CHM 372: Physical Chemistry for the Life Sciences

Mike Reppert

January 15, 2025

Contents

1 Introduction to CHM372

1.1 What is Physical Chemistry?

Basic question: How do the **basic principles of physics** produce the familiar features of the world around us?

PChem is typically divided into four broad categories:

1.1.1 Thermodynamics: Work and heat

Thermodynamics deals with the question of *how much useful work* can be produced by a given physical or chemical process – and, conversely, how much will necessarily be lost as heat.

1.1.2 Kinetics: How fast does it go?

Kinetics deals with the question of *how fast* a given process occurs.

1.1.3 Statistical Mechanics: Microscopic to macroscopic

Statistical mechanics gives a *microscopic* interpretation to the *macroscopic* laws of thermodynamics and kinetics.

1.1.4 Quantum Mechanics: The dreams stuff is made of

Quantum mechanics is *the physics of the very small* – it lays out the fundamental laws that govern all systems at the microscopic level and, ultimately, give rise to the everything else.

In this course, we'll work through the material "backwards", starting with the more familiar (but still difficult!) problems of thermodynamics and ending with the fundamental principles of quantum mechanics.

1.2 Why Is Physical Chemistry So Hard?

Two reasons:

- **The content is abstract.** There will be many terms you've never heard of or can't define even if you have – work, heat, energy, entropy, enthalpy, free energy, Hamiltonian, partition function, wave function, observable, etc.
  Advice: Keep things grounded in reality! Each time a new abstract term is introduced, make sure you have *at least one concrete example* you can reference to remind yourself what it means physically.

- **The teaching is lousy.** To quote a famous physical chemist:

  > The nihilist attitude of jumping into a discussion of terms whose meaning never has been explained obviously does not make sense in physical sciences. It has never been openly advocated, yet it has been silently adopted by all writers in thermodynamics without a single exception. – Otto Redlich [*Rev. Mod. Phys.* 40, 557 (1968)]

  Advice: I'll do my best! Philosophically, I'm much more concerned about *concepts making sense* than about teaching you to *solve specific problems.* My goal is to introduce the (relatively few) equations that you'll *really need* to use thermodynamics in the future and to make sure you *understand what they mean.*

2 The Ideal Gas Equation of State

Today we'll work through a (relatively) simple example of how relationships between *macroscopic* properties can be derived from *microscopic* physical models.

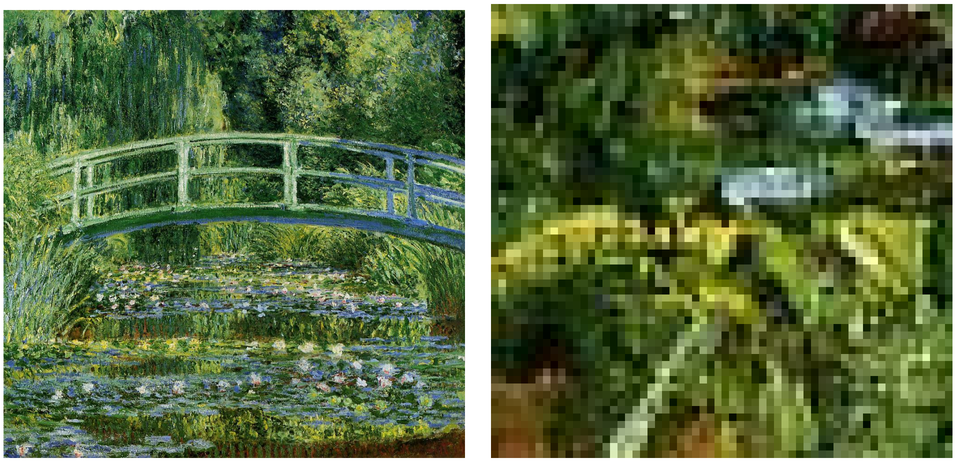
2.1 Macroscopic Properties and Coarse-graining

First, let's define some terms:

Microscopic quantities depend explicitly on *the coordinates and velocities of individual fundamental particles* (atoms, electrons, etc.) that make up a system.

Macroscopic quantities (e.g., pressure, volume, and temperature) are derived by *coarse-graining* (i.e., averaging) over microscopic quantities.

What exactly is coarse-graining? Coarse-graining refers to the process of representing a complicated system in a simplified way, usually by taking averages over some properties that you aren't particularly interested in. Paintings by Claude Monet provide a great example of coarse-graining.



By Claude Monet - the-athenaeum.org [1], Public Domain, `https://commons.wikimedia.org/w/index.php?curid=5749305`

When viewed from a distance, the picture looks realistic. But when we zoom in on a small patch, it looks like nonsense. This is because Monet's style *coarse-grains* over the finer details of the scene in order to present a desired "impression", i.e., to keep the viewer from getting lost in the details.

This is almost exactly what we do in physics: we usually don't have access to all microscopic degrees of freedom in a system and, even if we did, we'd get hopelessly lost trying to keep track of them all. Instead, thermodynamics deals with those coarse-grained, average properties that are both *accessible* and *interesting* in everyday processes.

2.2 The Ideal Gas

As an example of how this coarse-graining process works, let's consider a very simple model: the ideal gas. Specifically, we'll consider a gas of *monatomic* particles that **don't interact with each other**.

2.2.1 Setup

At a microscopic level, the dynamics of each particle is governed by its *Hamiltonian*[1] or energy function.

> The **Hamiltonian** of a material system is its *energy function*, i.e., a mathematical function that describes its energy as a function of the coordinates and velocities (or momenta) of all particles in the system.

For a single particle of a monatomic gas, the Hamiltonian takes the form

$$h = \frac{mv^2}{2} + u(x, y, z). \tag{1}$$

Here $\boldsymbol{v}$ represents the velocity of the particle,

$$\frac{mv^2}{2} = \frac{m\left(v_x^2 + v_y^2 + v_z^2\right)}{2} \tag{2}$$

is its kinetic energy, and $u(x, y, z)$ is its *potential energy* as a function of its three Cartesian coordinates x, y, z. Since the gas particles don't interact with each other, the potential energy depends only on the interactions of the particle with the walls of the box. To simplify the calculations, we'll assume that

- The box is rectangular, with dimensions L_x (length), L_y (width), and L_z (height).
- The potential energy of each particle is zero until it hits one of the walls. From there, the potential energy increases linearly, gradually repelling the particle until it bounces back off the wall.

2.2.2 Microscopic Dynamics

For example, take the *left wall* of the box to be located at position $x = 0$. When a particle reaches the coordinate $x = 0$, it begins to interact with the wall with energy

$$u(x) = -f_o x, \qquad\qquad x \leq 0. \tag{3}$$

Let's see what effect this has on the particle's dynamics.

According to Newton's laws of motion, a particle moving in one direction feels an acceleration $a = \frac{d^2x}{dt^2}$ directly proportional to the force $\boldsymbol{F}$ induced by its potential energy function

$$\boldsymbol{F} = m\boldsymbol{a} = -\frac{\partial u(x)}{\partial x}. \tag{4}$$

In the interior of the box, $u(x) = 0$, so $\boldsymbol{F}(x) = 0$ and $\boldsymbol{a} = 0$. When interacting with the wall,

$$\boldsymbol{a} = \frac{\boldsymbol{F}}{m} = \frac{f_o}{m}, \tag{5}$$

i.e., the particle feels a constant acceleration in the $+x$ direction that eventually pushes it back away from the wall. More specifically

$$\frac{d^2x}{dt^2} = \frac{f_o}{m} \tag{6}$$

---

[1] Named for William Rowan Hamilton, Irish Mathematician, physicist, and inventor of the quaternions used in video-game graphics engines and a whole bunch of other cool stuff.

or

$$\frac{dv_x}{dt} = \frac{f_o}{m}. \tag{7}$$

If the particle is initially moving (to the left, i.e., toward negative x) with x-component velocity $-v_o$ and begins to interact with the wall at time $t = 0$ (i.e., $v_x(t = 0) = -v_o$ and $x(t = 0) = 0$), then

$$v_x = -v_o + \int_0^t d\tau \, \frac{dv_x}{dt} = v_o + \frac{f_o}{m} t. \tag{8}$$

Since $\frac{dx}{dt} = v_x$, we further obtain

$$x(t) = \int_0^t d\tau \; v_x(\tau) = \frac{f_o t^2}{2m} - v_o t \qquad\qquad x \leq 0. \tag{9}$$

Note that this solution is valid *only* for $x \leq 0$ since outside of this interaction zone the particle feels no force and our starting equation for the acceleration is no longer valid.

Notice also that $x(0) = 0$ is one possible solution to the equation; this reflects the initial conditions that the particle begins to interact with the wall at $t = 0$. At first the $-v_o t$ term dominates the expression, indicating that the particle is continuing to move left; the value of x becomes negative as the particle enters the wall's interaction zone. For larger values of t, however, the positive $\frac{f_o t^2}{2m}$ term begins to dominate, so that the particle begins to move back out of the wall. Eventually, the value of $x(t)$ reaches zero again, indicating that the particle has stopped interacting with the wall; from this point, it proceeds to move with constant velocity until it runs into another wall.

At what time does the particle leave the wall's interaction zone? Our equation for $x(t)$ has two solutions to $x(t) = 0$. From the quadratic formula:

$$t = \frac{m}{f_o}\left(v_o \pm \sqrt{v_o^2}\right) = \begin{cases} 0 \\ \frac{2mv_o}{f_o}. \end{cases} \tag{10}$$

Thus at the interaction time

$$\tau_{\text{bounce}} = \frac{2mv_o}{f_o} \tag{11}$$

the particle leaves the wall's interaction zone and proceeds back out into the free space of the box. From here it travels again at velocity v_o (now in the $+x$ direction) until it hits the opposite wall after a free travel time of

$$\tau_{\text{free}} = \frac{L_x}{v_o}. \tag{12}$$

Note that this analysis of the x-component motion is completely independent of and analogous to the motion in the y and z directions. Parallel statements hold for interactions with the right wall of the box, as well as interactions with the top, bottom, front, and back.

2.3 Equation of State

What may come as a surprise now is that this *microscopic* analysis suffices to completely determine the *macroscopic* properties of the gas. For example, suppose we want to calculate the average *pressure* exerted on the walls of the container by the gas of particles. The pressure is simply *the average force per unit area* exerted by the particles on the walls of the container.

As it turns out, this average force is easy to calculate, within some very reasonable assumptions. First, assume that the positions and velocities of the gas particles are completely random within the box. (We don't need at this point to know what the velocity distribution is, just that the distribution is *isotropic*, e.g., that the particles are just as likely to move left as right.) Note that each particle completes a round-trip across the box in time

$$\tau_{\text{tot}} = 2\tau_{\text{bounce}} + 2\tau_{\text{free}}. \tag{13}$$

If the box is very large, $\tau_{\text{free}} \gg \tau_{\text{bounce}}$, so we can approximate

$$\tau_{\text{tot}} \approx 2\tau_{\text{free}}. \tag{14}$$

On average, then, the fraction of the time that the particle spends interacting with the wall is

$$\chi_{\text{bounce}} = \frac{\tau_{\text{bounce}}}{2\tau_{\text{free}}} = \frac{mv_o^2}{L_x f_o}. \tag{15}$$

The average force a single particle exerts on the left wall is then

$$f_{\text{1part}} = f_o \chi_{\text{bounce}} = \frac{mv_o^2}{L_x}. \tag{16}$$

The average force exerted by *all particles* is then

$$f_{\text{avg}} = \frac{mN \left\langle v_x^2 \right\rangle}{L_x}, \tag{17}$$

where N is the number of particles in the box, and the notation $\left\langle v_x^2 \right\rangle$ means that an average is taken over the x-velocities of all particles in the box. Since the area of the left wall is simply $L_y \cdot L_z$, the pressure is then

$$P = \frac{f_{\text{avg}}}{L_y L_z} = \frac{mN \left\langle v_x^2 \right\rangle}{L_x L_y L_z}. \tag{18}$$

Now, the quantity in the denominator of this expression is simply the box volume. The quantity in the numerator is closely related to another important thermodynamic parameter: the macroscopic energy.

The **energy** U of a macroscopic system is the average value of the microscopic Hamiltonian for the entire system.

For the ideal gas, the Hamiltonian for each individual particle is identical, and the average energy of the system is just

$$U = \langle h \rangle \approx \frac{3N}{2} m \left\langle v_x^2 \right\rangle, \tag{19}$$

The approximation here is from dropping the (usually small) contribution of the wall potential $u(x, y, z)$, while the factor of 3 comes from summing up the contributions from the x, y, and z velocities. (Since the particles move in random directions, the v_y and v_z terms add contributions identical in value to the $\left\langle v_x^2 \right\rangle$ term.) Putting these results together, we obtain

$$P = \frac{mN \left\langle v_x^2 \right\rangle}{V} = \frac{2}{3}\frac{U}{V}. \tag{20}$$

Rearranging, this result is often stated as

$$U = \frac{3}{2} PV. \tag{21}$$

Thus we can calculate the energy of an ideal gas (and thus its average particle speed) from the product of its pressure and volume.

3 Thermodynamic Work and Reversibility

3.1 Definitions

In its original context, thermodynamics was *all about work*. How could you maximize the amount of useful work you could get out of a given system? Exactly how much *could* you get out anyway?

Historically, "work" was easy to define. In mechanics, work is usually defined as *a force exerted over a distance*:

$$W = \int_{x_1}^{x_2} F(x)dx. \tag{22}$$

In modern thermodynamics, things are a little more complicated. Sometimes we talk about electrical work or chemical work, neither of which is easy to think of directly as "a force exerted over distance." More generally, we would say that:

> **Work** performed *on a system* is any change in the system energy caused by displacement of a *macroscopic coordinate.*

To understand what that means, it will be helpful to look at a few examples.

3.2 Examples

- **Pressure-Volume (PV) work**: Here the *force* is Pressure and the *coordinate* is volume. (NB: Technically, P should be multiplied by area to have units of force, and V should be divided by area to give units of distance; but in the product these factors cancel.) In equations:

  $$dW = -PdV. \tag{23}$$

  Note the minus sign, indicating that (in our convention) *the work performed on the system is positive when the system is compressed.*

- **Electrical Work**: Work performed by moving a charge against an electric field **E**. Here the field represents a force, and the position of the charges represents the coordinate. In equations:

  $$dW = Q\mathbf{E} \cdot d\mathbf{x}. \tag{24}$$

- **Chemical Work**: Work performed by changing the chemical composition of a system. Here the *coordinate* might be the mole fraction of a certain chemical component or the total number of moles in the sample. The *force* is something called the "chemical potential" that we'll describe more completely later in the course.

- **Gravitational work**: Work performed by displacing an object against the force of gravity. Here the force is gravity, and the coordinate might be the height of the object above the earth.

3.3 Units

3.3.1 Fundamental Units

In this class, we'll use the SI system which uses the fundamental units

- *kilograms* (abbreviated kg) for mass
- *meters* (abbreviated m) for length
- *seconds* (abbreviated s) for time
- *Kelvins* (abbreviated K) for temperature.

The remaining units we encounter will (for the most part) will be derived from these.

3.3.2 Force

The SI unit of force is the *Newton* (abbreviated N):

$$\mathrm{N} = \mathrm{kg\,m/s^2}. \tag{25}$$

If this jumble of units looks confusing, think about the force exerted by gravity on an object. Unless opposed by some other force, gravity tends to *accelerate* objects, i.e., to increase their velocity toward earth. *Acceleration* has SI units of $\mathrm{m/s^2}$ since it reflects how fast velocity (m/s) changes per unit of time (s). The *force* exerted on an object by the force of gravity is just

$$F_{\mathrm{grav}} = m\,g \tag{26}$$

where m is the object's mass and g is the constant (at the earth's surface) *acceleration due to gravity* $g \approx 9.8\,\mathrm{m/s^2}$. So the force due to gravity has units "mass times acceleration" or "mass times velocity per unit time":

$$\mathrm{kg(m/s^2)} = \mathrm{kg\,m/s^2} = \mathrm{N}. \tag{27}$$

3.3.3 Work

Work has the same units as *energy*, reflecting the fact that work performed *on* a system increases its energy, and that work performed *by* a system decreases its energy. In SI units, energy and work have units of *Joules* (abbreviated J):

$$\mathrm{J} = \mathrm{kg\,m^2/s^2}. \tag{28}$$

If you have trouble remembering these units, think again in terms of work performed by the force of gravity. The work performed when an object falls through a very small distance dx looks like

$$dW_{\mathrm{grav}} = m\,g\,dx \tag{29}$$

which has units of

$$\mathrm{kg}\,\left(\mathrm{m/s^2}\right)\mathrm{m} = \mathrm{kg\,m^2/s^2} = \mathrm{J}. \tag{30}$$

So you can think of work units as "mass times acceleration times distance."

3.3.4 Pressure

As we saw in the ideal gas example, pressure is a measure of *force per unit area*. So the units of pressure will be "force per area" or "work over distance over area":

$$\frac{\mathrm{N}}{\mathrm{m^2}} = \frac{\mathrm{J/m}}{\mathrm{m^2}} = \mathrm{J/m^3} = \mathrm{kg/(m\,s^2)} \equiv \mathrm{Pa}, \tag{31}$$

where the symbol Pa stands for the SI name *Pascal* given to this fundamental unit of pressure. Note that this can also be expressed as "energy per volume". Physically, this suggests (correctly) that pressure can be though of as an "energy density" for PV work: The pressure tells you how much capacity the system has to perform PV work in each unit of volume.

For reference, standard atmospheric pressure (sometimes defined as a unit of pressure, atm, in its own right) is equal to approximately 10^5Pa:

$$1\,\mathrm{atm} = 101,325\,\mathrm{Pa}. \tag{32}$$

Example: PV Work

Problem: Let's say we have a cylinder of gas capped by a piston, kept at (standard) atmospheric pressure. How much work is performed *on the system* if we heat up the gas, causing it to expand from a volume of 1 L to 2 L? (Recall that one L is equal to 1000 cm^3 or $1000 \cdot (0.01\text{m})^3 = 10^{-3}\text{m}^3$.)

Solution: The work performed *on* the system is just

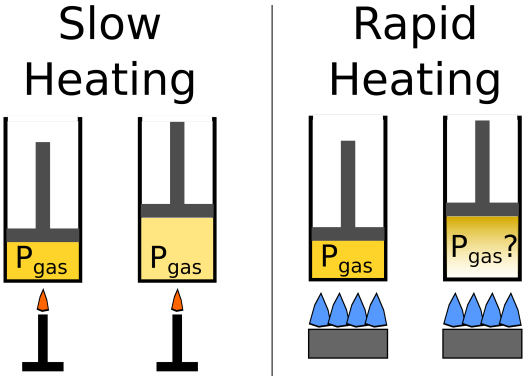
$$\begin{aligned} W &= -P\Delta V = -\left(101325\,\text{Pa}\right)\left(2\,\text{L} - 1\,\text{L}\right) = -101325\,\text{Pa}\,\text{L} \\ &= -\left(101325\,\text{Pa}\right)\cdot\left(10^{-3}\,\text{m}^3\right) \approx -101\,\text{J}. \end{aligned} \tag{33}$$

Note that the work is negative, indicating that the system energy is lowered by the expansion. Conversely, the work performed *by* the system would be just the opposite, i.e., +101 J.

3.4 Definition and Reversibility

A core question that's easy to overlook in discussing processes like gas expansion is the question of whether the quantities we're dealing with are even *well-defined*. For example, in the last example, we stated that the pressure of the system was held constant at atmospheric pressure. But in real life this isn't always easy to do.

Consider the two heating processes illustrated in the figure below. In the left panel, the gas is heated by a single bunsen burner; because the gas is heated slowly, the pressure in the piston stays uniform at all times. As the gas near the bottom of the piston gets heated by the flame, it has plenty of time to mix and diffuse with the gas near the top of the cylinder, keeping a uniform pressure and temperature throughout. In this case, the analysis we just completed would be perfectly appropriate.



In the right-hand panel, the gas is heated rapidly by a much stronger, hotter flame. If the heating takes place fast enough, the gas near the bottom of the container gets heated more rapidly than it can diffuse and mix with the gas at the top of the cylinder. In this case, a temperature and pressure gradient is established inside the cylinder. Neither the temperature nor the pressure are well-defined quantities. In this case, it would be meaningless to try to calculate the work performed using our usual $dW = -PdV$ formula because the pressure of the gas it itself no longer even well-defined![2]

This "definition" problem is closely related to the concept of *reversibility*. In thermodynamics:

---

[2] Very often we try to "cheat" a bit in problems like this and assume that the pressure *outside* the box is well-defined even if the pressure inside the box is not. Depending on the actual construction of our device (and the efficiency with which the gas outside the box is circulated), this may or may not be a valid approach.

A **reversible** process is one in which the system and surroundings are kept arbitrarily close to equilibrium with each other at all times.

This definition assumes a working knowledge of another important term in thermodynamics:

An **equilibrium** system is one whose macroscopic properties do not change with time, no matter how long the system is observed.[a]

---

[a] Strictly speaking, we should add also that an equilibrium state must robust against perturbations; it's possible to have *quasi-stable* states that don't change as long as they're left alone but which can be induced to transition irreversibly to other states by even the slightest perturbations. A simple example is a round boulder balanced at the very top of a tall hill. Such states are not considered to be true equilibrium states.

Both of these definitions are very much idealizations: real processes are never reversible, and real systems are never truly at equilibrium! They are, however, very useful approximations to real processes and real states.

For example, although the gas cylinder is clearly not really in equilibrium with the flame in *either* frame of the figure, it is much *closer* to equilibrium in the single-burner case. (In the multi-burner case, the cylinder isn't in equilibrium with itself, let alone its environment!) An even closer approximation to a reversible heating process would be obtained by heating the cylinder very slowly with a hotplate whose temperature is turned up gradually so that at all times it is kept only slightly higher than that of the cylinder. Similarly, if the flame were turned off and the cylinder allowed to cool down to room-temperature, it would not *really* be in equilibrium since eventually the oxygen and moisture in the surrounding air would corrode the walls of the cylinder and the whole apparatus would fall apart. But that process would happen over many years; over a very long time scale (relative to the heating and cooling of the gas), the system could effectively be said to be in equilibrium.

How is reversibility related to the "definition" problem that opened this section? Well, note first that a reversible process will always have well-defined properties. If the system is arbitrarily close to equilibrium at all times, its macroscopic properties must be arbitrarily well-defined! Conversely, a process in which *all components* have well-defined properties is also usually reversible. *Irreversible* processes (i.e., processes that are not reversible) virtually always end up leaving *some* macroscopic quantity poorly defined since the rapidity with which the transformation takes place prevents equilibration of the macroscopic properties. Since all real-world processes are to some extent irreversible, this term is sometimes used interchangeably with the word *spontaneous*.

4 Heat and Temperature

4.1 What is Heat?

Historically, thermodynamics has primarily been concerned with understanding the interplay between *work* and *heat* – in particular, with how *heat* can be used to perform *work*. You probably already have a pretty good intuition for what these two terms mean. Making that intuition *precise*, however, can be a little challenging. As described by H. S. Robertson:

> "Historically, the distinction between heat and work was made for entirely practical reasons and without any very consistent or objection-free prescription for distinguishing the two. Heat was transferred to a boiler by building a fire under it, and work was extracted from the steam engine *via* a movable piston, clearly and without subtleties, doubts, or confusion. The realization of the true nature of the distinction between heat and work has been rather slow in arriving, and the logical structure of the subject has suffered because authors have attempted distinctions based on spurious reasoning or circular definitions." – Harry S. Robertson *Statistical Thermophysics*, Prentice-Hall Inc., (1993) pg 61.

How would you define heat?

In this course, we'll define heat by what it is *not*:

Heat is the *transfer of energy* into or out of a system by any process that *cannot be identified as macroscopic work.*

Note carefully here that the term *heat* always refers to the *transfer* of energy from one system to another. Unlike energy, we don't consider heat to be a fundamental property of an object. The reason is that, as we'll see in the next lecture, energy transferred *into* a system as heat can be transferred *out of* the system in other forms, i.e., as work. Thus it doesn't make sense to speak of the "heat" of an object, since the amount of heat that can be released from an object depends very much on how the object is treated.

