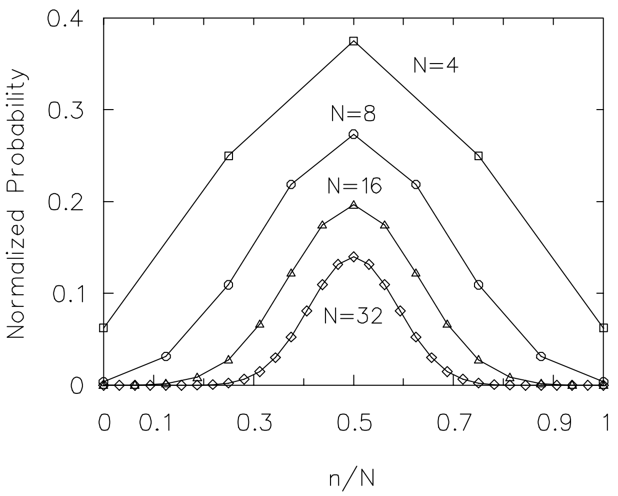

Figure 15.1 The *normalized binomial distribution* for several values of N. In order to put several cases on a single graph, the horizontal axis shows n divided by its maximum possible value N [giving the *fraction* of the total range] and the binomial coefficient $\Omega(n,N)$ given by Eq. (1) has been divided by the total number of possible fully specified N-particle states, 2^N, to give the "normalized" probability — *i.e.* if we add up the values of $\Omega(n,N)/2^N$ for all possible n from 0 to N, the total probability must be 1. [This is eminently sensible; the probability of n having *some* value is surely equal to unity!]

15.2.2 Entropy

> *"If we wish to find in rational mechanics an a priori foundation for the principles of thermodynamics, we must seek mechanical definitions of temperature and entropy."* — J.W. Gibbs

The function $\Omega(n,N)$ is called the MULTIPLICITY FUNCTION for the partially specified system. If N and n get to be large numbers (which is usually the case when we are talking about things like the numbers of electrons in a crystal), $\Omega(n,N)$ can get *really huge*.[8] It is so huge, in fact, that it becomes very difficult to cope with, and we do what one usually does with ungainly huge numbers to make them manageable: we take its *logarithm*. We define the [natural] logarithm of Ω to be the ENTROPY σ:

$$\sigma \equiv \ln \Omega \tag{3}$$

[5] The term "particle" is [in this usage] meant to be as vague as possible, just like "system:" the *particles* are "really simple things that are all very much alike" and the *system* is "a bunch of particles taken together."

[6] Actually what is plotted in Fig. 15.1 is the *probability* function

$$\mathcal{P}(n) \equiv \frac{1}{2^N} \cdot \Omega(n,N) = \frac{1}{2^N} \cdot \frac{N!}{n!\,(N-n)!}$$

vs. $\frac{n}{N}$, as explained in the caption. Otherwise it would be difficult to put more than one plot on the same graph, as the peak value of $\Omega(n,N)$ gets very large very fast as N increases!

[7] Recall your Physics Lab training on MEASUREMENT!

[8] A good estimate of the size of $N!$ for large N is given by *Stirling's approximation*:

$$N! \approx \sqrt{2\pi N} \cdot N^N \cdot e^{-N}$$

Let's say that again: the ENTROPY σ is the *natural logarithm* of the MULTIPLICITY FUNCTION Ω — *i.e.* of the *number of different ways we can get the partially specified conditions* in this case defined by n.

Is this all there is to the most fearsome, the most arcane, the most incomprehensible quantity of THERMODYNAMICS? Yep. Sorry to disappoint. That's it. Of course, we haven't played around with σ yet to see what it might be good for — this can get very interesting — nor have I told this story in an historically accurate sequence; the concept of ENTROPY preceded this definition in terms of "statistical mechanics" by many years, during which all of its properties were elucidated and armies of thermal physicists and engineers built the machines that powered the Industrial Revolution. But understanding THERMODYNAMICS the old-fashioned way is *hard*! So we are taking the easy route — sort of like riding a helicopter to the top of Mt. Everest.

15.3 Statistical Mechanics

Before we go on, I need to move away from our examples of binomial distributions and cast the general problem in terms more appropriate to Mechanics. We can always go back and generalize the paradigm[9] but I will develop it along traditional lines.

The owner of the parking lot described earlier is only interested in the total number of cars parked because that number will determine his or her profit. In Mechanics the "coin of the realm" is *energy*, which we have already said is always written U in thermal physics. The abstract problem in STATISTICAL MECHANICS involves a complex system with many possible states, each of which has a certain total energy U. This energy may be in the form of the sum of the kinetic energies of all the atoms of a gas confined in a box of a certain volume, or it may be the sum of all the vibrational energies of a crystal; there is no end of variety in the physical examples. But we are always talking about the *random, disordered energy* of the system, the so-called *internal* energy, when we talk about U.

Now, *given a certain amount of internal energy* U, the *number of different fully-specified states of the system* whose total internal energy is U [our *partial constraint*] is the conditional MULTIPLICITY FUNCTION $\Omega(U)$. Taking the binomial distribution as our example again, we could substitute *crystal lattice sites* for "parking places" and *defects* for "cars" [a defect could be an atom out of place, for instance]. If it takes an energy ε to create one defect, then the total internal energy stored in n defects would be $U = n\,\varepsilon$. Lots of other examples can be imagined, but this one has the energy U proportional to the number n of defects, so that you can see how the U-dependence of Ω in this case is just like the n-dependence of Ω before.

[9] Count on it!

So what?

Well, things start to get interesting when you put *two* such systems *in contact* so that U can flow freely between them through random statistical fluctuations.

15.3.1 Ensembles

One of the more esoteric notions in STATISTICAL MECHANICS is the concept of an *ensemble*. This has nothing to do with music; it goes back to the original meaning of the French word *ensemble*, which is a collection or gathering of things — much more general and abstract than the small band of musicians we tend to visualize. Anyway, the Statistical Mechanical "ENSEMBLE" is a collection of *all the possible fully specified states* of some *system*.

Of course, there are different *kinds* of ENSEMBLES depending upon what global *constraints* are in effect. For instance, the set of all possible states of an *isolated* system $\mathcal{S}$ consisting of a fixed number N of "particles"[10] with a well defined total energy U is called a MICROCANONICAL ENSEMBLE. This is what we have been discussing so far.

