

# Why Things Happen

Robert P. Bauman  
Professor of Physics Emeritus  
University of Alabama at Birmingham

“Once or twice I have been provoked and have asked the company how many of them could describe the Second Law of Thermodynamics. The response was cold; it was also negative. Yet I was asking something which is about the scientific equivalent of: *Have you read a work of Shakespeare?*”  
C.P. Snow, *The Two Cultures*

Imagine running a business without a bookkeeper. You might be able to stay in business for awhile, and perhaps even realize a profit – at least, until the government shut you down for failure to provide information for taxes. But it would be a foolish way to operate a business.

In much the same way, if you are concerned with things happening – operating machinery, heating or cooling goods or spaces, or converting one substance into another, for example—then it is important to understand at least the rudiments of the accounting system that tells you what changes are possible and what the costs are for changes. The accounting system that describes how the universe operates, and why, is called *thermodynamics*.

Although Snow’s complaint was published in 1959, it is not clear that during ensuing years the “intelligent layman” has been given a map to even the most basic ideas of thermodynamics. The thermodynamics story involves some unfamiliar terminology and information, along with some very familiar ideas, but it is not complex in comparison with other ideas regularly encountered today. (On a first reading, Shakespeare may seem obscure, but nonetheless worthwhile.) The ideas of thermodynamics, absorbed at a leisurely pace, provide fascinating insights into why things happen around us.

There are just two really important ideas that are the foundation of all of thermodynamics, or “thermo”. These two ideas are *energy* and *probability*. We consider them in turn.

Before we begin, however, we should look at a more recent generalization. Late in his career, Richard Feynman declared that, if he could impart to succeeding generations only one idea, he would tell them that everything is made of atoms. Accustomed though we are to the thought of molecules, atoms, and subatomic particles, this seems to reflect a strange bias. Surely when we examine matter at the sub-microscopic level we know we will encounter atoms, but is that really so important for “everyday” physics? Apart from a brief discussion on the kinetic theory of gases, most introductory physics textbooks largely ignore the structure of matter, at least before the final chapters.

Certainly if we wish to know how the universe operates, and why, thermodynamics is the effective approach. It is one of the axioms of thermodynamics that the detailed structure of matter is usually ignorable. It is a rare occasion that the atomic-molecular structure shows up in thermodynamic equations, except in its influence on certain constants, such as predictions of heat capacities and to find  $n$ , the number of moles (*e.g.*, for the gas equation) from masses.

Yet, in a deep sense, Feynman’s guide will prove valuable, even when we believe we are concerned only with bulk properties of matter. We will discover that link in the thermodynamic

properties that relate to probability.

## 1. Energy

The fundamental principle of accounting is that it should be possible to keep track of money as it passes into or out of the company. But this works only if “money” is defined broadly. The accountant must include everything of value. He must include not just bills and coins, but checks, credit and debit slips, raw materials, and finished goods; in short, anything that may be traded for, or converted to, “money”. Accounting, and business, is made more complicated because “value” can appear and disappear. A natural disaster, a scandal, or a false rumor can suddenly decrease the value of a business.

In much the same way, the accounting system for keeping track of changes, of any kind, must follow *energy* and its transfer. Again, it is necessary to follow energy in all its forms. Unlike “value”, energy is not subject to unexpected gains or losses, but like “value”, energy does like to run away and hide.

Because some forms of energy were not recognized for many years, it took a long time to recognize the fundamental rule for energy.

*As energy is transformed from one type to another, and transferred back and forth between objects, the total amount of energy does not change.*

This is so important, and so fundamental, that we call it our *first law of thermodynamics*.<sup>1</sup>

Before a bookkeeper can track movement of money to and from the company, quite obviously the bookkeeper must know what constitutes the company. Similarly, in keeping track of how energy is transferred to and from an object, the thermodynamic bookkeeper must know what is “inside” and what is “outside”. We could just refer to “the object”, or call it “it”, but neither label would be very satisfactory. By general agreement, we call the object, or the collection of objects, of interest “the system”. Everything else is called “the surroundings”. The surroundings may or may not “surround” the system. They may be intermixed. Because *system* and *surroundings* are simply labels, we should feel free to interchange those labels at any time. That is sometimes quite helpful. (From time to time we still choose conventional labels such as *objects* or *bodies* when we wish to emphasize characteristics of a homogeneous or clearly defined system.)

Any sort of change, whatever, is called a “process”. Generally we are interested primarily in a process, or change, within the system. But usually any change in the system involves some change in the surroundings. Consider a simple system such as a block of copper. To warm the block of copper, energy must flow from outside the block – that is, from the surroundings. If the block of copper expands, the volume of the block increases but the volume of the surroundings

---

<sup>1</sup> There is frequent confusion between the first law of thermodynamics, which states that the total amount of energy (for system plus surroundings) is constant, and the “first-law equation”, which says that the change of energy,  $\Delta E$ , for any system is the amount of energy put into the system as “heat” and the amount of energy put into the system as “work”, typically written as  $\Delta E = Q + W$ . This first-law equation is incomplete, and therefore not necessarily valid, but describes some of the ways that energy may be transferred between a system and its surroundings.

must then simultaneously decrease. A description of a process tells what happens to the system, and how, but it must therefore tell us at least something about changes in the surroundings.

