

1 Energy and the first law

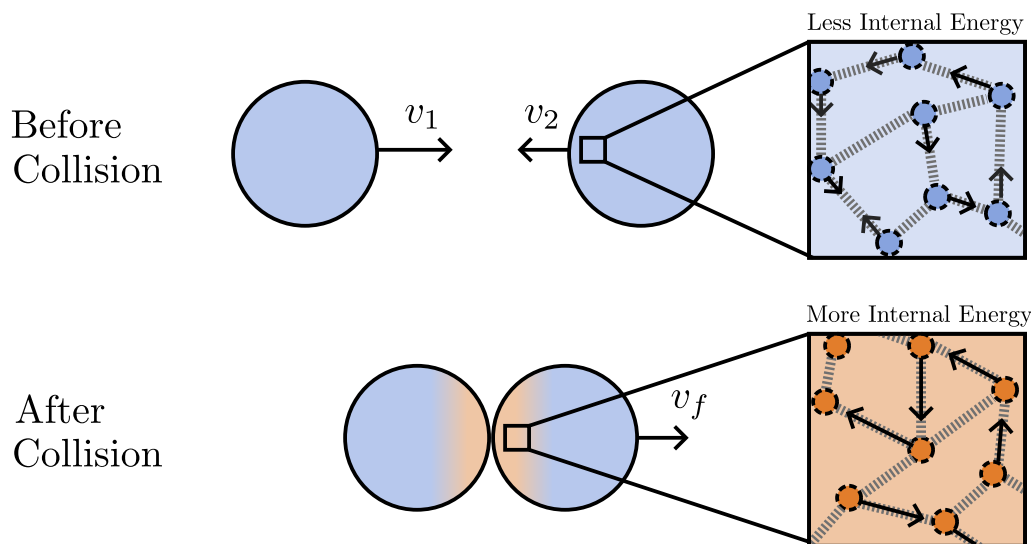


Figure 1: Energy flow in an inelastic collision. The lost kinetic energy turns into **internal** energy. We divide internal energy in two parts: work to deform the balls, and heat energy.

Summary of concepts:

- Partitioning of energy in an inelastic collision: $\sum_i \frac{1}{2} m_i v_{i,initial}^2 + \sum_i U_{i,initial} = \sum_i \frac{1}{2} m_i v_{i,final}^2 + \sum_i U_{i,final}$
- In PHYS 213, we are mainly concerned with changes in internal energy. We ignore all the v's.
- Internal energy changes are broken into two parts $dU = dQ - pdV$. The second term is "work" and represents deformation of the object. The first term is everything else, which goes into the microscopic motion of particles.
- The first law of thermodynamics is energy conservation. Total change in internal energy, is work plus heat.

You might recall that in your classical mechanics courses, you were told about *inelastic* and *elastic* collisions. The difference between them is that *kinetic energy* is conserved in an elastic collision, while it is not in an inelastic collision. Where does the energy go in the inelastic collision, like the one in Figure 1? It goes into the microscopic motion of atoms; they actually move faster after the inelastic collision happens than they were moving previously. We perceive this as an increase in temperature.

Thermodynamics is about the flow of **thermal energy**. Thermal energy is the microscopic kinetic and potential energy of atoms that we can't observe very easily in our day to day lives. This motion is very similar to the mechanics that you've already studied in physics, but it involves the motion of $\sim 10^{23}$ atoms for a roughly human-sized object. One could try to keep track of the velocities and positions of all these atoms,¹ but most of us would quickly run out of patience trying to do that. We'd prefer to instead use *collective* descriptions of the atoms, like how much total energy is in their motion, and how compressed they are. In future chapters, we will introduce other collective variables, like temperature and entropy. These collective variables are called **macrostates**.

¹It turns out that you can do computer simulations on millions of atoms which actually do act very similarly to 10^{23} particles; this is an entire field of research!

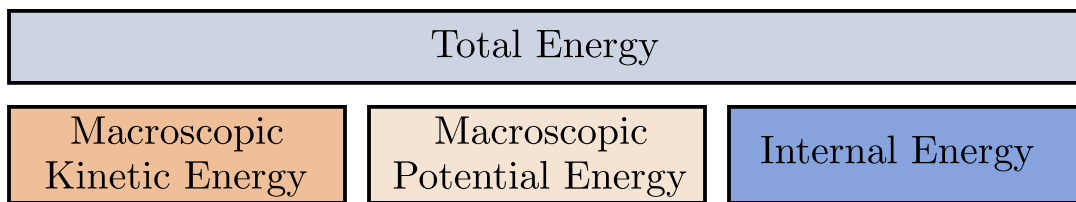


Figure 2: The total energy of a system is divided into macroscopic kinetic energy, macroscopic potential energy, and internal energy.

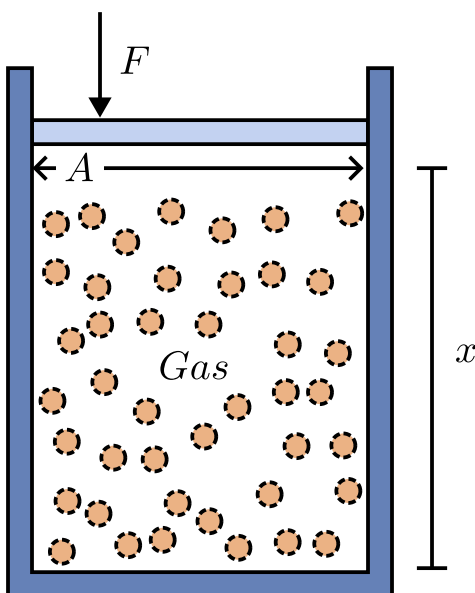


Figure 3: The relation of pdV to $Fd\ell$. The pressure is the force per area.

1.1 Internal energy

Internal energy is all the energy that is not macroscopic kinetic and potential energy, which is present in all the microscopic motions inside a given object. We typically write internal energy as U . A rough idea of the relationship between all the different energies is present in Figure 2. In reality, the internal energy is actually much larger than the macroscopic potential and kinetic energy. Energy can move between the different classifications. In this course, we will focus on the behavior of just the internal energy.

It is often very difficult to compute or measure the total value of the internal energy U . However, as you may recall from classical mechanics, all that matters for physics is not the total value of the energy, but the *change* in energy. So in thermodynamics, we typically compute only the change in internal energy during a process.

Example: Suppose that two balls collide inelastically on a flat plane (so potential energy is zero). The total kinetic energy before the collision is 40 J, while after the collision, the balls continue on but with only 20 J of kinetic energy. The internal energy of the balls increases by 20 J, the remaining energy; that is, $\Delta U = U_f - U_i = 20$ J. During this process, we know the change in U but not its total value.

1.2 Work

The balls in the inelastic collision may become damaged or deformed. It takes some energy to do this deformation, which is work. In other systems such as gas in a balloon, it takes energy to compress the

balloon. As we learned in classical mechanics, the work we do **on** an object is

$$W_{on} = \int_{x_i}^{x_f} F dx, \quad (1)$$

where we have applied force F over a distance from an initial position x_i to a final position x_f . In thermodynamics we will consider work on the object as changing its total internal energy.

To understand work a little more concretely, consider a piston (Figure 3) experiment. We will use these often in the course, so it's worth getting used to the setup. The chamber is filled with a gas and sealed, so no gas can get in or out. The sides and bottom are fixed in place. The top of the piston is allowed to move up and down. We will identify the position of the top of the piston as x from Equation 1.

Now suppose we apply a force F to the top of the piston. Using Equation 1, the work on the gas is

$$W_{on} = \int_{x_i}^{x_f} F dx \quad (2)$$

$$= \int_{x_i}^{x_f} \frac{F}{A} A dx \quad (3)$$

$$= \int_{V_i}^{V_f} \frac{F}{A} dV \quad (4)$$

$$, \quad (5)$$

where in the last step we used “u-substitution” to replace our integration variable with volume $V = Ax$. Finally, we substitute pressure $p = F/A$ (Pressure always pushes outward!) to get our final result

$$W_{on} = - \int_{V_i}^{V_f} p dV. \quad (6)$$

In thermodynamics we typically think of collective variables volume and pressure, rather than force on a given point and specific dimensions of the object. Writing work in this form allows us to use those collective variables.

Work on versus work by. In the previous paragraph, we imagined deforming a gas by doing work on it. If the gas were expanding, we might prefer to say that the gas itself is *doing* work. In thermodynamics, that is referred to as work **by** the gas. Very simply, $W_{on} = -W_{by}$ from Newton's third law, but it's very important to keep them straight. A simple mnemonic is that if the volume of the object is increasing, it is doing work and so W_{by} is positive, and vice versa if the volume is decreasing.

1.3 Heat

Heat energy is all the energy that enters the object that is not macroscopic kinetic energy (which for the most part we ignore in thermodynamics) or work. We often call this Q . For the most part this is the microscopic vibrations of bonds between atoms, and in the case of gases, the velocity of molecules/atoms. Later on in this course, we will explicitly count up the heat energy for an idealized gas.

Note that both heat and work are defined only for changes over a process. There is no such thing as “total heat” or “total work” independent of a process. On the other hand, in principle there is a total internal energy of a given object.

1.4 Differentials and the first law

The first law of thermodynamics really just follows from the definition of heat and work. It says that

$$\Delta U = W_{on} + Q. \quad (7)$$

In other words, the change in internal energy during a process is the work done on it plus the heat energy that flows in. This is true by definition of heat energy, since we defined heat energy as all the energy that wasn't covered by work! However, the law is not usually written quite like the above; it is usually written in terms of differentials.

To write the first law in terms of differentials, note that

$$U_f - U_i = \int_{U_i}^{U_f} dU = \int_0^Q dQ - \int_{V_i}^{V_f} p dV. \quad (8)$$

This is a very useful equation that you will use throughout the course. Note that the integrals are important here; any dependence of the pressure on the volume must be included. In thermodynamics, often this equation will be written as

$$dU = dQ - p dV, \quad (9)$$

where it is implied that the equation must be integrated over a process to get a change in internal energy.

2 Energy and temperature

The purpose of this chapter is to explore how we describe heat energy and its spontaneous flow. We don't really have a principle to explain **why** energy flows from hot to cold yet. For that we'll need entropy, which will be covered in the later chapters.

Summary of concepts:

- If two objects are placed into thermal contact, heat flows from high temperature to low temperature until the temperatures are the same.
- Heat conduction – how fast heat flows.
- Heat capacity – how much heat energy it takes to change the temperature.
- Equipartition – gives us a way to estimate heat capacity for a wide class of materials.

2.1 Temperature

Temperature is the propensity for an object to give up its heat energy. Heat spontaneously flows from higher temperature to lower temperature. In this course, it is almost always measured in Kelvin. It is recommended to always convert to Kelvin before performing any thermodynamic calculations.

2.2 Heat capacity

Heat capacity is a measure of how much heat it takes to change the temperature.

$$C = \frac{\partial Q}{\partial T} \quad (1)$$

You can imagine an experiment as follows: you put a small amount of heat into an object, and measure how much the temperature changed. If the temperature changes a lot, then the object has small heat capacity, while if it changes a little, the object has large heat capacity. The units in SI are J/K.

Specific heat capacity Particularly in chemistry it is useful to know the heat capacity *per mass* of a substance. To measure the specific heat of a substance, you just divide the heat capacity by the mass of the substance. The units in SI are J/K kg.

Molar specific heat capacity As it turns out, the heat capacity per atom is often fairly universal—at high temperatures most solids have the same heat capacity per atom. Typically, though, we work with moles of atoms. One mol is N_A atoms, where N_A is Avagadro's number. To measure the molar specific heat, you just divide the heat capacity by the number of moles in the object. If you know the atomic mass of the atoms, you can get that from the mass. For example, if an atom has an atomic mass of 40, then that means that its mass is 40 g/mol.

Heat capacity at constant volume If the volume is constant, that means that $dV = 0$, so all changes in the internal energy are due to heat. Therefore

$$C_V = \frac{dQ}{dT} = \frac{dU + pdV}{dT} = \frac{dU}{dT} \quad (2)$$

where the V subscript means constant volume.

Heat capacity at constant pressure If the pressure is constant, then

$$C_p = \frac{dQ}{dT} = \frac{dU + pdV}{dT} = \frac{dU}{dT} + p \frac{dV}{dT}, \quad (3)$$

where the p subscript means constant pressure. To compute this, one needs to know $V(T)$.

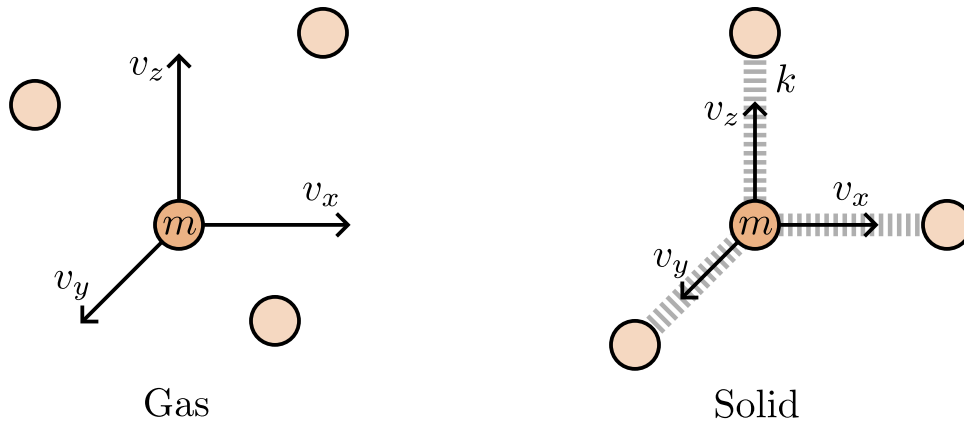


Figure 1: An atom of gas has a total energy equal to its kinetic energy $E_g = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2)$. An atom in the solid has kinetic energy *and* potential energy from its bonds with other atoms, for a total energy $E_s = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) + \frac{1}{2}k(x^2 + y^2 + z^2)$. The gas's energy is proportional to 3 squared quantities, while the solid's energy is proportional to 6 squared quantities. Each quadratic degree of freedom corresponds to one squared quantity.

2.3 Equipartition hypothesis

The equipartition hypothesis is not always true, but later in the class we will learn some more about why it works as well as it does. The statement is the following: for every **quadratic** degree of freedom, the internal energy per atom is $\frac{1}{2}k_B T + C$, where k_B is the Boltzmann constant and C is an unknown constant (remember that a constant offset of energy doesn't change physics). Each direction of travel or rotational mode counts as one degree of freedom from the kinetic energy, and each direction of vibration counts as two degrees of freedom from the kinetic and potential energy.

Example: monatomic gases Consider a monatomic gas (for example, He) in which the atoms are free to whiz around. Each atom can go in three directions, and the kinetic energy is given by $\frac{1}{2}mv^2$, which is quadratic. Therefore, the internal energy of a gas of N He atoms is $\frac{3}{2}Nk_B T + C$. Alternatively, the total internal energy of a gas of n moles of He atoms is $\frac{3}{2}nRT + C$, where R is the ideal gas constant. So the molar heat capacity at constant volume of a monatomic gas is very simply $\frac{3}{2}R$.

Example: diatomic gases In this case, we have two-atom molecules in the gas. This applies to air, which is mostly N_2 and O_2 . In this case, the number of degrees of freedom is 5, because the molecules now have two extra rotational degrees of freedom. It is two because rotation around axis of the molecule requires a lot of energy due to quantum mechanics. So the internal energy is $\frac{5}{2}Nk_B T + C$.

Example: solids In solids the atoms cannot move freely, but can vibrate against one another. Each atom has three directions in which it can vibrate (x,y,z), so that is a total of 6 degrees of freedom per atom. So the molar heat capacity at constant volume of a solid is $3R$.

The equipartition hypothesis is very useful since it allows us to estimate the heat capacity for many diverse objects! We should note that it becomes inaccurate at very low temperatures due to quantum mechanical effects; however, for room temperatures it is often very good.

2.4 Heat conduction

For the most part, this course deals with equilibrium, and not so much in how systems come into equilibrium. However, there is a simple model for heat conduction that is worth including. Conduction of heat can be approximately described using *heat conductivity*, which is very much like electrical conductivity. In this analogy, potential difference (voltage) corresponds to the difference in temperature and current is heat. Usually the symbol κ is used for the conductivity of a given material and the SI units are W/m K.