The set of all possible states of a system $\mathcal{S}$ consisting of a fixed number N of particles but in "thermal contact" with a *much, much larger* system $\mathcal{R}$ (called a "heat reservoir") so that the energy U of $\mathcal{S}$ can flow in or out of $\mathcal{R}$ at random is called a CANONICAL ENSEMBLE.

And the set of all possible states of a system $\mathcal{S}$ in contact with a reservoir $\mathcal{R}$ with which it can exchange *both* energy (U) *and* particles (N) is called a GRAND CANONICAL ENSEMBLE.

If the utility of these concepts is less than obvious to you, join the club. I won't need to use them to derive the good stuff below, but now you will be able to scoff at pedants that pretend you can't understand "Stat Mech" unless you know what the various types of Ensembles are.

[10] Remember, a "particle" is meant to be an abstract concept in this context!

15.4 Temperature

> *"The general connection between energy and temperature may only be established by probability considerations.* [Two systems] *are in statistical equilibrium when a transfer of energy does not increase the probability."*
> — M. Planck

When we put *two* systems $\mathcal{S}_1$ and $\mathcal{S}_2$ (with N_1 and N_2 particles, respectively) into "thermal contact" so that the (constant) total energy $U = U_1 + U_2$ can redistribute itself randomly between $\mathcal{S}_1$ and $\mathcal{S}_2$, the *combined* system $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ will, we postulate, obey the FUNDAMENTAL PRINCIPLE — it is equally likely to be found in any one of its accessible states. The *number* of accessible states of $\mathcal{S}$ (partially constrained by the requirement that N_1, N_2 and $U = U_1 + U_2$ remain constant) is given by

$$\Omega \;=\; \Omega_1(U_1) \;\times\; \Omega_2(U_2) \tag{4}$$

where Ω_1 and Ω_2 are the MULTIPLICITY FUNCTIONS for $\mathcal{S}_1$ and $\mathcal{S}_2$ taken separately [both depend upon their internal energies U_1 and U_2] and the overall multiplicity function is the *product* of the two individual multiplicity functions because the rearrangements within one system are statistically independent of the rearrangements within the other.[11] Since the ENTROPY is the log of the MULTIPLICITY and the log of a product is the sum of the logs, Eq. (4) can also be written

$$\sigma \;=\; \sigma_1(U_1) \;+\; \sigma_2(U_2) \tag{5}$$

— *i.e.* the entropy of the combined system is the *sum* of the entropies of its two subsystems.

15.4.1 The Most Probable

So what? Well, here's the thing: we know that all accessible states of the system are *a priori* equally likely; however, the *number* Ω of accessible states will depend upon the division of the total energy U between U_1 and U_2. Moreover, for a certain value of U_1 (and therefore of $U_2 = U - U_1$), Ω will be a *maximum* — *i.e.* that value of U_1 will make possible the largest variety of equally likely random states of the system and consequently we will be more likely, on average, to find the system in states with that value of U_1 than in other states[12] with different values of U_1.

This special value of U_1 is called (reasonably enough) the "most probable value" and is given the symbolic representation $\hat{U}_1$.

15.4.2 Criterion for Equilibrium

If our two systems are initially prepared separately with energies U_1 and U_2 *other than* the most probable, *what will happen* when we bring them into contact so that U can flow between them? The correct answer is, of course, "Everything that possibly *can* happen." But there is a *bigger variety* of possibilities for certain gross distributions of energy than for others, and this makes those gross distributions *more likely* than others. The overall *entropy* is thus a *measure* of this likelihood. It seems inevitable that one will eventually feel compelled to anthropomorphize this behaviour and express it as follows:[13]

> All random systems "like" *variety* and will "seek" arrangements that maximize it.

In any case, the tendency of energy to flow from one system to the other will *not* be governed by equalization of either energy or entropy themselves, but by equalization of the *rate of change* of entropy with energy, $\frac{\partial \sigma}{\partial U}$. To see why, suppose (for now) that more energy always gives more entropy. Then suppose that the entropy σ_1 of system $\mathcal{S}_1$ depends only *weakly* on its energy U_1, while the entropy σ_2 of system $\mathcal{S}_2$ depends *strongly* on its energy U_2. In mathematical terms, this reads

$$\text{Suppose} \qquad \frac{\partial \sigma_1}{\partial U_1} < \frac{\partial \sigma_2}{\partial U_2}$$

Then *removal* of a small amount of energy dU from $\mathcal{S}_1$ will *decrease* its entropy σ_1, but <u>*not by as much*</u> as the *addition* of that same energy dU to $\mathcal{S}_2$ will *increase its* entropy σ_2. Thus the *net* entropy $\sigma_1 + \sigma_2$ will be *increased* by the transfer of dU from $\mathcal{S}_1$ to $\mathcal{S}_2$. This argument is as convoluted as it sounds, but it contains the irreducible essence of the definition of temperature, so don't let it slip by!

The converse also holds, so we can combine this idea with our previous statements about the system's "preference" for higher entropy and make the following claim:

[11] If I flip my coin once and you flip your coin twice, there are $2^1 = 2$ ways my flip can go [h, t] and $2^2 = 4$ ways your 2 flips can go [HH, HT, TH, TT]; the total number of ways the *combination* of your flips and mine can go [hHH, hHT, hTH, hTT, tHH, tHT, tTH, tTT] is $2 \times 4 = 8$. And so on.

[12] Nothing precludes finding the system in states with other values of U_1, of course. In fact we *must* do so sometimes! Just less often.

[13] Perhaps the converse is actually true: human "wants" are actually manifestations of random processes whose variety is greater in the direction of perceived desire. I find this speculation disturbing.

> Energy U will flow spontaneously *from* a system with *smaller* $\dfrac{\partial\sigma}{\partial U}$ *to* a system with *larger* $\dfrac{\partial\sigma}{\partial U}$.

If the rate of increase of entropy with energy $\left(\frac{\partial\sigma}{\partial U}\right)$ is the *same* for $\mathcal{S}_1$ and $\mathcal{S}_2$, then the combined system will be "happy," the energy will stay where it is (on average) and a state of "thermal equilibrium" will prevail.