Example: Heating an Ideal Gas at fixed Volume

Problem: Suppose we heat an ideal gas in a closed cylinder at a fixed volume of 1 L. If during the heating process the pressure of the gas increases from 1 atm to 2 atm, how much heat was added to the system?

Solution: Since $\Delta V = 0$, the change in system energy is due entirely to the heat that enters the system. So to calculate the heat, we need only to calculate the change in energy. From relation (21), we can calculate

$$\Delta U = \frac{3}{2}\left(P_2 - P_1\right)V = \frac{3}{2}\left(1\text{atm}\right)\left(1\text{L}\right) \approx 152\,\text{J}. \tag{34}$$

4.2 Temperature: The Ideal Gas Scale

You may be surprised that we so far haven't used the words "thermal" or "temperature". In fact, "thermal energy" is often used interchangeably with "heat" to describe the transfer of energy through means *other* than work. The reason we've avoided these terms is that they're often difficult to define quantitatively. In fact, "thermal energy" is used by different textbooks to refer to different quantities; for this reason, we'll try to avoid using it here for technical purposes.

The status of the term "temperature" is a little better; although different books do indeed define temperature differently, they usually end up meaning the same thing. For practical purposes, the temperature is usually defined by reference to how some easily measured physical property of a system changes when heat is added to it. For example, in a glass thermometer, the measured temperature corresponds to how much the volume of an enclosed pillar of mercury or alcohol increases as it is heated. In principle, any physical property that varies systematically with the addition or subtraction of heat could be used to define a temperature scale, but by convention all such scales are standardized to each other to avoid confusion.

Conceptually, the *ideal gas temperature scale* is perhaps the easiest to understand. Experimentally, it has been observed that gases at very low concentrations really do (almost!) follow the idealized equation of state Eq. (21). As we saw in the example above, this means that when an ideal gas is heated at fixed volume, the pressure changes as

$$\Delta P = \frac{2}{3}\frac{\Delta U}{V} = \frac{2}{3}\frac{Q}{V}. \tag{35}$$

Here Q represents the amount of heat added to the system, which is equal to the total energy change since no work can be performed at constant volume.

This relationship suggests a very simple temperature scale, one that is directly proportional to the pressure of some fixed quantity of gas in a fixed volume, i.e.,

$$T \propto P. \tag{36}$$

To satisfy the intuitive notion that temperature should directly reflect the energy of the gas, we can go further: by choosing

$$T \propto PV, \tag{37}$$

we ensure that (for the ideal gas) T is directly proportional to the energy. Finally, to satisfy our intuitive notion of the temperature, we want to make that it's an *intensive* property like pressure – one that doesn't scale with the size of the system. In other words, if we put double the quantity of gas in a container of double the volume but with the same pressure, the energy should also double – but the temperature should stay the same. We can accomplish this by dividing by the *number of moles* of gas in the system:

$$T \propto \frac{PV}{n}. \tag{38}$$

Finally, we choose a proportionality constant R^{-1} and define

$$T \equiv \frac{PV}{nR}. \tag{39}$$

In this construction, the value of the constant

$$R \approx 8.314\,\mathrm{J/(mol\,K)} \tag{40}$$

defines the value of the temperature T. Rearranging this relationship, we obtain the familiar ideal gas law:

$$PV = nRT. \tag{41}$$

Note that this likewise implies that

$$U = \frac{3}{2}nRT, \tag{42}$$

showing that the energy of the ideal gas depends only on the temperature and not (independently) on its volume or pressure.

In our discussion, we briefly encountered one of a pair of new concepts:

Extensive properties *scale with system size.* If you were to combine two identical systems to make one larger system, the extensive properties of each system would double.

Intensive properties *do not* scale with system size. If you were to combine two identical systems to make one larger system, the intensive properties would be unchanged.

For example, n and V are extensive properties, while T and P are intensive properties. Although this distinction won't play as central a role in our development of thermodynamics as in some treatments, we certainly haven't encountered it for the last time.

4.3 Absolute Zero

One of the most interesting consequences of the ideal gas equation $PV = nRT$ is that it implies that the product PV tends to zero as $T \to 0$. And, indeed, for sufficiently low pressures, real gases actually do exhibit this behavior (see Figure 2.3, page 76 in the textbook). Since P and V can clearly never be negative, this seems to suggest that the point $T = 0$ in our temperature scale defines an *absolute minimum temperature*, below which no substance can ever go. Of course this is something of a speculation since even at very low pressures real gases eventually start to deviate from the ideal gas scale when the temperature goes low enough. As it turns out, though, the speculation turns out to be correct: The *absolute zero* temperature defined by the ideal gas scale (-273.15^{o} C) really does set a lower bound on attainable temperatures in the real world. We'll examine this idea more closely when we reach the Third Law of Thermodynamics.

4.4 Heat Capacity

You probably know from experience that it's easier to heat some objects than others. There are actually two factors at play in the "easiness" of heating:

The **Heat Capacity** of a substance characterizes *how much heat* must be added to bring about a particular change in temperature.

The **thermal conductivity** of a substance characterizies *how quickly heat flows* over a given distance through the material.

One of these properties is extensive, and the other is intensive. Which one is which? Do you know what property is defined by dividing this extensive property by the number of moles in the sample?[3]

In thermodynamics, it's often useful to distinguish between two different types of heat capacity: at constant volume and constant pressure. In equation form

$$C_V \equiv \left.\frac{dQ}{dT}\right|_V \tag{43}$$

$$C_P \equiv \left.\frac{dQ}{dT}\right|_P . \tag{44}$$

Here the subscript indicates which quantity is held constant during the heating (or cooling) process. Although these two quantities are usually almost the same for solids and liquids they can be very different for gases which expand dramatically when they're heated.

---

[3] **Answer**: The *Specific Heat Capacity* or often just *Specific Heat* $\bar{C}_{\text{V}} = \frac{C_{\text{V}}}{n}$.

5 The First Law: Energy is Conserved

5.1 Introducing the First Law

If you're familiar with any of the "laws of thermodynamics," you most likely know the first one, in some form or another. There are several equivalent statements of the law, but most boil down to the following:

First Law of Thermodynamics

Energy can neither be created nor destroyed. It can only be converted between heat and work. Mathematically

$$\Delta U = Q + W, \tag{45}$$

where Q represents the amount of heat *entering* a system, and W represents the work *performed on* the system.

Note that in this statement the quantities Q and W may both be negative: a negative heat "entering" the system means that heat has actually left into the environment, while negative work "performed on" the system means that the system has actually performed work on the environment.

5.2 Universal Principle or Cheap Cop-out?

If you're watching closely, you may have noticed that, in terms of logical structure, we've just fallen for one of the classic blunders: a circular definition! Recall that we defined heat itself to be any energy change *not* attributable to macroscopic work. In other words, we *defined*

$$Q \equiv \Delta U - W. \tag{46}$$

Our first "law" thus follows directly from this definition by simply moving the work term W to the other side of the equation! At this point, you might be suspicious that you've been duped. Why do we call this a "law" if it follows trivially from our definition of heat?

5.3 Historical Perspective: Energy as a State Function

You're right to be suspicious, but the reasoning isn't quite as malicious as it might appear. Recall from Lecture 4 the quote by Harry S. Robertson explaining that, historically, heat and work were were distinguished "for entirely practical reasons" through processes like heating and cooling a boiler or extracting work from a steam engine, without any "subtleties, doubts, or confusion." In other words, the definitions of "heat" and "work" were taken to be self-evident. It was thus a perhaps somewhat surprising – and tremendously significant – observation that, regardless of how a system was manipulated, the *sum total* of the heat and work *entering* a system was always opposite to the *sum total* of the heat and work *leaving* it. It thus became clear that, although heat and work could be interchanged – i.e., heat could be used to produce work and work could be used to produce heat – the sum of the two together was something immutable: what we now call energy.

The modern statement of this distinction between quantities like work and heat (that are in some sense inter-convertible) and those like energy (that are immutable) is that the latter are *state functions*, whereas the former are not. To be precise with our words here:

A **state function** is a physical quantity that depends only on the *current state* of a system and not on its history.

For example, the energy stored in a pot of boiling water is the same regardless of whether the water was heated over a gas range or stirred so vigorously with an egg beater that the temperature begins to rise.[4] If the final state of the water is the same (defined by the temperature, pressure, and volume), then the energy is the same.

In contrast, heat and work are *not* state functions: how much heat or work is produced in a process depends very much on the details of *how* that process is conducted. In its early days, thermodynamics was very much concerned with how one could most efficiently produce work by running some working material (e.g., the steam in a steam engine) through a cyclic process of heating and cooling. In this context, the realization that energy was a state function – and thus the steam would always contain the same amount of energy at the ending point of each cycle – was hugely important.

5.4 Applying the First Law

Since we've already defined heat *in terms* of the energy change to a system, the first law may seem like a bit of a conceptual let-down. But as a book-keeping tool for practical calculations, it has great value. "Follow the energy" turns out to be a great strategy for many thermodynamics problems. We'll see how this works in a series of examples that will be useful in our next lecture.

An **adiabatic** process is one in which *no heat is exchanged* with the environment.

Example: Ideal Gas Adiabatic Expansion (Pressure)

Problem: Suppose we have a volume V_1 of ideal gas at a pressure P_1 and (by very slowly decreasing the external pressure) we allow it to expand reversibly and adiabatically to a new volume V_2. What will be the final pressure of the gas?

Solution: What's happening to the energy in this system? Since the process is adiabatic, no heat is exchanged; according to the first law, the only energy change in the system must come from the work performed:

$$dU = dW = -PdV. \tag{47}$$

But from the ideal gas law [Eq. (21)], we also know that

$$dU = \frac{3}{2} d\left(PV\right). \tag{48}$$

This gives us a relationship

$$\frac{3}{2} d(PV) = -PdV \tag{49}$$

between differentials of the pressure and volume. Now, the differential of a product like PV is obtained from the product rule as

$$d(PV) = PdV + VdP \tag{50}$$

---

[4] Yes, this is possible; in the early days of thermodynamics, something very much like this was a critical demonstration of the equivalence between heat and work.

so that

$$\frac{3}{2}\left(PdV + VdP\right) = -\frac{2}{2}PdV \tag{51}$$

$$\rightarrow \frac{5}{2}PdV = -\frac{3}{2}VdP \tag{52}$$

$$\rightarrow \frac{5}{3}\frac{dV}{V} = -\frac{dP}{P}. \tag{53}$$

Integrating both sides gives

$$-\frac{5}{3}\ln\frac{V_2}{V_1} = \ln\frac{P_2}{P_1} \tag{54}$$

or exponentiating each side and solving for P_2:

$$P_2 = P_1\left(\frac{V_1}{V_2}\right)^{\frac{5}{3}}. \tag{55}$$

Example: Ideal Gas Adiabatic Expansion (Temperature)

Problem: What is the final temperature of the gas in the last problem?

Solution: From the last problem we know that

$$P_2 = P_1\left(\frac{V_1}{V_2}\right)^{\frac{5}{3}}. \tag{56}$$

From the ideal gas equation of state [Eq. (41)] we also know that

$$P_2 = \frac{nRT_2}{V_2}. \tag{57}$$

Combining the two equations, we obtain

$$\frac{nRT_2}{V_2} = P_1\left(\frac{V_1}{V_2}\right)^{\frac{5}{3}} \tag{58}$$

or

$$T_2 = \frac{P_1 V_2}{nR}\left(\frac{V_1}{V_2}\right)^{\frac{5}{3}} = T_1\left(\frac{V_1}{V_2}\right)^{\frac{2}{3}}. \tag{59}$$

Example: Ideal Gas Adiabatic Expansion (Work)

Problem: How much work is performed on the system during the adiabatic expansion just described?

Solution: Eq. (42) tells us that the final energy of the gas is just

$$U_2 = \frac{3}{2}nRT_2. \tag{60}$$

Likewise, the initial energy is

$$U_1 = \frac{3}{2}nRT_1 = \frac{3}{2}P_1V_1. \tag{61}$$

Since the process is adiabatic, the *difference* of these two energies is exactly the amount of work performed on the gas, i.e.,

$$W = \Delta U = \frac{3}{2}nR\left(T_2 - T_1\right) \tag{62}$$

$$= \frac{3}{2}nRT_1\left(\left(\frac{V_1}{V_2}\right)^{\frac{2}{3}} - 1\right). \tag{63}$$

Example: Ideal Gas Isothermal Expansion (Pressure)

Problem: Suppose we have a volume V_1 of ideal gas a pressure P_1 and we allow it to expand reversibly and *isothermally* at temperature T to a new volume V_2. What will be the final pressure of the gas?

Solution: In this case, the energy changes due both to the work performed by the system *and* the heat that enters to maintain the temperature. Since both the volume and temperature of the final state are specified, we can calculate the pressure directly. The ideal gas equation gives for the initial system

$$nRT = P_1V_1. \tag{64}$$

and for the final system

$$P_2 = \frac{nRT}{V_2} = P_1\frac{V_1}{V_2}. \tag{65}$$

Thus we can calculate the final pressure from only the initial pressure and the initial and final volumes. This relationship is often stated as

$$P_1V_1 = P_2V_2, \tag{66}$$

a relationship that is valid only for isothermal processes for a fixed quantity of ideal gas.

Example: Ideal Gas Isothermal Expansion (Heat)

Problem: In the last example, how much heat is added to the gas during the expansion?

Solution: To calculate the heat, we again apply the first law:

$$\Delta U = Q + W \tag{67}$$

$$\rightarrow Q = \Delta U - W. \tag{68}$$

If we can calculate independently both ΔU and W, we can thus calculate the heat Q.

Calculation of the energy change turns out to be trivial: because the ideal gas energy is directly proportional to temperature, ΔU is exactly zero for an isothermal process! Thus the *heat added* to the system is exactly the inverse of the *work performed* on the system:

$$Q = -W = \int_{V_1}^{V_2} P dV \tag{69}$$

$$= P_1 V_1 \int_{V_1}^{V_2} \frac{dV}{V} \tag{70}$$

$$= P_1 V_1 \ln \frac{V_2}{V_1} \tag{71}$$

$$= nRT \ln \frac{V_2}{V_1}. \tag{72}$$

6 The Carnot Engine: Heat Engine Efficiency

6.1 Heat Engines

We've stated several times that traditional thermodynamics was deeply concerned with the transformation of heat into useful work. In this lecture, we'll begin to see how the concepts we've developed can provide useful information on how – and how efficiently – such transformations take place.

Before we get "into the weeds" of how this works, let's do some big-picture thinking about how heat-work transformation happens. What basic inputs do you need?

- The first point is pretty easy: you need a **source of heat**! In a car engine, this is provided by igniting a gasoline/air mixture that combusts to produce a hot gas. In a steam engine, it's the coal fire that heats the boiler.
- The second thing you need is a **working substance**, i.e., the thing that gets heated by the source. In the car engine, the heat source and the working substance are in some sense the same thing – the explosion of the gasoline/air mixture creates a new gas of water and CO_2 whose rapid expansion drives the engine. In the steam engine, the working substance is the steam itself.
- The last essential ingredient may be a bit less obvious: a **cold sink**, i.e., a place for the heat to go. The reason this is necessary is that if all parts of the engine were at the same temperature, there could be no force to cause anything to move! For example, although the heat in a steam engine is critical, it's no less critical that cold (condensed) water be provided as an input. It's the *cyclic* process of heating the cold water and then cooling the hot steam that makes the engine operate.

Formally, a device that uses these principles to transform heat into work is called a *heat engine*:

A **Heat Engine** is a device that converts *heat* into *work* through a cyclic process in which heat is transferred through some working substance contained in the engine from a hot object (the source) to a cold object (the sink).

For a simple example, consider the Aeolipile: `https://en.wikipedia.org/wiki/Aeolipile`.

A critical performance metric for all heat engines is their *efficiency*

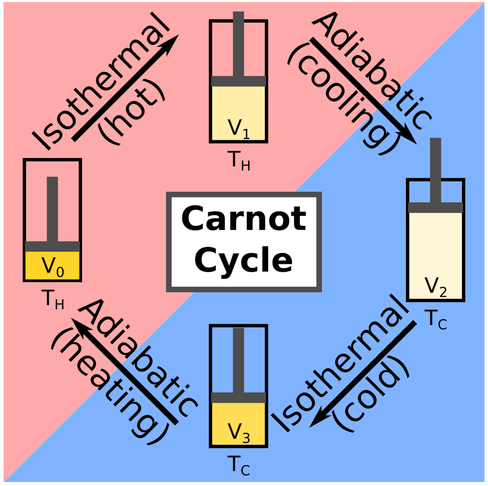
The **efficiency** of a heat engine is the *ratio* of the work output from the working substance to the *heat input* to the substance.

$$\eta = \frac{W_{\text{out}}}{Q_{\text{in}}}. \tag{73}$$

One of the chief objectives of thermodynamics in its early days was to maximize this efficiency, i.e., to extract as much work as possible by transporting thermal energy between a hot source and a cold sink.

6.2 The Carnot Cycle

Today we'll examine the archetype device for all heat engines: the *Carnot engine*. The Carnot engine is a heat engine that operates by a specific set of transformative steps in the working substance. For simplicity, we'll suppose the engine consists of a small gas cylinder with a movable piston and containing a fixed quantity of ideal gas. The cylinder is placed in between a hot object (the source) at temperature T_{H} and a cold object (the sink) at temperature T_{H}. In each step of the cycle, the cylinder is placed in one of three positions: touching the hot object, touching the cold object, or in isolation (not touching either object). Specifically, the Carnot cycle consists of the four steps outlined schematically in the figure and described in detail in the text that follows.



Let's suppose we begin our cycle in equilibrium with the hot reservoir (at temperature $T_{\rm H}$) and some initial volume V_0.

1. **Isothermal Expansion**: The gas is first allowed to expand reversibly at constant temperature $T_{\rm H}$ to some desired volume V_1. Both V_0 and V_1 may be chosen by the engineer to have any values so long as $V_1 > V_0$. The volume of gas in the remaining steps of the process are predetermined by these choices and the constraints of the Carnot process.

2. **Adiabatic Expansion**: Next, the gas is removed from contact with the hot reservoir and allowed to continue to expand, again reversibly, but this time adiabatically, until the temperature reaches $T_{\rm C}$, the temperature of the cold reservoir, at some new volume V_2. This new volume is determined by V_1 and the temperatures $T_{\rm H}$ and $T_{\rm C}$ via (see Eq. (59)):
$$V_2 = V_1 \left(\frac{T_{\rm H}}{T_{\rm C}}\right)^{\frac{3}{2}}. \tag{74}$$

3. **Isothermal Compression**: Now the gas is placed in contact with the cold reservoir and compressed, again isothermally but this time at temperature $T_{\rm C}$. The volume at this step is also predetermined by the cycle parameters, but it will be easier to calculate after we examine the remaining steps. For now, we'll take it as a new parameter V_3.

4. **Adiabatic Compression**: Finally, the gas is compressed adiabatically until its temperature reaches $T_{\rm H}$. As a function of V_3, the final volume V_4 is (see again Eq. (59))
$$V_4 = V_3 \left(\frac{T_{\rm C}}{T_{\rm H}}\right)^{\frac{3}{2}}. \tag{75}$$

Now, for this process to be cyclic, we must have that $V_4 = V_0$, i.e., the system returns to its original volume on the final step. Eq. (75) thus implies that
$$V_3 = V_0 \left(\frac{T_{\rm H}}{T_{\rm C}}\right)^{\frac{3}{2}}. \tag{76}$$

With this calculation, the state of the system (both temperature and volume) is completely specified at all points in the cycle.

6.3 Carnot Efficiency

From this starting point, we can easily calculate the efficiency of the cycle. The work and heat at each step can be calculated as follows:

1. **Isothermal Expansion**: Since the working substance is an ideal gas (whose energy depends only on temperature), the work performed *on* the system must be exactly opposite to the heat flow *into* the system. From Eq. (72)

$$Q_1 = -W_1 = nRT_{\mathrm{H}} \ln \frac{V_1}{V_0} \tag{77}$$

2. **Adiabatic Expansion**: For an adiabatic process, the heat is exactly zero, so the energy change comes entirely from the work. From Eq. (62):

$$W_2 = \frac{3}{2} nR \left(T_{\mathrm{C}} - T_{\mathrm{H}}\right). \tag{78}$$

3. **Isothermal Compression**: Similar to the first step (but now with T_{C} in the prefactor):

$$\begin{aligned} Q_3 = -W_3 &= nRT_{\mathrm{C}} \ln \frac{V_3}{V_2} && (79) \\ &= nRT_{\mathrm{C}} \ln \frac{V_0}{V_1} \left(\frac{T_{\mathrm{H}}}{T_{\mathrm{C}}}\right)^{\frac{3}{2}} \left(\frac{T_{\mathrm{C}}}{T_{\mathrm{H}}}\right)^{\frac{3}{2}} && (80) \\ &= nRT_{\mathrm{C}} \ln \frac{V_0}{V_1}. && (81) \end{aligned}$$

4. **Adiabatic Compression**: In exact opposition to step 2:

$$W_4 = \frac{3}{2} nR \left(T_{\mathrm{H}} - T_{\mathrm{C}}\right). \tag{82}$$

In total, we find that

$$\begin{aligned} W_{\mathrm{out}} &\equiv -\left(W_1 + W_2 + W_3 + W_4\right) && (83) \\ &= nRT_{\mathrm{H}} \ln \frac{V_1}{V_0} + nRT_{\mathrm{C}} \ln \frac{V_0}{V_1} && (84) \\ &= nR \left(T_{\mathrm{H}} - T_{\mathrm{C}}\right) \ln \frac{V_1}{V_0} && (85) \end{aligned}$$

and

$$Q_{\mathrm{in}} = Q_1 = nRT_{\mathrm{H}} \ln \frac{V_1}{V_0}. \tag{86}$$

As a result, the efficiency is

$$\begin{aligned} \eta \equiv \frac{W_{\mathrm{out}}}{Q_{\mathrm{in}}} &= \frac{T_{\mathrm{H}} - T_{\mathrm{C}}}{T_{\mathrm{H}}} && (87) \\ &= 1 - \frac{T_{\mathrm{C}}}{T_{\mathrm{H}}}. && (88) \end{aligned}$$

Thus the efficiency of the Carnot engine is *maximized* when temperature of the hot body is very large, and the temperature of the cold body is very low. In the limit that $\frac{T_{\mathrm{H}}}{T_{\mathrm{C}}} \to \infty$, the efficiency approaches its fundamental limit of unity.

6.4 The Carnot Refrigerator

A critical feature of the Carnot engine is that it is *reversible* – each step in the cycle can be run backwards just as well as forwards. What this means practically is that if work is provided *to* the engine, it can be used to drive heat *from* a cold reservoir *to* a hot reservoir. Although a real refrigerator or air conditioner uses a slightly different cycle than the Carnot engine, the physical basis is the same. When we study the Second Law, we'll also see that the reversibility of the Carnot engine/refrigerator has profound consequences for the *upper limit* of how efficiently a heat engine can run.

7 Entropy: The Thermal Displacement

7.1 Introducing Entropy

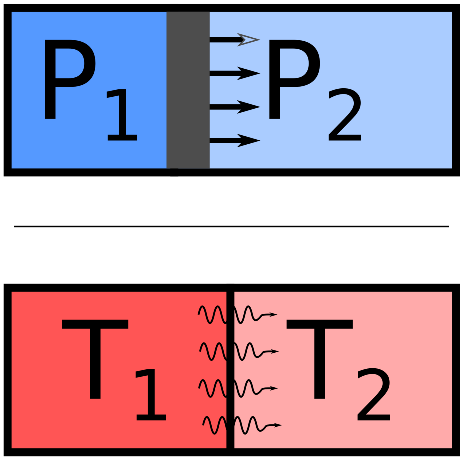
Finally, we're ready to approach what may be the most enigmatic concept in thermodynamics (and don't you love enigmas?): the *entropy*. The entropy has a near-mythical reputation for mystery,[5] but in a thermodynamic sense it's really no more mysterious than work, volume, or temperature. In fact, there's an extremely close analogy between *entropy* and *volume*; and this is where we'll begin our discussion.