Now we can refine our definition of the surroundings by saying that the surroundings are everything in the universe, *except the system*, that might be affected by the process under consideration.

Energy is a familiar idea. We eat food to get energy for our daily activities. We stand in front of a campfire, on a cool evening, to absorb energy from the fire to get warm. We store energy in a battery that will crank our car engine or run a golf cart. How, then, may we define energy?

The concept of energy began with *energy of motion*, generally now called *kinetic energy*. We may define energy as energy of motion, or anything that energy of motion may be converted to, or obtained from. A spinning weathervane has *rotational energy*. A spring with mass attached, bouncing up and down, has *vibrational energy*. There is *potential energy* associated with a skier poised at the top of a slope. A poker pulled red hot from the fire has *thermal energy*. The hot Sun transmits energy to us, across empty space, as *radiant energy*. A battery stores energy as *chemical energy*, which may, upon demand, be converted to *electrical energy*. Much of the energy in a block of wood or of copper is stored as *nuclear energy*.

It is easy to see that a moving car or baseball has kinetic energy. It is generally not easy to see energy totally contained within a body, which we call *internal energy*. Internal energy includes chemical energy and nuclear energy. It also includes *thermal energy*. Thermal energy is the portion of the total internal energy that changes when the temperature changes (or in a process such as melting or boiling at constant temperature). The thermal energy may come from outside the body (from the surroundings) or from inside the body by conversion of other forms of internal energy. So a thermometer must measure something that is related to thermal energy.

Thermal energy is essentially just kinetic energy, rotational energy, vibrational energy, and potential energy of individual atoms or molecules inside the material of the system. Because there are so many molecules in any object of ordinary size, there are many “bits” or “lumps” of energy constantly being traded among the molecules. We will find this is the important clue to the predicted behavior of thermal energy, which among other properties causes it to flow from one region, or object, to another, much like a liquid.<sup>2</sup>

Building blocks do not flow. Lumps of coal do not flow very well. Sand flows much better, but not as well as water. The critical difference is size. Because the “lumps” of energy are extremely small, energy flows very well, like water consisting of very small molecules not tightly bound to each other (as they would be in ice).

Imagine a lake, filled with energy that can flow like water. That lake is connected by a stream to another lake, filled with a different kind of energy that also can flow like water. For example, one lake might be imagined to be filled with the kinetic energy of molecules (*e.g.*, along the *x* axis). Another lake then might be imagined to be filled with the rotational energy of

---

<sup>2</sup> This similarity of flowing thermal energy to the flow of a liquid is so strong that it led Lavoisier, in the 18<sup>th</sup> century, to the conclusion that the quantity flowing could be considered a material substance that he called “heat” or, in French, “caloric”. Shortly afterward Benjamin Thompson, Count Rumford, showed that heat or caloric could not be a material substance, but his work was largely ignored for nearly half a century.

molecules (*e.g.*, about the  $x$  axis). And, with certain restrictions (to low-frequency vibrations, hence small lumps of vibrational energy) another set of (twin) lakes might be imagined to be filled with the vibrational energies (kinetic and potential energies of vibration) of molecules. For all the lakes that are connected, none can rise any higher than the others. (Lakes that are not connected may have their own levels, which may be higher or lower than that of the interconnected lakes.) There is a common height to all the interconnected energy levels. And even though the different lakes contain different kinds of energy, we know from our definition of energy that these forms can change into one another, so there is no problem of incompatibility between the contents of different lakes.

The higher the energy level within any object, the more likely it is that energy will flow from that body to another body. We say that of two objects, the one with the higher energy level is *warmer*. In other words, it is the common “energy height” (of the interconnected “lakes”) that a thermometer measures.

Because each energy lake has nearly the same cross-sectional shape and size, we can often make a good estimate of how much thermal energy a body contains just by knowing how many of the energy lakes it contains, but there are complications involved in changes of shape of lakes at the bottom. In any case, we can easily measure the important quantity, which is the *capacity* for thermal energy, at any selected temperature. We put a known quantity of thermal energy into a known quantity of substance (*e.g.*, water) and measure the change in temperature. Then

$$\text{amount of energy added} = \text{“heat capacity”} \times \text{temperature change}$$

(Heat capacity per unit of mass, or weight, is also called “specific heat”, which meant originally the ratio of the heat capacity of some other substance to the heat capacity of the same mass, or weight, of water.)

Many (but not all) forms of internal energy are in thermal equilibrium. Thus the average amount of kinetic energy (in any one direction, such as the  $x$  axis; that is, the average value of  $\frac{1}{2} m v_x^2$ ) is equal to the average amount of rotational energy (about any one axis) and so forth for any other forms of internal energy that are in equilibrium.

