An introduction to thermodynamics

Zhigang Suo, Harvard University, suo@seas.harvard.edu

I will keep updating this google doc, and draw figures in class. Please access the current version online, and keep good notes in class.

The play of thermodynamics

Leading role. 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*. Entropy plays the leading role in thermodynamics. Energy crisis is a timely topic; entropy crisis, timeless.

Supporting roles. In thermodynamics, energy plays a supporting role, along with space, matter, and charge. Indeed, these supporting roles are analogous to one another, and are of equal importance. Calling thermodynamics the science of energy distorts the structure of the subject, and neglects obviously significant roles of space, matter, and charge.

Children and grandchildren. Each of these supporting roles, together with entropy, produces a child. The four children—temperature, pressure, chemical potential, and voltage—are the second generation of supporting roles. They produce grandchildren: thermal capacity, compressibility, coefficient of thermal expansion, etc.

Extras. There are also some extras: enthalpy, Helmholtz function, Gibbs function, etc. They are called thermodynamic potentials, introduced by Gibbs (1875). They are shadows of entropy. Let no shadows obscure the real thing—entropy.

The Cast of Thermodynamics

Leading role:
Entropy

Supporting roles:
Energy
Space
Matter
Charge

Children of entropy and the supporting roles:
Temperature
Pressure
Chemical potential
Voltage

Grandchildren:
Thermal capacity
Compressibility
Coefficient of thermal expansion
…...

Extras:
Enthalpy
Helmholtz function
Gibbs function
…...

This course. This course will develop the *logic* of entropy from first principles, *intuition* of entropy from everyday experience, and *application* of entropy in many domains.

We will let everyday experience, along with discoveries and inventions, reveal the long arm of entropy, and its intriguing plays with the supporting roles.

I am writing this file for lectures that focus on thermodynamics itself, rather than its applications. Any one of the standard textbooks will fill the rest of the course with copious applications.

This course does not teach the history of thermodynamics. It is impractical to teach thermodynamics by tracing the steps of Carnot, Clausius, Boltzmann, and Gibbs, just as it is impractical to teach calculus by tracing the steps of Newton and Leibniz. A subject and its history are different things. Thus said, the history of thermodynamics is interesting, important, and well-documented, full of moments of triumph and despair. Many original works are available online in English. I will place a few names and years in the notes as landmarks. You can read the history of thermodynamics online, starting with the Wikipedia entry.

Entropy

The definition of entropy requires two ideas: isolated system and sample space.

Isolated system

System. We have met the cast. Now look at the stage—the *world*. Any part of the world is called a *system*. The rest of the world is called the *surroundings*.

We can regard any part of the world as a system. Even the empty space can be a system; the vacuum hosts electromagnetic field.

A proton and an electron constitute a system, called a hydrogen atom.

A half bottle of water is a system. The system is composed of water molecules and some other molecules, such as nitrogen, oxygen, and carbon dioxide. In the half bottle of water, liquid occupies some volume, and gas fills the rest. The liquid and the gas together constitute the system. Do we include the plastic bottle as a part of the system? Maybe, if we decide to study the permeation of water molecules through the plastic. The decision is ours.

Interaction between a system and its surroundings. A system and its surroundings can have many modes of *interaction*. The hydrogen atom changes the shape of its electron cloud when the atom absorbs or emits photons, or when the atom is subject to an electric field.

I hold the half bottle of water in my hand. I see the water because the liquid-gas interface refracts light. I feel moist because water molecules hit me. I warm up the water when the vibration of the molecules in my hand couples the vibration of the molecules in the water. When

I drink from the bottle, the bottle transfers water molecules to my body. I shake the bottle and hear the sound. I pour honey into water and watch them mix.

Isolated system. Our play—thermodynamics—shows all modes of interaction between a system and its surroundings. But our narrative begins with something simpler: an *isolated system*—a system that does not interact with its surroundings.

To make the half bottle of water an isolated system, I cap the bottle to prevent molecules from leaking in and out. I insulate the bottle in a thermos to block the vibration of the molecules in my hand from coupling with the vibration of molecules in the water. I make the bottle rigid to fix the volume. I do not shake the bottle. I am alert to any other modes of interaction between the water and the surroundings. Does the magnetic field of the earth affect the water? If it does, I will find a way to shield the bottle of water 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 I choose to study. For example, it may be too much trouble for me to isolate the water from gravity. Few people care to study water under the zero-gravity condition. Gravity is important if I move the bottle around, but unimportant if I study the vapor pressure in the bottle.

Exercise. Describe a system and what you need to do to make it an isolated system.

Sample space

Quantum states of an isolated system. When a hydrogen atom is isolated at the second energy level, the isolated system has eight quantum states.

Quantum mechanics governs all systems, however complicated. A quantum state of the half bottle of water is a cloud of electrons and positions of nuclei. Such a macroscopic isolated system has a large, but finite, number of quantum states.

Sample space. In the theory of probability, each trial of an experiment is assumed to result in one of multiple possible outcomes. Each possible outcome is called a *sample point*. The set of all possible outcomes of the experiment is called the *sample space*. The notion of sample space comes from Mises (1919).

An isolated system is an “experiment” in the sense of the word used in the theory of probability. The isolated system flips from one quantum state to another, rapidly and ceaselessly. Each quantum state is a possible outcome, or a sample point, of the isolated system. All the quantum states of the isolated system constitute the *sample space of the isolated system*.

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 possible outcomes do you get when you throw 1000 dies?

Exercise. What is the sample space of a hydrogen atom isolated at the second energy level? Sketch the electron clouds of the quantum states.

Definition of entropy

Now enters the leading role—entropy. Let Ω be the number of quantum states of an isolated system. Define the *entropy* of the isolated system by

$S = \log \Omega$.

Logarithm of any base will do. For convenience, we will use the natural base e. The number e simplifies the derivative of logarithm. Recall a fact of calculus: $d \log x / dx = 1/x$. For any other base b, recall that $\log_b x = (\log_b e)(\log_e x)$, so that $d \log_b x = (\log_b e)/x$. The prefactor clutters the formula and serves no purpose.

Entropy is an extensive quantity. Why do we hide Ω behind a log? Consider two isolated systems, A and B. Isolated system A has one sample space of Ω_A quantum states, labeled as

$\{a_1, a_2,..., a_{\Omega A}\}$.

Isolated system B has another sample space of Ω_B quantum states, labeled as

$\{b_1, b_2,..., b_{\Omega B}\}$.

The two systems are *separately isolated*. Together they constitute a *composite*, which is also an isolated system. Each quantum state of this composite is a combination of a quantum state of one isolated system, a_i, and a quantum state of the other isolated system, b_j. All such combinations together constitute the sample space of the composite. The total number of all such combinations is the product:

$\Omega_{\text{composite}} = \Omega_A \Omega_B$.

Recall a property of logarithm: $\log (\Omega_A \Omega_B) = \log \Omega_A + \log \Omega_B$. Thus, the entropy of a composite of two separately isolated systems is the sum of the entropies of the two individual isolated systems.

We now see the significance of logarithm: it turns a product to a sum. The entropy of a system is the sum of the entropies of its parts, each part being separately isolated. Such an additive quantity is called an *extensive quantity*.

Entropy of a pure substance. A pure substance is a collection of a large number of a single species of molecules (or atoms). The molecules can aggregate into various forms, called *phases*. For example, at room temperature and atmospheric pressure, 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.

Entropy is an extensive quantity. The entropy of a piece of a pure substance is proportional to the number of molecules in the piece. For a piece of a pure substance, having the number of molecules N and entropy S, the entropy of the substance per molecule is

$s = S/N$.

Later we will describe how to measure entropy experimentally. For now, we look at some measured numbers. At 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, at room temperature and atmospheric pressure, the entropies for O, O_2 and O_3 are 19.4, 24.7, 28.6, respectively.

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

Zero entropy. The entropy of a pure substance is often tabulated, at the end of textbooks and online, as a function of temperature and pressure. Such a table often assumes 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 the definition of entropy, $S = \log \Omega$. Zero entropy is not something arbitrary, but has physical significance: zero entropy corresponds to an isolated system of a single quantum state.

Exercise. The entropy of a throw of a fair die is log 6. The entropy of an isolated system is $\log \Omega$. 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. 12 grams of diamond has 6.02×10^{23} number of carbon atoms. How many quantum states are there in one gram of diamond at room temperature and atmospheric pressure?

Exercise. Carbon atoms can also aggregate in other forms, such as graphene, nanotube, and buckyball. Learn about these forms online, and find the entropy per atom in each form.

Fundamental postulate

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/Ω .

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.

Fact 3 of the world is called the *fundamental postulate*. 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.

The fundamental postulate links thermodynamics to probability. Our world acts like a compulsive gambler, ceaselessly and rapidly throwing fair dies, each having an enormous number of faces.

	Probability	Thermodynamics
Experiment	Roll a fair die	Isolate a system for a long time
Sample space	6 faces	Ω quantum states
Probability of a sample point	1/6	1/Ω
Subset	Event	Subset
Probability to realize a subset	(number of faces in the subset)/6	(number of quantum states in the subset)/Ω

A function from sample space to a set of values	Random variable	Internal variable

Exercise. A cheater makes an unfair die of six faces, labeled as a,b,c,d,e, f. Through many throws, the cheater 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?

Ignorance is bliss. In throwing a die, the gambler does not need to know the material that makes the die, or the symbols that mark the faces. All the gambler needs to know about the die is that it has six faces of equal probability.

The same is true in thermodynamics. In studying an isolated system, we do not need to know the quantum states themselves (the shape of the cloud of electrons, the positions of nuclei, or the number of protons). We just need to know how many quantum states that the isolated system has.

This enormous reduction of information is central to the success of thermodynamics. An isolated system is reduced to a pure number, the number of quantum states, Ω. Later we will learn how to count the number of quantum states of an isolated system experimentally.

Dispersion of ink

Empirical facts. Let us watch the fundamental postulate in action. Drip a drop of ink into a bottle of water, and the ink disperses over time. The dispersion of ink is readily observed at macroscopic scale, but can also be observed in a microscope, as described below.

The ink contains pigment particles of size less than a micron. Each pigment particle is bombarded by water molecules, rapidly and ceaselessly, from all directions. At any given time, the bombardments do not fully cancel out, but result in a net force that moves the pigment particle. This rapid, ceaseless, random motion of a particle in a liquid was first observed in a microscope by Brown (1827). Wiki Brownian motion.

Thermodynamic theory. Individual pigment particles move in all directions *randomly*. After some time, the pigment particles disperse in the bottle of water *homogeneously*. How can the random motion of individual pigment particles cause pigment particles collectively to do something directional—dispersion?

The answer is simple. When the concentration of pigment particles is inhomogeneous, more particles will diffuse from a region of high concentration to a region of low concentration. This bias continues until the pigment particles are distributed homogeneously.

Let us relate this everyday experience to the fundamental postulate. To make a definite calculation, we assume that the pigment particles in water are far apart, so that each particle is free to explore everywhere in the bottle of water, unaffected by the presence of other pigment particles. Consequently, the number of quantum states of each pigment particle is proportional to the volume of the water in the bottle, V. The number of quantum states of N pigment particles is proportional to V^N.

On the other hand, if the N pigment particles are localized in a small region, say, in the initial drop of ink of volume V/70, the number of quantum states of the N pigment particles is proportional to $(V/70)^N$.

After the ink is in water for a long time, all quantum states are equally probable. Thus, the ratio of the probability of finding the N pigment particles in volume V to the probability of finding the N pigment particles in volume V/70 is 70^N. This ratio is enormous because a drop of ink has a large number of pigment particles, N. This fact explains why the pigment particles much, much prefer dispersion to localization.

Exercise. The density of the pigment material is 1000 kg/m^3. Assume each pigment particle is a sphere, diameter 100 nm. How many pigment particles are there in 1g of dry ink? After the ink is in a bottle of water, volume 100 ml, for a long time, what is the ratio of the probability of finding all pigment particles in a volume of 10 ml to the probability of finding all pigments in the volume of 100 ml?

Equilibrium

The dispersion of ink illustrates several characteristics common to all isolated systems.

Right after a small drop of ink enters the bottle of water, all the pigment particles are localized in the drop. The pigment particles then start to diffuse into the pure water. After some time, the pigment particles disperse in the bottle of water homogeneously, and the system of the pigment particles in water is said to have reached *equilibrium*.

Right after isolation, the system has Ω quantum states, flips to some of them more often than others, and is said to be out of equilibrium. Out of equilibrium, the probability for the isolated system to be in a quantum state is time-dependent.

After being isolated for a long time, the system flips to every one of its quantum states with equal probability, 1/Ω, and is said to have reached equilibrium. In equilibrium, the probability for the isolated system to be in a quantum state is time-independent.

Several phrases are 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”. Whenever we speak of equilibrium, we identify a system isolated for a long time.

In or out of equilibrium, an isolated system flips from one quantum state to another, ceaselessly and rapidly.

In or out of equilibrium, an isolated system has a fixed sample space of Ω quantum states. The entropy of the isolated system, $S = \log \Omega$, is a fixed number and does not change, no matter whether the isolated system is in or out of equilibrium.

Irreversibility

So long as the bottle is isolated, the homogeneously dispersed pigment particles will be extremely unlikely to come back into a small volume.

Once in equilibrium, the isolated system will *not* go out of equilibrium. The isolated system out of equilibrium is said to approach equilibrium with *irreversibility*.

Thermodynamics uses 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 between “before” and “after”. Irreversibility gives time the direction, the *arrow of time*.

Fluctuation

In equilibrium, the pigment particles keep in ceaseless Brownian motion. The distribution of the pigment particles *fluctuates*. Possibly all pigment particles can move into a small volume in the bottle. However, the probability of finding a nonuniform distribution is exceedingly small. This course will ignore fluctuation.

Kinetics

A system isolated for a long time flips to every one of its quantum states with equal probability. How long is long enough? The fundamental postulate is silent on this question. The pigment particles disperse slower in honey than in water. The time needed to reach equilibrium scales with the viscosity of the liquid. The study of how fast a system evolves is called *kinetics*, which will not be studied in this course.

Constrained equilibrium

When an isolated system is in equilibrium, any part of the isolated system is also in equilibrium, so long as the part is macroscopic. It makes no sense to talk about equilibrium at the level of a few pigment particles, but equilibrium will prevail in a large number of pigment particles.

We can divide the isolated system into many parts. Each part is large compared to individual particles, but small compared to the entire isolated system. We regard each part as an isolated subsystem. Each isolated subsystem has its own sample space of quantum states. Thus, the entropy of an isolated system is the sum of the entropies of all parts of the system.

When an isolated system is out of equilibrium, we often divide the isolated system into parts. For example, before the ink is fully dispersed in the bottle of water, we may divide the bottle into many small volumes. Each small volume has a large number of pigment particles, which are approximately homogeneously distributed, so that we can think of each small volume as an isolated system, with its own sample space of quantum states. We say that the isolated system is in *constrained equilibrium.*

Separation of phases

Empirical facts. An isolated system in equilibrium can be heterogeneous. Here is the half bottle of water again. As I shake the bottle, water moves and bubbles pop. After I stop shaking, the half bottle of water becomes approximately an isolated system. Right after the isolation, the half bottle of water is still out of equilibrium. After being isolated for some time, the half bottle of water calms down at macroscopic scale. In equilibrium, some water molecules form the liquid, and others form the vapor. Water molecules in the bottle are said to separate into two *phases*, liquid and vapor.

Thermodynamic theory. We are not ready to develop a full theory of phase separation, but begin with a few ideas here, and pick them up later. The isolated system flips among a set of quantum states, which constitute the sample space of the isolated system. Denote the number of quantum states of the isolated system by Ω.

The half bottle of water has a total of M water molecules. The number of water molecules in the vapor, N, can take one of a set of values:

$\{0,1,...,M\}$.

When the number of water molecules in the vapor is fixed at N, the isolated system flips among quantum states in a *subset of the sample space*. Denote the number of quantum states in this subset by $\Omega(N)$. Thus,

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

Probability. After the bottle is isolated for a long time, every quantum state in the sample space is equally probable, so that the probability to observe N water molecules in the vapor is

$\Omega(N)/\Omega$.

In equilibrium, the most probable amount of molecules in the vapor, N, maximizes the function $\Omega(N)$.

From probability to (almost) certainty. The half bottle of water is a macroscopic isolated system, and has a sample space of a large number of quantum states. In equilibrium, the fluctuation in the number of water molecules in the vapor is exceedingly small, and the observed amount of water molecules in the vapor is well described by the amount N that maximizes the function $\Omega(N)$. The fluctuation in the number of molecules in the vapor is negligible compared to the total number of molecules in the vapor. This observation indicates that the function $\Omega(N)$ has a sharp peak.

Internal variable

Constraint internal to an isolated system. Fixing the number of water molecules in the vapor in the half bottle of water is an example of a *constraint internal to an isolated system*. The constraint can be made real by placing a seal between the liquid and vapor. With the seal, the two parts of the isolated system can separately reach equilibrium, but are not in equilibrium with each other. The isolated system is in constrained equilibrium.

When the constraint is removed, the number of water molecules in the vapor can change, and is called an *internal variable*. Let us abstract this example in general terms.

Sample space. An isolated system has a set of quantum states, which constitute the sample space. Denote the total number of quantum states of the isolated system by Ω.

Subset of sample space. Let X be a set of values:

$X = \{x_1,..., x_n\}$.

An internal variable is a function that maps the sample space to the set X. In the theory of probability, such a function is called a *random variable*.

When a constraint internal to the isolated system fixes the internal variable at a value x in the set X, the isolated system flips among quantum states in a *subset of the sample space*. Denote the number of quantum states in this subset by $\Omega(x)$.

The internal variable dissects the sample space into a family of subsets. Any two subsets in the family are disjoint. The union of all the subsets in the family is the sample space. Thus,

$\Omega(x_1) + \ldots + \Omega(x_n) = \Omega$.

Probability. After the constraint is removed for a long time, the isolated system flips to every one of its Ω quantum states with equal probability, and the internal variable can take any value in X. In equilibrium, the probability for the internal variable to take a particular value x in X is

$\Omega(x)/\Omega$.

Equilibrium. A macroscopic isolated system has a sample space of a large number of quantum states. The function Ω(x) has a sharp peak. After the constraint is removed for a long time, the observed value of the internal variable is well described by the value x that maximizes the function Ω(x).

Thermodynamics studies an isolated system in constrained equilibrium, and determines the most probable value of the internal variable. Thermodynamics neglects the fluctuation in an internal variable, and does not study the kinetics of how the isolated system approaches equilibrium.

Irreversibility. Right after being isolated, the system is out of equilibrium. As time moves forward, the isolated system evolves toward equilibrium, and the internal variable changes in a sequence of values that increase the function Ω(x). So long as the system is isolated, the change in the internal variable is irreversible.

Basic algorithm of thermodynamics

Function Ω(x). An internal variable is a function that maps the sample space of an isolated system to a set X. When the internal variable takes a value x in the set X, the isolated system flips among the quantum states in a subset of the sample space. Denote the number of quantum states in the subset by Ω(x).

Function S(x). Define $S(x) = \log \Omega(x)$. Because logarithm is an increasing function, maximizing Ω(x) is equivalent to maximizing S(x), and increasing Ω(x) is equivalent to increasing S(x).

The function S(x) stands for "the logarithm of the number of quantum states in the subset of the sample space of an isolated system when an internal variable is fixed at a value x". This function is central to the application of thermodynamics, but is unnamed. For brevity, I will call the function S(x) the *subset entropy*.

Basic algorithm of thermodynamics. Here is how we use entropy in thermodynamics.

1. Construct an isolated system with an internal variable x.

2. Identify the function S(x).
3. *Equilibrium*. Find the value of the internal variable x that maximizes the function S(x).
4. *Irreversibility*. Change the value of the internal variable x in a sequence that increases the function S(x).

The second law of thermodynamics

The basic algorithm is one of many alternative statements of *the law of the increase of entropy*, or *the second law of thermodynamics*.

For entertainment, we will later list some historical statements of the second law of thermodynamics. They may sound like ancient philosophical pronouncements. They sound mysterious not because they are more profound than the basic algorithm, but because they miss basic facts of the world (e.g., the rapid and ceaseless flips among quantum states, and the fundamental postulate).

Here is one such statement made by Clausius (1865), in the same paper in which he made up the word entropy:

The entropy of the universe tends to a maximum.

We will take the word *universe* to mean an isolated system. This statement is adopted by numerous textbooks. The statement is elegant but confusing. The entropy of an isolated system is S = log Ω, where Ω is the total number of quantum states in the entire sample space of the isolated system. The entropy of an isolated system is a fixed number, not a function that can change values.

To talk about a macroscopic change of an isolated system, we need to identify an internal variable. When a constraint internal to the isolated system fixes the internal variable at a value x, the isolated system flips among quantum states in a subset of the sample space. The number of quantum states in this subset is Ω(x), and the entropy of this subset is S(x) = log Ω(x). When the constraint internal to the isolated system is removed, the internal variable x can change values. It is the subset entropy S(x) that tends to a maximum. When elegance and clarity conflict, we go for clarity.