7.2 Entropy: The Thermal Volume

Recall that, for very small displacements dV in the volume of a system, the work performed on the system is given by the differential relation

$$dW = -PdV. \tag{89}$$

Physically, what this means is that, when work is performed, the *macroscopic coordinate* V either *pushes against* or *is pushed by* the *macroscopic force* P. Physically, this makes sense. After all, what is it that drives a system to increase its volume? Think about the diagram below, showing two gas chambers with a movable divider between them. What determines whether the divider moves to the left (increasing the volume of chamber 2 and decreasing that of chamber 1) or to the right (decreasing the volume of chamber 2 and increasing that of chamber 1)? It's the pressure, of course! If $P_1 > P_2$, the divider moves right. If $P_1 < P_2$, the divider moves left. Pressure is thus the *conjugate force* that controls changes in volume.



Now, according to our first law, the total (very small) change in energy of a system is the sum of the differentials of heat and work:

$$dU = dQ + dW. \tag{90}$$

The second term here is exactly the infinitesimal work which (as we just discussed) is driven by the pressure. What macroscopic force drives the first term, the heat?

Although the wording may be confusing, intuitively you already know the answer. Consider the lower panel in the figure, showing two blocks of metal at temperatures T_1 and T_2. What force

---

[5] So much so that, when Claude Shannon introduced a new measure of information content, John von Neumann is reputed (probably incorrectly) to have advised him to call it "entropy" as a safeguard against any future arguments about it: since nobody understands what entropy *is*, Shannon could never lose!

determines whether heat flows left (decreasing the temperature of block 2 and increasing that of block 1) or right (increasing the temperature of block 2 and decreasing that of block 1)? The temperature of course! If $T_1 > T_2$, the heat flows right. If $T_2 > T_1$, the heat flows left.

This naturally suggests a question: If the macroscopic force called *pressure* pushes against the macroscopic coordinate called *volume*, what macroscopic coordinate does *temperature* push against? The answer, as you may guess by now, is what we call the *entropy.*

Entropy (denoted by the symbol S) is the macroscopic property whose driving force is *temperature.*

Now, this definition may feel a bit unsatisfying, since it doesn't really tell you what entropy *is*, it just tells you how it *behaves.* Volume feels like a familiar concept since it's something you can feel with your hands and see with your eyes. The reason entropy remains enigmatic for so many of us is precisely that you *can't* feel it with your hands or see it with your eyes.

But, ultimately, this is exactly the point. Remember that we defined *heat* as any change to the energy of a system that is *not* describable as macroscopic work – i.e., any change to system energy that you can't describe as a displacement in a coordinate that you can see with your eyes or feel with your hands. That exchange of energy via heat, of course, is driven by *temperature.* And we've now defined entropy as exactly that coordinate that's driven by heat – i.e., exactly that macroscopic coordinate that you can't see with your eyes or feel with your hands!

Fortunately, there is some hope for building an intuitive picture. As it turns out, you can quite accurately understand entropy as a *thermal volume* – a measure of the amount of space accessible to a system *at the microscopic level.* Just as pressure pushes a compressed gas to increase the macroscopic volume it inhabits, temperature pushes particles *at the microscopic level* to inhabit new regions of "phase space", i.e., new configurations of the coordinates and momenta. At low temperatures, there are relatively few states that a physical system can occupy since thermal energy is simply not available to push the particles over energetic barriers. As the temperature increases, many more possible configurations are made accessible.

For example, at low temperatures, proteins tend to "collapse" into a compact folded (or sometimes globular) state, the lowest energy conformation available. At high temperatures, many more conformations are accessible; the protein could remain folded, but it could also unfold to produce the long, winding chain of a disordered polymer. (This thermal denaturation is exactly what happens when you cook an egg.) Ultimately, entropy is just a measure of the amount of *thermal volume* occupied by the system particles, i.e., *the number of different microscopic configurations accessible to the system.*

Later on in the course, we'll see that statistical mechanics (which seeks to derive the principles of thermodynamics from microscopic models) provides a more formal confirmation of this intuitive picture. In particular, statistical mechanics leads (or is founded on, depending on the logical framework) the result that

$$S = k_B \ln \Omega, \tag{91}$$

where k_B is *Boltzmann's constant* (the ideal gas constant R divided by Avogadro's number N_A) and Ω is a measure of the number of distinct microscopic states accessible to the system.

7.3 Measuring Entropy

The above definition of entropy is well and good, but its abstractness begs one important question: How do you measure entropy? By construction, entropy is a quantity that can't be directly measured, so it's not obvious that it even makes sense to talk about it as a meaningful physical quantity.

Fortunately, there's a straightforward solution. Although we can't readily measure the *absolute* entropy that a system possesses, we *can* measure *changes* in the entropy of a system. Just as Eq. (89)

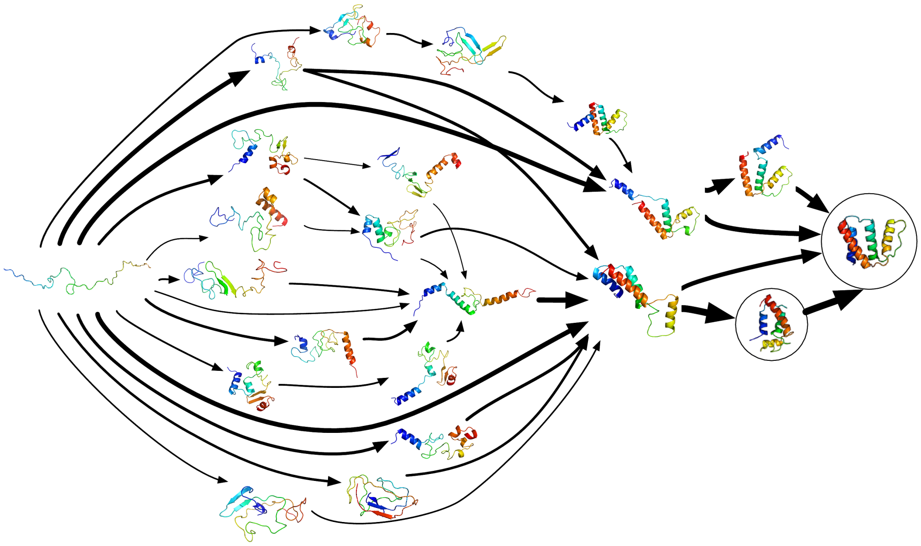


Figure 1: Image by Vincent Voelz, online at `https://commons.wikimedia.org/wiki/File:ACBP_MSM_from_Folding@home.tiff`. Related to publication Voelz et al. *J. Am. Chem. Soc.* 2012, 134, 12565-12577; `http://dx.doi.org/10.1021/ja302528z`

describes the infinitesimal work performed on a system in terms of an infinitesimal change in volume (i.e., the macroscopic volume), a conjugate relation

$$dQ = TdS \tag{92}$$

describes the infinitesimal heat added to a system in terms of infinitesimal displacements in entropy (i.e., the thermal volume). Rearranging this equation, we find that

$$dS = \frac{dQ}{T}. \tag{93}$$

Integrating both sides of the equation gives us a practical means of determining entropy changes experimentally:

Measuring Entropy Changes: In any *reversible process* the change in entropy of a system can be calculated as

$$\Delta S = \int \frac{dQ}{T}. \tag{94}$$

Note that, just like in our discussion of the "well-definedness" of PV work, it's necessary for this definition to make sense that the temperature remain well-defined at all points in the process – in other words, that the process is reversible.

7.4 Entropy as a State Function: ΔS for Irreversible Processes

There seems to be a logical loophole here: Our definition only works for reversible processes, which as we already stated, don't really exist! How do we measure entropy changes for real processes?

The solution to this dilemma is to recognize that – like energy, temperature, pressure, and volume – entropy is a *state function*. As a result, the entropy change involved in a process depends only on *the initial and final states* and not the specific path that connects them. Thus we can calculate the entropy change for a *real* process by calculating the entropy change for an idealized – reversible – process that connects the same two states. The example below illustrates more concretely how this works.

Example: Irreversible Expansion

Problem: Suppose we have an ideal gas contained in a volume V_1 and suddenly open a valve that allows the gas to expand into a larger chamber to some total volume is $V_2 > V_1$. Assume that the process is adiabatic, i.e., that both chambers are thermally insulated so that no heat can flow into the system during the process. What is the entropy change for this process?

Solution: The process is clearly *not* reversible since the gas is very much *not* in equilibrium while it flows between the two chambers. However, we can still calculate the entropy change if we can think of another reversible process that ends with the same result! Now, in the irreversible process, the system neither does work nor receives heat from the environment. Therefore the energy change is zero, i.e., $\Delta U = 0$. Since the gas is ideal, this means that the change in temperature is also zero, i.e., the initial and final temperatures are the same.

Thus we could achieve *the same change in state* by a *reversible* isothermal expansion to the new volume V_2. But we already solved this problem in Section 5! The result is that the heat required is exactly

$$Q = nRT \ln \frac{V_2}{V_1}. \tag{95}$$

Since the temperature is constant, the entropy change is thus

$$\Delta S = \frac{Q}{T} = nR \ln \frac{V_2}{V_1}. \tag{96}$$

The entropy change for the *irreversible* expansion is necessarily the same.

8 The Second Law: Entropy Increases

8.1 Entropy: The One, the Only

In the last section, I suggested that entropy is an enigmatic concept simply because it's not something that we can directly observe or measure. While this is true, it probably isn't the only reason entropy attracts so much interest and speculation. The second reason is that, among all conceivable thermodynamic coordinates, entropy plays a unique role in dictating not just *how* physical processes take place but also *which* processes take place. To be more precise, thermodynamics asserts that:

Second Law of Thermodynamics

The entropy of any *isolated system* is *conserved* during any reversible process and *increases* during any irreversible process.

Before discussing this remarkable claim in more detail, we should emphasize its one major limitation: it applies only to *isolated* systems. The entropy of a system *can* be decreased by the exchange of energy (either heat or work) between the system and its environment. Thus, for example, the entropy of a glass of lukewarm water can be decreased by placing it in a freezer. As the water freezes, the entropy of the glass decreases. (This could never, of course, happen to an *isolated* glass of water at equilibrium above the melting point.) The entropy of the *freezer*, however, increases during this process – so much so, in fact, that if the glass and the freezer were taken together as an isolated system, the total entropy of the *composite* would be positive, in agreement with the Second Law. Thus entropy *can* increase – but only via interactions with an environment.

With this limitation in mind, the uniqueness of the role played by entropy in this discussion should not be overlooked. Suppose that in the Second Law we replaced the word "entropy" with the name of another thermodynamic variable:

- The *volume* of an isolated system increases during any irreversible process.
- The *pressure* of an isolated system increases during any irreversible process.
- The *temperature* of an isolated system increases during any irreversible process.

None of these statements are true! In fact, there are many irreversible processes that can occur in isolation and at constant volume (e.g., the irreversible combustion of hydrogen and oxygen to create water in a sealed chamber). Likewise, the irreversible condensation of steam into liquid water will certainly decrease the pressure inside an isolated gas cylinder. And many irreversible chemical reactions consume heat, lowering the temperature of their container if carried out in isolation.

Only entropy carries the distinction of *never* decreasing in an isolated system. For this reason, entropy has been called the "arrow of time" – an indicator unlike any other that, in an isolated system, always increases with the passage of time. While microscopic laws like Newton's equations are *time-reversible*, the macroscopic principle embedded in the Second Law of Thermodynamics is not. It is perhaps this observation more than any other that has cemented entropy's unique and inscrutable place in our collective understanding of the universe.

8.2 The Relentless March of the Thermal Volume

All this discussion of the "uniqueness" of entropy begs the question: Why entropy? In the last lecture, we saw that entropy was very closely analogous to the macroscopic volume. Just as pressure drives expansion of the macroscopic (mechanical) volume V, temperature drives expansion of the microscopic (thermal) volume S. In this sense, the two quantities seem quite comparable. So what's so special about entropy?

Ultimately, entropy's uniqueness comes from its fundamental connection with heat. As we've repeated several times already, heat is uniquely that type of energy change that *can't* be described as macroscopic work. Now, in a macroscopic system, consisting of on the order of 10^{23} atoms, the few coordinates we recognize as corresponding to the performance work – say volume, height above the earth, chemical composition, or position in a (macroscopic) electric field – represent an infinitesimally small fraction of all the roughly 10^{23} possible coordinates that define the microscopic state. So while only a handful of coordinates describe all possible modes of macroscopic work, the entropy represents the thermal volume inhabited by *all the remaining degrees of freedom.*

In this light, the relentless march of increasing entropy in isolated systems shouldn't come as a surprise. In the absence of outside intervention (i.e., in isolation), it seems reasonable to expect that a system will tend to distribute its energy more or less randomly through all available coordinates. And, indeed, this is precisely what an increase in entropy represents: an increase in the "thermal volume" accessible to the system at a microscopic level. Once this thermal volume reaches a certain level, *decreasing* the volume (i.e., lowering the entropy) would require all the countless microscopic, non-work coordinates to *collectively* withdraw from those regions of phase space that they previously occupied. In the absence of outside intervention, the chances of this occurring (in a system of $\sim 10^{23}$ particles) is essentially zero.

8.3 Entropy and Spontaneity

From a practical perspective, the great utility of the Second Law is that it gives us a rigorous way to quantify whether or not a given process will actually proceed in the real world. Recall that we earlier stated that, in the real world, all processes are, to some extent, irreversible. Flipping the Second Law in reverse, then, one obtains the decree that *a process in an isolated system will only happen in the real world if it increases the entropy of the system.* Thus if we can calculate the entropy change for a given process, we can certify definitively whether or not it can take place in the real world.[6]

For many processes, this predictive power seems a bit unnecessary. For example, while we could use the second law to certify definitively that the air in a room won't spontaneously contract into one corner (by calculating the entropy change associated with the corresponding reversible compression, common sense already tells us that we probably don't need to worry about it! Where the Second Law really comes in handy is in dealing with less intuitive processes such as the progress of a chemical reaction. In this case, it's often not obvious at all ahead of time whether a reaction will proceed in the forward or reverse direction, or at what point (i.e., at what ratio of reactant:product concentrations) it will stop. As we'll see in the second section of our course (after Exam 1), the Second Law gives us exactly the tools we need to make such predictions quantitatively, often using experimental inputs that are much easier to collect than carrying out the actual reaction in the lab.

8.4 Alternative Statements of the Second Law

We should note here that there are several different ways of stating the Second Law. They all turn out to be equivalent, but the connections between them may not be obvious. We'll provide two of those statements here and show how they *follow from* our earlier statement. (They can also be shown to *imply* the Second law, but we won't worry about demonstrating that here.)

Clausius Statement of the Second law

Heat cannot of itself transfer from a colder body to a hotter body.

The phrase "of itself" here means "without work input from the environment". Recall that in our discussion of the Carnot engine we noted that the process can be run in reverse as a "Carnot

---

[6] Note that the Second Law doesn't guarantee that a given process *will* happen, only that certain processes *will not.*

refrigerator". So if work is supplied to the system from the outside, heat *can* be transferred from a colder body to a hotter one. Clausius' statement asserts that this can *only* happen with the input of external work.

Why does this follow from the Second Law? A fairly simple calculation (that we'll go through in class) shows that any process in which heat flows from a cold body to a hot body *decreases* the entropy of the combined system. According to the Second Law, this is impossible.

Carnot's Theorem

No heat engine operating between two temperature T_H and T_C can be more efficient than the Carnot engine operating between those temperatures.

This statement might seem a bit surprising. What is so special about Carnot's heat engine? The answer is that it's *reversible*: i.e., every step of the cycle can be carried out arbitrarily close to equilibrium. As a result, the engine can be run backwards as a refrigerator to pump heat from a hot to a cold resorvoir, and it operates with *the same efficiency* as a refrigerator as it does as a heat engine. As it turns out, this implies that if it *were* possible to build a heat engine with a *higher* efficiency than a Carnot engine, we could use the work produced by that engine to run a Carnot engine in reverse and *overall* to transport heat from a cold body to a hot body. This violates Clausius' statement and, therefore, the Second Law.

To see this, suppose we had a "better" engine that operates at an efficiency η_B greater than the Carnot efficiency $\eta_C = 1 - \frac{T_C}{T_H}$. We could then set up both engines to operate between the same hot and cold reservoirs, and connect the work *output* from the Better engine to the work *input* of the Carnot engine. The Better engine thus operates in the forward direction (i.e., producing work $W_{\text{out}}^{(B)}$ from heat input $Q_{\text{in}}^{(B)}$), while the Carnot engine operates in reverse as a refrigerator (i.e., transferring heat $Q_{\text{out}}^{(C)}$ into the hot body driven by the work input $W_{\text{in}}^{(C)} = W_{\text{out}}^{(B)}$ from the Better engine). Note that we've chosen our sign conventions here so that all the quantities just mentioned are positive. Now, the efficiency of each engine is

$$\eta_B = \frac{W_{\text{out}}^{(B)}}{Q_{\text{in}}^{(B)}} \tag{97}$$

$$\eta_C = \frac{W_{\text{in}}^{(C)}}{Q_{\text{out}}^{(C)}} \tag{98}$$

which can be rearranged to give

$$Q_{\text{in}}^{(B)} = \frac{W_{\text{out}}^{(B)}}{\eta_B} \tag{99}$$

$$Q_{\text{out}}^{(C)} = \frac{W_{\text{in}}^{(C)}}{\eta_C}. \tag{100}$$

The total heat transferred *into* the hot reservoir in this process is then

$$\begin{aligned} Q_H &= -Q_{\text{in}}^{(B)} + Q_{\text{out}}^{(C)} && (101)\\ &= -\frac{W_{\text{out}}^{(B)}}{\eta_B} + \frac{W_{\text{in}}^{(C)}}{\eta_C} && (102)\\ &= W_{\text{out}}^{(B)}\left(\frac{1}{\eta_C} - \frac{1}{\eta_B}\right). && (103) \end{aligned}$$

If $\eta_B > \eta_C$, as assumed, then $\frac{1}{\eta_C} > \frac{1}{\eta_B}$, and in total

$$Q_H > 0. \tag{104}$$

But this violates the Clausius statement of the Second Law! (Note that the combined system of hot reservoir, cold reservoir, and both engines constitues an isolated system – no work or heat came in from the outside during this process.) The only possible conclusion is that – if the Second Law holds true – *no engine can operate with an efficiency greater than that of the Carnot engine.*

9 The Third Law: Standardizing Entropy

9.1 The Third Law

Equation (94) gives us a straightforward means of calculating the *change* in the entropy of a system as a function of temperature. It doesn't, however, give us any means of establishing the *absolute* entropy at any given temperature. Although determining such absolute values isn't essential to most thermodynamic applications (entropy changes turn out to be much more significant), this situation nonetheless begs an important conceptual question: Does the entropy of a material even *have* a well-defined absolute value?

The *Third Law of Thermodynamics* answers this question in the affirmative:

Third Law of Thermodynamics: The entropy of all material systems *approaches zero* as the temperature approaches *absolute zero* on the ideal gas scale.

The third law thus establishes that the *absolute value* of the entropy of all materials approach zero at $T = 0\,\mathrm{K}$. From this fact, the absolute value of the entropy at any other temperature can be calculated from Eq. (94), using $T = 0\,\mathrm{K}$ as the lower limit of integration. For example, if a material is heated at constant pressure from $0\,\mathrm{K}$ to some finite temperature T', the absolute entropy of the material can be calculated as

$$S(T') = \int_0^{T'} \frac{dQ_P}{T} = \int_0^{T'} dT \frac{C_P(T)}{T}, \qquad P = \text{ constant.} \tag{105}$$

Equivalently, if the heating is carried out at constant volume, we obtain

$$S(T') = \int_0^{T'} dT \frac{C_V(T)}{T}, \qquad V = \text{constant.} \tag{106}$$

Now, it should be acknowledged that no real-world experiment can reach *exactly* absolute zero. As a result, any real-life application of these formulas must suffer some experimental uncertainty. However, the same statement could be made about any experimental measurement; and since low-temperature experiments now routinely reach temperatures much less than $1\,\mathrm{K}$, the error is in most cases not serious. Over the years, researchers have used these methods to establish the absolute entropy values of a vast array of different materials, providing a wealth of valuable thermodynamic reference data.

9.2 Why Zero?

Originally, it was exactly this sort of experimental tabulation that led to the establishment of the Third Law itself. Intuitively, however, the law is also justified by our understanding of entropy as the "thermal volume" of thermal states accessible to the particles of a system. As the temperature approaches absolute zero, the number of thermally accessible states decreases further and further until, at absolute zero, the system is unable to access any states except its absolute energetic minimum. It is exactly this "contraction" of all microscopic coordinates to their minimum energy state that causes the macroscopic entropy coordinate to tend toward zero.

9.3 Are There Exceptions?

This observation also suggests a possible reason that real-world measurements might deviate from the Third Law, even in the $T \to 0$ limit. Experimentally, it can often happen that, as a material is cooled, it becomes trapped in a local energy minimum that is *not* the global energy minimum. As a result, the system never releases the heat associated with the transition from this local minimum to the global minimum, and the absolute value of the entropy remains above zero. At the microscopic level, this corresponds to the fact that the microscopic coordinates have failed to completely relinquish their hold on the thermal volume.

A commonly cited example is the the carbon monoxide crystal. In a perfect CO crystal, the C-O bonds are all aligned in the same direction in a "head to tail" configuration, so that each C atom is adjacent to the O atom of its nearest neighbor, something like:

C=O··· C=O··· C=O··· C=O··· C=O··· C=O
C=O··· C=O··· C=O··· C=O··· C=O··· C=O
C=O··· C=O··· C=O··· C=O··· C=O··· C=O

However, the energy of the alternate "head-to-head" configuration – where one molecule's O atom is adjacent to its neighbor's O – is sufficiently close in energy to that of the heat-to-tail structure (and the potential energy barrier between them is sufficiently large) that inevitably in real-life experiments some of the CO molecules get trapped in the "wrong" state, like the red-colored molecules below:

C=O··· O=C··· C=O··· C=O··· C=O··· C=O
C=O··· C=O··· C=O··· C=O··· C=O··· C=O
C=O··· C=O··· C=O··· C=O··· O=C··· C=O

This leads to a small but measurable offset in the absolute value of the experimentally measured entropy compared to what would be expected from the entropies of the pure elements C and O.

Such errors do not invalidate the Third Law, so long as it is understood that the law applies only when the material is cooled slowly enough that the absolute-minimum configuration can be reached. And in most cases the errors involved are small enough to be of little practical significance.

9.4 Entropy of the Ideal Gas

Now that we have a way to define the absolute value of entropy, it becomes quite informative to inquire: What is the absolute entropy of the ideal gas?

To answer this question, we return to our "follow the energy" strategy of applying the First Law. Combining the differential form of the First Law

$$dU = dQ + dW, \tag{107}$$

with Eqs. (92) (relating the differential heat and entropy) and (89) (relating the differential work and volume), we find

$$dU = TdS - PdV \tag{108}$$

or, after some rearrangement

$$dS = \frac{dU}{T} + \frac{P}{T}dV. \tag{109}$$

Now, using Eq. (42) and (41) to simplify each term we have

$$dS = \frac{3}{2}nR\frac{dT}{T} + \frac{nR}{V}dV. \tag{110}$$

Since the two differentials are independent of each other, they can be integrated separately to obtain for a *finite* change

$$S_2 - S_1 = \frac{3}{2} nR \ln \frac{T_2}{T_1} + nR \ln \frac{V_2}{V_1}. \tag{111}$$

Intuitively, this result seems reasonable: larger temperatures and larger volumes both tend to increase the entropy since they permit the microscopic coordinates to explore a larger number of distinct states.