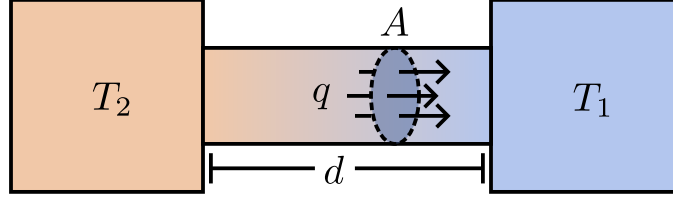


Figure 2: Heat flows from an object at a higher temperature T_2 to an object at a lower temperature T_1 . The heat flux q through an area A depends on the difference in temperature and the distance between the two objects according to Equation 4.

The total heat transport per second Q (W) is given by

$$q = -\frac{\kappa A}{d}(T_2 - T_1) \quad (4)$$

where d is the length of the connection, and κ is the thermal conductivity of the material. It is useful to define the conductance of a given object with cross-sectional area A and length d as $k = \frac{\kappa A}{d}$

If there are multiple paths with conductance k_1 and k_2 , then they combine very much like electrical conductivity. For example, for two objects in series,

$$k_{tot} = \frac{1}{\frac{1}{k_1} + \frac{1}{k_2}}. \quad (5)$$

And for parallel:

$$k_{tot} = k_1 + k_2. \quad (6)$$

Sometimes the thermal resistance is used $R = 1/k$, which helps as a mnemonic for the serial and parallel formulae: in serial the resistance is summed, while in parallel the conductance is summed.

2.5 Managing all the variables in thermodynamics.

In thermodynamics, we have many macrostate variables to keep track of. Already in this course, we've seen $T, p, V, C, U, W_{on}, Q, N$. You may have noticed that we often don't indicate that each of these variables can actually also be dependent on each of the others. For example, pressure can be written as a function of number, volume, and temperature $p(N, V, T)$. Indeed, that is precisely what the comforting ideal gas law does for us; for an ideal gas, $p = Nk_B T/V$. Other systems might have different relations between the macrostate variables. At the same time, one can write volume V as a function of N, p, T . The equipartition hypothesis allows us to write internal energy as a function of T : $U = \frac{N_{DOF}}{2} Nk_B T + C$.

For these problems a general strategy is to:

- Identify what is constant and what is changing. (example: constant volume, but temperature is changing as in the blocks example)
- Find the equilibrium condition (example: $T_1 = T_2$)
- Use the first law to compute changes in internal energy. This will require integration; make sure that all integrals are written in terms of constants and the integration variable.
- Do all integrals to find changes in the macrostate variables.

This strategy works for almost all problems in this course.

3 Entropy, microstates, macrostates, and the second law

- We describe a thermodynamic system using macrostates (volume, temperature, energy, number of particles)
- A complete (classical mechanics) description would specify the position and velocity of every single particle which is called the microstate. However, in a solid material this is not possible to do.
- Entropy tells us how much information we throw away when we use a macrostate. $S = k \ln \Omega$.
- Entropy is related to the probability of the macrostate; the highest entropy macrostate is the most probable by a very large margin.
- The above leads to the second law of thermodynamics: $S_f - S_i \geq 0$ for all processes, where S_i is the initial entropy and S_f is the final entropy.

3.1 Microstates and macrostates

We have already discussed macrostate variables in this course. They are collective quantities about an object, such as the total volume, internal energy, pressure, and temperature. On the other hand, to specify the **microstates** of a system, one needs to specify all the positions and velocities of all the atoms in the object. Obviously this is far too much information to track for any real system explicitly. Instead, we just want to keep count of how much information we lose by using macrostate variables only. Entropy is a measure of the information that we lose when we describe the system in terms of macrostate variables.

3.2 Definition of entropy

For example, suppose we describe our system in terms of the internal energy, volume, and number of particles U, V, N . Then the number of microstates for that value of U, V, N is called $\Omega(U, V, N)$, and the entropy is defined¹ as

$$S(U, V, N) = k_B \ln \Omega(U, V, N). \quad (1)$$

(aside) You might reasonably ask why this particular definition. Entropy is additive with this definition, while Ω is multiplicative (two dice have 36 total states, not 6). By making it \ln instead of some other base, the definitions of pressure and temperature which we will see later will work better. Otherwise there would be stray values around when derivatives and integrals are taken. So entropy is defined this way so that it has nice properties that we will explore throughout this course.

3.3 Entropy of flipping coins

Let's examine a toy system to get a handle on the concept of entropy. We all have an intuitive understanding of entropy already. For example, if you walked into a room and someone told you that they flipped a coin 1000 times and got 1000 heads, you would be understandably skeptical. However, if they said they got 522 heads, you might believe them.² This is an intuitive understanding of entropy! There are many more ways that one could get 522 heads in 1000 flips than 1000 heads in 1000 flips, and so it is more likely, and therefore higher entropy. Let's now analyze this situation in the language of thermodynamics.

First things first, we should identify the microstates and macrostates. For a sequence of N coin flips, the microstate is the particular sequence. For example, for $N = 5$, the sequence HHTHT is a different microstate than the sequence HTHHT, even though they have the same number of heads. The macrostate in our case will be the *fraction* of flips that came up heads. Both of the above microstates "belong" to the same macrostate of $N_H = 3$, $N = 5$, or $F_H = 3/5$.

¹Technically this is the statistical mechanics definition of entropy. Next time we will get to the thermodynamical definition of entropy. I think that the statistical mechanics definition makes a lot more sense!

²Other than questioning why they are flipping coins in a room by themselves!

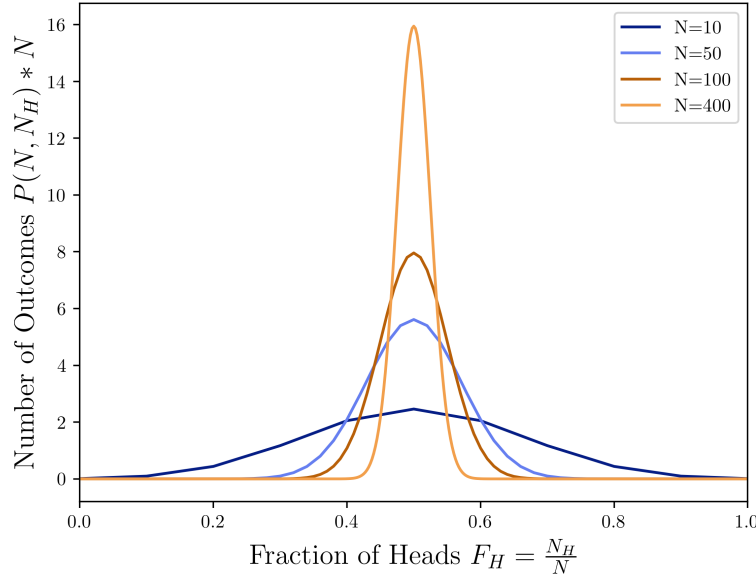


Figure 1: The probability of getting K outcomes out of N total events is given by $P(N, K) = \binom{N}{N_H} P^K (1 - P)^{N-K}$, where P is the probability of K . For our coin toss the odds of getting a heads is 50%, therefore $P = \frac{1}{2}$ and we write $K = N_H$. The plot shows $P(N, N_H)$ vs $F_H = N_H/N$ for different values of N ; the distribution of results gets closer and closer to the highest probability result as N increases.

To construct the entropy $S(N_H, N)$, we need to be able to compute $\Omega(N_H, N)$, how many ways one could flip N_H heads with N flips. This is given by the binomial coefficient

$$\Omega(N_H, N) = \binom{N}{N_H} = \frac{N!}{N_H!(N - N_H)!}. \quad (2)$$

Often this function is available on calculators as a “choose” function. It is also available in Python as `scipy.special.binom`.

In Fig 3.3, we show the number of coin toss sequences (microstates) with a given fraction of heads F_H (macrostate). The number of outcomes with F_H near 0.5 grows very quickly as N increases. As $N \rightarrow \infty$, with almost probability 1, the fraction of heads is 0.5. Therefore, for very large numbers of flips, the most likely macrostate is just the one that occurs.

Some physical systems actually do act as if they are coin flips. In the case of magnets, the magnetism is given by very small magnetic moments called spins³ which align at low temperatures to create ferromagnets. In this case, whether the spin is pointed up or down is equivalent to the coin flips.

3.4 Entropy of free particles in a volume

Consider the number of microstates $\Omega(U, V, 1)$ of a single atom in a volume V and total energy U . The atom could be in any location in the volume, and have velocity pointing in an direction in 3D. We will not try to compute Ω exactly, but for one atom, it is *proportional* to the volume: $\Omega(U, V, 1) = g(U, 1)V$, where g is an unknown function of U and N . Now, if there are N particles that we can approximate as if they don't interact (i.e., an ideal gas!),

$$\Omega(U, V, N) = (g(U, N)V)^N \quad (3)$$

³We will explore this later in the Boltzmann sections!

and

$$S(U, V, N) = k \ln \Omega = Nk \ln V + Nk \ln g. \quad (4)$$

3.5 Second law of thermodynamics

Given the definition of entropy in terms of the number of states, the second law becomes almost trivial. It simply says that once there are many particles in a system, the most likely value of macrostate variables is the one that happens. This is not because the other values are impossible (they are technically not), but because the most likely values are **so much more likely** that it is effectively impossible to see the unlikely occurrence. Imagine flipping 10^{23} coins; the fraction $f_H = N_H/N$ is extremely unlikely to be very far from 0.5. Thus, the **second law of thermodynamics**: in an isolated system, the entropy always stays the same or increases, or

$$S_f - S_i \geq 0, \quad (5)$$

where S_f is the final entropy and S_i is the initial entropy.

3.6 Internal equilibrium

For almost all cases in this course (we will let you know if this is not the case), we will consider systems internally in equilibrium. Internal equilibrium means that for the given macrostate variables (often U, V, N for example) entropy has been maximized already, keeping those macrostate variables fixed. That means that the objects, if left to their own devices, will not spontaneously undergo any changes. Most processes that we will consider in this course will be *quasistatic* processes, which means that the processes are slow enough that objects are in internal equilibrium the entire time. If a system is in internal equilibrium, we can use all of our equilibrium relations like the ideal gas law.

Given a system where the volume (V) and number of particles (N) do not change, how do we calculate a change in entropy (S) or energy (U)?

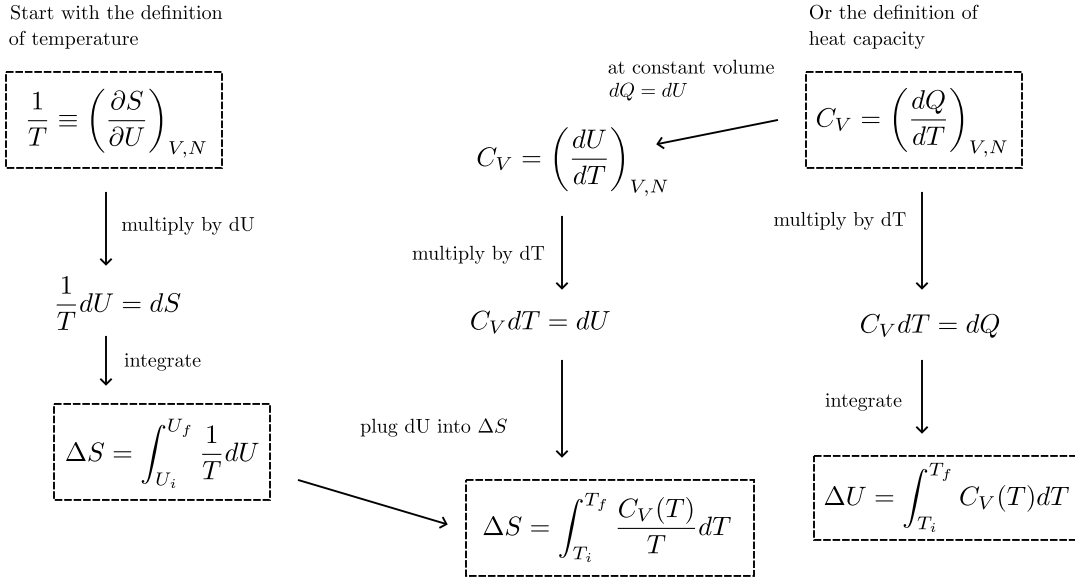


Figure 1: Entropy, energy, heat capacity, and temperature are all connected to one another through derivatives and integrals.

4 Entropy of physical systems

- Be able to compute the entropy change as two bricks come into equilibrium.
- Definition of temperature in terms of entropy: $\frac{1}{T} = \frac{dS}{dU}$
- Pressure in terms of entropy: $\frac{p}{T} = \frac{dS}{dV}$
- Fundamental relation: $dS = \frac{dQ}{T}$
- Heat capacity and entropy: $\Delta S = \int \frac{C}{T} dT$

4.1 The “two bricks” model

For this section, we will consider a very simple model of two bricks that are in internal equilibrium, but different temperatures. It doesn't matter that they are bricks! They could be any objects. We will explore how maximizing entropy leads us to a thermodynamic definition of temperature, and compute how much entropy increases as the two bricks come into equilibrium with each other.

4.2 Definition of temperature

The total entropy of the two-brick system is

$$S_{tot} = S_1(U_1) + S_2(U_2). \quad (1)$$

We are assuming that the volume and number of particles in the bricks stays the same, so we are only going to consider the dependence on internal energy. As we determined in the previous section, spontaneous processes stop happening when the entropy is maximized. Also due to conservation of energy, we know that $U_1 + U_2 = C$, where C is some constant.

To maximize entropy, we take the derivative and set it equal to zero:

$$\frac{dS_1}{dU_1} + \frac{dS_2}{dU_1} = 0. \quad (2)$$

Note that we had to take the derivative of the entropy of both bricks with respect to U_1 . However, we have the function $S(U_2)$. Using conservation of energy, we can find that $\frac{dU_1}{dU_2} = -1$. Using the chain rule, we have

$$\frac{dS_2}{dU_1} = \frac{dS_2}{dU_2} \frac{dU_2}{dU_1} = -\frac{dS_2}{dU_2}. \quad (3)$$

This makes some sense in retrospect; an increase in U_1 is a decrease in U_2 , so the derivatives should be opposite sign.

So finally we can find the **equilibrium relation**:

$$\frac{dS_1}{dU_1} = \frac{dS_2}{dU_2}. \quad (4)$$

Note that entropy has units of J/K and internal energy has units of J, so the derivative has units of 1/K. This is very suggestive, and indeed it turns out that if you define

$$\frac{dS_1}{dU_1} \equiv \frac{1}{T_1}, \quad (5)$$

then we get the same behavior as we noted empirically in this class—energy spontaneously flows from high temperature to low temperature until the two bricks are the same temperature.

The reason that maximization of entropy is such an important concept is that it allows us to derive many other **equilibrium conditions** like $T_1 = T_2$.

4.3 Fundamental relation for constant volume

Now we'll derive a very useful relationship for the change in entropy during a process. If during the process, everything is held constant other than the internal energy U , then using the chain rule:

$$\frac{dS}{dt} = \frac{dS}{dU} \frac{dU}{dt} = \frac{1}{T} \frac{dU}{dt}. \quad (6)$$

Here we used a time derivative to give a sense of a process, but often in thermodynamics we are lazy and instead write the same equation as

$$dS = \frac{1}{T} dU \quad (7)$$

4.4 Definition of pressure in terms of entropy

We can define pressure similarly to temperature. Consider a chamber with a movable membrane as noted in Fig 4.4. We can do the same trick as with energy to get

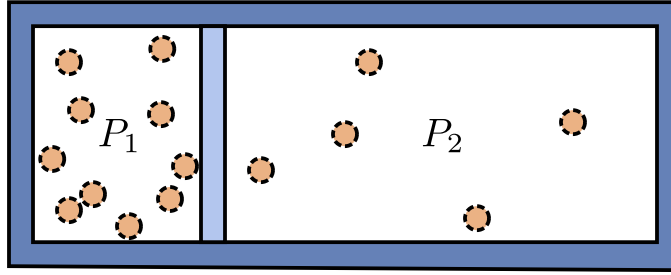
$$\frac{dS_1}{dV_1} = \frac{dS_2}{dV_2}. \quad (8)$$

The units of $\frac{dS}{dV}$ are J/K m³, which doesn't look like much of anything, but recall that J = N m, which means the units are also N/m² K. Pressure has units of force per area, and temperature has units of Kelvin. So one might guess that

$$\frac{dS}{dV} \equiv \frac{p}{T} \quad (9)$$

in analogy to temperature and energy.

