Review of thermodynamics

Basic algorithm of thermodynamics

Zhigang Suo @zhigangsuo

I describe an approach to thermodynamics, called the basic algorithm of thermodynamics. This approach does not center on the laws of thermodynamics, but on the use of thermodynamics.

Throughout this review, I have added links to tweets, where you and I can discuss thermodynamics point by point.

I will keep updating this review. If you have the google doc app on your phone, pad, and laptop, your copy of the review will be synced with mine.

I have taught the basic algorithm of thermodynamics in the following three courses.

This review is adapted from my notes for an undergraduate course, Thermodynamics.

This review supplements my graduate course Advanced Elasticity, which studies elastic deformation as a reversible thermodynamic process.

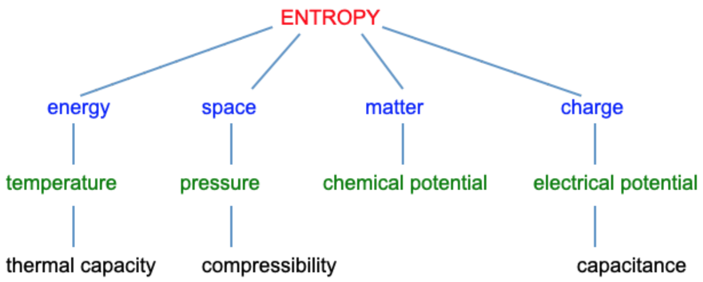
This review also supplements my graduate course Fracture, which studies crack growth as an irreversible thermodynamic process.

Ideal gas

The play of thermodynamics

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.*

In thermodynamics, entropy plays the leading role. Energy plays a supporting role, along with space, matter, and charge. These supporting roles are analogous, of equal importance. Calling thermodynamics the science of energy twists the structure of thermodynamics, and makes thermodynamics hard to understand.



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

Synthesis of thermodynamics

Steam power. Set wood on fire, the fire turns water into steam, and the steam raises a weight. The conversion was known to ancients. Steam, fueled by coal and housed in steel, helped to launch the Industrial Revolution in the eighteenth century.

Electrification. In the late nineteenth century, steam helped to launch electrification. Today, a power plant uses coal, nuclear fuel, petroleum, or natural gas to turn water into steam. The steam turns a turbine, the turbine turns a generator, and the generator generates electricity.

Exercise. Watch this video of a modern demonstration of Hero's engine. Explain how this setup converts fire to motion.

Exercise. Learn about turbines and generators. Explain how they work.

Twitter Carnot efficiency

Discovery of entropy. Carnot (1824) asked, How does a given amount of coal raise the heaviest weight to a height? In answering this question, Carnot conceived the notion of entropy. Clausius (1854) related entropy to energy and temperature. Clausius (1865) coined the word “entropy”. By the time of Gibbs (1870s), entropy had been applied to numerous situations. Entropy is measured by macroscopic experiments. But what is entropy?

Entropy links the microscopic and the macroscopic. Kelvin (1854) coined the word “thermodynamics”. Thermo means being thermal, and dynamics means macroscopic motion. At the time, the nature of “being thermal” was obscure. Clausius, Maxwell, Boltzmann, Gibbs, and others discovered that “being thermal” is about microscopic motion. Thermodynamics links the microscopic and the macroscopic. For an isolated system, entropy is the logarithm of the number of microstates. But what are microstates?

New synthesis. After Planck (1900) discovered quantum states, a new synthesis of thermodynamics started on the basis of a postulate: A system isolated for a long time flips to every one of its quantum states with equal probability. For an isolated system of many particles, this postulate of probability predicts averages of many quantities with (almost) certainty. The postulate predicts the zeroth, second, and third laws of thermodynamics. For an isolated system, entropy is the logarithm of the number of quantum states.

This new synthesis is described in many textbooks on thermal physics. Examples include Reif (1965), Kittel-Kroemer (1980), and Landau-Lifshetz (1980). These textbooks use the fundamental postulate to develop thermodynamics, as well as statistical mechanics.

Many textbooks on thermodynamics ignore the new synthesis. They do delve into diverse applications: engines, power plants, refrigerators, phases, mixtures, reactions, air conditioners, fuel cells, batteries…

Basic algorithm of thermodynamics. This review is devoted to the new synthesis of thermodynamics. I divide the new synthesis into two parts. One part formulates the basic algorithm of thermodynamics, and the other part applies the algorithm to various situations.

The basic algorithm of thermodynamics focuses exclusively on the leading actor—entropy. The algorithm does not mention any particular supporting actors, but provides a point for their entry. The supporting actors—energy, space, matter, charge—enter when the algorithm is applied to various situations. This basic algorithm runs all applications of thermodynamics.

Twitter New synthesis of thermodynamics

Supporting roles

In thermodynamics, energy, space, matter, charge play supporting roles, of equal importance. But they seem to make unequal impressions on students.

Energy. You know *energy* in various forms: thermal energy in a cup of coffee, potential energy of a weight at a height, kinetic energy of a mass at a velocity, electrical energy to power a phone, and chemical energy in food and fuel.

Here is a definition of energy: whatever that can raise a weight to some height. This definition enables the “discovery” of any form of energy. We ascertain if something is a form of energy by a test: raise a weight to some height.

By this definition, energy is conserved, and can convert from one form to another. Here are examples.

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

Thermodynamics is applied to various technologies, but is not centered on them. Do not let these technologies overwhelm you. In watching the play of thermodynamics, do not let the commotion of this pesky supporting actor–energy–divert your attention from the main actor: entropy. Thermodynamics is centered on entropy, not on energy.

Twitter Definition of energy

Work and heat are two methods to transfer energy. In the 19th century, the founders of thermodynamics used work and heat to introduce energy and entropy. Today, as energy and entropy become primary concepts, work and heat are superfluous ideas, and are not invoked in this review.

Twitter Work and heat are superfluous ideas

Space. You know *space* in various aspects: length, area, and volume. Students of thermodynamics are never distracted by this supporting actor, space. But space supports the play of thermodynamics on the same footing as energy.

Matter. You know *matter* in many details. Two hydrogen atoms and one oxygen atom react to form a water molecule, H_2O. In a half bottle of water, for example, H_2O molecules form water and steam. At a low temperature, water freezes into ice. An elastomer is a three-dimensional polymer network. A gel aggregates macromolecules and small molecules. This supporting actor, despite its enormous variation, also does not distract students.

Charge. You know *charge* in terms of electrons, protons, ions, and atoms, as well as current and voltage. The effect of charge will only be treated lightly in this review.

Principles of conservation. Energy, space, matter, and charge obey analogous principles of nature—*the principles of conservation*.

Exercise. Energy is liberated when hydrogen and oxygen react and explode. Does this observation violate the principle of the conservation of energy?

Exercise. Energy is liberated when water freezes. Does this observation violate the principle of the conservation of energy?

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

Exercise. Baby Lucas has discovered a fact: the number of people that enter a room equals the change in the number of people in the room plus the number of people that leave the room. As he grows up, he learns that people include mom, dad, and Lucas, as well as boys, girls, men, and women, but not cats and dogs. Then he learns another word: animal. He also learns about birth and death. Explain these facts to baby Lucas. Will the phrase “the principle of the conservation of the number of people” be useful in your explanation? Why do professors trouble their students with the phrase “the principle of the conservation of energy”? What are facts associated with this phrase?

Supporting roles classify systems

System. Any part of the world is called a *system*. The rest of the world is called the *surroundings*.

I bring to class a demo: a half bottle of water. In the bottle, water fills some volume, and steam fills the rest. We may regard the water as one system, and the steam as another. We may also regard all H_2O molecules in the bottle together as a single 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.

A classification of systems. A system changes its energy, space, matter, and charge by gaining from, or losing to, the surroundings. These changes classify systems into types. For the time being, we neglect the change of charge.

	transer 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

Isolated system. An *isolated system* does not interact with its surroundings.

A half bottle of water can be made into an isolated system. I cap the bottle so that water molecules do not leak. I insulate the bottle to stop any transfer of energy through atom vibration and light radiation. 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.

After all this trouble of insulating the system from transferring energy, space, matter, and charge, I say that the isolated system conserves energy, space, matter, and charge.

Isolated systems are used to formulate the basic algorithm of thermodynamics.

Thermal system. A *thermal system* and its surroundings transfer energy but nothing else.

A half bottle of water can be made into a thermal system. I cap the bottle to prevent molecules from leaking in or out. I make the bottle rigid to fix its volume. But the energy of the water can be changed by placing the bottle over a fire, by shaking the bottle, or by passing an electrical current through a resistor placed near the bottle.

The word *thermal* is an adjective associated with microscopic motion and interaction. Thermal energy is the kinetic and potential energy at microscopic scale. Thermal energy is also called *internal energy*.

In textbooks on thermodynamics, thermal systems are not singled out for study. Yet they function as an excellent pedagogical device. Their practical use is pervasive.

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. Describe the origin of the internal energy of the gas in a bottle.

Exercise. Can you cool a kitchen by opening the door of a refrigerator?

Closed system. A *closed system* and its surroundings transfer space and energy.

A cylinder contains a gas, and is sealed with a piston, so that no molecules leak. The gas in the cylinder 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 of the gas. Second, when the cylinder is brought over a fire, the fire transfers energy to the gas. The gas in the cylinder is a thermal system.

In addition to modeling a gas in a cylinder-piston setup, closed systems model a material carrying a load, a capacitor storing electricity, etc.

Exercise. Is an office a closed system?

Exercise. What is the difference between a closed system and an isolated system, and between a closed system and a thermal system?

Exercise. Leakproof *and* frictionless piston is achieved by a rolling diaphragm. Learn online and describe how it works.

Open system. An *open system* and its surroundings transfer matter, space, and energy.

A cup of coffee is an open system. I am its surroundings. I feel its warmth, smell it, and drink it.

Open systems model mixtures and reactions.

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.

Reviews and tweets of books on thermodynamics and statistical mechanics

Isolated system

Thermodynamics is divided into two parts. One part formulates the basic algorithm of thermodynamics, and the other part applies the algorithm to various situations.

This section completes the first part. To formulate the algorithm, we only need systems of one type—isolated systems.

Fundamental postulate

Of nature 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.

Fact 3 of nature is called the *fundamental postulate of the statistics of an isolated system*. This fact 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 a fact of nature, and use this fact to formulate the BAT.

A system isolated for a long time behaves like a fair die rolled many times:

1. The die has six faces.
2. The die is rolled from one face to another.
3. The die rolled many times lands on every one of its faces with equal probability, 1/6.

The facts of nature link two great sciences: probability and thermodynamics. Isolating a system for a long time is like rolling a fair die many times. Nature acts like a maniac gambler, ceaselessly and rapidly rolling numerous fair dies, each die having numerous faces.

	Probability	Thermodynamics
Experiment	Roll a fair die many times	Isolate a system for a long time
Sample point	A face	A quantum state
Sample space	6 faces	Ω quantum states
Probability of a sample point	1/6	1/Ω
Subset of sample space	Event	State of constrained equilibrium

	Probability	Thermodynamics
Probability of realizing a subset	(number of faces in the subset)/6	(number of quantum states in the subset)/Ω
Map from sample space to another set	Random variable	Internal variable

Exercise. A cheater makes an unfair die of six faces, labeled a,b,c,d,e,f. By 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 rolls?

Tweet Probability vs. thermodynamics

Sample space of rolling a die. In probability, several words are associated with an *experiment*. Each *trial* of the experiment results in an *outcome*, called a *sample point*. All *possible outcomes* of the experiment form a set, called the *sample space*.

Rolling a die is an experiment. The die has six faces, labeled a,b,c,d,e,f. Each roll of the die is a trial of the experiment. Each face is an outcome of the experiment, or a sample point. Write the sample space of the experiment—rolling a die—by listing all the possible outcomes, {a,b,c,d,e,f}.

A sample point is typically *not* a number. A roll of a die results in a face, which need not be marked with a number. The number of dots on a face of a common die is just one way to distinguish the faces. The faces can be just as well marked by any symbols without changing the nature of the experiment.

Throwing a coin is an experiment of two possible outcomes: head and tail. Write the sample space by listing all the possible outcomes, {head, tail}. The head and tail of a coin are certainly not numbers.

It does not matter if the die is made of wood, plastic, or metal. A course on material science is not a prerequisite for studying probability. So far as probability is concerned, the die has six faces, and every face is equally probable for a fair die.

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

Sample space of an isolated system. Isolated at the second energy level, a hydrogen atom has eight quantum states. Wiki atomic orbital.

A quantum state of an isolated half bottle of water is a cloud of many electrons and positions of many nuclei. Such an isolated system has a large, but finite, number of particles. The isolated system also has a large, but finite, number of quantum states.

The isolated system flips from one quantum state to another quantum state, rapidly and ceaselessly.

Only when a system is isolated can we speak of its quantum states.

An isolated system is an "experiment" in the sense of the word used in probability. The isolated system flips from one quantum state to another, rapidly and ceaselessly. Each flip is a trial of the experiment. Each quantum state is a possible outcome of the experiment, a sample point. All the quantum states of the isolated system constitute a set, the sample space. The size of the sample space is measured by the number of quantum states of the isolated system, Ω.

A die has six faces, and a coin has two sides. The two numbers are known. However, counting the number of quantum states of an isolated system, Ω, is a basic problem in thermodynamics. Later we will describe experimental methods to do so.

In formulating thermodynamics, details of quantum states of an isolated system do not matter. A course on quantum mechanics is not a prerequisite for studying thermodynamics. So far as thermodynamics is concerned, the isolated system has a certain number of quantum states, Ω, and every quantum state is equally probable after the system is isolated for a long time.

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

Equilibrium. I roll a fair die a few times, and the fraction of the times that the die lands on a particular face may not be 1/6. I roll the fair die many times, and the fraction of the times that the die lands on a particular face approaches 1/6.

I shake a half bottle of water. Water sloshes and bubbles pop. I stop shaking, and 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 H_2O molecules form water, and other H_2O molecules form steam.

Isolating a system for a long time is analogous to rolling a fair die many times. Several sentences mean the same:

- A system is isolated for a long time.

- An isolated system flipps to every one of its quantum states with equal probability.
- An isolated system is in *equilibrium*.

Whenever we speak of equilibrium, we must isolate a system for a long time. How long is long enough for an isolated system to attain equilibrium? This question is not studied in thermodynamics, but in kinetics.

In or out of equilibrium, an isolated system has a certain number of quantum states, Ω, and flips from one quantum state to another, ceaselessly and rapidly. Right after isolation, the system flips to some quantum states more often than others, and is said to be out of equilibrium. 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.

Exercise. Compare equilibrium in mechanics and in thermodynamics.

Internal variable

Random variable. In probability, a *random variable* is a map from a sample space to another set. A map, also called a function, is a mathematical term specified by three sets. The *domain* of the map is the sample space of an experiment. The *codomain* of the map is another set. The *graph* of the map is a set of ordered pairs, such that each ordered pair consists of a sample point and an element in the codomain, with every sample point appearing in one and only one pair.

Consider a fair die with faces labeled as {a,b,c,d,e,f}, which constitutes the sample space of the experiment, a roll of the die.

Before rolling the die, we agree on a rule of winning: $2 for face a, $6 for face b, $1 for face c, $4 for face d, $0 for face e, $2 for face f. This rule of winning is a *random variable*, a map that sends each face of the die to an amount of winning. The domain of the map is the sample space—the set of the six faces:

{a,b,c,d,e,f}

The codomain of the map is the set of five amounts of winning:

{$0, $1, $2, $4, $6}

The graph of the map is the rule of winning—a set of six ordered pairs:

{(a,$2), (b,$6),(c,$1), (d,$4), (e,$0), (f,$2)}

Whereas the rule of winning is deterministic, the face obtained from each roll of the die is random, and the amount of winning after each roll of the die is random.

Exercise. We agree on the above rule of winning. What is the probability of winning $7 after rolling a fair die three times?

The codomain—the set of winnings—need not be a collection of numbers. For example, the set of winnings can be {a candy, a cookie, a crayon}.

Constraint internal to an isolated system. Let us make a half bottle of water an isolated system. Inside the bottle, some H_2O molecules form water, and others form steam. The molecules can exchange between water and steam. We can fix the number of H_2O molecules in the steam in the bottle by placing a seal between water and steam. The seal is a *constraint internal to an isolated system*.

When the seal is removed, the isolated system can flip to every quantum state in the sample space. When the seal is in place, the number of H_2O molecules in the steam is fixed, and the isolated system can only flip to quantum states in a *subset* of the sample space.

When the seal is removed, H_2O molecules can go between the water and steam, and the number of H_2O molecules in the steam, x, can change. When the seal is in place, the number of H_2O molecules in the steam, x, is fixed. The number x is called an *internal variable*.

Internal variable. We now abstract the idea. In thermodynamics, a map from a sample space to another set is called an *internal variable*. Consider an isolated system, which has a sample space of quantum states. Let $\{x_1,..., x_n\}$ be another set. An internal variable is a map from the sample space to the set $\{x_1,..., x_n\}$. The sample space is the domain of the map, and the set $\{x_1,..., x_n\}$ is the codomain of the map.

The domain of the map, the sample space of an isolated system, is a set of quantum states, not a set of numbers. For the example of a half bottle of water, the elements in the set $\{x_1,..., x_n\}$ are numbers—the numbers of H_2O molecules in the steam. But in general, the codomain of the map, $\{x_1,..., x_n\}$, need not be a set of numbers, and can be a set of things. For example, in the Boltzmann distribution, the codomain is a set of things.

An internal variable dissects a sample space into a collection of subsets. An isolated system flips among a set of quantum states, which constitute the sample space. Denote the number of quantum states by Ω.

We have described the domain and codomain of the map. Let us turn to the graph of the map. An internal variable x can take any value in the codomain $\{x_1,..., x_n\}$. When a constraint internal to the isolated system fixes an internal variable at a value x, the isolated system flips among the quantum states in a *subset of the sample space*. Denote the number of quantum states in this subset by Ω(x), which measures the size of this subset of the sample space.

For example, for the isolated system of a half bottle of water, Ω(x) means the number of quantum states of the isolated system subject to the internal constraint that the number of H_2O molecules in the steam is x.

Probability distribution of an internal variable. The internal variable dissects the sample space into a family of subsets. Any two subsets in the family share no quantum states. The union of all the subsets in the family is the sample space. Thus,

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

After the constraint is removed, the system isolated for a long time flips to every one of its Ω quantum states in the sample space with equal probability, and the isolated system is said to have reached equilibrium. In equilibrium, the internal variable can take any value in the set $\{x_1,..., x_n\}$, and the probability for the internal variable to take a particular value x is

$\Omega(x)/\Omega$

This is the probability for the isolated system to flip to the quantum states in a subset of the sample space.

An isolated system has a sample space of Ω quantum states. An internal variable x maps the sample space to the codomain:

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

The internal variable directs the sample space into a collection of subsets. The subsets have various numbers of quantum states:

$\Omega(x_1),\ldots, \Omega(x_2)$

When the constraint internal to the isolated system is removed and the system is isolated for a long time, the system flips to every quantum state in the sample space with equal probability, but flips to quantum states in various subsets with the probability distribution:

$\Omega(x_1)/\Omega,\ldots, \Omega(x_2)/\Omega$

Such a probability distribution is calculated in statistical mechanics. For an example, see the Boltzmann distribution.

From probability to (almost) certainty. In principle, an internal variable can dissect the sample space into a collection of subsets, with each subset containing just one quantum state. But that will be pointless. In thermodynamics, the codomains of internal variables are quantities that can be measured experimentally. Examples include the cast of supporting actors: energy, amount of matter, volume, and amount of charge. These internal variables are analogous to the number of H_2O molecules in the steam in a half bottle of water. Associated with each value x in the codomain is a subset of the sample space having a large number of quantum states, Ω(x).

For such an internal variable x, the function Ω(x) has a sharp peak. After the constraint is removed and the system is isolated, both for a long time, the measured value of the internal variable is well described by the value x that maximizes the function Ω(x). That is, the fluctuation of x is exceedingly small compared to the average of x.

This behavior is a consequence of the fundamental postulate, as well as the large number of particles that constitute an isolated system and the type of internal variables. An example is found in the development of the Boltzmann distribution, in which energy is the internal variable.

The sharp peak in the probability distribution of an internal variable is taken as given in formulating the basic algorithm of thermodynamics.

Logic of thermodynamics

The logic of thermodynamics is framed in terms of an isolated system with a sample space of quantum states, an internal variable x that dissects the sample space into a collection of subsets, and the number of quantum states in a subset, Ω(x).

The principle of the increase of Ω(x). The above discussion suggests the following principle:

After a system is isolated and an internal variable x is allowed to vary, x changes such that Ω(x) increases, and attains equilibrium such that Ω(x) maximizes.

This is a statement of the second law of thermodynamics. We unpack this statement and turn it into the basic algorithm of thermodynamics. Numerous statements of the second law litter in the literature. There is no need to decipher them if they do not speak to you.

State of constrained equilibrium. In probability, an element in the sample space is called an *outcome*, and a subset of the sample space is called an *event*.

In thermodynamics, an element in the sample space is called a *quantum state*, and a subset of the sample space is called a *state of constrained equilibrium*.

When the seal in the half bottle of water is in place, the two parts of the bottle are two distinct isolated systems. They can separately reach equilibrium, but are not in equilibrium with each other. The isolated system is in a state of constrained equilibrium.

When a system is isolated and an internal variable is fixed at a value x = a, both for a long time, the system flips to every quantum state in a subset of the sample space with equal probability. The isolated system is said to be in a state of constrained equilibrium at x = a.

We now use the word “state” in two ways. A quantum state is a sample point in the sample space of an isolated system. A state of constrained equilibrium is a subset of the sample space when an internal variable is fixed at a value. For brevity, we often just use the word “state”, and let the context distinguish its meaning.

Process. When an internal variable x changes from one value, x = a, to another value, x = b, an isolated system is said to undergo a *process*.

Associated with the two values of the internal variable, x = a and x = b, are two states of constrained equilibrium, which mark the beginning and the end of the process. But the change *between* the beginning and the end need not be associated with any state of constrained equilibrium.

In speaking of a thermodynamic process, we always identify an isolated system, an internal variable, x, and two states of constrained equilibrium, x = a and x = b. The two states of constrained equilibrium correspond to different subsets of the sample space, with different numbers of quantum states, Ω(a) and Ω(b).

Impossible, reversible, and irreversible processes. Difference in the two numbers distinguishes three cases. A process from one state of constrained equilibrium x = a to another state of constrained equilibrium x = b is *impossible* if Ω(a) > Ω(b), *reversible* if Ω(a) = Ω(b), and *irreversible* if Ω(a) < Ω(b).

A measure of irreversibility of a process. An isolated system undergoing an actual process always changes from a subset having fewer quantum states to a subset having more quantum states. Once two values of the internal variable are identified, x = a and x = b, many actual processes may exist for the isolated system to change from the constrained state of equilibrium a to b. All these actual processes have the same degree of irreversibility, quantified by the ratio Ω(b)/Ω(a).

A reversible process is an idealization. A reversible process from a to b is synonymous to the condition Ω(a) = Ω(b), and to a sequence of states of constrained equilibrium, each having the same number of quantum states.

Every actual process is irreversible. A reversible process is an idealization.

Tweet The logic of thermodynamics

Entropy

The logic of thermodynamics has just been framed in terms of the function Ω(x). We next define entropy. As will become evident, entropy adds no physical content, but adds algebraic convenience.

Definition of entropy. Let Ω be the number of quantum states of an isolated system. Define the *entropy* of the isolated system by

$S = \log \Omega$

The entropy of an isolated system is defined as the logarithm of the number of quantum states of the isolated system.

Logarithm of any base will do. We will use the natural base e. Recall a fact of calculus

$d \log\Omega/d\Omega = 1/\Omega$.