Something strange happens, however, if we try to apply the Third Law to this expression. According to the Law, the entropy S_1 at $T_1 = 0\,\mathrm{K}$ should be exactly zero, regardless of the volume V_1. Applying this to our expression, we obtain

$$S_2 = \frac{3}{2} nR \ln \frac{T_2}{0\,\mathrm{K}} + nR \ln \frac{V_2}{V_1}. \tag{112}$$

But $\ln \frac{T}{0}$ is infinite! Physically, this implies that the entropy of the ideal gas at any positive temperature is *infinitely larger* than its entropy at $T = 0\,\mathrm{K}$. What's going on?

Ultimately, the paradox results from the fact that no real gas follows the ideal gas equation of state in the $T \to 0$ limit. Under the assumptions of the ideal gas law, the only thing that changes about an ideal gas as the temperature is lowered is that the particles move more and more slowly. At no temperature do they ever reach an "absolute minimum" since they always possess finite energy and thus can always move a little more slowly, leading to a further release of entropy as the temperature is lowered. Real gas particles, by contrast, eventually begin to interact with each other, collapsing into first a liquid and then solid phase, and eventually converging to an absolute-minimum crystalline form. As the absolute minimum state is approached, the entropy released with each decrease in temperature becomes smaller and smaller until the entropy finally converges to its absolute minimum of zero.

Although Eq. (111) fails to accurately describe real gases as they approach absolute zero, it does provide an extremely useful approximation at ambient temperatures. Later in the course when we encounter the *chemical potential*, this result will provide the key to our understanding of both ideal gases and (more practically useful) ideal solutions.

10 Phase Transitions and Phase Diagrams

So far we've limited our discussion of thermodynamics to single-component, single-phase systems, i.e., systems where the entire sample can be described uniformly as solid, liquid, or gas. Although single-phase systems are very useful for introducing basic concepts, a great deal of interesting thermodynamics takes place in multi-phase systems, i.e., where two or more phases coexist in equilibrium. In this lecture, we'll introduce some basic vocabulary to describe such systems and the processes that transform one phase to another. Let's start with some definitions:

A thermodynamic **phase** is a region of space, throughout which all physical properties of a material are essentially uniform.

Probably the most familiar examples of thermodynamic phases are the three classic states of matter solid, liquid, and gas. But nature – and especially biology – is full of other interesting examples. To name a few:

- Most chemical reactions in biology occur in the **solution phase** – the reacting species (e.g., proteins and substrates) are not themselves "liquid", but they are solvated by the solution around them. Very often biological function is lost when enzymes crash out of solution into a solid phase.
- **Cell membranes** can exist in a variety of distinct phases (characterized by different stiffness and permeability) depending particularly on temperature and the molecular composition (e.g., cholesterol content).
- Many proteins that interact specifically with cell membranes can exist in an equilibrium between **bound and unbound** phases.

Although we'll continue to use the simpler examples of solid/liquid/gas to introduce basic concepts, these more complex biological examples will become increasingly relevant as we discuss applications of these thermodynamic principles.

10.1 Phase Transitions and Phase Diagrams

A **phase transition** is a physical process that converts matter in one phase to another.

Familiar examples of phase transitions include:

- **Melting/Freezing:** A transition from solid to liquid / liquid to solid.
- **Vaporization/Condensation:** A transition from liquid to gas / gas to liquid.
- **Sublimation/Deposition:** A transition from solid to gas / gas to solid.

For pure substances, transitions between these various phases occur at well-defined conditions. For example, under 1 atm of pressure, water freezes at 0 °C and boils at 100 °C. In fact, the conditions under which these phase transitions take place are so well-characterized (and readily available) that the Celsius scale was *defined* to produce these values!

For most substances, of course, it's not quite so easy to keep track of the conditions under which they exist in solid, liquid, or gaseous states – or where they coexist between various states. For this reason, *phase diagrams* are often used to catalog this information

A **phase diagram** is a plot showing the physical state (i.e., phase) of a given pure substance as a function of its macroscopic parameters.

Most often, phase diagrams are plotted as a function of temperature volume, and/or pressure; but, in general, any macroscopic coordinate that defines the system state can be used. For example, a phase diagram for a cell membrane might be plotted as a function of cholesterol concentration.

For a single, pure substance, the most complete type of phase diagram consists of a three-dimensional plot that can be thought of as an "existence surface" for the material. Read as a function of P and V, for example, the plot below tells us what temperature would be required to bring the substance to a given set of values for pressure and volume. Alternatively, if we specify P and T, the plot tells us what volume the system would occupy at, say, 1 atm of pressure and a temperature of 50 °C. The labels on the surface tell us what phase the given point corresponds to. Note that in such a three-dimensional phase diagram, *the equilibrium system can never occupy points above or below the surface.*

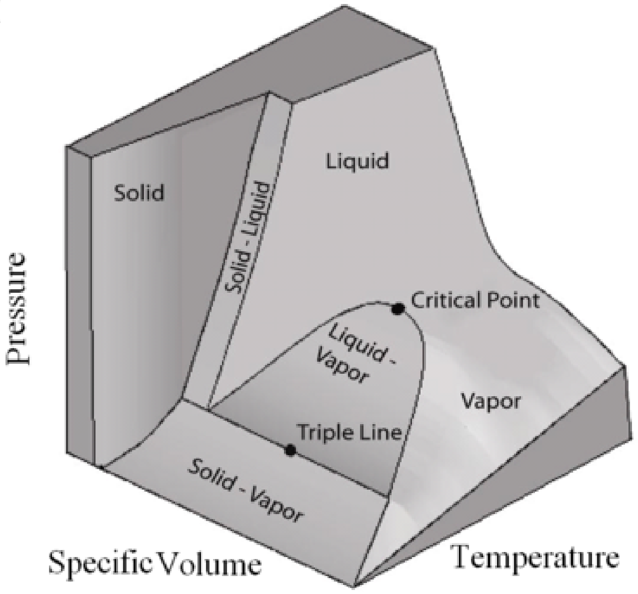


Image from: `http://www.thermalfluidscentral.org/encyclopedia/index.php/Properties_of_pure_substances`.

To simplify the visualization of such diagrams, we often project them down into only two dimensions. Let's look first at the Pressure/Temperature graph on the left. This graph is roughly what one would see if we were to view the 3D phase diagram above from a vantage point directly perpendicular to the PT plane. We no longer have any information on what the volume is at a given point, but we can still identify the various phases of the material as a function of pressure and temperature. Now, at most points in such a PT diagram, the volume is, in fact, uniquely specified in our 3D diagram as a function of pressure and temperature. The only exceptions are points along the *coexistence curves* that define the boundaries between the various phases.

A **coexistence curve** in a PT phase diagram indicates values of the pressure and temperature at which more than one phase can coexist.

Referring back to the our 3D diagram, note that there are certain values of pressure and temperature at which a line drawn parallel to the V axis would lie exactly within the surface. Physically, such a line represents the fact that, when a phase transition happens under constant pressure, the volume of the system changes but the temperature remains constant.

For example, water can exist at 1 atm and 0 °C either as a solid or as a liquid. Think of ice cubes added to a drink of water in a styrofoam cup, for example. When ice is added to the drink, the liquid and solid phases quickly establish an equilibrium, where some of the ice is melted, and the rest remains solid. If the system were perfectly insulated, this equilibrium would exist indefinitely, with both phases at exactly 0 °C. In reality, of course, the styrfoam cup does let in a small amount of heat; and, as a result, the ice gradually melts. Not until all the ice has melted, however, does the temperature of the drink increase above 0 °C, assuming, of course, that the heat enters slowly enough for the liquid and solid phases to maintain a well-defined equilibrium with each other. Although it may not be perceptible to our eyes, a very small volume change does occur within the system during this melting process. It is exactly this volume change that is represented in the 3D phase diagram by the lines parallel to the V axis: the system takes on a series of different volumes as it transitions from "mostly ice" to "all water." In the 2D projection of the PT diagram, all information on this volume change is lost. All we know is that at the given values of P and T (say 1 atm and 0 °C) the liquid and solid phases coexist.

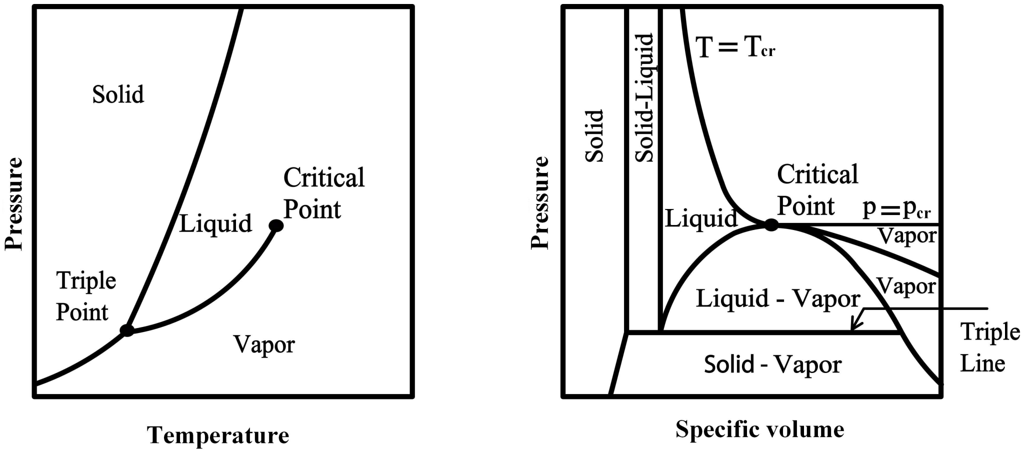


Image from: `http://www.thermalfluidscentral.org/encyclopedia/index.php/Properties_of_pure_substances`.

A similar equilibrium exists between liquid and gas phases (think, for example, of the gas/liquid equilibrium set up in an autoclave or pressure cooker). A unique feature of gas/liquid equilibrium, however, is the existence of a *critical point*, beyond which no distinction can be drawn between gas and liquid phases.

The **critical point** in a PT phase diagram is the point – defined by a **critical temperature** T_{crit} and **critical pressure** P_{crit} – at which liquid and gas phases can no longer be distinguished. If the temperature is higher than T_{crit} or the pressure is higher than P_{crit}, the volume of the substance changes smoothly as a function of P and T with no noticeable phase transition between gas and vapor.

A similarly unique feature of the PT diagram is the *triple point*:

The **triple point** in a PT phase diagram is the single point at which solid, liquid, and vapor phases all coexist with each other at equilibrium.

Thanks to the uniqueness of this phenomenon – three-way coexistence can occur only at a single temperature – the triple point is, in fact, the basis for our rigorous definition of the Kelvin temperature scale. By *definition* of the Kelvin scale, the triple point of water is *exactly* 273.16 K.

10.2 Entropy and Heat of Transition

One of the interesting features of phase transitions is that we can directly measure the *entropy change* associated with it. Intuitively, when a solid melts to a liquid, we expect it's entropy to increase – and indeed this is generally the case. In fact, we can quantify these entropy changes by measuring the *heat of transition* associated with the process.

A **heat of transition** is the amount of heat that must be added to a system to bring about a phase transition *at the transition temperature and constant pressure* (i.e., melting point, boiling point, etc.).

Some common examples include:

- Heat of fusion: the amount of heat required to melt a substance from a solid to a liquid.
- Heat of vaporization: the amount of heat required to vaporize a substance from liquid to gas.
- Heat of Sublimation: the amount of heat required to sublimate a substance from solid to gas.

Note that if the transition is carried out reversibly, this allows us to calculate the entropy change associated with the process by simply dividing the transition heat by the temperature:

$$\Delta S_{\text{trans}} = \frac{Q_{\text{trans}}}{T}. \tag{113}$$

11 Enthalpy

11.1 Thermodynamics in the Real World: Volume vs. Pressure

At the end of the last section, we introduced the concept of *transition heat* – the amount of heat released by a substance when it undergoes a reversible phase transition at the transition temperature and constant pressure. As we saw, once the transition heat for a given substance is known, it can be used to calculate the entropy change associated with the transition by simply dividing the transition heat by the transition temperature.

What's perhaps a little dissatisfying about this description (from a thermodynamics perspective) is that the transition heat can't be *directly* understood as a change in any particular state function. The heat entering or released from the system certainly contributes to changing the thermodynamic energy U, but ΔU also contains contributions from the work performed by whatever expansion or contraction of the system volume accompanies the transition.

If we were to allow the transition to occur at constant *volume*, of course, this wouldn't be a problem. Since no work can be performed at constant volume, we could always just define a "constant volume transition heat" which would directly measure the energy change associated with the transition.

The problem is that, experimentally, it's much easier to carry out processes at constant pressure than constant volume. For most of us in this course (whether chemist, biochemist, biologist, or pharmacologist) most of the processes we study occur at constant pressure: benchtop reactions; wet lab cultures; drug absorption, storage, or metabolism. Very few occur at constant volume. In other words, although the thermodynamics might be easier at constant V, the real world tends to operate at constant P.

11.2 Enthalpy – a better potential

In short, the energy turns out to be something of an inconvenient quantity to work with in the real world. And this problem is by no means limited to dealing with transition heats! As we'll see in the coming weeks, it's often extremely difficult to monitor energy changes directly in experimental measurements.

It's for this reason that a great deal of time is spent in thermodynamics introducing new *thermodynamic potentials*: quantities that behave like energy in certain circumstances but are easier to measure directly or are at least more closely related to experimental measurements. Typically how this works is that we *define* a new thermodynamic potential as the energy U plus or minus some quantity that we don't want to deal with in a particular experiment. In the context of transition heats, this leads us to define the *enthalpy*:

The **enthalpy** H of a substance is its energy plus the product of its pressure and volume:

$$H \equiv U + PV. \tag{114}$$

It may not seem obvious that this is an improvement. If our original problem was that U is hard to measure, how does it help to define a new potential in terms of that hard-to-measure quantity? But take a look at what happens when we monitor *changes* in enthalpy:

$$\begin{aligned} dH &\equiv d(U + PV) && (115)\\ &= dU + d(PV) && (116)\\ &= dU + PdV + VdP. && (117) \end{aligned}$$

In the first line, we simply applied the definition Eq. (114). In the second and third we just expanded the differential using the sum and product rules. Notice now what happens when we use Eq. (108)

to expand dU in terms of differentials in S and V:

$$\begin{aligned} dH &= dU + VdP + PdV && (118)\\ &= TdS - PdV + PdV + VdP && (119)\\ &= TdS + VdP. && (120) \end{aligned}$$

Thus, remembering that for reversible processes:

$$dQ = TdS, \tag{121}$$

we see that, for a reversible process, the enthalpy change is the sum of two terms: the *heat* TdS that enters the system and a differential pressure term VdP.

The reason this makes the enthalpy useful is that *if we carry out a process at constant pressure*, the last term (VdP) disappears, and the enthalpy change is *directly measured by the heat added to the system.*

Key Idea: The enthalpy is a *state function* whose change is, at constant pressure, *directly measured by the amount of heat entering a system.* So, for constant pressure processes, the enthalpy acts like heat – but it's also a state function!

11.3 Why does this matter?

Okay, so enthalpy is a state function. Why do we care?

Where this turns out to be really useful is in analyzing the thermodynamics of *chemical reactions.* Suppose we wanted to know, for example, how much heat is produced in the reaction

$$2\,\mathrm{C(graphite)} + 2\mathrm{H_2(g)} \longrightarrow \mathrm{C_2H_4(g)}\ .$$

This reaction isn't easy to carry out directly. (How do you get the hydrogens to interact with the graphite in such a way as to produce *only* ethylene?) What *is* easy to do, however, is to measure the enthalpy change (i.e., heat at constant pressure) for burning either the reactants

$$\begin{aligned} \mathrm{2C(graphite)} + \mathrm{2O_2(g)} &\longrightarrow \mathrm{2CO_2(g)} && \Delta H = \text{-787,365 J}\\ \mathrm{2H_2(g)} + \mathrm{O_2(g)} &\longrightarrow \mathrm{2H_2O(l)} && \Delta H = \text{-571,922 J} \end{aligned}$$

or the products $\mathrm{C_2H_2(g)} + \mathrm{3O_2(g)} \longrightarrow \mathrm{2CO_2(g)} + \mathrm{2H_2O(l)} \qquad \Delta H = \text{-1,411,577 J}.$

In each case, the value of ΔH given to the right of the equation is the measured molar enthalpy change when the reaction occurs at 298 K.

Now, because enthalpy is a state function:

- The enthalpy change for any *forward* reaction is just minus the enthalpy change for the *reverse* reaction. For example, the enthalpy change for the reaction

  $$\mathrm{2CO_2(g)} + \mathrm{2H_2O(g)} \longrightarrow \mathrm{C_2H_2(g)} + \mathrm{3O_2(g)}.$$

  is ΔH = +1,411,580 J, i.e., minus the enthalpy change for the combustion reaction shown earlier.

- The enthalpy change for a *combination* of two separate reactions is just the sum of the enthalpy changes for each individual reaction. So, for example, the enthalpy change for the reaction

  $$2\,\mathrm{C(graphite)} + \mathrm{2H_2(g)} + \mathrm{3O_2(g)} \longrightarrow \mathrm{2CO_2(g)} + \mathrm{2H_2O(l)},$$

which is just the combination of combustion of graphite and hydrogen gas (above), is $\Delta H = -787,365$ J $-571,992$ J $= -1,359,357$ J.

- The enthalpy change for a *sequence* of (balanced) reactions is again just the sum of the enthalpy changes for each individual reaction. This means that the enthalpy change for the reaction

  $$2\,\mathrm{C(graphite)} + 2\mathrm{H_2(g)} + 3\mathrm{O_2(g)} \longrightarrow 2\mathrm{CO_2(g)} + 2\mathrm{H_2O(l)} \longrightarrow \mathrm{C_2H_2(g)} + 3\mathrm{O_2(g)}\ .$$

  is just the combined enthalpy changes for each individual reaction, i.e., (from the last two bullet points) $\Delta H = -1,359,357\,\mathrm{J} + 1,411,580\,\mathrm{J} = 52,223\,\mathrm{J}$

Thus by combining the ΔH values for several *easy to measure* reactions, we can calculate ΔH for a single *difficult to measure* reaction, i.e., the production of ethylene from graphite and hydrogen gas. And all because enthalpy is a state function!

11.4 Here be dragons!

We've avoided two complicating factors in our discussion:

- First, I've been very careless in specifying the *initial and final states* of reactants and products in these reactions. In general, the enthalpy of all materials depends on their temperature and pressure, as well as their physical state (solid, liquid, gas, solution, etc.). The goal of this lecture was to introduce the *concept* of enthalpy as a state function without getting too far "down into the weeds" with specification of the state. We'll discuss this problem much more carefully in the next lecture.

- We started off our lecture with noting that transition heats (i.e., enthalpies) can be used to calculate the *entropy* change associated with a given phase transition. To head off any confusion in the future, let me point out that we usually **cannot** do the same with the enthalpies of reaction that we've just been discussing. Only in the special case that a reaction is performed *reversibly* (*and* at constant pressure) can we use the enthalpy change to calculate ΔS. Most experimental reactions are *not* carried out reversibly and thus the associated reaction enthalpies are useless for the purpose of calculating ΔS. In later lectures, we'll return to the idea of reversible chemical reactions and see how it is possible experimentally to determine ΔS.

12 Measuring and Tabulating Enthalpy Changes

In the last lecture, we introduced a new state function called the *enthalpy* H which (we claimed) is easier to measure experimentally than the energy U. We showed, in particular, how enthalpy's status as a state function makes it easier to tabulate ΔH values for reactions that may be hard to conduct directly under real-life experimental conditions. In this lecture, we'll talk about some of the practical issues associated with enthalpy measurements and conventions on how published enthalpy values are tabulated.

12.1 Reversibility

Let's first cover an important conceptual issue: reversibility in enthalpy measurements. We noted in the last lecture that if a process is carried out *reversibly and at constant pressure* then

$$dH = TdS = dQ. \tag{122}$$

Thus for reversible, constant-pressure reactions, the enthalpy change is exactly the heat input to the system:

$$\Delta H = Q. \tag{123}$$

In the real world, though, it's difficult to carry out reactions reversibly. It's easy to combust hydrogen gas with oxygen to produce water. But the reaction is highly explosive! It's not so easy to carry out the process reversibly. So how do we measure enthalpy changes in *irreversible* chemical reactions?

Actually, the situation is not so bad. As it turns out, Eq. (123) is valid even for *irreversible* processes (when Eq. (122) is *not* valid), so long as *the pressure remains well-defined and constant throughout the process.* Note from the definition $H = U + PV$ that for any change in state

$$\Delta H = \Delta U + \Delta\left(PV\right). \tag{124}$$

Now, if the initial and final pressures are the same, $\Delta(PV) = P\Delta V$, so that

$$\Delta H = \Delta U + P\Delta V. \tag{125}$$

This equation is valid regardless of the path that connects the two states! Now, using the First Law, we can expand

$$\begin{aligned} \Delta H &= Q + W + P\Delta V \quad &(126)\\ &= Q - P\Delta V + P\Delta V \quad &(127)\\ &= Q. \quad &(128) \end{aligned}$$

The only assumption we have made here is that the pressure remains well-defined and constant throughout the process. (This is a necessary assumption for replacing W with $-P\Delta V$.) This condition is always satisfied for reversible processes, but it is also satisfied for many *irreversible* processes.

As an example, suppose we want to know the enthalpy change for the reaction of NaOH with HCl (both in aqueous solution). A very reasonable measurement could be achieved by simply

1. Placing the NaOH solution in a styrofoam cup with a loose-fitting lid,
2. Measuring the initial temperature of the NaOH solution,
3. Bringing the HCl solution to the same initial temperature,
4. Adding the HCl to the cup,

5. Measuring the final temperature of the solution, and
6. Using the measured temperature change, along with the known heat capacity of water, to calculate the amount of heat generated in the reaction.

So long as the lid of the cup fits loosely enough that air is free to enter and leave our system, the pressure on the solution will remain essentially constant during this process. Thus the heat evolved during the action can accurately be interpreted as the enthalpy change for the reaction under the specified conditions.

Note that, in order for the evolved heat to be directly measured as heat, we ought really to carry out the measurement in a slightly different way: instead of monitoring the temperature change inside the cup, we should allow the warm cup to gradually heat up some other reference system (say a separate water bath, outside the cup) until it returns back to its initial temperature. Then the thermal energy evolved during the reaction would be properly measured as (negative) heat, i.e., through the flow of that energy out of the system. In practice, though, monitoring the temperature change is very nearly just as accurate since (unless the concentration of NaOH and/or HCl is very, very high, so that the heat capacity of the solution differs significantly from that of pure water) the known heat capacity of water allows us to calculate how much heat *would* flow out of the system if we allowed it to do so.

In describing the enthalpy changes associated with different chemical reactions, it's useful to define two short-hand descriptors:

An **exothermic** process is one that, at constant pressure, releases heat into the environment, i.e., for which $\Delta H < 0$.

An **endothermic** process is one that, at constant pressure, absorbs heat from the environment, i.e., for which $\Delta H > 0$.

12.2 Standard States

As discussed in the last lecture, the beautiful thing about entropy is that – because it's a state function – we can use tabulated values of ΔH measured for experiments that are *easy* to perform to calculate ΔH values for experiments that are *hard* to perform. For this process to be practically useful, however, we need to have some conventions on the conditions under which experiments will be performed and recorded. After all, there are an infinite number of different variations on temperature, pressure, concentration, etc. under which a reaction could be carried out. If all the tabulated values of ΔH applied to different reaction conditions, there would be no way to combine them, since we'd have no way of making sure that the product state for one reaction really matched the initial state for the next.