This course does not study the history of thermodynamics. We do not attempt to read the mind of Clausius and decipher old pronouncements. Rather, we will use the basic algorithm to direct calculation and measurement. In particular, the basic algorithm will let us count the number of quantum states experimentally.

Energy, space, matter, charge

The basic algorithm calls for internal variables. Now enters the particular showy supporting actor—energy. Energy serves as an internal variable in thermodynamics, along with space, matter, and charge.

Potential energy

An apple weighs about 1 Newton. When I pick up the apple from the ground, the apple reaches about 1 meter high and adds about 1 Joule of energy. This form of energy is called the *potential energy* (PE). From mechanics you have learned the fact:

PE = (weight)(height).

Potential energy has the unit of force times length, (Newton)(meter). This unit of energy is called the Joule.

The height is relative to some fixed point, such as the ground. Thus, potential energy is a *relative quantity*. For a given height, the potential energy is proportional to the amount of material. Thus, potential energy is also an *extensive quantity*. Indeed, all forms of energy are relative and extensive.

From mechanics you have learned another fact:

weight = (mass)(acceleration of gravity).

An apple has a mass about 0.1 kg. The acceleration of gravity is about 10 m/s^2. Thus, the weight of the apple is (0.1 kg)(10 m/s^2) = 1 Newton.

Definition of energy

Energy is whatever that can lift a weight to some height. By this definition, energy is conserved, relative, and extensive.

The definition of energy calls for action. It is up to people to discover energy in its various forms, and invent ways to convert energy from one form to another. How do we know that something offers a form of energy? Just test if this something can be converted to lift a weight to some height.

Forms of energy

Kinetic energy. My hand now releases the apple. Just after the release, the apple is 1 meter high and has zero velocity. The falling apple then loses height, but gains velocity. The energy associated with the velocity of a mass is called the *kinetic energy* (KE). The falling apple *convert* potential energy to kinetic energy. Mechanics tells us that

KE = (½)(mass)$(\text{velocity})^2$.

The conservation of mechanical energy. The friction between the apple and the air is negligible. Mechanics tells you that, as the apple falls, the sum of the potential energy and the kinetic energy is constant—that is,

PE + KE = constant.

The potential energy and the kinetic energy are two forms of *mechanical energy*. The above equation says that mechanical energy is *conserved* when friction is negligible.

Exercise. What is the velocity of the apple just before hitting the ground?

Exercise. A tiger jumps 1 m high. What is the velocity of the tiger just before it hits the ground?

Exercise. Derive the conservation of mechanical energy from Newton's second law.

Thermal energy (internal energy). Watch the apple fall again. The apple falls from a height, and gains a velocity just before hitting the ground. After hitting the ground, the apple bumps, rolls, and then stops. What happens to all that potential energy and kinetic energy?

Just before the apple hits the ground, all potential energy has converted to kinetic energy. After the apple hits the ground and comes to rest, all the kinetic energy of the apple disperses, or *dissipates*, into the motion of the molecules inside the apple and the ground. We say that the apple and the ground gains *thermal energy* (TE), also called *internal energy*. We generalize the principle of the conservation of energy to

TE + PE + KE = constant.

The word *thermal* is an adjective that describes phenomena related to microscopic interaction and motion. Thermal energy is just the potential energy and kinetic energy at microscopic scale. We designate PE and KE as the potential energy and kinetic energy at macroscopic scale.

In *Heat considered as a Mode of Motion*, published in 1863, John Tyndall described numerous experiments that tested the hypothesis of heat as a form of energy. Wiki Julius Robert von Mayer. Wiki James Prescott Joule. The following passage is taken from the book.

A bullet, in passing through the air, is warmed by the friction, and the most probable theory of shooting stars is that they are small planetary bodies, revolving round the sun, which are caused to swerve from their orbits by the attraction of the earth, and are raised to incandescence by friction against our atmosphere.

Electrical energy. Electrical energy (EE) takes many forms. One way to use electrical energy is the resistive heating. A voltage of an electric outlet moves electrons in a metal wire, and the resistance of the metal converts the electrical energy into thermal energy. Recall voltage = (resistance)(current). The electrical energy is

$$EE = (\text{resistance})(\text{current})^2(\text{time}) = (\text{voltage})(\text{current})(\text{time}).$$

Chemical energy. Chemical energy (CE) has always been familiar to humans in the form of fire and food. We will learn how to measure chemical energy later in the course.

Exercise. Describe how one can test if electrostatics offers a form of energy.

Exercise. Find the nutrition energy of a banana. If this nutrition energy is fully converted to the potential energy of the banana, what will be the height of the banana?

Energy belongs to many sciences

Energy plays parts 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, food, blood, wind, rivers, 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. These facts do not originate from thermodynamics, but we will use them just as we use facts in calculus. We do not steal lines from other sciences; we borrow.

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.

	to thermal	to mechanical	to electrical	to chemical
thermal	heat exchanger	engine	thermocouple	reaction
mechanical	friction	turbine	generator	fracture
electrical	resistor	compressor	capacitor	charging battery
chemical	fire, food	muscle	battery	reaction

Exercise. What is the function of a turbine? How does it work? Link your answer to a video online.

Exercise. What is the function of a generator? How does it work? Link your answer to a video online.

An isolated system conserves energy, space, matter, and charge

Beside energy, we now add a few other supporting roles: space, matter, and charge.

I make a half bottle of water into an isolated system. I close the bottle so that water molecules can neither enter nor leave the bottle. I thermally insulate the bottle to stop any energy transfer by heat. I do not squeeze the bottle, so that the volume of the bottle is fixed. There is also no transfer of charge between the system and its surroundings. We state the *principles of conservation*: an isolated system conserves energy, space, matter, and charge.

Thermodynamics will use these conserved quantities as internal variables. Of course, internal variables need not be restricted to conserved quantities. All conserved quantities obey similar mathematics, and are convenient to study in parallel.

Exercise. What are space, matter, and charge? Do you know why each is conserved?

A classification of systems

Depending on the modes of interaction between the systems and their surroundings, we classify systems into several types.

	transfer matter	transfer space	transfer energy
isolated system	no	no	no
thermal system	no	no	yes

closed system	no	yes	yes
open system	yes	yes	yes

Thermal system. A *thermal system* interacts with its surroundings in one mode: transfer energy.

For example, a bottle of water is a thermal system. We cap the bottle to prevent molecules to leak in or out. We make the bottle rigid to fix its volume. We can still change the energy of the water by placing the bottle over a flame, or by shaking the bottle.

Closed system. A *closed system* and its surroundings do not transfer matter, but transfer space and energy.

Consider a cylinder-piston setup that encloses water molecules. Some water molecules form a liquid, and others form a vapor. We can make the water molecules inside the cylinder closed system by sealing the piston, so that no molecules will leak in or out. The cylinder-piston setup interacts with the rest of the world in two ways. First, when weights are added on top of the piston, the piston moves down and reduces the volume inside the cylinder. Second, when the cylinder is brought over a flame, the flame heats up the water.

Open system. An *open system* and its surroundings transfer matter, space, and energy. A bottle of water, once the cap is removed, is an open system.

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?

Exercise. For each type of system listed above, give an example. In each example, describe all modes of interaction between the system and its surroundings.

Transfer energy between a closed system and its surroundings

A half bottle of water. Here is the half bottle of water again. I cap the bottle to prevent molecules from leaking in or out. I shake the bottle, or just touch it. In both cases, my hand transfers *energy* to the bottle. In shaking the bottle, my hand transfers energy to the bottle by *work*, through force times displacement. In touching the bottle, my hand transfers energy to the bottle by *heat*, through molecular vibration.

The half bottle of water is a closed system. Work and heat are two ways to transfer energy between the closed system and its surroundings.

Transfer energy by work. We have learned many ways of doing work in mechanics and electrodynamics. We now recall two examples.

The expansion of a gas. A weight is placed on top of a piston, which seals a cylinder of gas. For brevity, by a weight we mean the force acting on the piston due to both a block of mass and the pressure of the surrounding air:

weight = mg + P_0A.

Here m is mass, g is the acceleration of gravity, P_0 is the pressure of surrounding air, and A is the area of the piston.

Assume that the piston moves with no friction. The balance of the forces acting on the piston relates the force of the weight, F, to the pressure of the gas in the cylinder, P:

weight = PA.

When the piston raises its height by dz, the gas expands its volume by dV = Adz and does work to the weight:

(weight)dz = PdV.

The expansion of the gas raises the weight.

Resistive heating. As a second example, a resistor is placed in a container of water. When a voltage V is applied to the two ends of the resistor, an electric current I goes through the resistor, and the work done by the voltage per unit time is IV. The electrical work heats the water.

Transfer energy by heat. Energy transfers by heat in several ways.

Conduction. Energy can go through a material. At a macroscopic scale, the material remains stationary. At a microscopic scale, energy is carried by the flow of electrons and vibration of atoms.

Convection. Energy can go from one system to another with the flow of a fluid. This way of energy transfer involves the transfer of matter between systems and is present for an open system.

Radiation. Energy can be carried by electromagnetic waves. Because electromagnetic waves can propagate in vacuum, two systems can transfer energy without being in proximity.

Sign convention. Given a closed system, we adopt the following sign convention.

- $Q > 0$, energy transfer by heat to the closed system from the surroundings.
- $W > 0$, energy transfer by work from the closed system to the surroundings.
- $\Delta U > 0$, increase of the internal energy of the closed system.

Some authors adopt other sign conventions. You can adopt any sign convention, so long as you make your sign convention explicit in the beginning of a thought, and do not change the sign convention in the middle of the thought.

The first law of thermodynamics. For a closed system, the first law of thermodynamics states that

$Q = W + \Delta U$.

Internal energy is a property of the closed system. Neither transfer energy by work nor transfer energy by heat is a property of the closed system; they are methods of the transfer of energy between the closed system and its surroundings.

We already know how to measure the internal energy and the energy transfer by work. The first law of thermodynamics defines the energy transfer by heat. We do our best to identify various processes that transfer energy by work. Energy transfer by heat is then the transfer of energy that we do not bother to call work. Thus, the first law of thermodynamics is not really a law; it just defines the energy transfer by heat.

In old literature, heat was sometimes used as a synonym for thermal energy. We will stick to the modern usage. Heat is a method of energy transfer, whereas thermal energy is a form of energy. The two concepts are distinct. One can increase the thermal energy of a system by work or heat.

Entropy and energy

Now that you have seen both entropy and energy, you are ready to critique the following extract from a paper by a founder of thermodynamics (Clausius 1865).

We might call S the transformational content of the body, just as we termed the magnitude U its thermal and ergonal content. But as I hold it to be better to borrow terms for important magnitudes from the ancient languages, so that they may be adopted unchanged in all modern languages, I propose to call the magnitude S the *entropy* of the body, from the Greek word *τροπη*), transformation. I have intentionally formed the word *entropy* so as to be as similar as possible to the word *energy*; for the two magnitudes to be denoted by these words are so nearly allied in their physical meanings, that a certain similarity in designation appears to be desirable.

1. The energy of the universe is constant.
2. The entropy of the universe tends to a maximum.

Exercise. Critique this extract. Take the word “universe” to mean an isolated system. Do you agree that entropy and energy are nearly allied in their physical meanings? How can the entropy of an isolated system increase?

Thermal system

Now we have an all-star cast of actors. Let us watch them play, act by act.

Let act one start: the union of entropy and energy produces a child—temperature. The child is so prodigious that it is much better known than its parents, entropy and energy.

A family of isolated systems of a single independent variation

Characteristic function S(U). A thermal system and its surroundings interact in one mode only: transfer energy. When the thermal energy of the thermal system is fixed at a value U, the thermal system becomes an isolated system. Denote the number of quantum states of this isolated system by Ω(U).

As the energy U of the thermal system varies, the function Ω(U), or its equivalent, S(U) = log Ω(U), characterizes the thermal system as *a family of isolated systems*, capable of one *independent variation*.

We call S(U) the *characteristic function* of the thermal system. Later we will determine this function by experiment—that is, we will count the number of quantum states of each member isolated system experimentally.

Hydrogen atom. A hydrogen atom changes its energy by absorbing photons. When isolated at a particular value of energy, the hydrogen atom has a fixed set of quantum states. Each quantum state in the set is characterized by a distinct electron cloud and spin.

The hydrogen atom is a thermal system. Its characteristic function has been computed in quantum mechanics:

Ω(−13.6eV) = 2,
Ω(−3.39eV) = 8,
Ω(−1.51eV) = 18,
……

The domain of the function Ω(U) is a set of discrete values of energy: −13.6eV, -3.39eV, -1.51eV ,… The range of the function Ω(U) is a set of integers: 2, 8, 18,…. For the hydrogen atom, the values of energy have large gaps.

A half bottle of water. A half bottle of water is a thermal system. We can transfer energy to the water in many ways, by touch, fire, shake, and electric current, etc. We seal the bottle to prevent any transfer of matter. We make the bottle rigid to fix the volume.

For a complex system like a half bottle of water, the values of energy are so closely spaced that we regard the energy of the system as a continuous real variable. The characteristic function S(U) is a continuous function.

Exercise. Describe a few more examples of thermal systems.

General features of the function S(U)

Energy-entropy plane. In a plane, we draw energy and entropy as two perpendicular axes, with energy as the horizontal axis and entropy as the vertical axis.

Drawing two axes perpendicularly is a convention, and has no experimental significance. What does it even mean to say that energy is perpendicular to entropy? The world works well if we draw the two axes with an arbitrary angle, or not draw them at all. But we will follow the convention and draw the two axes perpendicularly.

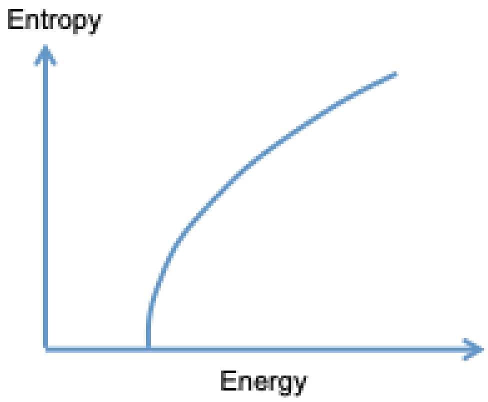
In linear algebra, we call energy U and entropy S two *scalars*. A pair of the values of energy and entropy (U,S) is called a *vector*. The energy-entropy plane is called a *two-dimensional vector space*.

Features of the S(U) curve. We characterize a thermal system with a function S(U), which is a *curve* in the energy-entropy plane.

Of course, different thermal systems have different characteristic functions. Several features are common to all thermal systems. We will list these common features in mathematical terms of the curve S(U) here, and will relate them to experimental observations as we progress.

1. Because energy is relative, the curve S(U) can translate horizontally without affecting the behavior of the thermal system.
2. Because entropy is absolute, the curve starts at S = 0, and cannot be translated up and down.
3. The behavior of a thermal system is often independent of the size of the system. For example, 1 kg of water behaves the same as 2 kg of water. Both energy and entropy is linear in the size of the thermal system. As we change the size of the thermal system, the curve S(U) changes size, but keeps the shape.
4. As entropy approaches zero, the curve S(U) approaches the horizontal axis vertically. That is, as $S \to 0$, $dS(U)/dU \to \infty$.
5. The more energy, the more quantum states. Thus, Ω(U) is an increasing function. Because logarithm is an increasing function, S(U), is also an increasing function. That is, the slope of the curve S(U) is positive, $dS(U)/dU > 0$.

6. The curve S(U) is *convex upward*. That is, the slope dS(U)/dU decreases as U increases, or equivalently, $d^2S(U)/dU^2 < 0$.



Phrases associated with a family of isolated systems

A single isolated system. For a single isolated system, as noted before, several phrases are 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”.

Whenever we speak of equilibrium, we should identify a system isolated for a long time. Associated with each isolated system is a set of quantum states—the sample space. The isolated system flips to its quantum states ceaselessly and rapidly. Out of equilibrium, the isolated system flips to some of its quantum states more often than others. In equilibrium, the isolated system flips to every one of its quantum states with equal probability.

A family of isolated systems. A thermal system is a family of isolated systems capable of one independent variation—energy. Each member in this family is a distinct isolated system, has a fixed value of energy, and flips among the quantum states in its own sample space.

We describe a family of isolated systems using several additional phrases. These phrases are applicable to any family of isolated systems, but here we introduce these phrases using a family of isolated systems that constitute a thermal system.

Thermodynamic state. In a family of isolated systems, a member isolated system in equilibrium is called a *thermodynamic state*, or a *state of equilibrium*, of the family.

We now use the word “state” in two ways. An isolated system has many quantum states, but a single thermodynamic state. A thermodynamic state is synonymous to an isolated system in equilibrium. In a thermodynamic state, the isolated system flips to every one of its quantum states with equal probability.

A thermal system is a family of isolated systems. Each member isolated system corresponds to one thermodynamic state of the thermal system, specified by a value of energy. As energy varies, the thermal system can be in many thermodynamic states. Each thermodynamic state corresponds to a point on the curve S(U).

Thermodynamic process. A sequence of thermodynamic states is called a *reversible thermodynamic process*, or *quasi-equilibrium process*. Each thermodynamic state in this process corresponds to a distinct isolated system in equilibrium. A thermodynamic process is synonymous to a family of isolated systems, each being in equilibrium.

A thermal system is capable of one type of thermodynamic process: changing energy. After each change of energy, we isolate the thermal system long enough to reach equilibrium.

Function of state. We specify a thermodynamic state of the thermal system by a value of energy. Energy is called a *function of state*. The word *state* here means thermodynamic state, not quantum state. A function of state is also called a *thermodynamic property*.

Entropy is also a function of state, so is the number of quantum states. For a thermal system, we will soon introduce four other functions of state: temperature, thermal capacity, Massieu function, and Helmholtz function.

Equation of state. Once the energy U is fixed, a thermal system becomes an isolated system of a fixed sample space, so that the entropy S is also fixed. Consequently, given a thermal system, the energy and entropy are related. The relation S(U) is called an *equation of state*. Again, the word *state* here means thermodynamic state, not quantum state.

In general, an equation of state is an equation that relates thermodynamic properties of a family of isolated systems. For a thermal system, we will soon introduce several equations of state, in addition to S(U).

Dissipation of energy

Empirical facts. When I shake a half bottle of water, energy transfers from my muscle to the water, and increases the temperature of the water.

I can also transfer energy to the water using an electric current. This transfer of energy goes through several steps. An electric outlet transfers energy by work to a metal wire, where the voltage of the electric outlet moves electrons in the metal wire. The resistance of the metal wire heats up the metal wire. Let us say that the metal wire is immersed in the water, and transfers energy by heat to the water.

I can also drop a weight into the bottle of water from some height. The weight comes to rest in the water and heats up the water.

In the three examples, the energy starts in the form of, respectively, chemical energy in the muscle, electrical energy in the outlet, and the potential energy in the weight at some height. After transferring into the water, chemical, electrical, and potential energy converts to thermal energy of the water. The chemical, electrical, and potential energy are said to *dissipate* into thermal energy.

Thermodynamic analysis. To speak of the dissipation of energy, we identify thermal energy and other forms of energy in a single isolated system. The isolated system conserves energy, but dissipates the other forms of energy into thermal energy. So long as the system is isolated, the direction of dissipation is irreversible, from other forms of energy to thermal energy, not the other way around. Thermal energy is called *low-grade energy*. We next analyze the falling weight into water using the basic algorithm.

Construct an isolated system with an internal variable. The bottle of water and the weight together constitute an isolated system. Before the weight drops, the thermal energy of water is U_0, and the potential energy of the weight is PE. According to the principle of the conservation of energy, the energy of the isolated system is fixed. After the weight comes to rest in water, the potential energy of the weight vanishes, and the thermal energy of the water is U_0 + PE. Thus, the thermal energy of the water, U, is an internal variable of the isolated system, increasing from U_0 to U_0 + PE.

Find the subset entropy as a function of the internal variable. The bottle of water is a thermal system, characterized by a function S(U), which is an increasing function. Let the entropy of the weight be S_{weight}, which is taken to be unchanged after falling into the water. The entropy of the isolated system is the sum

$S(U) + S_{weight}$.

Maximize the subset entropy to reach equilibrium. Before the weight drops, the entropy of the isolated system is $S(U_0) + S_{weight}$. After the weight comes to rest in the water, $S(U_0 + PE) + S_{weight}$. The law of the increase of entropy requires that

$S(U_0 + PE) + S_{weight} > S(U_0) + S_{weight}$.

Because S(U) is an increasing function, the above inequality holds if the potential energy of the weight changes to thermal energy in the water. The potential energy of the weight is said to *dissipate* into the thermal energy in the water. The isolated system maximizes the entropy when the potential energy of the weight fully changes to the thermal energy of the water, the weight comes to rest, and the isolated system reaches equilibrium.

Increase the subset entropy to see irreversibility. The reverse change would violate the law of the increase of entropy, and therefore violate the fundamental postulate. The weight, after rest in water, will not draw thermal energy from the water and jump up. Dissipation—the conversion of potential energy to thermal system—is irreversible. What make thermal energy low-grade energy is its high entropy.