Had we adopt any other base b, thermodynamics would be littered with a numerical factor.

The logarithm function is the inverse of the exponential function, so that

$\Omega = \exp(S)$

Thermodynamics is replete with logarithm and exponential functions, mostly originating from the definition of entropy.

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

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

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

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

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:

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

Here $\Omega_{composite}$ is the number of quantum states of the composite. Recall a fact of logarithm:

$\log (\Omega_A \Omega_B) = \log \Omega_A + \log \Omega_B$

Logarithm turns a product into 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*. Being extensive brings algebraic convenience, though it seems unfair for entropy to steal the show from the number of quantum states.

Exercise. In defining entropy, we need a function that obeys $f(xy) = f(x) + f(y)$. Why do we need such a function? Show that logarithm is the only candidate.

Entropy is dimensionless. By definition, $S = \log \Omega$, entropy is *dimensionless*.

Entropy is absolute. By definition, $S = \log \Omega$, zero entropy is not something arbitrary, but has empirical significance. Zero entropy corresponds to an isolated system of a single quantum state. Like volume, entropy is *absolute*.

Ice, water, and steam coexist in equilibrium. Here is a fact of nature. Ice, water, and steam coexist in equilibrium at a temperature of 0.01 Celsius and a pressure of 611.73 Pascal. Listed below are the entropies per molecule of ice, water, and steam in equilibrium. Recall the definition of entropy, $S = \log \Omega$. Given a value of entropy per molecule, the corresponding number of quantum states is $\Omega = \exp(S)$.

	entropy per molecule	number of quantum states
ice	4.976	144.89
water	7.619	2036.52
steam	27.44	8×10^{11}

In the ice, the molecules form a lattice and can not freely translate and rotate, giving a small number of quantum states. In the water, the molecules can translate and rotate, but still touch one another, giving a larger number of quantum states. In the steam, the molecules can fly, giving the largest number of quantum states: almost a trillion quantum states.

In equilibrium, the steam has so many more quantum states than ice and water. How can ice and water survive? Later we will interpret this equilibrium, and describe how these entropies are measured.

For a collection of H_2O molecules, the entropy of the collection is proportional to the number of molecules, N. In each thermodynamic state, the number of the quantum states in the collection is the number listed in the table raised to power N. Recall that 18 g of water has 6.022×10^{23} molecules. Need we say more?

Entropy and energy. Now that you have met both entropy and energy, you are ready to critique the passage in which Clausius (1865) coined the word “entropy”.

“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.”

Exercise. Critique this extract. Do you agree that entropy and energy are nearly allied in their physical meanings?

Tweet Entropy

Subset entropy S(x)

Let x be an internal variable of an isolated system, and Ω(x) be the number of quantum states in the subset of the sample space of the isolated system when the internal variable takes value x. Define

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

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 x”. For brevity, I will call the function S(x) the *subset entropy* of the isolated system. Let me know if you find a better name for this function. The subset entropy of the isolated system is a function of the internal variable.

Because logarithm is an increasing function, increasing Ω(x) is equivalent to increasing S(x).

Impossible, reversible, and irreversible processes in terms of S(x). A process from a state of constrained equilibrium at x = a to another state of constrained equilibrium at x = b is impossible if S(a) > S(b), reversible if S(a) = S(b), and irreversible if S(a) < S(b).

All actual processes are irreversible. For an irreversible process of an isolated system from a state of constrained equilibrium at x = a to another state of constrained equilibrium at x = b, the increase in subset entropy, S(b) - S(a), called *entropy generation*, measures the degree of irreversibility of the process.

Only an idealized process from a to b keeps the subset entropy constant, S(a) = S(b), and is called a reversible process.

Basic algorithm of thermodynamics

The above discussions suggest the basic algorithm of thermodynamics.

1. Construct an isolated system with an internal variable x.
2. Find the subset entropy of the isolated system as a function of the internal variable, S(x).
3. When the internal variable x is free to change, the isolated system equilibrates if the internal variable reaches the value that maximizes S(x).
4. Given two values of the internal variable, x = a and x = b, a process from the state of constrained equilibrium a to the state of constrained equilibrium b is *impossible* if S(a) > S(b), *irreversible* if S(a) < S(b), and *reversible* if S(a) = S(b).

An isolated system may have multiple internal variables. A list of values of the internal variables specifies a state of constrained equilibrium. The BAT runs just the same for an isolated system of any number of internal variables.

The BAT mentions no supporting actors. Thermodynamics is hard enough without the commotion of that particular energetic supporting actor. But BAT provides a point of entry for the supporting actors. They enter as internal variables when the algorithm is applied to various situations.

Thermodynamics invokes time in two ways. In Step 3, equilibrium is reached after the system is isolated and the constraint is lifted, both for a long time. In Step 4, the subset entropy increases in time. Neither step invokes exact duration of time. No quantity with the unit of time appears in thermodynamics.

The logic of thermodynamics leads to the basic algorithm of thermodynamics, which are steps of actions, like calculus. The underlying logic may be scrutinized for rigor, but the algorithm itself enables the application of thermodynamics to various situations. The remainder of the review illustrates the application of this algorithm. As we will show, all applications of thermodynamics run on this basic algorithm.

Exercise. It seems unfair for entropy to steal the show from the number of quantum states. The basic algorithm of thermodynamics works fine with the function Ω(x). Why do we hide it behind a log? On what basis does entropy prevail over the number of quantum states?

Tweet Basic Algorithm of Thermodynamics

Thermal system

We have met an all-star cast of actors. Let us watch them play, act by act. The remaining sections apply the basic algorithm of thermodynamics to various situations.

Act one: **The union of entropy and energy produces a child—temperature**. The child is so prodigious that it is better known than its parents. Let us watch entropy and energy (1) unite and (2) produce temperature. We watch these in thermal systems.

A thermal system is a family of isolated systems of a single variable

A thermal system is a system that transfers energy with its surroundings, but transfers neither space nor matter. A thermal system is called a thermal source on losing energy, and a thermal sink on gaining energy.

Characterize a thermal system by a function Ω(U). Let U be the energy of a thermal system. When the energy U is fixed, the thermal system becomes an isolated system. This isolated system has a fixed number of quantum states, denoted by Ω(U).

When the thermal system changes its energy U, the thermal system is *a family of isolated systems*. Each member isolated system in the family has a specific value of energy, U, flips among its own sample space of quantum states, and has its own number of quantum states, Ω(U). The family of isolated systems has a single h *independent variable*, the energy U.

The function Ω(U) fully characterizes the thermodynamics of the thermal system. Once the function Ω(U) is known, all other thermodynamic functions of the thermal system are known, as we will show. Later we will determine the function Ω(U) by experiment—that is, we will count the number of quantum states of each member isolated system experimentally.

Hydrogen atom. A hydrogen atom can be classified as a thermal system. The 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 characteristic function of a hydrogen atom, Ω(U), has been determined 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 are discrete, called energy levels.

A half bottle of water. A half bottle of water can be made into a thermal system. In the bottle, some H_2O molecules form water, and others form steam. Molecules can go from one phase to the other. Cap the bottle, so that molecules do not leak. Make the bottle rigid, so that the bottle does not change volume.

The molecules in the bottle store energy in the form of molecular motion and interaction. This form of energy is called thermal energy.

When we place the bottle over a fire, the chemical energy of the fuel converts to the thermal energy in the molecules in the bottle.

When we shake the bottle, the mechanical energy of the motion converts to the thermal energy of the molecules in the bottle.

When we pass an electric current in a resistor placed in the bottle, the electrical energy converts to the thermal energy of the molecules in the bottle.

A half bottle of water has many particles—protons, neutrons, and electrons. The energies of such a thermal system are so closely spaced that we regard the characteristic function, Ω(U), as a continuous function.

Exercise. Describe another thermal system and methods of transfer energy to the thermal system.

Characterize a thermal system by a function S(U). Define entropy by

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

The two functions, S(U) and Ω(U), have the same physical content, and characterize the same thermal system. But S(U) has the attribute that both S and U are extensive quantities. This attribute adds no physical content, but adds algebraic convenience.

Converting energy of various forms to thermal energy

Converting chemical energy to thermal energy. When I shake a half bottle of water, the chemical energy in my muscles transfers to thermal energy in H_2O molecules in the bottle. This transfer of energy goes through several steps. The chemical energy in the muscles moves my hand, the motion of my hand causes the water in the bottle to slosh, and the slosh of water transfers mechanical energy of water to thermal energy in H_2O molecules. The conversion of the chemical energy to thermal energy is irreversible: the thermal energy in the H_2O molecules will not become the chemical energy in the muscles.

Converting electrical energy to thermal energy. I can also transfer energy to the H_2O molecules in the bottle using an electric current. This transfer of energy goes through several steps too. The voltage of the electric outlet moves electrons in the metal wire. The moving elections bump atoms in the metal wire, and transfer the electrical energy to thermal energy. The metal wire is immersed in the water, and transfers thermal energy in the metal wire to the thermal energy in the H_2O molecules. The conversion of the electrical energy to thermal energy is irreversible: the thermal energy in the H_2O molecules will not become the electrical energy in the outlet.

Converting mechanical energy to thermal energy. I can also drop a weight from some height into the bottle of water. The weight comes to rest in the water, converting the mechanical energy of the weight to the thermal energy in the H_2O molecules. The conversion is irreversible: the thermal energy in the H_2O molecules will not return to the potential energy of the weight.

Thermal energy is low-grade energy. When the potential energy of the weight converts into thermal energy in water, we say that energy *dissipates*. The total amount of energy is conserved. The potential energy of the weight converts into thermal energy in water, but the thermal energy in water cannot raise the weight to the initial height. Thermal energy is called *low-grade energy*.

The irreversibility of the dissipation of energy is understood from the molecular picture. Thermal energy is stored in the motion and interaction at the molecular scale, 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.

Dissipation of energy

A falling apple. Facts about a falling apple are well known. As the apple loses height, it gains velocity. So long as the friction between the apple and the air is negligible, the sum of the potential energy and kinetic energy of the apple is constant.

We model the apple, together with a part of space around the apple, as an isolated system. All the quantum states of the isolated system constitute the sample space.

The height of the apple is the internal variable of the isolated system. When the apple is at height z, the isolated system flips among a subset of quantum states. Denote the number of the quantum states in this subset by $\Omega(z)$.

Say the process of falling keeps the number $\Omega(z)$ fixed, independent of the internal variable z. Such a 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 curved slide. So long as friction is negligible, the apple will return to the same height.

Of course, friction is inevitable in reality. An apple cannot go up and down a slide for a long time. The apple will stop after some time. But a frictionless process can be a useful idealization. For example, the planet Earth has moved around the Sun for a very long time.

Exercise. Describe another reversible process.

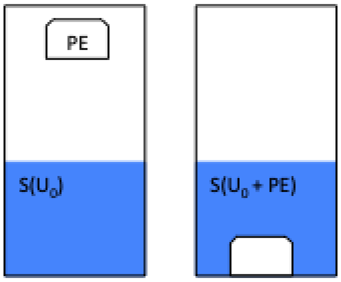
Athermal energy reservoir. To turn this tale of a falling apple into a general principle, we introduce the phrase *athermal energy reservoir*, by which we mean a system that changes energy but not entropy. A weight is an athermal energy reservoir. The weight times its height is the potential energy. When the weight is raised, its potential energy increases, but its entropy remains constant.

Various weights at various heights constitute an athermal energy reservoir. This athermal energy reservoir gains or loses energy without changing entropy of the reservoir itself.

A flying wheel is also an athermal energy reservoir. The kinetic energy of the wheel increases with the angular velocity of the wheel. When the wheel spins faster, the kinetic energy increases, but the entropy remains constant.

We next analyze the falling weight into water using the basic algorithm of thermodynamics (BAT). We analyze dissipation of energy in an idealized model: energy goes from an athermal energy reservoir to a thermal system. The former loses energy but does not change entropy. The latter gains energy and entropy. The athermal system and the thermal system together constitute an isolated system. During this process, the isolated system conserves energy, but increases entropy. The process is irreversible. Energy is dissipated in the sense it goes from the athermal energy reservoir to the thermal system.

Step 1. Construct an isolated system with an internal variable. Picture an athermal energy reservoir by a weight, and a thermal system by a half bottle of water. Together they constitute an isolated system. Before the weight drops, the thermal energy of the isolated system is U_0, and the mechanical energy of the isolated system is the potential energy of the weight PE. The isolated system conserves energy. After the weight comes to rest in water, the potential energy of the weight vanishes, and the thermal energy of the isolated system is U_0 + PE. The thermal energy, U, is an internal variable of the isolated system, increasing from the initial value U_0 to the final value U_0 + PE.



Step 2. Find the subset entropy of the isolated system as a function of the internal variable. As the weight drops, the potential energy lowers, but the weight itself does not change entropy. On gaining energy, the half bottle of water increases entropy. The half bottle of water is a thermal system characterized by entropy as a function of energy, S(U), which is an increasing function. This function is also the subset entropy of the isolated system.

Step 3. Maximize the subset entropy to equilibrate. Before the weight drops, the subset entropy is $S(U_0)$. After the weight comes to rest in water, the subset entropy is $S(U_0 + PE)$. Because S(U) is an increasing function, the isolated system maximizes the subset entropy when the mechanical energy fully changes to the internal energy, the weight comes to rest, and the isolated system reaches equilibrium.

Step 4. Increase the subset entropy to change irreversibly. The reverse change would violate the law of the increase of entropy. The weight, after coming to rest in water, will not draw thermal energy and jump up. Dissipation is irreversible. What makes thermal energy low-grade energy is its high entropy.

Thus, the dissipation of energy—the conversion of potential energy to thermal energy—results from a basic fact: S(U) is an increasing function.

The Kelvin-Planck statement of the second law of thermodynamics. This analysis confirms an empirical fact. It is impossible to produce no effect other than the raising of a weight by drawing thermal energy from a single thermal source. This fact 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. In *Heat considered as a Mode of Motion*, published in 1863, John Tyndall described numerous experiments that tested the hypothesis of thermal energy. Wiki Julius Robert von Mayer. Wiki James Prescott Joule. The following passage is taken from the book by Tyndall.

“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.”

Assignment. Use the basic algorithm of thermodynamics to analyze energy dissipation by friction. Identify an isolated system with an internal variable x. Identify the function S(x). Interpret equilibrium. Interpret irreversibility.

Thermal contact

Two thermal systems are said to be in *thermal contact* if they transfer energy to each other.

Two thermal systems in thermal contact transfer energy between each other through microscopic motion and interaction. For example, a half bottle of water is one thermal system, and a tank of oil is another thermal system. The former has been kept over a fire, and the latter has been left in the room. When the half bottle of water is submerged in the tank of oil, energy transfers from water to oil. The motion and interaction in the water molecules affect the motion and interaction in the molecules of the wall of the bottle, which affect the motion and interaction in the oil molecules. Thermal contact transfers neither matter nor space between the two thermal systems.

We next study thermal contact by running the basic algorithm of thermodynamics (BAT).

Step 1. Construct an isolated system with an internal variable. Let U_A be the energy of thermal system A, and U_B be the energy of thermal system B. Energy is an extensive quantity, so that the energy of the two thermal systems together is the sum: $U_A + U_B$.

The two thermal systems together constitute an isolated system if the sum of their energies is fixed:

$U_A + U_B =$ constant.

In other words, the principle of the conservation of energy requires the isolated system to have a fixed amount of energy.

When system A gains an amount of energy, system B loses the same amount of energy. Thermal contact is a zero-sum game. Consequently, the isolated system has a single independent internal variable, say, the energy of one of the thermal systems, U_A.

Step 2. Find the subset entropy of the isolated system as a function of the internal variable*.* Thermal system A is characterized by function $S_A(U_A)$, and thermal system B is characterized by function $S_B(U_B)$.

When the energy transfers without constraint between the two thermal systems, the isolated system flips among all quantum states in the sample space. When the internal variable U_A is fixed by insulating system A, the isolated system flips among quantum states in a subset of the sample space. Entropy is an extensive quantity, so that the subset entropy of the isolated system is a sum:

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

Thermal system A can be insulated at various values of energy, U_A. Consequently, the subset entropy of the isolated system is a function of a single independent internal variable, $S_{iso}(U_A)$.

Step 3. Maximize the subset entropy to equilibrate*.* Given a constant energy U_{iso}, how do the two thermal systems partition this energy in equilibrium?

Typically, $S_A(U_A)$ and $S_B(U_B)$ are both increasing functions. When energy transfers from system A to system B, $S_A(U_A)$ decreases but $S_B(U_B)$ increases. Consequently, the function $S_{iso}(U_A)$ is not monotonic.

According to calculus,

$dS_{iso}(U_A) = (dS_A(U_A)/dU_A)dU_A + (dS_B(U_B)/dU_B)dU_B$.

According to the principle of the conservation of energy, when system A gains an amount of energy dU_A, system B loses energy by the same amount, $dU_B = - dU_A$. Write the above equation as

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

After being isolated for a long time, the isolated system reaches equilibrium, and the two thermal systems in thermal contact are said to reach *thermal equilibrium*. The isolated system reaches equilibrium when the subset entropy of the isolated system maximizes, $dS_{iso}(U_A) = 0$, so that

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

This equation is the condition of thermal equilibrium. This condition, along with the conservation of energy, $U_A + U_B$ = constant, partitions the fixed amount of energy between the two thermal systems A and B in equilibrium.

A thermal system is characterized by a function S(U). The function is a curve in the energy-entropy plane. The analysis identifies the importance of the slope of the energy-entropy curve, dS(U)/dU. In thermal equilibrium, two thermal systems have an equal slope.

Step 4. Increase the subset entropy to change irreversibly. Prior to reaching equilibrium, the subset entropy of the isolated system increases in time, $dS_{iso}(U_A) > 0$. The increase of the subset entropy requires that

$(dS_A(U_A)/dU_A - dS_B(U_B)/dU_B)dU_A > 0$.

Distinguish two possibilities:

- If $dS_A(U_A)/dU_A > dS_B(U_B)/dU_B$, the increase of the subset entropy requires that $dU_A > 0$, so that system A gains energy from system B.
- If $dS_A(U_A)/dU_A < dS_B(U_B)/dU_B$, the increase of the subset entropy requires that $dU_A < 0$, so that system A loses energy to system B.

In thermal contact, the two thermal systems transfer energy in one direction: from a thermal system of a low slope to a thermal system of a high slope.

We next assume that $S_A(U_A)$ and $S_B(U_B)$ are *convex functions*. We will critique this assumption later, but accept it for the time being. When $dS_A(U_A)/dU_A > dS_B(U_B)/dU_B$, system A gains energy from system B. 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. The transfer of energy continues until the two slopes are equal, $dS_A(U_A)/dU_A = dS_B(U_B)/dU_B$.

Twitter Thermal contact

Temperature

Definition of temperature. A thermal system is characterized by a function S(U). Running the BAT on two thermal systems in thermal contact identifies the significance of the derivative of the function, dS(U)/dU.

Define temperature T by

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

This definition conforms to our intuition about temperature as follows.

Step 3 of the BAT shows that two thermal systems in thermal contact equilibrate when their temperatures are equal.

Step 4 of the BAT shows that, when two thermal systems are not in equilibrium, energy goes from a thermal system of higher temperature to a thermal system of lower temperature.

Clausius (1854) discovered the relation $1/T = dS(U)/dU$. He then took energy and temperature as given, and used this relation to define entropy. His presentation has been regurgitated in many textbooks to this day, and is a main source of confusion for students of thermodynamics.

Today, we define entropy as the logarithm of the number of quantum states of an isolated system, energy by the principle of the conservation of energy, and temperature by the relation, $1/T = dS(U)/dU$. Before these definitions, our ancestors had long associated temperature to the sense of hotness, but did not know the origin of temperature. Temperature was an orphan. This relation defines temperature as a *derivative* of entropy and energy. The child has found its parents.

What's important is the slope of the function, dS(U)/dU, not how we name the slope. But wait a minute! We associate temperature with a quantity that is equal for two thermal systems in thermal equilibrium. Any function of $dS(U)/dU$ also serves 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 of the entropy-energy curve, $dS(U)/dU$. Nature is indifferent to what humans call this slope.

Temperature is absolute. Entropy is an absolute property, but energy is a relative property. Because the definition of temperature only uses the difference in energy, temperature is an absolute property.

Temperature is positive. As we noted before, $dS(U)/dU > 0$ commonly holds for thermal systems. Thus, the temperature is positive. Furthermore, as $S \to 0$, $dS(U)/dU \to \infty$. Thus, as $S \to 0$, $T \to 0$. The ground state has temperature zero.

Temperature is intensive. Both S and U are extensive properties. When a thermal system scales by a factor b, entropy and energy also scale by the same factor b. Consequently, the slope $dS(U)/dU$ does not change. Temperature is an intensive property.

Temperature has the same unit as energy. Because entropy is a dimensionless number, $S = \log \Omega$, temperature has the same unit as energy, Joule.

Joule is a giant unit of temperature. By convention, people report temperature using another unit, Kelvin. The 2019 redefinition of the SI base units defines the conversion of the two units—Kelvin and Joule—by

1 Kelvin = 1.380649×10^{-23} Joules

The conversion factor is called the Boltzmann constant, k_B. For practical calculation, take $k_B = 1.38 \times 10^{-23}$ J/K.

Nature's unit vs. humans's unit. The Boltzmann constant k_B has no significance in nature; it converts two units of temperature, Joule and Kelvin. The unit of energy is nature's unit of temperature. Kelvin is a human's unit of temperature.

Nature is indifferent to how humans define a unit for temperature. To describe any behavior of nature, if temperature T is in the unit of Kelvin, the product k_BT must appear together. Thus, the introduction of the unit Kelvin clutters all facts of nature with the factor of conversion between the two units of temperature, the Boltzmann constant k_B.

Incidentally, many people call k_BT thermal energy. This designation is wrong. There is no need to give any other interpretation: when T is the temperature in the unit of Kelvin, k_BT is the temperature in the unit of energy.

Fake unit of entropy. When temperature is reported in Kelvin, to maintain the equation $1/T = dS(U)/dU$, the entropy need be in the fake unit of the Boltzmann constant:

$S = k_B \log \Omega$

The change of unit for temperature leads to a fake unit for entropy: Joule/Kelvin. If we report entropy S in this fake unit, only the ratio S/k_B can enter any fundamental result. Nature is indifferent to human's units.

Twitter What does the Boltzmann constant mean?

Exercise. Wiki Boltzmann constant. Describe the history of the fake unit, J/K.

Various temperatures form a scalar set. A set is called a *scalar set* if any particular nonzero element in the set times various numbers gives all elements in the set. Each element in the set is called a *scalar*. The particular nonzero element is called a *unit*.

Let u be a nonzero element in a scalar set. Let λ be a number. Their multiplication $x = \lambda u$ is a scalar in the set. The number λ is called the *magnitude* of the scalar x relative to the unit u.

Various amounts of energy form a scalar set. So do various amounts of charge and various amounts of mass. Energy, charge, and mass are extensive properties.

Various amounts of entropy form a scalar set. Consider an isolated system that has Ω quantum states. If we place two such isolated systems together, the new isolated system has Ω^2 quantum states. The entropy of the isolated system is $\log \Omega$, and the entropy of the new isolated system is $2 \log \Omega$. Entropy, too, is an extensive property.

Temperature T is defined by $1/T = dS(U)/dU$. Rewrite this definition as

$$dU = TdS$$

Various amounts of increment in entropy form a scalar set, and various amounts of increment in energy form another scalar set. A fixed temperature is a linear map from one scalar set to another scalar set. According to linear algebra, all linear maps from one scalar set to another scalar set form a scalar set. Thus, various temperatures form a scalar set.

