ENTROPY

Thermodynamics is often called the science of energy. This designation steals accomplishments from other sciences, and diminishes accomplishments of thermodynamics. Rather, thermodynamics is the science of *entropy*.

Thermodynamics Is the Science of Entropy

The play of thermodynamics. In the play of thermodynamics, entropy is the leading actor, and energy is a supporting actor, along with other supporting actors, such as space, matter, and charge. We would never say that thermodynamics is the science of charge. That designation belongs to electrodynamics.

Entropy behaves differently from energy, space, matter, and charge, but energy, space, matter, and charge behave similarly:

- An isolated system maximizes entropy.
- An isolated system conserves energy, space, matter, and charge.

This distinction is responsible for the structure of thermodynamics.

In thermodynamics, energy, space, matter, and charge are supporting actors of equal importance. Each of these supporting actors, together with entropy, produces a child: temperature, pressure, chemical potential, and electrical potential—the second generation of supporting actors. We will introduce some grandchildren: heat capacity, latent heat, elastic modulus, coefficient of thermal expansion, etc. We will also introduce some cousins: enthalpy, Helmholtz function, and Gibbs function.

Entropy is the mother of thermodynamics. If entropy were removed, thermodynamics would collapse, and would have absolutely nothing to say about energy or anything, except for lines stolen from other sciences.

Energy belongs to many sciences. Let us begin the play of thermodynamics with this loud, pesky, supporting actor—energy. Energy is an actor in many sciences. All have much to claim about energy: forms of energy, stores of energy, carriers of energy, conversion of energy from one form to another, and flow of energy from one place to another.

Associated with these words—forms, stores, carriers, conversion and flow—are a great variety of inventions and discoveries. Examples include fire, springs, capacitors, water wheels, windmills, steam, engines, refrigerators, turbines, generators, batteries, light bulbs, and solar cells. These you have learned, and will learn more, from many courses (including this one), as well as from daily life.

It has been common to let the supporting actor—energy—to dominate the play of thermodynamics. We will avoid this pitfall. We will let energy play its supporting role, along with space, matter, and charge.

This course. We will develop the *logic* of entropy from first principles, *intuition* of entropy from everyday phenomena, and *application* of entropy in many domains. In particular, we will spend much time on the application of entropy to devices that convert heat and motion. Examples include engines and refrigerators. This course will highlight how entropy links theory and experiment.

One subject, many titles. The science of entropy is often given different titles, such as thermodynamics, statistical thermodynamics, and statistical mechanics. Depending on domains of application, other titles include engineering thermodynamics, chemical thermodynamics, biological thermodynamics, and thermal physics. Entropy is also a main part of physical chemistry, electrochemistry, fluid and solid mechanics, as well as electrodynamics.

Exercise. Find all courses devoted to the science of entropy on campus. Compare the descriptions of the courses. Why do you take this course?

Thermodynamics vs. probability. All this energetic commotion merely prepares us for the entry of the leading actor: entropy. Let's cast light on the leading actor now, and return to supporting actors later.

Of our world we know the following facts:

1. An isolated system has a certain number of quantum states. Denote this number by Ω.
2. The isolated system flips from one quantum state to another, rapidly and ceaselessly.
3. A system isolated for a long time flips to every one of its quantum states with equal probability, $1/\Omega$.

Thus, a system isolated for a long time behaves like a *fair* die:

1. The die has six faces.
2. The die is rolled from one face to another.
3. The die is rolled to every face with equal probability, 1/6.

The similarity of thermodynamics and theory of probability is evident. Indeed, thermodynamics will borrow many ideas from the theory of probability. The following table adds additional examples. See supplements on the theory of probability and its comparison with thermodynamics.

	Theory of probability	Thermodynamics
Experimental setup	Die	Isolated system
Experiment	Roll the die	Isolate the system for a long time
Sample space	6 faces	Ω quantum states
Probability of each sample point	1/6	$1/\Omega$
Subset of sample points	Event (e.g., a roll resulting an even number)	Configuration Macrostate Thermodynamic state
Probability to realize a subset	3/6	(number of quantum states in the subset)/ Ω
A map from sample space to a real variable	Random variable	Internal variable Thermodynamic property

Entropy. We call the quantity $\log\Omega$ the *entropy*. Logarithm of any base will work; we will use the natural base. The entropy of a throw of a fair die is $\log 6$. The entropy of an isolated system is $\log\Omega$.

What is an isolated system? What is a quantum state? How long is a long time? How do we count the number of quantum states of a system in everyday life, such as a glass of wine? Why do we hide Ω behind the log? How does entropy help us understand the conversion between fire and motion? We will answer these questions in this course.

Exercise. What is the entropy of a throw of a coin? What is the entropy of a simultaneous throw of a coin and a die? What is the entropy of a simultaneous throw of 1000 dies?

Exercise. Carbon atoms can form a crystalline lattice, known as diamond. At room temperature and atmospheric pressure, the entropy of diamond per carbon atom is 0.3. What does this number mean?

Isolated System

System. We can regard any part of the world as a *system*. For example, a proton and an electron constitute a system, called a hydrogen atom. A glass of wine is also a system. This system is composed of water, ethanol, and other molecules. Do we include the glass as a part of the system? Maybe, if we decide to study the reaction between water and glass. The decision is ours. We can regard any part of the world as a system. Even the empty space can be a system; the vacuum hosts electromagnetic field.

System interacts with the rest of the world. The hydrogen atom interacts with the rest of the world, for example, by absorbing or emitting photons. When the hydrogen atom absorbs a photon of a certain frequency, the electron cloud changes its shape.

A system like the glass of wine may interact with the rest of the world in many ways. We see the wine because it reflects light. We smell the wine because molecules jump off the wine and enter our noses. We hear the wine when we move a wet finger around the edge of the glass. We warm up the wine when the vibration of the molecules in our hands transfers energy to the vibration of the molecules in the wine. We taste the wine when we drink it.

We can augment our five senses by using other things to interact with the wine. We subject the wine to a pressure to learn how tightly the molecules pack. We subject the wine to an electric field to find out how electric charges on the molecules move around. We pour fruit juice into the wine and watch they mix. We place the wine on a flame and smell the vapor.

Exercise. Describe various ways in which a piece of cheese interacts with the rest of the world.

Isolated system. Our subject—thermodynamics—deals with the interactions between a system and the rest of the world. But our narrative begins with something simpler: an *isolated system*—a system that does not interact with the rest of the world.

To make a glass of wine an isolated system, we seal the glass to prevent molecules from escaping, we place the glass in a thermos to prevent energy from escaping by heat, and we make the seal unmovable so that the external pressure cannot do work to the wine. We are alert to any other modes of interaction between the wine and the rest of the world. Does the magnetic field of the earth affect the wine? If it does, we will find a way to shield the glass of wine from the magnetic field also.

Of course, nothing is perfectly isolated. Like any idealization, the isolated system is a useful approximation of the reality, so long as the interaction between the system and the rest of the world negligibly affects a phenomenon that we choose to study.

For example, it may be too much trouble for us to isolate the wine from gravity. Few have the luxury to study the wine under the zero-gravity condition.

Gravity is important if we carry the wine around, but is unimportant if we want to study the interactions between the molecules in the wine.

Exercise. Describe a method to keep water hot for a long time. What can you do to prolong the time? What makes water eventually cool down?

Quantum States of an Isolated System

In high schools we were told that a particle is also a wave, and a wave is also a particle. Sometime later we were told that the wave is governed by a partial differential equation known as Schrodinger's equation. In this course we will not talk much about partial differential equations, but some solutions provide a useful perspective. A picture is worth a thousand words, a movie is worth a thousand pictures, and an equation is worth a thousand movies. With an equation, you can generate lots of movies, plot lots of pictures, and tell lots of stories.

Exercise. Study the supplement on the quantum states of a particle in a box. A particle confined in a cubic box constitutes an isolated system. One quantum state of the isolated system is $\left(n_x,n_y,n_z\right)=\left(1,2,3\right)$. List all quantum states of the isolated system.

Exercise. Read about quantum states of a hydrogen atom online. Sketch the shape of electron cloud of the eight quantum states when a hydrogen atom is at the second energy level. Describe these quantum states in words.

When a hydrogen atom is isolated at the third energy level, how many quantum states does the isolated system have?

An isolated system has a certain number of quantum states. We describe a quantum state of the glass of wine by a collective cloud of all electrons and positions of all nuclei. An isolated system like the glass of wine has a huge number of quantum states. We denote the number of quantum states of an isolated system by Ω. The number of quantum states of isolated system follows obvious trends:

- Warm cheese has more quantum states than cold cheese. An isolated system of higher energy has more quantum states.
- Two glasses of wine have more quantum states than one glass of wine. An isolated system of more molecules has more quantum states.

- Molecules in a gas have more quantum states than the same number of molecules in a liquid. An isolated system of larger volume has more quantum states.
- Molecules in a liquid have more quantum states than the same number of molecules in a crystal. Molecules in the liquid can be arranged in more ways than in a crystal.

Exercise. Suppose we know that a glass of wine has Ω number of quantum states. How many quantum states do three identical glasses of wine have?

Use quantum states without knowing much quantum mechanics. Wait a minute! Can we really tell apart individual quantum states of a glass of wine? What do we even mean we can tell them apart? We will answer these questions in due course. In particular, we will show how to determine the number of quantum states of a glass of wine by experimental measurement.

Given an isolated system, all its quantum states are determined by quantum mechanics. For example, when the hydrogen atom is isolated at the second energy level, the isolated system has eight quantum states.

In practice, however, computers are too slow to perform quantum-mechanical calculation for most systems. We can perform quantum-mechanical calculation for the hydrogen atom, but we cannot do so for the glass of wine.

We cannot wait for the brute-force quantum-mechanical computation to answer practical questions. We need to develop other methods. After all, we have all learned to ride bicycles without any use of Newton's mechanics. Our bodies and bicycles know enough mechanics for us to ride them. Similarly, our experimental instruments and the glass of wine know more quantum mechanics than computers do. We can readily see the color of the wine, measure the density of the wine, and analyze its molecular composition—all without performing any quantum-mechanical computation.

Count the number of quantum states of an isolated system by experimental measurement. To a large extent, thermodynamics is an algorithm to obtain and use information from experiments, without quantum-mechanical computation, even though all such information is rooted in quantum mechanics. These lectures teach you this algorithm.

In particular, we will describe how to determine experimentally the number of quantum states of a system such as a glass of wine. Such an experiment, in effect, counts the number of quantum states without identify them individually.