The irreversibility is understood from the molecular picture. Thermal energy corresponds to molecular motion, whereas the jumping up of the weight corresponds to all molecules adding velocity in one direction. The former corresponds to more quantum states than the latter. The isolated system changes in the direction that increases the number of quantum states. Consequently, the potential energy dissipates into thermal energy, not the other way around.

The Kelvin-Planck statement of the second law of thermodynamics. This analysis confirms a general empirical observation. It is impossible to produce no effect other than the raising of a weight by drawing thermal energy from a single thermal system. This observation is called the *Kelvin-Planck statement of the second law of thermodynamics*.

Exercise. Use the basic algorithm to analyze heating water by an electric current.

Exercise. Use the basic algorithm to analyze heating by friction.

Isentropic process. Reversibility

A falling apple loses height, but gains velocity. So long as the friction of air is negligible, the sum of the potential energy and kinetic energy of the apple is constant, and the thermal energy of the apple is also constant. We model the apple, together with a part of space around the apple, as an isolated system. The height of the apple is the internal variable of the isolated system. When the apple is at a particular height, z, the isolated system flips among a set of quantum states, and has a certain entropy, S(z). The process of falling keeps the subset entropy S(z) fixed, independent of the internal variable z. Such a process is called an *isentropic process*.

An isentropic process of an isolated system is *reversible*. We can arrange a setup to return the apple to its original height without causing any change to the rest of the world. For example, we can let the apple fall along a circular slide. So long as friction is negligible, the apple will return to the same height.

Of course, friction is inevitable in reality; we have never seen an apple go up and down a slide for a long time. The apple inevitably stops after some time. But a frictionless process can be a useful idealization. For example, the planet Earth has been moving around the Sun for a very long time.

For an isolated system, the two words, *isentropic* and *reversible,* are equivalent. Both adjectives describe an isolated system that keeps the number of quantum states unchanged when an internal variable changes. An isentropic (or reversible) process is also called a *quasi-equilibrium* *process*. To avoid increasing entropy, the process must be slow enough for the isolated system to have a long enough time to reach equilibrium at every point along the process.

Exercise. Describe another isentropic process.

Thermal contact

We have just analyzed the dissipation of energy from a high grade to a low grade. We now look the transfer energy from one thermal system to another thermal system. Two thermal systems are said to be in *thermal contact* if the following conditions hold.

1. The two thermal systems interact in one mode only: transfer energy.
2. The two thermal systems together form an isolated system.

Empirical facts. The principle of the conservation of energy requires that an isolated system should have a fixed amount of energy. It allows arbitrary partition of energy between the two thermal systems, so long as the sum of the energies of the two thermal systems remains constant. However, our everyday experience indicates two facts.

1. When two thermal systems are in thermal contact, energy transfers from one system to the other system, one-way and *irreversible*.
2. After some time, the energy transfer stops, and the two thermal systems are said to reach *thermal equilibrium*.

Thermodynamic analysis. We now trace these empirical facts to the fundamental postulate. We use the basic algorithm to analyze thermal contact.

Construct an isolated system with an internal variable. The two thermal systems, A and B, together constitute an isolated system. We call this isolated system the composite. Let the energies of the two thermal systems be U_A and U_B.The composite of the two thermal systems is an isolated system, and has a fixed amount of energy, denoted by $U_{composite}$. Energy is an extensive quantity, so that

$U_{composite} = U_A + U_B$.

Consequently, the isolated system has a single independent internal variable, say, the energy of one of the thermal systems, U_A.

Find the subset entropy as a function of the internal variable. The two thermal systems are characterized by two functions, $S_A(U_A)$ and $S_B(U_B)$. Once the internal variable U_A is fixed, the composite flips among a subset of its quantum states. Denote the subset entropy by $S_{composite}(U_A)$. Entropy is an extensive quantity, so that

$S_{composite}(U_A) = S_A(U_A) + S_B(U_B)$.

Equilibrium. When one thermal system gains energy dU_A, the other thermal system loses energy by the same amount, so that

$dS_{composite}(U_A) = (dS_A(U_A)/dU_A - dS_B(U_B)/dU_B)dU_A$.

After being isolated for a long time, the composite reaches equilibrium, and the subset entropy maximizes, $dS_{composite}(U_A) = 0$, so that

$dS_A(U_A)/dU_A = dS_B(U_B)/dU_B$.

This equation is the condition of *thermal equilibrium*, and determines the equilibrium partition of energy between the two thermal systems A and B.

Irreversibility. Prior to reaching equilibrium, the subset entropy of the composite increases in time, $dS > 0$, so that

If $dS_A(U_A)/dU_A > dS_B(U_B)/dU_B$, then $dU_A > 0$, and thermal system A gains energy from thermal system B;

If $dS_A(U_A)/dU_A < dS_B(U_B)/dU_B$, then $dU_A < 0$, and thermal system A loses energy to thermal system B.

These inequalities dictate the direction of energy transfer. In thermal contact, the direction of energy transfer is one way and irreversible.

This analysis confirms the two empirical facts of thermal contact.

1. Without loss of generality, assume that, right after thermal contact, $dS_A(U_A)/dU_A > dS_B(U_B)/dU_B$. Our analysis shows that system A grains energy from system B. The direction of the energy transfer is one-way and irreversible.
2. We further assume that $S_A(U_A)$ and $S_B(U_B)$ are convex functions. Consequently, as system A gains energy, the slope $dS_A(U_A)/dU_A$ decreases. As system B loses energy, the slope $dS_B(U_B)/dU_B$ increases. Our analysis shows that the transfer of energy continues until the two slopes are equal, when the two thermal systems reach thermal equilibrium.

Definition of temperature

Define the temperature T by

$1/T = dS(U)/dU$.

Both S and U are extensive thermodynamic properties. The definition ensures that T is an intensive thermodynamic property.

We can use the word *temperature* to paraphrase the above analysis of thermal contact.

1. When two thermal systems are brought into thermal contact, energy transfers only in one direction, from the system of high temperature to the system of low temperature. This observation is called the *Clausius statement of the second law of thermodynamics*.
2. After some time in thermal contact, energy transfer stops, and the two systems have the same temperature. This observation is called the *zeroth law of thermodynamics*.

But wait a minute! Any monotonically decreasing function of dS/dU will also serve as a definition of temperature. What is so special about the choice made above? Nothing. It is just a choice. Indeed, all that matters is the slope dS/dU. What we call the slope makes no difference.

Range of temperature. For a thermal system, the function S(U) is a monotonically increasing function. The more energy, the more quantum states, and the more entropy. Thus, the definition 1/T = dS(U)/dU makes temperature positive. Usually we only measure temperature within some interval. Extremely low temperatures are studied in the science of cryogenics. Extremely high temperatures are realized in stars, and other special conditions.

What can you do for temperature? An essential step to grasp thermodynamics is to get to know the temperature. We define temperature by analyzing an everyday experience—thermal contact. How does temperature rises up as an abstraction from everyday experience of thermal contact? How does temperature comes down from the union of energy and entropy?

Let me paraphrase a better-known Bostonian. And so, my fellow thermodynamicist: ask not what temperature can do for you—ask what you can do for temperature.

Exercise. Write an essay with the title, *What is temperature*? How does temperature rises up as an abstraction from everyday experience of thermal contact? How does temperature comes down from the union of energy and entropy?

Exercise. Hotness and happiness are two common feelings. We measure hotness by an experimental quantity, temperature. Can we do so for happiness? Why?

Two units of temperature

Recall the definition of temperature, 1/T = dS/dU. Because entropy is a dimensionless number, temperature has the unit of energy, Joule.

But a different unit for temperature, Kelvin, is commonly used. The conversion factor between the two units of temperature, Joule and Kelvin, is called the Boltzmann constant k, defined by

1.380649×10^{-23} Joule = 1 Kelvin.

Wiki 2019 redefinition of the SI base units. The old and new definitions of Kelvin start at 28:40 of this video.

Modern Celsius scale. Temperature is defined by the equation 1/T = dS(U)/dU.
Any monotonic function of this temperature defines an *arbitrary scale of temperature*. Here we will just mention a commonly used scale, the Celsius scale C, defined by

C = T − 273.15 K (T in the unit of kelvin).

This modern definition of the Celsius scale differs from the historical definition. Specifically, the melting point and the boiling point of water are no longer used to define the modern Celsius scale. Rather, these two temperatures are determined by experimental measurements. The experimental values are as follows: water melts at 0°C and boils at 99.975°C.

The merit of using Celsius in everyday life is evident. It feels more pleasant to hear that today's temperature is 20 degree Celsius than 293.15 Kelvin, or 404.34×10^{-23}Joule.

Exercise. How was Celsius scale originally defined? What made the international committee to redefine it?

Exercise. How would you defend the choice of the conversion factor, 1.380649×10^{-23} Joule = 1 Kelvin?

Unit of entropy

Dimensionless entropy. The entropy of an isolated system, S, is defined by the number of quantum states of the isolated system, Ω, as

S = log Ω.

The entropy is a dimensionless number.

A unit of entropy. By a historical accident, however, the entropy is given a unit. To preserve the equation 1/T = dS/dU, when we use temperature in the unit of Kelvin, we multiply the dimensionless entropy by k. The entropy is then reported in the unit of JK^{-1}. Write

S = k log Ω.

The Boltzmann constant, $k = 1.380649 \times 10^{-23}$ Joule/Kelvin, does no honor to the three great scientists. It is unfortunate that we associate an unsightly number with the three great names. The number is just the conversion factor between the two units of temperature, Joule and Kelvin.

This unit of entropy may give an impression that the concept of entropy depends on the concepts of energy and temperature. This impression is wrong. As we have seen, for a sample space of a finite number Ω of sample points, entropy is $\log \Omega$. Entropy and energy are two independent concepts. Whereas entropy is a measure of the size of a sample space of an isolated system, energy is just one of numerous quantities that can serve as internal variables. Temperature is the child of the union of entropy and energy, $1/T = dS/dU$.

Ideal gas law

kT is temperature in the unit of energy. The Boltzmann constant k has no fundamental significance; it merely defines a unit of temperature more manageable in everyday use. For any fundamental result, if T is in units of Kelvin, the product kT must appear together. Often people call kT thermal energy. This designation is in general incorrect. There is no need to give any other interpretation: kT is temperature in the unit of energy.

Ideal gas law. The change of unit is illustrated with the ideal gas law. The ideal gas law takes the form

$PV = NT$.

where P is the pressure, V the volume, N the number of molecules, and T the temperature in the unit of energy. Historically, the law of ideal gases was discovered empirically before the discovery of entropy. Later we will show that the two definitions of temperature, $PV = NT$ and $1/T = dS(U)/dU$, are the same.

When T is in the unit of Kelvin, the ideal gas law becomes that

$PV = NkT$.

Avogadro constant N_A. According to the 2019 redefinition of the SI base units, the Avogadro constant N_A is defined by

$6.02214076 \times 10^{23}$ items = 1 mole of items.

If we use mole at a unit for the amount of gas, and use Kelvin as a unit for temperature, the ideal gas law becomes that

$PV = nRT$,

where n is the amount of the gas in the unit of mole and

$R = kN_A = 8.314\ JK^{-1}mole^{-1}$.

This quantity R is called the universal gas constant, and is the product of two historical accidents: the Kelvin unit for temperature, and the Avogadro unit for amount of items. The name “the universal gas constant” sounds pretentious to the modern ear.

Vapor pressure

Partial pressure. The air around us is a mixture of many species of molecules, including nitrogen, oxygen, carbon dioxide, and water. Wiki atmosphere of the Earth.

Let V be the volume of a mixture of gas, and N_{water} be the number of water molecules in the volume. The *partial pressure* of water in the mixture, P_{water}, is defined by

$P_{water}V = P_{water}kT.$

Vapor pressure. The vapor pressure of water is the partial pressure of water at which the vapor equilibrates with a condensed phase of water, liquid water or ice. The vapor pressure depends on temperature. At room temperature, the vapor pressure is approximately 3 kPa. At 100 degrees Celsius, the vapor pressure is approximately 100 kPa. Wiki vapor pressure. Wiki vapor pressure of water.

When the partial pressure of water is below the vapor pressure, the liquid water or ice evaporates. When the partial pressure the vapor pressure, the water molecules in the gas will condense.

Relative humidity. At a given temperature, when a moist air is in equilibrium with the liquid water, we say that the air is saturated with water. If air contains fewer water molecules than the saturated air does, the number of water molecules in the air divided by the number of water molecules in the saturated air is called the *relative humidity*. Write

$RH = N/N_{satu}.$

When the vapor is modeled as an ideal gas, the relative humidity is also given by

$RH = P/P_{satu},$

Exercise. Estimate the number of water molecules in the vapor in the half bottle of water. What is the volume of the vapor divided by the number of molecules? Compare this value to the volume of an individual water molecule.

Specific quantities

Both mass M and volume V are extensive quantities. Their ratios define two intensive quantities: the ratio of mass to volume M/V is called the *mass density*, and the ratio of volume to mass V/M is called the *specific volume*.

We also report volume per molecule V/N, or volume per mole of molecules, V/n. Some textbooks give V/M, V/N, and V/n distinct symbols. We will denote them all by v, and let the context tell the difference.

We will similarly denote specific entropy S/M, entropy per molecule S/N, and entropy per mole of molecules S/n by s. We denote specific energy U/M, energy per molecule U/N, and energy per mole of molecules U/n by u.

The mass of a mole of water is 18 g. At room temperature and atmospheric pressure, the mass density of water is 1000 kg/m^3.

Exercise. How would you defend the choice of the conversion factor, $6.02214076 \times 10^{23}$ items = 1 mole of items?

Exercise. Calculate the mass density, specific volume, volume per molecule, and volume per mole of water molecules at a temperature of 200°C and a pressure of 100 kPa.

Exercise. The gravity of the Earth pulls molecules of gas toward the Earth, but the entropy of the gas disperses the molecules into the space. Assume that the temperature is constant, how does the density of a species of molecules in the gas change with the elevation?

General features of the function T(U)

Recall the definition of temperature, 1/T = dS(U)/dU. We plot the function T(U) on the energy-temperature plane. We use energy as the horizontal axis, and temperature as the vertical axis. Each point on the curve represents a thermal system isolated at a particular value of energy. That is, each point on the curve represents a thermodynamic state of the thermal system. Note several general features of the curve T(U).

1. The energy is defined up to an additive constant, so that the curve T(U) can be translated along the axis of energy by an arbitrarily amount.
2. The temperature T starts at absolute zero, is positive, and has no upper bound.
3. For the time being, we assume that T(U) is an increasing function.

These features of T(U) should be compared with those of S(U).

Exercise. A thermal system can also be characterized by function S(T). Discuss the general features of the curve S(T) on the entropy-temperature plane.

Thermal capacity

When the energy of a thermal system changes by dU, the temperature changes by dT. Define the *thermal capacity* by

$1/C = dT(U)/dU$.

Because energy is an extensive quantity and temperature is an intensive quantity, thermal capacity is an extensive quantity. Thermal capacity is a function of state. When we report temperature in the unit of energy, thermal capacity is dimensionless. When we report temperature in the unit Kelvin, the thermal capacity has the unit JK^{-1}.

The thermal capacity is a function of energy, C(U). We assume that the thermal capacity is positive. This assumption is equivalent to that the function T(U) has positive slope, and that the function S(U) is convex. Recall the definition of temperature, $1/T = dS(U)/dU$.

Thermal capacity is also called heat capacity. Thermal capacity is a function of state, and is independent of the method of energy transfer. For example, we can add energy to a half bottle of water by work, such as shaking the bottle, or electric heating. Once we commit to the modern usage of the word heat as a method of energy transfer, it is inappropriate to name a function of state using the word heat.

Specific thermal capacity. The energy needed to raise the temperature of a unit mass of a substance by a degree is called the *specific thermal capacity* of the substance. Liquid water has approximately a constant specific thermal capacity of 4.18 J/g/K. Ice has approximately a constant specific thermal capacity of 2.06 kJ/kg/K.

Exercise. (a) We immerse a 100 W light bulb in 1 kg of water for 10 minutes. Assume that all electric energy applied to the bulb converts to the internal energy of water. How much does the temperature of the water increase? (b) How much does the temperature of the water increase when an apple falls from a height of 1 m into the water and converts all its potential energy to the internal energy of the water? Similar experiments were conducted in 1840s to establish that heat is a form of energy.

Exercise. What are the changes in energy and entropy of 1 kg of water when raised from the freezing to the boiling temperature?

Exercise. When 1 kg of water at the freezing point is mixed with 2 kg of water at the boiling point, what is the temperature in equilibrium? What is the change of entropy associated with this mixing?

Exercise. For a solid or a liquid, the change of volume is small when temperature increases. We model the solid or liquid as a thermal system, and measure the temperature-energy curve experimentally. For a small range of temperature, the thermal energy is approximately linear in temperature, $U(T) = CT$, where the thermal capacity C is taken to be a constant. Derive the characteristic function $S(U)$.

Experimental thermodynamics

Division of labor

Some things are easy to calculate theoretically, others are easy to measure experimentally. A division of labor improves the economics of getting things done. Much of thermodynamics is only sensible in terms of the division of labor between theory and experiment, and between people and machines. For most isolated systems, counting quantum states experimentally is far more economic than computing them theoretically.

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.

Calorimetry

The art of measuring thermal energy is called *calorimetry*. A device that measures thermal energy is called a *calorimeter*. Calorimetry has become a fine art of high sophistication, and it is too much of a tangent to talk about current practice of calorimetry in a beginning course in thermodynamics. All we need to know is that thermal energy is measured routinely. If you wish to have a specific method in mind, just think of an electric heater. The electric energy is (current)(voltage)(time), assumed to be fully converted into thermal energy.

Thermometry

The art of measuring temperature is called *thermometry*. A device that measures temperature is called a *thermometer*. Temperature affects all properties of all materials. In principle, any property of any material can serve as a thermometer. The choice is a matter of convenience, accuracy, and cost. Here are two commonly used thermometers.

Mercury-in-glass thermometer. A mercury-in-glass thermometer relies on a property of mercury: the volume expands as temperature increases. Thus, a volume indicates a temperature.

Gas thermometer. An ideal gas obeys the equation of state:

$PV = NkT$,

This equation relates temperature T to measurable quantities P, V, and N. Thus, an ideal gas can serve as a thermometer, called the *gas thermometer*.

Exercise. Before you accept an ideal gas as a thermometer, describe a method to measure the number of molecules in a bottle of gas.

A division of labor. How does a doctor determine the temperature of a patient? Certainly she does not count the enormous number of quantum states of her patient. Instead, she uses a thermometer. Let us say that she brings a mercury-in-glass thermometer into thermal contact with the patient. Upon reaching thermal equilibrium with the patient, the mercury expands a certain amount, giving a reading of the temperature of the patient.

The manufacturer of the thermometer must assign a volume of the mercury to a temperature. This he can do by bringing the thermometer into thermal contact with a flask of an ideal gas. He determines the temperature of the gas by measuring its volume, pressure, and number of molecules. Also, by heating or cooling the gas, he varies the temperature and gives the thermometer a series of markings.

Any experimental determination of the thermodynamic temperature follows these basic steps:

1. For a simple system, formulate a theory that relates temperature to a measurable quantity.
2. Use the simple system to calibrate a thermometer by thermal contact.
3. Use the thermometer to measure temperatures of any other system by thermal contact.

Steps 2 and 3 are sufficient to set up an arbitrary scale of temperature. It is Step 1 that maps the arbitrary scale of temperature to the thermodynamic temperature.

Our understanding of temperature now divides the labor of measuring temperature among a doctor (Step 3), a manufacturer (Step 2), and a theorist (Step 1). Only the theorist needs to count the number of quantum states, and only for a very few idealized systems.

Experimental determination of entropy

This is a magic of thermodynamics. We can count experimentally the number of quantum states of an isolated system of any complexity, knowing nothing about the quantum states themselves. We illustrate the method using a thermal system.

Here is a statement of the task. Given a thermal system, measure its characteristic function S(U). Recall the definition of temperature, $dS(U) = T^{-1}dU$. Counting the number of quantum states requires a combination of thermometry to measure T and calorimetry to measure U.

Experimental measurement of the function T(U). We add energy to the thermal system by, say, an electric heater. We measure the change in energy U by $(\text{time})(\text{resistance})(\text{current})^2$. At each increment of energy, we isolate the system, wait until the system reaches equilibrium, and measure temperature T. These incremental measurements determine the function T(U).

Determination of entropy. Recall the definition of temperature:

$$dS(U) = T^{-1}dU.$$

This equation relates the function S(U) to experimentally measurable quantity, U and T. Once the function T(U) is measured experimentally, an integration determines the function S(U).

In this integration, set $S = 0$ as $T \rightarrow 0$. That is, at the ground state, the number of quantum states is low, and is set to be one. This is a statement of *the third law of thermodynamics*. On the energy-entropy plane, the curve S(U) approaches the horizontal axis vertically.