Let us see what these statements mean to a thermal system characterized by a function S(U). Temperature is a linear map in that, for any number λ,

$$T(\lambda dS) = \lambda(TdS)$$

At a state of the thermal system, of an energy U and a temperature T, when the increment of entropy changes from dS to λdS, the increment of energy changes from TdS to $T(\lambda dS)$.

Now consider the set of all temperatures. Let T be a temperature, $dU = TdS$. For any number λ, λT is also a temperature in that

$$(\lambda T)dS = \lambda(TdS)$$

Find two states of the thermal system at temperatures T and λT. The same increment of entropy dS will change the increment of energy from TdS to $(\lambda T)dS$.

Tweet Temperature is a scalar

Calorimetry

The art of measuring thermal energy is called *calorimetry*. A device that measures thermal energy is called a *calorimeter*.

Calorimetry has become an art of high sophistication. 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.

This said, it is good to have a specific method of calorimetry in mind. Just think of an electric heater in a tank of water. The electric energy is (current)(voltage)(time), assumed to be fully converted to the thermal energy in the water.

Wikipedia Calorimetry.

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 accuracy, ease, and cost. For example, our fingers are thermometers. So are many parts of our bodies. But these thermometers are unreliable and function in a narrow range of temperature. We next describe two commonly used thermometers.

Exercise. A block of steel and a block of wood have the same temperature, which is higher than our body temperature. On touching them with a finger, you feel the steel is hotter than the wood. Why?

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

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

$PV = Nk_BT$

This equation relates temperature T to P, V, and N. We will show later that the temperature defined by the ideal gas law coincides with the temperature defined by the equation $1/T = dS(U)/dU$. The quantities P, V, and N can be measured, so that an ideal gas serves as a thermometer, called the *gas thermometer*.

Exercise. Watch a video of a gas thermometer at constant volume. Explain how this thermometer works.

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

The manufacturer of the thermometer must assign a volume of the liquid 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 temperature defined by the relation $1/T = dS(U)/dU$.

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 very few idealized systems. Thermodynamics helps us understand temperature and design thermometers.

Exercise. YouTube temperature sensor. Explain how the thermocouple works.

H. Chang, Inventing Temperature, 2004.
W.E.K. Middleton, A History of the Thermometer and Its Uses in Meteorology, 1966.

Experimentally count the number of quantum states

Given an isolated system, counting the number of its quantum states is a basic problem in thermodynamics. A commonly used experimental method is described here, using a thermal system as an example.

When the energy U is fixed, a thermal system becomes an isolated system. Denote the number of quantum states in this isolated system by $\Omega(U)$. When U varies, the function $\Omega(U)$ characterizes the thermal system as a family of isolated systems. Denote $S(U) = \log \Omega(U)$.

We have just talked about the following experiments.

1. Temperature can be measured experimentally (thermometry).
2. Thermal energy U can be measured experimentally (calorimetry).

You can measure temperature as a function of internal energy, $T(U)$. Recall the definition of temperature T: $1/T = dS(U)/dU$. An integration gives the function $S(U)$.

The integration starts as the absolute zero, $T \to 0$, where $S \to 0$. It is challenging to measure S near the absolute zero. The Debye (1912) predicts that, at low temperatures, entropy takes the form:

$S = aT^3$

When the constant a is fit by experimentally measured entropies at low temperatures, the above equation extrapolates the experimental data to the absolute zero.

Absolute entropies have been measured for numerous systems. Once an absolute entropy is determined, the number of quantum states is

$\Omega = \exp(S)$

Twitter Experimentally count the number of quantum states

Thermal capacity

A thermal system is characterized by a function S(U). The inverse slope of this function defines another function T(U). Define the *thermal capacity* C by

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

Rewrite the definition by

$dU = CdT$

Thus, the thermal capacity is a linear map from the increment in temperature to the increment in energy.

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

By definition, thermal capacity is a function of energy, C(U). For many thermal systems, the function S(U) is convex upward, so that T(U) is a monotonically increasing function, and $C(T) > 0$.

Energy-entropy plane is a two-dimensional vector space

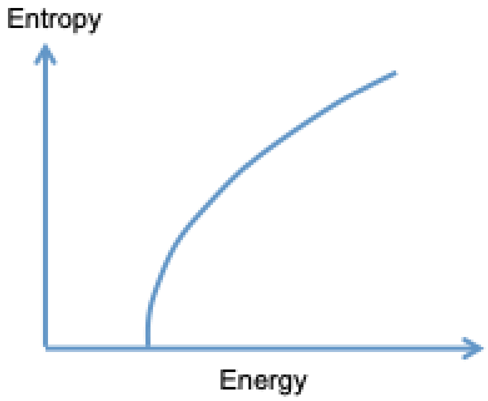
Energy and entropy are extensive variables. We add energy to energy, and multiply a real number to energy. We add entropy to entropy, and multiply a real number to entropy.

We do not add energy and entropy. Nor do we multiply them.

A pair of values (U,S) is a two-dimensional vector. The set of all such vectors forms a two-dimensional vector space. We can add two energy-entropy vectors, which means adding energy to energy, and entropy to entropy. We can also multiply an energy-entropy by a number. However, the length of an energy-entropy vector is undefined and meaningless. Also undefined is the angle between two energy-entropy vectors. In the language of linear algebra, the energy-entropy space is a vector space with no inner product.

We represent the vector space of energy and entropy by a plane. Draw energy as a horizontal axis, and entropy as a vertical axis. Each point in the plane represents a pair of values of energy and entropy. We usually draw the two axes perpendicularly, but this practice is unnecessary; two axes not perpendicular to each other are just as good. The angle between two energy-entropy vectors is undefined and has no significance.

A thermal system corresponds to a curve in the energy-entropy plane. A thermal system is characterized by a function S(U). The function is a *curve* in the energy-entropy plane. Thus, a thermal system is characterized by a curve on the energy-entropy plane.



A function S(U) is specific to one thermal system, but many thermal systems share common features. We list these features in mathematical terms of the curve S(U), and will relate these features to experimental observations as we progress. We will also comment if any of these features are absent in a specific thermal system.

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 curve S(U) is smooth. (There are significant exceptions to this feature, and will be discussed later.)

4. The more energy, the more quantum states—that is, $\Omega(U)$ is an increasing function. Because logarithm is an increasing function, $S(U)$ is also an increasing function. Thus, the slope of the curve $S(U)$ is positive, $dS(U)/dU > 0$. The slope defines temperature T by $1/T = dS(U)/dU$.
5. As entropy approaches zero, the curve $S(U)$ approaches the energy axis vertically. That is, as $S \rightarrow 0$, $dS(U)/dU \rightarrow \infty$.
6. The curve $S(U)$ is *convex upward*. That is, the slope $dS(U)/dU$ decreases as U increases, or equivalently, $d^2S(U)/dU^2 < 0$. (There are significant exceptions to this feature. Later we will examine thermal systems of which the curve $S(U)$ is *not* convex upward.) Define thermal capacity C by $1/C = dT(U)/dU$.

Tweet. Entropy-energy plane

Phase and state

Pure substance. I hold a half bottle of water. One half of the bottle is filled with liquid, and the other half of the bottle is filled with gas. The half bottle of water is *not* half empty.

The half bottle of water contains molecules other than H_2O, such as N_2 and O_2. For the time being, consider an idealization that the bottle only contains H_2O molecules, and all other species of molecules are absent.

By a *pure substance* we mean a large number of a single component (molecules or atoms). For example, a large number of H_2O molecules is a pure substance.

Phase. A pure substance can aggregate in three forms: solid, liquid, and gas. Each form is called a *phase*.

For H_2O molecules, the three phases have distinct names: *ice*, *water*, and *steam*. YouTube ice, water, steam.

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

In a liquid, molecules touch one another, but do not form a periodic lattice. Molecules change neighbors readily. They also rotate. Wiki water

In a gas, molecules on average are far apart. They fly, collide, and separate. Wiki steam.

A pure substance may form more than three phases. For example, H_2O molecules form multiple phases of ice at high pressure. Wiki ice. Each solid phase has a distinct crystalline lattice. This fact of multiple solid phases is neglected in this course, unless otherwise stated.

Exercise. Write a few sentences on the usage of the following words: gas, vapor, fluid, steam, and air.

Thermodynamic state. A pure substance isolated for a long time reaches equilibrium, called a *(thermodynamic) state*.

Phase and state are different concepts. A substance in a phase can change state. For example, steam—a phase—can be compressed to a smaller volume, or heated to a higher temperature. Steams in different volumes and different temperatures are different states, but all these states belong to the same phase—the gas phase.

A pure substance can equilibrate in a single-phase state, a two-phase state, or a three-phase state.

For example, H_2O molecules can be in a single-phase state of ice, water, or steam. In a half bottle of water, the H_2O molecules equilibrate in a two-phase state, a mixture of water and steam. As we will see, H_2O molecules can also equilibrate in a three-phase state, a mixture of ice, water, and steam.

A pure substance cannot equilibrate in a state of four or more phases.

Property

A property is a map. The domain of the map is a set of thermodynamic states. The codomain of the map is a set of values of a quantity.

Whenever we speak of a thermodynamic property, we must identify a system isolated for a long time—that is, an isolated system in equilibrium.

We will mostly talk about five properties: temperature, volume, pressure, energy, and entropy. All five properties would look unfamiliar to ancient people.

When thermodynamics was created in the mid-nineteenth century, the creators were familiar with temperature, volume, and pressure, but had to invent energy and entropy.

Most of you enter this course feeling comfortable with temperature, volume, and pressure, but not with energy and entropy. By the end of this course, you will be familiar with all five properties, and more. Familiarity breeds comfort. Civilization advances.

When a half bottle of water equilibrates, the water and steam have the same temperature. Such a property is called an *intensive property*. Pressure is also an intensive property.

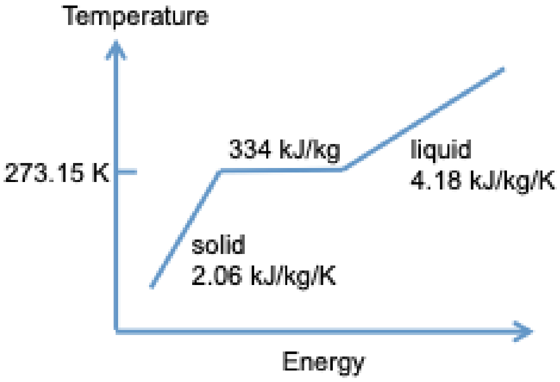
When the half bottle of water equilibrates, the volume of the water and the volume of steam add to give the volume of the bottle. Such a property is called an *extensive property*. Mass is also an extensive property. So are the number of molecules, energy, and entropy.

Thermodynamics of phase transition

Temperature-energy curve. Ice and water only change volume slightly. We often neglect the change in volume when ice and water change temperature and when ice melts into water. In describing ice and water, we often characterize the states of H_2O using a single independent variable, e.g., temperature.

Molecules in ice and water interact and move. The energy associated with the molecular interaction and motion is called *thermal energy*. Define the *specific thermal energy* u as the thermal energy per unit mass.

The art of measuring temperature is called thermometry. The art of measuring thermal energy is called calorimetry. Given a thermal system, such as a collection of water molecules, we can measure temperature as a function of energy, T(u). Plot temperature T as the vertical axis, and the specific energy u as the horizontal axis. In the plane, the function T(u) is a curve, consisting of three segments.



H_2O molecules melt at zero Celsius,or 273.15 Kelvin. At the melting temperature, the function T(u) is a horizontal, straight segment. Each point on the horizontal segment represents a mixture of ice and water. The left end of the segment represents the state of ice that just begins to melt, and has specific energy u_{ice}. The right end of the segment represents the state of water that just begins to freeze, and has specific energy u_{water}. The difference in energy, $u_{water} - u_{ice}$, defines the *energy of melting*. The energy of melting for H_2O molecules is 334 kJ/kg.

Below and above the melting temperature, the function T(u) is curved, but is commonly approximated by two straight segments with fixed slopes.

Below the melting temperature, the slope of the function T(u) defines the *specific thermal capacity of ice*, and is 2.06 kJ/kg-K.

Above the melting temperature, the slope of the function T(u) defines the *specific thermal capacity of water*, and is 4.18 kJ/kg-K.

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

Exercise. In 1845, James Joule presented experimental data to suggest that the water at the bottom of a waterfall should be warmer than at the top. His suggestion met great resistance from other scientists, partly because the change in temperature is small. For the Niagara Falls, about 50 m high, how much is the change in temperature?

Exercise. I hold a bottle of water in my hand. When I drop the bottle, it hits the ground. Assume that all the potential energy when the bottle is in my hand converts to thermal energy in the water molecules when the bottle is on the ground. How many times do I need to drop the bottle when the temperature of water rises by one centigrade?

Exercise. YouTube phase-change coffee mug. Explain the science of the phase-change coffee mug. What substance undergoes phase change? What are the phases? What characteristics of the phase change are significant to the coffee mug.

A model of phase transition. A large number of a species of molecule, such as H_2O, constitute a system. Consider a system of a fixed number of molecules, N. In a condensed state, such as ice or water, the molecules touch one another. As the system changes from one state to another, the change in volume is small, and is neglected in this model. Thus, we model the molecules as a thermal system, with a single independent variable, energy U.

Energy is an extensive property. Let u = U/N be the energy per molecule. When the energy is fixed, the thermal system becomes an isolated system, flipping among a set of quantum states. The logarithm of the number of quantum states defines entropy S. Entropy is an extensive property. Let s = S/N be the average entropy per molecule.

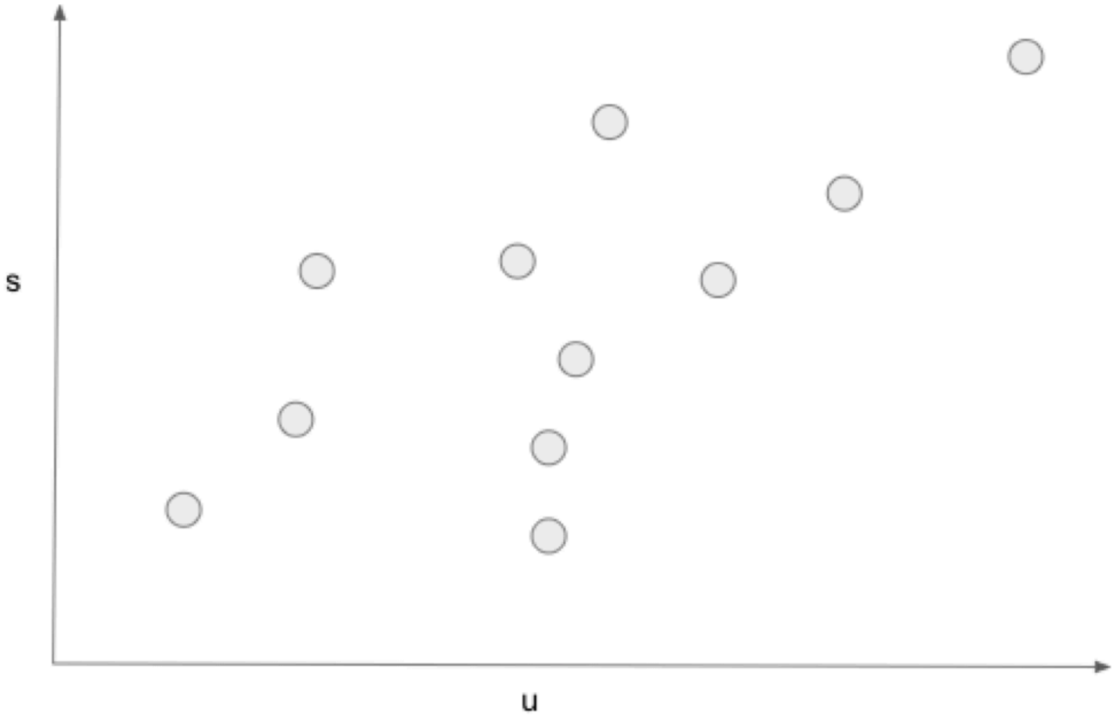
Isolated for a long time, the system is in equilibrium. The state of equilibrium can be a homogenous state, such as a state of ice, or a state of water. The state of equilibrium can also be a mixture of homogeneous states, such as a mixture of ice and water.

Even in a homogenous state, the isolated system rapidly and ceaselessly flips among its quantum states. Individual molecules move and interact. On average, the system is homogeneous over any length scale larger than the size of individual molecules.

In a mixture of homogeneous states, each piece of homogenous state is much larger than the size of individual molecules. The size, shape, and location of each piece of homogenous state do not concern this model. The model neglects molecules at the interfaces between pieces of homogenous states. The molecules at the interfaces have their own average energy and entropy, which are different from those in 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.

Energy-entropy plane. Draw a plane with u as the horizontal axis and s as the vertical axis. In the plane, a homogeneous state is a point, and so is a mixture of two or more homogeneous states. Some of the homogeneous states and some of the mixtures are states of equilibrium. Given a set of homogenous states in the (u,s) plane, the model locates (1) all mixtures, and (2) all states of equilibrium.

For example, mark a finite set of homogeneous states in the energy-entropy plane. Let us see how the model locates mixtures and thermodynamic states of equilibrium.



Mixtures of two homogeneous states. Let (u_A,s_A) and (u_B,s_B) be two homogeneous states. A mixture of the two homogeneous states has a total of N molecules, of which N_A molecules are in homogeneous state A, and N_B molecules are in homogeneous 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 fixed:

$N_A + N_B = N$.

Dividing the above equation by N, we obtain that

$y_A + y_B = 1$.

Let u be the average energy per molecule in the mixture, u = U/N. Energy is an extensive variable, so that the energy of the mixture is the sum of the energies of the two homogeneous states:

$Nu = N_A u_A + N_B u_B$.

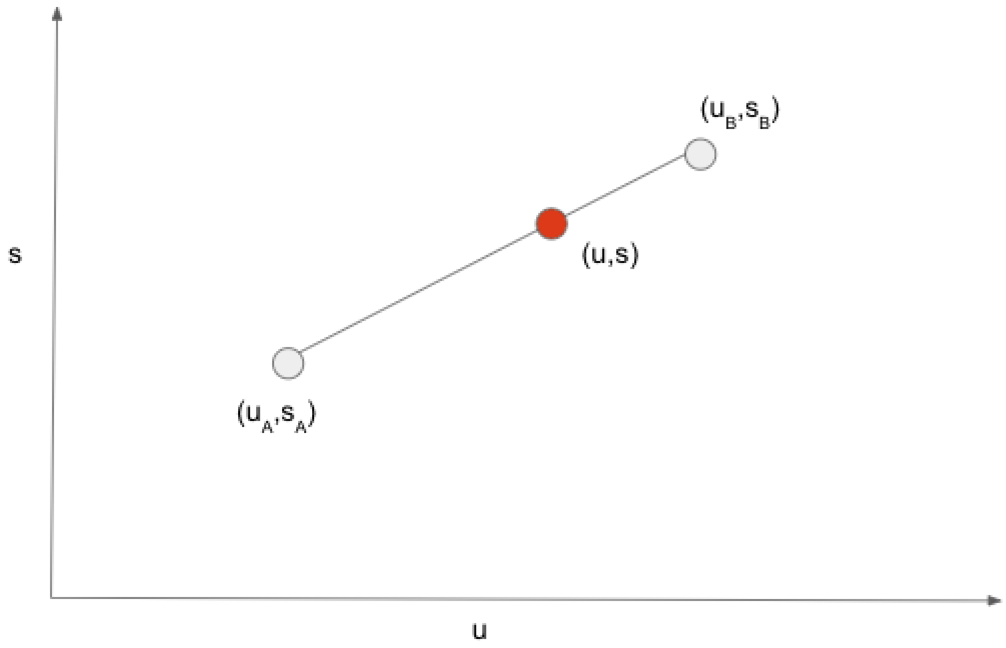
Dividing the above equation by N, we obtain that

$u = y_A u_A + y_B u_B$.

This equation is a *rule of mixture*. The same is true for entropy:

$s = y_A s_A + y_B s_B$.

Here s is the average entropy per molecule in the mixture.



The rules of mixture can be pictured on the (u,s) plane. The two homogeneous states are two points, (u_A,s_A) and (u_B,s_B), whereas the mixture is another point, (u,s). When $y_B = 0$, all molecules are in homogeneous state A, and the point (u,s) coincides with the point (u_A,s_A). When $y_B = 1$, all molecules are in homogeneous state B, and the point (u,s) coincides with the point (u_A,s_A). When $0 < y_B < 1$, some molecules are in homogeneous state A, and others in homogeneous state B, so that the point (u,s) is somewhere on the line segment joining the two points, (u_A,s_A) and (u_B,s_B).

Mixtures of three homogeneous states. A mixture 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 nonnegative numbers. The mixture has a fixed number of molecules:

$N_A + N_B + N_C = N$

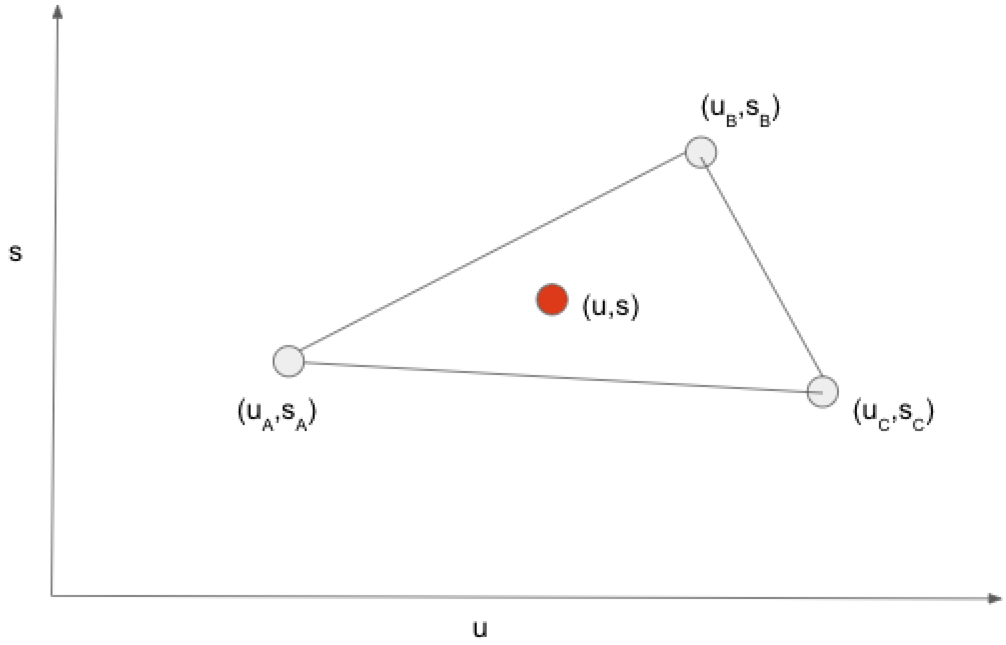
Divide this equation by N, and we obtain that

$y_A + y_B + y_C = 1$

The average energy and energy per molecule in the mixture are given by

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

The three homogeneous states, (u_A,s_A), (u_B,s_B), and (u_C,s_C), are vertices of a triangle. The mixture, (u,s), is a point in the triangle.