The above statement may sound mysterious at this point, but it should not be. Consider a crude analogy. We can "count" the number of carbon atoms in a piece of diamond without identifying the atoms individually. We simply measure the mass of the piece of diamond. A carbon atom contains 6 protons, 6 neutrons, and 6 electrons. The mass of each proton and each neutron is about $1.67\times 10^{-27}\,\mathrm{kg}$. By comparison, the mass of electron is negligible. Thus, the mass of each carbon atom is approximately $12\times 1.67\times 10^{-27}\,\mathrm{kg}$. The mass of the piece of diamond divided by the mass of a carbon atom gives the number of carbon atoms in the piece of diamond.

Fundamental Postulate

An isolated system flips from one quantum state to another, rapidly and ceaselessly. An isolated system is not static: it flips rapidly and ceaselessly from one quantum state to another.

For example, in the glass of wine, molecules move rapidly and ceaselessly. Even though you cannot see the movement of the molecules with your own eyes, you can smell the movement: molecules jump off the wine and enter your nose.

Fundamental postulate. Will the isolated system flip to some of its quantum states more often than others? The fundamental postulate says that

> *A system isolated for a long time flips to every one of its quantum states with equal probability.*

The fundamental postulate cannot be deduced from more elementary facts, but its predictions have been confirmed without exception by empirical observations. We will regard the fundamental postulate as an empirical fact, and use the fact to build thermodynamics.

Equilibrium. A system has just been isolated, and the isolated system has Ω quantum states. Right after isolation, the system flips to some of its quantum states more often than others. The isolated system is then said to be *out of equilibrium*. When the system is isolated for a long time, the fundamental postulate says that the system flips to every one of its quantum states with equal probability. The isolated system is then said to be *in equilibrium*.

The fundamental postulate connects two descriptions of a system: "isolated for a long time", and "flipping to every one of its quantum states with equal probability". We have just made the three phrases synonymous: a system "isolated for a long time" is a system "flipping to every one of its quantum states with equal probability", and is a system "in equilibrium".

An isolated system, in or out of equilibrium, is always dynamic and is never static. The isolated system flips rapidly and ceaselessly from one quantum state to another. The isolated system out of equilibrium differs from the isolated system in equilibrium in one respect. Out of equilibrium the isolated system flips to every one of its quantum states with unequal probability, whereas in equilibrium the isolated system flips to every one of its quantum states with equal probability.

Irreversibility. The fundamental postulate says that a system isolated for a long time flips to every one of its quantum states with equal probability. A newly isolated system is typically out of equilibrium, flipping to some quantum states more often than others. After being isolated for a long time, the system is in equilibrium, flipping to every one of its quantum states with equal probability. So long as the system is isolated, once it reaches equilibrium, it will not get out of equilibrium. The isolated system out of equilibrium is said to approach equilibrium with *irreversibility*.

Whereas all other fundamental laws of nature are time-reversible, the fundamental postulate of thermodynamics gives time a direction. We call it the arrow of time.

Kinetics. A newly isolated system has the whole set of quantum states, but flips to them with unequal probabilities. After a long time in isolation, the system flips to every one of its quantum states with equal probability. How long is a long time? The fundamental postulate is silent on this question. The time to attain equilibrium depends on how fast the molecules move. The study of how long an isolated system attains equilibrium is a subject known as *kinetics*. We regard kinetics as a subject distinct from thermodynamics. Kinetics studies the evolution of a system in time.

Thermodynamics is timeless. By contrast, thermodynamics cares about the *direction* of time, but not the *duration* of time. Thermodynamics makes no use of any quantity with dimension of time. Time enters thermodynamics merely to distinguish "before" and "after"—that is, to give a direction of irreversibility.

Thermodynamic State of an Isolated System

Sample space. An isolated system flips among a set of quantum states. Using the language of the theory of probability, we call each quantum state a *sample point*, and call the set of all quantum states the *sample space*. Thus, associated with an isolated system is a sample space. Denote the total number of quantum states of the isolated system by Ω.

Configuration. The sample space is the whole set of quantum states of an isolated system. Each *subset* of the sample space constitutes some quantum states of the isolated system. The sample space can form a total of 2^{Ω} subsets.

In the theory of probability, a subset of the sample space is called an *event*. This word "event", however, seems to be inappropriate for the isolated system. What do we mean by the phrase "an event of an isolated system"? This inconvenience of English should not confuse us. We call a subset of quantum states of by a different name: a *configuration*.

This important concept—a subset of quantum states of an isolated system—is sometimes given other names, such as a *conformation*, a *macrostate,* and a *thermodynamic state*. All these names mean the same thing: a subset of quantum states of the isolated system. The phrase is just too long for us to say it all the time. We need a word or a short phrase. But replacing a long phrase with a short one does not create any new idea. To identify a configuration of an isolated system, we can list all the quantum states that constitute the subset.

Exercise. A hydrogen atom is isolated at the second energy level. Describe the sample space of this isolated system. List all the subsets of the sample space.

Probability of a configuration. An isolated system flips among a set of Ω quantum states. A subset of the quantum states, designated A, consists of Ω_A number of quantum states. After being isolated for a long time, the system flips to every one of its quantum states with equal probability. Consequently, after the system is isolated for a long time, the probability for the system to be in subset A is

$$\frac{\Omega_A}{\Omega}.$$

A system isolated from a long time is more probable in a subset of quantum states if the subset consists of more quantum states.

Identify a subset through an experimental observation. An isolated system such as a glass of wine has a large number of quantum states, and an even larger number of subsets of quantum states. We will not be interested in that many subsets. Rather, we are interested in subsets that we can identify through an experimental observation. We next give several examples.

A Long Molecule in a Liquid

Consider an isolated system consisting of a long molecule in a liquid. The long molecule can be in two configurations: a loop or a chain. The long molecule is a loop when its two ends are bonded together, and is a chain when its two ends are separated. The two configurations can be differentiated by experiments. For example, we can synthesize the long molecule such that, when its two ends bond, the molecule produces a photon.

In this example, all the quantum states of the isolated system constitute a sample space. The two configurations are subsets of the sample space. Each configuration consists of many quantum states. For example, both a straight chain and a wiggled chain belong to the configuration of chain. Besides, the molecules in the surrounding liquid ceaselessly move.

We have dissected the whole set of the quantum states into two subsets: one associated with a chain, and the other associated with a loop. The individual quantum states are too numerous to interest us.

A Bouncing Ball

Exercise. A ball bounces on the ground and eventually stops. The ball loses its energy to the ground and air. But energy is conserved. How come the ball does not reverse the process, gain energy from air and ground, and bounce higher and higher? Explain this irreversibility in terms of fundamental postulate. Construct an isolated system. Use the energy of the ball as the internal variable to dissect the set of quantum states of the isolated system.

Energy is indeed conserved; it just transfers from the ball to the air and the ground. We say that energy is *dissipated* from the ball.

Half Bottle of Water

Throughout this course we will spend much time to study water molecules in liquid and gaseous phases. In particular, vaporization and condensation of water are central to the operation of steam engines. Here we start with a few empirical observations.

Consider a half bottle of water. We block all interactions between the inside and outside of the bottle. We cap the bottle, so that no water molecules will escape from the bottle, and no air from outside will enter the bottle. We wrap the bottle with a thermal insulator, so that the water in the bottle and the air outside do not exchange energy by heat. We make the bottle rigid, so that the pressure inside the bottle does no work to the environment. We make the half bottle of water an isolated system.

Irreversibility. In the beginning, we partition the bottle into two parts using a seal. The bottom part of the bottle contains water molecules, and the top part is a vacuum. When we puncture a hole in the seal, molecules can fly into the top part of the bottle.

Our experience indicates that this process is *irreversible*. Some molecules will fly from the bottom part of the bottle into the top part of the bottle. But the top part of the bottle will not become a vacuum again. What causes this irreversibility?

This question presented the founders of thermodynamics a paradox. The movements of the molecules obey the laws of mechanics, which take the same form when time is reversed. If the molecules can fly into the vacuum, the laws of mechanics predict that the same molecules can also fly back to the bottom part of the bottle. Then why the observed process is irreversible?

Equilibrium. Molecules take some time to find the small hole, and fly from the bottom part of the bottle into the top part of the bottle. Given enough time, the liquid and the vapor *equilibrate*. In equilibrium, molecules in the liquid still evaporate into the vapor, and molecules in the vapor condense into the liquid, but the number of molecules in the vapor and that in the liquid remains constant.

Kinetics. After the seal is punctuated, how much time is needed for the half bottle of water to equilibrate? The time to attain equilibrium depends on how fast the molecules jump off the liquid, fly through the hole, and spread in the top part of the bottle. This course will not spend much time on such kinetic questions.

Constraint Internal to an Isolated System

Internal variable. We have made the bottle of water molecules an isolated system. The seal in the middle of the bottle provides a *constraint internal to the isolated system*. The act of puncturing a hole in the seal lifts the constraint, making the number of molecules in the top part of the bottle, N, an *internal variable*.

A constraint internal to an isolated system is an essential idea, but the idea seems to be vague or unmentioned in most textbooks of thermodynamics. In the language of the theory of probability, the internal variable of an isolated system is a random variable—that is, a function that maps the sample space to a real variable. The *domain* of the function is the sample space—the whole set of quantum states of the isolated system. The *range* of the function is a set of numbers.

Use an internal variable to dissect sample space. The isolated system has M water molecules, and Ω quantum states.

Before the seal is punctured, the top part of the bottle has no water molecule, $N = 0$, and the isolated system flips among a subset of quantum states. Denote the number of quantum states in this subset by $\Omega(0)$.

When the top part of the bottle has N water molecules, the isolated system flips among a subset of quantum states. Denote the number of quantum states in this subset by $\Omega(N)$.

We have just used the number of molecules in the top half of the bottle, N, to dissect the set of quantum states of the isolated system into a family of subsets. These subsets are disjoint, and their union is the whole set of the quantum states of the isolated system. Write

$$\Omega = \Omega(0) + \Omega(1) + ... + \Omega(M).$$

Our notation may upset a purist. We have used the symbol Ω to represent two different objects: Ω is the number of quantum states in the sample space of the isolated system, and $\Omega(N)$ is the number of quantum states in the subset of the sample space when the internal variable is held at a value N. The former is a number, and the latter is a function.

Constrained equilibrium. If the hole in the seal is small, few molecules go through the hole in a given amount of time. Consequently, when the top part of the bottle has any particular number of molecules, N, there is a long enough time for the isolated system to be in every one of its $\Omega(N)$ quantum states with equal probability. We say that the isolated system is in *constrained equilibrium*, with N number of molecules in the top part of the bottle.