To avoid such problems, scientists have adopted a (nearly) universal convention on the definition of a thermodynamic *standard state*:

The thermodynamic **standard state** of a material is defined as follows:

- For **liquids and solids**, the standard state is the state of the pure substance in its most stable form at 10^5 Pa.
- For **gases**, the standard state is a *hypothetical* state in which the gas follows the ideal gas law at a pressure of 10^5 Pa. (See discussion below.)
- For **solutes** in solution, the standard state is a *hypothetical* state in which the solute behaves like an ideal solution at 1 molal concentration and 10^5 Pa pressure (see discussion below).

Note that in all cases the standard state defines only the form and the pressure; the temperature must be specified separately. (Thus one can speak of the standard state at 298 K, as distinct from the standard state at 77 K.)

The first of these conditions is straightforward enough. For example, the standard state of water at 298 K is liquid water under 10^5 Pa, while the standard state of water at 270 K is ice under 10^5 Pa of pressure.

The situation is more complicated for gases and solutes. In both cases, the standard state is defined as a *hypothetical state* where the material follows idealized equations of state. The first of these, the ideal gas law, we've already discussed in some detail; the laws for ideal solutions we'll cover later in the semester. The reason for choosing these hypothetical states (as opposed to the real state of the system at, say 10^5 Pa or 1 molal concentration) is that it creates a common reference state for all substances, regardless of how far they actually deviate from ideality. Since real gases obey the ideal gas law at sufficiently low pressure, and since real solutions obey the ideal solution laws at sufficiently low concentration, thermodynamic properties in the hypothetical "standard state" can be tabulated by

1. Measuring the *real* thermodynamic properties of the substance in the low pressure/low concentration limit and then
2. Using the ideal gas/ideal solution laws to extrapolate those properties to the standard conditions of 10^5 Pa and/or 1 molal concentration.

These extrapolated values define the hypothetical "standard state" for real gases and solutions.

Note that, under the ideal gas law, the enthalpy is actually a function *only* of temperature:

$$H = U + PV = \frac{3}{2}nRT + nRT = \frac{5}{2}nRT, \qquad \text{Ideal gas.} \tag{129}$$

Thus the enthalpy of an ideal gas at a fixed temperature and 10^5 Pa pressure is identical to the enthalpy in the low-pressure limit and the same temperature. As a result, enthalpy changes measured in the standard state are equivalent to enthalpy changes in the low-pressure limit at the same temperature. So, as far as enthalpy change is concerned, the thermodynamic standard state for a gas could just as well be defined as the real gas in the low-pressure limit and a fixed temperature.

Unless explicitly stated, we will in this course always assume that gases obey the ideal gas law and that solutions obey the ideal solution laws (to be introduced later). Thus, unless otherwise noted, you'll be able to ignore the distinction between these hypothetical standard states and the *real* state of a gas at 10^5 Pa or the *real* state of a solution at 1 molal concentration. We draw the distinction carefully here only so that, if you encounter situations in the future where the distinctions matter, you won't be taken by surprise.

In defining the standard state, we repeatedly encountered the quantity 10^5 Pa, nearly the same as 1 atm, which is often assigned its own name as a unit of pressure:

One **bar** of pressure is defined to be 10^5 Pa, a rough measure of average atmospheric pressure.

12.3 Standard Enthalpies

With these definitions for thermodynamic standard states, we can introduce two important forms in which enthalpy changes are commonly tabulated:

The **standard enthalpy of reaction** $\Delta H^{\circ}_{\text{rxn}}$, for any given chemical reaction, is the enthalpy change per molar equivalent of the reaction from the reactants in their standard states to the products in their standard states.

The **standard enthalpy of formation** $\Delta H_\mathrm{f}^\mathrm{o}$ for a compound is the enthalpy change involved in forming one mole of the compound at its standard state from its elemental components in their standard states.

The phrase "per molar equivalent" here means that if the stoichiometric coefficient of a given reactant is ν, then ν moles of that reactant are to be consumed during the reaction. E.g., in one molar equivalent of the reaction $O_2 + 2H_2 \longrightarrow 2H_2O$, one mole of O_2 and two moles of H_2 will be consumed to form two moles of H_2O.

12.4 Reaction Enthalpies from Enthalpies of Formation

As we've pointed out repeatedly, the useful thing about H being a state function is that we can calculate ΔH values for a given reaction as a combination of ΔH reactions that, in combination, produce the same transformation. One of the most useful such calculation schemes is to calculate the enthalpy change for a given reaction by taking a difference between the enthalpies of formation of its products and reactants. If the equation is balanced, the enthalpies of the pure elements will necessarily cancel each other out in this process, leaving only the ΔH_rxn for the desired reaction.

To be precise, consider the generic chemical reaction

$$\nu_1 \mathrm{X}_1 + \nu_2 \mathrm{X}_2 + ... + \nu_{N_\mathrm{r}} \mathrm{X}_{N_\mathrm{r}} \longrightarrow \nu_{N_\mathrm{r}+1} \mathrm{X}_{N_\mathrm{r}+1} + \nu_{N_\mathrm{r}+2} \mathrm{X}_{N_\mathrm{r}+2} + ... + \nu_{N_\mathrm{r}+N_\mathrm{p}} \mathrm{X}_{N_\mathrm{r}+N_\mathrm{p}} \, .$$

Here the integers ν_1, ν_2, etc. are the stoichiometric coefficients for the N_r reactants X_1, X_1, etc., while $\nu_{N_\mathrm{r}+1}$, $\nu_{N_\mathrm{r}+2}$, etc. are the stoichiometric coefficients for the products $X_{N_\mathrm{r}+1}$, $X_{N_\mathrm{r}+2}$, etc. The enthalpy of reaction can then be calculated as

$$\Delta H_\mathrm{rxn}^\mathrm{o} = \sum_i^{\text{products}} \nu_i \Delta H_{\mathrm{f},i}^\mathrm{o} - \sum_i^{\text{reactants}} \nu_i \Delta H_{\mathrm{f},i}^\mathrm{o}. \tag{130}$$

Here $\Delta H_{\mathrm{f},i}^\mathrm{o}$ is the standard enthalpy of formation of species i. Note that, strictly speaking, this relation is valid only if the various reactants and products are physically *separated from each other* both before and after the reaction. In general, the total enthalpy of a mixture of compounds (even at standard state) may differ from what is expected based on the various standard enthalpies of formation, due to interactions between the different species. Thus, Eq. (130) is strictly valid only if all reactants are separated from each other in the initial and final state and the products are separated from each other in the final state. For many systems, however, this distinction is not of great significance, and in this course we'll generally ignore such corrections.

13 The Gibbs Free Energy: Indicator for Spontaneity

13.1 The Second Law

The Second Law of Thermodynamics states that for an *isolated* system, irreversible processes are always associated with an *increase* in the system entropy. We mentioned at the time (though it wasn't of much practical use then) that this fact can be used as a criterion for determining whether or not a given physical process can actually happen in the real world: processes during which an isolated system increases its entropy are irreversible or *spontaneous*. In the real world, only spontaneous processes actually occur; *reversible* processes are an idealization of a spontaneous process in which the isolated-system entropy change goes to zero. Processes for which entropy *increases* can never happen in isolated systems.

Although the Second Law *could* be used directly to predict whether a process is spontaneous or not, this isn't usually very practical: first, because entropy changes aren't very easy to measure experimentally and, second, because most real experiments aren't conducted on isolated systems. Today's lecture will introduce a new thermodynamic potential (similar to the energy or enthalpy) that allows us to make predictions about spontaneity under the much more practical conditions of constant temperature and constant pressure.

13.2 Processes at Constant Temperature and Pressure

Before we get into the weeds, let's first make clear what exactly we mean by "constant temperature and pressure conditions." In thermodynamics, we typically divide the universe into two parts: a "system" under study and an "environment" that surrounds and may or may not interact with the system. When we refer to processes carried out "at constant temperature and pressure conditions", what we really mean is that the *environment* maintains a constant temperature and pressure. Specifically, we are interested in systems that can exchange heat and PV work with the environment and for which the temperature and pressure *of the environment* remain constant throughout the process. Thus if heat or PV work is performed *on the environment* the associated energy changes are reversible *for the environment*. As the system approaches equilibrium with its environment, its interior temperature and pressure will eventually equilibrate to the constant temperature and pressure of the environment; but the internal T and P may not necessarily stay constant during the course of the process, and the physical processes internal to the system may or may not be reversible. Only at the beginning and end are T and P required to be the same as the constant environment, and only with the environment are PV work and heat necessarily exchanged reversibly.

As a concrete example, consider the styrofoam cup we discussed in a previous lecture. If a chemical reaction carried out inside the cup produces heat, the temperture *inside* the cup will very likely *not* be constant throughout the reaction. However, since the styrofoam cup releases that heat to the environment very slowly – and because the large volume of air in the room is presumably heated only very slightly by the reaction – the temperature and pressure of the air *around* the cup stays essentially constant. Thus, as far as the environment is concerned, the process occurs reversibly at constant T and P. Within the cup – and within the thin boundary layer of styrofoam that forms the interface with the environment – the process may very well not be reversible and may very well not occur with constant temperature and pressure. As we'll see, however, it's possible to predict sponteneity of even such poorly defined processes, so long as the *environment* is well-behaved.

13.3 The Gibbs Free Energy

Now, as stated previously, the introduction of new thermodynamic potentials usually involves taking an existing potential and getting rid of some contributions to it that are either inconvenient or uninteresting. And this is exactly what we're going to do here. In looking for a "spontaneity predicting" potential for constant P/constant T processes, we might think at first of the enthalpy.

We know already that the change in enthalpy is equivalent to the heat passing into the system at constant P. Intuitively, we tend to associate spontaneous processes with the *production* of heat (as in the reaction of acids and bases), i.e., to the condition that ΔH be negative. Physically, this is grounded in the fact that heat *released* to the environment *increases* the entropy of the environment. Because heat is released to the environment reversibly, the entropy change for a constant-temperature environment

$$\Delta S_{\mathrm{env}} = -\frac{\Delta H_{\mathrm{sys}}}{T} \tag{131}$$

is positive when ΔH_{sys} is negative (i.e., for exothermic processes) and negative when ΔH_{sys} is positive (i.e., for endothermic processes). This increase in entropy of the *environment* helps to make the overall entropy change for the system + environment more positive. For this reason, exothermic processes tend to be spontaneous at constant T. But this correlation is not universal: there are exothermic reactions that are *not* spontaneous, and there are *endothermic* reactions that *are* spontaneous.

The problem, of course, is that the entropy increase of the environment in an exothermic process can be offset by a *decrease* in the entropy of the system itself. Taking the system and environment together to form an isolated "super system" the Second Law dictates that the process will be spontaneous if

$$\Delta S_{\mathrm{sys}} + \Delta S_{\mathrm{env}} = \Delta S_{\mathrm{sys}} - \frac{\Delta H_{\mathrm{sys}}}{T} > 0 \tag{132}$$

or, equivalently, if

$$\Delta H_{\mathrm{sys}} - T\Delta S_{\mathrm{sys}} < 0. \tag{133}$$

Since T is constant, we can bring the factor of T inside the Δ sign to find that

$$\Delta H_{\mathrm{sys}} - \Delta\left(TS_{\mathrm{sys}}\right) = \Delta(H - TS) < 0 \tag{134}$$

is a universal criterion for the spontaneity of processes in a system interacting with an environment at constant temperature and pressure. Effectively, in subtracting the factor TS, we're subtracting off the contribution of the *system* entropy change from the *system* enthalpy change, a metric for the entropy change of the environment. What's left behind is the *total entropy change of the system and its environment*, multiplied by an overall factor of $-T$.

The net result is the definition of a new thermodynamic potential:

The **Gibbs Free Energy** of a system is defined by the relation

$$G \equiv H - TS = U + PV - TS. \tag{135}$$

As we've seen, the Second Law implies that

In a system interacting with a constant temperature, constant pressure environment, a given process is spontaneous if and only if

$$\Delta G_{\mathrm{sys}} < 0. \tag{136}$$

The process is reversible if $\Delta G = 0$.

Just as we earlier introduced the terms "exothermic" and "endothermic" to describe the enthalpy change of a process, it's useful to introduce two new terms to describe Gibbs free energy changes:

An **exergonic** process is one for which $\Delta G < 0$.

An **endergonic** process is one for which $\Delta G > 0$.

13.4 Standard Free Energies

This property of the Gibbs free energy makes it exceptionally useful in characterizing the thermodynamics of chemical reactions. To make use of this property, however, we have to have simple ways to calculate ΔG for a given reaction under various circumstances. Just like with ΔH, it turns out that, in many cases, it's possible to calculate ΔG values from standardized ΔG values at standard conditions. To this end we define

The **Standard Gibbs Free Energy** $\Delta G^{\circ}_{\mathrm{rxn}}$ of a reaction is the Gibbs free energy change per molar equivalent of the reaction from the reactants in their standard states to the products in their standard states.

The **Standard Gibbs Free Energy of Formation** $\Delta G^{\circ}_{\mathrm{f}}$ for a compound is the Gibbs free energy change involved in forming one mole of the compound at its standard state from its elemental components in their standard states.

In the coming lectures, we'll see how reaction free energies under nonstandard conditions can be calculated from these standard values.

13.5 Beware the Temperature!

In calculating free energy (or enthalpy) changes from standard values, it's important to keep in mind that *standard enthalpy and free energy values are functions of temperature.* A frequent error in applying thermodynamic equations is to assume that a ΔG° value measured at 298 K can be applied to reactions at other temperatures. Although ΔH° and ΔG° values are often recorded at 298 K, the temperature is *not* part of the definition of a standard state. Standard enthalpy and free energy values can be recorded at any temperature, so long as the reactants and products are in their standard states *for that temperature.*

Part of the confusion here is that many literature measurements are reported at *standard temperature and pressure* (STP), a set of conditions that does define both a pressure (1 bar) and a standard reference temperature. Unfortunately, even the definition of STP conditions is somewhat controversial, with different organizations using different values of the temperature as "standard". (See, for example, the Wikipedia article on the topic: `https://en.wikipedia.org/wiki/Standard_conditions_for_temperature_and_pressure`.) In this course, we will never use the designation "STP" to define experimental conditions. Only the phrase "standard state" – as defined in Section 12.2 – will be used to define pressure and physical state; the temperature will always be specified explicitly.

14 The Gibbs Free Energy of Ideal Gases

14.1 ΔG under Nonstandard Conditions

Now that we know something about the physical significance of the free energy, let's look at how we can determine it for a given chemical reaction. In particular, we want to know: Given the *standard* free energy value $\Delta G^{\mathrm{o}}_{\mathrm{rxn}}$ for a given reaction, how can we calculate the free energy under other, *non*standard conditions? Two circumstances in which we can answer this question exactly are for ideal gases and ideal solutions. We'll examine first Gibbs free energies for ideal gas mixtures and then discuss the properties and free energy changes of ideal solutions.

14.2 Ideal Gas Mixtures and Partial Pressure

Before we discuss free energy changes, we need first to establish some ground rules for dealing with ideal gas mixtures. Recall that our fundamental assumption in dealing with ideal gases is that *the gas particles do not interact with each other.* As a result, in a fixed volume V that contains a mixture of several ideal gases at temperature T, the pressure exerted on the walls by each gas is exactly the same as the pressure it would exerted on the walls in the *absence* of all the other gases. Let's say, for example, that the volume contains n_1 moles of gas number 1, n_2 moles of gas number 2, and so on for the remaining gases. Then each gas exerts a pressure $P_i = \frac{n_i RT}{V}$ on the walls of the container, just as it would if the other gases were absent. Moreover, since the forces exerted by each gas are additive, the total pressure is just the sum of the partial pressures, i.e.,

$$P_{\mathrm{tot}} = \sum_i P_i. \tag{137}$$

In summary:

The **partial pressure** of a particular gas in an ideal gas mixture

$$P_i = \frac{n_i RT}{V} \tag{138}$$

is the contribution the particles of that gas make to the total pressure

$$P_{\mathrm{tot}} = \sum_i P_i = \frac{n_{\mathrm{tot}} RT}{V} \tag{139}$$

where $n_{\mathrm{tot}} = \sum_i n_i$ is the total number of gas particles in the system.

14.3 Free Energy of an Ideal Gas

With this in mind, let's see what we can say about the the difference between the free energy of an ideal gas mixture at arbitrary partial pressures, relative to the free energy at standard conditions. To begin, recall from Eq. (111) that, at constant temperature, we can calculate the entropy change for a single ideal gas between any two states as

$$S_2 - S_1 = \frac{3}{2} nR \ln \frac{T_2}{T_1} + nR \ln \frac{V_2}{V_1}. \tag{140}$$

If the two states have the same temperature, this simplifies to

$$S_2 = S_1 + nR \ln \frac{V_2}{V_1} \tag{141}$$

$$= S_1 - nR \ln \frac{P_2}{P_1}, \tag{142}$$

where in the second line we have used the ideal gas equation of state to make the replacement $\frac{V_2}{V_1} = \frac{nRT \cdot P_2}{P_1 \cdot nRT} = \frac{P_1}{P_2}$. In particular, if S_1 is the entropy of the gas under standard conditions, then at any other pressure (but the same temperature)

$$S = S^{\rm o} - nR \ln \frac{P}{P^{\rm o}} \tag{143}$$

where $P^{\rm o} = 1$ bar is the pressure at standard state.

Now, since the enthalpy of an ideal gas

$$H = U + PV = \frac{3}{2}nRT + nRT = \frac{5}{2}nRT \tag{144}$$

is a function *only* of temperature (i.e., is independent of pressure so long as T remains constant), we can calculate

$$G = H - TS = H^{\rm o} - T\left(S^{\rm o} - nR \ln \frac{P}{P^{\rm o}}\right) \tag{145}$$

$$= H^{\rm o} - TS^{\rm o} + nRT \ln \frac{P}{P^{\rm o}} \tag{146}$$

$$= G^{\rm o} + nRT \ln \frac{P}{P^{\rm o}} \tag{147}$$

In what follows, it will be convenient to work with the *molar* free energy of the gas

$$\bar{G} \equiv \frac{G}{n} \tag{148}$$

which can be calculated as

$$\bar{G} = \bar{G}^{\rm o} + RT \ln \frac{P}{P^{\rm o}}. \tag{149}$$

14.4 Reaction Free Energies of Ideal Gas Mixtures

Given this start, it's straightforward to calculate reaction free energies for ideal gas mixtures at arbitrary pressures. Consider the generic reaction

$$\sum_i^{\text{reactants}} \nu_i X_i \longrightarrow \sum_i^{\text{products}} \nu_i X_i, \tag{150}$$

where we assume that all species are in the gas phase. Since the ideal gas particles don't interact with each other, the total Gibbs free energy change for the reaction is just the total Gibbs free energy of one stoichiometric equivalent of products minus the total Gibbs free energy of one stoichiometric equivalent of reactants. The free energy for the reaction is then

$$\Delta G_{\rm rxn} = \sum_i^{\text{products}} \nu_i \bar{G}_i - \sum_i^{\text{reactants}} \nu_i \bar{G}_i. \tag{151}$$

Now, using Eq. (147) to calculate the free energy of each component, this implies that

$$\Delta G_{\mathrm{rxn}} = \sum_i^{\text{products}} \left(\nu_i \bar{G}_i^{\mathrm{o}} + RT \ln \frac{P_i}{P^{\mathrm{o}}} \right) - \sum_i^{\text{reactants}} \left(\nu_i \bar{G}_i^{\mathrm{o}} + RT \ln \frac{P_i}{P^{\mathrm{o}}} \right) \tag{152}$$

$$= \Delta G_{\mathrm{rxn}}^{\mathrm{o}} + RT \sum_i^{\text{products}} \nu_i \ln \frac{P_i}{P^{\mathrm{o}}} - RT \sum_i^{\text{reactants}} \nu_i \ln \frac{P_i}{P^{\mathrm{o}}} \tag{153}$$

$$= \Delta G_{\mathrm{rxn}}^{\mathrm{o}} + RT \ln \left(\frac{\prod_i^{\text{products}} \frac{P_i^{\nu_i}}{(P_i^{\mathrm{o}})^{\nu_i}}}{\prod_i^{\text{reactants}} \frac{P_i^{\nu_i}}{(P_i^{\mathrm{o}})^{\nu_i}}} \right). \tag{154}$$

If each partial pressure P_i is reported in bar, the factors P^{o} can be set to unity and we obtain simply

$$\Delta G_{\mathrm{rxn}} = \Delta G_{\mathrm{rxn}}^{\mathrm{o}} + RT \ln \frac{\prod_i^{\text{products}} [P_i]^{\nu_i}}{\prod_i^{\text{reactants}} [P_i]^{\nu_i}}, \qquad \text{all pressures in bar.} \tag{155}$$

Here the notation $[P_i]$ means "the numerical value without units," where the units must be specified explicitly in the context. Note that this would **not** be correct if some other reference were used for pressure, e.g., Torr or Pa, since then the reference pressure P^{o} would have a numerical value different from 1 and could no longer be dropped from Eq. (154).

14.5 Entropy of Mixing

Although not necessary to obtain these results, it's worth noting here that formula Eq. (142) can be applied very simply to calculate the *entropy of mixing* of a mixture of ideal gases. Imagine that we have initially two ideal gases in separate chambers and at the same pressure P_{tot} and temperature T. If we now open a stopcock that connects the two chambers, the gases will irreversibly mix with each other, as particles from each side diffuse into the other chamber. The total pressure remains constant during the mixing process at

$$\frac{n_1 RT}{V_1 + V_2} + \frac{n_2 RT}{V_1 + V_2} = \frac{n_1 + n_2}{V_1 + V_2} RT = \frac{\frac{P_{\mathrm{tot}} V_1}{RT} + \frac{P_{\mathrm{tot}} V_2}{RT}}{V_1 + V_2} RT = P_{\mathrm{tot}}. \tag{156}$$

What is the entropy change associated with this process?

Because the gases are assumed to be ideal, we know that the particles don't interact with each other. As a result, the entropy of the final mixture must be exactly the sum of the entropies of each gas *occupying the new total volume of the combined chambers*. Since each gas begins at pressure P_{tot} and ends at its own partial pressure $P_i = \frac{n_i RT}{V_{\mathrm{tot}}}$, Eq. (142) implies that

$$\Delta S_{\mathrm{mix}} = \Delta S_1 + \Delta S_2 = -n_1 R \ln \frac{P_1}{P_{\mathrm{tot}}} - n_2 R \ln \frac{P_2}{P_{\mathrm{tot}}}. \tag{157}$$

More generally:

The **entropy of mixing** of an ideal gas mixture is the change in system entropy when the separate gaseous components are combined at constant temperature and pressure:

$$\Delta S_{\mathrm{mix}} = \sum_i \Delta S_i = -\sum_i n_i R \ln \frac{P_i}{P_{\mathrm{tot}}}. \tag{158}$$

Expressed differently, this result can be written

$$\Delta S_{\mathrm{mix}} = \sum_i \Delta S_i = -\sum_i n_i R \ln x_i \tag{159}$$

where x_i is the mole fraction of species i:

The **mole fraction** x_i of a given species i in a mixture is

$$x_i = \frac{n_i}{n_{\text{tot}}}, \tag{160}$$

where $n_{\text{tot}} = \sum_i n_i$ is the total number of moles in the system.

We'll make use of this last result in the next lecture when we study the properties of ideally dilute solutions.

15 Gibbs Free Energies of Ideally Dilute Solutions

15.1 Motivation

The relation just derived for the Gibbs free energy of ideal gases as a function of pressure are extremely useful in making predictions about the spontaneity of chemical reactions in the gas phase. By measuring $\Delta G^{\mathrm{o}}_{\mathrm{rxn}}$ at standard pressures, the free energy change for a reaction at any other set of pressures (and the same temperature) can be calculated directly, allowing us to predict (remarkably accurately, in many cases) the conditions under which a reaction will or will not "go".

However, most of wet lab chemistry (and biochemistry and molecular biology) is performed in the condensed phase. This is particularly true in questions of cellular metabolism where nearly all of the key reactions occur in solution. The question then arises: Can we obtain a similar expression for reaction free energies as a function of *concentration* in solution?