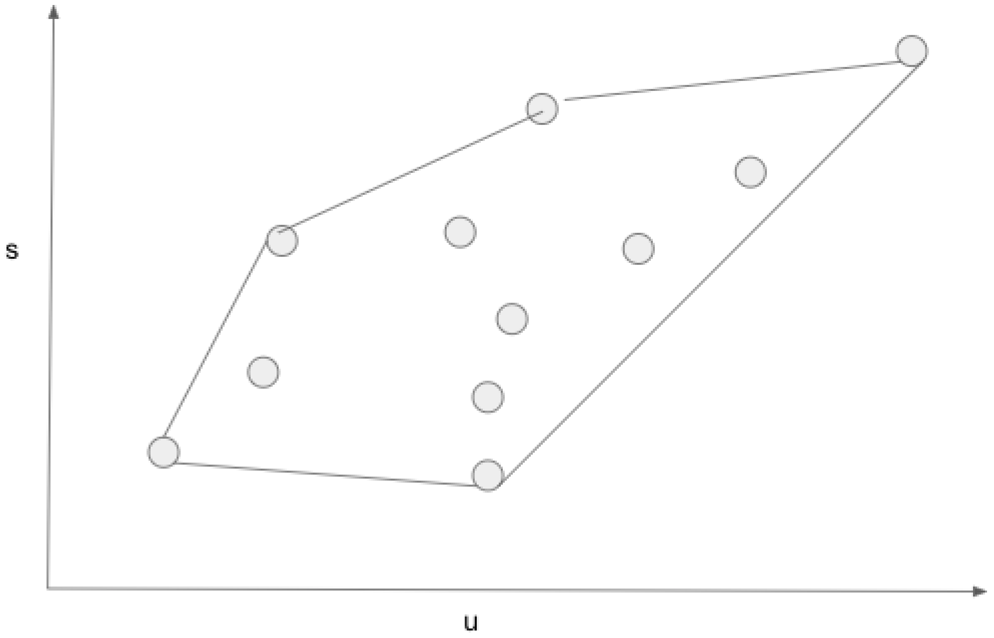


Mixtures of any number of homogenous states. The (u,v) plane is a *vector space*, and each point in the plane is a *vector*. A homogeneous state is a vector.

Because energy and entropy are extensive properties, a mixture of homogeneous states is a linear combination of the homogeneous states, with each coefficient being the fraction of molecules in a homogeneous state. All coefficients are nonnegative and sum to 1. Such a linear combination is called a convex combination in linear algebra, and a rule of mixture in thermodynamics.

Given a set of homogeneous states in the (u,s) plane, all mixtures form the convex hull of the homogeneous states.

For example, given a finite set of homogeneous states, the convex hull is a polygon.

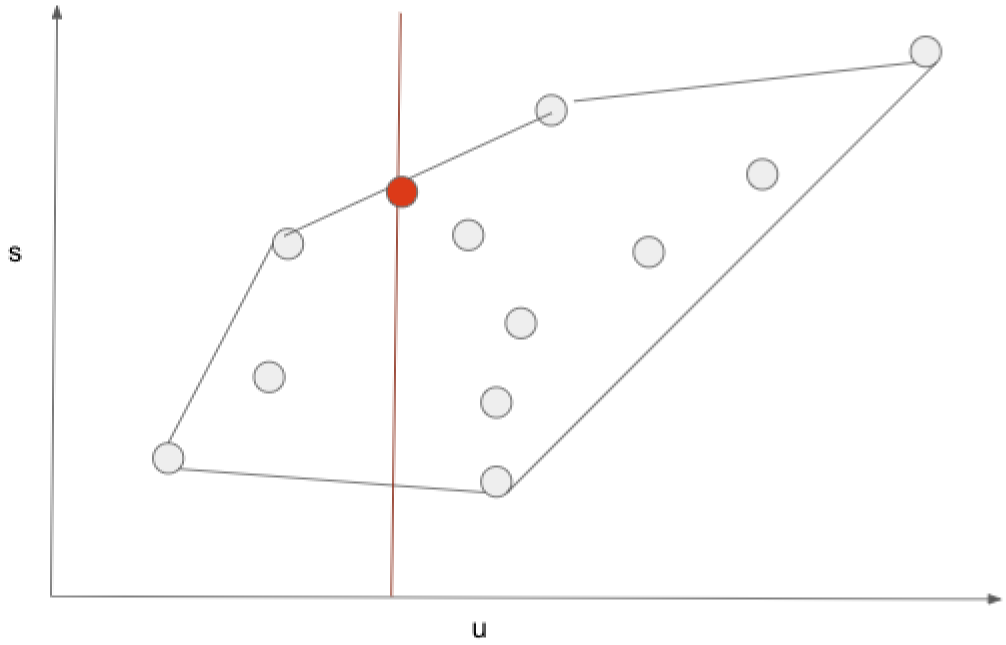


BAT on mixtures. So far energy and entropy play similar roles: they are extensive properties. Given a set of homogeneous states in the (u,s) plane, all mixtures form the convex hull of the homogeneous states. Not all homogeneous states and mixtures are thermodynamic states of equilibrium. To locate all thermodynamic states of equilibrium, we next apply the basic algorithm of thermodynamics.

The system of molecules is an isolated system if it has a fixed amount of energy, which is represented by a vertical line on the (u,s) plane.

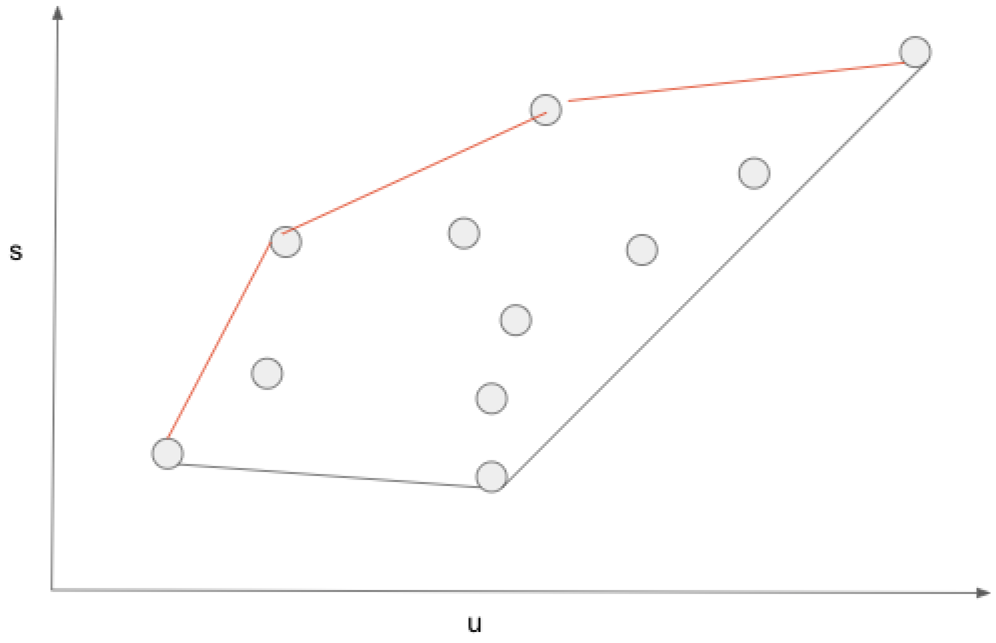
The isolated system has an enormous number of internal variables, including the number of homogeneous states, the location of each homogeneous state in the energy-entropy plane, and the number fraction of molecules allocated to each homogeneous state. The entropy of the mixture is the subset entropy, which is a function of all the internal variables.

For a fixed energy u, all mixtures fall in the convex hull, on the vertical line of the fixed energy. Each point on the line corresponds to some mixtures, which have a value of subset entropy, corresponding to the height of the point on the vertical line. The vertical line intersects the upper boundary of the convex hull at one point. This point maximizes the subset entropy, and is the state of equilibrium.



As the energy u changes, the upper boundary of the convex hull constitutes all thermodynamic states of equilibrium.

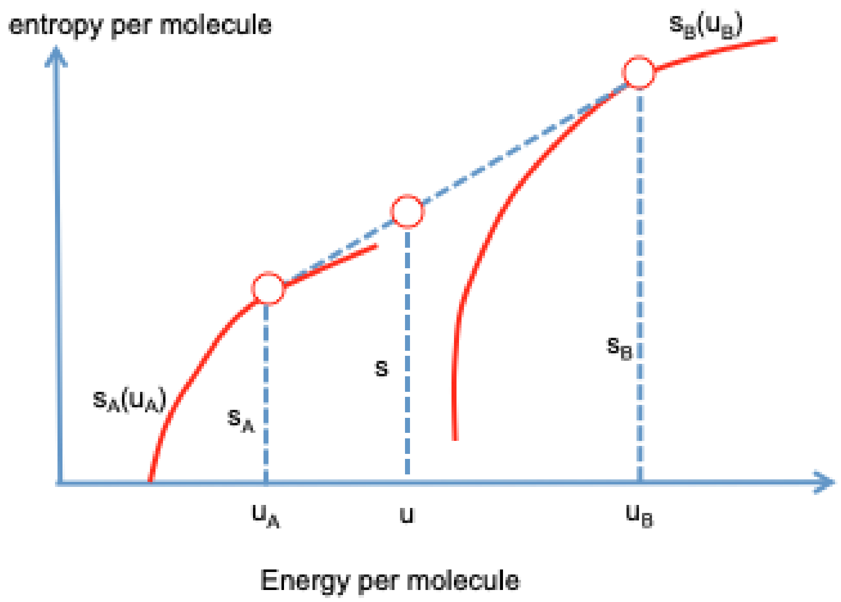
For a finite set of homogeneous states, the upper boundary of the convex hull consists of straight segments. Each segment connects two homogeneous states. Three homogeneous states on the same straight line are atypical, for a small modification of the system of molecules will make the three homogeneous states not on the same straight line. The system of molecules can equilibrate in either a homogeneous state or a mixture of two homogeneous states, but not in a mixture of three or more homogeneous states.



A system of molecules that forms two phases. We now apply the model to a system of one species of molecules that forms two phases. For example, a collection of H_2O molecules forms two condensed phases, ice and water. On the energy-entropy plane, all homogeneous states of each phase form a smooth curve.

Sketched on the u-s plane are two smooth curves, one for ice, $s_A(u_A)$, and the other for water, $s_B(u_B)$. The energy-entropy curve for each phase is taken to be convex. This decision is made on the basis of empirical observations: two homogeneous states of water do not equilibrate in a mixture, and two homogeneous states of ice do not equilibrate in a mixture. Incidentally, H_2O molecules form multiple solid phases, but each solid phase is represented by a distinct energy-entropy curve. This model is concerned with a single solid phase.

Wikipedia Phases of ice



Each of the two convex curves represents a phase of many homogeneous states. But the two covers of homogeneous states together do not contain the upper boundary of their convex hull, which is formed by rolling tangent lines on the curves of the two phases. The upper boundary of the convex hull consists of three pieces: part of the curve for phase A, part of the curve for phase B, and a straight-line segment tangent to both curves. The straight-line segment is called a *tie line*.

Every point on the upper boundary of the convex hull represents a state of thermodynamic equilibrium.

Equilibrium of a single homogeneous state. On the upper boundary of the convex hull, a point belonging to the curve of one phase corresponds to a homogeneous state. All such states are called thermodynamic states of *absolute stability*.

The tangent line can roll on the curve of the phase to change the slope of the tangent line. Thus, the curve of absolute stability has one degree of freedom.

Recall that temperature T is defined by $1/T = ds(u)/du$. The slope of each tangent line corresponds to a temperature.

Equilibrium of two homogeneous states. The ends of the tie line correspond to two homogenous states, called the *limits of absolute stability*. The two homogeneous states have the same temperature. A point on the tie line corresponds to a mixture of the two homogeneous states. The tie line cannot roll, is fixed in the energy-entropy plane, and has no degree of freedom.

Let the tie line touch curve $s_A(u_A)$ at point (u_A,s_A), and touch 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 tie line defines the melting temperature T_m. 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 curves for the two phases, $s_A(u_A)$ 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).

Equilibrium of three homogenous states does not exist in a thermal system. The three phases of H_2O—ice, water, and steam—can equilibrate. The three-phase equilibrium cannot be represented in the thermal-system model, but can be represented in the closed-system model, as we will say later.

In the thermal-system model, each phase is represented on the energy-entropy plane by a curve. Three phases are represented by three curves. A straight line rolling over the three curves can at once touch one curve or two curves, but not three curves. The exception is that all three curves are exactly positioned to have a common tangent line. The chance for such an exception is negligible.

Temperature-entropy curve. 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, for ice near the melting temperature, we have

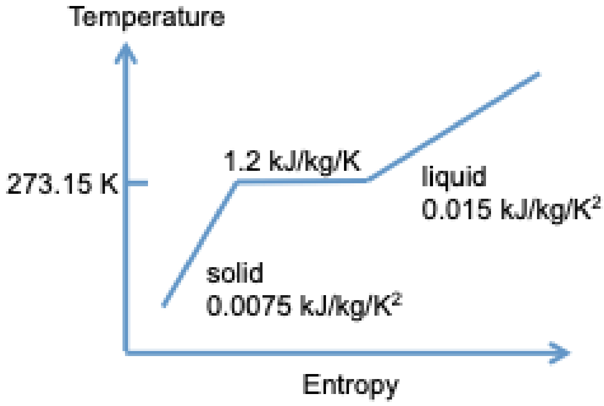
$$ds(T)/dT = c_s/T_m = (2.06 \text{ kJ/kg/K})/(273.15 \text{ K}) = 0.0075 \text{ kJ/kg/K}^2$$

At the melting temperature, the specific thermal capacity for liquid water is c_f = 4.18 kJ/kg/K. Thus, for liquid water near the melting temperature, we have

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

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

$$s_f - s_s = (334 \text{ kJ/kg})/(273.15 \text{ K}) =1.2 \text{ kJ/kg/K}$$



Exercise. Can the model of the thermal system describe 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 energy of phase transition 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. A convex function s(u) means that u(T) is an increasing function, and that thermal capacity is positive. In fact, a thermal system may have a nonconvex primitive curve s(u). We can form a tangent line touching two points on the curve s(u). The two points correspond to two homogeneous states in equilibrium.

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

Metastability. A curve for a phase 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 curve lies below the upper boundary of the convex hull, the part of the curve is beyond the limit of absolute stability. Each point of this part of the curve is called a *metastable state*. A metastable state is stable under continuous changes of state, but is unstable under discontinuous changes of state.

Exercise. Watch a video on supercooled water. What is supercooled water? How does it form? Indicate supercooled water in a sketch of the primitive curves on the energy-entropy plane. Indicate supercooled water in a sketch of the energy-temperature plane.

Tweet Thermodynamics of phase transition

Thermal reservoir

A *thermal reservoir* is a thermal system that stays at a fixed temperature when transferring energy with the surroundings.

A thermal system is characterized by entropy as a function of energy, $S(U)$. Define temperature T by

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

Let this thermal system be a thermal reservoir of a fixed temperature T, so that the function $S(U)$ is a straight line on the energy-entropy plane. On changing its energy from U to $U + Q$, the thermal reservoir changes its entropy by

$S(U + Q) - S(U) = Q/T$

A thermal reservoir is characterized by a fixed temperature T. On gaining energy Q from the surroundings, the thermal reservoir increases its entropy by Q/T.

A thermal reservoir is called a *thermal sink* on gaining energy, and a *thermal source* on losing energy.

Some authors say that the surroundings *transfer* entropy Q/T to the thermal reservoir. Do not let this language distract you from what happens. A thermal reservoir is a family of isolated systems. The thermal reservoir isolated at energy $U + Q$ has more quantum states than the thermal reservoir isolated at energy U by a factor of $\exp(Q/T)$.

Recall that thermal capacity C is defined by $dT(U)/dU = 1/C$. A thermal reservoir has infinite thermal capacity.

A thermal reservoir is a reservoir of energy, and is analogous to a reservoir of water. The level of a water reservoir remains nearly unchanged when a cup of water is taken out from the water reservoir. The amount of water in a water reservoir is analogous to the amount of energy in a thermal reservoir. The level of a water reservoir is analogous to the temperature of a thermal reservoir.

A thermal reservoir is approximately realized by using a large tank of liquid. When the liquid loses or gains a small amount of energy, the temperature of the liquid is nearly unchanged.

A mixture of ice and water is also approximately a thermal reservoir. At the atmospheric pressure, the temperature of the ice-water mixture is fixed at the melting point, zero Celsius. Upon gaining or losing energy, the mixture changes the ratio of ice and water, but keeps the temperature fixed.

A thermal reservoir commonly models the environment in which we do experiments. For example, when stretched, a material transfers thermal energy with the environment. When the stretch is slow enough, the material keeps the same temperature as the environment. We call the environment a thermal environment.

Thermal reservoirs also enter many models. Whenever you fix the temperature of something, do not forget the thermal reservoir that does the fixing. In speaking of a system in thermal contact with a thermal reservoir, to focus attention on the system, we often do not mention the thermal reservoir, and simply say that the system is held at a fixed temperature. This verbal change does not eliminate the need for the thermal reservoir. It is the thermal reservoir that holds the system at the fixed temperature by transferring energy between the system and the thermal reservoir. The phrase “a system held at a temperature” is synonymous to “a system in thermal equilibrium with a thermal reservoir of a fixed temperature”.

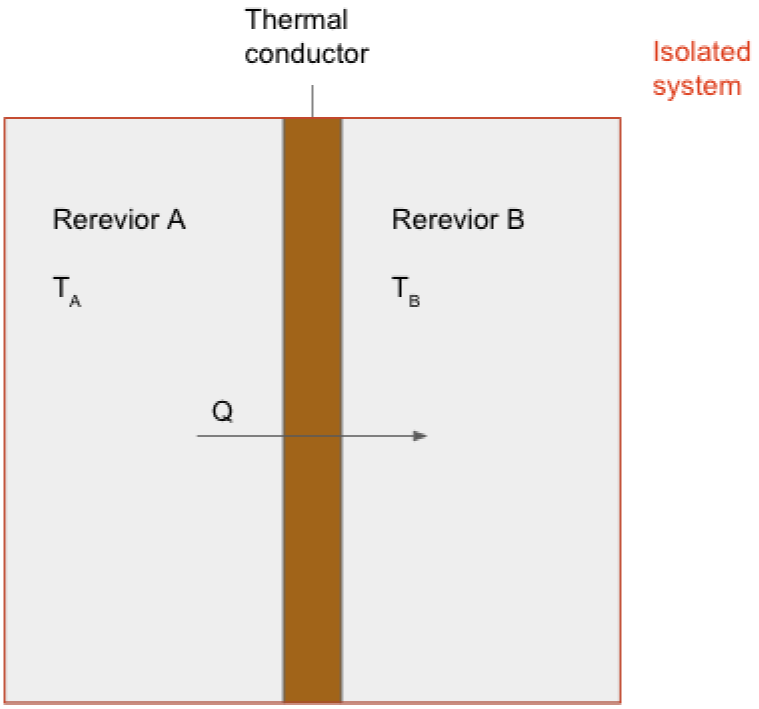
Twitter Thermal reservoir

Thermal conduction

We now analyze thermal conduction by applying the Basic Algorithm of Thermodynamics.

Step 1. Construct an isolated system with an internal variable. In this model, the isolated system consists of three parts: thermal reservoir A, thermal reservoir B, and a thermal conductor between them. The two reservoirs have fixed temperatures, T_A and T_B. Without loss of generality, assume that $T_A > T_B$, so that energy transfers from reservoir A, through the conductor, to reservoir B.

To identify an internal variable, assume that the energy flows through the conductor *slowly*, such that the temperature is homogeneous in reservoir A and in reservoir B. Further assume that the energy flows through the conductor *steadily*, such that the conductor itself changes neither energy nor entropy. The isolated system conserves energy, so that the energy lost by reservoir A is the same as the energy gained by reservoir B. Denote this transfer of energy by Q, which is the internal variable.



Step 2. Find the subset entropy of the isolated system as a function of internal variables. On losing energy Q, reservoir A lowers entropy by Q/T_A. On gaining energy Q, reservoir B raises entropy by Q/T_B. The subset entropy of the isolated system sums over its three parts. Recall that the conductor does not change its entropy. Thus, on transferring energy Q, the isolated system changes its subset entropy by

$Q/T_B - Q/T_A$

Step 3. Maximize the subset entropy to reach equilibrium. The subset entropy of the isolated system is maximized when the subset entropy stops changing. Because $T_A > T_B$, the subset entropy stops changing when energy stops transferring, $Q = 0$. Consequently, the two reservoirs of different temperature can only reach constrained equilibrium when the conductor approaches an insulator.

Step 4. Increase the subset entropy to undergo an irreversible process. The increase of the subset entropy requires that

$Q/T_B - Q/T_A > 0$

Because $T_A > T_B$, this inequality means that $Q > 0$, which means that energy transfers from reservoir A to reservoir B. On transferring energy Q, the isolated system generates entropy $Q/T_B - Q/T_A$. An isolated system does not conserve entropy, but generates entropy. The phrase “entropy transfer” is confusing, and should be eliminated.

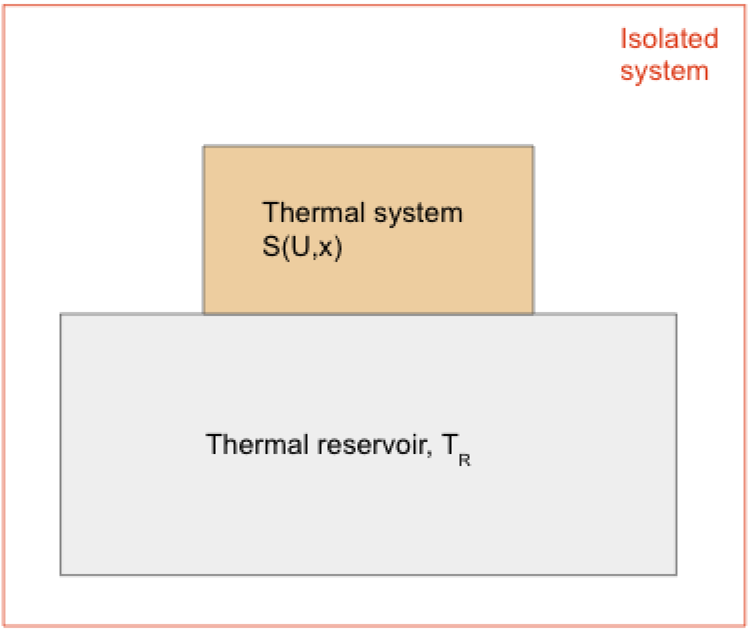
Free energy

A thermal system in thermal contact with a thermal reservoir. A half bottle of water is a thermal system, and a classroom is a thermal reservoir. The H_2O molecules in the bottle and the air molecules in the classroom are in thermal contact, through the vibration of molecules of the wall of the bottle. The thermal system and the thermal reservoir transfer energy, but nothing else.

In the bottle, H_2O molecules separate into two phases: steam and water. The number of H_2O molecules in the steam, x, is an internal variable of the thermal system.

This experiment illustrates a large class of phenomena. A thermal system has two independent variables, energy U and internal variable x. When U and x are fixed, the thermal system becomes an isolated system, flipping among a certain number of quantum states. Denote this number by $\Omega(U,x)$, and define $S(U,x) = \log \Omega(U,x)$. As U and x vary, the function $S(U,x)$ characterizes the thermal system as a family of isolated systems.

When the thermal system transfers energy with a thermal reservoir of a fixed temperature T_R, what happens to the internal variable x? We next analyze this class of phenomena using the Basic Algorithm of Thermodynamics.



Step 1. Construct an isolated system with two independent internal variables. A thermal system and a thermal reservoir together constitute an isolated system. The energy of the isolated system is a sum of the energy of its two parts:

$U_{iso} = U + U_R$

where U_{iso} is the energy of the isolated system, U the energy of the thermal system, and U_R the energy of the thermal reservoir.

An isolated system conserves energy. As the thermal system and the thermal reservoir transfer energy, the energy of the isolated system U_{iso} is constant. We choose the energy U of the thermal system as an independent variable, and the energy U_R of the thermal reservoir as a dependent variable.

As illustrated by a half bottle of water, the thermal system itself has an internal variable, x. Thus, the isolated system has two independent internal variables, U and x.