System is not in Equilibrium $P_1 \neq P_2$



System is in Equilibrium: $P'_1 = P'_2$

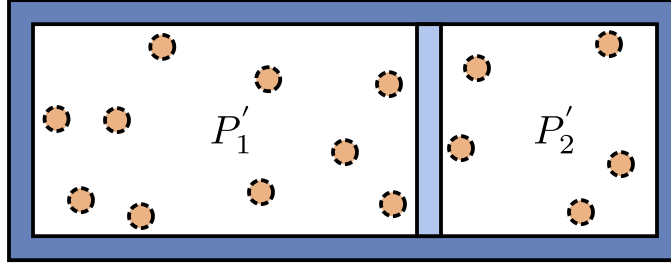


Figure 2: Consider a chamber with a free-moving partition between its two sides. The volume of each side changes as the partition slides back and forth. This system is in equilibrium when the wall between the two sides stops moving, or (equivalently) when the pressure on one side is equal to the pressure on the other. This condition maximizes the total entropy.

4.5 Fundamental relation for changing volume and energy

Now we can compute the change in entropy for any quasistatic process (even if both pressure and volume are changing) by using the equivalent of the chain rule for multiple dimensions:

$$dS = \frac{1}{T}dU + \frac{p}{T}dV = \frac{1}{T}dQ, \quad (10)$$

where we used the first law to substitute $dQ = dU + pdV$. This is the most general version of the fundamental relation, so long as the number of particles in the system is not changing. This was actually the original definition of entropy when thermodynamics was being developed. In the modern understanding, this relation is a result of the second law of thermodynamics, the definitions of temperature and pressure in terms of entropy, and some calculus.

4.6 Relationship between heat capacity and entropy

As we've mentioned before, it is often very difficult to measure the total entropy and total internal energy of an object. It actually is very easy to measure the heat capacity; simply transfer a known amount of heat Q into the object, and measure how much the temperature changes. Then $C = \frac{dQ}{dT} \simeq \frac{Q}{\Delta T}$, if the change in temperature is small enough. We can then compute the entropy change as follows:

$$\Delta S = \int dS = \int \frac{1}{T}dQ. \quad (11)$$

Remember that the temperature T changes as we put heat into the material, though, so we cannot just pull the temperature out of the integral. We have to either change our integration variable so we are integrating over T , or write T as a function of Q . The easier path is to change the integration variable from Q to T .

We do that using the definition of heat capacity: $C = \frac{dQ}{dT}$, so $dQ = C dT$. Then

$$\Delta S = \int_{T_i}^{T_f} \frac{C}{T} dT. \quad (12)$$

Keep in mind that sometimes the heat capacity can depend on temperature! If it does then you will need to include that dependence when integrating.

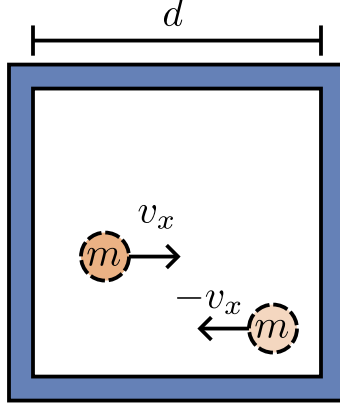


Figure 1: When the atom hits the wall of the box and bounces back with the same speed, its change in velocity is $-2v_x$

5 Kinetic theory of the ideal gas

- The ideal gas is one of the few cases where the connection between microstates and macrostates can be made explicitly.
- The pressure is connected to the average velocity of the gas molecules.
- The equipartition assumption and independence of molecules leads to the ideal gas law.
- More generically, the definition of temperature in terms of entropy can be used to derive other non-ideal gas laws.
- The ideal gas law is a relation that is true for a gas in internal equilibrium and only if the molecules are sufficiently diffuse so that they don't interact with each other too much.

5.1 Pressure from microscopic physics

Consider a single atom with mass m and velocity v_x in the x -direction. It is trapped in a box with reflecting walls a distance d apart, which means that when the atom hits the side of the box, it reverses its direction and heads the other way. If you recall from classical mechanics, the force imparted by such a collision is the change in the momentum *per time*.¹ So the average force imparted by a collision is

$$F_x = \frac{2mv_x}{t_0}, \quad (1)$$

where t_0 is how often the collision happens. t_0 is the time the atom takes to do one round trip in the box, which is $2d/v_x$. Therefore, the pressure on the sides of the box is

$$p = \frac{F_x}{A} = \frac{2mv_x^2}{2d \cdot d^2} = \frac{mv_x^2}{d^3} = \frac{mv_x^2}{V}, \quad (2)$$

because the area of the side of the box is d^2 and the volume of the box is $V = d^3$.

The kinetic energy is given by $\frac{1}{2}m|v|^2$, so we really want the total velocity of the atom $|v|^2 = v_x^2 + v_y^2 + v_z^2$. On average, the velocity is the same in every direction, so

$$\langle |v|^2 \rangle = 3\langle v_x^2 \rangle, \quad (3)$$

¹You can see this by rewriting $F = ma$ as $F = \frac{dp}{dt}$, where p is the momentum

where $\langle \rangle$ means to take the average of whatever's inside. That means that the average velocity squared is three times the average velocity squared in any given direction.

Combining Eqns 3 and 2, we get

$$p = \frac{m\langle |v|^2 \rangle}{3V} \quad (4)$$

for a single atom in a box. Now comes the **ideal gas approximation**—we assume that for N atoms,

$$pV = \frac{Nm\langle |v|^2 \rangle}{3}; \quad (5)$$

that is, the atoms don't interfere with each other. This equation relates a microscopic quantity on the right (v) with two macroscopic quantities on the left (pV).

5.2 Using equipartition to find the ideal gas law

For an ideal gas, the average kinetic energy is the internal energy. The ideal gas approximation says that the atoms don't interact with each other, so there is no potential energy in the system. We can use equipartition to say for a gas made of single atoms (monatomic gas):

$$U = \frac{3}{2}NkT = \frac{N}{2}m\langle |v|^2 \rangle. \quad (6)$$

It turns out that the constant normally in equipartition is actually zero in this case.

Combining Eqns 5 and 6, you can confirm that

$$pV = NkT, \quad (7)$$

the famous ideal gas law.

5.3 Using entropy to find the ideal gas law

We can also use the principle of maximum entropy to derive the ideal gas law. Given the macrostates of U , V , and N , we need to compute the entropy of the system. We use the previous result that $\Omega(U, V, N) = (g(U)V)^N$, where g is an unknown function of U and N .

For more N , we use the ideal gas approximation to determine that $\Omega(U, V, N) = \Omega(U, V, 1)^N$. Using the definition of entropy and that $\ln AB = \ln A + \ln B$, we get:

$$S(U, V, N) = Nk \ln(V) + f(U, N), \quad (8)$$

where f is an unknown function of U and N equal to $Nk \ln g(U)$. The important thing to know about f is that it doesn't depend on V .

Going back to Unit 4, we can use the definition of pressure in terms of entropy:

$$\frac{\partial S}{\partial V} = \frac{p}{T} = \frac{Nk}{V}, \quad (9)$$

which easily gives us the ideal gas law.

This derivation is particularly useful because it applies to any system in which we can compute the entropy! Other systems than ideal gases have different relationships between pressure, volume, and temperature.

5.4 Limitations of the ideal gas law

As you might imagine from the name, the ideal gas law is an idealization; the approximation that the atoms don't interact with each other is not precisely correct. Usually the approximation gets better as the density (number of atoms per volume) gets smaller, so on average each atom is very far from the others. Using the techniques in this chapter, it is possible to derive alternative gas laws that are somewhat more realistic, if a little bit more unwieldy.

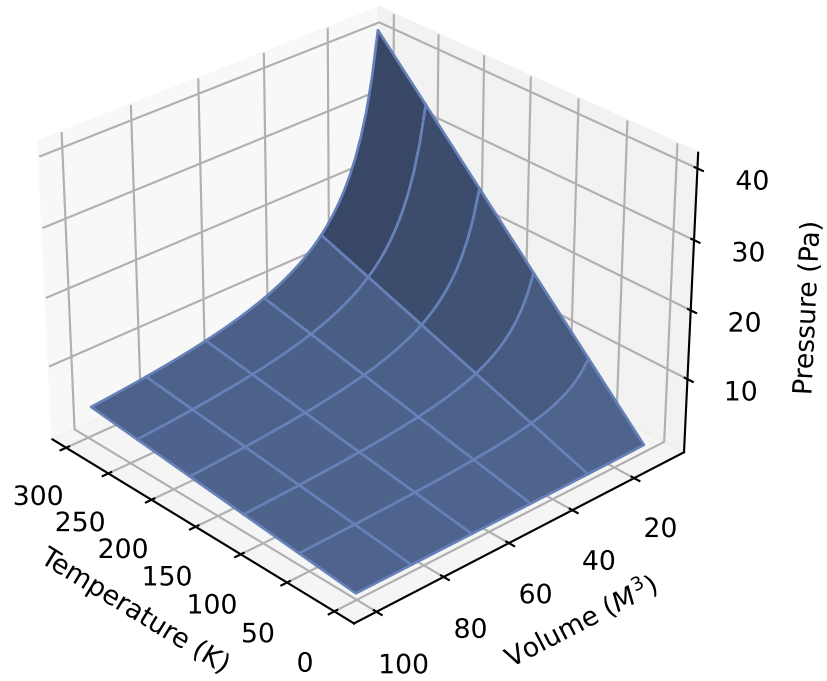


Figure 2: $pV = NkT$ means that when the gas is in internal equilibrium, the relation holds. This is a relationship between the macrostate variables, called an equation of state. If the system is in internal equilibrium then the state exists somewhere on this surface. Each light blue line on the surface represents a process where one variable is held constant. For example, the line at $T=50$ K shows the relationship between volume and pressure at 50 K.

6 Quasistatic thermodynamic processes

- Quasistatic processes are processes that proceed slowly enough that the system is in internal equilibrium.
- p-V diagrams are a useful way to visualize quasistatic processes.
- You should be able to compute change in internal energy, work done, and changes in entropy over a quasistatic process.
- Reversible vs irreversible processes: reversible processes don't change the total entropy, so $\Delta S = 0$.

6.1 Definition of quasistatic processes

Quasistatic processes are processes that proceed slowly enough that the system is in internal equilibrium. That is, given its macrostate variables such as U, V, N , its entropy has been maximized. The upshot of quasistatic processes is that at all times during the process, we can use our equilibrium conditions. That means that for an ideal gas, we can use the ideal gas law, we can use equipartition, and we can use the first (and second) law of thermodynamics, and the fundamental relation. If the system were changed too quickly then all of our equations from the first few chapters would not be applicable, and we would only be able to use overall conservation of energy, and other much more limited tools. Limiting ourselves to quasistatic processes allows us to compute quite a lot about systems.

6.2 Defining a process

Consider an ideal gas which is enclosed, so that its total number N is constant. We know that the ideal gas law holds so $pV = NkT$. There are many possible values of p, V, T that could satisfy this equation; however, every quasistatic process of an ideal gas must satisfy $pV = NkT$. We thus have three variables and one constraint. To define a process we need one more constraint. What follows is some of the common processes.

Isobaric process In this process the additional constraint is that $p = \text{constant}$. In the name, "iso" means same, and "baric" means relating to pressure. An example of an isobaric process might be to heat a gas from T_i to T_f in a piston which is allowed to expand freely. Then as the temperature increases, the volume is given by $V = \frac{Nk}{p}T$, where $\frac{Nk}{p}$ is constant.

Isochoric process In this process, the volume is constant. As you might imagine, "choric" is Greek having to do with volume. An example of an isochoric process might be to heat a gas from T_i to T_f in a container that cannot expand at all. Then as the temperature increases, the pressure is given by $p = \frac{Nk}{V}T$, where now $\frac{Nk}{V}$ is constant.

Isothermal process Shockingly, in this process, the temperature is constant. An example of an isothermal process might be to put a gas into a piston, and slowly expand the piston from V_i to V_f so the gas remains at the same temperature as the surroundings. Then, as the volume increases, the pressure is given by $p = \frac{NkT}{V}$, where NkT is constant.

Adiabatic process This process is a little different from the others; in this process the heat transfer dQ is zero. An example of an adiabatic process might be to put a gas into a piston, and quickly expand the piston from V_i to V_f so that no heat has a chance to flow into or out of the gas. For an ideal gas the derivation is a little tedious but,

$$pV^\gamma = \text{constant}, \quad (1)$$

where $\gamma = 5/3$ for a monatomic ideal gas like He, and $7/5$ for a diatomic ideal gas like N_2 , O_2 (which make up air).

6.3 Computing work done during a process

$$W_{by} = \int p dV. \quad (2)$$

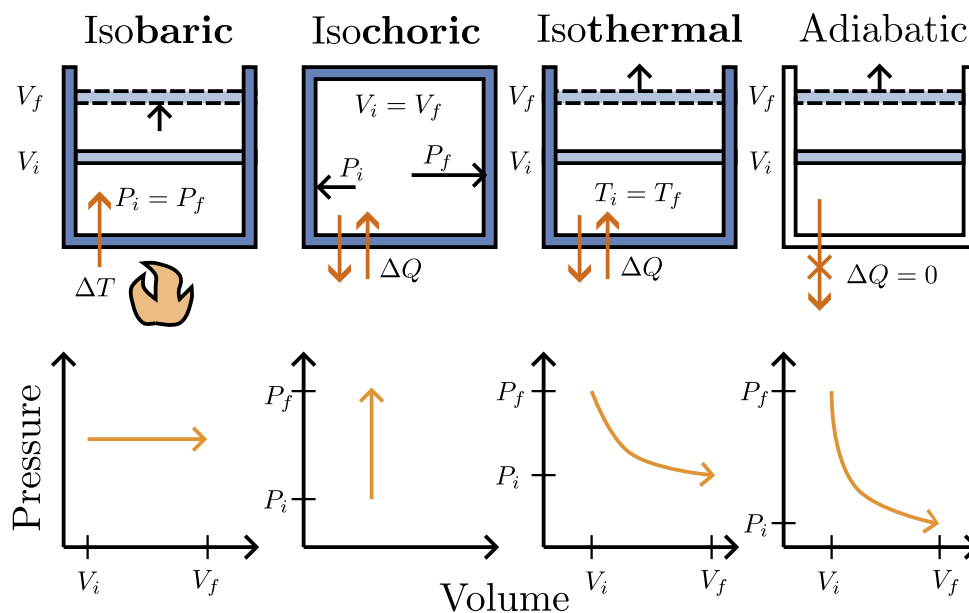


Figure 1: $p - V$ diagrams for common processes.

The general strategy is to write p as a function only of V and then do the integral.

For example, suppose we consider an isothermal process, in which a monatomic gas at high pressure is allowed to slowly push a piston out from V_i to V_f , which then rotates a crank and moves a car forward. Then $p(V) = \frac{NkT}{V}$, and the total work done by the gas is

$$W_{by} = \int_{V_i}^{V_f} \frac{NkT}{V} dV = NkT \ln \frac{V_f}{V_i}. \quad (3)$$

On the other hand, if the expansion is done suddenly so no heat could come in or out, that would be an adiabatic process and

$$W_{by} = \int_{V_i}^{V_f} \frac{C}{V^\gamma} dV = -\frac{C}{\gamma - 1} \left(\frac{1}{V_f^{\gamma-1}} - \frac{1}{V_i^{\gamma-1}} \right), \quad (4)$$

since $\gamma > 1$, where typically we will find the constant C from the initial conditions by using $p_i V_i^\gamma = C$.

The cases of isochoric and isobaric processes are much simpler than adiabatic and isothermal; try to apply the principles above to those cases.

6.4 p-V diagrams

As we saw in the previous section, for a given process it is very helpful to know $p(V)$, pressure as a function of volume, for a given process, since the area under that curve tells us the work done by the thermodynamic system. We will use p-V diagrams to understand multiple processes chained together, and ultimately we will use p-V diagrams to understand engines and other thermodynamic machines. In Fig 1 we show $p(V)$ curves for the main four types of processes we consider in this class. Note that almost any function $p(V)$ is technically possible, so long as the temperature changes accordingly so the ideal gas law (or equivalent for non-ideal systems) is satisfied.

6.5 Entropy and reversible processes

The basic equation is

$$\Delta S = \int \frac{1}{T} dQ = \int \frac{1}{T} dU + \int \frac{p}{T} dV. \quad (5)$$

We use the same strategy as for work to compute changes in entropy: write the integrand as a function of the integration variables, and perform one-dimensional integration.

For an ideal gas, the entropy change is fairly easy: from the ideal gas law $\frac{p}{T} = \frac{Nk}{V}$, and from equipartition $U = \frac{3}{2}NkT$. So

$$\Delta S = \frac{3}{2}Nk \int_{T_i}^{T_f} \frac{1}{T} dT + \int \frac{Nk}{V} dV = \frac{3}{2}Nk \ln \frac{T_f}{T_i} + Nk \ln \frac{V_f}{V_i}. \quad (6)$$

Here for the first integral, we substituted $dU = \frac{3}{2}NkdT$ since usually we know the temperature but not the internal energy. Hopefully this makes sense; if the temperature is raised, then the entropy increases because the gas particles are moving faster, and therefore have more velocities and directions that they could be moving in. Think about the surface of a sphere with radius $|v|$, v being the velocity of the particles; as $|v|$ increases, there is more surface area. Similarly, if the volume increases, then the gas particles have more locations they could be in the volume, which increases the entropy as well.