Evolution from one configuration to another. Each value N corresponds to a subset of the sample space—that is, each value N corresponds to a configuration, or a thermodynamic state, of the isolated system. As molecules go through the hole at a low rate, the isolated system evolves in the direction that changes the number N to increase the number $\Omega(N)$. The isolated system evolves from one configuration to another.

Distribute probability among the subsets in a dissection. In equilibrium, the isolated system flips to every of its Ω quantum states with equal probability. The probabilities to have no molecules, 1 molecule, or M molecules in the top part of the bottle are, respectively,

$$\frac{\Omega(0)}{\Omega},\frac{\Omega(1)}{\Omega},...,\frac{\Omega(M)}{\Omega}.$$

This list of numbers distributes probability among the subsets of dissection.

Arrow of time. We now understand the irreversibility in terms of the fundamental postulate. When the top part of the bottle is sealed from the bottom part, the top part of the bottle has no molecule, and the isolated system flips among a single subset of $\Omega(0)$ number of quantum states. When the seal is punctured, the isolated system flips among *all* quantum states. Of all subsets of $\Omega(0)$, $\Omega(1)$,..., $\Omega(M)$ numbers of quantum states, one subset has the largest number of quantum states. According to the fundamental postulate, the most probable number of molecules in the vapor, N, maximizes the function $\Omega(N)$. Consequently, the change from the initial subset to the final subset appears to be irreversible.

Basic Algorithm of Thermodynamics

From probability to (almost) certainty. After the seal is punctured for a long time, the vapor and the liquid equilibrate. In equilibrium, what is the number of water molecules in the vapor?

For an isolated system with a *large* number of quantum states, our experience indicates that the most probable partition is much, much more probable than other partitions. We therefore often focus our attention exclusively on the partition that maximizes the function, $\Omega(N)$, and not be bothered with less probable partitions. Thus, a large part of thermodynamics is a game of maximization or minimization.

Basic algorithm of thermodynamics. Water vaporizes into a vacuum. Ink disperses in water. Cheese smells. A ball bounces a few times, and then stops bouncing. These irreversible processes can be understood on the basis of the fundamental postulate explains in a unified way. Here is a summary of the basic algorithm of thermodynamics using general terms:

1. Construct an isolated system with an internal variable, x.
2. Use the internal variable to dissect the whole set of the quantum states of the isolated system into subsets. Call each subset a configuration of the isolated system. When the internal variable takes value x, the isolated system flips among a subset of its quantum states. Obtain the number of the quantum states in this subset, $\Omega(x)$.
3. Maximize the function $\Omega(x)$ to determine the (most probable) value of the internal variable x after the constraint is lifted for a long time.

The fundamental postulate implies this basic algorithm. The basic algorithm, however, is an incomplete expression of the fundamental postulate. In using the basic algorithm, we no longer seek probability of an individual configuration; rather, we seek the configuration with the maximum probability. For an isolated system of a large number of quantum states, the most probable value of the internal variable is much more probable than any other values. That is, the fluctuation in the value of the internal variable is negligible. The rest of this course shows how this algorithm links theory and experiment. In particular, the function $\Omega(x)$ is usually determined by experimental measurement.

Exercise. Consider a glass of wine and a piece of cheese together as a single system, isolated from the rest of the world. Initially the wine and the cheese are sealed from each other. Now we puncture the seal, so that water molecules (for example) can escape from the wine and move into the cheese, or the other way around. Use the fundamental postulate to make predictions.

Product and Sum

Two separately isolated systems. Now consider two isolated systems, such as a glass of wine and a piece of cheese. One isolated system has a total of Ω' number of quantum states, labeled as $\{\gamma'_1, \gamma'_2, ...,\}$. The other isolated system has a total of Ω'' number of quantum states, labeled as $\{\gamma''_1, \gamma''_2, ...,\}$.

We keep the two systems *separately isolated*. In particular, we do *not* allow the two systems to interact with each other. The two systems together, of course, constitute an isolated system. A quantum state of this composite is any combination of a quantum state chosen from the set $\{\gamma'_1, \gamma'_2, ...,\}$ and a quantum state chosen from the $\{\gamma''_1, \gamma''_2, ...,\}$. For example, one quantum state of the composite is when one system is in quantum state γ'_2 and the other system is in quantum state γ''_3. The total number of all such combinations is

$$\Omega'\Omega''.$$

This is the number of quantum states of the composite. Thus,

> *The number of quantum states of the composite of two separately isolated systems is the product of the quantum states of the individual isolated systems.*

Logarithm. Recall a property of the logarithm:

$$\log(\Omega'\Omega'') = \log\Omega' + \log\Omega''.$$

Thus, when two separately isolated systems are regarded as a composite, the logarithm of the number of quantum states of the composite is the sum of the logarithm of the number of quantum states of one isolated system and that of the other isolated system.

The base of the logarithm can be any number. Anticipating many uses of derivatives, we adopt Euler's number *e* as the base the natural base, so that

$$\frac{d(\log\Omega)}{d\Omega} = \frac{1}{\Omega}.$$

It is common to write logarithm with base *e* as ln, but we will just write log.

Exercise. We have two kinds of ice cream: vanilla and chocolate. We have three kinds of toppings: strawberry, walnut and almond. How many combinations of ice cream and toppings can we make?

Entropy

Entropy of an isolated system. If you are tired of the phrase "logarithm of the number of quantum states", you may as well join the crowd and replace the phrase with a cryptic word: *entropy*. That is, you call the quantity $\log\Omega$ entropy, and give it a symbol *S*. Thus, an isolated system of Ω number of quantum states has the entropy

$$S = \log\Omega.$$

By replacing a phrase with a word you should not trick yourself into believing that you have learned anything new. When you speak of entropy, you simply paraphrase what you already know about "logarithm of the number of quantum states".

Going from Ω to $\log\Omega$ is a trivial step. The step adds some mathematical convenience, but adds no physical content. It does seem to be unfair for entropy S to steal the show from the number of quantum states Ω.

Entropy is additive. The entropy of a composite system is the sum of the entropies of the individual systems:

$$S = S' + S''.$$

When several separately isolated systems are regarded as a composite, the entropy of the composite is the sum of the entropies of the individual systems.

Entropy of an isolated system subject to a constraint internal to the system. Consider an isolated system of an internal variable x. When the internal variable is constrained at a value x, the isolated system flips among a subset of its quantum states. Denote the number of quantum states in this subset by $\Omega(x)$. Denote the entropy of the isolated system subject to the internal constraint by

$$S(x) = \log\Omega(x).$$

The second law. The logarithm is an increasing function. Consequently, maximizing $\Omega(x)$ is equivalent to maximizing $S(x)$. Here is a paraphrase of the basic algorithm of thermodynamics:

> *When a constraint internal to an isolated system is lifted, the internal variable changes to maximize entropy.*

This statement is one version of the second law of thermodynamics. The fundamental postulate implies the second law. The second law is an incomplete expression of the fundamental postulate. The second law, in its various statements, sounds like ancient philosophical pronouncements. They sound mysterious because they miss basic facts in nature (e.g., the rapid and ceaseless motions of molecules, and the fundamental postulate). We will not waste time to decipher such pronouncements. Rather, we will start with fundamental postulate, and devise *algorithms* that link theory and experiment.

Entropy of a Pure Substance

Pure substance. By a pure substance we mean a collection of a large number of a single species of atoms (or molecules). The molecules aggregate into various forms, known as phases. For example, diamond is a crystalline lattice of carbon atoms, water is a liquid of H_2O molecules, and oxygen is a gas of O_2 molecules.

In a crystalline solid, molecules form a lattice. In a liquid, molecules touch one another, but change neighbors frequently. In a gas, molecules fly in space. The volume of a piece of the substance, V, is proportional to the number of molecules in the piece, N. The volume per molecule of the substance is

$$v = \frac{V}{N}.$$

We denote the volume of the piece of substance using the uppercase letter V, and denote the volume per molecule in the substance using the lower case letter v. In a solid and a liquid, v roughly gives the size of an individual molecule. In a gas, however, v is much larger than the size individual molecule.

Similarly, the mass of a piece of the substance, M, is proportional to the number of molecules in the piece, N. The mass per molecule of the substance is

$$m = \frac{M}{N}.$$

The mass per molecule of a substance is the same regardless whether the substance is a solid, liquid, or gas.

Exercise. Find the mass per water molecule. Find the volume per molecule in solid, liquid and gaseous water.

Entropy of a pure substance. The number of quantum states in a piece of a pure substance scales as

$$\Omega = \omega^N,$$

where the number ω is independent of N.

The entropy of a piece of the substance is given by $S = \log\Omega$, and is proportional to the number of molecules in the piece, $S = N\log\omega$. The entropy per molecule in the substance is

$$s = \frac{S}{N}.$$

At the room temperature and atmospheric pressure, the entropy of diamond, lead and water are 0.3, 7.8 and 22.70, respectively. A strong substance, such as diamond, has a small value of entropy, because individual atoms are held together by strong chemical bonds, which reduces the number of quantum states.

Complex substances generally have larger entropies than simple substances. For example, the entropies for O, O_2 and O_3 are 19.4, 24.7, 28.6, respectively.

When a substance melts, the substance transforms from a crystal to a liquid. Associated with this phase transition, the entropy typically increases by a number between 1 to 1.5.

Exercise. 12 g of diamond has a total of 6.02×10^{23} number of carbon atoms. What is the entropy in 1 g of diamond at room temperature?

Experimental determination of entropy. A thermodynamic state of a pure substance is specified by values of two independent thermodynamic properties. One can choose many pairs of thermodynamic properties as independent variables. Here is one choice: u and v.

For a given pure substance, once two thermodynamic properties are chosen as independent variables, all other thermodynamic properties are dependent variables. For example, entropy is a function of energy and volume:

$$s(u,v).$$

Later we will show a thermodynamic identity:

$$ds = \frac{1}{T}du + \frac{P}{T}dv,$$

where T is the temperature and P is the pressure. Thus, one can measure the function $s(u,v)$ incrementally.

Tables of entropy. The entropy of a substance is often tabulated as a function of two variables, such as temperature and pressure. You can find such tables for many substances at the end of a textbook. You can also find them online.

Zero entropy. Such a table of a substance often chooses an arbitrary state of temperature and pressure as a reference state. The table lists the entropy of the substance at this reference state as zero, and lists the entropy of the substance at any other state of temperature and pressure relative to the reference state.

We follow this practice with caution. Recall $S = \log\Omega$. Zero entropy is not something arbitrary, but has physical significance: it corresponds to an isolated system of a single quantum state.