Step 2. Find the subset entropy of the isolated system as a function of internal variables. The thermal system is characterized by a function of two variables, S(U,x). The thermal reservoir is characterized by a function of one variable, $S_R(U_R)$. The subset entropy of the isolated system is the sum of the entropies of its two parts:

$S_{iso} = S(U,x) + S_R(U_R)$

The thermal reservoir has a fixed temperature T_R, so that $S_R(U_R)$ is a linear function. By the conservation of energy, $U_{iso} - U_R = U$, so that $S_R(U_R) = S_R(U_{iso}) - U/T_R$. On losing energy U to the thermal system, the thermal reservoir lowers its entropy by U/T_R. Because U_{iso} is constant, $S_R(U_{iso})$ is also constant. The subset entropy of the isolated system, S_{iso}, is a function of two independent internal variables, U and x. A change in the subset entropy $S_{iso}(U,x)$ of the isolated system is the same as a change in the function

$S(U,x) - U/T_R$

The temperature of the thermal reservoir T_R is fixed, but both the internal energy U and the internal variable x of the thermal system are variables.

Step 3. Maximize the subset entropy to equilibrate. For the isolated system to reach equilibrium, change U and x to maximize the function $S(U,x) - U/T_R$. In this function, the quantities S, U and x are associated with the thermal system, but T_R is associated with the thermal reservoir. The thermal system and the thermal reservoir are in thermal contact, but may not be in thermal equilibrium. They may have different temperatures, and are undergoing an irreversible process to equalize temperatures.

Thermal equilibrium is a type of *partial equilibrium*, in which the isolated system reaches equilibrium with respect to change in U, but has not reached equilibrium with respect to change in x. For the isolated system to reach thermal equilibrium, fix x and change U to maximize the subset entropy of the isolated system, $S(U,x) - U/T_R$.

Sketch the function $S(U,x) - U/T_R$ on the energy-entropy plane for a fixed x. The term $S(U,x)$ is the entropy of the thermal system, and is a curve in the energy-entropy plane. The term U/T_R comes from the lowering of the entropy of the thermal reservoir, and is a straight line of slope $1/T_R$ in the energy-entropy plane. A given value of U corresponds to a line normal to the energy axis. The function $S(U,x) - U/T_R$ is the distance between the curve and straight line. The function maximizes when a tangent line of $S(U,x)$ parallels the line U/T_R.

For a fixed x, so long as the function $S(U,x)$ is convex, the function $S(U,x) - U/T_R$ has a unique maximum when

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

Recall the definition of the temperature T of the thermal system: $1/T = \partial S(U,x)/\partial U$. The above equation recovers a familiar condition: a thermal system and a thermal reservoir in thermal equilibrium have the same temperature, $T = T_R$. We drop the subscript in T_R, and denote the temperature of both the thermal system and thermal reservoir by T.

Given a thermal system of characteristic function $S(U,x)$, the above condition of thermal equilibrium solves the energy of the system as a function $U(T,x)$. Replacing U by $U(T,x)$, we can also convert the function $S(U,x)$ to the function $S(T,x)$.

A process taking place at a constant temperature is called an *isothermal process.* In an isothermal process, the subset entropy of the isolated system is

$$J = S(T,x) - T^{-1}U(T,x)$$

All quantities in J are now associated with the thermal system. To complete step 3 of BAT, we hold T constant and vary x to maximize $J(T,x)$ to reach equilibrium.

Step 4. Increase the subset entropy to change irreversibly. For the isolated system to undergo an irreversible process, change x to increase the function $J(T,x)$.

In the idealization that the isolated system undergoes a reversible process, change x to keep the function $J(T,x)$ constant.

Free energy. In an isothermal process, the temperature T is a positive constant, so that maximizing $S(T,x) - T^{-1}U(T,x)$ is the same as minimizing $U(T,x) - TS(T,x)$.

Call the quantity $F = U - TS$ *free energy*. Maximizing the function $J(T,x)$ is equivalent to minimizing the function $F(T,x)$. The latter is more often used in practice. Thus, free energy is entropy in disguise.

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

1. Construct a thermal system with an internal variable x. The thermal system is in contact with a thermal reservoir of temperature T.
2. Find the free energy as a function of the temperature and the internal variable, $F(T,x)$.
3. Change x to minimize $F(T,x)$ to equilibrate.
4. Change x to decrease $F(T,x)$ to undergo an irreversible process.

To conserve energy, or to minimize energy? An isolated system conserves energy and changes the internal variable x to maximize subset entropy $S(x)$.

A thermal system in contact with a thermal reservoir of temperature T changes the internal variable x to minimize free energy $F(T,x)$.

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

$$dF = - SdT.$$

Thus,

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

Recall the definition of the thermal capacity, $\partial T(U,x)/\partial U = 1/C$, and the definition of temperature, $\partial S(U,x)/\partial U = 1/T$. We obtain that

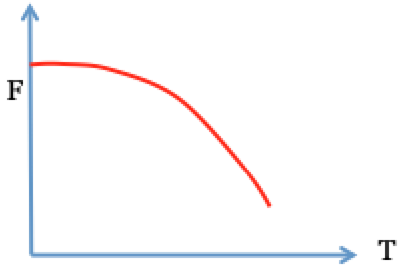
$$\partial S(T,x)/\partial T = C/T,$$
$$\partial^2 F(T,x)/\partial T^2 = -C/T.$$

Features of the free energy function. The free energy F is an extensive property, but temperature T is an intensive property. The F-T plane has different properties from the S-U plane.

For a fixed x, $F(T,x)$ is a function of T. We note the following features of this function as a curve on the F-T plane.

1. F is relative, so that the curve can translate up and down without affecting the behavior of the system.
2. T is absolute. The curve cannot translate left and right.
3. The function $F(T,x)$ is smooth. (We will look at exceptions later.)

4. S approaches zero as T approaches zero, so that the curve starts with a horizontal tangent.
5. Because S is positive, F decreases as T increases.
6. Because S increases as T increases, F is a convex function of T.



Twitter Free energy

Melting analyzed using free energy

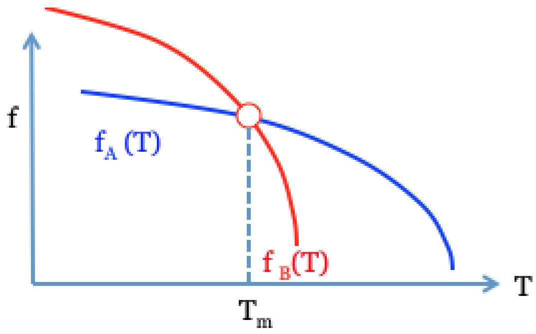
We now use the melting of a pure substance to illustrate the algorithm of thermodynamics for an isothermal process.

A system in an isothermal process. We characterize a system undergoing an isothermal process with a fixed temperature and an internal variable. Let $f_A(T)$ be the free energy per molecule of the solid, and $f_B(T)$ be the free energy per molecule of the liquid. The free energy per molecule of a solid-liquid mixture is

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

Recall that $y_A + y_B = 1$. The fraction of molecules in the solid, y_A, is the independent internal variable. In minimizing the free energy of the mixture, $f(T,y_A)$, T is constant and y_A is variable.

Equilibrium state. 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 free energy of the mixture minimizes if all molecules are in the solid phase. When $T > T_m$, the free energy of the mixture minimizes if all molecules are in the liquid phase.



Recall the definition of the free energy, $f = u - Ts$. The condition of equilibrium of the solid and liquid, $f_A(T) = f_B(T)$, gives that

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

This expression recovers what we have obtained before using the function s(u). The free energy gives us a slick derivation of the equilibrium condition, but pushes us a step away from the leading role—entropy.

Irreversible process. To illustrate an irreversible process, consider a supercooled liquid held at a temperature T below the melting point. The supercooled liquid is in thermal contact with a thermal environment of temperature T.

We model the substance as a system of a fixed temperature T and an internal variable y_A. In the initial state, the substance is a supercooled liquid, $y_A = 1$. The initial state does not minimize the free energy $f(T,y_A)$, and is not an equilibrium state. The substance minimizes the free energy at a fixed T by changing to solid, $y_A = 0$. The transition from the supercooled liquid to the solid changes the value of the internal variable y_A to reduce the free energy $f(T,y_A)$, and is an irreversible process.

Ferroelasticity

A phase transition that breaks a symmetry. A crystal is in a *paraelastic* phase above a critical temperature T_c, and is in a *ferroelastic* phase below the critical temperature. The phase transition breaks a symmetry and causes a remnant deformation. For example, a shape memory alloy is a ferroelastic material. Ferroelectric materials are usually ferroelastic. Wiki ferroelasticity.

Here we use ferroelasticity to illustrate the use of free energy. Consider a simplified example. In the paraelastic phase, the crystal has a rectangle lattice. In the ferroelastic phase, the crystal has a parallelepiped lattice. The phase transition breaks symmetry. The parallelepiped can be in two orientations of the same shape, but with a mirror symmetry. The two parallelepipeds are called two *variants* in the ferroelastic phase.

A symmetry-preserving free energy. Model the crystal by the free energy as a function of two variables F(T,D). Here T is the temperature, and D is the deformation from the rectangular symmetry (e.g., shear). This model regards D as an internal variable to minimize the free energy function.

We use two leading terms in the Taylor expansion in D:

$F(T,D) = A(T - T_c)D^2 + BD^4$,

where A and B are positive constants to fit experimental data. In the paraelectric phase, $D = 0$ when the crystal is subject to no stress. In the ferroelectric phase, $D \neq 0$ even when the crystal is not subject to any stress. This nonzero value of D is called the *remnant deformation*. To model two variants with a mirror symmetry, the free energy must be an even function of D. To keep the model simple, the function only contains two powers of D.

Paraelastic phase. At a fixed temperature above the critical temperature, $T > T_c$, the free energy function minimizes at $D = 0$, and the crystal is rectangular.

Ferroelastic phase. At a fixed temperature below the critical temperature, $T < T_c$, the free energy function has a local maximum at $D = 0$, and two minima at D_s and $-D_s$. The crystal has a parallelepiped symmetry, and D_s is the spontaneous distortion.

To minimize the free energy function, set $\partial F(T,D)/\partial D = 0$, giving

$2A(T - T_c)D + 4BD^3 = 0$.

The equation has three solutions. The solution at $D = 0$ corresponds to the local maximum of the free energy function. The other two solutions correspond to

$D = (A(T_c - T)/2B)^{½}$,
$D = -(A(T_c - T)/2B)^{½}$,

The two solutions correspond to the two minima of the free energy function.

Wiki Landau theory

Exercise. The model above does not include applied load. Add applied load to the free energy function. Identify a thermal system, with an internal variable, in a thermal environment.

Exercise. Design experiments to measure constants A, B, and T_c.

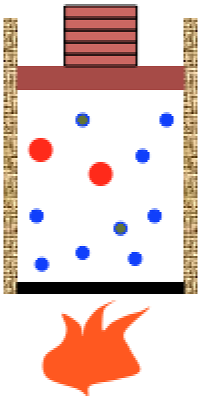
Closed system

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

Act two: **The union of entropy and volume produces a child—pressure**. The child coincides with the pressure in mechanics, defined as force per unit area.

A closed system is a family of isolated systems of two variables

A *closed system* is a system that does not transfer matter with the surroundings, but transfers energy and volume with the surroundings.



For example, a gas in a cylinder, sealed by a piston, is a closed system. The volume V of the gas changes when the piston slides relative to the cylinder. The energy U of the gas can be changed in many ways, for example, when a fire transfers energy to the gas, when an electrical current flows through a resistor placed in the cylinder, or when weights are added above the piston.

Thus, the closed system has two independent variables, U and V. When U and V are fixed by thermally insulating the cylinder and locking the position of the piston, a closed system becomes an isolated system, which flips among a set of quantum states. Denote the number of quantum states of this isolated system by $\Omega(U,V)$. Define $S(U,V) = \log \Omega(U,V)$. As U and V vary, the function $S(U,V)$ characterizes the closed system as a family of isolated systems.

Energy-volume-entropy space

Energy, volume, and entropy form a three-dimensional vector space without inner product. Drawing a three-dimensional vector space on a piece of paper is impossible, but the picture is so suggestive that we fake. We draw energy U and volume V as two horizontal axes. We draw entropy S as the vertical axis. Although being vertical is meaningless in a vector space without inner product, we fake.

Given a closed system, its characteristic function S(U,V) is a *surface* in the energy-volume-entropy space. Energy is relative, and the surface can be translated in the direction of the energy axis without affecting the closed system. Both entropy and volume are absolute, and the surface cannot be translated in the directions of the entropy and volume axes. The surface is typically smooth and convex upward. Entropy increases as energy and volume increase.

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

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

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 relate the two partial derivatives of the function to experimental measurements.

A closed system in a thermomechanical environment

Look again at a gas sealed in a cylinder by a piston. The closed system is in thermal contact with a thermal reservoir of a fixed temperature T_{env}. Placed above the piston is a weight, which exerts on the piston a fixed force AP_{env}, where A is the base area of the piston, and P_{env} is the applied force per area. When the piston is at height Z, the potential energy of the weight is $P_{env}AZ = P_{env}V$. The weight provides a mechanical environment for the closed system

For the closed system of the gas in the cylinder-piston setup, the thermal reservoir and the weight together provide a *thermomechanical environment*, of a fixed temperature T_{env} and a fixed pressure P_{env}. We now analyze this experiment using the BAT.

Step 1. Construct an isolated system with internal variables. Model this experiment by an isolated system of three parts: the gas in the cylinder, the weight above the piston, and the thermal reservoir.

The energy of the isolated system is the sum of the energies of its three parts:

$U_{iso} = U + P_{env}V + U_{env}$

where U the energy of the closed system, $P_{env}V$ the potential energy of the weight, and U_{env} the energy of the thermal environment.

The isolated system conserves energy, so that U_{iso} is constant. The isolated system has two independent internal variables: U and V. The energy U_{env} of the thermal reservoir is a dependent variable.

Step 2. Find subset entropy of the isolated system as a function of independent internal variables. The subset entropy of the isolated system is the sum of the entropies of its three parts:

$S_{iso} = S(U,V) + S_{weight} + S_{env}(U_{env})$

where S(U,V) is the entropy of the closed system, S_{weight} is the entropy of the weight, and $S_{env}(U_{env})$ is the entropy of the thermal reservoir. By the conservation of energy, $U_{iso} - U_{env} = U + P_{env}V$. Because T_{env} is constant, $S_{env}(U_{env})$ is a linear function, so that $S_{env}(U_{env}) = S_{env}(U_{iso}) - (U + P_{env}V)/T_{env}$.

Step 3. Maximize subset entropy to equilibrate. The entropy S_{weight} of the weight is constant. Because U_{iso} is constant, $S_{env}(U_{iso})$ is also constant. A change in the subset entropy of the isolated system is the same as a change in the following quantity:

$S(U,V) - (U + P_{env}V)/T_{env}$

In this function, U and V are independent variables of the closed system, but T_{env} and P_{env} are constants set by the thermomechanical environment.

Sketch the function $S(U,V) - (U + P_{env}V)/T_{env}$ in the energy-volume-entropy space. The term S(U,V) is the surface of the closed system. The term $(U + P_{env}V)/T_{env}$ comes from the entropy of the thermomechanical environment, and is a plane of slopes $1/T_{env}$ and P_{env}/T_{env}. A given pair of values of U and V corresponds to a line vertical to the energy-volume plane. The function $S(U,V) - (U + P_{env}V)/T_{env}$ is the distance between the curved surface and plane. The function maximizes when a tangent plane of S(U,V) parallels the plane $(U + P_{env}V)/T_{env}$.

This picture corresponds to a fact of calculus. Maximizing a function $S_{iso}(U,V)$ requires that

$\partial S(U,V)/\partial U = 1/T_{env}$
$\partial S(U,V)/\partial V = P_{env}/T_{env}$

The above equations of partial equilibrium recover the familiar conditions. In thermal equilibrium, the closed system and the thermal reservoir have the same temperature. In mechanical equilibrium, the pressure in the closed system equals the pressure applied by the weight.

When the isolated system is in thermal and mechanical equilibrium, we drop the subscripts in T_{env} and P_{env}, and write

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

Here T is the temperature for both the closed system and the thermal reservoir, and P is the pressure for both the closed system and the mechanical environment. The two equations interpret the two partial derivatives of the function S(U,V) of the closed system.

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

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

Write

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

This equation connects five properties of a closed system: U, V, S, T, and P.

Step 4. Increase subset entropy to undergo an irreversible process. Any actual process of an isolated system is irreversible. When the two internal variables change by dU and dV, the subset entropy of the isolated system, S_{iso}, must increase, so that

$dS - (dU + P_{env}dV)/T_{env} > 0$

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

$(1/T - 1/T_{env})dU + (P/T - P_{env}/T_{env})dV > 0$

This inequality holds for arbitrary combinations of dU and dV. When dV = 0, the inequality becomes

$(1/T - 1/T_{env})dU > 0$

When dU = 0, the inequality becomes

$(P/T - P_{env}/T_{env})dV > 0$

The irreversibility is caused by the transfer of energy between the gas in the cylinder and the thermal reservoir, and by the friction between the piston and the wall of the cylinder. We have studied the former, and let us examine the latter.

When the isolated system is in thermal equilibrium, the subset entropy does not change when the energy U changes, so that $T = T_{env}$. In the isothermal process, the entropy inequality becomes that

$(P - P_{env})dV > 0$

This inequality recovers our experience. When the piston moves up, $dV > 0$, the pressure in the cylinder must exceed the pressure due to the weight, $P > P_{env}$. The opposite is true when the piston moves down. The difference between the two pressures is caused by the friction between the piston and the wall of the cylinder.

As an idealization, a reversible process of an isolated system keeps the subset entropy constant. The conditions of reversibility are obtained by replacing all the entropy inequality with equality:

$(1/T - 1/T_{env})dU + (P/T - P_{env}/T_{env})dV = 0$

This inequality holds for independent changes dU and dV. Thus, a reversible process requires that

$T = T_{env}$
$P = P_{env}$

These conditions of reversibility coincide with those of equilibrium. In practice, a reversible process represents a limiting case when the difference in temperature and the difference in pressure between the closed system and the thermomechanical environment are negligibly small.

Tweet BAT on a closed system

Ideal gas

A collection of molecules is called an *ideal gas* if on average the distance between nearby molecules is much larger than the size of an individual molecule. The molecules move rapidly and ceaselessly. They fly, collide, and separate.

In an ideal gas, molecules are far apart on average. Each molecule can fly anywhere in the volume V as if all other molecules were absent. Consequently, the number of quantum states of each molecule is proportional to V, and the number of quantum states of N molecules is proportional to V^N. The proportional factor depends on N and energy U, but not on V:

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

where f(U,N) stands for a function of U and N, but not V.

The entropy of the gas, $S = \log \Omega$, is

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

Inserting this expression into $P/T = \partial S(U,V,N)/\partial V$, we obtain that

$PV = NT$.

This equation is the ideal gas law.

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 T is a function of U and N, and is independent of V. We invert the function T(U,N) to U(T,N). The internal energy is an extensive property, proportional to the number of molecules, N. Write

$U = Nu(T)$.

Here u(T) is the energy per molecule. Thus, an ideal gas is characterized by $PV = NT$, along with a function of a single variable, u(T). The functions u(T) for many species of molecules are listed in the ideal gas tables.

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, $dU = N(du(T)/dT)dT$ and $PV = NT$, we obtain that

$dS = N(1/T)(du(T)/dT + 1)dT - (N/P)dP$.

Entropy is a function of state. A state of the closed system is specified by two independent variables, T and P. In the above equation, the first term is due to the change in temperature, and the second term is due to the change in pressure. Integrating, we write

$S(T, P) = Ns(T,P_0) - N \log(P/P_0)$.

For many species of molecules, the ideal gas tables list the absolute entropy per molecule, $s(T,P_0)$, at a reference pressure P_0, as a function of T.

Twitter Ideal gas

Exercise. Assume a constant thermal capacity for an ideal gas. Calculate the function S(U,V,N).

Exercise. Use the ideal gas tables to find the energy, enthalpy, and entropy of O_2 at 1000 K, 1 MPa.

Exercise. Use the ideal gas tables to find the energy, enthalpy, and entropy of H_2O at 1000 K, 100 kPa.

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 basketbook is a bag of air. The ball bounces because 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*.

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, $P = NT/V$. Taking the partial derivative, we obtain the isothermal elastic modulus:

$B_{isothermal} = P$.

Adiabatic elastic modulus. Under the adiabatic condition, entropy is constant, $dS = 0$, so that

$N(1/T)(du(T)/dT + 1)dT - (N/P)dP = 0$

Assume that the thermal capacity per molecule, $c = du(T)/dT$, is constant. Let $b = (c + 1)/c$. Recall that $PV = NT$, so that $PdV + VdP = NdT$. The above equation becomes that

$(b/V)dV - (1/P)dP = 0$

Integrating, we obtain that PV^b = constant. Taking the partial derivative, we obtain the adiabatic elastic modulus:

$B_{adiabatic} = bP$.

A gas is stiffer in the adiabatic expansion than in the isothermal expansion.

Water and steam

When water boils into steam, the same number of H_2O molecules occupy a greatly expanded volume. We characterize states of H_2O using two independent variables, e.g., temperature and volume.

Specific volume. Both mass m and volume V are extensive properties. The ratio of mass over volume, m/V, defines the *mass density*. The ratio of volume over mass, V/m, defines the *specific volume*.

At room temperature and atmospheric pressure, the mass density of water is 1000 kg/m^3, and the specific volume of water is 0.001 m^3/kg.

We also report volume per molecule V/N, or volume per mole of molecules, V/n. Textbooks like to give V/m, V/N, and V/n distinct symbols. We will denote them all by v, and let the unit tell the difference. It is also common to use capital V to denote volume per molecule, per mole, or per unit mass. The context and unit will make it clear.

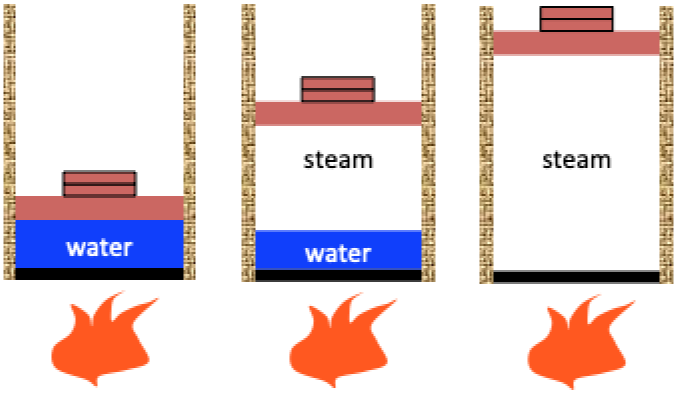
Exercise. At 25 Celsius, the pressure of H_2O molecules in steam is 3.170 kPa. A mole of water molecules has a mass of 18 g. Use the ideal gas law to estimate V/m, V/N, and V/n.

An experimental setup to study water and steam. Let us study water and steam using an experimental setup consisting of a cylinder, piston, weight, and fire.

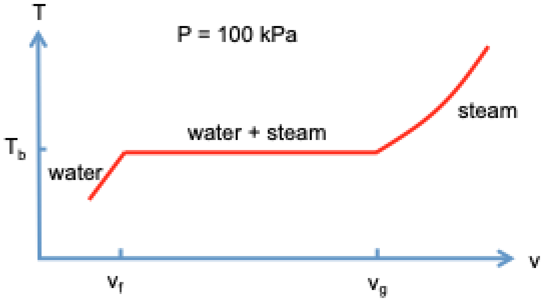
Fix the number of H_2O molecules in the cylinder, N, by sealing the piston. Assume that the cylinder only contains H_2O molecules. As the piston slides up, the cylinder increases its volume, V. Vary the pressure inside the cylinder, P, by the weight above the piston. Vary the temperature inside the cylinder, T, by the fire under the cylinder.