Independent of the ideal gas law, we can also say something about the entropy change of two of our processes: adiabatic and isothermal. Neither of these processes change the total entropy, and so we call them **reversible**. They are called reversible because due to the second law, only processes that increase the entropy or keep it the same can proceed forward. So if a gas undergoes, for example, isobaric expansion, the entropy increases. That means that isobaric compression, which would decrease the entropy of the gas, cannot occur unless the entropy of something else increases more than the entropy of the gas decreased. On the other hand, a reversible process can be run either way and still satisfy the second law of thermodynamics.

One can show that the isothermal and adiabatic processes have zero change in entropy. For the adiabatic process, dQ is zero, so according to Eqn 5, ΔS is zero. For the isothermal process, some amount of heat Q is transferred from the environment to the system during the process (negative Q means that the heat goes the other way). We have:

$$\Delta S = \Delta S_{system} + \Delta S_{env} = \frac{Q}{T} - \frac{Q}{T} = 0. \quad (7)$$

Because the temperature was constant, the integral becomes very simple, and because in an isothermal process the environment and the system are at the same temperature, the entropy change is equal in magnitude and opposite in sign for the environment and system. As it turns out, *only* isothermal and adiabatic processes are reversible.

7 Thermodynamic cycles: Engines

- You should be able to compute net work done over a thermodynamic cycle
- Understand the concept of a heat bath and an abstract heat engine
- Carnot efficiency, and how it relates to an engine that produces work from heat flow and also how it affects the creation of temperature differences using work.
- Understand the relationship between entropy and efficiency

7.1 Constructing a simple engine out of thermodynamic cycles

The basic idea of an engine is to construct a sequence of thermodynamic process such that net work is done. For concreteness, we will consider an ideal gas that is inside a cylinder. Imagine that we can heat and cool the gas by putting the cylinder in either a hot environment (could be as simple as boiling water) or a cold environment (could be ice water for example). We idealize both environments as being so large compared to the piston that their temperature doesn't change when we heat or cool the cylinder. Such systems are typically called **reservoirs or baths**. Imagine that we can also either lock the piston in place so that the volume cannot change (isochoric process) or we can allow the piston to move so that the volume can change but at a fixed pressure (isobaric process).

For a very simple example of this, consider the following set of processes, done :

1. Starting at p_1, V_1 , allow the piston to move, and heat the gas. The volume will increase, and we can use the force from the piston to push something and do some work.¹
2. Starting at $p_2 = p_1, V_2 > V_1$, fix the piston's position and cool the gas down by putting it in the ice bath. Now the pressure will decrease but no work is done since the volume doesn't change. Heat flows from the piston to the ice water here.
3. Starting at $p_3 \neq p_2, V_3 = V_2$, keeping it in the ice bath, apply a force to the piston until its volume is the original one V_1 . This requires us to do work on the gas.
4. Starting at $p_4 = p_3, V_4 = V_1$, fix the piston in place and put the piston back in the boiling water. Since the volume is fixed, the pressure will increase back to p_1 and the gas has ended up where we started. No work is done because the volume doesn't change. Heat flows from the boiling water to the piston here.

We can count up how much work was done by the gas during this whole cycle. In step 1, the gas expanded at a constant pressure p_2 , so the work done was $p_2(V_2 - V_1)$. In Step 3, we compressed the gas back to its original volume at a lower temperature, so the work **on** the gas was $p_3(V_2 - V_1)$. Since $p_2 > p_3$, the total work done **by** the gas was

$$W_{by} = p_2(V_2 - V_1) - p_3(V_2 - V_1) = (p_2 - p_3)(V_2 - V_1). \quad (1)$$

The energy for the work was gotten by transferring heat from the hot reservoir (the boiling water) to the cold reservoir (the ice water). Note that if the temperature of the reservoirs were the same, then it would not be possible to get work out of the setup. It's very important that heat spontaneously would move from one reservoir to another. We are extracting some of that heat to repurpose into work. One might consider the **efficiency** of the engine as follows:

$$\epsilon = \frac{W_{by}}{Q_H}, \quad (2)$$

where W_{by} is the work done in the cycle and Q is the heat transferred from the hot reservoir.

¹In a car for example it might turn a crank that turns the wheels.

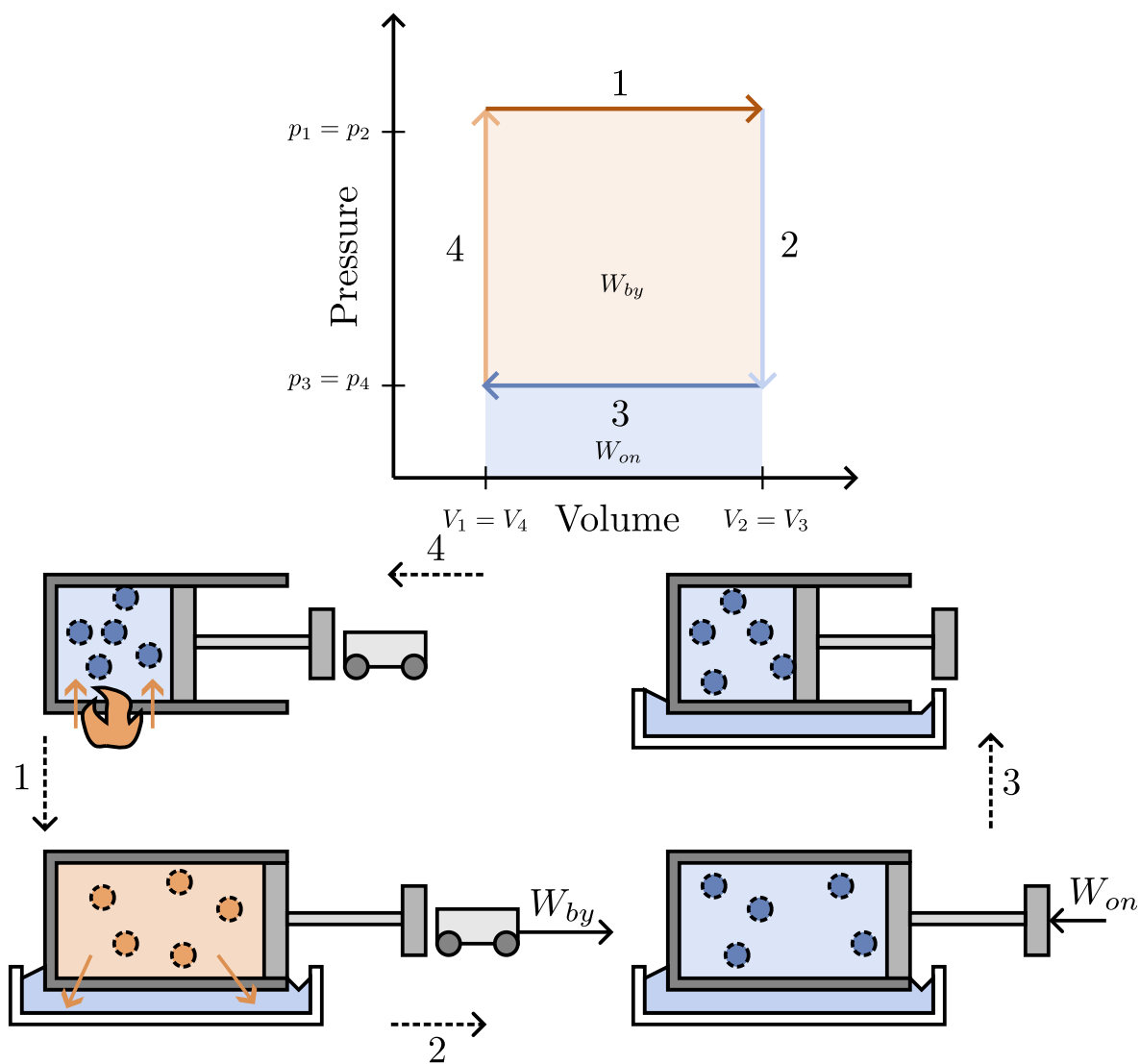


Figure 1: (above) $p - V$ diagram of the thermodynamic cycle in the text 7.1. (below) A diagram of how each process in the cycle might be implemented.

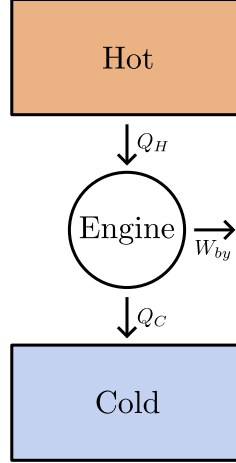


Figure 2: An abstract diagram of a heat engine. The engine uses the energy from heat flow from the hot reservoir to create net work. Waste heat energy is dumped into the cold reservoir.

The idealized engine presented here is not very practical; in reality the gas is typically transported in pipes, rather than moving a piston back and forth. As we will discover, it is also not that efficient. In the next section, we will compute how efficient the engine **could** be while still not violating the second law of thermodynamics.

7.2 Carnot efficiency

The Carnot efficiency is an upper limit on the efficiency attainable. Consider the abstract picture of an engine in Fig ?? . Over the engine cycle, some heat flows from the hot reservoir into the engine, some of that heat energy is turned into work, and finally some heat flows into the cold reservoir.

There are two main equations here. Energy conservation:

$$Q_H = W_{by} + Q_C, \quad (3)$$

and entropy change:

$$\Delta S_{total} = -\frac{Q_H}{T_H} + \frac{Q_C}{T_C} \geq 0 \quad (4)$$

The entropy of the hot reservoir goes down, but the entropy of the cold reservoir must go up at least as much, so

$$\frac{Q_C}{T_C} \geq \frac{Q_H}{T_H} \quad (5)$$

Substituting Eqn 3 into Eqn 5 and rearranging some terms, you can get

$$\epsilon = \frac{W_{by}}{Q_H} \leq 1 - \frac{T_C}{T_H}, \quad (6)$$

which is called the Carnot efficiency.

There are a few assumptions in the Carnot efficiency calculation: the engine itself is perfect, so it does not increase entropy at all. This must mean it is made up only of reversible processes. The engine we designed in the previous section is not perfect! Secondly, the hot and cold reservoirs are large enough that their temperatures do not change; then $\Delta S = \int \frac{dQ}{T} = \frac{Q}{T}$. Relaxing these assumptions generally makes the efficiency go down, so the Carnot efficiency remains an upper theoretical bound to any device that tries to use heat energy flowing between a hot and cold object.

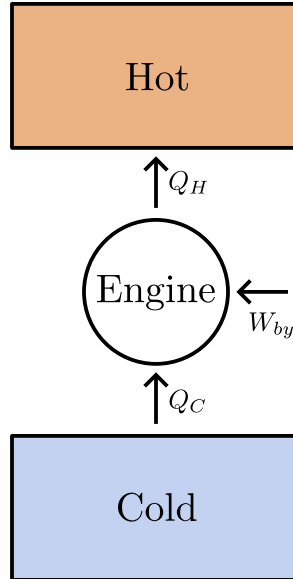


Figure 3: An abstract diagram of a heat pump. Here, work is input to draw heat from the cold reservoir and put it into the hot reservoir. It is exactly the reverse of the heat engine with a different figure of merit.

Relationships like Eqn 6 are very important because they tell us what is possible according to the laws of physics. They also tell us how to get close to a Carnot efficiency; it's very important that the engine itself creates as little entropy as possible. This means that the engine should use reversible or near-reversible processes. The Carnot analysis works not just for engines such as you'd find in a car, but also for power plants based on nuclear heat sources and even solar cells, which use the sun as a hot reservoir and the Earth as a cold reservoir.

7.3 Heat pumps and other applications of engine theory

While an engine is a device that turns heat motion into work, you can turn the system around and turn work into heat motion. This is a heat pump. In the winter, you can use a heat pump to move heat from the outside to the inside. Even though it's cold outside, there is still a lot of internal energy available; usually on Earth it does not get anywhere close to absolute zero. On the other hand, in the summer, you can use a heat pump to move heat from the house to outside. In each case, you are interested in how much heat you added or removed from the house compared to how much work you put into it. Practically the work is usually in the form of electricity running a motor, so the work is essentially the cost of the heating/cooling. The efficiency is typically called the coefficient of performance. For a heater, the coefficient is

$$COP = \frac{Q_H}{W_{on}} \quad (7)$$

while for a refrigerator it is

$$COP = \frac{Q_C}{W_{on}}. \quad (8)$$

All of these efficiencies are basically a ratio between what you get and what you pay.

We can use the exact same analysis as we did for the Carnot efficiency, except some of the signs are reversed (note that $W_{by} = -W_{on}$). Energy conservation is

$$Q_H = W_{on} + Q_C. \quad (9)$$

Then the second law reads

$$-\frac{Q_C}{T_C} + \frac{Q_H}{T_H} \geq 0 \quad (10)$$

or

$$\frac{T_H}{T_C} \geq \frac{Q_H}{Q_C}. \quad (11)$$

Taking the reciprocal, you get

$$\frac{T_C}{T_H} \leq \frac{Q_C}{Q_H}. \quad (12)$$

Let's do the analysis for a heat pump used to heat a home. The coefficient of performance is

$$\frac{Q_H}{W_{on}} = \frac{Q_H}{Q_H - Q_C} = \frac{1}{1 - \frac{Q_C}{Q_H}}. \quad (13)$$

Since $\frac{Q_C}{Q_H} \leq \frac{T_C}{T_H}$, $1 - \frac{Q_C}{Q_H} \geq 1 - \frac{T_C}{T_H}$, and finally

$$\frac{Q_H}{W_{on}} \leq \frac{1}{1 - \frac{T_C}{T_H}}. \quad (14)$$

Another way to do this is just to focus on the maximum efficiency possible and not worry about the inequalities. Note also that this is just the reciprocal of the Carnot efficiency, which makes sense!

Consider the use of electricity for heating, such as a space heater you might buy. One Joule of electrical energy gets turned into one Joule of heat, so the coefficient of performance is exactly 1. Now consider using a heat pump. Let's suppose that the outside temperature T_C is near freezing, 273 K, and inside it is 298 K. Then the maximum coefficient of performance is near 12, which means you get 12 J of heat for every one of electrical energy. In realistic implementations, the coefficient of performance is closer to 4-5 for a high quality device in the 2020's. The fact that no one has made a heat pump with such a large coefficient of performance (12) is further evidence that thermodynamics works.

8 Equilibrium in contact with a bath: Gibbs free energy

- $G = U - TS + pV$
- G is minimized in equilibrium (is there a good example?)
- Concept of 'fixed' variables.

In this section, we are changing our consideration to the computation of equilibrium when a thermodynamic system is connected to a bath, reservoir, or environment. (In thermodynamics, these terms are generally used roughly interchangeably.) The most common situation in our everyday life is that we are in a fixed (or specified) pressure and temperature environment, as opposed to a fixed energy and volume environment. That is to say, energy flows in or out of objects until the temperature is in equilibrium with the environment, and the volume increases/decreases until the pressure is equilibrated. We see this occur in our daily life—water freezes when the outside temperature goes below 0 C, for example. The phases of matter are predicted by finding the state that maximizes the entropy of **both** the environment and the system under consideration.

8.1 Minimization of the free energy is maximization of total entropy

Let's consider maximizing the total entropy of a system (S subscripts for system) + reservoir (R subscripts for reservoir):

$$S_t = S_R + S_S \quad (1)$$

with respect to some macrostate variable x . (x could be the internal energy, volume, etc of the system) Then we maximize the entropy by setting the derivative to zero:

$$\frac{dS_t}{dx} = \frac{dS_R}{dx} + \frac{dS_S}{dx} = 0. \quad (2)$$

We can use the fundamental relation:

$$dS_R = \frac{1}{T}dU_R + \frac{p}{T}dV_R. \quad (3)$$

We'd like to get the equation in terms of just system variables. We do this by noting that if the reservoir internal energy increases, then it must have come from the system, so $dU_R = -dU_S$, and the same for volume, $dV_R = -dV_S$. So

$$dS_R = -\frac{1}{T}dU_S - \frac{p}{T}dV_S. \quad (4)$$

Finally, plugging everything back in, we get the following equation for the maximization of total entropy:

$$\frac{dS_R}{dx} + \frac{dS_S}{dx} = \frac{dS_S}{dx} - \frac{1}{T} \frac{dU_S}{dx} - \frac{p}{T} \frac{dV_S}{dx} = 0. \quad (5)$$

Multiplying both sides by $-T$, we get

$$\frac{dU_S}{dx} - T \frac{dS_S}{dx} + p \frac{dV_S}{dx} = 0. \quad (6)$$

For convenience, we define the **free energy** $G = U - TS + pV$. Then the equilibrium condition is

$$\frac{dG}{dx} = 0, \quad (7)$$

where T and p are understood to be the temperature and pressure of the reservoir, which don't change when the system changes. This is a minimum because we multiplied by $-T$ above, which is a negative number.