Mathematical Derivation

Is there any way to *derive* a formal (mathematical) criterion for the condition of thermal equilibrium, starting from a hypothetical knowledge of Ω_1 as a function of U_1 and Ω_2 as a function of $U_2 = U - U_1$? Of course! Why else would I be doing this? The thing about a *maximum* of a function (or a minimum, for that matter; either type of *extremum* obeys the same rule) is that the *slope* of the function must be *zero* at the extremum. [Otherwise it would still have further up or down to go!] Since the slope is given by the *derivative*, this reads

$$\text{Criterion for an } \underline{\text{extremum}}: \qquad \frac{\partial\Omega}{\partial U_1} = 0 \tag{6}$$

In this case, since $\Omega = \Omega_1 \cdot \Omega_2$, the PRODUCT RULE for derivatives gives

$$\frac{\partial\Omega}{\partial U_1} = \frac{\partial\Omega_1}{\partial U_1} \cdot \Omega_2 + \Omega_1 \cdot \frac{\partial\Omega_2}{\partial U_1} = 0 \tag{7}$$

Now, Ω_2 is a function of U_2, not U_1; but we can get around that by using the CHAIN RULE,

$$\frac{\partial\Omega_2}{\partial U_1} = \frac{\partial\Omega_2}{\partial U_2} \cdot \frac{\partial U_2}{\partial U_1}.$$

where $U_2 = U - U_1$ and U is a constant, so

$$\frac{\partial U_2}{\partial U_1} = -1$$

We can therefore substitute $-\dfrac{\partial\Omega_2}{\partial U_2}$ for $\dfrac{\partial\Omega_2}{\partial U_1}$ in Eq. (7):

$$\frac{\partial\Omega_1}{\partial U_1} \cdot \Omega_2 - \Omega_1 \cdot \frac{\partial\Omega_2}{\partial U_2} = 0$$

or

$$\frac{\partial\Omega_1}{\partial U_1} \cdot \Omega_2 = \Omega_1 \cdot \frac{\partial\Omega_2}{\partial U_2}$$

If we now divide both sides by the product $\Omega_1 \cdot \Omega_2$, we get

$$\frac{1}{\Omega_1} \cdot \frac{\partial\Omega_1}{\partial U_1} = \frac{1}{\Omega_2} \cdot \frac{\partial\Omega_2}{\partial U_2}. \tag{8}$$

Now we need to recall the property of the *natural logarithm* that was so endearing when we first encountered it: $\ln(x)$ is the function *whose derivative is the inverse*,

$$\frac{d}{dx}\ln(x) = \frac{1}{x}$$

and, by the CHAIN RULE,

$$\frac{d}{dx}\ln(y) = \frac{1}{y} \cdot \frac{dy}{dx}$$

In this case "y" is Ω and "x" is U, so we have

$$\frac{\partial}{\partial U}\ln(\Omega) = \frac{1}{\Omega} \cdot \frac{\partial\Omega}{\partial U}$$

which means that Eq. (8) can be written

$$\frac{\partial}{\partial U_1}\ln(\Omega_1) = \frac{\partial}{\partial U_2}\ln(\Omega_2)$$

But the logarithm of the MULTIPLICITY FUNCTION Ω is the definition of the ENTROPY σ, so the equation can be simplified further to read

$$\frac{\partial\sigma_1}{\partial U_1} = \frac{\partial\sigma_2}{\partial U_2} \tag{9}$$

where of course we are assuming that all the other parameters (like N_1 and N_2) are held constant.

Note that we have recovered, by strict mathematical methods, the same criterion dictated by common sense earlier. The only advantage of the formal derivation is that it is rigourous, general and involves no questionable assumptions.[14]

15.4.3 Thermal Equilibrium

Eq. (9) establishes the criterion for the MOST PROBABLE CONFIGURATION — *i.e.* the value of $\hat{U}_1$ for which the combined systems have the maximum total entropy, the maximum total number of accessible states and the highest probability. This also defines the condition of THERMAL EQUILIBRIUM between the two systems — that is, if $U_1 = \hat{U}_1$, any flow of energy from $\mathcal{S}_1$ to $\mathcal{S}_2$ or back will *lower* the number of accessible states and will therefore be *less likely* than the configuration[15] with $U_1 = \hat{U}_1$. Therefore if we leave the systems alone and come back

[14] Or, at least, none that are readily apparent....

[15] Note the distinction between the words *configuration* and *state*. The latter implies we specify *everything* about the system — all the positions and velocities of all its particles, *etc.* — whereas the former refers only to some gross *overall macroscopic* specification like the total energy or how it is split up between two subsystems. A *state* is *completely specified* while a *configuration* is only *partly specified*.

later, we will be *most likely* to find them in the "configuration" with $\hat{U}_1$ in system $\mathcal{S}_1$ and $(U - \hat{U}_1)$ in system $\mathcal{S}_2$.

This seems like a pretty weak statement. Nothing certain, just a *bias* in favour of $\hat{U}_1$ over other possible values of U_1 all the way from zero to U. That is true. STATISTICAL MECHANICS has nothing whatever to say about what *will* happen, only about what is *likely* to happen — and *how* likely! However, when the numbers of particles involved become very large (and in Physics they do become *very* large), the fractional width of the binomial distribution [Eq. (2)] becomes *very* narrow, which translates into a probability distribution that is *incredibly sharply peaked* at $\hat{U}_1$. As long as energy conservation is not violated, there is *nothing but luck* to prevent all the air molecules in this room from vacating the region around my head until I expire from asphyxiation. However, I trust my luck in this. A quotation from Boltzmann confirms that I am in distinguished company:

> *"One should not imagine that two gases in a 0.1 liter container, initially unmixed, will mix, then again after a few days separate, then mix again, and so forth. On the contrary, one finds ... that not until a time enormously long compared to $10^{10^{10}}$ years will there be any noticeable unmixing of the gases. One may recognize that this is practically equivalent to never...."*
>
> — L. Boltzmann

15.4.4 Inverse Temperature

What do we expect to happen if the systems are *out of equilibrium*? For instance, suppose system $\mathcal{S}_1$ has an energy $U_1 < \hat{U}_1$. What will random chance "do" to the two systems? Well, a while later it would be more likely to find system $\mathcal{S}_1$ with the energy $\hat{U}_1$ again. That is, energy would tend to "spontaneously flow" from system $\mathcal{S}_2$ into system $\mathcal{S}_1$ to *maximize the total entropy*.[16] Now stop and think: is there any *familiar, everyday property* of physical objects that governs whether or not internal energy (HEAT) will spontaneously flow from one to another? Of course! Every child who has touched a hot stove knows that heat flows spontaneously from a *hotter* object [like a stove] to a *cooler* object [like a finger]. We even have a *name* for the quantitative measure of "hotness" — we call it TEMPERATURE.