Units of Entropy

All units are irrational, but some are necessary evils. Think about meters and inches. Units for entropy, however, are irrational, unnecessary evils.

Entropy is a dimensionless number. We have defined entropy of an isolated system, *S*, by the number of quantum states of the isolated system, Ω:

$$S = \log \Omega ,$$

Entropy is a pure number, and has no unit. By historical accidents, however, entropy is given many units.

Unit used in physics. Let us begin with a unit popular among physicists. The conversion factor between the dimensionless entropy and the entropy with the unit is called the Boltzmann constant, $k = 1.38 \times 10^{-23} \mathrm{J/K}$, where J stands for the unit of energy (Joule), and K stands for the unit of temperature (Kelvin).

$$S = k \log \Omega .$$

This conversion is even more pointless than that between inch and meter. The ugly unit does, however, remind us of three early developers of thermodynamics, Joule, Kelvin, and Boltzmann. Our culture celebrates their brilliance with silliness. Rationalizing silliness takes effort. Later in this course we will offer some rationalization for this silly unit.

Unit used in chemistry. Chemists further compound silliness. They give another pure number a unit! 12 g of diamond has a total of 6.02×10^{23} number of carbon atoms. Chemists take this number to be a unit, and call it a *mole*. The number is known as Avogadro's number, is designated as $N_A = 6.02 \times 10^{23} \mathrm{mole}^{-1}$, and serves as the factor of conversion from the number of individual molecules to the number of moles:

$$6.02 \times 10^{23} \text{ molecules} = 1 \text{ mole}.$$

This silliness is so transparent that no rationalization needs to be offered.

Entropy is then reported in the unit of $\mathrm{JK^{-1}mole^{-1}}$. Write

$$S = R \log \Omega .$$

The conversion factor is $R = kN_A = 8.31 \mathrm{JK^{-1}mole^{-1}}$. Now you draw upon the wisdom of crowd, and name *R* the universal gas constant. It makes you sound more learned than naming another universal gas constant: 1, as *R* should be if you take entropy to be a pure number, and count molecules by another pure number.

Exercise. Being silly has cost. Silly units generate silly questions. How come entropy has anything to do with gas? What is the significance of 12 g of diamond?

Unit used in engineering. In engineering, entropy is commonly reported in the unit of $\mathrm{kJK^{-1}kg^{-1}}$. Write

$$S = R_{\text{specific}} \log \Omega .$$

The conversion factor R_{specific} is known as the specific gas constant, and relates to the universal gas constant as $R_{\text{specific}} = R/M$, where M is the mass per mole of a given substance. For water, M = 18 g/mole and $R_{\text{specifc}} = 0.46\mathrm{kJK^{-1}kg^{-1}}$.

Exercise. Find the entropy of ice and the entropy of liquid water, both at the melting point. Find the entropy of liquid water and the entropy of gaseous water, both at the boiling point and 1 atmospheric pressure. Give all values in the dimensionless form.

Supplement: History of Thermodynamics

Compared to geometry and mechanics, thermodynamics is a young science. Entropy was discovered in analyzing the efficiency of steam engines (Carnot 1824). This domain of application has remained to be of singular importance to humanity since the Industrial Revolution. As they say, thermodynamics learned more from steam engines than steam engines learned from thermodynamics. Indeed, Kelvin (1850s) coined the word thermodynamics: thermo means heat, and dynamics means motion.

It took a few decades and a dozen people to relate entropy to microscopic states of molecules, and to relate entropy to energy, space, matter, and charge. Clausius (1954) discovered the relation between entropy, energy, and temperature. Clausius (1865) made up the word entropy after the Greek word $\tau\rho o\pi\eta$, transformation. His justification for this name looks contrived, but the name sticks. There is no point to look for scientific meanings from an ancient language. If you ever discover a new concept, do not name it using existing words in a popular language, such as English or Chinese. It would confuse the new concept with existing meanings of the word. Bad examples abound: pressure, stress, strain, power, force, work, heat and energy.

Boltzmann (1877) linked entropy to the microstates of molecules. Gibbs (1873, 1901) gave the science of entropy the form as we know today. Even in his time, the science of entropy had grown far beyond its earlier concern over thermomechanical conversion. Indeed, entropy exists in all physical systems, and some non-physical ones. The vast subject is no longer supported by the root of the word “thermodynamics”. We still call the subject thermodynamics, but the word no longer literally means heat and motion.

Supplement: Variety of Perspectives

Steam engine was a transformative invention two hundred years ago, just as the Internet is a transformative invention today. Cannot (1824) began his little book, *Reflections on the Motive Power of Fire*, with the following paragraphs.

Every one knows that heat can produce motion. That it possesses vast motive-power no one can doubt, in these days when the steam-engine is everywhere so well known.

To heat also are due the vast movements which take place on the earth. It causes the agitations of the atmosphere, the ascension of clouds, the fall of rain and of meteors, the current of water which channel the surface of the globe, and of which man has thus far employed but a small portion. Even earthquake and volcanic eruptions are the result of heat.

From this immense reservoir we may draw the moving force necessary for our purposes. Nature, in providing us with combustibles on all sides, has given us the power to produce, at all times and in all places, heat and the impelling power. To develop this power, to appropriate it to our uses, is the object of heat engines.

The study of these engines is of the greatest interest, their importance is enormous, their use is continually increasing, and they seem destined to produce a great revolution in the civilized world.

Already the steam engine works our mines, impels our ships, excavates our ports and our rivers, forges iron, fashions wood, grinds grains, spins and weaves our cloths, transports the heaviest burdens, etc. It appears that it must some day serve as a universal motor, and be substituted for animal power, waterfalls, and air currents.

Clausius’s bloodstained palms and pasterns (C. Truesdell, *The tragicomical history of thermodynamics 1822-1854*). Clausius' results are of supreme importance for thermodynamics. His way to them was like crawling through a thorn hedge. Mumbo-jumbo with mysterious differentials became and is still the hallmark of thermodynamics. The tradition, handed down in thousands of books, piously follows the traces left by Clausius' bloodstained palms and pasterns.

Simple premises, broad applicability, and permanence (Einstein, *Aerobiological Notes*, 1949). A theory is the more impressive the greater the simplicity of its premises is, the more different kinds of things it relates, and the more extended is 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 the applicability of its basic concepts.

Historical significance (*Feynman's Lectures on Physics*, 1963). What Feynman said about electrodynamics also applies to thermodynamics.

From a long view of the history of mankind—seen from, say, ten thousand years from now—there can be little doubt that the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics. The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade.

The Second Law and Shakespeare (C.P. Snow, *The Two Cultures*, 1959). A good many times I have been present at gatherings of people who, by the standards of the traditional culture, are thought highly educated and who have with considerable gusto been expressing their incredulity at the illiteracy of scientists. 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's?*

Unscientific use of the word "entropy" (N.G. van Kampen, *Views of a Physicist*, 2000). The word "entropy" is in the process of degenerating into a fashionable term. It is being used by authors, biologists, and economists in order to give a scientific odor to their deep thoughts. But whoever wants to utilize physics for his own purpose ought to accept the discipline of exact formulating and exact thinking. This article is an attempt to limit the unchecked growth of the use of entropy. Some examples:

The internal entropy can be transformed into the internal warmth and warmth of Goethianism.

Love is not anti-entropic as some would like.

Such nonsense is not taken seriously by anyone but there are cases that do have the appearance of being scientific.

Logic, intuition, and application (Feller, *An Introduction to Probability Theory and Its Applications*, 1968). The following two paragraphs remain valid when we replace the word probability with thermodynamics.

Probability is a mathematical discipline with aims akin to those, for example, of geometry or analytical mechanics. In each field we must carefully distinguish three aspects of the theory: (a) the formal logical content, (b) the intuitive background, (c) the applications. The character, and the charm, of the whole structure cannot be appreciated without considering all three aspects in their proper relation.

The history of probability (and of mathematics in general) shows a stimulating interplay of theory and applications; theoretical progress opens new fields of applications, and in turn applications lead to new problems and fruitful research. The theory of probability is now applied in many diverse fields, and the flexibility of a general theory is required to provide appropriate tools for so great a variety of needs. We must therefore withstand the temptation (and the pressure) to build the theory, its terminology, and its arsenal too close to one particular sphere of interest. We wish instead to develop a mathematical theory in the way which has proved so successful in geometry and mechanics

First principles (Aristotle, *Metaphysics*). And the things best to know are first principles and causes. For through them and from them all other things may be known but not they through the things covered by them…But these things, the most universal, are perhaps the most difficult things for men to grasp, for they are farthest removed from the senses.

Experiment and theory (Bryan, *Thermodynamics*, 1907). It is maintained by many people (rightly or wrongly) that in studying any branch of mathematical physics, theoretical and experimental methods should be studied simultaneously. It is however very important that the two different modes of treatment should be kept carefully apart and if possible *studied from different books*, and this is particularly important in a subject like thermodynamics.

Cathedral vs. workshop (Lewis and Randall, *Thermodynamics*, 1923). There are ancient cathedrals which, apart from their consecrated purpose, inspire solemnity and awe. Even the curious visitor speaks of serious things, with hushed voice, and as each whisper reverberates through the vaulted nave, the returning echo seems to bear a message of mystery. The labor of generations of architects and artisans has been forgotten, the scaffolding erected for the toil has long since been removed, their mistakes have been erased, or have become hidden by the dust of centuries. Seeing only the perfection of the completed whole, we are impressed as by some superhuman agency. But sometimes we enter such an edifice that is still partly under construction; then the sound of hammers, the reek of tobacco, the trivial jests bandied from workman to workman, enable us to realize that these great structures are but the result of giving to ordinary human effort a direction and a purpose.

Science has its cathedrals, built by the efforts of a few architects and of many workers. In these loftier monuments of scientific thought a tradition has arisen whereby the friendly usage of colloquial speech give way to a certain severity and formality. While this way sometimes promote precise thinking, it more often results in the intimidation of the neophyte. Therefore we have attempted, while conducting the reader through the classic edifice of thermodynamics, into the workshops where construction is now in progress, to temper the customary severity of the science in so far as is compatible with clarity of thought.

Mental training and discipline (Dixon, *A Treatise on Heat*, 1849). The object of these pages, therefore, is not only to convey a knowledge of the leading facts and general laws of the Science of which they treat, but also to assist in the task of mental training and discipline, which forms an important feature of University education. With this object in view, the author has described, as fully as the limits of the work permitted, the details of the principal experimental methods which has been employed in the examination of the phenomena of heat, has pointed out the disturbing causes complicating their results, and, where their object has been to obtain numerical values, he has explained the mode of investigating the formulas which connect such values with the data of experiment.