From the above, we have a new concept: the free energy, which is minimized in equilibrium, and includes the effects of the reservoir. For fixed pressure and temperatures, this is called the Gibbs free energy and is defined as $G = U - TS + pV$. Note that the Gibbs free energy is only minimized when both pressure and temperature are fixed. In another situation, such as a fixed volume, the Gibbs free energy is not minimized.

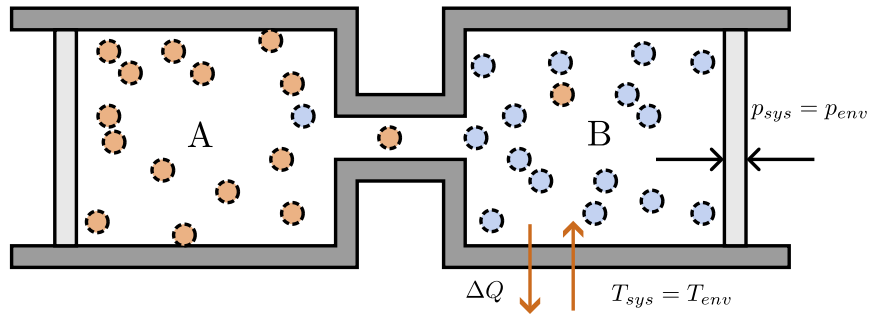


Figure 1: Two systems A and B that can exchange particles in the presence of an environment at a given pressure and temperature. Particles will flow between the systems until $\mu_A = \mu_B$, where $\mu_A = \frac{\partial G_A}{\partial N_A}$

8.2 Example: finding the equilibrium volume of an ideal gas in an environment

Consider an ideal gas with N atoms at a fixed pressure and temperature p, T . What is the equilibrium volume?

$$G = U - TS + pV. \quad (8)$$

We'd like to find the value of V such that

$$\frac{dG}{dV} = \frac{dU}{dV} - T \frac{dS}{dV} + p \frac{dV}{dV} = 0 \quad (9)$$

Using the fact that for an ideal gas, $S = Nk \ln V + f(U, N)$ and that U does not depend on V ,

$$\frac{dG}{dV} = -NkT/V + p = 0, \quad (10)$$

which results in the familiar by now ideal gas law. Note that p and T do not have a derivative because they refer to the pressure and temperature of the environment/reservoir. The important thing here is that as long as we can write down the entropy and internal energy of a system, we can find the equilibrium volume in an environment with a given pressure and temperature by taking the derivative of the Gibbs free energy, even if it's not an ideal gas.

8.3 Definition of chemical potential

In the last section, we found that the quantity $\frac{\partial G}{\partial V}$ is useful for finding equilibrium when the volume can change. When particles can move between systems, such as in Fig 1, $\frac{\partial G}{\partial N}$ is very useful. In fact, this is such a useful quantity that it gets a special name:

$$\mu = \frac{\partial G}{\partial N} \quad (11)$$

is called the chemical potential. The chemical potential is to the number of particles in a system as temperature is to internal energy. Just like heat flows from high temperature to low temperature until the temperatures are equilibrated, particles flow from high chemical potential to low chemical potential until the chemical potentials are equilibrated.

The reason that particles spontaneously flow from high chemical potential to low chemical potential is the same as spontaneous heat flow! This spontaneous flow maximizes the entropy or equivalently minimizes the free energy. To see the reason, consider the minimization of free energy:

$$\frac{G_{tot}}{dN_A} = \frac{dG_A}{dN_A} + \frac{dG_B}{dN_B} \frac{dN_B}{dN_A} = 0. \quad (12)$$

$\frac{dN_B}{dN_A}$ is simply -1 because any increase in N_B must be taken from system A . So therefore, our equilibrium condition is simply

$$\frac{dG_A}{dN_A} = \frac{dG_B}{dN_B}, \quad (13)$$

or

$$\mu_A = \mu_B. \quad (14)$$

8.4 $G = \mu(T, p)N$

Now let's consider a pure substance, for example, water in a gas phase, kept at fixed temperature and pressure. A general principle is that if we double the number of particles, then G also doubles. That is,

$$G(T, p, \alpha N) = \alpha G(T, p, N). \quad (15)$$

Immediately from this perspective we can see that

$$\frac{\partial G(T, p, \alpha N)}{\partial(\alpha N)} = \frac{\partial G(T, p, N)}{\partial N} = \mu(T, p, N), \quad (16)$$

which means that μ does not depend at all on α , and therefore does not depend on N . Therefore,

$$G(T, p, N) = \mu(T, p)N. \quad (17)$$

There are a few intuitive ways to understand Eqn 17. One way to think of it is that if the pressure and temperature are fixed, then the density (number of particles per volume) is also fixed, and since the chemical potential can only depend on the local environment of a molecule, it cannot depend on the total number of particles. The reason that the density is only dependent on temperature and pressure is that if we increase the number of particles, then the volume will just increase. You can see that for an ideal gas; if $pV = NkT$, then for a given pressure and temperature, increasing the number of particles just increases the volume by an equivalent amount, and therefore N/V (the density) is the constant p/kT . We know that there will be no net particle flow between two containers of gas at the same temperature and pressure, and so therefore their chemical potentials are the same, even if one of the containers is much larger than the other one.

Note on multicomponent systems. A similar relation exists for systems with multiple components labeled N_1, N_2, \dots , which is $G = \mu_1 N_1 + \mu_2 N_2 + \dots$. The simple arguments we make later for the stability of phases need to be extended when there are multiple components.

8.5 Phases

To understand why $G = \mu(T, p)N$ is useful, let's consider the following situation. We have N molecules of water in an environment with a given pressure p and temperature T . From our everyday experience, we know that if we wait long enough at atmospheric pressure (101 kPa) and temperature below 0°C then all the water will eventually freeze and become ice. On the other hand, between 0°C and 100°C it will be liquid, and above 100°C a gas (water vapor). These observations are typically summarized in a *phase diagram* as in Fig 2. Note that for this situation, the only way we can have coexistence of two phases in equilibrium is to have the temperature and pressure exactly at one of the phase transition lines.¹

Since pressure and temperature are fixed by the environment, we expect our observations from above to be reproduced by minimizing the free energy G . Let's consider just liquid and gas phases for concreteness. The free energy is simply

$$G_{tot} = \mu_\ell(p, T)N_\ell + \mu_g(p, T)N_g, \quad (18)$$

with the constraint that $N_\ell + N_g = N_w$, where N_w is the total number of water molecules. For a given pressure and temperature, $\mu_\ell(p, T)$ and $\mu_g(p, T)$ are just constants. The number of water molecules in the liquid phase can change, if they evaporate and enter the gas, and vice versa. Minimizing G_{tot} is achieved by:

¹The reason that we do often see coexistence of water vapor and liquid water, and liquid water and solid ice can vary. One is that most of the time, things are not really in equilibrium. The other is that we are considering a fixed pressure situation, which implies that the volume can vary without bound. That is rarely exactly true.

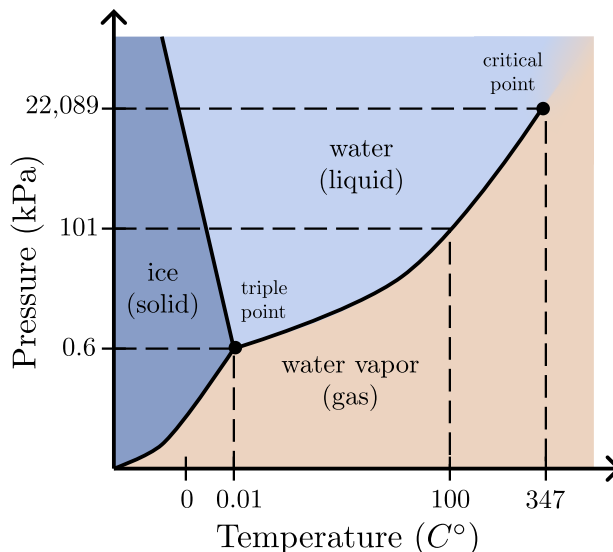


Figure 2: The phase diagram of water, taken from [<https://openstax.org/books/chemistry-2e/pages/10-4-phase-diagrams>], and used under the Creative Commons License.

- If $\mu_\ell < \mu_g$, $N_\ell = N_w$, $N_g = 0$
- If $\mu_\ell > \mu_g$, $N_g = N_w$, $N_\ell = 0$
- If $\mu_\ell = \mu_g$, then N_ℓ and N_g are free to change and G_{tot} remains at a minimum.

The conditions to minimize G exactly predict what we observed in the phase diagram. Evidently, the liquid phase occurs when $\mu_\ell < \mu_g$ and $\mu_\ell < \mu_s$ (the solid), and the same for each of the phases. The phase areas represent the values of p and T such that that particular phase has the lowest chemical potential. The lines indicate values of p and T such that two chemical potentials are equal. The triple point is a particular value of p and T such that *three* chemical potentials are equal.

We thus have a mechanism to determine phase diagrams: if we can compute the difference between chemical potentials (or free energies) for all the phases, we can determine the equilibrium phases by taking the lowest chemical potential, and the phase transition temperature and pressures by finding the values of T and p such that the chemical potentials are equal.

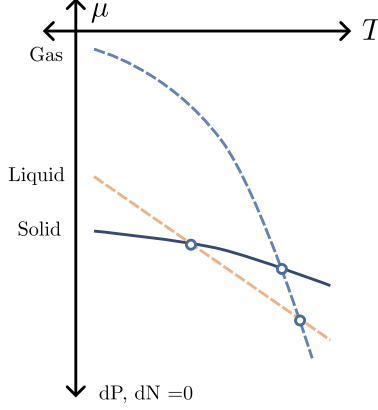


Figure 1: The chemical potential of a solid liquid and gas. The lowest chemical potential is the equilibrium phase. Where the chemical potentials cross is the phase transition.

9 Phase transitions

Last time, we learned that the phase with the lowest chemical potential (or free energy per particle) is the equilibrium phase. Let's explore that idea a little more.

$$G = U - TS + pV = \mu N \quad (1)$$

That means that

$$\mu(p, T) = \frac{U(p, T, N)}{N} - T \frac{S(p, T, N)}{N} + p \frac{V(p, T, N)}{N}. \quad (2)$$

In other words, the chemical potential is determined by the internal energy per particle, the entropy per particle, and the volume per particle (which is the inverse of the particle density). By changing T and p , the relative importance of the entropy per particle and volume per particle changes.

9.1 Higher temperature \rightarrow higher entropy

Let's consider the case where we start at a $T = 0$ K, keep the pressure constant, and increase the temperature T . According to Eqn 2,

$$\mu = \frac{U}{N} - T \frac{S}{N} + p \frac{V}{N}, \quad (3)$$

where we have left off the functional dependence. In general we know that as T increases, U increases roughly linearly (like equipartition), S increases logarithmically (if the heat capacity is roughly constant), and V increases roughly linearly for a gas, and less for a solid or a liquid. A detailed analysis of this is certainly possible, but let's simplify thing by assuming that U, S, V do not change as a function of temperature. This is a good approximation in any small enough region; we'll mostly be interested in the regions around the phase transitions. To simplify our equations a little bit, let's define some lower-case variables: $u = U/N$, $s = S/N$, and $v = V/N$, which are the internal energy, entropy, and volume *per particle*. These are useful quantities because they don't depend on how many particles there are in the system.

In the constant U, S, V approximation, we can determine at what temperature a phase transition occurs fairly simply. Let's consider a solid and liquid. Then for the liquid

$$\mu_\ell = u_\ell - Ts_\ell + pv_\ell \quad (4)$$

and for the solid

$$\mu_s = u_s - Ts_s + pv_s. \quad (5)$$

To find the temperature of the phase transition (the melting temperature), we set the difference in chemical potentials equal to zero:

$$\mu_\ell - \mu_s = (u_\ell - u_s) - T_{melt}(s_\ell - s_s) + p(v_\ell - v_s) = 0. \quad (6)$$

As you can see, only the difference in internal energies, entropies, and volumes per particle matters. For a liquid and solid, we expect:

$$u_\ell - u_s > 0 \quad (7)$$

$$s_\ell - s_s > 0 \quad (8)$$

$$v_\ell - v_s \simeq 0. \quad (9)$$

(while we all know that ice has a larger volume than liquid water, that will not matter too much here). Given those approximations,

$$T_{melt} \simeq \frac{u_\ell - u_s}{s_\ell - s_s}. \quad (10)$$

As you can see, the phase transition occurs when the internal energy cost and the entropy gain upon melting are in balance with each other.

You can probably already see that increasing the temperature tends to favor phases with higher entropy, since that will lower the chemical potential. This is generally the case. As temperature increases, the phases that appear are higher in entropy. Note that as the system crosses the phase transition between solid and liquid, its internal energy will increase; in fact it is drawing heat from its surroundings spontaneously in order to make the transition. This effect is used to make ice cream – ice melting into water draws quite a lot of heat from the ice cream for fast cooling.

9.2 Higher pressure → higher density

We can perform a very similar analysis as temperature for the pressure of phase transitions. Starting from the same equation, keeping T constant and changing just the pressure:

$$\mu_\ell - \mu_s = (u_\ell - u_s) - T(s_\ell - s_s) + p_{melt}(v_\ell - v_s) = 0. \quad (11)$$

Here, if the volume difference is zero then pressure cannot create a phase transition; it can only do so if the volumes are different between phases. In general you can solve this for p_{melt} just like T_{melt} if you know that volume difference.

9.3 Latent heat, enthalpy, and entropy

Now let's consider the heat that the system uses to undergo a phase transition. Let's consider the case in which we are melting ice. What we observe is that the temperature and pressure do not change, but some amount of heat must be transferred to melt all the ice before the temperature can rise again. This means that the heat capacity during the phase transition is infinite, so we must take some care in treating this situation. The total amount of heat it takes to cross the phase transition is called the latent heat L .

Recall the fundamental relation:

$$dS = \frac{dQ}{T}. \quad (12)$$

Since the temperature is constant during the phase transition, then we can integrate simply and determine that the latent heat is quite closely related to the change in entropy

$$L = T\Delta S = T(s_\ell - s_s), \quad (13)$$

where we have specialized to the liquid and solid. Alternatively, we can write this in terms of the quantities per particle

$$\frac{L}{N} = T(s_\ell - s_s) \quad (14)$$

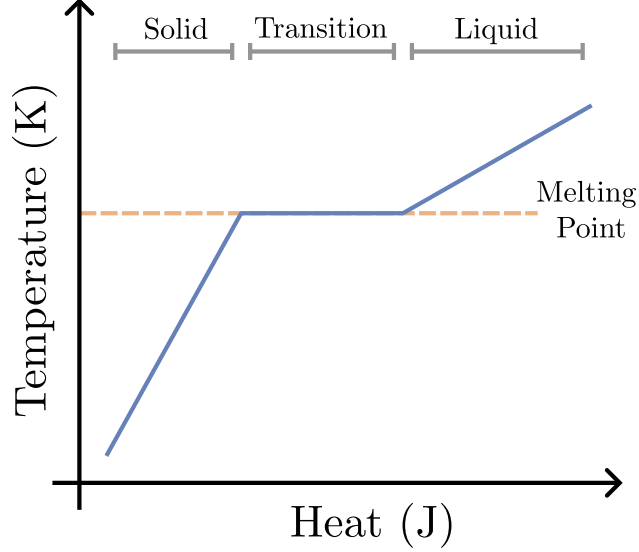


Figure 2: Temperature as a function of adding heat to a system, melting it. In the first rise of temperature, the system is a solid, and the inverse of the slope is the heat capacity dQ/dT . In the flat part, the system is melting, and the horizontal distance traveled is the latent heat. Finally, the third rise is the system in the liquid phase, with a different heat capacity.

We also know that at the phase transition,

$$\mu_\ell - \mu_s = (u_\ell - u_s) - T(s_\ell - s_s) + p(v_\ell - v_s) = 0. \quad (15)$$

Substituting for $T(s_\ell - s_s)$,

$$(u_\ell - u_s) + p(v_\ell - v_s) = \frac{L}{N}, \quad (16)$$

then multiplying by N ,

$$(U_\ell - U_s) + p(V_\ell - V_s) = L. \quad (17)$$

All these equations can be useful; it's often the case that the latent heat per kg is known, which can be converted to the latent heat per particle.