Going back to Eq. (9), we have a mathematical expression for the criterion for THERMAL EQUILIBRIUM, whose familiar everyday-life equivalent is to say that *the two systems have the same temperature*. Therefore we have a compelling motivation to associate the quantity $\frac{\partial \sigma}{\partial U}$ for a given system with the TEMPERATURE of that system; then the equation reads the same as our intuition. The only problem is that we expect heat to flow *from* a system at *high* temperature *to* a system at *low* temperature; let's check to see what is predicted by the mathematics.[17] Let's suppose that for some initial value of $U_1 < \hat{U}_1$ we have

$$\frac{\partial \sigma_1}{\partial U_1} > \frac{\partial \sigma_2}{\partial U_2}.$$

Then adding a little extra energy dU to $\mathcal{S}_1$ will increase σ_1 by *more* than we decrease σ_2 by subtracting the same dU from $\mathcal{S}_2$ [which we must do, because the total energy is conserved]. So the *total* entropy will *increase* if we move a little energy *from* the system with a *smaller* $\frac{\partial \sigma}{\partial U}$ *to* the system with a *larger* $\frac{\partial \sigma}{\partial U}$. The region of *smaller* $\frac{\partial \sigma}{\partial U}$ must therefore be *hotter* and the region of *larger* $\frac{\partial \sigma}{\partial U}$ must be *cooler*. This is the *opposite* of what we expect of TEMPERATURE, so we do the obvious: we define $\frac{\partial \sigma}{\partial U}$ to be the INVERSE TEMPERATURE of a system:

$$\frac{\partial \sigma}{\partial U} \equiv \frac{1}{\tau} \qquad (10)$$

where (at last) τ is the TEMPERATURE of the system in question. We can now express Eq. (9) in the form that agrees with our intuition:

Condition of THERMAL EQUILIBRIUM:

$$\tau_1 = \tau_2 \qquad (11)$$

— *i.e.* if the *temperatures* of the two systems are the same, then they will be in *thermal equilibrium* and everything will be most likely to stay pretty much as it is.

As you can see, TEMPERATURE is not quite such a simple or obvious concept as we may have been led to believe! But now we have a universal, rigorous and valid *definition* of temperature. Let's see what use we can make of it.

15.4.5 Units & Dimensions

I have borrowed from several authors the convention of expressing the ENTROPY σ in explicitly *dimensionless* form [the logarithm of a pure number is another pure number]. By the same token, the simple definition of TEMPERATURE τ given by Eq. (10) automatically gives τ dimensions of *energy*, just like U. Thus τ can be

[16] This is the same as maximizing the probability, but from now on I want to use the terminology "maximizing the entropy."

[17] We have already done this once, but it bears repeating! To avoid complete redundancy, this time we will reverse the order of hot and cold.

measured in joules or ergs or other more esoteric units like electron-volts; but we are accustomed to measuring TEMPERATURE in other, less "physical" units called *degrees*. What gives?

The story of how temperature units got invented is fascinating and sometimes hilarious; suffice it (for now) to say that these units were invented *before anyone knew what temperature really was*![18] There are two types of "degrees" in common use: Fahrenheit degrees[19] and Celsius degrees (written °C) which are moderately sensible in that the interval between the freezing point of water (0°C) and the boiling point of water (100°C) is divided up into 100 equal "degrees" [hence the alternate name "Centigrade"]. However, in Physics there are only one kind of "degrees" in which we measure temperature: degrees *absolute* or "Kelvin"[20] which are written just "K" without any ° symbol. One K is the same size as one °C, but the *zero* of the Kelvin scale is at *absolute zero*, the coldest temperature *possible*, which is itself an interesting concept. The freezing temperature of water is at 273.15 K, so to convert °C into K you just add 273.15 degrees. Temperature measured in K is always written T.

What relationship does τ bear to T? The latter had been invented long before the development of Statistical Mechanics and the explanation of what temperature really *was*; but these clumsy units never go away once people have gotten used to them. The two types of units must, of course, differ by some constant conversion factor. The factor in this case is $k_{\rm B}$, BOLTZMANN'S CONSTANT:

$$\tau \;\equiv\; k_{\rm B}\, T \quad \text{where}$$
$$k_{\rm B} \;\equiv\; 1.38066 \times 10^{-23} \text{ J/K} \tag{12}$$

[18]Well, to be fair, people had a pretty good working knowledge of the *properties* of temperature; they just didn't have a *definition* of temperature in terms of nuts-and-bolts mechanics, like Eq. (10).

[19]These silly units were invented by an instrument maker called Fahrenheit [1686-1736] who was selling thermometers to meteorologists. He picked *body temperature* [a handy reference, constant to the precision of his measurements] for one "fiducial" point and for the other he picked the *freezing point of saturated salt water*, presumably from the North Sea. Why not pure water? Well, he didn't like negative temperatures [neither do we, but he didn't go far enough!] so he picked a temperature that was, for a meteorologist, as cold as was worth measuring. [Below that, presumably, it was just "damn cold!"] Then he (sensibly) divided up the interval between these two fiducials into $96 = 64 + 32$ equal "degrees" [can you see why this is a pragmatic choice for the number of divisions?] and *voilá!* he had the Fahrenheit temperature scale, on which pure water freezes at 32°F and boils at 212°F. A good system to forget, if you can.

[20]Named after Thomson, Lord Kelvin [1852], a pioneer of thermodynamics.

By the same token, the "conventional entropy" S defined by the relationship

$$\frac{1}{T} \;=\; \frac{\partial S}{\partial U} \tag{13}$$

must differ from our dimensionless version σ by the same conversion factor:

$$S \;\equiv\; k_{\rm B}\, \sigma \tag{14}$$

This equivalence completes the definition of the mysterious entities of classical thermodynamics in terms of the simple "mechanical" paradigms of Statistical Mechanics. I will continue to use σ and τ here.

15.4.6 A Model System

Some of the more peculiar properties of temperature can be illustrated by a simple example:

Certain particles such as electrons have "spin $\frac{1}{2}$" which (it turns out) prevents their spins from having any orientation in a magnetic field $\vec{\boldsymbol{B}}$ other than parallel to the field ("spin up") or antiparallel to it ("spin down"). Because each electron has a magnetic moment $\vec{\mu}$ (sort of like a tiny compass needle) lined up along its spin direction, there is an energy $\varepsilon = -\vec{\mu} \cdot \vec{\boldsymbol{B}}$ associated with its orientation in the field.[21] For a "spin up" electron the energy is $\varepsilon_\uparrow = +\mu B$ and for a "spin down" electron the energy is $\varepsilon_\downarrow = -\mu B$.