The very important discovery that thermal energy (a portion of the total internal energy) is linked like our imaginary lakes, is called the principle of *equipartition of energy*. That is, thermal energy is divided equally among each form of thermal energy. A thermometer will give the same reading for different bodies in equilibrium, and the heat capacity of any substance is just the number of connected energy lakes, times a universal constant.

Other forms of internal energy, such as the energy of electrons, the chemical energy, and the nuclear energy, are not involved in these equilibria, and therefore do not change when the temperature is raised or lowered. Therefore a thermometer *does not* tell us how much energy is in the glass of water, even though it does give us *some* information about how much thermal energy is in the water. (At very low temperatures, and hence low average energies, or at very high temperatures, and thus high average energies, the calculation methods for average energies are somewhat more complex, but the principle is similar.)

You will probably have recognized that “thermal energy” is often called “heat”, and perhaps you will have recognized that transfer of thermal energy (the amount added to the glass of water, for example) is also called “heat”. And, to cap it off, the temperature is often called “heat” (as in, “It’s not the heat, it’s the humidity”, or “Raise the iron to white heat”.) We will avoid the term “heat” because you wouldn’t have any way of being sure which of the three meanings was

intended. “Heat” is a good catchall term when you don’t need to be specific about meanings. It is not an adequate technical term.<sup>3</sup>

## 2. Probability

A special branch of mathematics, with which we all have some familiarity, is *probability*. It is fundamentally simple for it is based only on counting. It can be difficult because we do not always see how to count. We will start with two relatively transparent examples: counting the ways in which two dice can be thrown, and the values obtained, and the operation of a simple pinball machine.

*Two Dice.* A simple, and familiar, example of probability is illustrated with a pair of dice. For each of the six values that may appear on one die, there are six possible values for the other die. Thus there are 36 (6 times 6) possible outcomes of throwing the pair, but only 11 possible values (sums 2 through 12), as shown in Table 1.

Table 1

Die A	Die B	Sum	Die A	Die B	Sum	Die A	Die B	Sum
1	1	2	3	1	4	5	1	6
1	2	3	3	2	5	5	2	7
1	3	4	3	3	6	5	3	8
1	4	5	3	4	7	5	4	9
1	5	6	3	5	8	5	5	10
1	6	7	3	6	9	5	6	11
2	1	3	4	1	5	6	1	7
2	2	4	4	2	6	6	2	8
2	3	5	4	3	7	6	3	9
2	4	6	4	4	8	6	4	10
2	5	7	4	5	9	6	5	11
2	6	8	4	6	10	6	6	12

What is generally important to the players is the value (the sum of values for the two dice). Table 2 shows the ways in which a given sum may be obtained, in terms of what we may call the “internal states”, or specific choices of values for each die. Clearly, not every sum is equally probable.

The sum of values we can then call (for lack of a better term) the “external state”. The “external state” 7 can be a consequence of any one of the six “internal states”: 1 + 6, 2 + 5, 3 + 4, 4 + 3, 5 + 2, or 6 + 1.

For the relatively few choices afforded by two dice, the range of probabilities is not very large (a factor of 6). If the number of possibilities is substantially increased, the range of probabilities

---

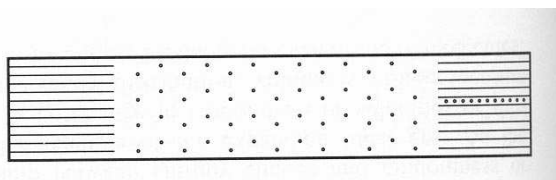
<sup>3</sup> In conventional bookkeeping terms we might say that “heat” is a measure of value per share (temperature) or net value (thermal energy) or earnings per quarter (cash flow or growth), very different quantities but undistinguished by the terminology. The term “heat” does still appear in technical writing, where it is almost always defined as a transfer of thermal energy ( $Q$ , or cash flow), then applied to mean thermal energy (or net value). This causes unnecessary confusion.

will dramatically increase. We will find that most “real life” situations deal with extremely large numbers of possibilities, and therefore extremely large differences in probabilities. Or, more precisely, because a probability of one is generally understood to mean certainty, we should say that there are extremely large differences in the ratios of probabilities.

Table 2

Sum	“Internal States”	No. Int. States
2	1+1	1
3	1+2, 2+1	2
4	1+3, 2+2, 3+1	3
5	1+4, 2+3, 3+2, 4+1	4
6	1+5, 2+4, 3+3, 4+2, 5+1	5
7	1+6, 2+5, 3+4, 4+3, 5+2, 6+1	6
8	2+6, 3+5, 4+4, 5+3, 6+2	5
9	3+6, 4+5, 5+4, 6+3	4
10	4+6, 5+5, 6+4	3
11	5+6, 6+5	2
12	6+6	1

*Pinball Machine.* A very simple pinball machine consists of a set of nails in a flat board, with channels at each end. Start with a set of balls in the center channel on the right, with the board horizontal. Tip the board gently, by raising the right end, and you would probably not be surprised to find the balls do not all go to the center channel on the left, but end up distributed among all the channels on the left. Each ball hits nails as it travels from right to left, and may go either left or right relative to the direction of travel (up or down in the figure), until it strikes the next nail.