Dangerous subject (Goodstein, *States of Matter*, 1975). Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hands. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics.

Supplement: Theory of Probability

Sample space. We idealize an experiment as follows. The experiment has many possible outcomes. Every outcome is distinct. A trial of the experiment results in one and only one outcome. Each outcome is called a *sample point*, and the set of all outcomes is called the *sample space*.

A die has six faces, labeled as $\{a,b,c,d,e,f\}$. The die is rolled to flip from one face to another. Each face is a sample point. The set of all the six faces constitute the sample space.

In general, we can designate a sample space by a symbol Γ or by listing all the sample points. Write

$$\Gamma = \{\gamma_1, \gamma_2, ..., \gamma_\Omega\} .$$

The sample space is not a set of numbers, but a set of objects. Each face of a die does not have to be associated with a number, but is something physical. In general, the sample points do not follow the arithmetic rule for numbers. For example, the addition of two faces of a die is meaningless.

Exercise. What is the sample space of a throw of a coin? What is the sample space of throw a coin and a die simultaneously? What is the sample space of a throw of two dies simultaneously? How many sample points do you get when you throw 1000 dies?

Distribution of probability. A distribution of probability is a *function* that maps each sample point to a number. The domain of the function is the sample space $\Gamma = \{\gamma_1, \gamma_2, ..., \gamma_\Omega\}$. The range of the function is the set of real number in the interval [0,1]. Write the map P as

$$P : \Gamma \to [0,1] .$$

Alternatively, write

$$P(\gamma) = p .$$

That is, the distribution of probability P maps every sample point $\gamma \in \Gamma$ to a number $p \in [0,1]$. The probabilities of all sample points sum to unity:

$$P(\gamma_1) + ... + P(\gamma_\Omega) = 1 .$$

In general, different outcomes can have different probabilities.

A die is said to be fair if a rolling results in every face with equal probability. When a fair die is rolled, the probability for the die to settle to a particular face is 1/6.

Exercise. A cheater makes a die of six faces, labeled as $\{a,b,c,d,e,f\}$. Through many throws, she finds that the probability of face a is ½, the probability of face f is 1/10, and the other four faces have an equal probability. What is the probability of getting face b twice in two throws?

Event. Given a sample space, a subset of the sample space is called an *event*. A die has six faces, labeled as $\left\{a,b,c,d,e,f\right\}$. The sample space can form a total of 2^6 subsets. For example, the subset of the sample space, $\left\{b,d,f\right\}$, corresponds to the event that a roll of the die results in either face *b*, or face *d*, or face *f*.

Exercise. An experiment is rolling a die six times. The event "winning" means that the six rolling results in every one of the six faces. How many sample points does the sample space of the experiment have? How many sample points does the event "winning" have?

Probability of an event. A die has six faces, labeled as $\left\{a,b,c,d,e,f\right\}$. When the die is rolled, what is the probability to obtain a face in the subset $\left\{b,d,f\right\}$? We can discover this probability experimentally, by rolling the die many times, and recording the result every time. The probability to obtain a face in the subset $\left\{b,d,f\right\}$ is the recorded number times of faces in the subset divided by the total number of rolling.

Alternatively, we can calculate the probability to observe a face in the subset $\left\{b,d,f\right\}$. We dissect the sample space $\left\{a,b,c,d,e,f\right\}$ into two subsets: $\left\{b,d,f\right\}$ and $\left\{a,c,e\right\}$.

The probability to obtain a face in the subset $\left\{b,d,f\right\}$ = $\frac{3}{6}$.

As another example, consider a sample space $\Gamma=\left\{\gamma_1,\gamma_2,...,\gamma_\Omega\right\}$ with distribution of probability $P(\gamma_1),P(\gamma_2),...,P(\gamma_\Omega)$. Let an event be the subset $A=\left\{\gamma_1,\gamma_6,\gamma_{10}\right\}$. The probability for event A is

$$P(A)=P(\gamma_1)+P(\gamma_6)+P(\gamma_{10}).$$

Study an experiment at a suitable level of detail. The sample space is a special event, so is every sample point. Also, we can regard an event as a sample space: a subset may be considered as a set in its own right; we speak of a subset merely to indicate that we have a larger set in the back of our minds. Furthermore, we can divide a sample space into a family of disjoint events, and then regard each event as a sample point.

This last change of view is particularly important because individual outcomes of an experiment are often too numerous to interest us; we would rather focus on some aggregates of outcomes. When tossing a coin, a gambler is interested in whether the outcome is a head or a tail. Of course, the coin is made of atoms, each atom is made of electrons, protons and neutrons, and each neutron is made of… Associated with a head, the coin has numerous quantum states. But this mode of description is too detailed to interest the gambler, unless she happens to be a physicist.

Dissection. Denote the sample space of an experiment by $\Gamma = \{\gamma_1, \gamma_2, ...\}$. By a *dissection* of the sample space is meant a family $\mathcal{A} = \{A_1, A_2, ..., A_i, ...\}$ of disjoint subsets whose union is the sample space.

For example, a roll of a die has a sample space $\{a,b,c,d,e,f\}$. We can dissect the sample space into two subsets: $\{b,d,f\}$ and $\{a,c,e\}$. The two subsets are disjoint, and their union is the sample space. We can dissect the same sample space in many ways. Here is another dissection: $\{a,b\}$ and $\{c,d,e,f\}$.

Distribute probability among the events in a dissection. The subsets in the dissection in general have a nonuniform distribution of probability, $P(A_1), P(A_2), ... P(A_i), ...$ The subsets in a dissection are disjoint, and the union of the subsets is the sample space. Consequently, the probabilities of all the subsets in the dissection $\mathcal{A}$ sum to unity:

$$P(A_1) + P(A_2) + ... + P(A_i) + ... = 1 .$$

Random variable. A versatile way to dissect a sample space is through the use of a *random variable*. A random variable is a *map* from a sample space to a set of numbers. Let the sample space of an experiment be Γ, and a set of numbers be R. Let X be a map from the sample space Γ to the set of numbers R. Write

$$X : \Gamma \rightarrow R .$$

When a trial of the experiment produces an outcome $\gamma \in \Gamma$, the random variable X takes the value $x \in R$. We also write the map from set to the other set as

$$X(\gamma) = x .$$

The *domain* of the map X is the sample space Γ, which consists of all the outcomes, objects that are usually not numbers. The *range* of the map is a set of numbers R, which obey usual arithmetic rules, such as addition and multiplication. The map is deterministic, but the outcome of a trial of the experiment is random. Consequently, each trial produces a random value of the variable.

Consider a die with faces labeled as $\{a,b,c,d,e,f\}$. Before rolling, we agree on a rule of winning: \$200 for face *a*, \$600 for face *b*, \$100 for face *c*, \$400 for face *d*, \$700 for face *e*, \$0 for face *f*. This rule maps each face to an amount of winning. The domain of the map is the sample space—the set of the six faces. The range of the map is a set of numbers—the amounts of winning. Whereas the rule is deterministic, the face obtained from each roll of the die is random, and the amount of winning after each roll of the die is random.

Exercise. We agree on the above rule of winning. What is probability of winning \$700 after rolling a fair die three times?

From objects to numbers. Instead of using a natural language (English, Chinese, etc.), we now use a real variable to dissect the set of quantum states of an isolated system into a family of configurations. Using a real variable has an obvious advantage: the variable will allow us to use calculus.

Mean of a random variable. Let $\Gamma = \{\gamma_1, \gamma_2, ...\}$ be the sample space of an experiment, and $P(\gamma_1), P(\gamma_2), ...$ be the corresponding probabilities of the sample points. When the outcome of the experiment is γ_i, a variable X takes value $X(\gamma_i)$. Define the *mean* or the *expectation* of the random variable X by

$$\langle X \rangle = \sum P(\gamma_i) X(\gamma_i).$$

The sum is taken over all sample points.

Variance of a random variable. Define the *variance* of a random variable X by

$$Var(X) = \sum P(\gamma_i) \left[X(\gamma_i) - \langle X \rangle \right]^2.$$

The sum is taken over all sample points. The calculation of fluctuation is often aided by the following identity:

$$Var(X) = \langle X^2 \rangle - 2\langle X \rangle \langle X \rangle + \langle X \rangle^2 = \langle X^2 \rangle - \langle X \rangle^2.$$

A dimensionless measure of the fluctuation of a random variable. The unit of the mean is the same as that of the random variable. The unit of the variance is the square of that the random variable. The fluctuation of the random variable X can be measured by a dimensionless ratio

$$\frac{\sqrt{Var(X)}}{\langle X\rangle}.$$

If we double the random variable, the mean will double and the variance will quadruple, but the fluctuation remains the same.

Example. A fair die has six faces labeled as $\{a,b,c,d,e,f\}$. A random variable maps the sample space to a set of six numbers as follows:

$$X(a)=1^2, X(b)=2^2, X(c)=3^2, X(d)=4^2, X(e)=5^2, X(f)=6^2,$$

Thus,

$$\langle X\rangle=\frac{1^2}{6}+\frac{2^2}{6}+\frac{3^2}{6}+\frac{4^2}{6}+\frac{5^2}{6}+\frac{6^2}{6}=\frac{91}{6}$$

$$\langle X^2\rangle=\frac{1^4}{6}+\frac{2^4}{6}+\frac{3^4}{6}+\frac{4^4}{6}+\frac{5^4}{6}+\frac{6^4}{6}=\frac{2275}{6},$$

$$Var(X)=\langle X^2\rangle-\langle X\rangle^2=\frac{2275}{6}-\left(\frac{91}{6}\right)^2=149.08.$$

$$\frac{\sqrt{Var(X)}}{\langle X\rangle}=0.81.$$

Supplement: Entropy of a Distribution of Probability

A sample space of outcomes of unequal probability. Consider a sample space $\Gamma=\{\gamma_1,\gamma_2,...,\gamma_i,...\}$ of an arbitrary distribution of probability $P(\gamma_1),P(\gamma_2),...$ Define the *entropy* S of the distribution of probability by

$$S=-\sum_i P(\gamma_i)\log P(\gamma_i).$$

The sum is over all sample points. As a convention we will use the logarithm of the natural base. Because $x\log x\to 0$ as $x\to 0$, we set $0\log 0=0$. Because $0\le P(\gamma_i)\le 1$, the number $-\log P(\gamma_i)$ is nonnegative.