This experiment studies one pure substance (H_2O), two phases (water and steam), three properties (pressure P, temperature T, and volume V), and infinitely many states. Each state of a fixed number of H_2O molecules is specified by a value of temperature and a value of volume.

We do not need to study the effect of the number of molecules, N. Extensive properties, such as mass and volume, are linear in N. Given a species of molecules, the relation between N and mass is fixed: the molar mass of H_2O is 18 g/mol. Intensive properties, such as temperature and pressure, are independent of N. At fixed P and T, we will report volume per unit mass, i.e., the specific volume v.



Temperature-volume plane. A pure substance has three phases, which we name individually: solid, liquid, and gas. A pure substance has infinitely many states, too many to name individually. Instead, we label each state by values of properties.



For a fixed amount of a pure substance, a state is specified by the values of *two* independent properties, such as volume and temperature.

The values of the two independent properties serve as an address—a name—of the state.

Consider a plane with volume and temperature as axes. Volume is an extensive property, and temperature is an intensive property. Given an extensive property and an intensive property, it is a common practice to plot the extensive property as the horizontal axis, and the intensive property as the vertical axis. Each point in the T-v plane specifies a state of the pure substance.

Process. In the experimental setup, we now fix the pressure of the H_2O molecules in the cylinder by fixing the weight placed above the piston, and transfer energy to the H_2O molecules over the fire. We transfer energy slowly, so that H_2O molecules in the cylinder reach equilibrium every time we collect a data point, a pair of values of T and v.

The temperature increases, the volume expands, and the H_2O molecules change state. A sequence of states is called a *process*.

When the pressure is fixed, the sequence of states is called a *constant-pressure process*. All states under a fixed pressure correspond to a curve in the T-v plane. Each point on the curve is a state of H_2O molecules. Let us follow a curve for a fixed pressure of 100 kPa. The curve has three segments.

Single-phase states of water. Below the boiling temperature, T_b, the H_2O molecules form water. As the temperature increases, the volume of the water expands.

Two-phase states of water-steam mixtures. At the boiling temperature, some of the H_2O molecules form water, and other H_2O molecules form steam. In coexistence, water and steam are said to be *saturated*. Upon receiving more energy from the fire, more molecules evaporate from water and join the steam: the mixture fixes the temperature at the boiling temperature, and expands volume.

Single-phase states of steam. Above the boiling temperature, all the H_2O molecules form steam. As the temperature increases, the volume of the steam expands.

When we remove the fire, the temperature-volume curve reverses, and the internal energy inside the cylinder dissipates into the surroundings.

The liquid, gas, and liquid-gas are collectively called a *fluid*. A fluid flows because molecules in the fluid readily change neighbors.

Water expands on evaporation. In this example, the pressure is fixed at P = 100 kPa, and water boils at the temperature

T = 99.62 Celsius (about 100 Celsius).

At the boiling temperature, the specific volume of the saturated water is

$v_f = 0.001043\ m^3/kg$,

and the specific volume of the saturated steam is

$v_g = 1.69400\ m^3/kg$.

The subscript f stands for *flüssigkeit,* the German word for liquid.

The specific volumes of the two phases, water and steam, differ enormously. At 100 kPa, v_g/v_f is on the order of 10^3. Thus, in the saturated steam, molecules are, on average, at a distance of about the size of 10 molecules.

A fire causes the liquid to evaporate. The enormous expansion on evaporation converts fire to motion. This phenomenon is essential to steam engines and power plants.

Because of this enormous difference in specific volumes of water and steam, our sketch of the Tv diagram is unrealistic. To fit the diagram on a page, we can plot the specific volume in the logarithmic scale.

Exercise. In the liquid, molecules touch one another. Use the value of v_f to estimate the size of each H_2O molecule. What is the average distance between two H_2O molecules in the saturated steam at 100 C. How does the average distance between molecules compare with the size of a molecule?

Liquid-gas dome. We can change the pressure by changing the weight above the piston. We now fix the pressure at a new value, say P = 200 kPa. The state of H_2O molecules in the cylinder also changes as the fire transfers energy to the molecules. At this constant pressure, all the states of H_2O molecules fall on a curve in the temperature-volume plane. At P = 200 kPa, we find that water boils at the temperature

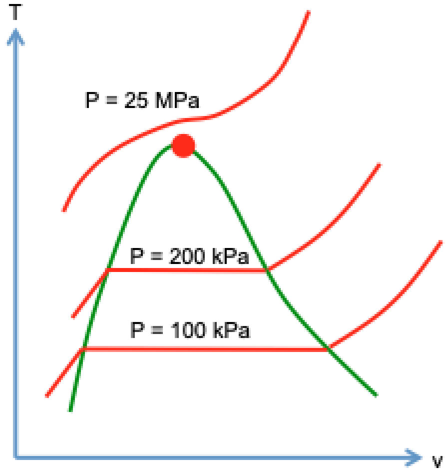
T = 120.21 Celsius.

The specific volume of the saturated water and the specific volume of the saturated steam are

$v_f = 0.001060\ m^3/kg$,
$v_g = 0.88568\ m^3/kg$.

As the pressure increases, the difference in the specific volume of the saturated water and the specific volume of the saturated steam reduces. When the pressure is fixed at a value above a *critical pressure*, the difference in the two phrases disappears, and the states of H_2O molecules under a constant pressure fall on a smooth curve.

In the T-v plane, the region of water-steam mixture is inside a *dome*.



Critical state. The top of the liquid-gas dome is called the *critical state*, or *critical point*. The top of the dome is a thermodynamic state. For H_2O molecules, the critical state has the following properties:

T = 374.1 Celsius
P = 22.089 MPa
v = 0.003155 m^3/kg

Subject to a constant pressure above the critical pressure, a liquid state can change to a gaseous state without crossing the liquid-gas phase boundary.

Exercise. YouTube critical state. Mark the process observed in the experiment on the T-v plane. Does the experiment demonstrate the critical state?

Exercise. How would you demonstrate the existence of the critical state of a pure substance experimentally?

Rule of mixture. Inside the dome, each state—a pair of values of T and P—corresponds to a mixture of water and steam of some proportion. The volume of the mixture, V, is the sum of the volume of the water, V_f, and the volume of the steam, V_g:

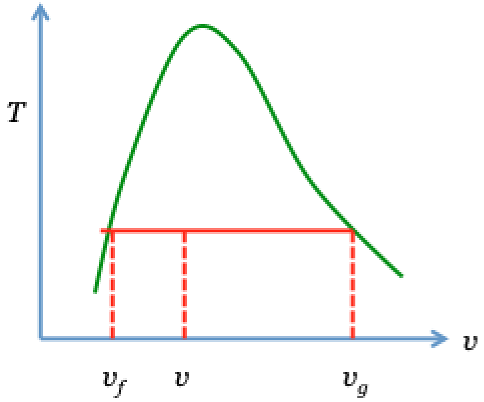
$V = V_f + V_g$.

The mass of the mixture, m, is the sum of the mass of the water, m_f, and the mass of the steam, m_g. The mass fraction of steam in the mixture, $x = m_g/(m_f + m_g)$, is called the *quality of the mixture*.

Define the specific volume of the water by $v_f = V_f/m_f$, the specific volume of the steam by $v_g = V_g/m_g$, and the specific volume of the mixture by $v = V/(m_f + m_g)$. Divide the equation $V = V_f + V_g$ by $m_f + m_g$, and these definitions lead to

$v = (1 - x)v_f + xv_g$.

This expression is called the *rule of mixture*, or the *lever rule*.



The rule of mixture applies to every extensive property. For example,

$s = (1 - x)s_f + xs_g$.

The quality x itself is an intensive property. We can specify a state of water-steam mixture by a pair of values of P and x, for example.

Exercise. At 100 kPa, the specific volume of the saturated water is $v_f = 0.001043$ m^3/kg, and the specific volume of the saturated steam is $v_g = 1.69400$ m^3/kg. A tank of a volume of 1 m^3 contains a mass of 2 kg of H_2O. What is the mass of the steam in the tank?

Many functions of two variables. A pure substance has numerous properties. Six examples are T, P, U, V, S, and x. We will introduce more properties, including enthalpy, thermal capacity, Helmholtz function, Gibbs function, compressibility, coefficient of thermal expansion, and Joule-Thomson coefficient.

A state of a pure substance is specified by values of two independent properties. Once a state is specified, every property takes a unique value. Thus, each property is a function of the two independent properties.

For example, for the six properties, T, P, U, V, S, and x, a total of 15 pairs exist. Each pair can serve as independent properties, with caveats to be noted later. Thus, for each pure substance, a total of 60 functions of two variables exist. A great accomplishment of thermodynamics is to show that all these functions can be derived from a single function, S(U,V), called a *characteristic function*. For now, let us focus on the task of presenting many functions of two variables.

A function of two variables, f(x,y), can be presented in several ways.

An equation expresses a property as a function of two other properties, and is called an *equation of state*. An example is the ideal gas law, kT = PV/N. This method of presentation is practical for idealized models with simple equations of state.

A function f(x,y) is a surface in three dimensional space (x, y, f). This method is rarely useful in practice, because our common media of presentation—paper and screen—are two-dimensional.

Steam tables. Given values of x and y, list the value of f. This method of presentation has prevailed for over two centuries, but has started to fade. For many functions, printed tables take many pages, and online tables are difficult to use. And these are just for one substance.

The thermodynamic properties of water and steam are presented using tables, called the *steam tables*. On-paper steam tables are fading out, just as on-paper maps are. Properties of steam and water live online and in apps.

Steam apps. Steam Tables Lite is a free iPhone app. To specify a state, the user enters the values of two properties, and the app displays the values of the other properties. The available pairs to be specified are PT, PH, PS, Px, Tx, HS.

Another iPhone app, Steam Pad, displays TS, HS, PH, and TH diagrams. The user can add and remove constant-property lines, and zoom in and out. The app also lets the user input any pair of properties in the list P, T, V, H, S, x. Why is U not in the list?

Yet another iPhone app, International Steam Tables, lets the user enter values of PT, Px, PH, PS, HS, UV, HV. The app outputs the other properties. Also see the associated online property calculator.

Why does an app limit to a few pairs of input? Why not let the user input any pair of properties? Does the memory of the phone place a list? The list of properties can be as long as that can fit on the screen of a phone. In coming years, steam apps will surely compete and excel.

Steam tables (old fashioned). To build intuition for the data of properties, we next describe the use of the NIST tables:

- NIST PDF (for water)
- NIST software (for many substances)
- YouTube steam table, Example 1, Example 2.

Inside the dome. The dome is formed by a smooth curve. The curve on the left is the specific volume of saturated water as a function of temperature, $v_f(T)$. The curve on the right is the specific volume of saturated steam as a function of temperature, $v_g(T)$.

For each pressure, the NIST table for saturation (pressure) lists the boiling temperature T, the specific volume of saturated water, v_f, and the specific volume of saturated steam, v_g. Also listed are specific entropies and enthalpies of the saturated water and saturated steam.

For example, at P = 100 kPa, the table lists the boiling temperature of 99.606 Celsius. The specific volumes and entropies of the saturated water and saturated steam are

$v_f = 0.001043\ m^3/kg$, $v_g = 1.6939\ m^3/kg$
$s_f = 1.3028$ kJ/kg-K, $s_g = 7.3588$ kJ-kg/K

Inside the dome, each point (T, v) corresponds to a state, a mixture of a state of saturated water and a state of saturated steam. In the mixture, the water and steam have the same temperature and the same pressure. A horizontal line through the point (T, v) intersects the curve of saturated water on the left and the curve of saturated steam on the right. The state (T,v) is a mixture of saturated water and saturated steam of some proportion.

Once a pressure is given, the table gives the boiling temperature. Thus, temperature is not an independent property. One more independent property is needed to specify a state. For example, a state can be specified by a pair of values P and v.

Given a value of P and a value of v, the quality x is determined by the rule of mixture $v = (1 - x)v_f + xv_g$, where v_f and v_g are found from the steam table for the given pressure.

Once x is determined, the entropy of the mixture is determined by $s = (1 - x)s_f + xs_g$, where s_f and s_g are found from the steam table for the given pressure. One can also determine the values of energy and enthalpy.

Exercise. 2 kg of H_2O has a pressure of 100 kP and a volume of 1 m^3. Use the NIST table (dome, pressure) to determine the energy and entropy.

Exercise. 1 kg of H_2O molecules has a temperature of 50 Celsius and a volume of 0.5 m^3. Use the NIST table (dome, temperature) to determine energy and entropy.

Exercise. On-paper steam tables use temperature and/or pressure as independent properties. This practice saves paper but wastes user's time. Of course, apps use no paper and aim to save user's time. Here is an illustration. 2 kg of H_2O has a volume of 1 m^3 and an energy of 4000 kJ. Determine the pressure. First try on-paper steam tables. Then try the app International Steam Tables, or the associated online property calculator.

Outside the dome. On a T-v plane, whereas a point inside the dome represents a *two-phase state*, a point outside dome represents a *single-phase state*.

Outside the dome, states of water and states of steam form a continuous region: no boundary divides states into liquid phase and gas phase. Some labels are used to name various parts of this region. Examples include compressed water, superheated steam, and supercritical fluid. Any labeling of a continuous region is arbitrary and should not distract us.

Outside the dome, the state can change continuously by independent change of temperature and pressure. Thus, temperature and pressure are independent properties. A pair of values (T,P) specifies a state outside the dome. Give a pair of values (T,P), the NIST table (outside the dome) lists volume, enthalpy, and entropy.

For example, at T = 400 C and P = 100 kPa, water molecules has the specific volume, specific enthalpy, and specific entropy:

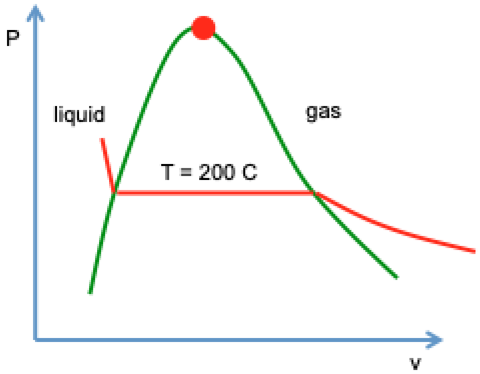
v = 3.1027 m^3/kg
h = 3278.6 kJ/kg
s = 8.5452 kJ/kg-K

Exercise. Use the NIST steam tables to find specific volume, enthalpy, and entropy at the state of 200 C and 300 kPa.

Pressure-volume plane. Pressure is an intensive variable, and 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 state of a unit mass of a pure substance. Note a single-phase state of water, a single-phase state of steam, and a two-phase state of water-steam mixture.

All states of a fixed temperature correspond to a curve in the P-v plane, called an *isotherm*. This sequence of states of a constant temperature is called an *isothermal process*.

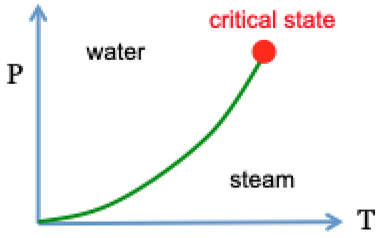
Consider an isotherm of T = 200 Celsius. At a high pressure, water is in the liquid phase. The volume of the liquid increases as the pressure drops.



At 200 C, the NIST table (dome, temperature) lists

P = 1555 kPa
v_f = 0.001157 m^3/kg
v_g = 0.12722 m^3/kg

Temperature-pressure plane. Temperature and pressure are both intensive properties. Each point on the temperature-pressure plane represents a pair of temperature and pressure.

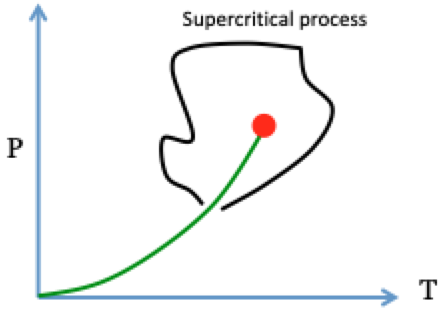


When the substance equilibrates in a single-phase state, the state corresponds to a point in the temperature-pressure plane. In the vicinity of a single homogeneous state in equilibrium, the state can change continuously by independent change of temperature and pressure. All single-phase states form a continuous region.

When the substance equilibrates in a two-phase state, each phase is a saturated state, and the two saturated states have the same temperature and the same pressure. In the temperature-pressure plane, all two-phase states fall on a curve, called a *phase boundary*. A point on the phase boundary represents many two-phase states, each state being a mixture of the two saturated states of some proportion. For the phase boundary, given a pressure, the steam tables list the temperature, as well as the specific volumes, energies, and entropies of the two saturated states.

As the pressure increases, the liquid-gas phase boundary terminates at the critical state. Beyond the critical state, no boundary divides water and steam.

Consider a point on the liquid-gas boundary. Near the point, the substance forms a liquid on one side of the boundary, and forms a gas on the other side of the boundary. The liquid can change to the gas through a sequence of single-phase states, without crossing the liquid-gas phase boundary. Such a sequence of single-phase states is called a *supercritical process*.



Exercise. Search for a temperature-pressure phase diagram for ammonia. Ammonia can be safely stored in a tank as a liquid at room temperature. Explain the science of this practice.

Exercise. Nitrogen cannot be stored as a liquid at room temperature. Why?

Exercise. Sketch a supercritical process on a T-v diagram and on a P-v diagram.

Exercise. Wiki supercritical fluid extraction. Explain the science of this process.

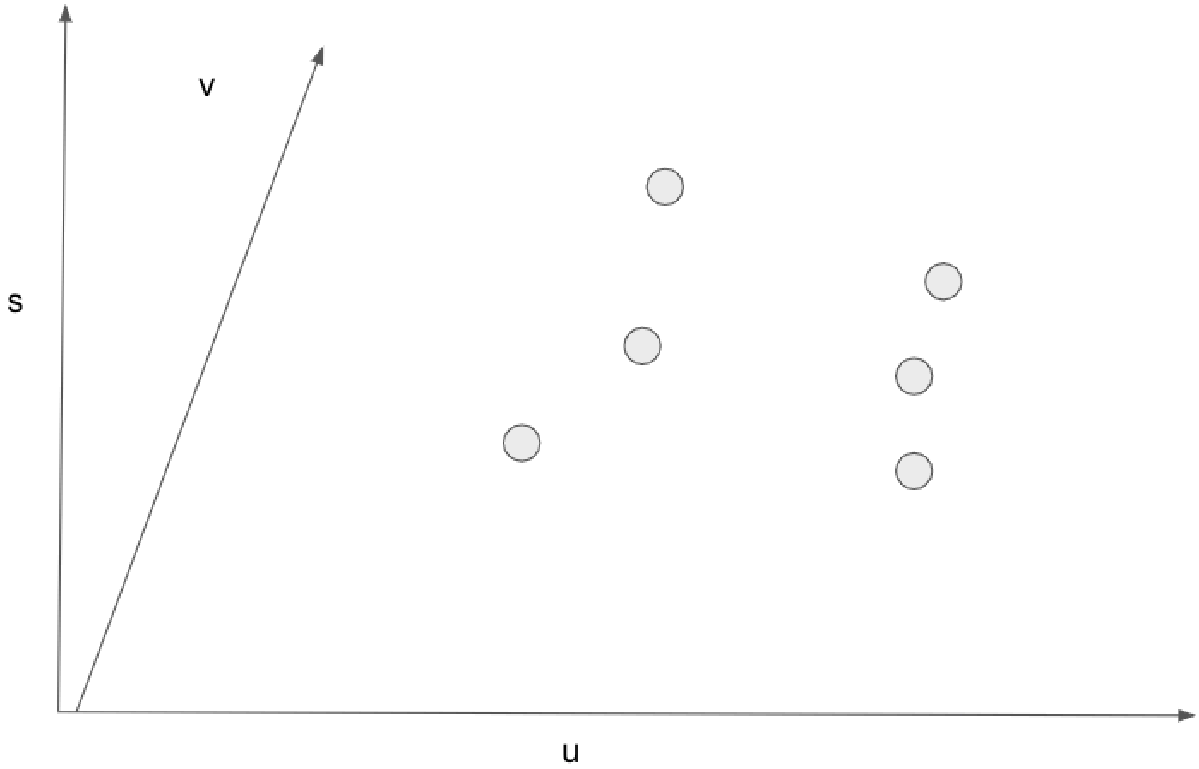
Twitter Steam Tables

Thermodynamic theory of water and steam

A fixed amount of a pure substance is a closed system. Its thermodynamic states are specified by two independent properties. For now, let the two independent properties be energy U and volume V. Once U and V are fixed, a fixed amount of the pure substance becomes an isolated system, which 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). This model neglects many effects, such as gravitational field, elastic field, electrical field, and magnetic field.

Energy-volume-entropy space. We now describe the thermodynamic theory of phases of a pure substance due to . Before Gibbs developed this theory, the empirical facts of pure substances described above were known. The facts were so numerous and called for a theory to organize them.

For a pure substance of N molecules, let u = U/N, v = V/N, and s = S/N be the average energy, volume, and entropy per molecule. The variables u, v, s form the axes of a three-dimensional vector space. A point in this space represents a state of the substance. For example, mark a few homogeneous states in the (u,v,s) space. Next consider mixtures of these homogeneous states.



Rule of mixture. Consider two homogeneous states, A and B. A mixture of the two homogeneous states has a total of N molecules, of which N_A molecules are in homogeneous state A, and N_B molecules are in homogeneous state B. 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

$$N = N_A + N_B.$$

Divide this equation by N, and we obtain that

$$y_A + y_B = 1.$$

The number fractions y_A and y_B are nonnegative and sum to one.

Let the average energy, volume, and entropy per molecule be (u_A,v_A,s_A) in homogeneous state A, be (u_B,v_B,s_B) in homogeneous state B, and be (u,v,s) in the mixture. Energy is an extensive property, 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

$u = y_A u_A + y_B u_B$.

The same holds true for volume and entropy:

$v = y_A v_A + y_B v_B$,
$s = y_A s_A + y_B s_B$.

These equations are called a rule *of mixture* in materials science, and a *convex combination* in linear algebra. The vector (u,v,s) is a linear combination of two vectors (u_A,v_A,s_A) and (u_B,v_B,s_B), with the coefficients y_A and y_B being nonnegative and summing to one.

The rule of mixture has a graphic interpretation in the energy-volume-entropy space. The mixture, (u,v,s), is a point on the segment joining the two homogeneous states, (u_A,v_A,s_A) and (u_B,v_B,s_B). The location of the point (u,v,s) on the segment depends on the fraction of molecules y_A and y_B allocated to the two homogeneous states.

Next 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 $y_C = N_C/N$. The number of molecules in the mixture is the sum of molecules in the three homogeneous states:

$N_A + N_B + N_C = N$.

Divide this equation by N, and we obtain that

$y_A + y_B + y_C = 1$.