Often, the measurement only extends to a temperature much above absolute zero. Assume that the measurement gives the energy-temperature curve in the interval between T_0 and T. Upon integrating, we obtain $S(T) - S(T_0)$. Such an experiment leaves $S(T_0)$ undetermined.

Recall that the function S(U) characterize the thermal system as a family of isolated systems. Thus, we can count the number of quantum states of each member isolated system in the family.

Debye model (1912). To illustrate the determination of the characteristic function S(U), consider the Debye model. Near absolute zero, the internal energy of a solid takes the form

$$U = aT^4,$$

where a is a constant. Debye obtained this expression from a microscopic model. Here we regard the U(T) as a curve obtained from experimental measurement.

Invert the above equation, and we have $T = (U/a)^{1/4}$. Integrating $dS = T^{-1}dU$, we obtain that

$S(U) = (4/3)a^{1/4}U^{3/4}$.

We have used the condition $S = 0$ at $T = 0$.

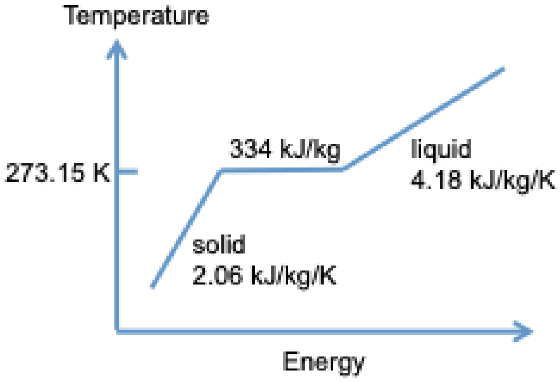
Exercise. Calculate the function C(U) for the Debye model.

Melting

Empirical facts

Ice and liquid are called two *phases* of water. The melting of ice is called a *phase transition*.

A fixed number of water molecules is a thermal system. Its thermodynamic states are capable of one independent variation. We represent the thermodynamic states by a curve, T(U), on the temperature-energy plane, with energy as the horizontal axis, and temperature as the vertical axis.



Solid. In a block of solid water, ice, water molecules form a crystal—a periodic lattice of water molecules. Individual water molecules vibrate around their lattice sites, but change neighbors rarely. The block of ice usually contains many grains of the same crystal. Between two grains, a layer of water molecules do not belong to the lattice of either grain, and is called a grain boundary. The grain boundary is thin, and the thickness is no more than a few molecules. Consequently, the water molecules in the grain boundary contribute negligibly to the extensive quantities (i.e., entropy, energy, and volume) of the block.

As a block of ice receives energy, its temperature increases. Ice has an approximately constant value of specific thermal capacity:

$c_s = 2.06$ kJ/kg/K.

The solid phase corresponds to a line of this slope in the energy-temperature plane. Each point on the line corresponds to one thermodynamic state of the solid phase. In this thermodynamic state, once the grain boundaries are neglected, all parts of the block are similar. The block of ice is said to be in a *homogeneous thermodynamic state*.

Liquid. In a bottle of liquid water, water molecules do not form a lattice. The water molecules touch one another and change neighbors readily. As the liquid receives more energy, the temperature increases again. Water has an approximately constant value of specific thermal capacity:

$c_f = 4.18$ kJ/kg/K.

The subscript f stands for *flüssigkeit,* the German word for liquid. The liquid phase corresponds to a line of this slope in the energy-temperature plane. Each point on the line corresponds to one thermodynamic state of the liquid phase. In this thermodynamic state, all parts of the liquid in the bottle are similar. The bottle of water is said to be in a homogeneous thermodynamic state.

Mixture of solid and liquid. When the temperature increases to a certain level, as the block of ice receives more energy, its temperature remains fixed, but some water molecules become liquid water. The fixed temperature is called the *melting temperature*, T_m. For water, the melting temperature is

$T_m = 273.15$ K.

A mixture of ice and water, when isolated, conserves the amount of energy and the number of water molecules. After being isolated for a long time, the two phases *coexist in equilibrium*: some water molecules form the solid, and the remaining water molecules form the liquid. The solid and the liquid are two thermal systems, in thermal contact and in thermal equilibrium. The coexistent solid and the liquid have the same temperature, T_m. A mixture of solid and liquid in equilibrium is called a *heterogeneous thermodynamic state.*

When a mixture of solid and liquid equilibrate, denote the energy per molecule in the solid by u_s, and the energy per molecule in the liquid by u_f. Of a total number of N molecules in the mixture, N_s molecules form the solid, and N_f molecules form the liquid. The total energy of the mixture is

$U = N_s u_s + N_f u_f$.

This equation is called the *rule of mixture*. We can add energy to the mixture, while keeping the total number of water molecules in the mixture fixed. As the energy of the mixture increases, some water molecules go from the solid to the liquid.

Thus, melting corresponds to a horizontal line in the energy-temperature plane. The level of the line is fixed at the melting temperature, T_m. The line ends on the left when all molecules form the solid, and ends on the right when all molecules form the liquid.

Latent energy. Define the *latent energy* per molecule by the jump in energy, $u_f - u_s$. Often, we report latent energy per unit mass, or specific latent energy. We will use the same symbol u for both the energy per molecule and energy per unit mass. The two quantities are converted by recalling that 6×10^{23} water molecules = 1 mole of water molecules = 18 grams of water molecules. The specific latent energy of water is

$u_f - u_s = 334$ kJ/kg.

The latent energy is commonly known by its historical name, *latent heat* (Black, 1750). This historical name is no longer appropriate. Latent energy is a jump in energy (a function of state), and should not be named using the word heat, a method of energy transfer. We can melt a block of ice by work, such as moving the block against friction. That is, we can form the concept of latent energy and measure it experimentally without ever invoking heat.

Function T(U). In summary, a pure substance near melting is characterized by a curve on the temperature-energy plane using four quantities: the melting temperature T_m, the latent energy $u_f - u_s$, the specific thermal capacity of the solid c_s, and the specific thermal capacity of the liquid c_f. Below the melting temperature and above the temperature, the function T(U) is curved, and both c_s and c_f change with temperature. In many estimate involving water, we will assume approximate constant values, $c_s = 2.06$ kJ/kg/K and $c_f = 4.18$ kJ/kg/K. The approximate function T(U) summarizes a thermodynamic model of water.

Exercise. Calculate the energy needed to bring 1 kg of ice at -50 degree Celsius to liquid water at 50 degree Celsius.

Exercise. 1 kg of ice at the freezing temperature is mixed with 1 kg of water at the boiling temperature. The mixture is insulated. What will be the temperature in equilibrium?

Primitive curves

We now trace the experimental observation of melting back to the fundamental postulate. We model fixed amount of a pure substance as a thermal system. This model assumes that the pure substance can change energy, but ignores that the pure substance can also change volume. The model is a good approximation for solid and liquid, but not for gas.

In phase A, the energy per molecule is u_A, the entropy per molecule is s_A, and the characteristic function is $s_A(u_A)$. Similarly, we model phase B as a thermal system of characteristic function $s_B(u_B)$. The functions $s_A(u_A)$ and $s_B(u_B)$ are two curves in the energy-entropy plane. The two curves are called the *primitive curves* of the pure substance. Each point on a primitive curve corresponds to a *homogeneous thermodynamic state* of the pure substance.

Rule of mixture

A mixture of two homogeneous states. We now consider a mixture of two homogeneous states: state (u_A,s_A) is a point on one primary curve, and state (u_B,s_B) is a point on the other primary curve. The mixture has a total of N molecules, of which N_A molecules are in state A, and N_B molecules are in state B. Denote the number fractions of the molecules in the two homogeneous states by

$y_A = N_A/N, \ y_B = N_B/N.$

Both y_A and y_B are nonnegative numbers. The number of molecules in the mixture is conserved:

$N = N_A + N_B.$

Dividing the above equation by N, we obtain that

$y_A + y_B = 1.$

The mixture is also a thermal system. Let u and s be the energy and entropy of the mixture divided by the total number of molecules. Energy is an extensive variable, so that the energy of the mixture is the sum of the energies of the two homogeneous states:

$Nu = N_Au_A + N_Bu_B.$

Dividing the above equation by N, we obtain that

$u = y_Au_A + y_Bu_B.$

The same is true for entropy:

$s = y_As_A + y_Bs_B.$

Graph the rules of mixture. These rules of mixture can be graphed on the energy-entropy plane. Homogeneous state A is a point (u_A,s_A) on the primitive curve $s_A(u_A)$. Homogeneous state B is a point (u_B,s_B) on the primitive curve $s_B(u_B)$. The mixture is a point (u,s) on the line segment joining the two points (u_A,s_A) and (u_B,s_B), located at the center of gravity, depending on the fraction of molecules y_A and y_B allocated to the two phases. In general, the mixture (u,s) is a point off the primitive curves, and may not be a state of equilibrium.

A mixture of any number of homogeneous states. Now consider a mixture of any number of homogenous states. The homogeneous states can be on one primitive curve, or on both primitive curves. For example, consider a mixture of three homogeneous states, A, B, and C. A pure substance has a total of N molecules, of which N_A, N_B, and N_C molecules are in the three homogeneous states. Denote the number fractions by $y_A = N_A/N$, $y_B = N_B/N$, and $y_C = N_C/N$. Here y_A, y_B and y_C are non-negative numbers. The mixture conserves the number of molecules:

$N_A + N_B + N_C = N$.

Divide this equation by N, and we obtain that

$y_A + y_B + y_C = 1$.

The three homogeneous states, (u_A,s_A), (u_B,s_B), and (u_C,s_C), are three points on one or two primitive curves. The three points form a triangle in the energy-entropy plane. The energy and energy per molecule of the mixture (u,s) are given by

$u = y_Au_A + y_Bu_B + y_Cu_C$,
$s = y_As_A + y_Bs_B + y_Cs_C$.

The mixture, (u,s), is a point in the energy-entropy plane, located at the center of gravity in the triangle, depending on the fraction of molecules y_A, y_B and y_C allocated to the three homogeneous states at the vertices of the triangle. In general, the mixture (u,s) is a point off the primitive curves.

Neglect the spatial arrangement of pieces of homogeneous states. The rules of mixture depend on the number of molecules in each homogenous state, but not on how the pieces of homogenous states are arranged in space.

The rules of mixture also neglect molecules at the interfaces between pieces of homogenous states. The molecules at the interfaces have their own thermodynamic properties, different from those of the homogeneous states. The interfaces contribute to energy and entropy negligibly, so long as the pieces of the homogenous states are much larger than the size of individual molecules.

Convex hull. Each point on the primitive curves corresponds to a homogenous state. A mixture corresponds to a point at the center of gravity of some number of homogeneous states. All possible mixtures of arbitrary numbers of homogeneous states constitute a region in the energy-entropy plane. Each mixture in the region may not be a state of equilibrium.

Incidentally, in the language of convex analysis, the energy-entropy plane is called a *vector space*, and each point in the plane is called a *vector*. All the homogeneous states on the primary curves define a set of vectors. A mixture is called a *convex combination* of the homogenous states, and the set of all mixtures is called the *convex hull* of the homogenous states.

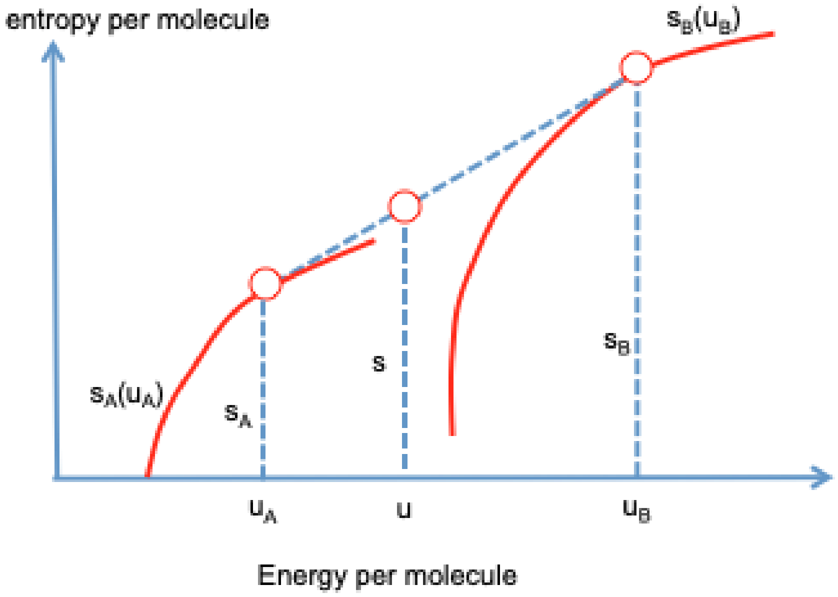
Derived curve

The region of all possible mixtures (i.e., the convex hull) is bounded from above by a single curve, called the *derived curve*. The derived curve is formed by rolling tangent lines on the primitive curves. A tangent line can touch the primitive curves at one point or two points, but not three or more points. The derived curve is convex.

So far energy and entropy play similar roles. All we have invoked is that they are extensive quantities. We next apply the basic algorithm. A fixed amount of mixture of a fixed amount of energy is an isolated system. The isolated system has an enormous number of internal variables: the number of homogeneous states, the location of each homogeneous state on one of the primitive curves, and the number fraction of molecules allocated to each homogeneous state.

The isolated system has a fixed amount of energy u, represented by a vertical line on the energy-entropy plane. The vertical line intersects the derived curve at one point, which maximize the subset entropy.

As the energy of a pure substance varies, the derived curve represents all thermodynamic states of the pure substance.



Equilibrium of a single homogeneous state

If a line tangent to one point on a primitive curve does not cut any primitive curves, this point belongs to the derived curve. The set of all such points is called the *curve of absolute stability*. The tangent line can roll on the primitive curve to change the slope of the tangent line. Thus, the curve of absolute stability has one degree of freedom. Recall $1/T = ds(u)du$. The slope of the tangent line corresponds to temperature.

Equilibrium of two homogeneous states

If a line tangent to two points on the primitive curves does not cut any primitive curves, the straight-line segment connecting the two tangent points belongs to the derived curve. The straight-line segment is called a *tie line*. The common tangent line cannot roll on the primitive curves, is fixed in the energy-entropy plane, and has no degree of freedom. The two tangent points at the ends of the tie line are called the *limits of absolute stability*.

Let the tangent line touch one primitive curve $s_A(u_A)$ at point (u_A,s_A), and touch the other primitive curve $s_B(u_B)$ at point (u_B,s_B). The two tangent points correspond to the two homogeneous states in equilibrium. The slope of the tangent line defines the melting temperature. Thus,

$$1/T_m = (s_B - s_A)/(u_B - u_A) = ds_B(u_B)/du_B = ds_A(u_A)/du_A.$$

Given the two primitive curves, $s_A(u_A)$ and $s_B(u_B)$, the above equations solve for the melting temperature T_m, as well as the two homogeneous states in equilibrium, (u_A,s_A) and (u_B,s_B).

Temperature-entropy curve

We now sketch the curve T(S) near the melting temperature. Recall

$$dS = T^{-1}dU = C(T)T^{-1}dT.$$

Ice melts at temperature 273.15 K. At the melting temperature, the specific thermal capacity for ice is c_s = 2.06 kJ/kg/K. Thus, the slope of the temperature-entropy curve for ice near the melting temperature is

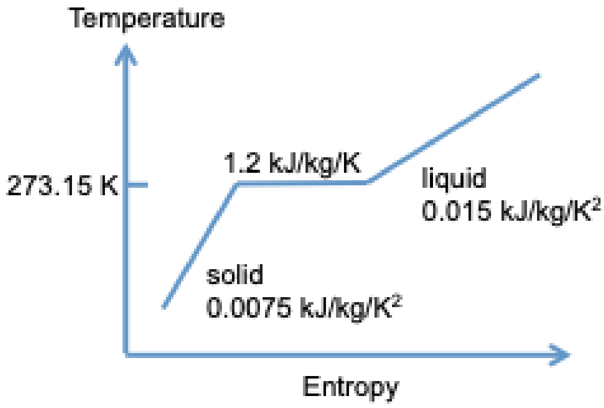
$$c_s/T_m = (2.06 \text{ kJ/kg/K})/(273.15 \text{ K}) = 0.0075 \text{ kJ/kg/K}^2.$$

At the melting temperature, the specific thermal capacity for liquid water is c_f = 4.18 kJ/kg/K. Thus, the slope of the temperature-entropy curve for liquid water near the melting temperature is

$$c_f/T_m = (4.18 \text{ kJ/kg/K})/(273.15 \text{ K}) = 0.015 \text{ kJ/kg/K}^2.$$

Recall that $s_f - s_s = (u_f - u_s)/T_m$. The specific latent energy is 334 kJ/kg. The change in specific entropy is

$s_f - s_s$ = (334 kJ/kg)/(273.15 K) =1.2 kJ/kg/K.



Exercise. Can the model of thermal system describes three homogenous states in equilibrium?

Exercise. For water, we have sketched curves T(u) and T(s) near the melting temperature. Now calculate and sketch the curve s(u) near the melting temperature.

Exercise. For water, the melting temperature is 273.15 K, the specific latent energy is 334 kJ/kg. Calculate the ratio of the number of quantum states in the liquid to that in the solid.

Thermal system of a nonconvex characteristic function s(u).

In hindsight, we should not have accepted so readily that the characteristic function of a thermal system, s(u), is a convex function, or that u(T) is an increasing function, or that thermal capacity is positive. In fact, a thermal system may have a nonconvex primary curve s(u). We can form a tangent line touching two points on the curve s(u). The two points correspond to equilibrium of two homogeneous states.

It turns out that a solid-liquid transition is modeled with two convex primitive curves, but a liquid-gas transition is modeled with a single nonconvex primitive curve. We will see this effect clearly later in a model that allows the pure substance to vary both energy and volume.

Metastability

A primitive curve may contain a convex part and a non-convex part. The point separating the two parts is called the *inflection point* in calculus, and is called the *limit of metastability* in thermodynamics. If a convex part of the primitive curve lies below the derived curve, the part of the primitive curve is beyond the limit of absolute stability. Each point of this part of the primitive curve is called a *metastable state*. A metastable state is stable in regard to continuous changes of state, but is unstable in regard to discontinuous changes of state. Watch a video on supercooled water.

Temperature as an independent variable

A thermal system has a single independent variation. So far we have specified the thermodynamic states of the thermal system using energy as an independent variable. Any one of the functions of state can serve as an independent variable. A popular choice is temperature.

U(T) and T(U)

As described above, we can measure the energy-temperature curve of a thermal system, T(U). For the time being, we assume that the function T(U) is an increasing function. That is, on receiving energy, a thermal system increases temperature. From calculus we have learned that any monotonic function is invertible. Write the inverse function as U(T). The two functions, U(T) and T(U), correspond to the same curve on the energy-temperature plane, and contain the same information.

Thermostat

A thermostat is a device that measures temperature and switches heating or cooling equipment on, so that the temperature is kept at a prescribed level. Thermostats are widely used in refrigerators and home-heating and -cooling units.

Sous-vide (/suːˈviːd/; French for "under vacuum") is a method of cooking. Food (e.g., a piece of meat) is sealed in an airtight plastic bag, and placed in a water bath for a longer time and at a lower temperature than those used for normal cooking. The temperature is fixed by a feedback system. Because of the long time and low temperature, sous-vide cooking heats the food evenly; the inside is properly cooked without overcooking the outside. The airtight bag retains moisture in the food.

Thermal reservoir

A thermal reservoir is a thermal system of a fixed temperature. We use the thermal reservoir to fix the temperature of another thermal system by thermal contact. The other thermal system has a much smaller thermal capacity than the thermal reservoir, so that the temperature of the reservoir, T_R, is fixed as the thermal reservoir and the other thermal system exchange energy. In thermal equilibrium, the other thermal system has the same temperature as the thermal reservoir.

We can realize a thermal reservoir by using a large tank of water. When water loses or gains a small amount of thermal energy, the temperature of water is nearly unchanged.

The situation is analogous to a water reservoir. A water reservoir has a lot of water. Its water level remains nearly unchanged when we take a small cup of water from the reservoir.

The thermal reservoir is a thermal system, and interacts with the rest of the world in one mode: transfer energy. When the energy of the reservoir is fixed at a value U_R, the reservoir becomes an isolated system, and has a certain number of quantum states, $\Omega_R(U_R)$. As U_R varies, the function $\Omega_R(U_R)$, or its equivalent, $S_R(U_R) = \log \Omega_R(U_R)$, characterizes the reservoir as a family of isolated systems.

When the reservoir is in thermal contact with a small thermal system, the composite of the reservoir and the system is an isolated system and has a constant energy, $U_{composite}$. Let U be the energy of the small thermal system. The energy of the composite is a sum of parts:

$U_{composite} = U_R + U$.

Recall the definition of temperature, $1/T_R = dS_R(U_R)/dU_R$. Because T_R is constant, integrating, we find that

$S_R(U_R) = S_R(U_{composite}) - (U_{composite} - U_R)/T_R$.