A sample space of outcomes of equal probability. We have defined the entropy for sample space of a general distribution of probability. We now specialize the definition for a sample space of an equal distribution of probability. Consider a sample space of Ω outcomes:

$$\Gamma=\{\gamma_1,\gamma_2,...,\gamma_\Omega\}.$$

Every outcome is equally probable. Thus, the probability of each event is

$$P(\gamma_i) = 1/\Omega .$$

The entropy is

$$S = \log \Omega .$$

Entropy in the theory of probability. To calculate the entropy, all we need is a distribution of probability. No more, no less. The physical nature of the experiment is irrelevant, be it tossing a coin, rolling a die, or enjoying a glass of wine.

We do not introduce entropy by following the accidents in the history of thermodynamics. Entropy is a mathematical concept independent of the concept of energy and temperature. The entropy of rolling a die is just as valid a concept as the entropy of a glass of wine. Entropy is a number calculated from a distribution of probability.

The situation is like integrating a function. An integral is a number calculated from a function. You may integrate a function in engineering, or a function in economics. There is no difference between an engineer's integration and an economist's integration. The difference is in the applications, not in the integration.

Supplement: Theory of Probability vs. Thermodynamics

We now translate words in the theory of probability of rolling a fair die into the words in the thermodynamics of an isolated system.

Lost in translation. In developing the theory of probability, we do not need to mention who rolls the die. The material of the die is also unimportant. Whether the die is made of wood or silver does not concern us. Woodworking and metallurgy do not concern us, either. Also unimportant is the nature of the individual faces of the die. We label them with numbers 1, 2, 3, 4, 5 and 6, or with letters *a*, *b*, *c*, *d*, *e* and *f*, or with marks of any design. What matter are that the die has six faces, that we can tell them apart, and that a roll of the die results in any one of the faces with equal probability 1/6.

Similarly, in developing thermodynamics, we do not need to mention how an isolated system flips. The material of the isolated system is also unimportant. Whether the system is a glass of wine or a piece of cheese does not concern us. The arts of making wine and cheese do not concern us, either. Also unimportant is the nature of the individual quantum states of the isolated system. What matter are that the isolated system has a certain number of quantum states, Ω, that we can tell them apart, and that the system isolated long enough can flip to any one of its quantum states with equal probability $1/\Omega$.

Sample space. An isolated system has a total of Ω number of quantum states, labeled as $\left\{\gamma_1,\gamma_2,...,\gamma_\Omega\right\}$. The isolated system flips from one quantum state to another rapidly and ceaselessly. We may as well call each quantum state a sample point, and the set of all quantum states the sample space.

In general, the sample points do not follow the arithmetic rule for numbers. For example, the addition of two quantum states is meaningless in general.

Exercise. Describe the sample space of a hydrogen atom isolated at its second energy level.

Fundamental postulate. According to the fundamental postulate, when the system is isolated for a long time, the probability for the isolated system to be in any one of its quantum states is the same, given by

$$\frac{1}{\Omega}.$$

A glass of wine has a huge number Ω of quantum states, so that the probability for a glass of wine to be in any one particular quantum state is tiny.

For a system isolated for a long time, all quantum states are equally probable, and the single number Ω describes the distribution of probability.

Configuration. An isolated system flips among a set of quantum states. What should we call a subset of the quantum states? Calling a subset of quantum states an event is unsuitable. We will call a subset of quantum states a *configuration*.

Probability of a configuration. An isolated system flips among a set of Ω quantum states. Consider a subset of the quantum states. The subset is designated A, and consists of Ω_A number of quantum states. After the system is isolated for a long time, the system can flip to in any one of its quantum states with equal probability. Consequently, the probability for the isolated system to be in subset A is

$$\frac{\Omega_A}{\Omega}.$$

A subset is more probable if the subset consists of more quantum states.

Internal variable. In developing thermodynamics, we often define a random variable by considering a constraint internal to an isolated system. Associated with this constraint is an internal variable. The internal variable maps the set of quantum states of the isolated system to a set of numbers. The internal variable is a random variable.

For example, consider a half bottle of water, and make it an isolated system. This isolated system has a total of Ω number of quantum states, and a total of M number of water molecules. Let the sample space be the set of the quantum states of the isolated system, $\Gamma = \left\{ \gamma_1, \gamma_2, ..., \gamma_\Omega \right\}$. Let the set of numbers be the numbers of water molecules in the top part of the bottle, $R = \left\{ 0, 1, ..., M \right\}$.

In the beginning, we partition the bottle into two parts using a seal. The bottom part of the bottle contains water molecules, and the top part is a vacuum. We then puncture the seal and still make the bottle as a whole an isolated system.

The seal in the bottle provides a *constraint* internal to the isolated system. The act of puncturing a hole in the seal lifts the constraint, making the number of molecules in the top part of the bottle, N, an *internal variable*.

The internal variable is a random variable. When the isolated system is in a quantum state, the top part of the bottle has a certain number of molecules. Thus, the internal variable maps the set of quantum states of the isolated system $\Gamma = \left\{ \gamma_1, \gamma_2, ..., \gamma_\Omega \right\}$ to the set of numbers $R = \left\{ 0, 1, ..., M \right\}$.

Use an internal variable to dissect a sample space. A random variable is often a many-to-one map. It maps many sample points to one number.

Before the seal is punctured, the top part of the bottle has no water molecule, $N = 0$, and the isolated system flips among a subset of quantum states. Denote the number of quantum states in this subset by $\Omega(0)$.

When the top part of the bottle has N water molecules, the isolated system flips among a subset of quantum states. Denote the number of quantum states in this subset by $\Omega(N)$.

We have just used the internal variable N to dissect the set of quantum states of the isolated system into a family of subsets (configurations). These subsets are disjoint, and their union is the whole set of the quantum states of the isolated system. Write

$$\Omega = \Omega(0) + \Omega(1) + ... + \Omega(M).$$

Distribute probability among the configurations in a dissection. In equilibrium, the isolated system flips to every of its Ω quantum states with equal probability. The probabilities to have no molecules, 1 molecule, or M molecules in the top part of the bottle are, respectively,

$$\frac{\Omega(0)}{\Omega}, \frac{\Omega(1)}{\Omega}, \ldots, \frac{\Omega(M)}{\Omega}.$$

This list of numbers constitutes the probability distribution of the dissection.

Supplement: Exponential and Logarithm

Exponential. Consider the relation

$$u = b^x,$$

where b is a positive number. We call u the exponential of x to the base b. If $b = 1$, this relation reduces to $u = 1$. If $b \neq 1$, this relation is a one-to-one map between a number x in the interval $(-\infty, +\infty)$ and a number u in the interval $(0, +\infty)$.

The following identities hold:

$$b^x b^y = b^{x+y},$$

$$\left(b^x\right)^p = b^{px},$$

for any real numbers x, y and p.

Logarithm. For every number b satisfying $b > 0$ and $b \neq 1$, the relation $u = b^x$ is a one-to-one map. Any one-to-one map can be inverted. Given a number u in the interval $(0, +\infty)$, we can find a number x in the interval $(-\infty, +\infty)$ to satisfy the relation $u = b^x$. We write this inverse relation as

$$x = \log_b u.$$

We call x the logarithm of u to the base b.

The following identities hold:

$$\log_b(uv) = \log_b u + \log_b v$$

for any two positive numbers u and v, and

$$\log_b\left(u^p\right) = p \log_b u$$

for any positive number u and any real number p.

The base b can be changed to any other positive number a by using

$$\log_a u = \frac{\log_b u}{\log_b a}.$$

Logarithm converts a product to a sum. The logarithm has the following property:

$$\log_b \left(uv \right) = \log_b u + \log_b v.$$

Is this property unique to logarithm? The answer is yes. To prove this statement, consider a function *f* of the property

$$f\left(uv \right) = f\left(u \right) + f\left(v \right)$$

for any two positive numbers *u* and *v*. We next look for all functions that have this property.

The property $f\left(uv \right) = f\left(u \right) + f\left(v \right)$ implies that

$$f\left(u^m \right) = mf\left(u \right)$$

for any positive real numbers *u* and any integer *m*. In the above formula, write $v = u^m$ or $u = v^{1/m}$, and we obtain that

$$f\left(v^{1/m} \right) = \frac{1}{m} f\left(v \right)$$

for any positive real number *v* and any nonzero integer *m*. From these two expressions, we confirm that

$$f\left(b^{m/n} \right) = \frac{m}{n} f\left(b \right)$$

for any positive real number *b* any integers m and *n*, where $n \neq 0$.

Any real number *x* may be approximated by a rational number *m*/*n*. Consequently, we can re-write the above equation as

$$f\left(b^x \right) = xf\left(b \right).$$

Let $u = b^x$, so that $x = \log_b u$. The above equation is written as

$$f\left(u \right) = A \log_b u,$$

where the number $A = f\left(b \right)$ is independent of *u*. Thus, the function has to be logarithm, and the arbitrary constant *A* allows us to choose any number as the base.

Natural logarithm. Why do we prefer one base over another? Our preference has to do with taking derivatives. The derivative of a function $u(x)$ is defined as

$$\frac{du(x)}{dx} = \lim_{h \to 0} \frac{u(x+h) - u(x)}{h}.$$

Consider the derivative of the exponential function $u(x) = b^x$. Note that

$$\frac{u(x+h) - u(x)}{h} = \frac{b^h - 1}{h} b^x.$$

Consequently, to calculate the derivative of the exponential function, we need to know the number

$$\lim_{h \to 0} \frac{b^h - 1}{h}.$$

This number is a function of b. For a given base b, we can use a calculator to obtain an approximate value of this limit by calculating $(b^h - 1)/h$ with a small number h, say $h = 0.00001$. For example, $(b^h - 1)/h \approx 0.92$ when b = 2.5, and $(b^h - 1)/h \approx 1.10$ when b = 3. By try and error, we find that the above limit is 1 for a particular base, $b \approx 2.718$. This particular base is denoted as $b = e$, and is known as Euler's number. That is, the number e is defined such that

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1.$$

When the number e is used as the base, the derivative of the exponential function takes a particularly simple form:

$$\frac{d(e^x)}{dx} = e^x.$$

By contrast, when an arbitrary number b is used as the base, the derivative of the exponential function is

$$\frac{d(b^x)}{dx} = b^x \log b.$$

When the number e is used as the base, we write the logarithm function as $x = \log u$, and call it the natural logarithm. Thus, in the above equation, $\log b$ means $\log_e b$. The derivative of the natural logarithm also takes a particularly simple form:

$$\frac{d(\log u)}{du}=\frac{1}{u}.$$

When an arbitrary number b is used as the base, the derivative of the logarithm function is

$$\frac{d(\log_b u)}{du}=\frac{1}{u\log b}.$$

If you loathe cluttering many formulas with the number $\log b$, you set the base to be *e*. The choice of the base *e* keeps formulas tidy, but adds no substance.