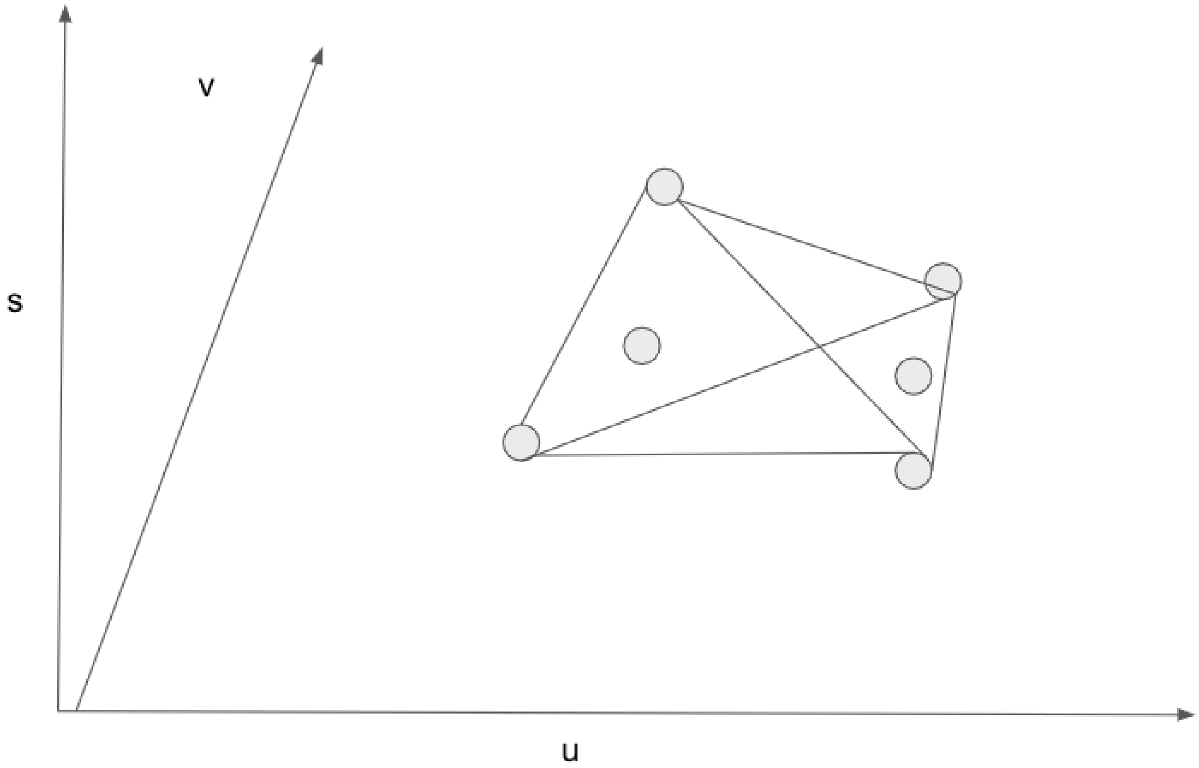
The number fractions y_A, y_B and y_C are nonnegative and sum to one. The three homogeneous states, (u_A,v_A,s_A), (u_B,v_B,s_B), and (u_C,v_C,s_C), are vertices of a triangle in the energy-volume-entropy space. The energy, volume, and entropy per molecule of the mixture (u,v,s) are given by a rule of mixture:

$u = y_A u_A + y_B u_B + y_C u_C$,
$v = y_A v_A + y_B v_B + y_C v_C$,
$s = y_A s_A + y_B s_B + y_C s_C$.

The mixture, (u,v,s), is a point in the energy-volume-entropy space, on the triangle formed by the three homogeneous states, (u_A,v_A,s_A), (u_B,v_B,s_B), and (u_C,v_C,s_C). The location of the point of mixture on the triangle depends on the fraction of molecules y_A, y_B and y_C allocated to the three homogeneous states.

Convex hull. Further consider a mixture of any number of homogenous states. The homogeneous states correspond to a set of vectors in the energy-volume-entropy space. A mixture is a linear combination of two or more homogenous states, with coefficients being nonnegative and summing to one. All mixtures of the homogeneous states constitute a solid figure, called the *convex hull* of the homogenous states.

The convex hull in linear algebra generalizes the rule of mixture in materials science. For a set of discrete homogeneous states, the convex hull is a polytope of triangular faces. A vertex represents a homogeneous state, an edge represents mixtures of two homogeneous states, and a face represents mixtures of three homogeneous states. A point in the interior of the polytope represents a mixture of two or more homogeneous states.

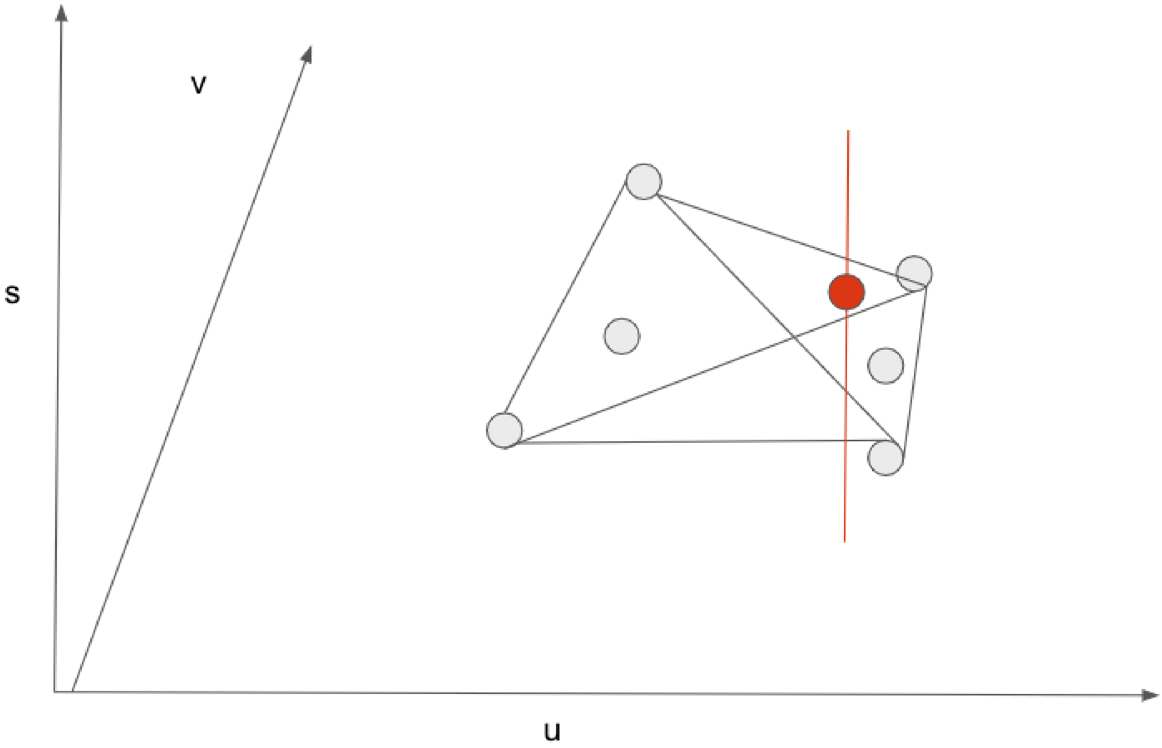


BAT on mixtures. So far, the three properties—energy, volume, and entropy—play similar roles. All we have invoked is that they are extensive properties. A fixed amount of a pure substance becomes an isolated system when energy and volume are fixed. The isolated system corresponds to a vertical line in the energy-volume-entropy space.

This isolated system has an enormous number of internal variables. For a given set of homogenous states, a mixture can allocate any fraction of molecules to each homogeneous state. All mixtures form the convex hull of the homogeneous states.

A point on the segment of the vertical line in the convex hull represents a mixture. Such a mixture in general is *not* in equilibrium. The height of the point represents the subset entropy of the isolated system. The isolated system equilibrates when the subset entropy maximizes—that is, when the vertical line intersects the upper boundary of the convex hull.

For a discrete set of homogenous states, the upper boundary of the convex hull is a sheet of many triangles. The pure substance can equilibrate in a homogeneous state, a mixture of two homogeneous states, or a mixture of three homogeneous states. The pure substance cannot equilibrate in four or more homogeneous states.



A pure substance of three phases. A collection of H_2O molecules can form ice, water, and steam. For the time being, represent each phase by its own energy-volume-entropy relation, s(u,v), corresponding to a smooth, convex surface in the energy-volume-entropy space.

A point on each surface represents a homogeneous state. These homogeneous states, together with all mixtures of two or more of the homogeneous states, form the convex hull. The upper boundary of the convex hull is derived by rolling a plane tangent to the surfaces of homogeneous states. A tangent plane can touch the surfaces at one, two, or three points, but not four or more points.

When a tangent plane touches the three 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. At fixed slopes, the tangent plane touches two 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 the surface of one phase.

The upper boundary of the convex hull has a single sheet. Every point on the upper boundary of the convex hull corresponds to a state of equilibrium. We next list distinct types of equilibrium.

Equilibrium of a single homogeneous state. If a point on the upper boundary of the convex hull belongs to the surface of a phase, the point represents 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 surface of a phase to change the two slopes of the tangent plane independently. The slopes of the tangent plane determine 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 point on the upper boundary of the convex hull belongs to a tie line, the point represents an equilibrium mixture of two homogeneous states. The common tangent plane can roll on the surfaces of two phases 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 point on the upper boundary of the convex hull belongs to the triangle, the point represents an equilibrium mixture of three homogeneous states. The tangent plane has no degree of freedom to roll.

Equilibrium of four homogeneous states is impossible for a pure substance. A pure substance can have more than three phases. For example, H_2O molecules can form many crystalline structures at high pressure. Each crystalline structure corresponds to a distinct phase. Can four or more homogeneous states equilibrate? The answer is no, so long as the states of the pure substance are specified by two independent variables.

In the two-variable model, each phase corresponds in the entropy-volume-energy space to a surface. Four phases correspond to four surfaces. A plane rolling over the four surfaces can at once touch one surface, two surfaces, or three surfaces, but not four surfaces. The exception is that all four surfaces are exactly positioned to have a common tangent plane. The chance for such an exception is negligible.

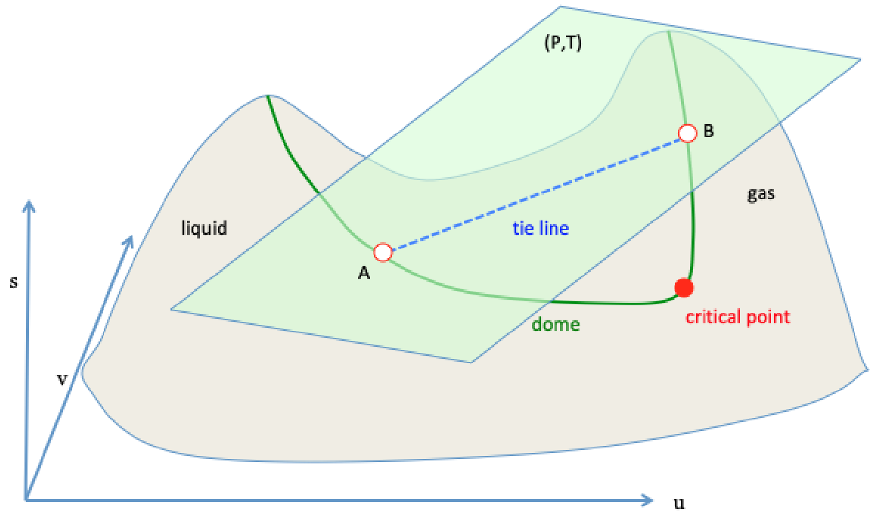
Critical state. Andrews (1869) reported the experimental observation that a substance can change from a liquid state to a gaseous state through continuous change of homogeneous states. That is, on the temperature-pressure plane, the liquid-gas phase boundary terminates at a point, called the critical state.

On the basis of this experimental observation, Gibbs (1873) introduced the theory of critical state. He 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 comes to an end. the point where the two points of contact fall together is the critical state."

His primitive surface is the surface of homogeneous states, and his derived surface is the upper boundary of the convex hull.

Gibbs modeled a pure substance with two surfaces: a convex surface for the solid phase, and a nonconvex surface for the liquid and gas phases. Google Books: Scientific Papers of J. Willard Gibbs



Exercise**.** YouTube Gibbs surface and Maxwell cast.

A. In the energy-volume-entropy space, sketch the primitive water-steam surface. Describe the surface with words.
B. Sketch the derived water-steam surface. Describe the surface with words.
C. Derive equations that determine a tie line.

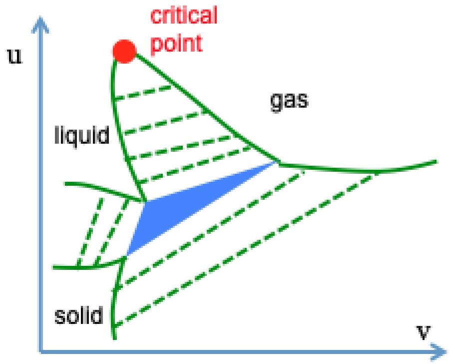
Metastability. A surface of homogeneous states may contain a convex part and a non-convex part. The curve separating the two parts is called the *limit of metastability*. If a convex part of the surface lies below the upper boundary of the convex hull, the part of the surface is beyond the limit of absolute stability. Each point of this part of the surface is called a metastable state. 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 upper boundary of the convex hull 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 state. He did not, however, draw the surfaces of phases and the upper boundary of the convex hull in the energy-volume-entropy space.

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

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

Here I sketch the projection of the upper boundary of the convex hull to the energy-volume plane. Each thermodynamic state corresponds to a distinct point in the energy-volume plane. Projected onto this plane are the tie lines in each region of mixtures of two homogeneous states



Exercise**.** Find the data to draw the ice-water-steam triangle on the energy-volume plane.

Exercise. Find the data to draw the water-steam dome on the energy-entropy plane. Include the critical point, several tie lines in the dome. Also include the ice-water-steam triangle, the ice-steam belt, and ice-water belt.

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 property, and obeys the rule of mixture. We next discuss the roles of the other two functions of state: temperature and pressure. They are intensive properties.

For a smooth function s(u,v), recall the meanings for the slopes:

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

When the substance equilibrates in a mixture of two homogeneous states, a tangent plane contacts the surfaces of homogeneous states 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:

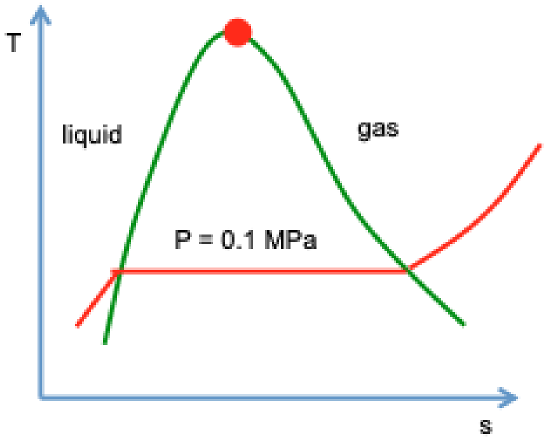
$1/T = \partial s_A(u_A,v_A)/\partial u_A = \partial s_B(u_B,v_B)/\partial u_B$,
$P/T = \partial s_A(u_A,v_A)/\partial v_A = \partial s_B(u_B,v_B)/\partial v_B$.

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. Once the surfaces of the two phases are given, $s_A(u_A,v_A)$ and $s_B(u_B,v_B)$, for a given temperature, the above equations solve the pressure P and the two tangent points (u_A,v_A,s_A) or the state (u_B,v_B,s_B).

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 Celsius, x = 0.9.
(b) T = 100 Celsius, P = 10 kPa.

(c) $T = 100$ Celsius, $P = 500$ kPa.
(d) $T = 50$ Celsius, $v = 1\ \mathrm{m^3/kg}$.

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

Exercise. YouTube Gibbs surface and Maxwell cast.

A. In the energy-volume-entropy space, sketch the primitive water-steam surface. Describe the surface with words.
B. Sketch the derived water-steam surface. Describe the surface with words.
C. Derive equations that determine a tie line.

Twitter Critical state

Thermodynamic relations

Function U(S,V). Write the equation $dS = (1/T)dU + (P/T)dV$ as

$dU = TdS - PdV$.

A closed system is characterized by a function $U(S,V)$. The above equation implies that

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

The second equation characterizes the elasticity of a closed system undergoing isentropic deformation.

Function F(T,V). Recall the definition of free energy, $F = U - TS$. We obtain that

$dF = -SdT - PdV$.

The free energy is a function $F(T,V)$. The above equation implies that

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

The second equation characterizes the elasticity of a closed system undergoing isothermal deformation.

Function G(T,P). Define $G = U - TS + PV$. Confirm that

$dG = -SdT + VdP$.

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

Twitter Gibbs free energy is a shadow of entropy

Van der Waals model

The ideal gas model represents real gases well at high temperature and low pressure, where individual molecules are far apart on average. However, at low temperature and high pressure, when the molecules nearly condense, the ideal gas model does not describe real gases well.

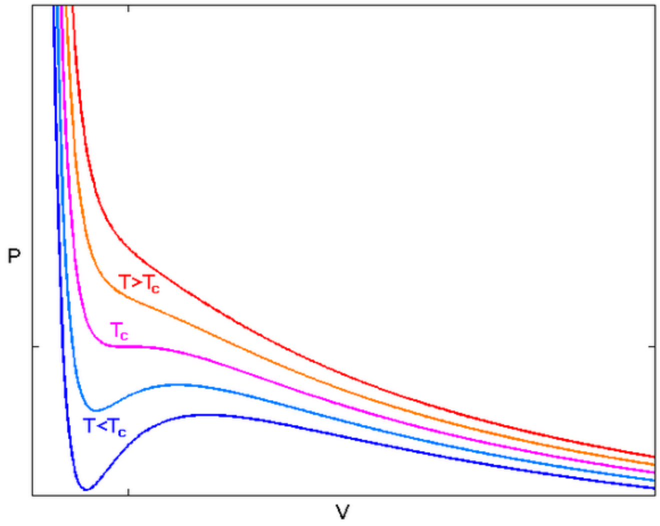
In a half bottle of water, some H_2O molecules form water, and others form steam. What is the molecular origin for this phase separation? We now understand this phenomenon in terms of the competition of two effects. First, H_2O molecules attract one another through hydrogen bonds, and the intermolecular attraction favors aggregation. Second, the open space in the bottle lets the molecules fly, and the increase of the subset entropy favors dispersion. This competition is represented in the van der Waals model (1873).

Equation of state, P(T,V). van der Waals modified the ideal gas law as follows:

$P = k_B T/(V - b) - a/V^2$.

Here V is the volume of the bottle divided by the number of molecules. As we will see, a represents the attraction between molecules, and b represents the finite volume of each individual molecule. The model explains the origin of the liquid-gas phase change.

When a = 0 and b = 0, the van der Waals equation of state reduces to the ideal gas law, $PV = k_B T$.



(Figure taken from the Wikipedia page on van der Waals equation, plotted for some values of a and b.)

Critical state. For a constant temperature T, the van der Waals equation corresponds to a curve on the pressure-volume plane. An isotherm at a high temperature is a monotonic curve. An isotherm at a low temperature has a minimum and a maximum. When the minimum and the maximum collide, the fluid reaches the critical state.

Recall a fact in calculus. At the minimum and the maximum of an isotherm, the following condition holds:

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

Recall another fact in calculus. Near the minimum, the isotherm is convex down, $\partial^2 P(T,V)/\partial V^2 > 0$. Near the maximum, the isotherm is convex up, $\partial^2 P(T,V)/\partial V^2 < 0$. The minimum and maximum collide when

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

Thus, the critical state satisfy three equations:

$P = kT/(V - b) - a/V^2$,
$- kT/(V - b)^2 + 2a/V^3 = 0$,
$2kT/(V - b)^3 - 6a/V^4 = 0$.

Solving these three equations, we obtain that

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

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

Free energy. The van der Waals model prescribes the function P(T,V). Recall that T and V are also the independent variables of the free energy function, F(T,V). Integrate the identity $-P = \partial F(T,V)/\partial V$, and we obtain that

$F(T,V) = - k_B T \log(V - b) - a/V + Z(T)$.

The integration is over V, but F is a function of T and V. Consequently, the constant of this integration is a function of temperature, Z(T).

Using the free energy function F(T,V), we can determine all other thermodynamic properties as functions of T and V.

Entropy. Recall the identity $-S = \partial F(T,V)/\partial T$. We obtain that

$S(T,V) = k_B \log(V - b) - dZ(T)/dT$.

At a fixed temperature, the number of quantum states per molecule is proportional to (V - b). Thus, b represents the size of an individual molecule.

Energy. Recall the definition of free energy, $F = U - TS$. We obtain that

$U(T,V) = - a/V + Z(T) - TdZ(T)/dT$.

Note that V is the volume of the container divided by the number of molecules. Thus, 1/V is the number density of molecules in the container. The van der Waals model assumes that the intermolecular attraction reduces the internal energy by an amount proportional to the number density of the molecules. This seems to be a reasonable first-order approximation. Whereas the internal energy of an ideal gas is independent of volume, the internal energy of a van der Waals gas depends on both volume and temperature.

Thermal capacity. Recall the definition of thermal capacity, $C_V = \partial U(T,V)/\partial T$. In general, the thermal capacity of a pure substance is a function of temperature and volume, $C_V(T,V)$. Inserting the expression for internal energy, we obtain that

$C_V = - Td^2Z(T)/dT^2$.

In the van der Waals model, the thermal capacity C_V is a function of temperature only, independent of volume.

Competition between entropy and energy. Van der Waals model lets us analyze this competition quantitatively. We next fix the temperature, and regard F(T,V) as a function of V. Entropy and energy compete to bend the curve F(T,V). The term $- k_BT \log(V - b)$ comes from the entropy, and is convex downward, which stabilizes a homogeneous state. The term $-a/V$ comes from the intermolecular attraction, and is convex upward, which destabilizes a homogeneous state.

Entropy tends to disperse molecules and favors gas. Energy (i.e., intermolecular attraction) tends to aggregate molecules and favors liquid.

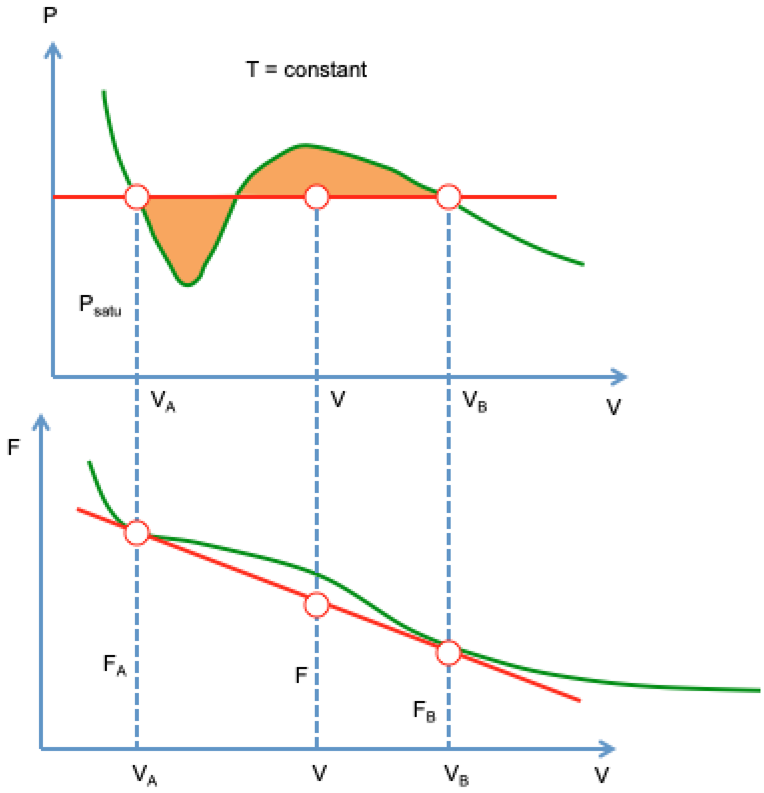
At a high temperature, the entropy prevails for all values of volume, so that the entire curve F(T,V) is convex downward, and the entire curve P(T,V) is monotonic.

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

The thermodynamics of a pure substance is fully characterized by the free energy function, F(T,V). 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.

Equilibrium of two phases. For a fixed T, the function F(T,V) is a curve on the F-V plane. Both F and V are extensive properties. The F-V plane is a two-dimensional vector space. Each point on the curve represents a homogeneous state. The curve of homogenous states is called the primitive curve. A mixture of homogeneous states obeys the rules of mixture. All mixtures correspond to the convex hull of the curve. The boundary of the convex hull is called the derived curve.