Because going left and going right are equally probable, the *most* likely paths lead to the center channel on the left, but the adjacent channels are *almost* as likely. The result, as you may very well have expected, will be a symmetric curve, with the largest number of balls in the center channel on the left and decreasing numbers as you move to the edges. If the number of balls is large and the board is wide, we can predict the shape by a well-known formula called a *binomial distribution*. (As the numbers become infinitely large, the limiting form is known as a *Poisson distribution*.)

To a very good approximation, each collision of a ball with a nail is *elastic*, meaning that the ball does not lose energy to the nail. The path for a ball moving from right to left should look very much like the path for a ball moving from left to right. If we tip the board back (raising the left end), will the balls then all end up where they came from, in the central channel on the right? Of course not, but what is different? Primarily, the starting conditions! Each set of balls starting from a channel on the left will end up distributed among the channels on the right, symmetrically

distributed *with respect to the starting channel on the left*. Balls starting from the third channel on the left will be symmetrically distributed about the third channel on the right, and so forth. Thus the more times the board is tipped, the more nearly equal will be the distribution of the balls in the channels.

*Generalization.* The principle underlying the pinball machine is like that for the pair of dice. There are more ways of distributing the balls equally among all the channels than ways of putting them all in the same channel, so the *most probable* arrangement has equal numbers of balls in each of the channels. If there are 10 channels at one end, then there are 10 choices for placement of the first ball and an independent choice among 10 channels for placement of the second ball for *each* of the first 10 choices. (Compare the two dice. For *each* value on the first die, there are 6 choices for the second die, and thus 36 choices for two dice.) Thus there are 100 choices ( $10 \times 10 = 10^2$ ) for placing two balls, 1000 choices ( $10 \times 10 \times 10 = 10^3$ ) for placing three balls, and if there are 50 balls,  $10 \times 10 \times 10 \times \dots \times 10$  (50 times) for all the balls, or  $10^{50}$ , choices for all the balls. That is a very large number. The likelihood is  $10^{50}$  times as great for the random distribution among all the channels as for all the balls to go into a single channel.

But we can say that better. The chance that all the balls will go into the middle channel is exactly the same as the chance that the balls will be distributed in any one particular way among all the channels. Most of the  $10^{50}$  arrangements of the random distribution will look like there are equal numbers in all the channels, so *collectively* the equal distributions will be almost  $10^{50}$  times as probable as putting all the balls in any one single channel.

The generalization we may draw is as follows. When the number of choices is large, the most likely *collection* of choices will be the random distribution, which will be *very* much more likely than any other arrangement that differs *significantly* from the most likely distribution.

If we look at a sample made up of molecules, then the positions of the molecules are much like putting balls into channels, except that there are many more molecules (than balls in the pinball machine) and many more locations available (than channels in the pinball machine). We don't try to place molecules in various positions by hand, one by one, but the molecules may be able to rearrange themselves. It is thermal energy, or "internal" motion of the particles within the sample object, that allows particles within an object to rearrange themselves between states. Such rearrangements favor a shift in the particles toward a more likely, or more probable, distribution.

In summary:

For any system in which there is sufficient thermal energy to allow the system to rearrange itself frequently, the system, left undisturbed, will always change to a more probable arrangement (or set of arrangements) of its parts (which we call a more probable *state*), unless it is already in a most probable state.

This generalization turns out to be so important that we call it our *second law*.

### 3. Idealizations

The universe is very complex. No two objects (at least of visible size) are truly identical. It would be an impossible job to write a description for the behavior of each and every object, and a thankless job, for the results would be of little value. We therefore write our descriptions for certain *models*, which serve as representations of large groups of objects. We treat all balls as equivalent to an ideal round ball (for many purposes), we treat all gases (for many purposes) as

equivalent to an *ideal gas* that obeys a simple equation relating temperature, pressure, volume, and quantity of gas. If we know how our models behave, then we know how real objects will behave to the extent they resemble the models.

Other models concern *processes*, or changes. One type of model for a process is a change that is *reversible*. To understand reversibility we contrast reversible processes with some that are not reversible.

There are some changes that are just not reversible, or are *irreversible*. Some “Oh-Oh” events, like dropping Grammy’s favorite teapot onto a concrete floor, or eating the piece of pie saved for Great Aunt Emma, cannot be reversed, even with great effort. Such completely irreversible changes are generally of less interest than those that can be reversed, either easily or with substantial effort.

If a ball bearing rolls off the table, it will not return by itself to the table top. Only with appreciable effort by you (for example, bending and lifting) will the ball bearing return to its previous state. Falling off the table is then called a (thermodynamically) *irreversible process*. It cannot be “undone” with vanishingly small effort.