Consider a *system* consisting of N electrons in a magnetic field and neglect all other interactions, so that the total energy U of the system is given by

$$U = (N_\uparrow - N_\downarrow)\, \mu B$$

where $N_\uparrow$ is the number of electrons with spin up and $N_\downarrow$ is the number of electrons with spin down. Since $N_\downarrow = N - N_\uparrow$, this means

$$U = (2N_\uparrow - N)\, \mu B \quad \text{or}$$

$$N_\uparrow = \frac{N}{2} + \frac{U}{2\mu B} \tag{15}$$

— that is, $N_\uparrow$ and U are basically the same thing except for a couple of simple constants. As $N_\uparrow$ goes from 0 to N, U goes from $-N\mu B$ to $+N\mu B$.

This system is another example of the *binomial distribution* whose multiplicity function was given by Eq. (1),

[21]The rate of change of this energy with the angle between the field and the compass needle is in fact the torque which tries to align the compass in the Earth's magnetic field, an effect of considerable practical value.

with $N_\uparrow$ in place of n. This can be easily converted to $\Omega(U)$. The *entropy* $\sigma(U)$ is then just the logarithm of $\Omega(U)$, as usual. The result is plotted in the top frame of Fig. 15.2 as a function of energy. Note that the entropy has a *maximum* value for equal numbers of spins up and down — *i.e.* for zero energy. There must be some such peak in $\sigma(U)$ whenever the energy is *bounded above* — *i.e.* whenever there is a *maximum possible energy* that can be stored in the system. Such situations do occur [this is a "real" example!] but they are rare; usually the system will hold as much energy as you want.

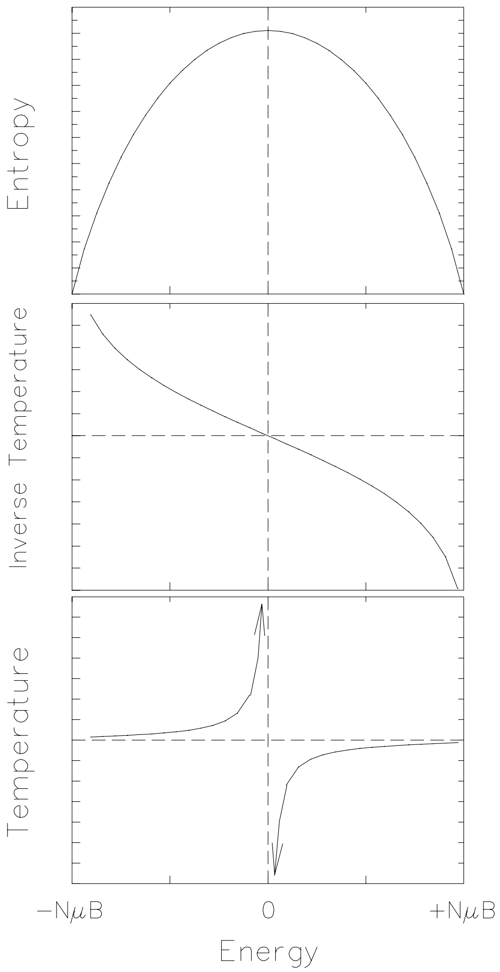

Figure 15.2 Entropy, inverse temperature and temperature of a system consisting of $N = 32$ spin-$\frac{1}{2}$ particles (with magnetic moments μ) in a magnetic field B.

Negative Temperature

The "boundedness" of U and the consequent "peakedness" of $\sigma(U)$ have some interesting consequences: the *slope* of $\sigma(U)$ [which, by Eq. (10), defines the *inverse temperature*] decreases steadily and smoothly over the entire range of U from $-N\mu B$ to $+N\mu B$, going through zero at $U = 0$ and becoming negative for positive energies. This causes the *temperature* itself to diverge toward $+\infty$ as $U \to 0$ from the left and toward $-\infty$ as $U \to 0$ from the right. Such discontinuous behaviour is disconcerting, but it is only the result of our insistence upon thinking of τ as "fundamental" when in fact it is $1/\tau$ that most sensibly defines how systems behave. Unfortunately, it is too late to get thermometers calibrated in inverse temperature and get used to thinking of objects with *lower* inverse temperature as being *hotter*. So we have to live with some pretty odd properties of "temperature."

Consider, for instance, the whole notion of *negative temperature*, which is actually exhibited by this system and can actually be prepared in the laboratory.[22] What is the behaviour of a system with a negative temperature? Our physical intuition, which in this case is trustworthy, declares that one system is *hotter* than another if, when the two are placed in thermal contact, heat energy spontaneously flows *out of* the first *into* the second. I will leave it as an exercise for the reader to decide which is most hot — infinite positive temperature or finite negative temperature.

15.5 Time & Temperature

Let's do the following *Gedankenexperiment*: Suppose I show you a movie of a swimming pool full of waves and splashes; suddenly (in the movie) all the waves come together and squirt a diver out of the pool. She flies gracefully through the air to land on the diving board while the pool's surface has miraculously returned to mirror smoothness. What is wrong with this picture? Wait! Before you answer, you also get the following movie: A box full of 100 black and 100 white marbles sits on a table; the marbles are arranged randomly. An anonymous assistant picks up the box, closes the lid, shakes the box for a while, puts it down and opens the lid. All the white marbles are now on the left side and all the black marbles are on the right side. Why do you keep thinking there is a problem? Try this: The same box, the same assistant, the same story; except this time there are only 4 marbles, two of each. Not so sure, hmmm? How about 2 marbles, one black and one white? Now we can't tell a thing about whether the movie is being shown forward or backward, right? What is going on here?

Our concept of the "arrow of time" is somehow bound

[22] [by reversing the direction of the magnetic field before the spins have a chance to react]

up with statistical mechanics and is alarmingly *fragile* — we can lose our bearings completely just by confining our attention to *too small* a system! As we will see later, the "fundamental" laws governing the microscopic interactions of matter will be no help at all in clarifying this mystery.

15.6 The Boltzmann Distribution

In defining the concept of *temperature*, we have examined the behaviour of systems in thermal contact (*i.e.* able to exchange energy back and forth) when the total energy U is fixed. In the real world, however, it is not often that we *know* the total energy of an arbitrary system; there is no "energometer" that we can stick into a system and read off its energy! What we often *do* know about a system it its *temperature*. To find this out, all we have to do is stick a calibrated thermometer into the system and wait until equilibrium is established between the thermometer and the system. Then we read its temperature off the thermometer. So what can we say about a *small* system[23] $\mathcal{S}$ (like a single molecule) in thermal equilibrium with a *large* system (which we usually call a "heat reservoir" $\mathcal{R}$) at temperature $\tau = k_{\rm B}T$?