The quantity $U + pV$ is so useful that it gets a new name: **enthalpy**, which is often referred to as H . The latent heat is then

$$L = \Delta H = T\Delta S. \quad (18)$$

It's interesting to think about this last relationship and what it means. On the enthalpy side, we are determining how much energy it takes to increase the internal energy from solid to liquid (in a solid the molecules are in a lower potential energy configuration than the liquid so the internal energy is lower), and to make space in the environment for the liquid phase (the $p\Delta V$ term). In other words, it's how much energy it takes to change the solid into the liquid in that environment. The right hand side $T\Delta S$ is exactly equal to how much heat flows out of the environment into the system, due to the fundamental relation. So this equation is really just following from energy conservation. All these effects happen spontaneously because they minimize the free energy and thus maximize the total entropy.

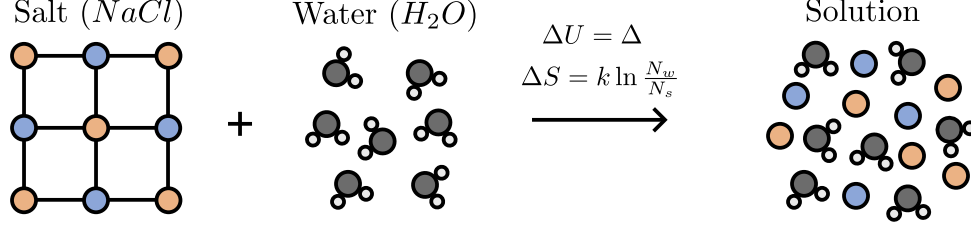


Figure 1: The model for an ideal solution

10 Ideal solutions

Let's consider the following question: in equilibrium, how much salt can be dissolved into water? The details here are very complicated; however, the general trends are very simple. Let's consider the following very simple model:

1. The **change** in total energy for moving one salt molecule from a solid to the water is Δ , which is a currently unknown parameter.
2. Once in the water, the salt molecule occupies the 'space' of one water molecule. This will change the entropy, and is called the entropy of mixing.
3. The volume of the water approximately does not change in this process. (this is good if the solution is dilute)
4. Note also that we are ignoring the fact that NaCl turns into Na^+ and Cl^- in solution. While this does matter to get everything perfectly correct, it does not change the basic behavior.

We would like to compute

$$\frac{dG_{tot}}{dN_S} = G_{tot}(N_S + 1) - G_{tot}(N_S) = 0, \quad (1)$$

where N_S is the number of salt molecules *in solution*. This equation is true from the definition of the derivative, and is accurate since there are many molecules. The derivative equals zero in equilibrium, since we are minimizing the free energy.

Since $G_{tot} = G_w + G_s$, where w and s refer to the liquid with salt in it, and the solid salt respectively, we can reduce Eqn 1 to a quantity that only has differences in it. We reduce the equation to differences because we don't know the absolute internal energy of either the liquid or solid, nor do we know the absolute entropy. Using the fact that $G = U - TS + pV$, we then get:

$$G_{tot}(N_S + 1) - G_{tot}(N_S) = \underbrace{(U_w(N_S + 1) - U_w(N_S) + U_s(N_S - 1) - U_s(N_S))}_{\Delta} \quad (2)$$

$$- T \underbrace{(S_w(N_S + 1) - S_w(N_S) + S_s(N_S - 1) - S_s(N_S))}_{\Delta S} \quad (3)$$

$$+ p \underbrace{(V_w(N_S + 1) - V_w(N_S) + V_s(N_S - 1) - V_s(N_S))}_{\Delta V}. \quad (4)$$

This may look intimidating, but from our assumptions we already know most of these terms. The -1 in the s terms is because when we add a salt molecule from the solid to the liquid, we lose one molecule from the solid (and gain one in the liquid). From assumption 1, we know that the first term involving the internal energy we have defined as Δ . The third term involving the volume is zero from our assumption 3. That leaves us with the second term; the **entropy of mixing**.

To compute the entropy of mixing, we want to know how much entropy increases when adding a salt molecule to the water. We assume that the solid does not change its entropy very much. Then we need to

compute $S_w(N_S + 1) - S_w(N_S)$. If there are N_S salt molecules in the water, and each salt molecule displaces one water molecule, then the number of ways to add them is:

$$\Omega_w(N_S) = \binom{N_w + N_S}{N_S}. \quad (5)$$

That is, we have $N_w + N_S$ possible positions and we choose N_S of them to be salt molecules. Then,

$$S_w(N_S + 1) - S_w(N_S) = k \ln \left(\frac{\Omega_w(N_S + 1)}{\Omega_w(N_S)} \right) = k \ln \left(\frac{(N_S + N_w + 1)!}{(N_S + 1)!N_w!} \frac{N_S!N_w!}{(N_S + N_w)!} \right) \quad (6)$$

Canceling out all the like terms, we get

$$S_w(N_S + 1) - S_w(N_S) = k \ln \left(\frac{N_S + N_w + 1}{N_S + 1} \right) \simeq k \ln \frac{N_w}{N_S}. \quad (7)$$

The last approximation is good because the number of water molecules is much much larger than the number of salt molecules, and both N_w and N_S are much much larger than 1.

Finally, we can put all these pieces together to get the equilibrium N_S :

$$\Delta - kT \ln \frac{N_w}{N_S} = 0. \quad (8)$$

Solving this for the concentration $\frac{N_S}{N_w}$:

$$\frac{N_S}{N_w} = e^{-\frac{\Delta}{kT}}. \quad (9)$$

In reality, we can describe most solutions by making a small modification to this equation:

$$\frac{N_S}{N_w} = n_s e^{-\frac{\Delta}{kT}}, \quad (10)$$

where n_s accounts for the fact that we might be able to fit more or less salt molecules into the solution than the number of water molecules.

10.1 Free energy versus internal energy

In Fig 2, we have the free energy, entropy and internal energy versus concentration $\frac{N_S}{N_w}$ for an ideal solution. The equilibrium concentration is the minimum in free energy. Note that the internal energy increases with the concentration, while the entropy also increases, which makes the $-TS$ term decrease. The minimum in free energy occurs as a result of competition between those two terms. Note that if you increase the temperature, the entropy term becomes more important and will push the minimum to higher concentration. On the other hand, lowering the temperature decreases the importance of the entropy and pushes the minimum to lower concentration. At zero temperature, the equilibrium concentration is zero.

The reason that higher temperature tends to favor higher entropy macrostates is that there is an environment at that temperature. The environment can increase total entropy by adding heat to the system, if it increases the entropy of the system enough. The breaking point is whether adding a given amount of heat Q increases the entropy of the system by more than Q/T , which is how much the entropy of the environment decreases when heat is transferred. As the temperature of the environment increases, the breaking point gets lower, and so the environment is more willing to give up heat to increase the entropy of the system. The free energy nicely captures all those effects!

Optional note: To compute the entropy and internal energy, you must integrate the derivatives:

$$U = \int \frac{dU}{dN_S} dN_S = \int \Delta dN_S = \Delta N_S, \quad (11)$$

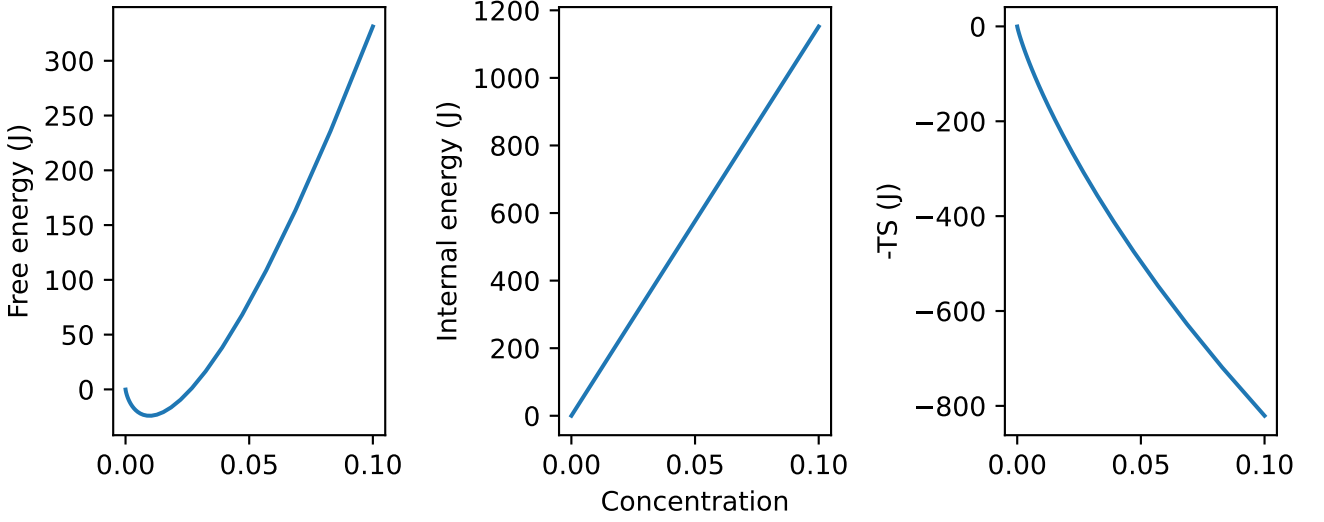


Figure 2: Free energy for $\Delta = 0.12$ eV, $T = 300$ K, and $N_w = 6 \times 10^{23}$. The minimum free energy is the equilibrium concentration, and is formed through a competition between internal energy and entropy.

and

$$S = -k \int \ln \frac{N_S}{N_w} dN_S = -k \left(N_S \ln \frac{N_S}{N_w} - N_S \right). \quad (12)$$

We assume the pV term does not change with concentration. There is an arbitrary constant in both integrals that we set to zero; only differences matter.

10.2 Arrhenius plots

A common way of analyzing systems is to take the log of Eqn 10:

$$\ln \left(\frac{N_S}{N_w} \right) = \ln n_s - \frac{\Delta}{kT}. \quad (13)$$

This equation says that if we plot the log of the concentration versus $\frac{1}{kT}$, then the slope is $-\Delta$. In this way, we can use bulk measurements to estimate the microscopic energy of a solute model. This also explains why one can dissolve more solute into a liquid when the liquid is hot. This, for example, is how sugar syrups are made.

11 Microscopic probabilities of systems connected to a bath: Boltzmann factors

We will end the course by coming full circle. Remember that in Unit 3 we talked about microstates and how they lead to the second law of thermodynamics and thus maximizing entropy to find equilibration. Then in Unit 8 we studied free energy, which when minimized, maximizes entropy of the combined system and reservoir. It is useful to understand what happens to the microstates of a system in this situation—when the system is connected to a reservoir. Here we will study the probability of microstates of a system connected to a reservoir at a given temperature T . We will first just start with the answer and some examples, and then explain why the Boltzmann factors are correct in a later unit.

11.1 Boltzmann factors

For a system in constant with a reservoir at temperature T , the probability of a microstate with energy E is proportional to $e^{-\frac{E}{kT}}$.

11.2 Example: Boltzmann atmosphere

Consider an idealization of our atmosphere in which the system is in complete thermal equilibrium at temperature T . Both of these approximations are quite large actually! In reality the temperature varies with location and altitude because the atmosphere is typically an out of equilibrium system. However, here we will study the equilibrium expectation for an atmosphere of gas.

We are going to be interested in the pressure as a function of altitude (h). We will do this by computing the *ratio* of pressures at different altitudes h_2, h_1 . Assuming that the atmosphere is an ideal gas,

$$\frac{p(h_1)}{p(h_2)} = \frac{N(h_1)/V(h_1)}{N(h_2)/V(h_2)}, \quad (1)$$

where we have assumed that the temperature is the same at both altitudes. The important insight here is that $N(h_1)/V(h_1)$ is the density of atoms at height h_1 , which is also proportional to the probability that any given atom appears at height h_1 . So the ratio of pressures is the ratio of probabilities:

$$\frac{p(h_1)}{p(h_2)} = \frac{P(h_1)}{P(h_2)}, \quad (2)$$

where P is the probability of an atom appearing at h_1 or h_2 . Note the difference in case. The Boltzmann factors allow us to compute that ratio.

Consider a single atom connected to a bath at temperature T . The microstate is given by its position, \mathbf{r} , and its velocity \mathbf{v} . The Boltzmann factor is then

$$P(h) \propto \exp\left(-\frac{\frac{1}{2}m\mathbf{v}^2 + mgh}{kT}\right), \quad (3)$$

where g is the force of gravity and h is the height. Then,

$$\frac{P(h_1)}{P(h_2)} = \frac{\exp\left(-\frac{\frac{1}{2}m\mathbf{v}(h_1)^2 + mgh_1}{kT}\right)}{\exp\left(-\frac{\frac{1}{2}m\mathbf{v}(h_2)^2 + mgh_2}{kT}\right)}. \quad (4)$$

The trick here is to realize that the velocity is completely independent of the altitude, so the velocity parts cancel.

$$\frac{P(h_1)}{P(h_2)} = \frac{\exp\left(-\frac{mgh_1}{kT}\right)}{\exp\left(-\frac{mgh_2}{kT}\right)} = \exp\left(-\frac{mg}{kT}(h_1 - h_2)\right) = \frac{p(h_1)}{p(h_2)}. \quad (5)$$

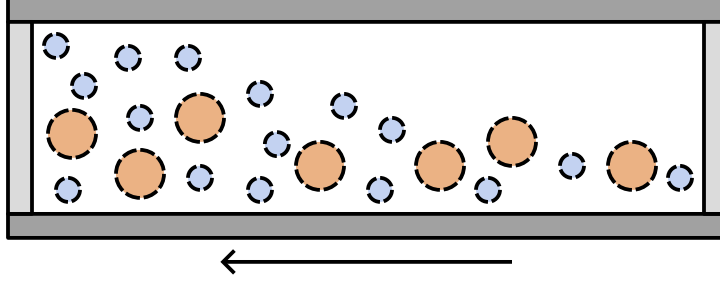


Figure 1: A centrifuge (or an atmosphere). The arrow indicates the direction of acceleration. The orange dots represent molecules with a larger mass, while the blue dots represent a higher mass type of molecule. The “height” increases to the right. There are relatively more blue dots at larger height than orange dots.

Now suppose we set h_2 to be sea level. Then the pressure (and density) of gas decreases exponentially as h_1 (the altitude) increases. Note that how much the pressure decreases depends on the mass of the atom. So atoms that are light like He end up being much more plentiful (relatively speaking) high in the atmosphere than heavy atoms.

11.3 Centrifuge

Consider a centrifuge, as you might have used in a chemistry lab. Such a device is used to separate particles of different mass by spinning them in a tube. Here’s the setup—suppose there are two different types of molecules in the centrifuge, call them A and B . They have different masses m_A and m_B . For the sake of discussion, let’s assume that $m_A > m_B$. We expect that the heavier molecules will be pushed to the bottom of the tube and the lighter ones pushed to the top. We can understand this approximately using Boltzmann factors. Note that here we are considering centrifuging gases; in chemistry often liquids are used which have a bit more of a complicated analysis. We will be interested in the ratio

$$\frac{n_A(h)}{n_B(h)}, \quad (6)$$

where $n_A(h)$ is N_A/V , the molecular density.

By spinning the centrifuge such that the centrifugal acceleration is a ,

$$\frac{n_A(h_1)}{n_A(h_2)} = \exp\left(-\frac{m_A a}{kT}(h_1 - h_2)\right). \quad (7)$$

Now we can compute the ratio of the densities at height h to another reference height h_{ref} .

$$\frac{n_A(h)}{n_B(h)} = \frac{n_A(h)}{n_A(h_{ref})} \frac{n_B(h_{ref})}{n_B(h)} \frac{n_A(h_{ref})}{n_B(h_{ref})} \quad (8)$$

$$= \frac{\exp\left(-\frac{m_A a}{kT}(h - h_{ref})\right) n_A(h_{ref})}{\exp\left(-\frac{m_B a}{kT}(h - h_{ref})\right) n_B(h_{ref})} \quad (9)$$

$$= \exp\left(-\frac{a}{kT}(m_A - m_B)(h - h_{ref})\right) \frac{n_A(h_{ref})}{n_B(h_{ref})}. \quad (10)$$

11.4 Quantum mechanics: a minimum of what you need to know

Boltzmann factors are particularly useful in analyzing very small systems. These systems are typically best described by quantum mechanics. But do not fear! You don’t need to know too much about quantum mechanics to understand this.