Thus, a thermal reservoir is a thermal system characterized by a linear function $S_R(U_R)$.

Isothermal process

A process that occurs at a constant temperature is called an *isothermal process*. For example, the sous-vide cooking in an isothermal process. The temperature is fixed while the food cooks. We can fix the temperature by a thermostat, or by a thermal reservoir. So far as the food is concerned, the method of fixing temperature makes no difference.

We model an isothermal process by a system of two variables: energy U and an internal variable x. When both U and x are fixed, the system is an isolated system having a certain number of the quantum states, $\Omega(U,x)$. Let $S(U,x) = \log \Omega(U,x)$.

We fix the temperature of the system by thermal contact with a reservoir. The composite of the system and the reservoir is an isolated system. The entropy of the composite is the sum of the entropies of the reservoir and the system:

$S_{composite} = S_R(U_{composite}) - U/T + S(U,x)$.

The isolated system has two internal variables: the energy U in the system, and the internal variable x. The total energy of the composite $U_{composite}$ is fixed.

The isolated system is in thermal equilibrium, so that $\partial S_{composite}/\partial U = 0$. This condition recovers a familiar condition:

$1/T_R = \partial S(U,x)/\partial U$.

That is, in thermal equilibrium, the reservoir and the thermal system have the same temperature. We will drop the subscript R and write the temperature as T in the following. The function S(U,x) is known, the above condition of thermal equilibrium defines the function U(T,x). We can also write the entropy as a function S(T,x).

Massieu function. The energy of the composite $U_{composite}$ is fixed, so that $S_R(U_{composite})$ is a constant. The composite is an isolated system of a single internal variable, x. The process proceeds to change x to increase $S_{composite}$, or equivalently, increase the function

$J = S(T,x) - U(T,x)/T$.

This function is written as J(T,x). This function contains quantities of the thermal system alone, and was introduced by Massieu (1869). In the isothermal process, the temperature is not a variable, but is fixed by the thermal reservoir. Aside from an additive constant, the Massieu function is the subset entropy of an isolated system: the composite of a thermal system and a thermal reservoir.

Algorithm of thermodynamics for isothermal process

We now paraphrase the basic algorithm of thermodynamics for an isothermal process.

1. Construct a thermal system with an internal variable x.
2. Identify the function J(T,x).
3. *Equilibrium*. Find the value of the internal variable x that maximizes the function J(T,x).
4. *Irreversibility*. Change the value of the internal variable x in a sequence that increases the function J(T,x).

Helmholtz function

Because the temperature is constant and positive, maximizing the J(T,x) is the same as minimizing the following function:

$F = U - TS$.

This function is written as F(T,x). This function contains quantities of the thermal system alone, and is called the *Helmholtz function*, or *free energy*. The function was introduced by Gibbs (1865). Note that U has an arbitrary additive constant, which also appears in F.

Observe that $F = - J/T$. When the system is held at a fixed temperature (i.e., in thermal equilibrium with the reservoir), of all values of the internal variable x, the most probable value minimizes the Helmholtz function F(T,x). In this minimization, temperature is not a variable, but is fixed by the thermal reservoir.

Shadow of the real thing. The Helmholtz function comes from the subset entropy of an isolated system: the combination of a system and a thermal reservoir. The Helmholtz function contains no new fundamental principle, and is a shadow of entropy. In the practice of thermodynamics, the Helmholtz function is so commonly used that many people are enamored by the shadow, and forget the real thing—entropy.

Derivative of Helmholtz function. At fixed x, $dF = dU - TdS - SdT$. Recall the definition of temperature, $dU = TdS$. We have that

$$dF = - SdT.$$

Thus,

$$- S = \partial F(T,x)/\partial T.$$

Because S is positive, F decreases as T increases. Because S increases as T increases, F is a convex function of T.

We can paraphrase the algorithm of thermodynamics for isothermal process in using the Helmholtz function F(T,x). Let us use the algorithm to analyze melting.

Melting analyzed using the Helmholtz function

Let the Helmholtz function per molecule in the solid be $f_A(T)$, and that in the liquid be $f_B(T)$. The Helmholtz function per molecule in a mixture of solid and liquid is

$$f = y_A f_A(T) + y_B f_B(T).$$

The change of phase is modeled as an isothermal process. Recall that $y_A + y_B = 1$.
The fraction of molecules in the solid, y_A, is the independent internal variable, which is varied to minimize the Helmholtz function of the mixture.

The two curves $f_A(T)$ and $f_B(T)$ are decreasing and convex functions. The equation

$$f_A(T) = f_B(T) .$$

determines the melting temperature T_m. When $T < T_m$, the Helmholtz function of the mixture minimizes if all molecules freeze. When $T > T_m$, the Helmholtz function of the mixture minimizes if all molecules melt.

Recall the definition of the Helmholtz function, $f = u - Ts$. The condition of equilibrium of the solid and liquid gives that

$$u_A - T_m s_A = u_B - T_m s_B.$$

This expression recovers what we have obtained before. The Helmholtz function gives us a slick derivation of the equilibrium condition, but pushes us a step away from the leading role, entropy.

Closed system

Now enters a second supporting role—*volume*. In thermodynamics, energy and volume play analogous supporting roles.

A family of isolated systems of two independent variations

A closed system and its surroundings do not transfer matter, but transfer energy and volume. Consider a half cylinder of water sealed with a piston. Above the piston is a weight, and beneath the cylinder is a fire. Inside the cylinder, liquid occupies some volume, and vapor occupies the rest. The water molecules inside the cylinder constitute a closed system.

Characteristic function S(U,V). Let U be the energy and V be the volume of a closed system. When the energy U and volume V are fixed, the closed system becomes an isolated system. For the half cylinder of water, we fix U and V by thermally insulate the cylinder and lock the position of the piston. Denote the number of quantum states of this isolated system by $\Omega(U,V)$. As U and V vary, the function $\Omega(U,V)$, or its equivalent, $S(U,V) = \log \Omega(U,V)$, characterizes the closed system as a family of isolated systems, capable of two independent variations, U and V.

Thermodynamic state. Each member isolated system, in equilibrium, defines a *thermodynamic state of the closed system*, specified by fixed values of the two thermodynamic properties, U and V. In the energy-volume plane, a point represents a thermodynamic state. A thermodynamic state is synonymous to a member isolated system in equilibrium.

Thermodynamic process. A curve in the energy-volume plane represents a thermodynamic process. A thermodynamic process is synonymous to a sequence of member isolated states, each being in equilibrium. A closed system is a family of isolated systems. A thermodynamic process of the closed system is a subfamily of isolated systems. Because the closed system is capable of two independent variations, there are infinitely many types of thermodynamic processes.

Energy-volume-entropy space.

The characteristic function of a closed system, S(U,V), is a *surface* in the energy-volume-entropy space, with energy and volume as the horizontal axes and entropy as the vertical axis. Characterizing a closed system by a surface S(U,V) in the energy-volume-entropy space starts with Gibbs (1873).

For a function of two variables, S(U,V), recall a fact of calculus:

$dS(U,V) = (\partial S(U,V)/\partial U)dU + (\partial S(U,V)/\partial V)dV.$

Further recall the meanings of the two partial derivative in calculus:

$\partial S(U,V)/\partial U = (S(U + dU,V) - S(U,V))/dU,$
$\partial S(U,V)/\partial V = (S(U,V + dV) - S(U,V))/dV.$

Draw a plane tangent to the surface at a point on the surface S(U,V). The tangent plane has the slope $\partial S(U,V)/\partial U$ with respect to the U axis, and the slope $\partial S(U,V)/\partial V$ with respect to the V axis. We next interpret the two partial derivatives in terms of experimental measurements.

Constant-volume process

When the volume V is fixed and the energy U is varied, the closed system becomes a thermal system. For the half cylinder of water, we lock the position of the piston, so that the volume in the cylinder is fixed. We allow the water to receive energy from the fire, so that the energy in the cylinder is varied. In the energy-volume plane, the constant-volume process corresponds to a line parallel to the energy axis.

For a thermal system, we have defined temperature T by

$1/T = \partial S(U,V)/\partial U.$

This equation relates one partial derivative to an experimentally measurable quantity—temperature T.

Adiabatic process

We next look at the other partial derivative, $\partial S(U,V)/\partial V$. We make both the cylinder and the piston using materials that block the transfer of matter and block the transfer of energy by heat. But we can move the piston and transfer energy by work.

When a closed system transfers energy with its surroundings by work but not by heat, the closed system is said to undergo an *adiabatic process*.

When we add weights, the gas loses volume but gains energy. Adding weights compresses the gas and causes *adiabatic heating*. When we remove weights, the gas gains volume but loses energy. Removing weights expands the gas and causes *adiabatic cooling.*

The half cylinder of water and the weight together constitute an isolated system. The isolated system has two internal variables, the volume of the cylinder, V, and the energy in the water molecules, U. The two internal variables are related. The piston is assumed to move with no friction. Let P be the pressure inside the cylinder, and A be the area of the piston. The weight applies a force to the piston, PA. When the piston moves up by a distance dz, the volume inside the cylinder increases by dV = Adz, the potential energy of the weight increases by PAdz = PdV, and the energy of the water increases by dU. The energy of the isolated system is the sum of the thermal energy of the water and the potential energy of the weight. The isolated system conserves energy, dU + PdV = 0. This equation relates the two internal variables U and V.

When the piston moves, the height of weight changes, but the entropy of the weight, S_{weight}, does not change. At fixed U and V, the entropy of the isolated system is the sum of the entropy of the water molecules and the entropy of the weight: S(U,V) + S_{weight}. When U and V vary, the subset entropy maximizes in equilibrium, so that $T^{-1}dU + (\partial S(U,V)/\partial V)dV = 0$. The isolated system undergoes an isentropic process.

This isentropic condition, together with the conservation of energy, dU + PdV = 0, yields that

$P/T = \partial S(U,V)/\partial V$.

This equation relates the partial derivative to experimentally measurable quantities—pressure P and temperature T. The ratio P/T is the child of the union of entropy and volume, just as the temperature is the child of the union of entropy and energy. Pressure is an intensive thermodynamic property.

Exercise. When I stretch a rubber band rapidly, energy has no time to diffuse out, and the rubber band approximately undergoes adiabatic heating. My lips can feel an increase of the temperature of the rubber band. After some time, some energy does diffuse out, and the temperature of the rubber band becomes the same as that of my lips. When I rapidly release the stretch, the rubber band approximately undergoes adiabatic cooling, and my lips feel a decrease in temperature. Try these experiments yourself.

Experimental determination of the function S(U,V)

The above equations give

$dS = (1/T)dU + (P/T)dV$.

This equation relates the characteristic function $S(U,V)$ to experimentally measurable quantities, U, V, P, T. By incrementally changing the energy and volume, we can measure the function $S(U,V)$.

For example, we can prescribe a value of the volume, V_1, regard the closed system as a thermal system, and measure the curve $S(U,V_1)$ using the procedure the same as that for any thermal system. That is, we measure the function $T(U,V_1)$ by incrementally adding energy but fixing volume V_1. We then integrate $1/T = \partial S(U,V_1)/\partial U$, starting with $S(U,V_1) = 0$ as $T(U,V_1) \rightarrow 0$.

We then prescribe another value of the volume, V_2, etc. These measurements determine a family of curves on the energy-entropy plane, $S(U,V_1)$, $S(U,V_2)$,... Each curve characterizes a thermal system of a fixed volume.

Because the energy is a relative quantity, each curve can translate horizontally by an arbitrary amount without affecting the behavior of the thermal system. But two curves of different volumes, $S(U,V_1)$ and $S(U,V_2)$, should not translate relative to each other by an arbitrary amount. Rather, the family of curves constitute a single surface, $S(U,V)$, in the energy-volume-entropy space. The entire surface can translate along the axis of energy by one arbitrary amount without affecting the behavior of the closed system.

We can eliminate the arbitrary translation between the two curves $S(U,V_1)$ and $S(U,V_2)$ by integrating $P/T = \partial S(U,V)/\partial V$. This procedure requires measurements of both pressure and temperature.

The function $S(U,V)$ characterize the closed system as a family of isolated systems. Thus, we can experimentally count the number of quantum states for each member isolated system in the family.

General features of the function S(U,V)

Note the following features of the surface $S(U,V)$ common to all closed systems.

1. Because energy is relative to an arbitrary reference, the surface $S(U,V)$ can translate in the direction of energy by an arbitrary amount without affecting the behavior of the closed system.
2. Because entropy is absolute, the surface starts at $S = 0$, and cannot be translated up and down.
3. Volume is also absolute and positive.
4. When a plane is tangent to the surface $S(U,V)$ at a point, the two slopes of the tangent plane represent $1/T$ and P/T.
5. For each fixed V, as S approaches zero, $S(U,V)$ is a curve that approaches the U axis vertically. That is, $\partial S(U,V)/\partial U$ approaches infinity as S approaches zero.

6. The more energy and volume, the more quantum states. Thus, $\Omega(U,V)$ is an increasing function with respect to both U and V. Because logarithm is an increasing function, $S(U,V)$, is also an increasing function. That is, the slopes $\partial S(U,V)/\partial U$ and $\partial S(U,V)/\partial V$ are positive.
7. For the time being, we assume that $S(U,V)$ is a *convex function*.

Ideal gas

The model of ideal gas

A bottle of volume V contains N molecules and energy U. The molecules are called an *ideal gas* if they are far apart on average. The molecules fly, collide, and separate.

When U,V, N are fixed, the gas is an isolated system. Denote the number of quantum states of this isolated system by $\Omega(U,V,N)$. Because the molecules are far apart, each molecule can fly to anywhere in the bottle as if all other molecules are not there. Consequently, the number of quantum states of each molecule is proportional to the volume V, and the number of quantum states of N molecules is proportional to V^N. The proportional factor depends on U and N, but not on V. Write

$$\Omega(U,V,N) = V^N f(U,N),$$

where $f(U,N)$ is a function of U and N.

By definition, the entropy is $S = k \log \Omega$, so that

$$S(U,V,N) = Nk \log V + k \log f(U,N) .$$

Ideal gas law

Inserting the above expression for entropy into the general equation $P/T = \partial S(U,V,N)/\partial V$, we obtain that

$$PV = NkT.$$

This equation of state is called the ideal gas law. Historically, the ideal gas law was discovered experimentally, before the discovery of entropy. After the discovery of entropy, the ideal gas law is derived from this simple model.

Energy of an ideal gas

Inserting the expression for entropy into the definition of temperature, $1/T = \partial S(U,V,N)/\partial U$, we obtain that $1/T = \partial(\log f(U,N))/\partial U$. This equation shows that the energy U is a function of temperature T and number N, and is independent of the volume V. The energy is an extensive property, and is linear in the number of molecules N. Write

$U = Nu(T)$.

Here u(T) is the thermal energy per molecule. This function has been determined for many species of molecules.

Define the thermal capacity per molecule under the condition of constant volume by

$c_V = du(T)/dT$.

In general, the thermal capacity is a function of temperature, $c_V(T)$. Thus,

$dU = Nc_V(T)dT$.

As an approximation, the energy is taken to be linear in temperature, and c_V is a constant independent of temperature.

Entropy of an ideal gas

Recall the general equation for a closed system

$dS = (1/T)dU + (P/T)dV$.

Inserting the equations specific to an ideal gas, $dU = Nc_V(T)dT$ and $PV = NkT$, we obtain that

$dS = (Nc_V(T)/T)dT + (Nk/V)dV$.

Assuming that c_V is independent of temperature and integrating, we obtain that

$S(T,V) = Nc_V \log(T/T_0) + Nk \log(V/V_0) + S(T_0,V_0)$.

Entropy is a function of state. A thermodynamic state of the closed system is specified by two independent variables, say T and V. In the above equation for ideal gases, the first term is due to the change in temperature, and the second is due to the change in volume.

Energy transfer by work and by heat for a fixed amount of an ideal gas

A fixed amount of an ideal gas is a closed system, characterized by two equations:

$PV = NkT$,
$dU = Nc_V(T)dT$.

The energy transfer by work from the gas to the surroundings is

$dW = PdV$.

The first law of thermodynamics defines the energy transfer by heat from the surroundings to the gas as

$dQ = dU + PdV$.

The temperature, pressure, volume, energy, and entropy are functions of state. Neither the energy transfer by work nor the energy transfer by heat is a function of state; they depend on process. We next consider several processes.

Constant-volume process. Subject to a constant volume, the gas does no work. The energy transfer by heat equals the change in internal energy:

$dQ = Nc_V(T)dT$.

When the temperature changes from T_1 to T_2, assuming the thermal capacity is independent of temperature, the energy transfer by heat is

$Q = Nc_V(T_2 - T_1)$.

Constant-pressure process. When the volume changes by dV, the gas does work $dW = PdV$. When the volume changes from V_1 to V_2 under a constant pressure P, the gas goes work

$W = P(V_2 - V_1)$.

In general, the energy transfer by heat both changes the internal energy and volume: $dQ = dU + PdV$. Subject to a constant pressure, this expression becomes that

$dQ = d(U + PV)$.

For an ideal gas under a constant pressure, the energy transfer by heat is

$dQ = N(c_V + k)dT$.

The quantity $c_V + k$ is the thermal capacity per molecule under constant pressure, and is designated as c_P.

Isothermal process. A constant-temperature process is also called an isothermal process. Because of the ideal gas law, the isothermal process is characterized by a curve PV = constant. For an ideal gas, the isothermal process does not change the internal energy, so that the energy transfer by heat to the closed system is the same as the work done by the closed system:

$dW = dQ = (NkT/V)dV$.

When the volume changes from V_1 to V_2, the energy transfer by heat is

$W = Q = NkT \log (V_2/V_1)$.

Adiabatic process. A closed system is said to undergo an adiabatic process if the system is thermally insulated, so that no energy transfers by heat between the closed system and the surroundings:

$dU + PdV = 0$.

For an ideal gas, this equation becomes

$c_V(T)dT + (kT/V)dV = 0$.

We further assume that the thermal capacity is independent of temperature. Integration yields

$c_V \log T + k \log V = \text{constant}$,

or

$TV^b = \text{constant}$,

where $b = k/c_V$. The constant is determined by one state in the process, say the initial state (T_i,V_i). Thus, constant = $T_iV_i^b$. The result TV^b = constant explains adiabatic cooling when a gas expands in an adiabatic process.

Recall the ideal gas law PV = NkT. The adiabatic process also obeys that

$PV^{b + 1} = \text{constant}$.

When the volume changes by dV, the adiabatic process transfers energy by work:

$dW = PdV = (\text{constant})\ V^{-b-1}dV$.

Integration yields

$W = (\text{constant}/b)\ (V_1^{-b} - V_2^{-b})$

Exercise. Assume that an ideal gas has a constant thermal capacity per molecule c_V. One mole of the ideal gas changes from an initial state (P_i,V_i) to a final state (P_f,V_f) along a straight line on the (P,V) plane. Calculate the energy transfer by work and energy transfer by heat. Calculate the changes in internal energy and entropy.

Exercise. An ideal gas undergoes an adiabatic process from an initial state (P_i,V_i) to a final state of volume V_f. Calculate the volume of the final state, V_f. Calculate the energy transfer by work in the adiabatic process.

Exercise. In a helium gas, each molecule consists of a single helium atom. For such a gas of single-atom molecules, $c_V = 1.5k$. Calculate how temperature changes when pressure doubles in an adiabatic process.

Entropic elasticity

When a spring made of steel is pulled by a force, the spring elongates. When the force is removed, the spring recovers its initial length. This elasticity is due to a distortion of the electron cloud of atoms. Such elasticity is called *energetic elasticity*.

A bag of air acts like a spring. The volume decreases when the pressure increases, and recovers when the pressure drops. This elasticity clearly does not result from distortion of atomic bonds in the molecules, but from the change of the number of quantum states with volume. Such elasticity is called *entropic elasticity*.

Thermoelastic coupling

The bag of air also illustrates thermoelastic coupling. Define *elastic modulus* by

$B = - V\partial P/\partial V$.

This definition is incomplete; we need to specify what is taken to be constant when we take the partial derivative. Let us consider two examples.

Isothermal elastic modulus. Under the isothermal condition, temperature is constant. Recall the ideal gas law, $P = NkT/V$. Taking the partial derivative, we obtain the isothermal elastic modulus

$B_T = P$.

Adiabatic elastic modulus. Under the adiabatic condition, PV^{b+1} = constant. Taking the partial derivative, we obtain the adiabatic elastic modulus

$B_{ad} = (b + 1)P$.

A gas is stiffer under the adiabatic condition than under the isothermal condition.

Osmosis

Consider N particles dispersed in a bag of water of volume V. The particles are different from water molecules, and can be of any size. The particles can be of many species. When the particles are molecules, we call them solutes. When the particles are somewhat larger, say from 10 nm to 10 μ m, we call them colloids. The bag is immersed in a tank of pure water. The bag is made of a semipermeable membrane: water can permeate through the membrane but the particles cannot.