Well, the small system can be in any one of a large number of fully-specified states. It is convenient to be invent an abstract *label* for a given fully-specified state so that we can talk about its properties and probability. Let's call such a state $|\alpha\rangle$ where α is a "full label" — *i.e.* α includes *all the information there is* about the state of $\mathcal{S}$. It is like a complete list of which car is parked in which space, or exactly which coins came up heads or tails in which order, or whatever. For something simple like a single particle's spin, α may only specify whether the spin is up or down. Now consider some particular fully-specified state $|\alpha\rangle$ whose *energy* is ε_α. As long as $\mathcal{R}$ is *very big* and $\mathcal{S}$ is *very small*, $\mathcal{S}$ can — and sometimes will — absorb from $\mathcal{R}$ the energy ε_α required to be in the state $|\alpha\rangle$, no matter how large ε_α may be. However, you might expect that states with really *big* ε_α would be excited somewhat less often than states with *small* ε_α, because the extra energy has to come from $\mathcal{R}$, and every time we take energy out of $\mathcal{R}$ we decrease its entropy and make the resultant configuration that much less probable. You would be right. Can we be quantitative about this?

Well, the *combined system* $\{\mathcal{S} + \mathcal{R}\}$ has a multiplicity function Ω which is the *product* of the multiplicity function $\Omega_{\mathcal{S}} = 1$ for $\mathcal{S}$ [which equals 1 because we have already postulated that $\mathcal{S}$ is in a specific *fully specified* state $|\alpha\rangle$] and the multiplicity function $\Omega_{\mathcal{R}} = e^{\sigma_{\mathcal{R}}}$ for $\mathcal{R}$:

$$\Omega \; = \; \Omega_{\mathcal{S}} \times \Omega_{\mathcal{R}} \; = \; 1 \times e^{\sigma_{\mathcal{R}}}$$

Moreover, the *probability* $\mathcal{P}_\alpha$ of finding $\mathcal{S}$ in state $|\alpha\rangle$ with energy ε_α will be proportional to this net multiplicity:

$$\mathcal{P}_\alpha \; \propto \; e^{\sigma_{\mathcal{R}}}$$

We must now take into account the effect on this probability of removing the energy ε_α from $\mathcal{R}$ to excite the state $|\alpha\rangle$.

The energy of the reservoir $\mathcal{R}$ before we brought $\mathcal{S}$ into contact with it was U. We don't need to know the value of U, only that it was a fixed starting point. The entropy of $\mathcal{R}$ was then $\sigma_{\mathcal{R}}(U)$. Once contact is made and an energy ε_α has been "drained off" into $\mathcal{S}$, the energy of $\mathcal{R}$ is $(U - \varepsilon_\alpha)$ and its entropy is $\sigma_{\mathcal{R}}(U - \varepsilon_\alpha)$.

Because ε_α is *so tiny* compared to U, we can treat it as a "differential" of U (like "dU") and estimate the resultant *change* in $\sigma_{\mathcal{R}}$ [relative to its old value $\sigma_{\mathcal{R}}(U)$] in terms of the *derivative* of $\sigma_{\mathcal{R}}$ with respect to energy:

$$\sigma_{\mathcal{R}}(U + dU) \; = \; \sigma_{\mathcal{R}}(U) \; + \; \left(\frac{\partial \sigma_{\mathcal{R}}}{\partial U}\right) \cdot dU$$

or in this case (with $dU \equiv -\varepsilon_\alpha$)

$$\sigma_{\mathcal{R}}(U - \varepsilon_\alpha) \; = \; \sigma_{\mathcal{R}}(U) \; - \; \left(\frac{\partial \sigma_{\mathcal{R}}}{\partial U}\right) \cdot \varepsilon_\alpha$$

But this derivative is by definition the *inverse temperature* of $\mathcal{R}$: $\dfrac{\partial \sigma_{\mathcal{R}}}{\partial U} \equiv \dfrac{1}{\tau}$. Thus

$$\sigma_{\mathcal{R}}(U - \varepsilon_\alpha) \; = \; \sigma_{\mathcal{R}}(U) \; - \; \frac{\varepsilon_\alpha}{\tau}$$

and thus the probability of finding $\mathcal{S}$ in the state $|\alpha\rangle$ obeys

$$\mathcal{P}_\alpha \; \propto \; e^{\sigma_{\mathcal{R}}(U - \varepsilon_\alpha)} \; = \; \exp\left[\sigma_{\mathcal{R}}(U) - \frac{\varepsilon_\alpha}{\tau}\right]$$

$$\text{or} \quad \mathcal{P}_\alpha \; \propto \; e^{\sigma_{\mathcal{R}}(U)} \cdot \exp\left(-\frac{\varepsilon_\alpha}{\tau}\right)$$

Since $e^{\sigma_{\mathcal{R}}(U)}$ is a constant independent of either ε_α or τ, that term will be the same for any state $|\alpha\rangle$ so we may ignore it and write simply

$$\mathcal{P}_\alpha \; \propto \; \exp\left(-\frac{\varepsilon_\alpha}{\tau}\right) \tag{16}$$

This is the famous BOLTZMANN FACTOR that describes exactly how to calculate the *relative probabilities* of different states $|\alpha\rangle$ of a system in thermal contact with a heat reservoir at temperature τ. It is probably the single most *useful* rule of thumb in all of thermal physics.

[23] A "small system" can even be a "particle," since both terms are intentionally vague and abstract enough to mean anything we want!

15.6.1 The Isothermal Atmosphere

The gravitational potential energy of a gas molecule of mass m at an altitude h above sea level is given approximately by $\varepsilon = mgh$, where $g = 9.81$ m/s^2. Here we neglect the decrease of g with altitude, which is a good approximation over a few dozen miles. Next we pretend that the *temperature* of the atmosphere does not vary with altitude, which is untrue, but perhaps relative to 0 K it is not all that silly, since the difference between the freezing (273.15 K) and boiling (373.15 K) points of water is less than 1/3 of their average. For convenience we will assume that the whole atmosphere has a temperature $T = 300$ K (a slightly warm "room temperature").