By contrast, if the ball bearing rolls very slowly toward the east on a hard, flat surface, then it can easily be stopped and turned around to roll back toward the west. If we are patient, regardless of the initial speed, a very, very small force toward the west will stop the initial eastward motion and produce the westward motion. Such a limiting case, in which it takes almost no effort to reverse the motion, is of particular importance. We describe a process as *thermodynamically reversible* if an infinitesimal decrease in the driving force, or an infinitesimal increase in the opposing force, is sufficient to reverse the direction of the process.

It can be shown that the more closely we approach reversibility, the longer it will require to carry out the process. True reversibility is essentially unattainable, but reversibility is a very important *model*, or *idealization*.

If a change in a system can be reversed with vanishingly small effort, we can describe the system as being *in equilibrium*. Equal weights in the pans of a well-constructed equal-arm balance are a familiar example of a system (very nearly) in equilibrium. Water in a cup, covered with plastic, is in equilibrium with the water vapor above the liquid. Water molecules are constantly passing from liquid to vapor while others are passing from vapor to liquid.

Reversibility and equilibrium – two descriptions of essentially the same thing – are among the most important idealizations, or approximations, for describing changes in systems and, when desired, changes in the surroundings.

#### **4. The Direction of Change**

Before introducing new ideas, it may be helpful to look at some specific examples of how the probability principle determines the direction of change in processes with which we are already familiar. In particular we will look at some examples of energy dissipation and examples of expansion.

*Energy Dissipation.* A ball placed on a sloping surface will always roll downhill, never uphill. This has led to the frequent assumption that energy is the most important consideration in determining the direction of change. Energy is important. It is true that when a system has available two states of different energy, the lower energy state is more likely to be occupied. Given a certain amount of energy, you could put all of it in one place, giving a state with an

energy difference  $\Delta E$  above a second state, or you could spread that amount of energy elsewhere. Thus finding the system in the upper energy state is less probable than finding it in the lower energy state.<sup>4</sup> But in practice, the ratio of probabilities given by energy differences (the Boltzmann factor) is typically not large enough to explain the direction of change, so we must look beyond energy.

For example, a ball at the top of a hill will roll downhill, whereas a ball at the bottom will not roll uphill. But the ball starting at the top will roll downhill, then roll right back up the next hill *unless or until it loses some of its energy of motion*. We know the ball *will* lose energy of motion, to its own internal energy and to its surroundings, because of *friction*. But that energy is not really lost – we cannot create or destroy energy. That energy “lost” by friction appears as thermal energy. The ball and track get a little bit warmer.

Why does energy move from the rolling ball to the surroundings, but not from the surroundings to the moving ball. To answer this we must look again at probabilities. Energy of rolling represents a single type of motion, or energy storage. There are many other places that rotational energy can go — about  $10^{20}$  or more other places within the ball itself. The energy of rolling, that the ball originally had, will get spread more or less equally among some billions of billions of molecules in the ball and in the track and the one initial rotational motion of the entire ball. Thus, left to itself, the probability of the energy reappearing as rotational energy of the ball is very small, smaller than one by a factor with at least 20 zeroes after the decimal place (smaller by a factor of  $10^{-20}$ ).

This spreading of energy, roughly equally, among all the molecules (*i.e.*, the equipartition of energy), is the reason you cannot drive a steamship across the ocean by withdrawing energy from the ocean water. It is the reason a hot piece of iron will cool down, but a cold piece of iron will not warm up, although either process would be possible without creating or destroying energy. (Either process would satisfy our first law.)

This energy spreading among the enormous number of independent molecules would be sufficient to justify Feynman’s comments concerning the importance of all matter being made up of atoms and molecules. But there is more to the question of why things happen.

*Expansion.* There is another type of process governed by probability that is very important. We may call these processes *expansion processes*. Consider first a gas in a container, but restricted to one-half the volume of the container by a thin diaphragm. What happens to the gas when a hole occurs in the diaphragm? Is it more likely that all the gas will be found in half the container or that the gas will be found equally distributed throughout the entire container? We all know the gas will expand into the whole container. There are twice as many microscopic “states”

---

<sup>4</sup> The difference in probability depends on how large  $\Delta E$  is compared to the average amount of energy available. (As we have previously hinted, the average energy available is  $kT$ , the Boltzmann constant  $k = 1.38 \times 10^{-23}$  joule per kelvin times the absolute temperature. The absolute temperature is zero at the lowest possible temperature. The temperature in kelvin is equal to temperature on the Celsius scale plus 273.15:  $T(\text{K}) = T(^{\circ}\text{C}) + 273.15$ .) Mathematically, the relative probability that the upper state will be occupied is given by the Boltzmann factor, which is an exponential. The mathematical constant  $e = 2.78\dots$  is raised to a power that is minus  $\Delta E$  divided by  $kT$ , hence  $e^{-\Delta E/kT}$ .