A convex-upward part of the primitive curve corresponds to a phase transition. Draw a line tangent to the curve at two points A and B. 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 corresponds to a point on the tangent line, located at the center of gravity according to N_A and N_B. For a fixed volume V, an equilibrium mixture minimizes the free energy function and corresponds to a point on the tangent line.



The two tangent 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(T,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 equates the two shaded areas. This construction is called the *Maxwell rule*.

Recall the definition of free energy, F = U - TS. The above equation of equilibrium is the same as

$U_A - U_B - T_{satu}(S_A - S_B) = P_{satu}(V_B - V_A)$.

This condition of equilibrium does not hide entropy.

Exercise. In the van der Waals model, assume a constant thermal capacity C_V. Find the function S(U,V). Plot this function around the critical state using some values of a, b, and C_V.

Exercise. The van der Waals model does not fit experimental data accurately. Clausius proposed a different model: $P = k_BT/(V - b) - a/[T(V + c)^2]$. Here V is the volume of the container divided by the number of molecules. Calculate the functions F(T,V), S(T,V), and U(T,V). Interpret a, b, c in terms of energy and entropy. Express the critical temperature, pressure, and volume in terms of a, b, c.

From ideal gas to real fluid, multicomponent materials, and materials genome

Ours is the age of molecules and the age of data. Molecules generate big data of thermodynamics. The BAT guides us to measure, curate, and use the big data.

Thermodynamic data are measured in many disciplines, including thermometry, calorimetry, thermochemistry, electrochemistry, and biochemistry.

Ideal gas. The thermodynamics of an ideal gas is fully specified by two equations:

$PV = Nk_BT$
$dU = Nc_V(T)dT$.

But the ideal gas law describes a substance only at high temperatures and low pressures.

Van der Waals model. The Helmholtz function F(T,V) lets us calculate all other thermodynamic properties as functions of T and V. For example, the relation $P = - \partial F(T,V)/\partial V$ calculates the function P(T,V).

The van der Waals model can be expressed as a Helmholtz function:

$F(T,V) = - k_BT \log(V - b) - a/V + Z(T)$.

The function contains two constants a and b, and a function Z(T). The constants a and b and the function Z(T) are left to fit experimental data. For example, a and b can be used to fit the measured P(T,V) function, and Z(T) can be used to fit the measured thermal capacity.

Steam tables then and now. The International Association for the Properties of Water and Steam (IAPWS) is an international association of national organizations concerned with the properties of water and steam. The following short documents describe the history and contemporary practice of constructing steam tables.

- Allen H. Harvey, Steam Tables
- Revised release on the IAPWS formulation 1995
- A Twitter thread on Steam Tables as a Big Data project

Big data. Of course, the van der Waals model is still too simplistic to fit data accurately. The IAPWS prescribes a form of the Helmholtz function F(T,V) with eighty-some coefficients, which are fit to experimental data. Constructing the steam tables dates back more than 200 years, and is perhaps the first big-data project. Thermodynamics teaches us how to obtain, curate, and use properties.

The Helmholtz function of a similar form has been adopted to fit data of numerous other pure substances (NIST software).

The approach has been extended to multi-component substances, and is called the CALPHAD method. Here, a form of the Gibbs function is assumed as a function of T, P, and composition. The function contains coefficients to fit experimental data.

In 2011, President Obama announced the Material Genome Initiative.

Exercise. A well-insulated, rigid tank is divided into two rooms by a rigid membrane. Room A has a volume of 1 m^3, and contains a water-seam mixture at a temperature of 200 C and a quality of 50%. Room B has a volume of 2 m^3, and contains a water-steam mixture at a temperature of 300 C and a quality of 90%. The membrane now ruptures and the H_2O molecules come to a final state of equilibrium. Determine the volume, energy, entropy, pressure, temperature, and quality of the final state. Compare the entropies of the initial and final state. You may wish to use a property calculator.

Open system

Now enters a third supporting actor—matter. In thermodynamics, matter, energy, and volume play analogous supporting roles, of equal importance.

Act three: **The union of entropy and the number of molecules of one species produces a child—chemical potential**. Whereas energy transfers from a place of high temperature to a place of low temperature, a species of molecules transfers from a place of high chemical potential to a place of low chemical potential. The analogy between energy and matter, as well as between temperature and chemical potential, plays out in this act.

An open system is a family of isolated systems of many variables

An open system and its surroundings transfer energy, volume, and matter. The open system may contain many species of molecules. To illustrate ideas, assume that only two species—called species 1 and species 2—transfer between the open system and surroundings.

The open system may contain molecules other than species 1 and 2. We assume that these other species do not transfer between the open system and the surroundings. For example, a polymer network is swollen with water and ethanol molecules. When the swollen polymer is submerged in a solution of water and ethanol, both water and ethanol molecules transfer between the swollen polymer and the solution, but the polymer network does not. As another example, the open system is enclosed by a semipermeable membrane, which lets molecules of species 1 and 2 pass, but not any other species of molecules.

For an open system, let the energy be U, the volume be V, the number of molecules of species 1 be N_1, and the number of molecules of 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 $\Omega(U,V,N_1,N_2)$. Denote the entropy of this isolated system by

$S(U,V,N_1,N_2) = k_B \log \Omega(U,V,N_1,N_2)$.

As U, V, N_1, N_2 vary, the function $S(U,V,N_1,N_2)$ characterizes the open system as a family of isolated systems.

Definition of chemical potentials

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

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

Between the open system and its surroundings, when we block the transfer of molecules, but allow the transfer of energy and volume, the open system becomes a closed system. We have already named two partial derivatives:

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

The other two partial derivatives of the function $S(U,V,N_1,N_2)$ 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 chemical potentials are intensive properties.

The unit for a chemical potential is energy per molecule. Often, the amount of molecules is measured in moles, so that the unit for a chemical potential is energy per mole. Occasionally, a chemical potential is reported in the unit of energy per unit mass. So, watch out for the unit used for a chemical potential.

Write

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

This equation applies to an open system, and generalizes the equation for a thermal system, and the equation for a closed system.

For an open system, temperature measures the tendency of energy escaping from the open system, pressure measures the tendency of volume adding to the open system, and chemical potential of a molecular species measures the tendency of the species escaping from the open system.

Temperature vs. chemical potential

Usage of words. When speaking of a chemical potential, we should identify both the molecular 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. 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 identify 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 fundamental: our world has many species of molecules, but only one species of energy. So long as no chemical reaction takes place, the number of each species of molecules is conserved; for example, the number of oxygen molecules is conserved, and the number of nitrogen molecules is conserved. By contrast, all forms of energy can convert to one another: it is the total energy that is conserved.

Experimental significance of temperature. The temperature of a system measures the change in the entropy of the system associated with the change in energy. Energy flows from a place of high temperature to a place of low temperature. Temperature corresponds to the daily experience of hotness. Indeed, our daily experience of hotness is so pervasive that it may even impede our learning of thermodynamics. It takes significant effort for us to sort out, of all our feelings for hotness, what is essential and what is incidental. Only then we can link our daily experience of hotness to the thermodynamics of temperature.

Experimental significance of a chemical potential. The chemical potential of ethanol in a glass of wine measures the escaping tendency of ethanol from the wine. Ethanol goes from a place of high chemical potential to a place of low chemical potential. Chemical potential corresponds to the daily experience of the smell of the wine.

Of all species of molecules, water is particularly significant to us. Indeed, our experience of humidity maps to the thermodynamics of the chemical potential of water, just as our experience of hotness maps to the thermodynamics of temperature.

In everyday language, some of us may say that a wine is concentrated with alcohol when we feel a strong smell. Indeed, the concentration of a solute in a solution maps to the chemical potential of the solute in the solution.

Why are we familiar with temperature, but not chemical potential? Given similarly pervasive roles of temperature and chemical potentials in nature, we may wonder why we do not have the same level of familiarity with the two types of quantities.

I can think of two reasons: cultural and technical. The cultural reason has to do with what our parents tell us. When we touch hot water, our parents tell us that the temperature of water is high. When we smell a strong wine, however, our parents miss the teachable moment, and fail to tell us that the chemical potential of ethanol in the wine is high. The phrase—the chemical potential of ethanol in the wine—is simply not in the everyday language. So blame our parents: they tell us temperature, but not chemical potential.

The technical reason is that, when we feel that something is hot, we attribute the sense of hotness to a single quantity: temperature. However, when we smell something, we may sense multiple quantities: a list of chemical potentials of a list of species of molecules. It is simpler to deal with one quantity than deal with multiple quantities. So blame our world: it confuses us with too many species of molecules.

Two open systems in contact

Two open systems, A and B, are said to be in contact if they exchange energy, space, and matter. We next analyze the two open systems in contact using the basic algorithm of thermodynamics (BAT).

Step 1. Construct an isolated system with internal variables. Characterize system A by a function $S_A(U_A,V_A,N_{1,A},N_{2,A})$, and system B by another function $S_B(U_B,V_B,N_{1,B},N_{2,B})$. We have to introduce a few quantities of two subscripts; for example, $N_{1,A}$ 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 isolated system conserves energy, space, and matter:

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

Here the conservation of matter means the following. 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, $N_{1,A}$, $N_{2,A}$. Changes of the internal variables satisfy the following relations:

$dU_A + dU_B = 0$
$dV_A + dV_B = 0$
$dN_{1,A} + dN_{1,B} = 0$
$dN_{2,A} + dN_{2,B} = 0$

Step 2. Find the subset entropy of the isolated system as a function of the internal variables. When the independent internal variables $U_A,V_A,N_{1,A},N_{2,A}$ are fixed at particular values, the isolated system flips 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 the subset entropy is a sum:

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

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

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

Step 3. Maximize the subset entropy to reach 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

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

Equilibrium requires four conditions. The two open systems have the equal temperature, equal pressure, equal chemical potential of molecular species 1, and equal chemical potential of molecular species 2.

Step 4. Change the internal variable# to increase the subset entropy. Out of equilibrium, the BAT requires that the subset entropy should increase as time progresses, $dS_{composite} > 0$.

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 $dS_{composite} > 0$ reduces to

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

Thus, molecular species 2 transfers 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. The chemical potential of a molecular species in an open system measures the tendency of the species to escape from the open system.

Equivalent definitions of chemical potentials

Gibbs (1875) used the function $U(S,V,N_1,N_2)$ to define the chemical potential:

$dU = TdS - PdV + \mu_1 dN_1 + \mu_2 dN_2$

This definition is equivalent to that by the function $S(U,V,N_1,N_2)$, which can be confirmed by solving dU from the expression $dS = (1/T)dU + (P/T)dV - (\mu_1/T)dN_1 - (\mu_2/T)dN_2$.

Some authors define chemical potentials using the Gibbs function, $G = U - TS + PV$. Here the Gibbs function is a function of four variables, $G(T,P,N_1,N_2)$. Thus,

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

This definition is equivalent to that using the entropy function.

Twitter Chemical potential
Slides Chemical potential

Chemical potential of molecules in a pure substance

Definition. A pure substance is a collection of molecules of a single species. Model a piece of a pure substance as an open system characterized by entropy as a function of energy, volume, and the number of molecules: S(U,V,N). Define temperature T, pressure P, and chemical potential μ by the equation

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

This equation defines T, P, and μ through the three partial derivatives of the function S(U,V,N).

Energy U, volume V, and entropy S are extensive quantities, proportional to the number of molecules N. Denote

$u = U/N$, $v = V/N$, $s = S/N$

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

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

Rewrite the above equation as

$\mu = u + Pv - Ts$

The chemical potential of molecules in a pure substance coincides with the Gibbs function per molecule.

Recall the definition of enthalpy per molecule, $h = u + pv$. Write

$\mu = h - Ts$

Chemical potential of molecules is the same in phases in equilibrium. Recall the theory of pure substance. When two or three phases are in equilibrium, the Gibbs function per molecule is the same in every phase, so that the chemical potential of molecules is the same in every phase. Thus, when water and steam equilibrate, the chemical potential of H_2O molecules in the two phases are the same:

$\mu_{water} = \mu_{steam}$

When water, ice, and steam equilibrate, the chemical potential of H_2O molecules in the three phases are the same:

$\mu_{water} = \mu_{ice} = \mu_{steam}$

Experiment determination of chemical potential. The chemical potential of molecules in a pure substance is a function, $\mu(T,P)$. For H_2O molecules, for example, h and s have been measured at various states of (T,P), and are listed in the steam tables. The expression $\mu = h - Ts$ gives the function $\mu(T,P)$.

In the steam tables, however, h and s are listed relative to an arbitrary state of (T,P). This practice is inconsistent with how μ plays in chemical reactions. To calculate μ, we modify the steam tables as follows. For water at T = 25 Celsius and P = 100 kPa, the enthalpy of formation and absolute entropy are

$h = - 285830$ J/mol
$s = 69.950$ J/mol K

These values differ from those in the steam tables. We add the differences to the values of h and s in all states of (T,P).

With the function $\mu(T,P)$ known, we can use pure H_2O molecules in various states of (T,P) as a "meter" to measure the chemical potential of H_2O molecules in any other system, such as a piece of cheese. At a given T, measure the partial pressure P of water vapor around the cheese. Read the value of chemical potential from the function $\mu(T,P)$.

Exercise. What is the chemical potential of water at T = 25 Celsius and P = 100 kPa?

Exercise. What is the chemical potential of water at T = 60 Celsius and P = 100 kPa?

Twitter Chemical potential of molecules in a pure substance

Chemical potential is a function of temperature and pressure. Recall that $ds = (1/T)du + (P/T)dv$. This equation, along with $\mu = u + Pv - Ts$, gives that

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

The chemical potential is a function of temperature and pressure, $\mu(T,P)$. The above equation identifies the partial derivatives:

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

Incompressible pure substance. In many applications of liquids and solids, the pressure is small, so that the volume per molecule in the system v, is taken to be independent of the pressure, and the system is called *incompressible*.

Recall the chemical potential of a species of molecules in a pure substance, $\mu = u + Pv - Ts$. For an incompressible substance, u, v, and s are functions of T, and are independent of P. Let P_0 be a reference pressure. Write

$$\mu(T,P) = \mu(T,P_0) + (P - P_0)v(T)$$

At a fixed temperature, the chemical potential of a species of molecules in an incompressible liquid is linear in pressure.

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

$PV = Nk_BT.$
$dU = Nc_V(T)dT$
$dH = Nc_P(T)dT, \quad c_P(T) = c_V(T) + k_B$
$dS = (Nc_P(T)/T)dT - (Nk_B/P)dP$

Also write these equations as

$PV = Nk_BT$.
$U = Nu(T)$
$H = Nh(T)$, $h(T) = u(T) + k_BT$
$S = Ns(T, P)$, $s(T, P) = s(T, P_0) - k_B \log(P/P_0)$

Here P_0 is a reference pressure. The two functions, $h(T)$ and $s(T,P_0)$, are listed in the ideal gas tables for various species of molecules.

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

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

Ideal gas mixture. The chemical potential of molecules of species 1 in an ideal gas mixture is

$\mu_1(T,P_1) = \mu_1(T,P_0) + k_BT \log(P_1/P_0)$

Here P_1 is the partial pressure of species 1 in the mixture. The function $\mu_1(T,P_0)$ is the same as that of a pure ideal gas. Thus, to determine the chemical potential of a species of molecules in an ideal gas mixture, we just measure the partial pressure of the species in the mixture.

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

$RH = P/P_{satu}$,

where P is the partial pressure of H_2O in the unsaturated gas, $P < P_{satu}$.

Write the chemical potential of water in the air as

$\mu(T,P) = \mu(T,P_{satu}) + k_BT \log(P/P_{satu})$.

The chemical potential of H_2O in moist air relates to the relative humidity. Our parents are not totally negligent. They do mention humidity.

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.

Chemical equilibrium

Chemical reaction. Consider a reaction

$H_2 + 0.5O_2 = H_2O$

Hydrogen and oxygen form water. Water split into hydrogen and oxygen. In equilibrium, the reactor will have all three species of molecules. How many molecules are in each species?

Denote the *degree of reaction* by ε. For the reaction $H_2 + 0.5O_2 = H_2O$, the degree of reaction ε is the number of hydrogen molecules consumed by the reaction. When the reaction advances by dε, the amounts of the three species of molecules change by

$dn_{H2} = - d\varepsilon$,
$dn_{O2} = - 0.5d\varepsilon$,
$dn_{H2O} = + d\varepsilon$.

The negative sign means that the reaction reduces the amount of each reactant, and the positive sign means the reaction increases the amount of the product. Each stoichiometric coefficient proportions the change of the amount of a species of molecules.

A closed system in a constant-temperature, constant-pressure process. The condition of equilibrium depends on the experimental setup. Here we assume that the reaction takes place in a cylinder-piston setup. The piston is well-sealed, so that no molecules leak in or out. The cylinder is in thermal contact with a thermal reservoir of constant temperature T. A constant weight is placed above the piston, and the piston can slide up and down without friction, so that the pressure in the cylinder is fixed at P. Consequently, the cylinder-piston setup is a closed system, and the reaction is a constant-temperature, constant-pressure process.

BAT on a chemical reaction. We identify an isolated system with internal variables. The isolated system has three parts: the mixture of molecules, the thermal reservoir, the weight. The isolated system has three internal variables: the energy transfers from the mixture to the thermal reservoir, the height of the weight, and the degree of reaction.

The isolated system conserves energy:

$U + PV + U_R = \text{constant}$,

where U is the internal energy of the mixture, PV is the potential energy of the weight, and U_R is the internal energy of the reservoir.

Up to an additive constant, the subset entropy of the isolated system is the Planck function:

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

where S is the entropy of the mixture.

That is, the BAT identifies a shadow of entropy, the Planck function, or equivalently, the Gibbs function, $G = -TY = U + PV - TS$. The isolated system minimizes the Gibbs function to reach equilibrium.

Condition of equilibrium. Let the Gibbs function of the mixture be $G(T, P, n_{H2}, n_{O2}, n_{H2O})$. Recall that

$dG = - SdT + VdP + \mu_{H2}dn_{H2} + \mu_{O2}dn_{O2} + \mu_{H2O}dn_{H2O}$.

When the reaction advances by dε under constant temperature and constant pressure, the Gibbs function changes by

$dG = (- \mu_{H2} - 0.5\mu_{O2} + \mu_{H2O})d\varepsilon$.

The change in the Gibbs function is expressed in terms of the chemical potentials of the three species of molecules in the mixture.

Under constant temperature and constant pressure, the BAT requires that the reaction move in the direction that reduces the Gibbs function. In chemical equilibrium, dG = 0 for any small change in the degree of reaction, dε. Consequently, the reaction reaches equilibrium when

$\mu_{H2} + 0.5\mu_{O2} = \mu_{H2O}$.

This condition of equilibrium is applicable in general, and is not limited to ideal gases. For example, H_2O molecules can be in any phase or any mixture of phases.

Exercise. Run the basic algorithm of thermodynamics on the reaction in the four steps and find the shadow of entropy.

Ideal gas reaction

Ideal gas reaction. A reaction is called an ideal gas reaction if, at each degree of reaction, the molecules in the reactor form an ideal gas mixture.

The merit of being ideal. A mixture of three species of molecules has two independent mole fractions:

$y_1 = n_1/(n_1 + n_2 + n_3)$,
$y_2 = n_2/(n_1 + n_2 + n_3)$.

The chemical potential for each species in the mixture is a function of temperature, pressure, and mole fractions, $\mu_1(T,P,y_1,y_2)$, $\mu_2(T,P,y_1,y_2)$, and $\mu_3(T,P,y_1,y_2)$. This information is rarely available when the mixture is not an ideal gas mixture. Consequently, the condition of equilibrium is hard to apply in general.

By contrast, the chemical potential of each species of molecules in an ideal gas mixture takes the form

$\mu_1(T,P,y_1,y_2) = \mu_1(T,P_0) + RT \log(y_1P/P_0)$.

Here R is the universal gas constant, P_0 is the reference pressure. The dependence on pressure and mole fraction takes a simple form. The reference chemical potential, $\mu_1(T,P_0) = h_1(T) - Ts_1(T,P_0)$, is a function of a single variable, temperature. Both functions, $h_1(T)$ and $s_1(T,P_0)$, are listed in the ideal gas tables for each species of commonly used molecules.

Equilibrium constant. Consider a reaction in an ideal gas. For example, at a high temperature and low pressure, H_2O molecules form an ideal gas. Thus, all species of molecules in the reaction $H_2 + 0.5O_2 = H_2O$ are ideal gases.

The chemical potentials of the three species molecules in the mixture are

$\mu_{H2}(T,P_{H2}) = \mu_{H2}(T,P_0) + RT \log(y_{H2}P/P_0)$,
$\mu_{O2}(T,P_{O2}) = \mu_{O2}(T,P_0) + RT \log(y_{O2}P/P_0)$,
$\mu_{H2O}(T,P_{H2O}) = \mu_{H2O}(T,P_0) + RT \log(y_{H2O}P/P_0)$.

Insert these chemical potentials into the condition of equilibrium, $\mu_{H2} + 0.5\mu_{O2} = \mu_{H2O}$, and write

$(y_{H2O}{}^1)/(y_{H2O}{}^1 y_{O2}{}^{0.5})(P/P_0)^{1-1-0.5} = K$,

where

$RT \log K = \mu_{H2}(T,P_0) + 0.5\mu_{O2}(T,P_0) - \mu_{H2O}(T,P_0)$.

The dimensionless number K is called the *equilibrium constant*. This name is not quite right: K is not a constant, but is a function of temperature, K(T).

The condition of equilibrium of the ideal gas reaction has the following trends.

- When K << 1, the reactants dominate the mixture in equilibrium.
- When K >> 1, the products dominate the mixture in equilibrium.
- When K is on the order of unity, comparable amounts of reactants and products constitute the mixture in equilibrium.
- K is a function of T. At room temperature, H_2O molecules dominate the mixture in equilibrium. As the temperature increases, the proportion of hydrogen and oxygen increases in the mixture in equilibrium.
- The power of pressure is counted using the stoichiometric coefficients: 1 - 1 - 0.5 = - 0.5. The reaction reduces the number of molecules, so that an increase in pressure increases the number of H_2O molecules.

Equilibrium composition. The equilibrium composition depends on temperature and pressure. Let us find this dependence. Let n^0_{H2}, n^0_{O2}, and n^0_{H2O} be the amounts of the three species of molecules in the cylinder before the reaction begins. Assume that we know these numbers. At a degree of reaction ε, the amounts of the three species of molecules become

$n_{H2} = n^0_{H2} - \varepsilon$,
$n_{O2} = n^0_{O2} - 0.5\varepsilon$,
$n_{H2O} = n^0_{H2O} + \varepsilon$.

Inserting the above equations into the condition of equilibrium, we obtain a nonlinear algebraic equation for ε. The equation can be plotted as a function $P(\varepsilon,T)$. Given T and P, this function determines the degree of reaction in equilibrium, ε, as well as the equilibrium composition, y_{H2}, y_{O2}, and y_{H2O}.

Exercise. Use chemical potentials to calculate the equilibrium constant for the reaction $H_2 + 0.5O_2 = H_2O$ at 1000 K.

Exercise. Derive the condition of equilibrium for an ideal gas reaction, $H_2 + 0.5O_2 = H_2O$, in a reactor of constant volume and constant temperature.

Exercise. Derive the condition of equilibrium for an ideal gas reaction, $H_2 + 0.5O_2 = H_2O$, in an adiabatic reactor of constant volume.

Boltzmann distribution

In a separate google doc, the fundamental postulate is used to formulate the basic algorithm of statistical mechanics and derive the Boltzmann distribution. The statistical perspective furthers the understanding of thermodynamics.

Gibbs distribution

A separate google doc derives the Gibbs distribution.

Ideal gas

Results on ideal gas are collected in a separate google doc, Ideal gas