In this approximation, the probability $\mathcal{P}(h)$ of finding a given molecule of mass m at height h will drop off exponentially with h:

$$\mathcal{P}(h) \;=\; \mathcal{P}(0)\,\exp\left(-\frac{mgh}{\tau}\right)$$

Thus the *density* of such molecules per unit volume and the *partial pressure* p_m of that species of molecule will drop off exponentially with altitude h:

$$p_m(h) \;=\; p_m(0)\,\exp\left(-\frac{h}{h_0}\right)$$

where h_0 is the altitude at which the partial pressure has dropped to $1/e$ of its value $p_m(0)$ at sea level. We may call h_0 the "mean height of the atmosphere" for that species of molecule. A quick comparison and a bit of algebra shows that

$$h_0 = \frac{\tau}{mg}$$

For *oxygen molecules* (the ones we usually care about most) $h_0 \approx 8$ km. For *helium atoms* $h_0 \approx 64$ km and in fact He atoms rise to the "top" of the atmosphere and disappear into interplanetary space. This is one reason why we try not to lose any helium from superconducting magnets *etc.* — helium is a non-renewable resource!

15.6.2 How Big are Atoms?

Wait a minute! How did I calculate h_0? I had to know m for the different molecules, and that requires some knowledge of the *sizes* of atoms — information that has not yet been set forth in this book! In fact, empirical observations about how fast the pressure of the atmosphere *does* drop off with altitude could give a pretty good idea of his big atoms are; this isn't how it was done historically, but let's give it a try anyway:

Suppose that, by climbing mountains and measuring the density of oxygen molecules (O_2) as a function of altitude, we have determined *empirically* that h_0 for O_2 is about 8,000 m. Then, according to this simple model, it must be true that the mass m of an O_2 molecule is about

$$m \;\approx\; \frac{\tau}{h_0\,g} \;=\; \frac{300 \times 1.38 \times 10^{-23}}{8 \times 10^3 \times 9.81}\ \text{kg}$$

$$\text{or} \quad m \;\approx\; 5.3 \times 10^{-26}\ \text{kg}$$

This is a mighty small mass!

Now to mix in just a pinch of actual history: Long ago, chemists discovered (again empirically) that different pure substances combined with other pure substances in fixed ratios of small integers times a certain characteristic mass (characteristic for each pure substance) called its *molecular weight* A. People had a pretty good idea even then that these pure substances were made up of large numbers of identical units called "atoms,"[24] but no one had the faintest idea how *big* atoms were — except of course that they must be pretty small, since we never could see any directly. The number N_0 of molecules in one *molecular weight* of a pure substance was (correctly) presumed to be the same, to explain why chemical reactions obeyed this rule. This number came to be called a "mole" of the substance. For oxygen (O_2), the molecular weight is roughly 32 grams or 0.032 kg.

If we now combine this conventional definition of a *mole* of O_2 with our previous estimate of the mass of one O_2 molecule, we can estimate

$$N_0 \;\approx\; \frac{0.032}{5.3 \times 10^{-26}} \;\approx\; 6 \times 10^{23}$$

The exact number, obtained by quite different means, is

$$N_0 \equiv 6.02205 \times 10^{23} \tag{17}$$

molecules per mole. This is known as AVOGADRO'S NUMBER.

Turning the argument around, the mass of a molecule can be obtained from its molecular weight A as follows: One *mole* of any substance is defined as a mass $A \times$ 1 gram, and contains N_0 molecules (or atoms, in the case of monatomic molecules) of the substance. Thus helium, with $A = 4$, weighs 4 gm (or 0.004 kg) per mole containing N_0 atoms, so one He atom weighs $(0.004/N_0)$ kg or 6.6×10^{-27} kg.

[24] I will cover the history of "Atomism" in a bit more detail later on!

15.7 Ideal Gases

We have argued on an abstract basis that the state of highest entropy (and hence the most probable state) for any complicated system is the one whose macroscopic properties can be obtained in the largest possible number of different ways; if the model systems we have considered are any indication, a good rule of thumb for how to do this is to let each "degree of freedom" of the system contain (on average) an equal fraction of the total energy U. We can justify this argument by treating that degree of freedom as a "system" in its own right (almost anything can be a "system") and applying Boltzmann's logic to show that the probability of that microsystem having an energy ε while in thermal equilibrium at temperature τ decays exponentially as $\exp(-\varepsilon/\tau)$. This implies a mean ε on the order of τ, if we don't quibble over factors comparable to 1.

The Equipartition Theorem, which is more rigourously valid than the above hand-waving would suggest,[25] specifies the factor to be exactly 1/2:

> *A system in thermal equilibrium with a heat reservoir at temperature τ will have a* **mean energy** *of $\frac{1}{2}\tau$ per degree of freedom.*

In an ideal monatomic gas of N atoms at temperature τ each atom has three degrees of freedom: left–right (x), back–forth (y) and up–down (z). Thus the average internal energy of our monatomic ideal gas is

$$U = \frac{3}{2} N \tau \tag{18}$$

[25] If you want the details, here they are: Suppose that p_i is the CANONICAL MOMENTUM characterizing the i^{th} degree of freedom of a system and that $\varepsilon(p_i) = bp_i^2$ is the energy associated with a given value of p_i. Assume further that p_i can have a *continuous* distribution of values from $-\infty$ to $+\infty$. Then the probability of p_i having a given value is proportional to $\exp(-bp_i^2/\tau)$ and therefore the average energy associated with that degree of freedom is given by

$$\langle \varepsilon(p_i) \rangle = \frac{\int_{-\infty}^{+\infty} bp_i^2 e^{-bp_i^2/\tau} dp_i}{\int_{-\infty}^{+\infty} e^{-bp_i^2/\tau} dp_i}$$

These definite integrals have "well known" solutions:

$$\int_{-\infty}^{+\infty} x^2 e^{-ax^2} dx = \frac{1}{2}\sqrt{\frac{\pi}{a^3}}, \qquad \int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}},$$

where in this case $a = b/\tau$ and $x = p_i$, giving

$$\langle \varepsilon(p_i) \rangle = \frac{\tau}{2}. \qquad \mathcal{QED}$$

In spite of the simplicity of the above argument[26] this is a profound and useful result. It tells us, for instance, that the energy U of an ideal gas *does not depend upon its pressure*[27] p! This is not strictly true, of course; interactions between the atoms of a gas make its *potential energy* different when the atoms are (on average) close together or far apart. But for most gases at (human) room temperature and (Earth) atmospheric pressure, the ideal-gas approximation is extremely accurate!