(*i.e.*, positions) available for a gas molecule in the whole container as in half the container, and that factor of two, for *each* molecule, gives an enormous advantage to filling the entire container. Remember that independent probabilities are multiplicative, like the 6 x 6 for two dice and 10 x 10 x ... for distributing balls among ten channels. The product 2 x 2 x 2 x ..., or 10<sup>20</sup> factors of two, gives a huge number.<sup>5</sup>

What happens to a small puddle of water on the sidewalk? Small amounts of energy are collected from all around to allow the molecules to break away from each other and evaporate (what we wanted to do with our steamship but couldn't). The molecules spread out into a very large volume from a small volume, so the problem is much like the gas that is initially trapped by a diaphragm, then released to expand.

If you drop a small crystal of salt into a pan of water, the salt dissolves. Why? Salt molecules like each other, so they like to get together to form a crystal. But salt molecules (or the parts, called *ions*, that make up a molecule) are also attracted to water, so there isn't very much difference in energy for the salt whether it sits as a crystal or wanders around through the pan of water. The *expansion* into the volume of the water wins out (if there is enough energy, as there will be at room temperature). We may think of any kind of expansion like a dissipation in space. Instead of a tidy arrangement, things get spread "all over the place", which is a state (or really, a collection of essentially equivalent "internal states") with substantially higher probability.

## 5. Measuring the Increase of Spreading, S

We found we could predict, quite accurately, the probabilities of different values for two dice, simply by counting the numbers of states. Clearly there should be an advantage if we could apply the same technique to more complex problems, involving billions and billions of molecules. Whether we are looking at dissipation of energy among atoms and molecules, or the spreading of atoms and molecules in space, it is the relative number of states that is important, and the number of states (energy or in space) may be measured as a "spreading" function, *S*.

*The Quantity of States, Q.* We will skip the mathematical techniques that allow us to calculate just how many states there are and simply assume that we can find the *quantity of states*, *Q*, among which the system gets spread. We will call this number the "spread" of the system, or rather, we will express *Q* as a power and take just the exponent, which is a very much smaller number. (For example, 1000 = 10 x 10 x 10 = 10<sup>3</sup>. The *logarithm* of 1000, to base 10, is 3, which is the exponent. We can let 3 represent 1000, let 6 represent a million, 9 a billion, and so forth. For practical reasons, arising from calculus, rather than taking powers of 10 we prefer to take powers of *e* = 2.78 ..., but the principle is the same.) Then we multiply by a constant (the Boltzmann constant, *k*) to get suitable units, and represent this "spread" function<sup>6</sup> by *S*. This

---

<sup>2</sup> 10<sup>20</sup> factors of two give 2 raised to the 10<sup>20</sup>, which is 10 raised to the 3 x 10<sup>19</sup>. Thus there are 3 x 10<sup>19</sup> digits in the answer before the decimal point.

<sup>6</sup> The function is written  $S = k \ln Q$ , where *Q* is the number of states,  $\ln Q$  represents the logarithm, to base *e*, of *Q* and  $k = 1.38 \times 10^{-23}$  is the Boltzmann constant. This spread function is called *entropy*. The name was chosen by Clausius from the Greek verb *entrepo* (*εντρεπω*), meaning "to turn," implying "change", and because it somewhat resembles the word *energy*, one of the important quantities involved in finding the number of states. (This *Q* is *not* the same as

spread function then turns out to be a convenient number, in convenient units, for measuring how widely distributed the system is among energy states and volume. The greater the value of  $S$ , the greater the probability the system will be found in such an overall state, so  $S$  is also a measure of probability. It is known by a derived Greek name, *entropy*.

*Measurement of Changes of Entropy.* The actual value of  $Q$  is seldom, if ever, important, but *changes* in  $Q$  are very important, because a system (left alone) will *always* change from its initial state to a *more probable* state (*probably!*). The increase in probability is typically by a factor so large that it is effectively beyond imagination. Hence probability and fact become equivalent statements.

To look for a method of calculating the increase in the spread function, entropy, we might well take advantage of what we saw for the rolling ball. The effective driving force, that prevents the ball from rolling up hill and down indefinitely, is that some of the energy will get converted to thermal energy. That spreads out the energy so much that it cannot get collected together again, without outside intervention (or an even larger probability factor arising from expansion in space). So we start by looking at a measure of the increase in thermal energy in the system.

The best measure of the increase in thermal energy of the system turns out to be the smallest possible amount of thermal energy transferred to the system that will get the system from its initial state to its final state. This minimum amount is the amount transferred under optimum conditions, meaning by any reversible path.

Furthermore, another consideration is the total amount of thermal energy in the system (and surroundings) originally. If that amount is small (*i.e.*, if the system is at a low temperature), even a small increase in thermal energy will be important. If there is lots of thermal energy in the system initially (a high temperature), a small increase will be much less important. So we want the *ratio* of the amount of thermal energy transferred to the amount of thermal energy “hanging around”, which is measured by the temperature. Thus to find a measure of how much the spread function, or probability, increases, we find the amount of thermal energy transferred to the system, under reversible conditions, and divide by the (absolute) temperature.