Theory of osmosis

The physics of this situation is analogous to the ideal gas, provided that the concentration of the particles is dilute. Every particle is free to explore the entire volume in the bag. The number of quantum states of the water-particle system scales as

$\Omega \propto V^N$.

As water permeates through the membrane, the volume of the bag V changes.

Recall $S = k \log\Omega$ and the $P/T = \partial S(U,V)/\partial V$, we obtain that

$P = kTN/V$.

This pressure is called the *osmotic pressure*. The law of osmosis is identical to the law of ideal gas.

Balance of osmosis

The osmotic pressure can be balanced in several ways. For example, the tension in the membrane can balance the osmosis pressure. One can also disperse particles in the tank of water outside the bag. The difference in the concentration of particles in the bag and that of particles in the tank causes a difference in the pressures in the bag and in the tank. The difference in pressure can be balanced by the tension in the membrane.

As yet another example, we place a rigid, semi-permeable wall in the liquid, with the particles on one side, but not the other. Water is on both sides of the wall, but alcohol is only on one side. The molecules of the liquid can diffuse across the wall, but the particles cannot. For the particles to explore more volume, the liquid molecules have to diffuse into the side where particles are. If this experiment were carried out in the zero-gravity environment, the infusion would continue until the pure liquid is depleted. In the gravitational field, however, the infusion stops when the pressure in the solution balances the tendency of the infusion.

Phases of a pure substance

Empirical facts

Phases. A pure substance consists of a single species of molecules, and can form three phases: solid, liquid, and gas. (A pure substance may have more than three phases. For example, water has multiple solid phases. For the time being, ignore this fact.) In each phase, the pure substance can change thermodynamic state by changing two independent thermodynamic properties, such as temperature and pressure.

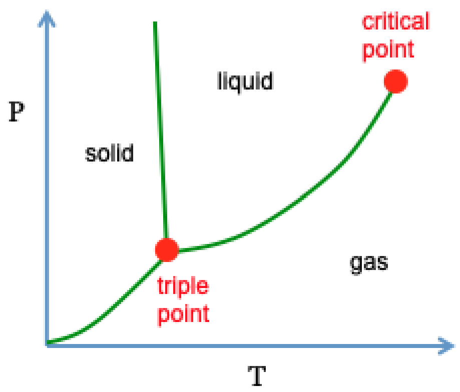
In the solid phase, molecules form a periodic lattice, called a crystal. Many small grains of the crystal form a bulk solid. Individual molecules vibrate near their sites in the lattice, and rarely jump out of the sites.

In the liquid phase, molecules still touch one another, but do not form a periodic lattice. Molecules change neighbors readily.

In the gas phase, molecules on average are far apart. They fly, collide, and separate.

Thermodynamic states of two independent variations. A fixed amount of a pure substance is a closed system. Its thermodynamic states are capable of two independent variations. For example, a fixed amount of a pure substance can be an isolated system of a fixed volume V and a fixed energy U. The isolated system has a certain number of quantum states. Denote the number of quantum states by Ω(U,V), and write S(U,V) = log Ω(U,V). Recall the identity dS(U,V) = (1/T)dU + (P/T)dV.

Steam tables. The data for water molecules are presented by tables, called the steam tables. In constructing the steam tables, the energy and entropy of the liquid at the triple point are typically set to be zero. The extensive quantities, S, U, V, are listed as values per unit amount of the substance.



Temperature-pressure plane. Temperature and pressure are both intensive quantities. Each point on the temperature-pressure plane represents a pair of temperature and pressure. In the temperature-pressure plane, we mark triple point and critical point, along with three phase boundaries. Such a diagram is called a *phase diagram* of a pure substance. See phase diagram of water. The experimental significance of the phase diagram is explained as follows. We will ignore many phases of ice at high pressure.

Equilibrium of a single homogeneous state. When the substance equilibrates in a homogenous state, the thermodynamic state corresponds to a point in the temperature-pressure plane. In the vicinity of a single homogeneous state in equilibrium, the thermodynamic state can change continuously by independent change of temperature and pressure. The steam tables then list volume, energy, and entropy as functions of two independent variables, temperature and pressure.

For example, at T = 400 K and P = 100 kPa, water molecules form a gas, with specific volume, specific energy, and specific entropy:

u = 2967.85 kJ/kg
v = 3.10263 m^3/kg
s = 8.5434 kJ/kg/K

Equilibrium of two homogeneous states. When the substance equilibrates as a mixture of two homogeneous states, they have the same temperature and the same pressure. In the temperature-pressure plane, all thermodynamic states of two homogeneous states fall on a curve, called a *phase boundary*. A point on the phase boundary represents many thermodynamic states; each thermodynamic state is a mixture of the two homogeneous states of some proportion. There are three phase boundaries: gas-liquid, liquid-solid, and solid-gas. For such a phase boundary, given a pressure, the steam tables list the temperature, as well as the specific volumes, energy, and entropy of the two homogeneous states.

For example, at P = 100 kPa, liquid water and gaseous water equilibrate at temperature 99.62 degree Celsius. The specific volumes, energies, and entropies of the two phases are

$u_f = 417.33$ kJ/kg, $u_g = 2506.06$ kJ/kg
$v_f = 0.001043$ m^3/kg, $v_g = 1.69400$ m^3/kg
$s_f = 1.3025$ kJ/kg/K, $s_g = 7.3593$ kJ/kg/K

Equilibrium of three homogenous states.When the substance equilibrates as a mixture of three homogeneous states, they have the same temperature and the same pressure. In the temperature-pressure plane, all thermodynamic states of three homogeneous states collapse to a single point, called the *triple point*. The triple point represents many thermodynamic states; each thermodynamic state is a mixture of the three homogeneous states of some proportions.

For water molecules, the triple point occurs at

T = 273.16 K
P= 611.657 Pa

The specific volumes, energies, and entropies of the three homogeneous states in equilibrium are listed in the following table.

	volume m^3/kg	energy kJ/kg	entropy kJ/kg/K
solid	0.001091	-334	-1.2
liquid	0.001000	0	0
gas	206.132	2375.33	9.1562

Critical point. We add a few more facts of the phase boundaries in the temperature-pressure plane. The three phase boundaries meet at the triple point. The liquid-solid phase boundary extends indefinitely as the pressure increases. The gas-solid phase boundary extends as temperature and pressure reduces, but must terminate when temperature approaches absolute zero. The gas-liquid phase boundary terminates at a point, called the *critical point*. Thus, a liquid state can continuously change to a gaseous state, without crossing the liquid-gas phase boundary. The critical point is a thermodynamic state.

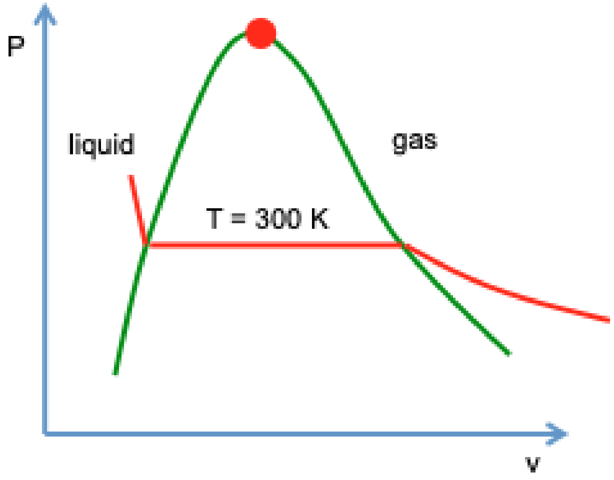
For water molecules at the critical point, the functions of state take the following values:

T = 374.1 degree Celsius
P = 22.089 MPa
u = 2029.58 kJ/kg
v = 0.003155 m^3/kg
s = 4.4297 kJ/kg/K

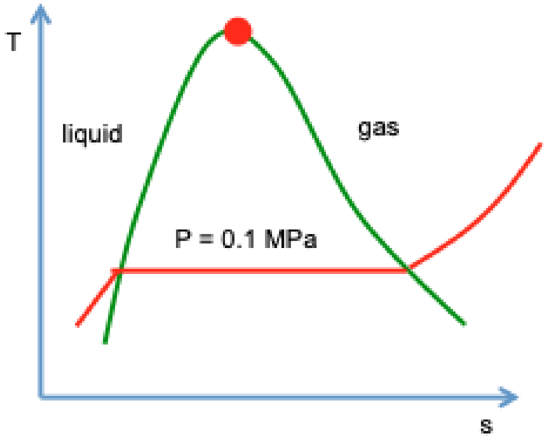
No equilibrium of four or more homogeneous states. A pure substance cannot equilibrate as a mixture of four or more homogeneous states.

Pressure-volume plane. Pressure is an intensive variable, but volume is an extensive variable. We use the specific volume v as the horizontal axis, and the pressure P as the vertical axis. Each point in the P-v plane corresponds to a thermodynamic state of a unit mass of a pure substance, say water. All thermodynamic states of a fixed temperature correspond to a curve in the P-v plane, called an *isotherm*.

Consider an isotherm of T = 300 K. At a high pressure, water is in the liquid phase. The volume of the liquid increases as the pressure drops.



Temperature-entropy plane.



Exercise. For a unit mass of water molecules, a thermodynamic state is specified by two of thermodynamic properties among many, such as temperature T, pressure P, specific energy u, specific volume v, specific entropy s, specific enthalpy h, and quality x. For each of the thermodynamic states specified below, use the steam tables to determine the other thermodynamic properties.

(a) T = 100 degrees Celsius, x = 0.9.
(b) T = 100 degrees Celsius, P = 10 kPa.
(c) T = 100 degrees Celsius, P = 500 kPa.
(d) T = 50 degrees Celsius, v = 0.00050 m^2/kg.

Primitive surfaces

We now describe the thermodynamic theory of phases of a pure substance due to Gibbs (1873). Before Gibbs developed this theory, the empirical facts of pure substances described above were known. They were so numerous and called for a theory.

Energy, volume, and entropy are extensive properties. Let u, v, and s be the energy, volume, and entropy of the substance per molecule. The variables u, v, s form the axes of a three-dimensional space.

Because the thermodynamic state of a fixed amount of a pure substance has two independent variations, once the substance is specified in a state by the values of energy u and volume v, the value of entropy s is fixed. For the time being, let us represent each phase by its own energy-volume-entropy relation, s(u,v), corresponding to a surface in the energy-volume-entropy space. The three phases correspond to three surfaces. Gibbs called them the *primitive surfaces*. A point on one of the three primitive surfaces corresponds to a homogeneous state of the substance.

Rule of mixture

Consider two homogeneous states, A and B, which can be two points either on one primitive surface, or on two primitive surfaces. A mixture of the two homogeneous states has a total of N molecules, of which N_A molecules are in one homogeneous state, and N_B molecules are in the other homogeneous state. Denote the number fractions of the molecules by $y_A = N_A/N$ and $y_B = N_B/N$. The total number of molecules in the mixture is conserved:

$N = N_A + N_B$.

Divide this equation by N, and we obtain that

$y_A + y_B = 1$.

Both y_A and y_B are non-negative numbers.

Denote the energies, volumes, and entropies of the two homogeneous states by (u_A,v_A,s_A) and (u_B,v_B,s_B). Let u, v, and s be the energy, volume, and energy of the mixture divided by the total number of molecules. Energy is an extensive variable, so that the energy of the mixture is the sum of the energies of the two homogeneous states:

$Nu = N_Au_A + N_Bu_B$.

Divide this equation by N, and we obtain that

$u = y_Au_A + y_Bu_B$.

The same rule applies to volume and entropy:

$v = y_Av_A + y_Bv_B$,
$s = y_As_A + y_Bs_B$.

The rules of mixture have a graphic interpretation in the energy-volume-entropy space. The two homogeneous states, (u_A,v_A,s_A) and (u_B,v_B,s_B), are two points on one or two primitive surfaces. The mixture, (u,v,s), is a point on the line joining the two points (u_A,v_A,s_A) and (u_B,v_B,s_B), located at the center of gravity, depending on the fraction of molecules y_A and y_B allocated to the two homogeneous states. In general, the mixture (u,v,s) is a point off the primitive surfaces.

We can generalize the rules of mixture to a mixture of three homogeneous states. A pure substance has a total of N molecules, of which N_A, N_B, and N_C molecules are in the three homogeneous states. Denote the number fractions by $y_A = N_A/N$, $y_B = N_B/N$, and $y_C = N_C/N$. Here y_A, y_B and y_C are non-negative numbers. The mixture conserves the number of molecules:

$N_A + N_B + N_C = N$.

Divide this equation by N, and we obtain that

$y_A + y_B + y_C = 1$.

The three homogeneous states, (u_A,v_A,s_A), (u_B,v_B,s_B), and (u_C,v_C,s_C), are three points on one or two primitive curves. The three points form a triangle in the energy-volume-entropy space. The energy, volume, and entropy per molecule of the mixture (u,v,s) are given by

$u = y_Au_A + y_Bu_B + y_Cu_C$,
$v = y_Av_A + y_Bv_B + y_Cv_C$,
$s = y_As_A + y_Bs_B + y_Cs_C$.

The mixture, (u,v,s), is a point in the energy-volume-entropy space, located at the center of gravity in the triangle, depending on the fraction of molecules $y_{A,}$ y_B and y_C allocated to the three homogeneous states at the vertices of the triangle. In general, the mixture (u,v,s) is a point off the primitive surfaces.

We can further generalize the rules of mixture to a mixture of any number of homogenous states. Now consider all possible mixtures of arbitrary numbers of homogeneous states. Given the primitive surfaces, the rules of mixture create a set of points, which constitute a solid figure in the energy-volume-entropy space. In general, each point in the solid figure represents a mixture.

In the language of convex analysis, the energy-volume-entropy space is called a vector space, and each point in the vector space is called a vector. All the homogeneous states on the primitive surfaces form a set of vectors. A mixture is called a *convex combination* of the homogenous states, and the set of all mixtures is called the *convex hull* of homogenous states.

Derived surface

Gibbs called the upper-bound surface of all mixtures the *derived surface*. This surface is derived by rolling planes tangent to the primitive surfaces. In the energy-volume-entropy space, a tangent plane can touch the primitive surfaces at one, two, or three points, but not four or more points.

When a tangent plane touches the three primitive surfaces at three points, the three tangent points are the vertex of a triangle.

From each edge of the triangle we roll out the tangent plane to touch two primitive surfaces at two points. The two tangent points are the ends of a straight-line segment, called the tie line.

From each vertex of the triangle, we retain a convex part of a primitive surface. The derived surface has a single sheet, and is a *convex surface*.

So far, the three quantities—energy, volume, and entropy—play similar roles. All we have invoked is that they are extensive quantities. We next isolate a fixed amount of a pure substance by fixing energy and volume. This isolated system has an enormous number of internal variables: the number of homogeneous states, the location of each homogeneous state on a primitive surface, and the number fraction of molecules allocated to each homogeneous state.

When a fixed amount of a substance is isolated with a fixed energy and a fixed volume, the substance in general is a mixture, corresponding to a point on a vertical line. Such a mixture in general is *not* in a thermodynamic state. The isolated system reaches equilibrium—that is, reach a thermodynamic state—at the point where the vertical line intersects the derived surface.

Equilibrium of a single homogeneous state

If a plane tangent to one point on a primitive surface does not cut any primitive surfaces, this point of the primitive surface belongs to the derived surface. The poi nt corresponds to a single homogenous state in equilibrium. Gibbs called the set of all such points the *surface of absolute stability*. The tangent plane can roll on the primitive surface to change the two slopes of the tangent plane independently. The slopes of the tangent plane determines the temperature and pressure of the state. Thus, the surface of absolute stability has two degrees of freedom.

Equilibrium of two homogeneous states

If a plane tangent to two points on the primitive surfaces does not cut any primitive surfaces, the straight-line segment connecting the two points belongs to the derived surface. The straight-line segment is called a tie line. The common tangent plane can roll on the two primitive surfaces to change its slope by a single degree of freedom. As the common tangent plane rolls, the tie lines form a developable surface, and the two tangent points trace out two curves on the primitive surfaces. Gibbs called the two curves the *limits of absolute stability*.

Equilibrium of three homogeneous states

If a plane tangent to three points on the primitive surfaces and does not cut any primitive surfaces, the triangle connecting these three points belongs to the derived surface. The tangent plane has no degree of freedom to roll, and is fixed in the energy-volume-entropy space.

Critical point

Gibbs (1873) introduced the theory of critical point. He cited a paper by Andrews (1869), which reported the experimental observation of a substance changing continuously from a liquid to a gas. Gibbs then wrote, “...the derived surface which represents a compound of liquid and vapor is terminated as follows: as the tangent plane rolls upon the primitive surface, the two points of contact approach one another and finally fall together. The rolling of the double tangent plane necessarily come to an end. the point where the two points of contact fall together is the critical point.”

In summary, Gibbs modeled a pure substance with two primitive surfaces: a convex surface for the solid phase, and a nonconvex surface for the liquid and gas phases. You can watch a video on the Gibbs surface.

Metastability

A primitive surface may contain a convex part and a non-convex part. The curve separating the two parts is called the limit of metastability. If a convex part of the primitive surface lies below a tangent plane of the derived surface, the part of the primitive surface is beyond the limit of absolute stability. Each point of this part of the primitive surface is called a metastable state. Gibbs noted that such a state is stable in regard to continuous changes of state, but is unstable in regard to discontinuous changes of state.

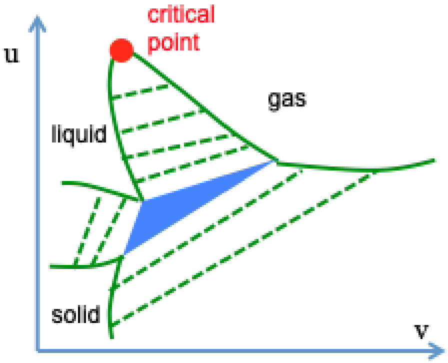
Energy-volume plane

Gibbs projected the derived surface onto the volume-entropy plane. He drew the triangle for the states of coexistence three phases, limits of absolute stability, limits of metastability, and critical point. He did not, however, draw the primitive surfaces and the derived surface in the energy-volume-entropy space.

Maxwell drew the surface in the energy-volume-entropy space in a later edition of his textbook, *Theory of Heat*.

Planck projected the derived surface on the energy-volume plane (Figure 4) in his textbook, *Treatise on Thermodynamics*. Planck made a mistake of adding a critical point for the solid-liquid transition. Such a critical point does not exist.

Here I sketch the projection of the derived surface to the energy-volume plane. Both energy and volume are extensive quantities, so that each thermodynamic state corresponds to a distinct point in the energy-volume plane.



Exercise. Find the data for to draw the water-steam dome on the energy-volume plane. Include the equilibrium ice-water-steam triangle, the critical point, and several tie lines.

Temperature and pressure

In the above, we have developed the entire theory using only three functions of state: energy, volume, and entropy. Each is an extensive quantity, and obeys the rule of mixture. We next discuss the roles of the other two functions of state: temperature and pressure. They are intensive quantities, and obey different mathematical rules.

Write the derived surface as a function s(u,v). Recall the meanings for the slopes of the surface s(u,v):

$1/T = \partial s(u,v)/\partial u$,
$P/T = \partial s(u,v)/\partial v$.

When the substance equilibrates in a mixture of two homogeneous states, a tangent plane contacts the primitive surfaces at two points, (u_A,v_A,s_A) and (u_B,v_B,s_B). The tangent plane has the same slopes at the two points, so that the two homogeneous states have the same temperature and pressure:

$1/T = \partial s_A(u_A,v_A)/\partial u_A = \partial s_B(u_B,v_B)/\partial u_B$,
$P/T = \partial s_A(u_A,v_A)/\partial v_A = \partial s_B(u_B,v_B)/\partial v_B$.

The tangent plane cuts the vertical axis of entropy at some point. The intercept can be calculated using the quantities at either the state (u_A,v_A,s_A) or the state (u_B,v_B,s_B), so that

$s_A - (1/T)u_A - (P/T)v_A = s_B - (1/T)u_B - (P/T)v_B$.

The above equations transcribe the geometrical expressions of the condition of equilibrium into analytical expressions.

Exercise. We have described the conditions of the equilibrium of three homogeneous states in geometric terms. Transcribe these conditions in analytic expressions.

Alternative independent variables

Entropy S(U,V)

We have so far used energy U and volume V as independent variables to specify thermodynamic states of a closed system. With this choice, the thermodynamics of a closed system is characterized by the function S(U,V).

The other two functions of state—temperature T(U,V) and pressure P(U,V)—are determined from the slopes:

$1/T = \partial S(U,V)/\partial U$,
$P/T = \partial S(U,V)/\partial V$.

Thus, once the function S(U,V) is determined by experimental measurement, the function T(U,V) is determined by calculation. The other way around is untrue. Once we measure the function T(U,V), the function T(U,V) does not let us calculate S(U,V). The function S(U,V) lets us calculate all functions defined by it, including temperature, pressure, thermal capacity, latent energy, enthalpy, compressibility, and many more. This is an enormous reduction of experimental work. This fact reveals the long arm of entropy.