Supplement: A Particle in a Box

Consider a particle of mass M in a cubic box of side L. The particle is taken to have no internal structure. We describe a point in the three-dimensional space by three coordinates (x,y,z). We associate the particle with a complex-valued function $\Psi(x,y,z)$, called the *wavefunction*. Born (1926) gave the wavefunction an interpretation. The probability for the particle to be in a small volume $dxdydz$ around a point (x,y,z) is $|\Psi(x,y,z)|^2 dxdydz$.

Schrödinger (1926) discovered that the wavefunction is governed by an equation:

$$-\frac{\hbar^2}{2M}\left(\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}+\frac{\partial^2\Psi}{\partial z^2}\right)=E\Psi,$$

where $\hbar \approx 10^{-34}\mathrm{J}\cdot\mathrm{s}$ is the Planck constant, and E is the energy of the particle. In this form, the Schrödinger equation is an eigenvalue problem: the wavefunction $\Psi(x,y,z)$ is the eigenfunction, and E is the eigenvalue.

In the three-dimensional space, the cubic box locates at $0 \le x \le L, 0 \le y \le L, 0 \le z \le L$. The cubic box confines the particle, so that the wavefunction of the particle vanishes outside the box. The wavefunction is continuous in space, so that the wavefunction also vanishes on the faces of the box. These boundary conditions, along with the Schrödinger's equation, give the form of the wavefunction:

$$\Psi = A\sin\left(\frac{n_x \pi x}{L}\right)\sin\left(\frac{n_y \pi y}{L}\right)\sin\left(\frac{n_z \pi z}{L}\right).$$

where n_x, n_y, n_z are any positive integers. Negative integers do not give independent solutions, and a zero does not give a solution. As common in eigenvalue problems, the constant A is undetermined. The wavefunction satisfies the boundary conditions. The wavefunction also satisfy Schrödinger's equation if the energy relates to the integers by

$$E = \frac{\pi^2 \hbar^2}{2ML^2}\left(n_x^2 + n_y^2 + n_z^2\right).$$

Because n_x, n_y, n_z are integers, we cannot prescribe the value of energy arbitrarily. The energy of the system is *quantized*.

Given a particle of mass M and a cubic box of side L, once a value of energy E is prescribed, the particle in the box is an isolated system. The three integers $\left(n_x, n_y, n_z\right)$ together specify a quantum state of the isolated system. For example, we set energy to be consistent with $\left(n_x, n_y, n_z\right) = \left(1,1,2\right)$. This isolated system has three quantum states: $\left(1,1,2\right)$, $\left(1,2,1\right)$ and $\left(2,1,1\right)$. We can picture the three quantum states using their wavefunctions.

Supplement: Hydrogen Atom

A hydrogen atom consists of a proton and an electron. In high schools we were told that the electron "appears" as a cloud around the proton. Sometime later we were told that the electron cloud is governed by a partial differential equation known as Schrodinger's equation.

A quantum state of a hydrogen atom is specified by an orbital and a spin. Even when isolated from the rest of the world, the hydrogen atom can still be in one of multiple quantum states. For example, when the hydrogen atom is isolated at the ground energy level, the atom has two quantum states, characterized by an electron cloud of spherical symmetry and two states of spin. As another example, when the hydrogen atom is isolated at the second energy level, the atom has eight quantum states.

Wavefunction. The electron cloud lives in the same three-dimensional space as we do. The electron cloud of a hydrogen atom is much smaller than us, but is much larger than the proton. Thus, in modeling the hydrogen atom, we regard the proton as a point in the three-dimensional space, and the electron cloud as a smooth distribution in the three-dimensional space.

We describe a point in the three-dimensional space by a radial vector **r** from the proton. We represent the electron cloud by a complex-valued function $\Psi(\mathbf{r})$, known as the *wavefunction*. Born (1926) gave the wavefunction an interpretation. The probability for the electron to be in a small volume dV around a point **r** is $|\Psi(\mathbf{r})|^2 dV$.

Schrödinger equation. Schrödinger (1926) discovered that the wavefunction is governed by an equation:

$$H\Psi = E\Psi,$$

where H is a differential operator, and E is the energy of the electron. You can find the explicit expression of the differential operator H online. In this form, the Schrödinger equation is an eigenvalue problem of the differential operator H. The wavefunction is the eigenfunction, and E is the eigenvalue.

Orbital. Let the three components of the radial vector **r** be x, y and z. The three components are the axes of rectangular coordinates. We can also describe the space using spherical coordinates, with

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta,$$

where r is the radial distance, the angle θ is the colatitude, and the angle ϕ is azimuth.

For the hydrogen atom, we solve the eigenvalue problem using the method of separation of variables. In the spherical coordinates, the wavefunction takes the form

$$\Psi_{nlm}(\mathbf{r}) = R_{nl}(r)Y_{lm}(\theta,\phi).$$

The function $R_{nl}(r)$ describes the radial variation of the wavefunction, and the function $Y_{lm}(\theta,\phi)$ describes the angular variation of the wavefunction. The numbers n, l, m are called quantum numbers. The number n is known as the principal quantum number, and takes values 1, 2, 3… The number l is known as the angular momentum number; for a given n, the number l takes values $l = 0,1,\ldots,(n-1)$. The number m is known as the magnetic quantum number; for given n and l, the number m takes values $m = 0,\pm1,\ldots,\pm l$.

Associated with a given set of the quantum numbers (n,l,m) is a distinct eigenfunction $\Psi_{nlm}(\mathbf{r})$, which is called an *orbital*. For given values of n and l, the number m takes $(2l+1)$ distinct values, giving $(2l+1)$ distinct orbitals. For a given value of n, the numbers l and m form n^2 distinct pairs, corresponding to a total of n^2 orbitals.

The explicit forms of the orbitals are available online. Here are some examples:

$$\Psi_{100} = \frac{1}{a^{3/2}\pi^{1/2}} \exp\left(-\frac{r}{a}\right),$$

$$\Psi_{200} = \frac{1}{a^{3/2}\sqrt{2\pi}} \left(2-\frac{r}{a}\right) \exp\left(-\frac{r}{2a}\right),$$

$$\Psi_{210} = \frac{1}{a^{3/2}4} \left(\frac{3}{2\pi}\right)^{1/2} \left(2-\frac{r}{a}\right) \exp\left(-\frac{r}{2a}\right) \cos\theta,$$

$$\Psi_{21(\pm1)} = \mp \frac{1}{a^{3/2}8\sqrt{\pi}} \left(\frac{r}{a}\right) \exp\left(-\frac{r}{2a}\right) \sin\theta \exp(\pm i\phi).$$

where $a = 0.53\times10^{-10}\mathrm{m}$. This length characterizes the size of the hydrogen atom. Note the dimension of the wavefunction, $\Psi(\mathbf{r}) \sim a^{-3/2}$. This dimension is consistent with the interpretation that $|\Psi(\mathbf{r})|^2 dV$ is a probability, a pure number.

Energy level. For a given value of the principle quantum number n, the energy of the hydrogen atom is

$$E_n = -\frac{13.6\mathrm{eV}}{n^2}.$$

For the hydrogen at the ground energy level (n = 1), the orbital is of spherical symmetry, Ψ_{100}. For the hydrogen atom at the second energy level (n = 1), the electron can be in four: Ψ_{200}, Ψ_{210}, Ψ_{211} and $\Psi_{21(-1)}$. That is, the four eigenfunctions are associated with the same eigenvalue. These eigenfunctions are said to be *degenerate*. Any linear combination of the four orbital is also an orbital associated with the second energy level.

Spin. The Earth not only moves around the Sun, but also spins around its own axis. By analogy, in a hydrogen atom, the electron moves around the proton and also spins. But this analogy is crude. The electron can only have two states of spin, known as up and down.

Quantum state. A quantum state of a hydrogen atom is specified by an orbital and a spin. Thus, a quantum state of the hydrogen atom is specified by four quantum numbers, three associated with the orbital, and one associated with the spin.

Even when isolated from the rest of the world, the hydrogen atom can still be in one of multiple quantum states. For example, when the hydrogen atom is isolated at the ground energy level, the atom has two quantum states, characterized by an electron cloud of spherical symmetry and two states of spin. As another example, when the hydrogen atom is isolated at the second energy level, the atom has eight quantum states.

Supplement: A Drop of Ink in a Glass of Wine

Brownian motion. To have some feel for numbers, consider a drop of ink in a glass of white wine. The ink contains pigment—particles of size less than a micron. The molecules of the wine move ceaselessly, and bombard each ink particle from all directions randomly. At any given time, the bombardment does not cancel out, but results in a net force acting on the ink particle, and causes the ink particles to move. This motion of the particle in the liquid is random, and is known as the *Brownian motion.*

Dispersion. Right after the small drop of ink is put in the glass of wine, the ink particles are all localized in a small volume. After some time, the ink particles disperse throughout the glass of wine. Individual ink particles are in perpetual Brownian motion, but the ink particles distribute in the glass of wine uniformly.

Our experience indicates that dispersion of ink in a liquid is an irreversible process: the ink particles will spontaneously disperse in the liquid, but the dispersed ink particles will not spontaneously all come back to form a drop of ink. What causes the irreversible dispersion?

Dispersion is irreversible. We view the wine and the ink as a single system, and isolate the system from the rest of the world. We call a spatial arrangement of all the ink particles a placement. A placement that all ink particles localize in a small region in the glass is just as probable as a placement that the ink particles disperse in the entire glass. However, there are many more placements that the ink particles disperse in the entire glass than the placements that the ink particles localize in a small region. Consequently, dispersion is more likely than localization.

How much more likely? Let V be the volume of the glass of wine, and N be the number of the ink particles. We have a dilute concentration of the ink particles suspended in the wine. The interaction between the ink particles is negligible, so that each particle is free to explore everywhere in the wine. Consequently, the number of placements of each ink particle is proportional to V. The number of placements of the N ink particles is proportional to V^N. On the other hand, if the N particles localize in a small region, say of volume $V/7$, the number of placements is proportional to $(V/7)^N$. Since all placements are equally likely, the probability to find the N ink particles in a given volume is proportional to the number of placements. Thus,

$$\frac{\text{probability for N particles in volume V}}{\text{probability for N particles in volume V/7}} = \frac{V^N}{(V/7)^N} = 7^N .$$

This ratio is huge if we have more than a few ink particles, a fact that explains why the ink particles much, much prefer dispersion to localization.