There is a very important restriction, however, on this calculation. We know, from experience, that a cold gas has less energy spread around than a hot gas, so the hot gas has greater entropy. But does this really mean the hot gas is “more probable”. If so, what does that mean? We know we can cool a gas, taking it to a lower entropy state. Similarly, a compressed gas occupies a smaller volume than the same gas at a lower pressure. The gas at low pressure (high volume) has higher entropy, and thus “higher probability”, but it is possible to compress a gas, thus decreasing its entropy. A crystal is an orderly arrangement with less dissipation in terms of energy and especially in terms of space, but it is possible to produce a crystal from a liquid or even a gas. Thus we can convert the system to a lower entropy state, or a “less probable” state. Clearly we have overlooked some important consideration.

To predict what will be possible, or “more probable”, it is not sufficient to look at changes in the system. We must also take into consideration changes in the surroundings. Thus we must also find the amount of thermal energy transferred to the surroundings, under reversible conditions, and divide by the (absolute) temperature (of the surroundings). Then we add the

---

the amount of thermal energy transfer,  $Q$ . We simply ran out of letters.)

changes in the spread function for the system and for the surroundings.

Perhaps surprisingly, that simple formula works! It gives the increase in the spread function,  $S$ , for system plus surroundings when the system goes from an initial state to a final state. If this change is positive, *for system plus surroundings*, the change is likely to occur. If the change is negative (for system plus surroundings), the change is quite unlikely to occur. We are quite safe in saying it will not happen. Some specific examples will make the calculation process clearer, and show why the result is in agreement with what we should expect. Thus far we have largely avoided mathematics beyond addition (including multiplying to get a number of states). Now we will have to consider some ratios

*Thermal Energy Transfer.* Although the amount of energy transfer,  $\Delta E$ , can, in principle, always be measured, it turns out that it is generally not possible to measure the amount of thermal energy transfer,  $Q$ , except when the thermal energy transfer is reversible. However, if a hot body is placed in contact with a cooler body, we will very quickly get a smooth drop in temperature across the boundary and thermal energy will flow reversibly across the boundary.<sup>7</sup> The *amount* of thermal energy transferred to the system under these conditions is then  $Q$ . (Note that this is *not*  $\Delta Q$ , which would suggest a *change* in  $Q$ . We simply want an amount of thermal energy. Also, this  $Q$  is quite different from the number of states,  $\mathcal{Q}$ .) We want  $Q$  for this reversible process, or “path” and therefore we write it as  $Q_{\text{rev}}$ .

Then, representing by  $T$  the temperature on an absolute scale, we can write the change in the spread function (entropy) for the system as the ratio of thermal energy (“heat”) transferred reversibly to the absolute temperature,

$$\Delta S = Q_{\text{rev}}/T$$

and for the surroundings,

$$\Delta S_{\text{surr}} = (Q_{\text{rev}}/T)_{\text{surr}}$$

where  $Q_{\text{rev}}$  is here the amount of thermal energy transferred to the surroundings (*from* the system). Thus for the universe (system plus surroundings),

$$\Delta S_{\text{total}} = Q_{\text{rev}}/T + (Q_{\text{rev}}/T)_{\text{surr}}$$

*Melting Ice.* To make the idea more explicit, consider the process of melting one gram of ice at 0°C (= 273.15 K). We know that this requires a transfer of 333.5 joule of thermal energy (about 80 calorie). Dividing 333.5 J by 273.15 gives  $\Delta S = 1.221$  J/K. That is, for the system (the gram of ice),

$$\Delta S = Q_{\text{rev}}/T = 333.5 \text{ J}/273.15 \text{ K} = 1.221 \text{ J/K}$$

---

<sup>7</sup> Because a very small amount of thermal energy flows before equilibrium is established at the boundary, the  $\Delta E$  is not precisely  $Q$ . That is, we can never, in a finite time, precisely measure a (non-zero) reversible flow of thermal energy. Thus  $Q$  is almost never known, but we can come close enough for nearly all purposes if we are careful and keep the temperature differences small.

If we make the same calculation for the surroundings, which give up 333.5 J, or absorb  $-333.5$  J of thermal energy at the same temperature, the answer is  $-1.221$  J/K.

$$\Delta S_{\text{surr}} = (Q_{\text{rev}}/T)_{\text{surr}} = -333.5 \text{ J}/273.15 \text{ K} = -1.221 \text{ J/K}$$

Adding together  $\Delta S$  for the system and for the surroundings tells us that  $\Delta S = 0$  for melting the gram of ice. That should not be surprising because we explicitly assumed the process was totally reversible – reversible for the system and reversible for the surroundings. The process can go in either direction, equally well, melting ice or freezing water.