We next consider other choices of independent variables. The mathematics of change of variables may bring convenience, but adds no new physics.

Energy U(S,V)

Recall that S(U,V) is an increasing function with respect to U. We can thus invert this function to obtain the function U(S,V). This inversion is purely mathematical, and adds or loses no information. Both functions S(U,V) and U(S,V) characterize the same closed system, and correspond to the same surface in the energy-volume-entropy space.

Rearrange the equation $dS = (1/T)dU + (P/T)dV$ as

$dU = TdS - PdV$.

This equation interprets the partial derivatives of the function U(S,V):

$T = \partial U(S,V)/\partial S, \quad -P = \partial U(S,V)/\partial V$.

One partial derivative reproduces the definition of temperature, and the other partial derivative comes from the mechanics of adiabatic process. This choice of independent variables, S and V, places the temperature and pressure in the symmetric roles of the two slopes of the surface U(S,V). The choice comes from Gibbs (1873) and has been adopted in many textbooks. This choice of independent variables is convenient to discuss an isentropic process, but entropy is rarely used as an independent variable in experiments. We will not consider this choice any further.

Helmholtz function F(T,V)

Many experiments choose temperature and volume as independent variables. As noted before, we can set the temperature of a system as an independent variable by bringing the closed system in thermal contact with a thermal reservoir.

Define the Helmholtz function by

$$F = U - TS.$$

Note that

$$dF = dU - TdS - SdT.$$

This equation, together with the equation $dS = (1/T)dU + (P/T)dV$, gives

$$dF = -SdT - PdV.$$

This equation suggests that the Helmholtz function of the closed system should be a function of temperature and volume, F(T,V), with partial derivatives

$$-S = \partial F(T,V)/\partial T, \quad -P = \partial F(T,V)/\partial V.$$

Thus, once the function F(T,V) is known, the above equations determine the two function of states S(T,V) and P(T,V).

Maxwell relation. For a function of two independent variables, F(T,V), recall an identity in calculus:

$$\partial(\partial F(T,V)/\partial T)/\partial V = \partial(\partial F(T,V)/\partial V)/\partial T.$$

We obtain that

$$\partial S(T,V)/\partial V = \partial P(T,V)/\partial T.$$

This equation is called a Maxwell relation.

Thermal capacity $C_V(T,V)$. When the volume of a closed system is fixed, the closed system becomes a thermal system. Recall the definition of the thermal capacity:

$C_V = \partial U(T,V)/\partial T$.

The subscript V indicates that the volume is fixed.

Function S(T,V). We have just regard the entropy as a function of temperature and volume, S(T,V). Recall a fact in calculus:

$dS = (\partial S(T,V)/\partial T)dT + (\partial S(T,V)/\partial V)dV$

Recall the definition of temperature: $dS = T^{-1}dU$ at constant volume. Thus, $\partial S(T,V)/\partial T = T^{-1}\partial U(T,V)/\partial T = C_V(T,V)/T$. This relation, along with the Maxwell relation $\partial S(T,V)/\partial V = \partial P(T,V)/\partial T$, gives that

$dS = (C_V(T,V)/T)dT + (\partial P(T,V)/\partial T)dV$.

This relation indicates that we can determine the function S(T,V) by measuring the two functions U(T,V) and P(T,V).

Function U(T,V). The above equation, along with the equation $dS = (1/T)dU + (P/T)dV$, gives that

$dU = C_V(T,V)dT + (T\partial P(T,V)/\partial T - P)dV$.

This equation suggests an identity:

$\partial U(T,V)/\partial V = T\partial P(T,V)/\partial T - P$.

Enthalpy H(S,P)

Once again consider a half cylinder of water. A weight is placed on top of the piston, and the water is in thermal contact with a fire.

Constant-pressure process. Let P be the pressure inside the cylinder, and A be the area of the piston. The piston moves without friction. The pressure in the cylinder pushes the piston up. The balance of forces requires that the weight above the piston should be PA. When the weight is constant, the pressure inside the cylinder is also constant.

When the weight is at a height z, the volume of the cylinder is $V = Az$, and the potential energy of the weight is $PAz = PV$. The water and the weight together constitute a thermal system. The energy of the composite is the sum of the internal energy of the water molecules in the cylinder, U, and the potential energy of the weight, PV:

$$H = U + PV.$$

The quantity H is called the *enthalpy* of the water molecules in the cylinder.

Thus, the same experimental setup can be viewed as a closed system or a thermal system. Enthalpy is used to measure energy transfer by heat to a closed system in a constant-pressure process.

Function H(S,P). Now let the pressure P be an independent variable. Recall the product rule in calculus, and we obtain that

$$dH = dU + PdV + VdP.$$

Combining with $dS = (1/T)dU + (P/T)dV$, we obtain that

$$dH = TdS + VdP.$$

We regard the enthalpy as a function of entropy and pressure, $H(S,P)$. The above equation interprets the partial derivatives:

$$T = \partial H(S,P)/\partial S, \quad V = \partial H(S,P)/\partial P.$$

Entropy is rarely used as an independent variable in practice, so the function $H(S,P)$ is seldom useful.

Gibbs function G(T,P)

Gibbs function of a closed system. The half cylinder of water and the weight above the piston together constitute a thermal system of energy $U + PV$. The Helmholtz function of this thermal system is

$$G = U + PV - TS.$$

This quantity is called the Gibbs function. The energy U has an arbitrary additive constant, which also appears in G. The Gibbs function is an extensive quantity, and is a function of state.

Partial derivatives of the Gibbs function. Note that

$dG = dU + PdV + VdP - TdS - SdT$.

Combining with $dS = (1/T)dU + (P/T)dV$, we obtain that

$dG = - SdT + VdP$.

We regard the Gibbs function as a function of pressure and temperature, $G(T,P)$. The above equation interprets the partial derivatives:

$-S = \partial G(T,P)/\partial T$, $V = \partial G(T,P)/\partial P$.

These equations suggest a Maxwell relation:

$-\partial S(T,P)/\partial P = \partial V(T,P)/\partial T$.

Thermal capacity $C_P(T,P)$. When the pressure of a closed system is fixed by a weight, the composite of the closed system and the weight becomes a thermal system. Define the enthalpy $H = U + PV$. Recall the definition of the thermal capacity:

$C_P = \partial H(T,P)/\partial T$.

The subscript P indicates that the pressure is fixed.

Function S(T,P). We have just regard the entropy as a function of temperature and pressure, $S(T,P)$. Recall a fact in calculus:

$dS = (\partial S(T,P)/\partial T)dT + (\partial S(T,P)/\partial P)dP$

Recall the definition of temperature: $dS = T^{-1}dH$ at constant pressure. Thus, $\partial S(T,P)/\partial T = T^{-1}\partial H(T,P)/\partial T = C_P(T,P)/T$. This relation, along with the Maxwell relation $\partial S(T,P)/\partial P = -\partial V(T,P)/\partial T$, gives that

$dS = (C_P(T,P)/T)dT - (\partial V(T,P)/\partial T)dP$.

This relation indicates that we can determine the function $S(T,P)$ by measuring the two functions $H(T,P)$ and $V(T,P)$.

Function H(T,P). The above equation, along with the equations $H = U + PV$ and $dS = (1/T)dU + (P/T)dV$, gives that

$dH = C_P(T,P)dT + (V - T\partial V(T,P)/\partial T)dP$.

This equation suggests an identity:

$\partial H(T,P)/\partial P = V - T\partial V(T,P)/\partial T$.

Constant-pressure and constant-temperature process

Consider a closed system with an internal variable x. For example, the closed system can be a half cylinder of water sealed by a frictionless piston, and the internal variable can be the number of water molecules in the vapor inside the cylinder.

The water molecules are in mechanical equilibrium with the weight above the piston. The weight is constant, and is related to the pressure inside the cylinder as PA. The water molecules are in thermal equilibrium with a thermal reservoir of constant temperature T. We identify the composite of the water molecules, the weight, and the thermal reservoir as an isolated system with three internal variables: the internal energy of the water molecules U, the volume enclosed by the cylinder V, and x.

The energy of the isolated system is

$U_{composite} = U + PV + U_R$.

Here U is the internal energy of the water molecules, PV is the potential energy of the weight, and U_R is the internal energy of the thermal reservoir. The isolated system conserves energy, so that $U_{composite}$ = constant.

The water molecules constitute a closed system, characterized by a function S(U,V,x). The entropy of the weight S_{weight} is constant The thermal reservoir reservoir is a thermal system of constant temperature T, characterized by a function $S_R(U_R)$, so that $S_R(U_R) = S_R(U_{composite}) + (U_R - U_{composite})/T$. The entropy of the composite is

$S_{composite} = S(U,V,x) + S_{weight} + S_R(U_{composite}) - (U +PV)/T$.

The isolated system is in thermal and mechanical equilibrium, so that

$\partial S_{composite}/\partial U = 0$ and $\partial S_{composite}/\partial V = 0$.

When pressure P and temperature T are constant, the two conditions recover the familiar conditions:

$1/T = \partial S(U,V, x)/\partial U$ and $P/T = \partial S(U,V, x)/\partial V$.

Given the function S(U,V, x), these two equations solve the function U(T,P,x) and V(T,P,x). We can also write the entropy as a function S(T,P,x).

Thus, x is the only remaining internal variable. The basic algorithm requires x to change to increase $S_{composite}$, or equivalently, to increase the function

$Y = S - (U + PV)/T$.

In maximizing this function Y(T,P,x), T and P are fixed by the thermal reservoir and the weight, and only x is variable. Aside from additive constants, this the function Y(T,P,x) is the subset entropy of an isolated system: the composite of a closed system, a weight that fixes the pressure P, and a thermal reservoir that fixes the temperature T.

Both the function $J = S - U/T$ and the function $Y = S - (U + PV)/T$ were introduced by Massieu (1869), and the function Y was extensively used later by Planck. We will call J the Massieu function, and Y the Planck function.

Algorithm of thermodynamics for constant-pressure and constant-temperature process

We now paraphrase the basic algorithm of thermodynamics for a constant-pressure and constant-temperature process.

1. Construct a closed system with an internal variable x.
2. Identify the function Y(T,P,x).
3. *Equilibrium*. Find the value of the internal variable x that maximizes the function Y(T,P,x).
4. *Irreversibility*. Change the value of the internal variable x in a sequence that increases the function Y(T,P,x).

Maximizing the Massieu function is equivalent to minimizing

$U + PV - ST$.

This is the Gibbs function G(T,P,x). In this minimization, T and P are constant, and x is variable. The above algorithm can be paraphrased in terms of minimizing G(T,P,x).

Equilibrium of two homogeneous states by equating the Gibbs function

Two homogeneous states in equilibrium have equal values of temperature T and pressure P, but have different values of energy, volume, and entropy per molecule.

The phase boundary between two phases of a pure substance has a single degree of freedom. We can regard T as the independent variable. Along the phase boundary, P, s_A, s_B, u_A, u_B, v_A, and v_B are functions of T.

Let the Gibbs function per molecule in homogeneous state A be

$g_A(T,P) = u_A + Pv_A - Ts_A$.

Let the Gibbs function per molecule in homogeneous state B be

$g_B(T,P) = u_B + Pv_B - Ts_B$.

The Gibbs function per molecule of a mixture of the two homogeneous states is

$g = y_A g_A(T,P) + y_B g_B(T,P)$.

The change of phase is modeled as a process of constant temperature and constant pressure. Recall that $y_A + y_B = 1$. The fraction of molecules in one homogenous state, y_A, is the independent internal variable, which is varied to minimize the Gibbs function of the mixture. The condition of equilibrium is

$g_A(T,P) = g_B(T,P)$

This equation is the same as

$u_A + Pv_A - Ts_A = u_B + Pv_B - Ts_B$.

This condition reproduces what we have obtained by maximizing entropy.

Clapeyron equation

Recall the identities:

$dg_A = -s_A dT + v_A dP$,
$dg_B = -s_B dT + v_B dP$.

Along the phase boundary, the two phases have the equal value of the Gibbs function, so that $dg_A = dg_B$. Thus,

$-s_A dT + v_A dP = -s_B dT + v_B dP$.

Rearranging, we obtain that

$dP/dT = (s_B - s_A)/(v_B - v_A)$.

This result is called the *Clapeyron equation*.

Liquid-solid phase boundary. The right side of the equation is approximately independent of temperature. Thus, the liquid-solid phase boundary is approximately a straight line, with the slope given by the right side of the Clapyron equation. For water molecules, the specific volume is smaller in liquid water than in ice,

$v_l - v_s = - 0.09\ m^3/kg$.

The specific entropy is larger in liquid water than in ice:

$s_l - s_s = 0.09$ kJ/K/kg.

Inserting these values into the Clapyron equation, we obtain the slope for the ice-water phase boundary:

$dP/dT = 13$ MPa/K.

Liquid-gas and solid-gas phase boundaries.

Regelation

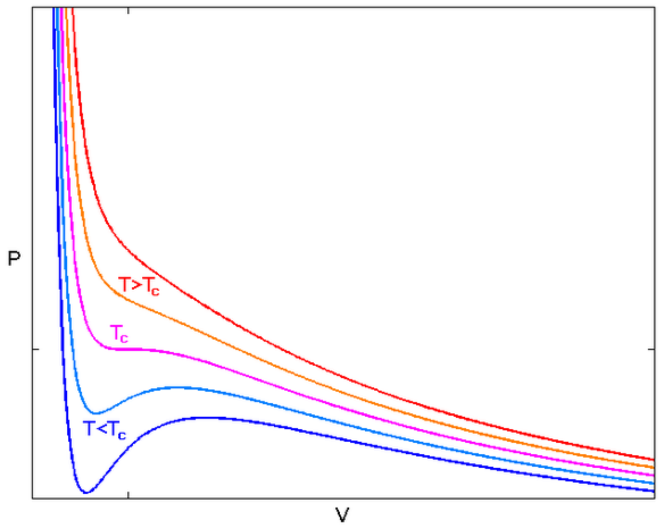
Van der Waals model of liquid-gas phase transition

The ideal gas model represents real gases well at high temperatures and low pressures, when individual molecules are far apart on average. However, at low temperatures and high pressures, when the molecules are near condensation, the ideal gas model greatly deviates from the behavior of real gases.

We now describe an equation of state due to van der Waals (1873):

$(P + aN^2/V^2)(V - Nb) = NkT$,

where a and b are constants for a given substance. When a = 0 and b = 0, the van der Waals equation reduces to the equation of ideal gases. We next examine the physical significance of the two modifications.



(Figure taken from the Wikipedia page on van der Waals equation)

Isotherms

Write the van der Waals equation as

$P = NkT/(V - Nb) - aN^2/V^2$.

For a fixed number of molecules N and a constant temperature T, this equation corresponds to a curve on the pressure-volume plane. The curve is called an *isotherm*. An isotherm at a high temperature is a monotonic cure. An isothermal at a low temperature has a minimum and a maximum.

Critical point

The critical point takes place on the isotherm where the minimum and the maximum collide, so that

$\partial P(T,V)/\partial V = 0, \partial^2 P(T,V)/\partial V^2 = 0.$

Thus, the critical point satisfy three equations:

$P = NkT/(V - Nb) - aN^2/V^2,$
$- NkT/(V - Nb)^2 + 2aN^2/V^3 = 0,$
$2NkT/(V - Nb)^3 - 6aN^2/V^4 = 0.$

Solving these three equations, we obtain that

$V_c = 3Nb$, $kT_c = 8a/(27b)$, $P_c = a/(27b^2)$.

These equations express the critical volume, temperature, and pressure in terms of the constants a and b.

Energy

Recall an identity for a closed system: $\partial U(T,V)/\partial V = T\partial P(T,V)/\partial T - P$. We obtain that

$\partial U(T,V)/\partial V = aN^2/V^2$.

Integrating, we obtain that

$U(T,V) = - aN^2/V + Nu(T)$.

Here u(T) is the same as the thermal energy per molecule for an ideal gas. The term aN^2/V represents the effect of cohesion between the molecules. The van der Waals model assumes that the cohesion reduces the thermal energy by an amount proportional to the number density of the molecules. This seems to be a reasonable first-order approximation.

In the van der Waals model, the thermal capacity is independent of volume:

$C_V = \partial U(T,V)/\partial T = Ndu(T)dT = Nc_V(T)$,

where $c_V(T)$ is the same as the thermal capacity per molecule in an ideal gas.

Entropy

Recall another identity for a closed system: $dS = (C_V(T,V)/T)dT + (\partial P(T,V)/\partial T)dV$. For the van der Waals model, this equation reduces to

$dS = (Nc_V(T)/T)dT + (kN/(V-Nb))dV$.

The term Nb represents the effect of finite volume of the molecules. At a fixed temperature, the number of quantum states is proportional to $(V - Nb)^N$.

In the above, we have started from the van der Waals equation, and examined its consequences for energy and entropy. Since the physical interpretation of the two terms aN^2/V and Nb are quite reasonable, we may as well use them as a starting point to derive the van der Waals equation.

Entropy-energy competition

A fixed number of molecules forms a closed system, which we characterize using the Helmholtz function F(T,V). Recall that

$dF = -SdT - PdV$.

At a fixed temperature, the Helmholtz function is a function of volume, F(V). Plot the function F(V) as a curve in the plane with V as the horizontal axis and F as the vertical axis. When the volume change by dV, the Helmholtz function changes by

$dF = - PdV$.

Thus, -P is the slope of the curve F(V). Because P is positive, F decreases as V increases.

When the pressure decreases as the volume increases, the curve F(V) is convex downward. When the pressure increases as the volume increases, the curve F(V) is convex upward. That is, a nonmonotonic P(V) curve corresponds to a nonconvex F(V) curve

How does the van der Waals model produce a nonconvex Helmholtz function? Recall the definition of the Helmholtz function, F = U - TS. For the van der Waals model, the Helmholtz function takes the form

$F(V) = - aN^2/V - NkT \log(V - Nb)$.

Here we have dropped additive terms which are purely functions of T.

Entropy and energy compete to bend the curve F(V). The second term comes from the entropy of the molecules, and is convex downward, which stabilizes a homogeneous state. The first term comes from the cohesion of the molecules, and is convex upward, which destabilizes a homogeneous state. The entropy tends to disperse molecules, but the energy tends to attract molecules together.

At a high temperature, the entropy prevails for all values of volume, so that the entire curve F(V) is convex downward, and the entire curve P(V) is monotonic.

At a low temperature, the entropy prevails for small and large values of volume, so that only these parts of the curve F(V) is convex downward, and only these parts of the curve P(V) is monotonic.

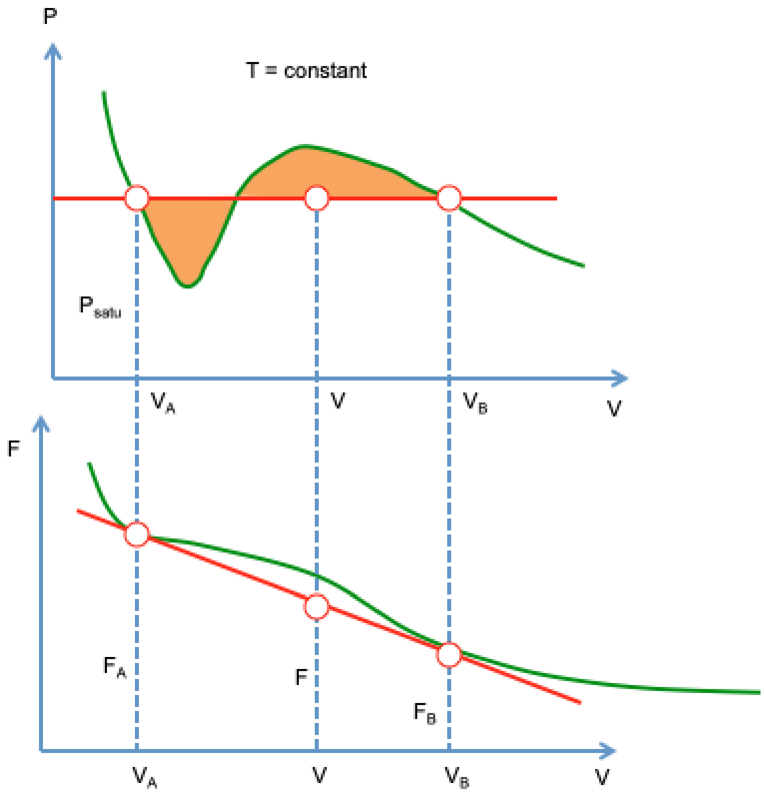
Maxwell rule

A convex-upward part of the F(V) curve corresponds to a phase transition. Draw a line tangent to the F(V) curve at two points A and B. The two points correspond to two homogeneous states in equilibrium at the same pressure, P_{satu}. This pressure corresponds to the slope of the tangent line:

$$F_A - F_B = P_{satu}(V_B - V_A).$$

This equation can be interpreted on the pressure-volume plane: $F_A - F_B$ is the area under the curve P(V) between states A and B, and $P_{satu}(V_B - V_A)$ is the area of a rectangle. The equality of the two areas requires that P_{satu} be placed at the level that equate the two shaded areas. This construction is called the *Maxwell rule*.