The answer turns out to be "yes", with some caveats. Molecular dynamics in the condensed phase are *much* more difficult to treat from a "first principles" perspective, since (in contrast to our relatively simple derivation of the ideal gas law earlier in the semester) intermolecular interactions most certainly cann*not* be neglected in solution. In fact, intermolecular interactions are what prevents a solution from evaporating into the gas phase! Nonetheless, with some simple assumptions about *how* the various species in solution interact with each other, we can derive a relation for the Gibbs free energy of a solution that is remarkably similar to that for the ideal gas.

15.2 Ideally Dilute Solutions

The fundamental assumption we'll make in handling solution thermodynamics is that the solution is *very dilute*:

An **Ideally Dilute Solution** is a solution in which the concentration of one single species, termed the *solvent* is much greater than the concentration of any of the other species, each of which is termed a *solute*. In particular, the concentration of each solute is assumed to be sufficiently low that the energy and volume of the solution are *additive* in the number of moles of each species:

$$U = \sum_i n_i u_i(T) \tag{161}$$

$$V = \sum_i n_i v_i(T) \tag{162}$$

and that the entropy is additive up to an ideal-gas correction for the entropy of mixing

$$S = \sum_i n_i \sigma_i(T) - R \sum_i n_i \ln x_i \tag{163}$$

where x_i is the **mole fraction** of species i. Here $u_i(T)$, $v_i(T)$, and $\sigma_i(T)$ are (respectively) molar energy, volume, and entropy functions for each species i that depend on temperature but not on concentration.

The first two of these equations are straightforward to understand physically: Each species in the solution carries with it an associated molar energy and volume reflecting the local physical properties of the particles of each species. Averaged solute-solvent interactions are lumped into these terms and, as a result, they are specific to the solvent/solute system under consideration. What we're neglecting in these equations is any interaction *between solute particles*; each solute is assumed to interact independently with the solvent around it, and this is why the equations are *linear* in n_i.

The last equation [Eq. (163)] looks a bit more complicated but also has a simple physical explanation. The first term is a sum of molar entropy functions, just like the energy and volume functions in the preceding equations. The last term represents the *entropy of mixing* for distributing the various particles throughout the system. In fact, this term is identical to the ideal gas entropy of mixing in Eq. (159).

Note that, since these expressions are to be applied in the limit of infinite dilution, the molar functions u_{solv}, v_{solv}, and σ_{solv} must be identical to the corresponding functions for the *pure* solvent. This is *not* necessarily the case for the *solutes.* In general, these functions may differ substantially from their values in the pure solute.

15.3 Reaction Free Energies

With these definitions, we can easily calculate the Gibbs free energy of an ideally dilute solution:

$$\begin{aligned} G &\equiv U + PV - TS && (164) \\ &= \sum_i n_i u_i(T) + P \sum_i n_i v_i(T) - T \sum_i n_i \sigma_i(T) + \sum_i RT n_i \ln x_i && (165) \\ &= \sum_i n_i \bar{G}_i, && (166) \end{aligned}$$

where on the last line we have introduced the *molar Gibbs free energy* of each species:

$$\bar{G}_i = u_i + P v_i - T\sigma_i + RT \ln x_i. \qquad (167)$$

Now, the first three terms in this expression are independent of concentration and are thus identical (for a fixed temperature) at standard and nonstandard conditions. Only the last term varies with concentration and so determines the difference between standard and nonstandard reaction free energies. We can thus calculate the Gibbs free energy of reaction as

$$\begin{aligned} \Delta G_{\text{rxn}} - \Delta G^{\text{o}}_{\text{rxn}} &= \sum_i^{\text{products}} \nu_i \bar{G}_i - \sum_i^{\text{reactants}} \nu_i \bar{G}_i - \left(\sum_i^{\text{products}} \nu_i \bar{G}^{\text{o}}_i - \sum_i^{\text{reactants}} \nu_i \bar{G}^{\text{o}}_i \right) && (168) \\ &= \sum_i^{\text{products}} \nu_i RT \left(\ln x_i - \ln x^{\text{o}}_i \right) - \sum_i^{\text{reactants}} \nu_i RT \left(\ln x_i - \ln x^{\text{o}}_i \right) && (169) \\ &= RT \ln \frac{\prod_i^{\text{products}} \left(\frac{x_i}{x^{\text{o}}_i} \right)^{\nu_i}}{\prod_i^{\text{reactants}} \left(\frac{x_i}{x^{\text{o}}_i} \right)^{\nu_i}}. && (170) \end{aligned}$$

Using this formula, we can calculate reaction free energies in ideal solutions at *arbitrary concentrations*, assuming we know the value of the reaction free energy at standard conditions.

15.4 Concentration Units

For real-world applications, it's often more useful to deal in concentration units like molarity and molality than in the mole fractions x_i in the last equation. Recall the definitions:

The **molarity** C_i^{molar} of solute i in solution is the number of moles of solute per liter of solution.

The **molality** C_i^{molal} of a solute i in solution is the number of moles of solute per kilogram of solvent.

In dilute solutions (where the total number of moles differs only negligibly from the number of moles of solvent), the three concentration units can be interconverted as follows:

$$\frac{1000 \cdot x_i}{M_{\text{solv}}} = C_i^{\text{molal}} = \frac{C_i^{\text{molar}}}{\rho_{\text{soln}}} \tag{171}$$

where M_{solv} is the molecular weight of the solvent in g/mole, and ρ_{soln} is the density of the solution in kg/L. Now, since in our free energy formula, the mole fraction always appears in the ratio $\frac{x_i}{x_i^{\text{o}}}$, the conversion factors in switching between concentration units always cancel with each other. Thus Eq. (170) can be written in the three equivalent forms:

$$\begin{aligned} \Delta G_{\text{rxn}} &= \Delta G^{\text{o}}_{\text{rxn}} + RT \ln \frac{\prod_i^{\text{products}} \left(\frac{x_i}{x_i^{\text{o}}}\right)^{\nu_i}}{\prod_i^{\text{reactants}} \left(\frac{x_i}{x_i^{\text{o}}}\right)^{\nu_i}} && (172) \\ &= \Delta G^{\text{o}}_{\text{rxn}} + RT \ln \frac{\prod_i^{\text{products}} \left(\frac{C_i^{\text{molar}}}{C_i^{\text{o,molar}}}\right)^{\nu_i}}{\prod_i^{\text{reactants}} \left(\frac{C_i^{\text{molar}}}{C_i^{\text{o,molar}}}\right)^{\nu_i}} && (173) \\ &= \Delta G^{\text{o}}_{\text{rxn}} + RT \ln \frac{\prod_i^{\text{products}} \left(\frac{C_i^{\text{molal}}}{C_i^{\text{o,molal}}}\right)^{\nu_i}}{\prod_i^{\text{reactants}} \left(\frac{C_i^{\text{molal}}}{C_i^{\text{o,molal}}}\right)^{\nu_i}}, && (174) \end{aligned}$$

where in each case the denominator contains the concentration at standard state. Since solution standard states are taken to be of unit molality, the factor of $C_i^{\text{o,molal}} = 1\,\text{mol/liter}$ can be omitted in the last equation to give

$$\Delta G_{\text{rxn}} = \Delta G^{\text{o}}_{\text{rxn}} + RT \ln \frac{\prod_i^{\text{products}} \left[C_i^{\text{molal}}\right]^{\nu_i}}{\prod_i^{\text{reactants}} \left[C_i^{\text{molal}}\right]^{\nu_i}}, \qquad \text{all concentrations in molality} \tag{175}$$

so long as all concentrations are expressed in molal units. Similarly, since the density of pure water

$$\rho_{\text{H}_2\text{O}} \approx 1\,\text{kg/L} \tag{176}$$

is a very good approximation to the density of dilute aqueous solutions, the molal and molar units are equivalent for ideal aqueous solutions. Thus

$$\Delta G_{\text{rxn}} = \Delta G^{\text{o}}_{\text{rxn}} + RT \ln \frac{\prod_i^{\text{products}} \left[C_i^{\text{molar}}\right]^{\nu_i}}{\prod_i^{\text{reactants}} \left[C_i^{\text{molar}}\right]^{\nu_i}}, \qquad \begin{array}{c}\text{aqueous solution}\\ \text{all concentrations in molarity}\end{array} \tag{177}$$

15.5 Biochemical Standard State

In applying these formulas to biological systems, it's often useful to work with a slightly different set of standard states than are conventional in more general thermodynamic systems. The motivation for these changes is that biochemical reactions happen almost university in *aqueous solution*, typically with solute concentrations low enough that the mole fraction of the water is very nearly unity, and the density of the solution is very close to that of pure water. As a result, mol*ar* concentrations become equivalent to mol*al* concentrations, and, most importantly, *the mole fraction of water remains always very close to unity.* Because the water concentration never changes, it is both irrelevant and somewhat inconvenient to include the concentration of water in most calculations.

For example, in the hydrolysis of ATP,

$$\text{ATP} + \text{H}_2\text{O} \longrightarrow \text{ADP} + \text{P}_{\text{i}},$$

we should in principle calculate the ΔG for the reaction as

$$\Delta G = \Delta G^{\rm o} + RT \ln \frac{[\mathrm{ADP}][\mathrm{P_i}]}{[\mathrm{ATP}][\mathrm{H_2O}]}, \tag{178}$$

since water participates as a reactant. But since the reaction occurs in dilute solution, the consumption of H_2O by the process changes the mole fraction of water only negligibly. Thus the molar concentration of H_2O would be left constant at 55.5 moles/liter regardless of the concentration of the other species or the extent of the reaction. Although there's nothing incorrect about including this factor of 55.5 in the calculation, it's more convenient to simply redefine $\Delta G^{\rm o}$ to include it. In particular, if $[H_2O]$ is *always* 55.5 moles/liter (under both standard and non-standard conditions), we can rewrite the last expression as

$$\Delta G = \Delta G^{\rm o} - RT \ln(55.5) + RT \ln \frac{[\mathrm{ADP}][\mathrm{P_i}]}{[\mathrm{ATP}]}. \tag{179}$$

If we now *define*

$$\Delta G^{\rm o\prime} \equiv \Delta G^{\rm o} - RT \ln(55.5), \tag{180}$$

we can calculate ΔG as

$$\Delta G = \Delta G^{\rm o\prime} + RT \ln \frac{[\mathrm{ADP}][\mathrm{P_i}]}{[\mathrm{ATP}]}. \tag{181}$$

Thus, we can *drop* the water concentration in all subsequent calculations of ΔG and – so long as the water concentration is effectively 55.5 moles/liter under both standard *and* non-standard conditions, the relationship between ΔG and $\Delta G^{\rm o\prime}$ will remain unchanged.

This reformulation is the basis for the **biochemical standard state** convention:

Biochemical Standard States are characterized as follows

- All reactions are assumed to occur in dilute aqueous solution.
- The concentration of each solute is 1 molar.
- The pH of the solution is 7.0. Further, all species capable of protonation/deprotonation occur in the ratio of conjugate acid/conjugate base that occurs naturally at pH 7.
- The concentration of water is *by convention* set to 1 (i.e., ignored) in all ΔG calculations in which water participates as a reactant or product.

Thermodynamic quantities at biochemical standard state are denoted with a superscript o and a prime, e.g., $\Delta G^{\rm o\prime}$.

Note that "setting $[H_2O]$ to one" in this convention is **not** because water doesn't participate or because the concentration of water is *actually* one molar. It's simply a convention, enabled by the fact that the true concentration of water is "absorbed" into the $\Delta G^{\rm o\prime}$ value. This is why in general

$$\Delta G^{\rm o\prime} \neq \Delta G^{\rm o}. \tag{182}$$

16 Chemical Equilibrium

16.1 Equilibrium at Constant Temperature and Pressure

We noted earlier that the Gibbs free energy of a process can be used as an indicator of spontaneity at constant pressure and temperature. Processes for which ΔG is negative were seen to be spontaneous, while processes for which ΔG is positive were seen to be physically forbidden. It is in some respects even more interesting to examine the case $\Delta G = 0$. What can we say for such processes?

In this case, neither the forward process nor the reverse process is spontaneous. The only conclusion is that the process is *at equilibrium*: it can progress neither in the forward nor the reverse direction and must therefore remain unchanged with time. This principle is an exceptionally useful tool for characterizing the equilibrium states of chemical reactions and other everyday physical processes.

At constant temperature and pressure, a process is at equilibrium if and only if the ΔG associated with it is exactly zero.

16.2 Equilibrium in Chemical reactions

Beginning from our molar expression for reaction free energies in dilute solutions [Eq. (175)], we can apply this principle by setting $\Delta G = 0$ to obtain at equilibrium:

$$\Delta G_{\mathrm{rxn}} = \Delta G^{\mathrm{o}}_{\mathrm{rxn}} + RT \ln \frac{\prod_i^{\mathrm{products}} \left[C_i^{\mathrm{molal}}\right]^{\nu_i}}{\prod_i^{\mathrm{reactants}} \left[C_i^{\mathrm{molal}}\right]^{\nu_i}}, \qquad \text{all concentrations in molality} \tag{183}$$

This leads us to define the *equilibrium constant*

The **equilibrium constant** for a chemical reaction with stoichiometric coefficients ν_i is

$$K_{\mathrm{eq}} \equiv \frac{\prod_i^{\mathrm{products}} \left[C_i^{\mathrm{molal}}\right]^{\nu_i}}{\prod_i^{\mathrm{reactants}} \left[C_i^{\mathrm{molal}}\right]^{\nu_i}}, \qquad \text{all concentrations in molality} \tag{184}$$

From this value, the standard Gibbs free energy of reaction can be calculated as

$$\Delta G^{\mathrm{o}}_{\mathrm{rxn}} = -RT \ln K_{\mathrm{eq}}. \tag{185}$$

Conversely, given $\Delta G^{\mathrm{o}}_{\mathrm{rxn}}$, the equilibrium constant can be calculated as

$$K_{\mathrm{eq}} = e^{-\frac{\Delta G^{\mathrm{o}}_{\mathrm{rxn}}}{RT}}. \tag{186}$$

Equivalent results hold for gas-phase reactions if we replace all molar concentrations with pressures in units of bar.

Intuitively, the equlibrium constant tells us *how far* a reaction will progress. A $K_{\mathrm{eq}} \gg 1$ indicates that the equilibrium concentrations in the numerator (i.e., those of the reaction products) are much larger than those in the denominator (i.e., those of the reactants). Thus a very large K_{eq} indicates that a reaction strongly favors formation of products or will "go to completion". Conversely, if $K_{\mathrm{eq}} \ll 1$, then the equilibrium concentration of products must be much less than the equilibrium concentration of reactants; thus the reaction favors the reactants at the given temperature.

A qualitative statement of these results is given by the statement of Le Chatelier:

Le Chatelier's Principle states that chemical equilibrium will respond to a *change in external constraints* in such a way as to partially offset the external change.

So, for example, if we begin with a generic reaction

$$\mathrm{A} + \mathrm{B} \leftrightarrow \mathrm{C}$$

that is initially at equilibrium and then add a large quantity of reagent B to the mix, the reaction will progress *forwards* to create more of product C. Conversely, if we add a large quantity of C, the reaction will proceed *backwards* to produce A and B. If all chemicals are in the gaseous state, increasing the *total pressure* on the system will push the reaction toward the product C since there are fewer moles of gas on this side of the reaction and so the numerator of K_{eq} (indicating product pressure) is increased by a smaller factor than the denominator (the *product* of the pressures of both reactants). Finally, the expressions we'll derive in the next section show that if the reaction produces heat (i.e., is exothermic) then raising the temperature pushes the reaction *away from the product.* Conversely, if the reaction consumes heat (i.e., is endothermic), then raising the temperature pushes the equilibrium *toward* products. Thus, for the purpose of analyzing shifts in equilibrium with temperature, heat consumed in a reaction can loosely be thought of as a reactant, and heat produced can be thought of as a product.

16.3 Driving Forces in Chemical Reactions

This discussion naturally leads us to a more quantitative discussion of the question: What determines whether a reaction favors products or reactants at a given temperature? In other words, what are the *driving forces* behind chemical reactions?

There are several levels at which we can answer this question. Ultimately, of course, all real-world processes are driven by the Second Law requirement to increase entropy. Since the universe itself is (so far as we know) an isolated system, the Second Law implies that all real processes must increase the total entropy of the universe. As we've seen, in the specific context of a system held at constant temperature and pressure, this implies that ΔG for any spontaneous process must be negative. As we've just seen, for ideally dilute solutions and ideal gases, this principle can be further extended to show that ΔG^{o} predicts *how far* a reaction will proceed at a given temperature.

To get a more physically intuitive interpretation of this statement, let's split apart ΔG^{o} into its component parts:

$$\Delta G^{\mathrm{o}} = \Delta H^{\mathrm{o}} - T\Delta S^{\mathrm{o}}. \tag{187}$$

Thus we can identify two distinct contributions or "driving forces" that can push a process toward completion at constant temperature and pressure:

1. The standard enthalpy change ΔH^{o} characterizes how much heat is released by the process. A negative ΔH^{o} indicates that heat is released into the environment, increasing the environmental entropy and tending to make the process more spontaneous. Conversely, a positive ΔH^{o} indicates that heat is absorbed during the process, lowering the entropy of the environment and thus tending to oppose the spontaneity of the process.

2. The standard entropy change ΔS^{o} characterises how much the *entropy of the system* increases during the process. A positive ΔS^{o} indicates that (at the specified temperature), the process increases the entropy of the system, increasing the overall entropy of the universe and thus tending to make the process spontaneous. Conversely, a negative ΔS^{o} indicates that the entropy of the system decreases, lowering the overall entropy of the universe and tending to oppose spontaneity.

How far any specific process will go toward completion depends on the balancing of these two factors. A strongly endothermic reaction can be made spontaneous by a sufficiently large ΔS^{o}, and a strongly exothermic reaction can be prohibited by a sufficiently negative ΔS^{o}. The precise balance between the two is determined by the temperature. Note that, in general, both ΔH^{o} and ΔS^{o} may themselves depend on temperature, so the balance between the two is, in general, more complex than a simple linear temperature-dependence. Nonetheless, the sign of each term usually stays consistent over a significant range of temperatures, so that the intuitive picture of balancing the contribution of each term to the overall ΔG^{o} remains applicable.

16.4 Energy Storage in Chemical Bonds

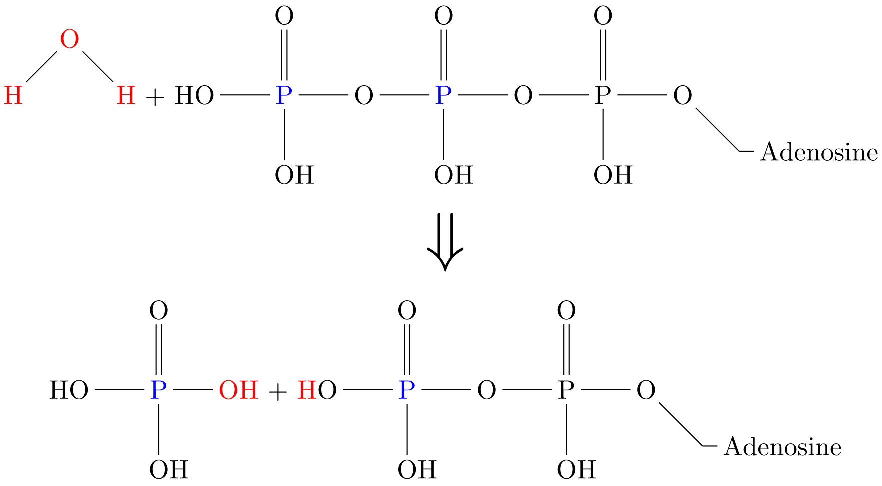
In this discussion of "driving forces" in chemistry, it's worth commenting on the mantra of "energy storage in chemical bonds" that is often repeated in discussions of biological metabolism. One common statement is that living systems store energy in "high-energy bonds" such as the phosphate-phosphate bonds of ATP. Similarly, it's often stated that in ATP hydrolysis or glucose metabolism, energy is released through "the breaking of chemical bonds". Informally, these statements have some utility, conveying correctly that the release of (Gibbs free) energy is what drives these reactions to completion.

If interpreted literally, however, both statements are a bit misleading and technically incorrect. The idea of a "high-energy bond" is itself something of a misnomer. By definition, a chemical "bond" is an interaction between two atoms that tends to hold the atoms close to each other. This can only happen if the molecular energy is *lower* when the two atoms form the bond than when they dissociate to separate atomic species. (The entropy change for formation of a bond is almost always negative since the system has far fewer accessible microstates under the constraint of inter-atomic binding.) For example, the reaction

$$2\mathrm{O}(\text{atomic}) \rightarrow \mathrm{O}_2(\text{gas}), \tag{188}$$

is spontaneous at standard conditions precisely because oxygen gas is *more stable* when the individual molecules associate into a bonded, diatomic configuration. This increase in molecular stability is reflected in the fact that the reaction is exothermic. In fact this is the only driving force for the reaction: the entropy change for the process is negative, since two moles of gaseous reactants combine to form a single mole of gaseous product. Thus it is actually the *formation* of the O=O chemical bond that leads to the release of energy.

Similar principles apply to the ATP hydrolysis reaction:



Here we *break* a single P —— P bond and an O —— H bond. However, we also *form* a new O —— H bond, along with another new O —— P bonds. It is actually the *formation* of the new bonds that releases energy; if we had stopped only with the breaking of the phosphate bond, the reaction would be strongly endothermic and very much disfavored.

The take-home message here is not that it's necessarily incorrect to use the phrase "high-energy bond" or to talk about energy being "stored in chemical bonds": it's just important to keep in mind what these terms really mean. All chemical bonds lower the energy to some extent relative to the "no bond" case; otherwise, we wouldn't call them bonds! It's just that some bonds lower the energy more than others. A "high-energy bond" is one that provides relatively *less* stabilization at the molecular level; if a chemical reaction replaces some of these "high-energy" bonds with "low-energy" bonds that provide *more* stabilization, the reaction will be exothermic and, if the entropy change is either positive or not too negative, exergonic. This is what drives many metabolic reactions, including ATP hydrolysis and glucose oxidation.

16.5 Variation of K_{eq} with Temperature: The Van't Hoff Equation

Eq. (187) above makes obvious that temperature plays a critical role in determining the balance between reactants and products in chemical reactions. Because of the explicit factor of T in the $-T\Delta S^{\text{o}}$ term, increased temperatures tend to enhance the significance of the ΔS^{o} value of a reaction. When T is close to zero, reactions are driven almost exclusively by the value of ΔH^{o}, since the $T\Delta S^{\text{o}}$ term is suppressed. Conversely, at high temperatures, the system entropy change tends to dominate spontaneity.

This relationship can be made more explicit by calculating the derivative of the equilibrium constant with respect to temperature. The derivative of Eq. (187) with respect to temperature can be calculated as

$$\frac{d\Delta G^{\text{o}}}{dT} = \frac{d\Delta H^{\text{o}}}{dT} - T\frac{d\Delta S^{\text{o}}}{dT} - \Delta S^{\text{o}}. \tag{189}$$

Now, in general ΔH^{o} and ΔS^{o} both depend on temperature, so that the first two terms on the right hand side of this equation are in general nonzero. However, in many cases these terms depend rather weakly on temperature. As a result, over a relatively small range of temperatures, one can usually make the approximation

$$\frac{d\Delta G^{\text{o}}}{dT} \approx -\Delta S^{\text{o}}. \tag{190}$$

A more experimentally useful version of the same result is obtained by examining the variation of $\ln K_{\text{eq}}$ with the *inverse* temperature $1/T$. Taking the natural log of Eq. (186) gives

$$\begin{aligned}\ln K_{\text{eq}} &= -\frac{\Delta G^{\text{o}}_{\text{rxn}}}{RT} & (191)\\ &= -\frac{\Delta H^{\text{o}}_{\text{rxn}}}{R}\frac{1}{T} + \frac{\Delta S^{\text{o}}_{\text{rxn}}}{R}. & (192)\end{aligned}$$

Differentiation of this expression with respect to $1/T$ gives the *Van't Hoff Equation*

Van't Hoff Equation:

$$\frac{d\ln K_{\text{eq}}}{d(1/T)} = -\frac{\Delta H^{\text{o}}_{\text{rxn}}}{R} \tag{193}$$

which is valid when $\Delta H^{\text{o}}_{\text{rxn}}$ and $\Delta S^{\text{o}}_{\text{rxn}}$ are both independent of temperature.