It also means that if we change the temperature of a container of gas, the *rate of change* of the internal energy U *with temperature*, which is the definition of the HEAT CAPACITY

$$C \equiv \frac{\partial U}{\partial T}, \tag{19}$$

is extremely simple: since $\tau \equiv k_B T$ and $U = \frac{3}{2}N\tau$, $U = \frac{3}{2}Nk_BT$ and so the heat capacity of an ideal gas is *constant*:

$$C \text{ [ideal gas]} = \frac{3}{2} N k_B \tag{20}$$

Now let's examine our gas from a more microscopic, "mechanical" point of view: picture *one atom* bouncing around inside a cubical container which is a length L on a side. In the "ideal" approximation, atoms never hit each other, but only bounce off the walls, so our consideration of a *single* atom should be independent of whether or not there are other atoms in there with it. Suppose the atom in question has a velocity $\vec{v}$ with components v_x, v_y and v_z along the three axes of the cube.

Thinking only of the wall at the $+x$ end of the box, our atom will bounce off this wall at a rate $1/t$ where t is the time taken to travel a distance $2L$ (to the far wall and back again) at a speed v_x: $t = 2L/v_x$. We assume *perfectly elastic* collisions — *i.e.* the magnitude of v_x does not change when the particle bounces, it just changes sign. Each time our atom bounces off the wall in question, it imparts an *impulse* of $2mv_x$ to that wall. The average *impulse per unit time* (*force*) exerted on said wall by said atom is thus $F_1 = 2mv_x/t$ or $F_1 = mv_x^2/L$. This force is (on average) spread out all over the wall, an area $A = L^2$, so that the *force per unit area* (or *pressure*) due to that one particle is given by $p_1 = F_1/A = mv_x^2/L^3$. Since $L^3 = V$, the volume of the container, we can write $p_1 = mv_x^2/V$ or

$$p_1 V = m v_x^2$$

[26] We can, of course, make the explanation more elaborate, thus satisfying both the demands of rigourous logic and the Puritan conviction that nothing of real value can be obtained without hard work. I will leave this as an exercise for other instructors.

[27] Unfortunately, we use the same notation (p) for both *momentum* and *pressure*. Worse yet, the notation for *number density* (number of atoms per unit volume) is n. Sorry, I didn't set up the conventions.

The average pressure p exerted by *all N atoms together* is just N times the mean value of p_1: $p = N\langle p_1 \rangle$, where the "$\langle \cdots \rangle$" notation means *the average of* the quantity within the angle brackets. Thus

$$pV = N m \langle v_x^2 \rangle \tag{21}$$

Now, the kinetic energy of our original atom is explicitly given by

$$\frac{1}{2}mv^2 = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2)$$

since $\vec{v}$ is the vector velocity. We expect each of the *mean square* velocity components $\langle v_x^2 \rangle$, $\langle v_y^2 \rangle$ and $\langle v_z^2 \rangle$ to average about the same in a random gas, so each one has an average value of $\frac{1}{3}$ of their sum.[28] Thus $\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle = \frac{1}{3}\langle v^2 \rangle$ and the mean kinetic energy of a *single* particle is $U_1 = \frac{3}{2}\, m \langle v_x^2 \rangle$. The kinetic energy of all N atoms is just $U = NU_1$, or

$$U = \frac{3}{2} N m \langle v_x^2 \rangle \tag{22}$$

But according to Eq. (18) we have $U = \frac{3}{2} N\tau$; so we may write[29]

$$m \langle v_x^2 \rangle = \tau \tag{23}$$

Combining Eqs. (21) and (23), we obtain the famous IDEAL GAS LAW:

$$pV = N\tau \tag{24}$$

Despite the flimsiness of the foregoing arguments, the IDEAL GAS LAW is a quantum mechanically correct description of the interrelationship between the pressure p, the volume V and the temperature $\tau \equiv k_{\rm B}T$ of an ideal gas of N particles, as long as the only way to store energy in the gas is in the form of the kinetic energy of individual particles (usually atoms or molecules). Real gases can also store some energy in the form of rotation or vibration of larger molecules made of several atoms or in the form of potential energies of interaction (attraction or repulsion) between the particles themselves. It is the latter interaction that causes gases to spontaneously condense, below a certain *boiling point* T_b, into *liquids* and, at a still lower temperature T_m (called the *melting point*), into *solids*. However, in the gaseous phase even carbon [vaporized diamond] will behave very much like an ideal gas at sufficiently high temperature and low pressure. It is a pretty good Law!

[28]We may say that the average kinetic energy "stored in the x *degree of freedom*" of an atom is $\frac{1}{2}m\langle v_x^2 \rangle$.

[29]This is equivalent to saying that the average energy stored in the x degree of freedom of one atom [or, for that matter, in any other degree of freedom] is $\frac{1}{2}\tau$ — which is just what we originally claimed in the EQUIPARTITION THEOREM. We could have just jumped to this result, but I thought it might be illuminating to show an explicit argument for the equality of the mean energies stored in several different degrees of freedom.

15.8 Things I Left Out

As you can tell by the length of this chapter, I find it hard to stop talking about this wonderful subject. Thermal Physics is like an old but vibrantly modern city with a long, fabulous and meticulously preserved history: around every corner there is a host of fascinating shops, theatres, galleries and restaurants offering the latest goodies from a cosmopolitan state of the art, intermixed with libraries and museums that tell stories of heroic acts and world-changing events. "Shop till you drop!" Still, I have to stop somewhere.

The foregoing has been a rather unusual *introduction* to Thermal Physics. I have completely left out THE LAWS OF THERMODYNAMICS — the traditional *starting point* for the subject — in favour of a strictly conceptual (though often painfully formal, I know) explanation of the meaning of entropy and temperature, in the conviction that these notions can be *generalized* to provide tools for quantitative analysis of random statistical processes in realms where no one ever dreamed of applying the paradigms of Physics. In my zeal to convey this conviction, I have also omitted any discussion of the profound *practical applications* of Thermodynamics, like ENGINES and REFRIGERATORS. Worst of all, I have not told any stories of the bizarre spontaneous behaviour of large numbers of similar atoms under different conditions of temperature and pressure — the so-called EQUATIONS OF STATE and PHASE DIAGRAMS of gases, liquids and solids, from FERMI GASES to SUPERFLUIDS and SUPERCONDUCTORS. Part of the reason for this is that you need a bit more introduction to the phenomenology of Physics — QUANTUM MECHANICS in particular — before you can fully appreciate (or even, in some cases, *describe*) much of the above-mentioned behaviour. All I can hope to have done in this *HyperReference* is to have unlocked the door (and perhaps opened it a crack) to a world of wonder and magic where analytical thinking and mathematics play the role of spells and incantations. I urge you to continue this adventure beyond the limits (and end) of this *HyperReference*!