Of the total of N molecules, N_A molecules are in homogeneous state A, and N_B molecules are in homogeneous state B. The mixture of the two homogeneous states correspond to a point on the tangent line, located at the center of gravity according to N_A and N_B.



Open system

Now enters another supporting role—the number of a species of molecules. In thermodynamics, the number of a species of molecules, energy, and volume play analogous supporting roles.

A family of isolated systems of many independent variations

An open system and its surroundings transfer energy, volume, and molecules. The molecules can be of many species. To illustrate basic ideas, consider that only two species of molecules, 1 and 2, transfer between the open system and its surroundings. All other species of molecules are blocked.

Let the energy be U, volume be V, the number of molecular species 1 be N_1, and the number of molecular species 2 be N_2. When U, V, N_1, N_2 are fixed, the open system becomes an isolated system. Denote the number of quantum states of this isolated system by $\Omega(U,V, N_1, N_2)$. As U, V, N_1, N_2 vary, the function $\Omega(U,V,N_1,N_2)$, or its equivalent, $S(U,V,N_1,N_2) = \log \Omega(U,V,N_1,N_2)$, characterizes the open system as a family of isolated systems. The family of isolated systems has four independent variations, U, V, N_1, N_2.

Definition of chemical potentials

For the function of four variables, $S(U,V,N_1,N_2)$, recall a fact of calculus:

$dS= (\partial S(U,V,N_1,N_2)/\partial U)dU + (\partial S(U,V,N_1,N_2)/\partial V)dV + (\partial S(U,V,N_1,N_2)/\partial N_1)dN_1 + (\partial S(U,V,N_1,N_2)/\partial N_2)dN_2$.

When we block the transfer of the molecules between the open system and its surroundings, but allow the transfer of energy and volume, the open system becomes a closed system. We have already related two partial derivatives to experimentally measurable quantities:

$1/T = \partial S(U,V,N_1,N_2)/\partial U$, $P/T = \partial S(U,V,N_1,N_2)/\partial V$.

The other two partial derivatives are used to define the *chemical potentials*:

$- \mu_1/T = \partial S(U,V,N_1,N_2)/\partial N_1$,
$- \mu_2/T = \partial S(U,V,N_1,N_2)/\partial N_2$.

The ratio μ_1/T is the child of the union of the entropy and the number of molecular species 1, and the ratio μ_2/T is the child of the union of the entropy and the number of molecular species 2, just as the temperature is the child of the union of entropy and energy, and as the ratio P/T is the child of the union of entropy and volume. The chemical potentials are intensive functions of state.

Write

$dS = (1/T)dU + (P/T)dV - (\mu_1/T)dN_1 - (\mu_2/T)dN_2$.

Flexibility in defining chemical potentials. We have already mentioned the flexibility in defining temperature: any monotonically decreasing function of the derivative $\partial S(U,V,N_1,N_2)/\partial U$ can be used to define temperature. This enormous flexibility comes about because temperature has no definition outside thermodynamics. The choice adopted here, $1/T = \partial S(U,V,N_1,N_2)/\partial U$, is a historical accident.

We do not have any flexibility in defining pressure; we insist that the definition of pressure should recover that in mechanics, force/area. We have shown that $P/T = \partial S(U,V,N_1,N_2)/\partial V$.

In defining chemical potentials, once again we have enormous flexibility, because chemical potentials have no definition outside thermodynamics.

What really matters is that the derivatives $\partial S(U,V,N_1,N_2)/\partial U$, $\partial S(U,V,N_1,N_2)/\partial V$, $\partial S(U,V,N_1,N_2)/\partial N_1$, and $\partial S(U,V,N_1,N_2)/\partial N_2$ play analogous roles. All these partial derivatives are equally significant because thermodynamics is a play of maximization. Even if we choose not to call these derivatives by any name, we will still be doing the same experiment and the same calculation.

The particular definition of chemical potentials adopted here comes from Gibbs (1875), and is just a name given to a partial derivative. We do not need any reason to give a particular name to a child. For this definition of chemical potential, we will find a reason for the presence of the negative sign, but we cannot find any reason for the presence of temperature.

Note that $\partial S(U,V,N_1,N_2)/\partial N_1$ is a pure number. The number means the increase of the number of quantum states associated with adding one molecule of species 1, while keeping the energy, volume, and number of molecules of species 2 fixed. The quantity has clear significance. Gibbs twisted this number into a quantity to have the unit of energy/amount. He was perhaps too enamored with the supporting actor, energy. His reason was twisted, but his definition has stuck.

Usage of words. When we speak of a chemical potential, we should name both the molecule species and the open system. For example, we speak of the chemical potential of water molecules in a piece of cheese, or the chemical potential of water molecules in a glass of wine. We also speak of the chemical potential of carbon dioxide in a bottle of Coca Cola. We denote the chemical potential of molecular species 1 in open system A by $\mu_{1,A}$.

When we speak of temperature, we only need to name the place. For example, we speak of the temperature of a piece of cheese, or the temperature of a glass of wine.

This difference in usage comes from something basic: the world has many species of molecules, but only one species of energy.

Two open systems in contact

Two open systems, A and B, exchange energy, volume, and two molecular species, 1 and 2. The open system A is characterized by a function $S_A(U_A,V_A,N_{1,A},N_{2,A})$, and the system B is characterized by another function $S_B(U_B,V_B,N_{1,B},N_{2,B})$. Note that $N_{1,A}$ denotes the number of molecules of species 1 in system A.

We make the composite of the two open systems into an isolated system. The principles of conservation require that

$U_A + U_B$ = constant,
$V_A + V_B$ = constant,
$N_{1,A} + N_{1,B}$ = constant,
$N_{2,A} + N_{2,B}$ = constant.

Here we assume that the two species of molecules do not undergo a chemical reaction, so that the number of molecules in each species is conserved. The composite is an isolated system of four independent internal variables, U_A, V_A, $N_{1,A}$, $N_{2,A}$.

When the internal variables are fixed at particular values, the isolated system can only flip in a subset of the sample space. Denote the subset entropy by $S_{composite}(U_A,V_A,N_{1,A},N_{2,A})$. Entropy is an extensive quantity, so that

$S_{composite}(U_A,V_A,N_{1,A},N_{2,A}) = S_A(U_A,V_A,N_{1,A},N_{2,A}) + S_B(U_B,V_B,N_{1,B},N_{2,B})$.

When the internal variables change by dU_A, dV_A, $dN_{1,A}$, $dN_{1,B}$, the subset entropy changes by

$dS_{composite} = (1/T_A - 1/T_B)dU_A + (P_A/T_A - P_B/T_B)dV_A + (- \mu_{1,A}/T_A + \mu_{1,B}/T_B)dN_{1,A} + (- \mu_{2,A}/T_A + \mu_{2,B}/T_B)dN_{2,A}$.

Equilibrium. The four internal variables $U_A,V_A,N_{1,A},N_{2,A}$ can change independently. In equilibrium, the subset entropy maximizes, $dS_{composite} = 0$, so that

$T_A = T_B$,
$P_A = P_B$,
$\mu_{1,A} = \mu_{1,B}$,
$\mu_{2,A} = \mu_{2,B}$.

Irreversibility. Out of equilibrium, the subset entropy increases in time, $dS_{composite}(U_A,V_A,N_A) > 0$. Consider a situation where the two open systems are in partial equilibrium, $T_A = T_B$, $P_A = P_B$, $\mu_{1,A} = \mu_{1,B}$, but not in equilibrium with respect to the transfer of molecular species 2. The inequality $dS_{composite}(U_A,V_A,N_A) > 0$ reduces to

$(- \mu_{2,A} + \mu_{2,B})dN_{2,A} > 0$.

Thus, molecular species 2 transfer from the system of high chemical potential to the system of low chemical potential. The presence of the negative sign in the definition of chemical potential leads to this verbal convenience.

Experimental determination of the chemical potential of a species of molecules in a complex system

How do we experimentally measure the chemical potential of a species of molecules in a complex system? When two systems can exchange energy and a species of molecules, the fundamental postulate dictates that the two systems reach equilibrium when they have the same temperature and the same chemical potential of the species of molecules. Consequently, once the chemical potential of a species of molecules in one system is determined, the system can be used to determine the chemical potential of the same species of molecules in other systems.

For example, we can determine the chemical potential of water molecules in a flask containing a pure water vapor as a function of temperature and pressure, $\mu(T,P)$. We now wish to measure the chemical potential of water molecules in a glass of wine. We can bring the wine into contact with a flask of water vapor. The contact is made with a semipermeable membrane that allows water molecules to go through, but blocks all other species of molecules. We then allow the wine to equilibrate with the water vapor in the flask, so that both energy and water molecules stop exchanging between the wine and the flask. The two systems have the same temperature and the same chemical potential. A reading of the chemical potential of water in the flask gives the chemical potential of water in the wine.

Molecular reservoir. We can fix the chemical potential of a species of molecules in a system by letting it transfers the species of molecules to a molecular reservoir. The situation is analogous to fixing temperature. The molecular reservoir has a fixed chemical potential of the species of molecules. For example, a large tank of an aqueous solution of salt is a molecular reservoir of water. A small amount of water molecules can go in and out of the tank, though the vapor. The salt evaporates negligibly. The chemical potential of water in the solution is fixed.

Extensibility

In mathematics, a function $Z(X,Y)$ is called a *homogeneous function* if for any number a the following relation holds:

$aZ(X,Y) = Z(aX,aY)$.

Take derivative with respective to a, and we obtain that

$Z(X,Y) = X\partial Z(X,Y)/\partial X + Y\partial Z(X,Y)/\partial Y$.

This mathematical identity holds for any homogeneous function.

We next apply this mathematical identity to an open system of two species of molecules, characterized by the function $S(U,V,N_1,N_2)$. Note that S, U, V, N_1, N_2 are extensive properties. If we amplify every extensive property by a factor of a, the characteristic function obeys the relation:

$aS(U,V,N_1,N_2) = S(aU,aV,aN_1,aN_2)$.

Thus, the characteristic function $S(U,V,N_1,N_2)$ is a homogeneous function of four independent variables.

Take derivative with respective to a, and we obtain that

$S = U/T + PV/T - N_1\mu_1/T - N_2\mu_2/T$.

The four partial derivatives of the function $S(U,V,N_1,N_2)$ define four intensive properties. When we amplify every extensive property by a factor of a, the open system increases size proportionally, but all the intensive properties remain unchanged.

Gibbs function

Recall the definition of the Gibbs function:

$G = U - TS + PV$.

Of the five quantities on the right-hand side of the above equation, only energy U has an arbitrary additive constant. The same additive constant appears in the Gibbs function.

The definition of the Gibbs function, along with the equation $dS = (1/T)dU + (P/T)dV - (\mu_1/T)dN_1 - (\mu_2/T)dN_2$, gives that

$dG = -SdT + VdP + \mu_1 dN_1 + \mu_2 dN_2$.

This equation suggests that the Gibbs function be regarded as a function $G(T,P,N_1,N_2)$, with the partial derivatives

$-S = \partial G(T,P,N_1,N_2)/\partial T$,
$V = \partial G(T,P,N_1,N_2)/\partial P$,
$\mu_1 = \partial G(T,P,N_1,N_2)/\partial N_1$,
$\mu_2 = \partial G(T,P,N_1,N_2)/\partial N_2$.

Gibbs-Duhem relation

For an open system of four extensible variables, U, V, N_1, N_2, once the size of the system (i,e., the total number of molecules) is fixed, the system has only three independent variations. Thus, the four intensive quantities, T, P, μ_1, μ_2, cannot be independent variables. We next derive a relation among the four intensive quantities.

The definition of the Gibbs function, $G = U - TS + PV$, along with the equation $S = U/T + PV/T - N_1\mu_1/T - N_2\mu_2/T$, gives that

$G = \mu_1 N_1 + \mu_2 N_2$.

Taking derivative, we obtain that

$dG = \mu_1 dN_1 + \mu_2 dN_2 + N_1 d\mu_1 + N_2 d\mu_2$.

This equation, along with the equation $dG = -SdT + VdP + \mu_1 dN_1 + \mu_2 dN_2$, gives that

$-SdT + VdP - N_1 d\mu_1 - N_2 d\mu_2 = 0$.

This equation, called the *Gibbs-Duhem relation*, relates the changes in the four intensive quantities, T, P, μ_1, μ_2.

Chemical potential of a species of molecules in a pure substance

Chemical potential is an intensive property. We have specified a thermodynamic state of a pure substance by two intensive properties, temperature and pressure. Thus, for a pure substance, the three intensive properties, chemical potential, temperature, and pressure, are not independent properties. We now derive a relation between them.

We model a piece of a pure substance as an open system of a characteristic function $S(U,V,N)$. Thus,

$$dS = (1/T)dU + (P/T)dV - (\mu/T)dN$$

This equation defines the temperature T, the pressure P, and the chemical potential μ.

We can increase the number of molecules the piece without changing the functions of state per molecule, u, v, s, and without changing T and P. When we add dN number of molecules to the piece, the extensive functions of state change by $dS = sdN$, $dU = udN$, and $dV = vdN$. The equation $dS = (1/T)dU + (P/T)dV - (\mu/T)dN$ becomes

$$s = (1/T)u + (P/T)v - (\mu/T).$$

Rearranging, we find that

$$\mu = u + Pv - Ts.$$

This equation shows that the chemical potential of a species of molecules in a pure substance equals the Gibbs function per molecule of the pure substance.

We have already learned how to measure the function $s(u,v)$ for a pure substance experimentally. Once $s(u,v)$ is determined, so is the chemical potential. Of the five quantities on the right-hand side of the above equation, only energy u has an arbitrary additive constant. The same additive constant appears in the chemical potential of the molecular species in the pure substance.

Recall that $ds = (1/T)du + (P/T)dv$. This equation, along with $\mu = u + Pv - Ts$, gives that

$$d\mu = -sdT + vdP.$$

This equation suggests that we regard the chemical potential as a function of temperature and pressure, $\mu(T,P)$, with the partial derivatives

$$\partial\mu(T,P)/\partial T = -s, \quad \partial\mu(T,P)/\partial P = v.$$

These results recover the similar equations when we regard the chemical potential of a species of molecules in a pure substance as the Gibbs function per molecule of the substance.

Chemical potential of a species of molecules in an ideal gas

An ideal gas, of N number of a single species of molecules in a flask of volume V, is subject to pressure P and temperature T. Recall the equations of the ideal gas model:

$PV = NkT$.
$dU = Nc_V(T)dT$
$dS = (Nc_V(T)/T)dT + (Nk/V)dV$.

The chemical potential $\mu = u + Pv - Ts$ can be written as

$\mu(T,P) = \mu(T,P_0) + kT \log(P/P_0)$.

Here P_0 is an arbitrary pressure. At a fixed temperature, this expression determines the chemical potential of a species of molecules in an ideal gas up to an additive constant.

Humidity

At a given temperature, when a moist air is in equilibrium with the liquid water, we say that the air is saturated with water. If air contains fewer water molecules than the saturated air does, the number of water molecules in the air divided by the number of water molecules in the saturated air is called the *relative humidity*. Write

$RH = N/N_{satu}$.

When the vapor is modeled as an ideal gas, the relative humidity is also given by

$RH = P/P_{satu}$,

where P is the *partial pressure* of water in the unsaturated gas, $P < P_{satu}$.

Write the chemical potential of water in the air as

$\mu = kT \log(P/P_{satu}) = kT \log(RH)$,

with the understanding that the chemical potential is relative to that of the water molecules in a saturated water at the same temperature.

The lung is always saturated with water vapor at the body temperature (37C), but the atmospheric air may not be. In winter, the cold air outside has low water content even at 100% relative humidity. When the cold air enters a warm room, the relative humidity in the room will reduce below 100% at room temperature. We will feel uncomfortable. Also, water inside the warm room will condense on cold window panes.

Incompressible pure substance

In many applications of liquids, the pressure is small, so that the volume per molecule in the liquid, v, is taken to be independent of the pressure, and the liquid is called *incompressible*.

Recall the chemical potential for a pure substance, $\mu = u + Pv - Ts$. For an incompressible substance, u, v, and s are functions of T, and are independent of P. We obtain that

$\mu(T,P) = u(T) - Ts(T) + Pv(T)$.

The chemical potential of a species of molecules in an incompressible liquid is linear in pressure. Often, we assume that the specific volume is also independent of temperature.

The ascent of sap

How does a tree transport liquid water from the ground to the top? Unlike an animal, the tree does not have a heart to pump liquid. The tree does have a system of tubes, called xylums, to conduct liquid water.

The tensile stress. Let us view a column of water as a free body. Let A be the cross-sectional area of the column. The pressure in the water at the top of the tree, P_{top}, exerts a force AP_{top}. The pressure in the water at the bottom of the tree, P_{bottom}, exerts a force AP_{bottom}. The column of water weighs ρghA, where ρ is the mass density, g is the acceleration of gravity, and h is the height of the tree. The balance of the three forces requires that

$P_{top} = P_{bottom} - \rho gh$.

The bottom of the tree is taken to equilibrate with the water in the soil through the roots of the tree. Assume that the soil is saturated with water, so that

$P_{bottom} = P_{satu}$.

We estimate the other term for a 100 m tall tree.

$\rho gh = (1000\ kg/m^3)(10\ m/s^2)(100\ m) = 1\ MPa$.

The pressure at the top of the tree is enormous and negative:

$P_{top} = -0.9\ MPa$.

A negative pressure means a tensile stress. Thus, the water rises by a tensile force applied at the top of the tree.

What applies the tensile stress? The low humidity! At the top of the tree, the chemical potential of water in the liquid inside the tree is

$\mu = (P_{top} - P_{satu})v$.

The chemical potential of water molecules in the air is

$\mu = kT \log(RH)$.

When the liquid water equilibrates with the gaseous water in the air, the two chemical potentials are equal, so that

$\rho gh = -(kT/v) \log(RH)$.

The volume per water molecule is v = . At temperature T = 300K, kT = 1.38 . thus, kT/v = …

Mixture of ideal gases

Chemical reaction

Stoichiometric coefficients

At atmospheric pressure, above 100 degree Celsius, water molecules form a gas. At this pressure and temperature, oxygen molecules also form a gas, so do hydrogen molecules. When the three species of molecules are enclosed in the same container, molecules fly and collide. After collision, two molecules separate on some occasions, but form new molecules on other occasions. Let us focus on the chemical reaction

$O_2 + 2H_2 = 2H_2O$.

Each oxygen molecule consists of two oxygen atoms, each hydrogen molecule consists of two hydrogen atoms, and each water molecule consists of two hydrogen atoms and one oxygen atom. In this reaction, the oxygen molecules and hydrogen molecules are called the *reactants*, and the water molecules are called the *reaction products*.

A chemical reaction recombine atoms from one set of molecules, the reactants, to another set of molecules, the reaction products. The reaction conserves the number of each species of atoms, but changes the number of each species of molecules.

One oxygen molecule and two hydrogen molecules react to form two water molecules. The coefficients in front of the molecules ensure that every species of atoms is conserved. These coefficients are called the *stoichiometric coefficients*.

Once atomic composition of every molecule is known, the stoichiometric coefficients are determined by conserving every species of atoms.

Exercise. A reaction fundamental to life on Earth is photosynthesis. Plants absorb carbon dioxide and water to produce glucose. Determine the stoichiometric coefficients of the reaction.

Degree of reaction

Consider a reaction of four species of molecules, A, B, C, and D:

$2A + 3B = 5C + 7D$.

The equation indicates the stoichiometric coefficients of the reaction.

Denote the *degree of reaction* by dx, such that the numbers of the four species of molecules A, B, C, and D change by

$dN_A = -2dx$, $dN_B = -3dx$, $dN_C = 5dx$, $dN_D = 7dx$.

Chemical equilibrium

We assume that the reaction takes place in a cylinder sealed with a piston, so that both the temperature T and the pressure P are fixed as the reaction proceeds. Let the Gibbs function be $G(T,P,N_A,N_B,N_C,N_D)$.

When the reaction advances by dx under the condition of constant temperature and constant pressure, the Gibbs function changes by

$dG = (5\mu_C + 7\mu_D - 2\mu_A - 3\mu_B)dx$.

Reaction goes in the direction that reduces the Gibbs function. Thus, the reaction moves toward, $dx > 0$, when

$2\mu_A + 3\mu_B > 5\mu_C + 7\mu_D$.

The reaction moves backward, dx < 0, when

$2\mu_A + 3\mu_B < 5\mu_C + 7\mu_D$.

The reaction reaches equilibrium when

$2\mu_A + 3\mu_B = 5\mu_C + 7\mu_D$.

Reaction of ideal gases

When all species of molecules form ideal gases, each species obeys the ideal gas law:

$P_A = c_A kT$,

where $c_A = N_A/V$ is the concentration of species A. The chemical potential of species in the cylinder is

$\mu(T,P_A) = \mu(T,P_0) + kT \log(P_A/P_0)$.

$(c_C^5 c_D^7)/(c_A^2 c_B^3) = K$.