Now, if a series of values for K_{eq} is measured experimentally as a function of temperature (by actually monitoring the concentration of reactants and products at equilibrium), a very informative figure can be obtained by plotting $\ln K_{\text{eq}}$ versus the inverse temperature $1/T$. Such a figure is often referred to as a *Van't Hoff Plot*. According to Eq. (192), if $\Delta H^{\text{o}}_{\text{rxn}}$ and $\Delta S^{\text{o}}_{\text{rxn}}$ are temperature-independent, then $\ln K_{\text{eq}}$ will vary linearly with $1/T$ in a Van't Hoff plot. The slope of the line will be

$$\text{Slope} = -\frac{\Delta H^{\text{o}}_{\text{rxn}}}{R}, \qquad \text{Van't Hoff Plot} \tag{194}$$

and the intercept (where $1/T$ goes to zero, i.e., T tends to infinity) gives

$$\text{Intercept} = \frac{\Delta S^{\text{o}}_{\text{rxn}}}{R}. \qquad \text{Van't Hoff Plot} \tag{195}$$

Thus by simply measuring a series of equilibrium concentrations, the standard enthalpy and entropy changes for the reaction can be calculated. If $\ln K_{\text{eq}}$ is found *not* to vary linearly with $1/T$, this means that $\Delta H^{\text{o}}_{\text{rxn}}$ and/or $\Delta S^{\text{o}}_{\text{rxn}}$ themselves vary significantly over the temperature range considered.

17 General Criteria for Equilibrium

17.1 Differential Relations

Our application of the Gibbs free energy to characterize equilibrium processes has so far been limited to a rather specific context: Chemical reactions in the gas phase or in ideally dilute solutions that are maintained at constant temperature and pressure. But in fact, the usefulness of G as a metric for equilibrium extends far beyond these simple cases – even for systems that are *not* at constant temperature and pressure. In this lecture, we'll derive a general mathematical criterion for using the Gibbs free energy to characterize equilibrium states.

To see how this works, let's start with a simple example. Suppose we're interested in a chemical reaction

$$\mathrm{A} \longrightarrow \mathrm{B}$$

involving the conversion of a single chemical species A to some other species B. The total Gibbs free energy, $G_{\text{tot}}(T, P, n_A, n_B)$, for the system depends in general on the temperature, pressure, and number of moles n_A and n_B of the two species. Now, if the system is at equilibrium, the conversion of a small amount of species A to species B must *not* cause G_{tot} to decrease – otherwise the process would be spontaneous and would cause the system state to change. Likewise for the conversion of a small about of species B to species A. If the system is closed (i.e., cannot exchange matter with the environment), then the total number of moles is constant, so that if n_A changes by a small amount dn, then n_B must change by $-dn$. Thus we can express our equilibrium condition mathematically as

$$dG = G(T, P, n_A + dn, n_B - dn) - G(T, P, n_A, n_B) = 0. \tag{196}$$

In other words, at constant temperature and pressure, the differential change in G when some small quantity of species A is converted to species B must be identically zero.

Exactly this same argument can be generalized to characterize the dependence of G on *any* macroscopic coordinate in a system at equilibrium at constant temperature and pressure:

In a system at constant temperature and pressure, **the Gibbs free energy is minimized at equilibrium**. In other words, for an equilibrium system at constant temperature and pressure

$$dG = 0. \tag{197}$$

17.2 Arbitrary Systems

So far, this principle allows us only to characterize equilibrium for systems that are held at constant temperature and pressure. The remarkable thing is that the finding can be extended to characterize the equilibrium states even of systems that are *not* constrained to have constant temperature and pressure.

To see how this works, let's look at a concrete example. Suppose we're interested in a chemical reaction taking place in a closed, thermally isolated vessel at constant volume. Neither pressure nor temperature is fixed during this process. If the reaction proceeds in a direction that produces heat, the system temperature will increase. If the reaction proceeds in a direction that adds more gas to the system, the pressure will increase. If the reaction is irreversible, temperature and pressure may not even be well-defined during the course of the reaction!

Nonetheless, the system *will* eventually reach equilibrium. Eventually, the system will reach a point where the reaction ceases to progress and the temperature and pressure both stabilize at some set of values T_{final} and P_{final} that, at present, we can't predict *a priori*. What can we say about the system in this final, equilibrium state?

To answer this question, think about what would happen if we added one more step to our experiment. Suppose that, after the reaction reached equilibrium, we were to put the system in thermal contact with an outside reservoir at constant temperature T_{final} and inside of a piston with a constant external pressure of P_{final}. In other words, we now allow the system to interact with an isothermal and isobaric environment, with parameters that exactly match the final equilibrium state. What would happen? *Nothing!* Because the system already has the same temperature and pressure as the environment, the equilibrium state of the system will not be modified by the transition from complete isolation to isothermal/isobaric conditions.

Now think about what this tells us about the equilibrium state. In the isothermal/isobaric case, we know already that equilibrium is characterized by the fact that the dG must be exactly zero for any differential change that doesn't modify T or P. But, since G is a state function and since the equilibrium state is identical in the two cases, this statement must also be true in the original experiment. Thus, even for the isolated system, it remains true that, at equilibrium, G is minimized with respect to all coordinates *other than* temperature and pressure. Thus our previous result actually implies the much more general statement:

If the Gibbs free energy for *any system* can be expressed as a function $G(T, P, x_1, ..., x_N)$ of temperature, pressure, and the N independent variables $x_1, ..., x_N$, then at equilibrium

$$dG = 0 \tag{198}$$

for any infinitesimal displacement of the variables dx_i with P and T held fixed.

17.3 The Chemical Potential

Having arrived at this general differential criterion for equilibrium, it's worth putting this finding in the context of our earlier differential expression of the First Law of Thermodynamics. Earlier in the course, we saw that the energy of a single-component thermodynamic system could be expressed as

$$dU = dQ + dW = TdS - PdV, \tag{199}$$

so long as all processes are carried out reversibly, so that both the TdS and PdV terms remain well-defined. Now, for reversible processes, this expression is valid for *closed* thermodynamic systems; but what about a system that can exchange matter with the environment? In this case, the energy U depends also on the number of moles n_i of each chemical species in the system. If an infinitesimal amount dn_i of species i is added to the system, a small energy change occurs as well. Mathematically, this leads to the more general law

$$dU = TdS - PdV + \sum_i \mu_i dn_i, \tag{200}$$

where the quantity μ_i defines the variation of energy with displacements in n_i (with all other variables held constant) and is termed the *chemical potential*:

The **chemical potential** of a thermodynamic system is defined as

$$\mu_i \equiv \left(\frac{\partial U}{\partial n_i}\right)_{S,V,n_{j\neq i}}. \tag{201}$$

What exactly is this new creature, the "chemical potential"? From the definition just given, μ_i represents how much the energy increases when a small amount of species i is added to the system *at constant volume and entropy.* While mathematically precise, this definition isn't very physically informative – or very easy to measure. Experimentally, it's quite difficult to devise a process in which a system exchanges matter with its environment but, at the same time, maintains the same entropy. Fortunately, the situation becomes much clearer if our differential relation is rewritten in terms of Gibbs free energies.

Recall that, by definition $G \equiv U - TS + PV$. As a result:

$$\begin{aligned} dG &= dU - SdT - TdS + PdV + VdP && (202)\\ &= TdS - PdV + \sum_i \mu_i dn_i - SdT - TdS + PdV + VdP && (203)\\ &= VdP - SdT + \sum_i \mu_i dn_i. && (204)\end{aligned}$$

Mathematically, this expression defines the variation of Gibbs free energy with small changes in P, T, or dn_i. In complete analogy to Eqs. (200) and (201), μ_i in this case represents the variation of G with respect to n_i with all other variables held constant. Thus, in place of Eq. (201), we could just as well have defined

$$\mu_i \equiv \left(\frac{\partial G}{\partial n_i}\right)_{T,P,n_{j\neq i}}. \qquad (205)$$

In this case, μ_i represents the change in Gibbs free energy when a small amount of species i is added at constant temperature and pressure. Practically, this is a *much* easier process to control experimentally than conducting measurements at constant entropy!

17.4 The Chemical Potential as a Metric for Equilibrium

With these definitions, we can, for processes involving the transformation of matter from one form to another, give a general criterion for equilibrium in terms of the chemical potential. Return, for instance, to our earlier example of a closed system involving conversion of one chemical species A into another species B. If a small quantity dn_A of species A is created from species B, the differential free energy is

$$\begin{aligned} dG &= \mu_A dn_A + \mu_B dn_B && (206)\\ &= (\mu_A - \mu_B)\, dn && (207)\end{aligned}$$

since $dn_A = -dn_B$. At equilibrium, this variation must be identically zero, regardless of the actual value of dn. Thus our condition for equilibrium becomes

$$\mu_A = \mu_B. \qquad (208)$$

Thus, for any two species in a one-to-one interconversion equilibrium, *the chemical potentials must be identical.* This principle can be applied immediately to systems of *phase equilibrium* where a single species is transferred back and forth between two distinct physical phases:

> At equilibrium, the *chemical potential* for a given molecular species *must be the same in all phases.*

More generally, the same argument shows that

> A necessary and sufficient condition for the generic process
>
> $$\nu_1 \mathrm{X}_1 + \nu_2 \mathrm{X}_2 + ... + \nu_{N_\mathrm{r}} \mathrm{X}_{N_\mathrm{r}} \longrightarrow \nu_{N_\mathrm{r}+1} \mathrm{X}_{N_\mathrm{r}+1} + \nu_{N_\mathrm{r}+2} \mathrm{X}_{N_\mathrm{r}+2} + ... + \nu_{N_\mathrm{r}+N_\mathrm{p}} \mathrm{X}_{N_\mathrm{r}+N_\mathrm{p}}$$

to be at equilibrium is that

$$\sum_i^{\text{products}} \nu_i \mu_i - \sum_i^{\text{reactants}} \nu_i \mu_i = 0. \tag{209}$$

17.5 The Chemical Potential is the Molar Gibbs Free Energy

An even more intuitive interpretation is offered for μ_i by conducting a simple thought experiment. Suppose we were to build a thermodynamic system up from scratch, starting with nothing and adding one atom at a time up to some final composition of n_i moles of each molecular species i. At all times during the process, we keep all species in equilibrium with each other in fixed proportions and at constant temperature and pressure. At each infinitesimal step in this process, the Gibbs free energy increases by a quantity μ_i when the infinitesimal quantity dn_i is added to the system, i.e.,

$$dG = \sum_i \mu_i dn_i. \tag{210}$$

Thus at the end of the the total Gibbs free energy can be written as

$$G(n_1, ..., n_N) = \sum_i \int_0^{n_i} dn_i \mu_i. \tag{211}$$

Now, because the chemical composition, temperature, and pressure remain constant throughout the process μ_i is also constant for each species. The chemical potential can thus be pulled out of the integral expression, and we obtain simply

$$G(n_1, ..., n_N) = \sum_i \mu_i n_i. \tag{212}$$

The chemical potential μ_i can thus be understood rigorously as a *molar Gibbs free energy* of species i, i.e., the free energy contributed to the system by each mole of species i.

17.6 Differential Relations

While we're discussing the differential relation [Eq. (204)] for the Gibbs free energy, it's worth pointing out one other useful result that follows immediately from this equation. As a general rule of calculus, whenever we have a well-behaved function $f(x_1, ..., x_N)$ whose differential can be expressed in the form

$$df = g_1 dx_1 + g_2 dx_2 + ... + g_N dx_N, \tag{213}$$

the coefficients g_i *must* satisfy the relation

$$g_i = \left(\frac{\partial f}{\partial g_i}\right)_{x_{j \neq i}}. \tag{214}$$

Intuitively, this just means that, if g_i defines how f varies with small displacements in x_i, then it must *by definition* be the differential of f with respect to x_i.

This statement may seem obvious, but it leads to some otherwise non-obvious relations in thermodynamics. In particular, if we apply it to Eq. (204) we obtain the relations

$$V = \left(\frac{\partial G}{\partial P}\right)_{T, n_i} \tag{215}$$

$$S = -\left(\frac{\partial G}{\partial T}\right)_{P, n_i}. \tag{216}$$

Both of these will be very useful in the next section when we discuss phase equilibrium.

18 Colligative Properties

In the last section, we derived a general criterion for thermodynamic equilibrium, namely that the Gibbs free energy must be minimized with respect to the displacement of any coordinates at constant temperature and pressure. In the particular context of phase equilibria, we showed that a general criterion for equilibrium is that the chemical potential for a given species be the same in all phases. In this section, we'll use this reuslt to derive explicit laws for the *colligative properties* of dilute solutions, including vapor pressure suppression, boiling point elevation, and freezing point depression.

A **colligative property** of a solution is a property that depends only on the ratio of the *total mole fraction* of solute molecules and not on the specific chemical properties of each solute.

Colligative properties end up being very useful in real-life applications. For example, maple syrup is produced by boiling down sap from the sugar maple tree until the boiling point of the increasingly concentrated solution reaches a fixed reference temperature. In this section, we'll find a formula that allows us to calculate directly (if approximately) the concentration of sugar in maple syrup at the final boiling temperature.

18.1 Vapor Pressure

Recall from our earlier discussion of phase diagrams that a P/T phase diagram features *coexistence curves* where the various phases exist in equilibrium with each other. In particular, the liquid/vapor coexistence curve defines the set of temperatures and pressures at which a liquid will coexist in equilibrium with its vapor. If we view the sequence of pressures defined by this curve as a *function* of the temperature, we can read the coexistence curve as a plot of the solvent's *vapor pressure*:

The **vapor pressure** of a solvent is the pressure at which its liquid and vapor phases coexist in equilibrium at a given temperature.

Suppose, for example, that we were to take 1 liter of liquid water at 20 °C and place it in an evacuated cylinder (i.e., under vacuum) in a total volume of 2 liters. What would happen? The water will not, of course, stay all in the liquid state under a total vacuum. Instead, some of the water will evaporate until the pressure of the water vapor above the solution reaches equilibrium with the evaporation of water from the liquid phase. The pressure at which the water liquid and vapor reach equilibrium is, by definition, the vapor pressure.

Now, what would happen if instead of pure water the liquid contained a low concentration of some solute, say table sugar? A liquid/gas equilibrium will again be established, but, as it turns out, the vapor pressure above the solution would be diminished. Physically, this effect can be understood as a result of the fact that the sugar (which itself has a negligible vapor pressure) interferes with the evaporation of the water, preventing some of it from leaving the solution and, as a result, lowering the overall vapor pressure. Our goal now is to provide a *quantitative* description of this effect, using our recently established results on the free energies of dilute solutions.

Equilibrium is established when the chemical potential of the solvent is the same in both liquid and vapor phases. We can use this condition to derive an expression for liquid/vapor equilibrium. For simplicity, let's denote the number of sugar molecules as n_{sug} and the number of water molecules in solution as n_{liq}; similarly, we'll let n_{vap} be the number of water molecules in the gas phase.

$$\mu^*_{\mathrm{liq}} + RT \ln x_{\mathrm{liq}} = \mu^{\mathrm{o}}_{\mathrm{vap}} + RT \ln \frac{P}{P^{\mathrm{o}}} \tag{217}$$

Now, suppose P^* is the vapor pressure of the pure solvent. Then

$$\mu^*_{\mathrm{liq}} = \mu^{\mathrm{o}}_{\mathrm{vap}} + RT \ln \frac{P^*}{P^{\mathrm{o}}} \tag{218}$$

Combining the two equations gives

$$RT \ln \frac{P^*}{P^{\mathrm{o}}} + RT \ln x_{\mathrm{liq}} = RT \ln \frac{P}{P^{\mathrm{o}}} \tag{219}$$

or

$$\ln x_{\mathrm{liq}} = \ln \frac{P}{P^*}. \tag{220}$$

Exponentiating both sides then gives

$$P = x_{\mathrm{liq}} P^*. \tag{221}$$

Thus *the vapor pressure of the water is decreased in proportion to its mole fraction.* Note that this effect depends only on the total mole fraction $1 - x_{\mathrm{liq}}$ of the solute, not on its particular chemical identity. The same result would be obtained if we had multiple different solute species; only the total solute concentration matters. This is the defining feature of a *colligative property.*

18.2 Boiling Point Elevation

A similar effect is observed in monitoring the boiling point of our solution. At a fixed pressure P, the boiling temperature T_b is defined by the condition that the vapor pressure is equal to P. Since the vapor pressure is suppressed by the presence of the solute, the boiling point will, in general, be higher than for the pure solution. To describe this elevation quantitatively, we'll need to consider explicitly the dependence of the chemical potential for the pure solvent (liquid and vapor) as a function of temperature.

Our condition for equilibrium is that

$$\mu^*_{\mathrm{liq}}(T_b) + RT_b \ln x_{\mathrm{liq}} = \mu^*_{\mathrm{vap}}(T_b). \tag{222}$$

Since the solution is dilute, the boiling point T_b will not differ very much from the boiling temperature T_b^* for the pure liquid. Thus, to first order in the difference we have

$$\mu^*_{\mathrm{liq}}(T_b) \approx \mu^*_{\mathrm{liq}}(T_b^*) + \frac{d\mu^*_{\mathrm{liq}}}{dT}\left(T_b - T_b^*\right). \tag{223}$$

Now, according to Eq. (212), μ^*_{liq} is just the molar Gibbs free energy of the liquid; its differential with temperature is therefore given by Eq. (216) as

$$\frac{d\mu_{\mathrm{liq}}}{dT} = -s_{\mathrm{liq}}, \tag{224}$$

where s_{liq} is the molar entropy of the liquid solvent. A similar result holds for the variation of $\mu^{\text{o}}_{\text{vap}}$ with temperature, so that we obtain

$$\mu^*_{\text{liq}}(T_b) \approx \mu^*_{\text{liq}}(T^*_b) - s_{\text{liq}}\,(T_b - T^*_b) \tag{225}$$

$$\mu_{\text{vap}}(T_b) \approx \mu_{\text{vap}}(T^*_b) - s_{\text{vap}}\,(T_b - T^*_b)\,. \tag{226}$$

Now, at the boiling point T^*_b of the pure solvent, our equilibrium condition gives us the relation

$$\mu^*_{\text{liq}}(T^*_b) = \mu_{\text{vap}}(T^*_b). \tag{227}$$

Introducing approximations [Eqs. (225) and (226)] into our equilibrium condition Eq. (222) and then using Eq. (227), we obtain

$$\mu^*_{\text{liq}}(T^*_b) - s_{\text{liq}}\,(T_b - T^*_b) + RT_b \ln x_{\text{liq}} = \mu^*_{\text{vap}}(T^*_b) - s_{\text{vap}}\,(T_b - T^*_b) \tag{228}$$

$$\to RT_b \ln x_{\text{liq}} = [s_{\text{liq}} - s_{\text{vap}}]\,(T_b - T^*_b) \tag{229}$$

or, after some rearrangement,

$$\ln x_{\text{liq}} = -\frac{\Delta s}{RT_b}\,(T_b - T^*_b) \tag{230}$$

where

$$\Delta s = s_{\text{vap}} - s_{\text{liq}} \tag{231}$$

is the molar entropy of vaporization. Since the solution is dilute, the mole fraction of the solvent x_{liq} is very close to one; thus the Taylor series expansion

$$\ln 1 - x \approx -x \tag{232}$$

can be used to express the result as[7]

$$-\Delta s\,(T_b - T^*_b) \approx RT_b \ln\left(1 - x_{\text{solute}}\right) \approx -RT^*_b x_{\text{solute}} \tag{233}$$

or

$$T_b = T^*_b + x_{\text{solute}}\frac{RT^*_b}{\Delta s}. \tag{234}$$

Since the molar entropy of vaporization is related to the molar heat of vaporization by

$$Q_{\text{vap}} = \frac{\Delta s}{T^*_{\text{b}}}, \tag{235}$$

this is commonly expressed as

$$T_b = T^*_b + \frac{R(T^*_b)^2}{Q_{\text{vap}}}x_{\text{solute}}. \tag{236}$$

Note here that – as with all colligative properties – the effect depends only on the *total concentration* of dissolved species, regardless of their chemical identities. Thus if we wanted to calculate the boiling point of a solution of NaCL (which dissociates in solution), the solute mole fraction x_{solute} must refer to the *total* concentration of ions, i.e., *twice* the concentration of Na or Cl individually.

---

[7] Note that we've replaced T_b with T^*_b on the right hand side; since the right hand side is already proportional to a small quantity, the difference between T_b and T^*_b is negligible.

18.3 Freezing Point Depression

A similar effect occurs when we examine the freezing point of a dilute solution. We won't go through the derivation since it's virtually identical to the calculation above for boiling point elevation. The final result is the formula

$$T_f = T_f^* - \frac{R\left(T_f^*\right)^2}{Q_{\text{fus}}} x_{\text{solute}}, \tag{237}$$

where T_f^* is the freezing point of the pure solvent, and Q_{fus} is the heat of fusion of the pure solvent. Note that Q_{vap} and Q_{fus} are both positive for most solvents; thus the last term in this expression has an overall negative sign, so that the freezing point is *lowered* in the presence of a solute.

18.4 Osmotic Pressure

Finally, let's consider the problem of *osmotic pressure* across a semipermeable membrane. Suppose we have a solvent contained in two chambers; one contains pure solvent and the other a dilute solution of some solute molecule dissolved in the same solvent. The two chambers are separated by a semipermeable membrane that allows passage of solvent molecules but not the solute. Experimentally it is observed that the presence of the solute induces a *pressure difference* between the two chambers at equilibrium. This difference is the *osmotic pressure.*

Osmotic pressure is the difference in pressure across a semipermeable membrane induced by the presence of one or more solutes restricted to one side of the membrane.

At equilibrium, the chemical potential of the pure solvent $\mu^*(T, P)$ must be equal to that of the higher-pressure solvent that contains the solute $\mu^*(T, P + \delta P) + RT \ln x_{\text{solv}}$. Here δP is the osmotic pressure. Our condition for equilibrium is thus

$$0 = \mu^*(T, P + \delta P) + RT \ln x_{\text{solv}} - \mu^*(T, P). \tag{238}$$

Expanding the dependence of μ^* on P in a Taylor series, this condition becomes

$$\left(\frac{\partial \mu^*}{\partial P}\right)_T \delta P = -RT \ln x_{\text{solv}}. \tag{239}$$

But from Eq. (215), this partial derivative is just the molar volume v_{solv} of the pure solvent! Expanding the ln term via Eq. (232), we obtain

$$v_{\text{sol}} \delta P = RT x_{\text{solute}} \tag{240}$$

or simply

$$V \delta P = n_{\text{solute}} RT. \tag{241}$$

Thus, perhaps surprisingly, the osmotic pressure due to the solute follows simply the ideal gas law!

19 The Gibbs Free Energy as Available Work

Before we leave our discussion of the Gibbs free energy, it's worth recognizing that it can be understood in yet one other conceptual framework: as the available non-PV work stored in a system.