Dispersion takes time. Right after the small drop of ink is put in the glass of white wine, the ink particles are all localized in a small volume. The molecules of the liquid move ceaselessly, and bombard each ink particle from all directions randomly. At any given time, the bombardment does not cancel out, but results in a net force acting on the ink particle. This force causes the ink particle to undergo the Brownian motion.

After some time, the ink particles disperse throughout the entire volume of the wine, and the isolated system reaches equilibrium. The time to attain equilibrium depends on how fast the ink particles move, and how large is the volume of the wine. Indeed, if the drop of ink is in a glass of a more viscous liquid, honey, the ink particles will disperse after a longer time that they disperse in the glass of wine.

Einstein (1905) turned this picture into equations. Here we summarize his main results without derivation. The mean displacement L of a particle depends on time t as

$$L = \sqrt{2Dt} .$$

Here D is known as the diffusivity of the particle in the liquid, given by

$$D = \frac{kT}{6\pi a\eta} ,$$

where $k = 1.38\times 10^{-23}\,\mathrm{J/K}$ is known as Boltzmann's constant, T is the temperature in the unit of Kelvin, a is the radius of the ink particle, and η is the viscosity of the liquid.

The viscosity of wine is on the same order as that of water, $\eta = 10^{-3}\,\mathrm{Pa\,s}$. Take the radius of an ink particle to be $a = 1\mu\mathrm{m} = 10^{-6}\mathrm{m}$. At room temperature $T = 300\mathrm{K}$, the diffusivity of the ink particle in water is

$$D = \frac{\left(1.38\times 10^{-23}\,\mathrm{J/K}\right)\left(300\mathrm{K}\right)}{6\pi\left(10^{-6}\mathrm{m}\right)\left(10^{-3}\,\mathrm{Pa\,s}\right)} = 2\times 10^{-13}\,\mathrm{m^2/s}.$$

In one year, the ink particle travels a mean distance of

$$L = \sqrt{2\left(2\times 10^{-13}\,\mathrm{m^2/s}\right)\left(365\times 24\times 3600\mathrm{s}\right)} = 3.6\mathrm{mm}.$$

It takes years for the drop of ink to disperse in the glass of wine.

Exercise. How do you accelerate the dispersion of the ink in the glass of white wine?

Exercise. Estimate the time needed for a drop of ink to disperse in a jar of honey.

References

Textbooks that begin with statistical description. Our presentation above deviates significantly from how the second law was discovered historically. Our presentation, however, is not new. Similar elements of presentation can be found in many textbooks, such as

- T.M. Cover and J.A. Thomas, Information Theory. Second Edition, Wiley, 2006.
- K.A. Dill and S. Bromberg, Molecular Driving Forces, 2nd edition, Garland Science, 2010.
- D.L. Goodstein, States of Matter. Prentice-Hall, 1975.
- H. Gould and J. Tobochnik, Statistical and Thermal Physics, Princeton University Press, 2010.
- D.J. Graham, Chemical Thermodynamics and Information Theory. CRC press, 2011.
- C. Kittel and H. Kroemer, Thermal Physics, W. H. Freeman and Company, New York, 1980.

- L.D. Landau and E.M. Lifshitz, Statistical Physics, Butterworth-Heinemann, 1980.
- D.S. Lemons, A Student's Guide to Entropy. Cambridge University Press, 2013.
- R.K. Pathria, Statistical Mechanics, Pergamon Press, 1972.
- F. Reif, Fundamentals of Statistical and Thermal Physics, McGraw-Hill, Inc., New York, 1965.
- D.V. Schroeder, Thermal Physics, Addison Wesley Longman, 2000.

Classical thermodynamics. There are many carefully written textbooks of classical thermodynamics. These textbooks do not rely on the fundamental postulate, but on historical development of thermodynamics. Here are a few such textbooks:

- C.J. Adkins, Equilibrium Thermodynamics, 3rd edition, Cambridge University Press, 1983.
- G.H. Bryan, Thermodynamics, 1907.
- H.B. Callen, Thermodynamics and Introduction to Thermostatistics, 2nd edition, Wiley, 1985.
- E. Fermi, Thermodynamics, Dover Publications, 1956.
- J. Kestin, A Course in Thermodynamics. Blaisdell Publishing Company, 1966.
- D.S. Lemons, Mere Thermodynamics, The Johns Hopkins University Press, 2009.
- A.B. Pippard, The elements of Classical Thermodynamics, Cambridge University Press, 1957.
- M.W. Zemansky and R.H. Dittman, Heat and Thermodynamics, McGraw-Hill, Inc. 7th edition, 1997.

Engineering thermodynamics. A large number of texts exist for undergraduate students in mechanical engineering. Emphasis is on the applications of thermodynamics to engines and refrigerators. Historically, the invention and development of these devices created the science of thermodynamics.

- R.T. Balmer, Modern Engineering Thermodynamics, Academic Press, 2011.
- C. Borgnakke and R.E. Sonntag, Fundamentals of Thermodynamics, 8th edition, Wiley, 2013.
- Y.A. Cengel and M.A. Boles, Thermodynamics, 7th edition, McGraw Hill, 2010.

- J.H. Keenan, Thermodynamics, Wiley, 1941.
- S. Klein and G. Nellis, Thermodynamics, Cambridge University Press, 2012.
- M.J. Moran, H.N. Shapiro, D.D. Boettner and M.B. Bailey, Fundamentals of Engineering Thermodynamics, 7th edition, 2011.
- W.C. Reynolds and H.C. Perkins, Engineering Thermodynamics, McGraw Hill, 1977.
- G. Rogers and Y. Mayhew, Engineering Thermodynamics, 4th edition, Longman, 1992.
- E. Schmidt, Thermodynamics. Dover, 1966.
- O. Singh, Applied Thermodynamics, 3rd edition, New Age, 2009.
- H. Struchtrup, Thermodynamics and energy conversion. Springer, 2014.

Chemical thermodynamics. Reactions. Mixtures. Phases. Electrochemistry.

- K. Denbigh, The Principles of Chemical Equilibrium, 4th edition, Cambridge University Press, 1981.
- H. Devoe, Thermodynamics and Chemistry, Prentice Hall, 2001. The author has made second edition freely available online.
- J.D. Fast, Entropy. Palgrave Macmillan, 1970.
- G.N. Lewis and M. Randall, Thermodynamics and the Free Energy of Chemical Substances, 1923.
- A.P.H. Peters, Concise Chemical Thermodynamics, 3rd edition, CRC Press, 2010.
- S.I. Sandler, Chemical, Biological, and Engineering Thermodynamics, 4th edition, Wiley, 2006.
- E.B. Smith, Basic Chemical Thermodynamics, 5th edition, Oxford University Press, 2004.

Biological thermodynamics.

- D.T. Haynie, Biological Thermodynamics, 2nd edition, Cambridge University Press, 2008.
- G.G. Hammes, Thermodynamics and Kinetics for Biological Sciences, Wiley, 2000.

Nonequilibrium thermodynamics.

- D. Kondepudi and I. Prigogine, Modern Thermodynamics. 2nd edition, Wiley, 2014.

- S.R. De Groot and P. Mazur, Non-equilibrium Thermodynamics, Dover publications.
- I. Prigogine, Introduction to Thermodynamics of Irreversible Processes.

Popular accounts of thermodynamics.

- P.W. Atkins, The 2nd Law: Energy, Chaos, and Form, W.H. Freeman, 2nd edition, 1994.
- P.W. Atkins, Four Laws That Drive the Universe, Oxford University Press, 2007.
- A. Ben-Naim, Entropy Demystified: The Second Law Reduced to Plain Common Sense, World Scientific Publishing Company, 2007.
- P.W. Bridgman, The Nature of Thermodynamics, Harper & Row, 1941.
- J.B. Fenn, Engines, Energy, and Entropy, W.H. Freeman, 1982.
- I. Prigogine and I. Stengers, Order out of Chaos, Bantam Books, 1984.
- J.E.J. Schmitz, The Second Law of Life, William Andrew Publishing, 2007.
- H.C. von Baeyer, Maxwell's Demon. Random House, 1998.

Old books. Some are written by the founders of thermodynamics. Some are textbooks that influenced the founders. These old books are freely available at Google Books. They are more authentic in conveying the magic of discovery.

- S. Carnot, Reflections on the Motive Power of Fire, 1824.
- R.J.E. Clausius, The Mechanical Theory of Heat, 1867. This edition consists of the English translation of nine papers published by Clausius on thermodynamics.
- R.J.E. Clausius, The Mechanical Theory of Heat, 1879. This edition reworked the original papers into a form of textbook.
- R.V. Dixon, A Treatise on Heat, 1849.
- J.W. Gibbs, The Scientific Papers of J. Willard Gibbs, Volume 1, Thermodynamics, 1906.
- J.C. Maxwell, Theory of Heat, 1871.
- M. Planck, Treatise on Thermodynamics, 1922.
- W.J.M. Rankine, A Manual of the Steam Engine and Other Prime Movers, 1908.
- J. Tyndall, Heat Considered as a Mode of Motion, 1865.

History of thermodynamics. Thermodynamics was established within the span of the lifetimes of Carnot (1796-1832) and Boltzmann (1844-1906). The history is short and interesting, involving about a handful of main characters. You can have a quick introduction to the history of thermodynamics on Wikipedia. The following items will lead you to anecdotes and original papers.

- M. Bailyn, A Survey of Thermodynamics. AIP press, 1994.
- S.G. Brush, The Kind of Motion We Call Heat. North-Holland Publishing Company, 1976.
- S.G. Brush, The Kinetic Theory of Gases. An Anthology of Classic Papers with Historical Commentary. Imperial College Press (2003).
- D.S. Cardwell, From Watt to Clausius. Cornell University Press (1971).
- J. Kestin, ed., The Second Law of Thermodynamics. A selection of original papers by Carnot, Clausius, Kelvin, Gibbs, etc.
- S.I. Sandler, L.V. Woodcock, Historical observations on laws of thermodynamics. Journal of Chemical Engineering Data 55, 4485-4490 (2010).
- C. Smith, The Science of Energy, The University of Chicago Press, 1998.
- I. Muller, A History of Thermodynamics. Springer, 2006.

Unclassified items.

- J. Israelachvili, Intermolecular & Surface Forces. 3rd ed., Academic Press, 2011.
- J.M. Marciano, Whatever Happened to the Matric System? How America Kept Its Feet. Bloomsbury, 2014.
- P. Perrot, A to Z of Thermodynamics. Oxford University Press, 1998.