However, we have not described a very practical situation. How long would it take to melt a gram of ice if we packed it in ice at  $0^\circ\text{C}$ ? Certainly a very long time (an infinite amount of time). If we are serious about melting the ice, we must increase the temperature of the surroundings, say by  $\delta T$  (any finite temperature difference, which we may take to be quite small). Then we find

$$\Delta S = Q_{\text{rev}}/T = 333.5 \text{ J}/273.15 \text{ K} = 1.221 \text{ J/K}$$

for the system, but

$$\Delta S_{\text{surr}} = Q_{\text{rev}}/(T + \delta T) = -333.5 \text{ J}/(273.15 \text{ K} + \delta T) > -1.221 \text{ J/K}$$

The entropy change for the surroundings is negative, but smaller in magnitude than the positive entropy change for the system, so the sum of the entropy changes, for system plus surroundings, is positive.<sup>8</sup>

Similarly, if we wish to freeze water, the surroundings must be colder than  $0^\circ\text{C}$ , by some amount  $\delta T$ . Then, for the system,

$$\Delta S = Q_{\text{rev}}/T = -333.5 \text{ J}/273.15 \text{ K} = -1.221 \text{ J/K}$$

while for the surroundings,

$$\Delta S_{\text{surr}} = Q_{\text{rev}}/(T - \delta T) = +333.5 \text{ J}/(273.15 \text{ K} - \delta T) > 1.221 \text{ J/K}$$

and the sum, for system plus surroundings, is again positive.

These examples illustrate the general principle:

For any truly reversible (or equilibrium) process, the sum of the entropy changes for system plus surroundings is zero. For any real process (one that will proceed in a finite period of time), the sum of the entropy changes for system plus surroundings will be greater than zero.

This is an alternative, more complete way of stating the second law of thermodynamics. It is equivalent to saying that real processes go in the direction of reaching a more probable state, for system plus surroundings. A process is fully reversible only if the equilibrium states are equally

<sup>8</sup> A nit-picker could argue that if  $\delta T > 0$ , the process cannot be truly thermodynamically reversible. But if  $\delta T$  is small,  $Q$  is approximately equal to  $Q_{\text{rev}}$  and  $T$  is approximately equal to  $T + \delta T$ , so the difference from reversibility is negligible, and certainly close enough for our illustrative purposes.

probable.

## 6. Conclusion

Fundamental rules for the type of accounting called thermodynamics allow us to determine what things will happen, or why things happen. To summarize these rules, we look first at energy transfers, from the surroundings to the system. However, the total energy, of system plus surroundings, is constant, so measuring the energy transfer does not tell us whether a process will or will not actually happen (unless we are proposing a process that would change the total amount of energy, and therefore could not happen). Therefore, to choose which of the possible processes will occur, we look next at probability — which process is more likely to happen, based on mathematical probability.

The most important factor in determining the direction of change is whether the number of available states for the system plus surroundings increases, decreases, or remains the same. The relative number of states available determines the relative probability. Only those processes for which the number of available states, for system plus surroundings, increases will actually happen (although, in the limit for very slow processes, the number of states may remain essentially constant).

The two primary causes of an increase in number of states is conversion of energy to thermal energy, the importance of which is measured by  $\Delta S = Q_{\text{rev}}/T$ , and an expansion, for which more locations in space become available. But even an expansion process requires  $Q_{\text{rev}}$  to be positive for the expanding system, so it is sufficient to find  $\Delta S = Q_{\text{rev}}/T$  for the system, and a similar calculation for the surroundings, to include both dissipation of energy, as thermal energy, and expansion processes. If this total increase in the spread function (entropy) is greater than zero, the process may go as described. If it is negative, the process cannot go. If it is zero, the process describes an equilibrium, and therefore may go in either direction, but will require a very long time.<sup>9</sup>

In any case, it is clearly probability that decides what will happen. For real systems, made up of very, very large numbers of atoms and molecules, the probability argument is sufficient to determine with great certainty what will actually happen. In this sense, Feynman clearly was right when he said the most important information about the universe is that things are made up of atoms and molecules.

An increase in the spread function (entropy) may be described as an increase in *chaos*. The

---

<sup>9</sup> If we restrict our considerations to processes that occur at constant volume and constant temperature (or, with slight modifications, at constant pressure and constant temperature), the system will tend to go to a state of lower energy, as well as tending to go to states of greater spread function ( $S$ , or entropy). Therefore, if we measure the energy change and subtract the product of temperature and  $\Delta S$ , we obtain a function, or mathematical description, that tells us whether a process will or will not go solely in terms of changes in the system, without having to make calculations for the surroundings. For serious thermodynamic calculations, this is a clear advantage. (The change in this function must be negative. This quantity is generally known as “free energy”, a name that is somewhat misleading.)

second law of thermodynamics tells us that the universe always tends to become more chaotic. However, we recognize that systems can become more ordered. We may arrange molecules neatly in crystals and compress gases to smaller volumes. Life itself is a striking example of creation of ordered systems from randomly distributed chemicals. To explain this, we saw that we must consider changes in the surroundings as well as changes to the system.

Thus thermodynamics tells us that

*We can decrease the chaos in a system (lower its entropy), but only by increasing the chaos elsewhere in the universe, even more.*