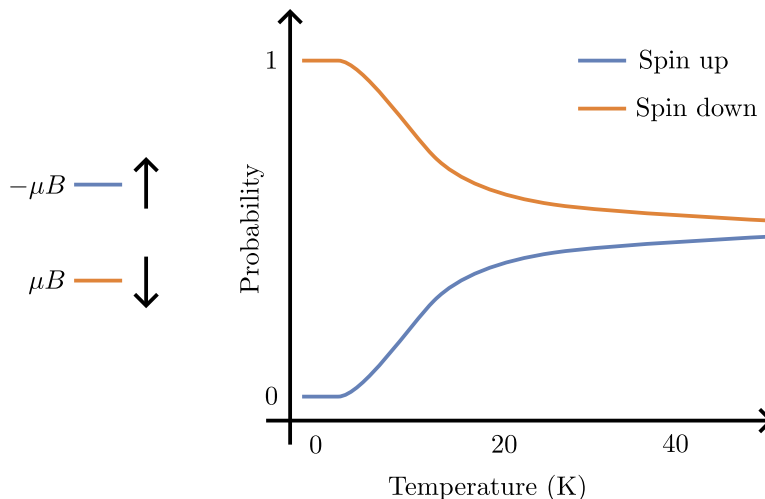


Figure 2: (left) The energy level diagram for a two-state system of an electron in a magnetic field. (right) The probability of each state as a function of temperature.

You simply need to know that unlike classical mechanics, in quantum mechanics under certain conditions (such as electrons in an atom or molecule, or vibrations in solids), energy is quantized. This means that only particular energies are allowed. Typically this is charted using *energy level diagrams* as a guide.

In quantum mechanics, the system is in a state, which we will label $i = 1, 2, 3, \dots$. The energy of that state is E_i . Often these energies will be measured in the unit eV, which is approximately equal to 1.6×10^{-19} J. For that reason we often use the Boltzmann constant in eV/K units for problems like this.

11.5 Computing average energy with Boltzmann probabilities

Define the Boltzmann factor for state i as

$$f_i = e^{-E_i/kT}. \quad (11)$$

The partition function is the sum of all the Boltzmann factors:

$$Z = \sum_i f_i. \quad (12)$$

Then the probability of a state is given by

$$P_i = \frac{f_i}{Z}. \quad (13)$$

To compute the internal energy, we simply take the average over all possible states:

$$U = \langle E \rangle = \sum_i P_i E_i(T) \quad (14)$$

11.6 An electron in a magnetic field

For example, consider an electron in a magnetic field. Electrons have magnetic moments μ that due to quantum mechanics can only either point up or down. The energy of these two states in a magnetic field pointing in the z direction is $E_{\uparrow} = -\mu B$, and $E_{\downarrow} = \mu B$.¹ Confusing the matter slightly is that μ is actually

¹We are glossing over some details; if you want to know exactly how this can be computed/derived, you can find that in a quantum mechanics class! In this class we will always assume that a quantum mechanic has computed the energy values for us!

negative. These magnetic moments (often called spins) line up in ferromagnets, but for the moment let's just consider a single electron connected to a reservoir at temperature T .

The Boltzmann factors are

$$f_{\uparrow} = \exp(\mu B/kT) \quad (15)$$

and

$$f_{\downarrow} = \exp(-\mu B/kT). \quad (16)$$

The partition function is

$$Z = f_{\uparrow} + f_{\downarrow} = \exp(\mu B/kT) + \exp(-\mu B/kT). \quad (17)$$

Using the above facts, the probability of the magnetic moment pointing up is

$$p_{\uparrow} = \frac{\exp(\mu B/kT)}{\exp(\mu B/kT) + \exp(-\mu B/kT)}. \quad (18)$$

Multiplying the top and bottom by $\exp(-\mu B/kT)$, we get

$$p_{\uparrow} = \frac{1}{1 + \exp(-2\mu B/kT)}. \quad (19)$$

Doing the same thing for the \downarrow state,

$$p_{\downarrow} = \frac{1}{1 + \exp(2\mu B/kT)}. \quad (20)$$

The internal energy is

$$U = -\mu B p_{\uparrow} + \mu B p_{\downarrow} \quad (21)$$

We now have completely understood the microstate physics of the two state system. We now can compute the macrostate variables from the microstate variables. As the temperature goes to zero, $p_{\downarrow} \rightarrow 1$ and $p_{\uparrow} \rightarrow 0$, remembering that μ is negative. This minimizes the internal energy, just as we saw in the free energy section. In fact, the Boltzmann factors are just the microscopic probabilities that correspond to a system connected to a bath at temperature T , so the value of U that minimizes the free energy must be the same as the value of U computed using Boltzmann factors.

On the other hand, as the temperature goes to infinity, $p_{\uparrow} \rightarrow 0.5$ and $p_{\downarrow} \rightarrow 0.5$. The average internal energy thus increases from μB (which is negative) to 0 as the temperature increases, as you can see in Fig 3

There is no volume to speak of for an electron, so the heat capacity is given by the derivative of the internal energy with respect to the temperature:

$$C = \frac{dU}{dT}. \quad (22)$$

The exact form is complicated and not particularly enlightening; we will just discuss the qualitative features here. The heat capacity has an odd shape, in which it is near zero at zero temperature, then jumps suddenly as the higher energy state starts to have non-zero probability, and then once both states are equally occupied and the internal energy no longer increases with temperature, the heat capacity goes back to zero. This is very different from what we saw in equipartition (which had a constant heat capacity), and in fact was one of the ways in which quantum mechanics was confirmed experimentally.

We can also compute the entropy as

$$S = \int_0^T \frac{C(T')}{T'} dT'. \quad (23)$$

The entropy S starts at zero and approaches a constant. That constant is exactly equal to $S = k \ln 2$. Why is this? At high temperature, both states are equally occupied, and so $\Omega = 2$! The reason that we cannot use the original $S = k \ln \Omega$ formula is that the probability of the states is not exactly equal unless the temperature is very high. Next time, we will sort out this apparent tension.

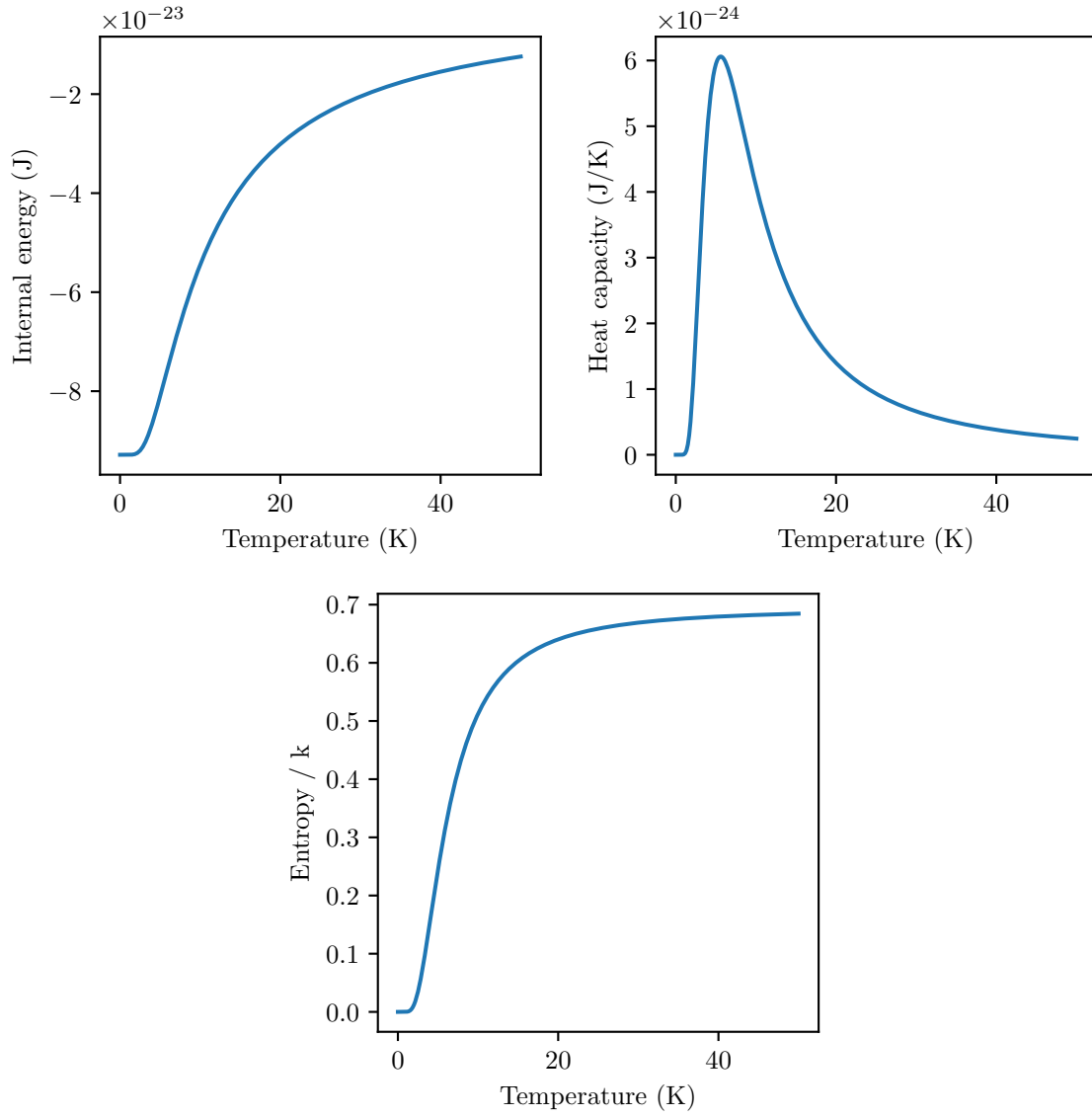


Figure 3: Macrostate variables for an electron in a magnetic field of 10 Tesla.

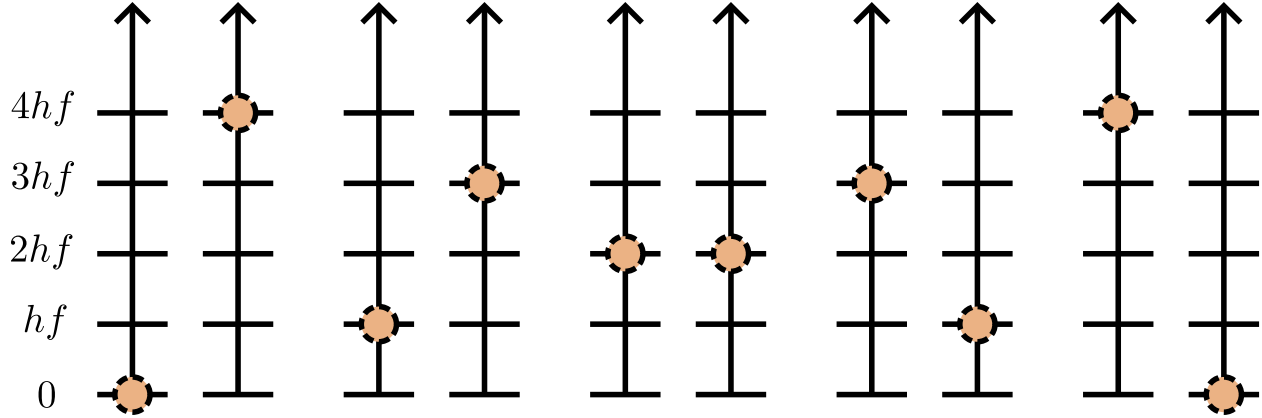


Figure 1: Possible microstate configurations for $U = 4hf$ and two oscillators. Note that in this graph, the factor of $\frac{1}{2}$ was dropped.

12 Oscillators and Boltzmann factors

12.1 Model system: quantum oscillators

Consider a molecule like O_2 or N_2 , which comprise the majority of air. These are diatomic molecules, in which the atoms are bound together via a chemical bond. In the section on equipartition, we mentioned that such a gas could have degrees of freedom either 3 (at very low temperatures), 5 (at room temperature), or 7 (at very high temperatures). The 5 were translational and rotational degrees of freedom, and the remaining two of seven were the kinetic and potential energy of vibration. However, we said that due to quantum mechanics the vibrations need a lot of energy to activate. We will explore this effect in this chapter.

A pretty good model for vibrations in molecules and solids is the quantum simple harmonic oscillator. In this system, the energy levels are given by

$$E_n = (n + \frac{1}{2})\epsilon, \quad n = 0, 1, 2, 3, \dots, \quad (1)$$

where $\epsilon = hf$, and

$$f = \frac{1}{2\pi} \sqrt{\frac{k}{m}}, \quad (2)$$

k being the spring constant and m being the mass of the oscillator. Note that in some references, the factor of $\frac{1}{2}$ is dropped and energy is written as $E_n = nhf$. Since any shift in energy does not change the physics, either option will give the same results.

For molecules and solids, typically ϵ is of order a few meV to a few tenths of an eV for light molecules.¹ Recall that 1 eV is about 1.6×10^{-19} J, which sounds like a small amount of energy. However, kT at room temperature is only about 0.025 eV, so the Boltzmann factors for larger values of ϵ can be rather small at room temperature. This is why the vibrations in N_2/O_2 are mostly inactive at room temperature.

12.2 Derivation of Boltzmann factors for collections of oscillators

12.2.1 Two oscillators with fixed energy U

As a thought experiment, consider an **isolated** system of two identical oscillators with $\epsilon = hf$. This could be two N_2 molecules out in space, not interacting with anything else. The most relevant macrostate variable

¹Light because of the mass in f .

is the internal energy U . To compute the entropy, we use

$$S(U) = k \ln \Omega(U). \quad (3)$$

Since the system can only take on integer amounts of energy, we can count the total *quanta* q in the system. That is, for oscillator A and B , we have n_A and n_B that represent the microstate. U is given simply by the sum $n_A + n_B = q$.

As illustrated in Fig 1, for q quanta, there are $q + 1$ possible configurations. Then

$$U = (q + 1)hf \quad (4)$$

$$S = k \ln(q + 1) \quad (5)$$

We would like to find the remaining thermodynamic quantities by taking the derivative; however, since q is an integer, we can only take the finite difference approximation to the derivative:

$$\frac{1}{T} = \frac{\partial S}{\partial U} \simeq \frac{S(q + 1) - S(q)}{U(q + 1) - U(q)} = \frac{k \ln \left(\frac{q+2}{q+1} \right)}{hf}. \quad (6)$$

Solving for T ,

$$kT = \frac{hf}{\ln \left(\frac{q+2}{q+1} \right)}. \quad (7)$$

This is a bit of a strange equation, but one can see that as q increases, then $\frac{q+2}{q+1}$ gets closer to 1, making the logarithm closer to zero, thus increasing the temperature.

12.2.2 N oscillators with q quanta

Now let's consider many oscillators. This is a model of a solid originally formulated by Einstein, which has many atoms that can oscillate.

The claim is that

$$U = \left(q + \frac{N}{2} \right) hf \quad (8)$$

and

$$\Omega(q) = \binom{N - 1 + q}{q} \quad (9)$$

from this, we can compute

$$S(q + 1) - S(q) = k \ln \left(\frac{q + N}{q + 1} \right) \quad (10)$$

We can then get the temperature as a function of q :

$$\frac{1}{T} = \frac{S(q + 1) - S(q)}{U(q + 1) - U(q)} = \frac{k \ln \left(\frac{q+N}{q+1} \right)}{hf}. \quad (11)$$

12.2.3 Boltzmann factors

Now let's zoom into one of the oscillators. We want to find out what the probability of this oscillator having a given number n_1 of quanta, given that the entire collection has Q quanta. For two total oscillators (one "system" and one "environment"), then given n_1 , we know that $n_2 = Q - n_1$ so there is just one microstate for each value of n_1 . Therefore, the probability is equal for all n_1 up to a maximum of Q . For three total oscillators, given n_1 , we have $q - n_1$ quanta to distribute among two "environment oscillators. This can be done in $Q - n_1 + 1$ ways, as noted previously. So we can see that now the probability of the first oscillator having n_1 quanta decreases as n_1 increases.

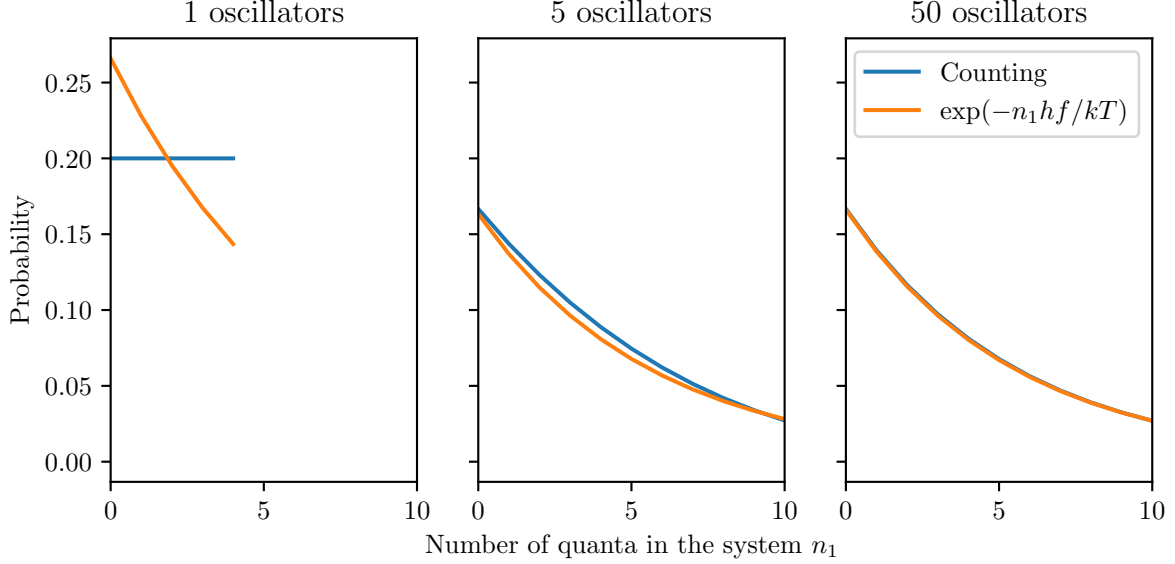


Figure 2: Convergence of the probability distribution of a single oscillator in an environment of other oscillators. The total energy per oscillator is kept constant, and as the environment becomes large compared to the system, the distribution approaches the Boltzmann distribution.

The general formula for N oscillators with q quanta is:

$$\Omega(q, N) = \binom{N-1+q}{q} \quad (12)$$

Substituting $Q - n_1$ for q , we get

$$\Omega(Q, N, n_1) = \binom{N-1+Q-n_1}{Q-n_1}, \quad (13)$$

for $n_1 < Q$.

Given a fixed value of Q and N , we can plot the probability of a given state n_1 as

$$P(n_1, Q, N) = \frac{\Omega(Q, N, n_1)}{\sum_{n_1} \Omega(Q, N, n_1)}. \quad (14)$$

Note that n_1 is a property of the system, while Q and N are properties of the environment. Figure 2 does this plot for increasing N . You can see that as N gets larger, then the probability curve approaches

$$P(n_1, T_{env}) = \frac{\exp(-n_1 hf / kT_{env})}{\sum_{n_1} \exp(-n_1 hf / kT_{env})}, \quad (15)$$

where T_{env} can be computed from Q and N using Eqn 11. Note that here we don't actually care about the total number of oscillators in the environment or the total energy of the environment; all that matters is the *temperature* of the environment. This turns out to be generally true. This is what allows us to abstract away the environment to just having a temperature.

12.3 Heat capacity of oscillators: Einstein's models of solids

Here you get to find out why equipartition sometimes is not applicable.

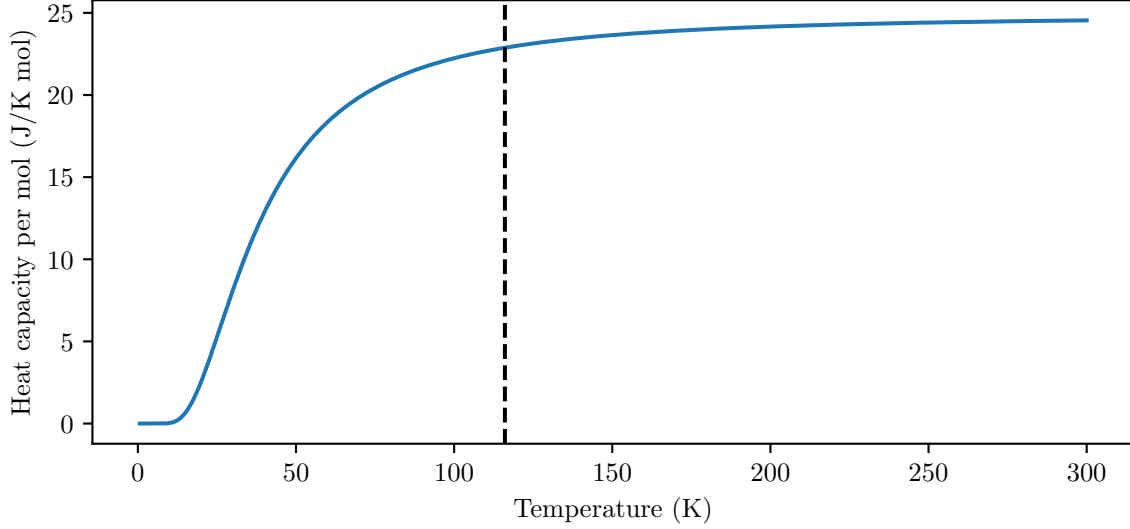


Figure 3: The heat capacity per mol of a collection of quantum oscillators with $hf = .01$ eV. The vertical dotted line is the temperature where $T = hf/k$, and is about where quantum effects are apparent (below this temperature, quantum effects reduce the heat capacity).

First we write down the internal energy as a function of temperature.

$$U = \sum_i E_i P_i(T) \quad (16)$$

At constant volume

$$C_V = \frac{dU}{dT} = \sum_i E_i \frac{dP_i}{dT}. \quad (17)$$

Skipping over quite a few algebraic manipulations and applications of chain rules, etc, we can find that for one oscillator,

$$C_V = k \frac{x^2 e^x}{(e^x - 1)^2}, \quad (18)$$

where $x = \frac{hf}{kT}$. Therefore, for N oscillators in 3 dimensions,

$$C_V = 3kN \frac{x^2 e^x}{(e^x - 1)^2}. \quad (19)$$

Let's consider the limits of this function. It is useful to multiply the numerator and denominator by e^{-x} , and bring it inside the square, which results in

$$C_V = 3kN \frac{x^2}{e^x - 2 + e^{-x}}, \quad (20)$$

As $T \rightarrow \infty$, $x \rightarrow 0$, and $e^x \rightarrow 1$. This would give us $0/0$, but using L'Hopital's rule and differentiating (twice), we get

$$\lim_{T \rightarrow \infty} C_V = \lim_{x \rightarrow 0} 3kN \frac{2}{e^x + e^{-x}} = 3kN. \quad (21)$$

This is the same result we got from equipartition for a collection of N atoms in 3D! So this model has the same high temperature limit as equipartition.

Now let's consider $T \rightarrow 0$. $x \rightarrow \infty$. This is a little easier since we know that e^x increases faster than anything else, so as x becomes very large then it will dominate the other terms. Therefore,

$$\lim_{T \rightarrow 0} C_V = 0. \quad (22)$$

This is very different from the equipartition result, and is a direct result of quantum mechanics! The fact that at low temperatures, the heat capacity goes to zero was one of the first hints that classical mechanics is not the whole story.

We can see the entire function plotted in Fig 3. Note that at room temperature the heat capacity flattens out, but as the temperature is decreased, at around the point of $T = hf/k$, the heat capacity starts to decrease. This is roughly the temperature at which the ratio of the Boltzmann factor between n_1 and $n_1 + 1$ quanta, $\exp(-hf/kT)$, starts to become apparent. Temperatures higher than that don't "see" the discrete energy levels, while for lower temperatures,

Table 1: Resistivity of a few materials at different temperatures

Material	Temperature (K)	resistivity
Copper	60	$9.7 \times 10^{-10} \Omega\text{m}$
Copper	273 K	$1.5 \times 10^{-8} \Omega\text{m}$
Copper	400 K	$2.4 \times 10^{-8} \Omega\text{m}$
Silicon	298 K	$0.5 \Omega\text{m}$
Silicon	1375 K	$9.1 \times 10^{-5} \Omega\text{m}$

13 Semiconductors

13.1 Phenomenology of semiconductors

We can classify most materials as insulating, which means that they do not conduct electricity very well¹, or conducting, in which the electrons are free to move around easily and electricity can be conducted easily. The main difference between them is that a conductor has lower electrical resistivity at lower temperatures, while the insulator has higher electrical resistivity at lower temperatures.

In Table 1 are some values of electrical resistivity versus temperature for a semiconductor (silicon) and a conductor (copper). There are both quantitative and qualitative differences between these values of resistivity. First, note that near room temperature, the silicon is a very poor conductor, with a resistivity about 10^8 larger than copper. However, raising the temperature of silicon lowers the resistivity, while raising the temperature of copper only slightly **increases** the resistivity. Despite the name, a semiconductor is a type of insulator that happens to conduct electricity a little bit at room temperature. In this section we will understand why the semiconductor behaves in the way that it does.

13.2 Simple model of a semiconductor.

This is not a very realistic model of semiconductors, but it captures the basic behavior of them. If you take classes that specialize on semiconductors then you will use quantum mechanics to make more sophisticated models of them. In this chapter, we will model of semiconductors as a bunch of atoms, each of which has two possible states for electrons. The lower energy state is “filled,” which means that in the lowest energy state it has an electron in it. The rule is that electrons can only move to unoccupied states, and there is only one electron allowed per state.

If an electron moves to an unoccupied energy state (called the conduction states, or conduction band) then there’s room for it to move around. On the other hand, if the electron is removed from the filled level (called the valence states, or valence band) then the ‘hole’ can move around as well through collective motion of the electrons in the filled level.

The conductivity is proportional to the number of electrons in the conduction band and holes in the valence band of the material. These come from the metal which is The number of electrons n_e that migrate from the metal is:

$$n_e = NP(e), \quad (1)$$

where $P(e)$ is the probability for an electron to migrate from the metal to a single atom. We will define the chemical potential of electrons in the metal as μ , so the probability of atom i taking on an electron is:

$$P(e) = \frac{1}{Z} e^{-(E_c - \mu)/kT}, \quad (2)$$

where N is the number of atoms.

We also have to consider another process: an electron migrating from the atom to the metal to create a hole. The number of holes n_h is:

$$n_h = NP(h), \quad (3)$$

¹With enough voltage, anything can be a conductor!

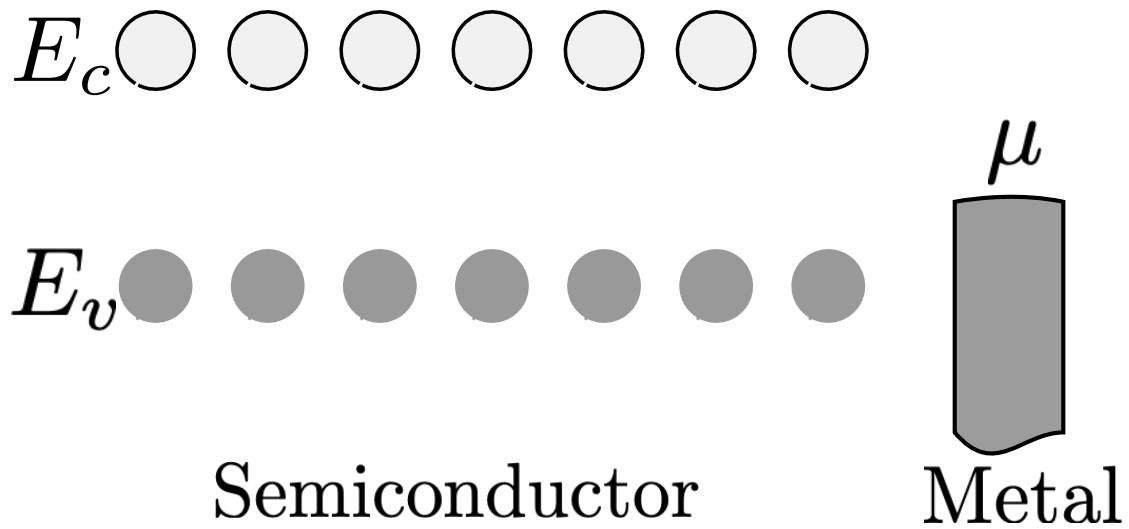


Figure 1: A sketch of the model we are using to understand semiconductors. Each atom in the semiconductor has a low energy state with energy E_v that is filled with an electron (with probability close to 1, similar to our two-state systems!) at low temperature, and a high energy state with energy E_c that is not filled at low temperature. Nearby there is a metal which can either absorb electrons creating a hole, or donate electrons, creating an electron in the conduction states. The gap is the energy difference $E_c - E_v$.

and similarly for atom i losing an electron:

$$P(h) = \frac{1}{Z} e^{-(\mu - E_v)/kT} \quad (4)$$

13.3 Neutral limit

Our previous equations are perfectly fine to use; however, they are not intrinsic to the semiconductor, since they depend on the type of metal attached to it through μ . In most applications, the amount of current we send through the semiconductor is very small compared to the number of atoms in the material itself. In that situation, we can very accurately say that the number of holes and electrons are equal because the system is neutral; that is, $n_h = n_e$.

Using neutrality,

$$n_e = n_h = \sqrt{n_e n_h} = \frac{1}{Z} e^{-(E_c - \mu)/2kT} e^{-(\mu - E_v)/2kT}. \quad (5)$$

Combining the exponentials,

$$n_e = n_h = \frac{1}{Z} e^{-(E_c - E_v)/2kT}, \quad (6)$$

which only depends on the properties of the semiconductor; in particular the so-called **gap** $\Delta = E_c - E_v$, the energy difference between the conduction and valence bands.

Since the conductivity is proportional to the density of electrons n_e , the resistivity ρ is given as follows:

$$\rho \propto 1/n_e \propto e^{\Delta/2kT}. \quad (7)$$

Therefore the resistivity is

$$\rho = C e^{\Delta/2kT}, \quad (8)$$

where C is a constant that is determined by the specifics of the material. This constant, with some work, is computable by more advanced methods. Note the factor of 2 here, as opposed to a standard Boltzmann-like relationship.

This explains the observation we made at the beginning of the chapter; that semiconductors have the opposite behavior of resistivity with respect to temperature as compared to conductors. Conductors (or metals) have zero gap, so the model we made does not apply to them. For a conductor, temperature inhibits electrical conductivity because the vibrations in the atoms get in the way of the electrons' motion. On the other hand, for a semiconductor, increased temperature is helpful to conductivity, because the electrons are helped into the conduction band by the temperature fluctuations.

The gap is one of the primary ways that we characterize semiconductors. Depending on how exactly you would like the material to behave, you may select a different gap. For example, silicon which has a gap of 1.1 eV works quite well at room temperature; however, at very high temperatures it is essentially a conductor of electricity and therefore not typically useful. So at high temperatures or high voltages, you would probably want to use a larger gap material like GaN with a gap of 3.4 eV.

13.4 Measuring the gap: Arrhenius plots of semiconductors

Taking the logarithm of Eqn 8,

$$\ln \rho = \ln C + \frac{\Delta}{2kT} \quad (9)$$

Therefore, when plotting the logarithm of the resistivity versus $1/(2kT)$, the slope is the gap. This is a quick and easy way to estimate the gap of semiconductors.

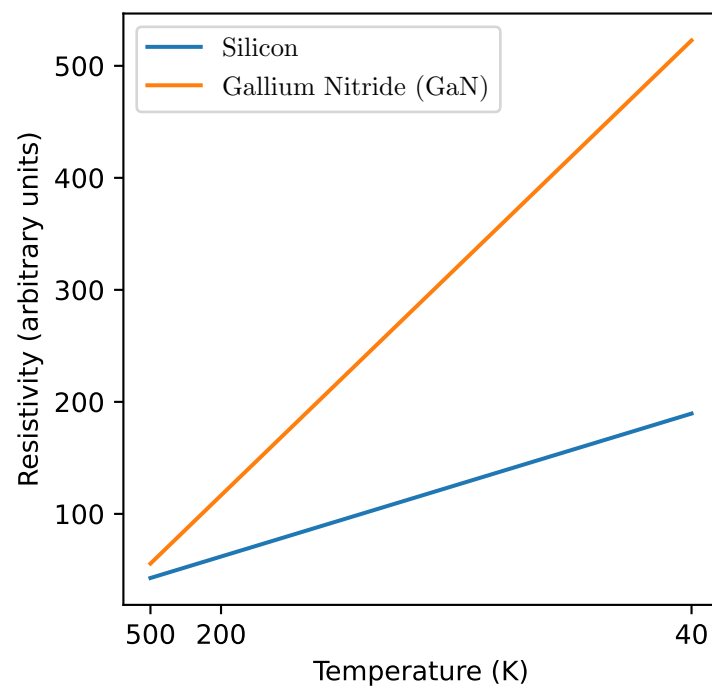


Figure 2: Arrhenius plots for materials with different gaps. The larger gap material has a larger slope on this plot. Note that we are actually plotting $1/2kT$ on the x-axis.