Suppose we have a system whose thermodynamic energy U is a function of N independent coordinates $x_1, ..., x_N$, in addition to S and V. The coordinates $x_1, ..., x_N$ could be the quantities n_i of various chemical species, but they could also be any other macroscopic work coordinate. For example, if the system under consideration is a muscle cell, one of the coordinates x_i might be the length of the cell, the variable associated with the performance of physical work by the corresponding muscle. The differential relation for energy is then

$$dU = TdS - PdV + \sum_i \left(\frac{\partial U}{\partial x_i}\right)_{S,V,j\neq i} dx_i. \tag{242}$$

Now, for any reversible change in the state of the system, the differential amount of non-PV work dW_{npv} is simply the last N terms in this expression, i.e.,

$$dW_{\text{npv}} = \sum_i \left(\frac{\partial U}{\partial x_i}\right)_{S,V,j\neq i} dx_i. \tag{243}$$

A similar relation holds for the Gibbs free energy. The total differential is

$$dG = VdP - SdT + \sum_i \left(\frac{\partial G}{\partial x_i}\right)_{P,T,j\neq i} dx_i. \tag{244}$$

Now, since the x_i are independent variables from P and T, the derivatives $\left(\frac{\partial g}{\partial x_i}\right)_{P,T,j\neq i}$ in this expression must represent exactly the same quantities as the derivatives $\left(\frac{\partial U}{\partial x_i}\right)_{P,T,j\neq i}$ in the differential for energy. Thus the total differential of non-PV work for a *reversible process at constant pressure and temperature* is similarly just

$$dW_{\text{npv}} = \sum_i \left(\frac{\partial G}{\partial x_i}\right)_{T,P,j\neq i} dx_i. \tag{245}$$

Integrating this equation across some macroscopic displacement of the coordinates $x_1, ..., x_N$, we obtain the following observation:

> The Gibbs free energy change associated with any process at constant temperature and pressure is equal to **the total amount of non-PV work** that could be obtained **if the process were carried out reversibly.**

In particular, if the coordinates x_i are mole numbers, the Gibbs free energy change associated with a process quantifies the total amount of chemical work that could be achieved if the process were carried out reversibly. Thus, for example, the ΔG for the oxidation of glucose to carbon dioxide and water defines the total amount of chemical work (e.g., formation of ATP from ADP or reduction of NADP^+ to NADPH) that can be generated by its metabolism.

As we've stated many times, of course, all process in the real world are irreversible to some extent. This means that the actual non-PV work generated from a constant P/constant T process is never actually equal to the Gibbs free energy change. Rather ΔG sets an *upper bound* on the total work that *could* be performed.

20 Statistical Mechanics: Entropy and Information

20.1 Motivation and Background

The last section closes our formal introduction of thermodynamics in this course. This lecture opens our foray into *statistical mechanics*:

> **Statistical mechanics** is the field of physical chemistry that uses explicit averages over microscopic physical models to derive thermodynamic relations.

You might wonder why we're bothering with statistical mechanics, given that its ultimate goal is to derive the *thermodynamic* relations that (to a considerable extent) we've already covered! In this course, there are two primary motivations, one academic and the other practical:

- Statistical mechanics allows for a *deeper fundamental understanding* of how the world works by explaining *why* macroscopic systems behave the way they do. At a more practical level:
- Statistical mechanics allows us to *devise and test physical models* for how a system works at the *microscopic* level based on only *macroscopic* measurements of physical quantities.

The latter possibility – the ability to devise and test microscopic models against macroscopic measurements – is what has made statistical mechanics an essential tool in molecular biophysics, including problems ranging from enzyme function (How many substrates does an enzyme bind? How many enzymes function in a given pathway?) to protein structure (Which structural states are most stable? What is the likelihood of finding one of the less stable structures?).

In fact, it was statistical mechanics that originally convinced the scientific world of the most basic of chemistry models: the existence of atoms. Prior to the statistical arguments formulated by scientists such as James Clerk Maxwell, Ludwig Boltzmann, and Josiah Willard Gibbs, much of the scientific world was convinced that thinking about "atoms" was something of a foolish game. It was only when statistical mechanics was proven capable of *predicting macroscopic physical laws with quantitative accuracy* – but only if one assumed the existence of atoms! – that the scientific world started really to take the idea seriously.

20.2 Technical Approach

In its early days, statistical mechanics was an exceptionally difficult subject, particularly since its pioneers were really trying to solve two problems at once: to develop the statistical tools to describe the world at the microscopic level *and* to come up with microscopic models that (when subjected to statistical analysis) gave rise to the right macroscopic physics. Our study will be made substantially easier by more than a century of work on the subject, offering both a wealth of accurately parameterized microscopic models and a variety of different statistical approaches to "get through the math".

The approach that we'll adopt toward statistical mechanics is often termed the *information theory* or *maximum entropy* formulation. In this method, information about the *microscopic* properties of a system are "guessed" based on its observed *macroscopic* properties and the assumption that – in a very precisely defined, technical sense – we can assume *no* information about the microscopic state of the system *other than* the constraints offered by its macroscopic properties. Specifically, the information theory approach uses these assumptions to assign *probabilities* to every possible *microstate* the system can inhabit:

> A **microstate** μ is a specific configuration of a physical system in which the positions and momenta of all particles are precisely defined.

Conversely, a **macrostate** M is the thermodynamic state characterized by a given set of macroscopic properties (e.g., temperature, pressure, volume, etc.).

Using these microstate probabilities p_μ, *macroscopic* properties can be calculated as average values. Note that, unless explicitly noted, the symbol μ will throughout the rest of this course be used to enumerate system microstates **not** chemical potentials.

Suppose, for example, that we know a gas cylinder contains exactly N ideal gas particles and that the macroscopic energy of the gas is *exactly* 1 J; these macroscopic properties define the *macrostate.* Statistical mechanics then asks: Given these constraints, what are the principles governing the *microscopic* distribution of gas molecules throughout the box? In particular, what is the probability for any given system microstate, i.e., a particular configuration of gas particles in the box? The information theoretic approach answers immediately: All microstates involving N gas particles and an energy of exactly 1 J should have equal probability, since we have no information that could be used to assign one state as more likely than any others. Any microstate with an energy *other than* 1 J should have zero probability since it contradicts our macroscopic constraints. And this argument turns out to work quite well: If these assigned probabilities are used to calculate macroscopic quantities like temperature and pressure, this *statistical model* leads directly to the ideal gas law.

20.3 Information Theory

In many cases, however, it isn't nearly as obvious how to "assume no information" other than that afforded by the macroscopic constraints. To see how to do this rigorously, we need some background in *information theory*:

Information Theory is the field of mathematics that deals with the quantification and analysis of *information* or, its opposite, *uncertainty.*

The aspect of information theory with which we'll be concerned deals with the amount of uncertainty embedded in a set of *discrete probabilities*:

A **discrete probability** p_μ is a numerical assignment of *how likely* a given situation or event μ is to occur. Discrete probabilities must always be greater than zero (the outcome is impossible) and less than one (the outcome is guaranteed). For example, in a coin flip, there are two possible outcomes; if the coin is "fair" (i.e., equally balanced on each side), a reasonable discrete probability for the outcome "heads" would be $p_{\text{heads}} = \frac{1}{2}$.

A **discrete probability distribution** is a complete set $\{p_\mu\}$ of discrete probabilities, i.e., an assignment of how likely *each* of a range of possible situations or events is to occur. If there are N possible situations or events, then

$$\sum_{\mu=1}^{N} p_\mu = 1. \tag{246}$$

For example, in a coin flip, a reasonable set of possibilities would be $p_{\text{heads}} = p_{\text{tails}} = \frac{1}{2}$ so that $p_{\text{heads}} + p_{\text{tails}} = 1$.

In statistical mechanics, we're concerned about assigning probabilities p_i to each of the possible *microstates* that a system might inhabit under the constraints imposed by a given set of macroscopic constraints.

The foundational development in information theory was the introduction[8] by Claude Shannon of a rigorous mathematical metric for quantifying the *amount of uncertainty* embedded in a set

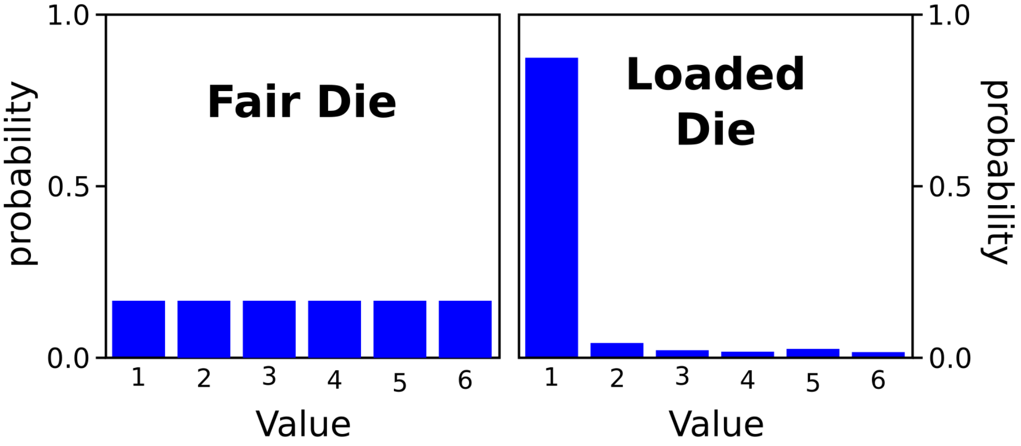
---

[8] See [C. E. Shannon, "A mathematical theory of communication," in *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, July 1948] online at `https://ieeexplore.ieee.org/document/6773024`.

of discrete probabilities. Intuitively, it should be reasonable that certain probability distributions provide more information than others. For example, the bar charts below compare two different sets of probabilities for a six-sided die (singular of “dice”, as used in board games and gambling).

The left chart shows the probabilities for a *fair die*; in this case the probabilities of all six possible outcomes of a roll have the same weight. This probability distribution gives us very little information about the possible result of a roll of the die; its *uncertainty* is thus high. All we know is that the die will show one of the six possible values. On any given roll, we’re equally likely to obtain a 6 as a 1, a 2, or any other possible value.

The chart on the right shows the probability distribution for a *loaded die*. Here (presumably because of a weight added to one side) we’re very much more likely to obtain a 1 than any other number. This probability distribution contains quite a bit of information about the outcome of a roll: we’re almost certain to roll a 1. The *uncertainty* of the distribution is thus very low. If this die were to be used in a game, we’d certainly want to know this information before we play!



Shannon’s seminal contribution – and the starting point for modern information theory – was to introduce a *quantitative* way to measure the amount of uncertainty associated with a probability distribution. It’s not immediately obvious that a well-defined measure of the “amount of uncertainty” even exists. But Shannon reasoned that *if it exists* the uncertainty measure $s(\{p_i\})$ should satisfy three postulates:

Shannon’s Uncertainty Postulates:

- The uncertainty $s(\{p_\mu\})$ should be a continuous function of each probability p_μ. This just means that there should be no “sudden jumps”: a small change in the probabilities should give rise to a correspondingly small change in uncertainty.
- When each of N possible outcomes are equally likely, the uncertainty $s(\{p_\mu\})$ must increase monotonically with increasing N. Intuitively, this just says that the *more possible outcomes* there are, the *less we know* about what’s going to happen!
- If a choice between two outcomes can be broken down into successive choices, the total uncertainty should be just the total uncertainty of *making* the choice plus the weighted sum of the individual uncertainties *within* each choice.

The last of these assumptions may be a little confusing, so let’s work through an explicit example. Suppose that, instead of getting information about the probabilities of each die roll individually we were given two separate pieces of information:

1. The relative probabilities p_{even} and p_{odd} for rolling even and odd numbers, and also
2. The relative probabilities for rolling a 2, 4, or 6 *given* that the outcome is even, along with the relative probabilities for rolling a 1, 3, or 5 *given* that the outcome is odd.

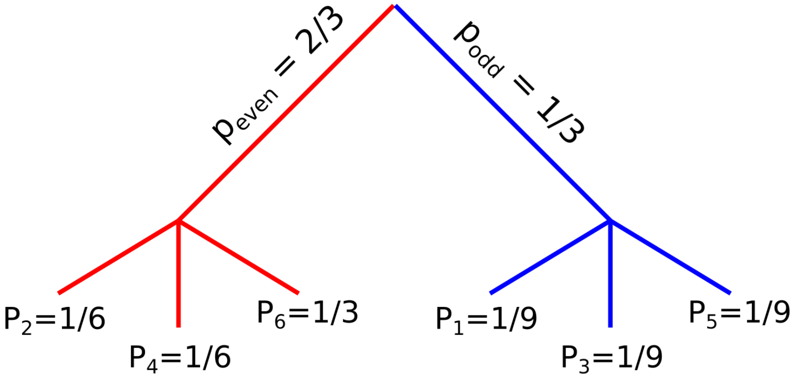
Shannon's assumption is that we can break down the *total* uncertainty of the distribution into separate contributions from these two statements:

$$s(\{p_1, p_2, p_3, p_4, p_5, p_6\}) = s(\{p_{\text{even}}, p_{\text{odd}}\}) + p_{\text{even}} \cdot s(\{p_{2|\text{even}}, p_{4|\text{even}}, p_{6|\text{even}}\}) \tag{247}$$

$$+ p_{\text{odd}} \cdot s(\{p_{1|\text{odd}}, p_{3|\text{odd}}, p_{5|\text{odd}}\}), \tag{248}$$

where $p_{i|\text{even}}$ is the probability for roll i *given* that the outcome is even, and $p_{i|\text{odd}}$ is the probability for roll i *given* that the outcome is odd.

A concrete example is illustrated in the figure below. Here probability of rolling an even number is $p_{\text{even}} = \frac{2}{3}$, while the probability of rolling an odd number is only $p_{\text{odd}} = \frac{1}{3}$. Further *within the possible odd outcomes*, all values (1, 3, or 5) are equally likely, while *within the possible even outcomes*, we're twice as likely to roll a 6 as to roll either a 2 or 4.



In this case, Shannon's assumption implies that

$$s\left(\left\{\frac{1}{9}, \frac{1}{6}, \frac{1}{9}, \frac{1}{6}, \frac{1}{9}, \frac{1}{3}\right\}\right) = s\left(\left\{\frac{1}{3}, \frac{2}{3}\right\}\right) + \frac{2}{3} \cdot s\left(\left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right\}\right) + \frac{1}{3} \cdot s\left(\left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}\right). \tag{249}$$

20.4 Shannon's Information Entropy

Shannon's critical insight was that (up to an overall scaling factor) *there is only one mathematical function that can satisfy all three of these postulates.* That function is

The Shannon Information Entropy: Shannon's information entropy is a measure of the uncertainty associated with a given probability distribution $\{p_\mu\}$ and is defined as:

$$s(\{p_\mu\}) = -\sum_\mu p_\mu \ln p_\mu. \tag{250}$$

To understand what this new quantity represents, let's look at two limiting cases. Given a set of N possible outcomes to an event:

- If one outcome is certain to happen and all others are impossible, then the information entropy is exactly zero.

- If all outcomes are equally likely, i.e., $p_\mu = \frac{1}{N}$ for each μ, then

$$s(\{p_\mu\}) = -\sum_{\mu=1}^{N} \frac{1}{N} \ln \frac{1}{N} \tag{251}$$

$$= -N \cdot \frac{1}{N} \ln \frac{1}{N} \tag{252}$$

$$= \ln N. \tag{253}$$

The first case tells us correctly that if the outcome is certain, there is no uncertainty! The second says that, with equally probable outcomes, the information entropy is just the natural log of the number of possibilities.

This last formula should look rather familiar. In fact, apart from multiplication by the prefactor k_B, this is exactly Boltzmann's formula [Eq. (91)] for the thermodynamic entropy! This connection seems at first rather surprising since Boltzmann's formula was developed through a heroic effort to directly calculate the thermodynamic properties of physical systems, while Shannon's formula was obtained through an entirely abstract process of reasoning about probability distributions. In the next section, we'll begin to explore the connection between these two results.

21 Information Theory and Statistical Mechanics

21.1 The Maximum Entropy Approach

In the last section, we introduced the concept of *information entropy* $s(\{p_\mu\})$ of a probability distribution and stated that we would use this information-theoretic tool to "infer" the correct statistical description of thermodynamic systems. To do this, we'll follow the *maximum entropy* method introduced by E. T. Jaynes:

The Maximum Entropy Method: To assign microstate probabilities to a thermodynamic system, proceed as follows:

1. Identify the external constraints on the system. For example, is the system constrained to have a fixed volume? What about pressure or temperature? Is the total energy fixed (as for a totally isolated system)?

2. Identify what possible *microstates* are consistent with the macroscopic constraints just defined. For example, for an isolated system at fixed energy U, we know that only those microstates μ that have energy U are allowed.

3. Identify the set of probabilities p_μ that, when assigned to the microstates μ, *maximizes the information entropy* of the distribution, *subject to the macroscopic constraints.*

The resulting probabilities p_μ can then be used to calculate any unknown thermodynamic properties by averaging over the probability-weighted microstates.

Before we go into the technical details of how this works, let's back up and talk about the big picture. What are we trying to accomplish here, and why does our approach make sense?

21.2 Why are we doing this?

Ultimately, what we're trying to accomplish is to find a way to calculate *macroscopic* physical properties from *microscopic* models. This approach is immensely useful in both physical and biological applications since it allows us to *test mechanistic molecular models using only macroscopic measurements.* For example, suppose we want to know how many binding sites a particular enzyme has. It would be extremely inconvenient if the only way to answer this question were to crystallize the protein and analyze its structure at the microscopic level. A much easier approach would be to *build a statistical model* for how the thermodynamic and kinetic properties of the enzyme would differ between a system with only one binding site versus a system with two or more sites. Then we could simply perform the appropriate macroscopic binding measurements in solution and check which model (single-site or multi-site) more accurately matched the data. No crystal structure necessary!

Alternatively, if we're in the business of trying to design *new* molecualr systems (e.g., photovoltaic cells, bioactive polymers, biomimetic scaffolds, or enzyme-inhibiting drugs), statistical mechanics gives us a way to *predict* which designs will most efficiently achieve our objectives. Although (given the complexity of biological macromolecules) theoretical models usually can't predict macroscopic properties *exactly*, they do offer a great deal of *qualitative* insight into what designs are *likely* to work. For example, statistical principles are routinely used to predict Gibbs free energy change of ligand binding in computational drug-discovery studies. Although such calculations carry substantial uncertainty, they *can* rapidly screen much larger numbers of possible drug molecule designs than could be tested experimentally, selecting only the most likely candidates for experimental testing.

21.3 Why does maximum entropy make sense?

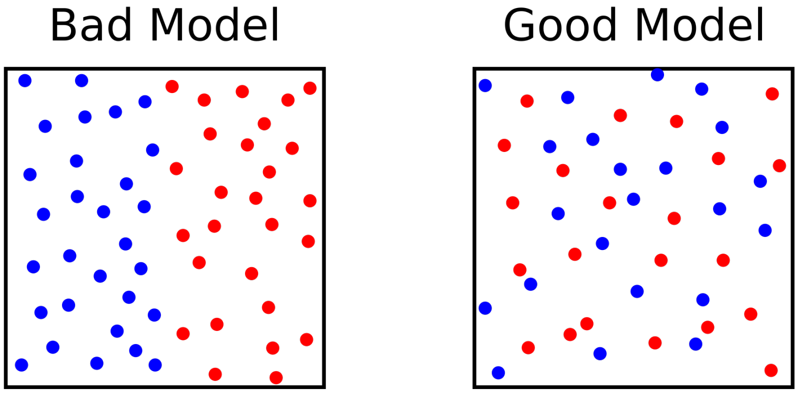
These examples hopefully make clear why it's useful to be able to build (for a given microscopic model) a statistical description that lets us calculate macroscopic properties from microscopic probabilities p_μ. What may be less clear at this point is why it makes sense to pick the *particular* probability distribution that maximizes the information entropy $s(\{p_\mu\})$.

Although there are many different arguments that can be used to justify this choice, the simplest one is simply that *it works*. In a few lines of equations, the maximum entropy approach can be used to obtain mathematical results that would take a rigorous first-principles approach many volumes of work to derive explicitly. And, when applied correctly, the method agrees with other "first principles" approaches in all known cases.

More fundamentally, the reason this works goes back to our observation in the first weeks of the course that *the thermodynamic state of most systems can be described accurately in terms of only a very small number of macroscopic coordinates.* In fact, almost the first step in our treatment of thermodynamics was to distinguish between *work* and *heat*. Work, you may recall, is any change to the system energy resulting from displacement of one of only a handful of macroscopic coordinates – volume, electrical charge, chemical concentration, etc. Heat consists of any change to the system energy that is *not* describable as macroscopic work. Thus, the vast majority of the $\sim 10^{23}$ microscopic coordinates of a macrosocpic system are lumped together into a single coordinate (the entropy) which is related to the flow of heat, while only a few select coordinates are observable as macroscopic work.

Jaynes' key insight was to realize that a very similar distinction can be applied to assign microscopic probabilities: Since only a few macroscopic coordinates suffice to determine thermodynamic state, we'll always obtain correct results *if we assume no information* about those coordinates that do *not* define the macroscopic state. If the true microstate probabilities *did* contain information about some other macroscopic coordinate (information contrary to that already implied by the existing constraints), then this would necessarily constitute a new macroscopic constraint and should be incorporated into our definition of the thermodynamic state. Thus we'll never run into contradictions as long as we *assume as little information as possible* about the microscopic state of the system.

Suppose, for example, that we wanted to build a probability distribution for a system that consists of a mixture of Ar and Xe gas in a sealed box. We *could* construct a set of microstate probabilities in which all the Xe was always on the left side of the box and all the Ar was always on the right side (see figure below). But this set of probabilities implies a macroscopic separation of our gases that isn't really present in our system; in fact, thermodynamically, this would be a completely different system. A much more reasonable model is one in which the two gases are evenly distributed across both sides of the box, i.e., one that assumes no information about the relative locations of Ar and Xe.



In short, the maximum entropy approach is based on the fact that, if probabilities are assigned in a way that both *respects the macroscopic constraints* and *carries no information about unconstrained coordinates*, our predictions will *always be correct.* This doesn't necessarily mean that the system actually inhabits all of the states in our model; in fact, it would take a fantastically long time even for relatively simple systems to sample all of the microstates consistent with their respective thermodynamic states. It does mean, however, that any deviations between which microstates are *actually inhabited* by the system and which microstates our model *predicts will be inhabited* will *not lead to errors* in calculating macroscopic quantities. And, ultimately, this is what thermodynamics is all about: describing systems in *coarse-grained* terms that allow us to make macroscopic predictions without worrying about microscopic details.

21.4 Thermodynamic Ensembles

The maximum entropy approach can be applied to systems under any set of thermodynamic constraints – whether based on volume, energy, pressure, or anything else. The statistical model describing any such system is referred to as a *thermodynamic ensemble*:

A **thermodynamic ensemble** is a hypothetical collection of physical microstates μ, along with assigned probabilities p_μ, used to model a real macroscopic system subject to a fixed set of thermodynamic constraints. If the ensemble is constructed correctly, then all macroscopic properties X for the real system can be calculated as *ensemble averages*

$$X = \langle x \rangle = \sum_\mu p_\mu x(\mu) \tag{254}$$

of the corresponding microscopic functions $x(\mu)$.

We performed a very simple ensemble average at the very beginning of this course, when we calculated the average energy of a monatomic ideal gas in terms of the macroscopic pressure and volume. In that case, we didn't have the specific probabilities over each microstate; we simply showed that the average energy could be related to the quantity PV.

In the coming lectures, we'll use the maximum entropy approach to construct specific microstate probabilities for each possible microstate in an ideal gas, along with several other systems. Each of these calculations will be performed in one of two thermodynamic ensembles:

The **Microcanonical Ensemble** is the thermodynamic ensemble corresponding to a system held at constant number of particles N, volume V, and energy E. It is sometimes referred to as the NVE ensemble.

The **Canonical Ensemble** is the thermodynamic ensemble corresponding to a system held at constant number of particles N, volume V, and temperature T. It is sometimes referred to as the NVT ensemble.

There exist many other thermodynamic ensembles (e.g., the "Gibbs Ensemble" corresponding to a system held at constant pressure or the "Grand Canonical Ensemble" describing open systems), but these two will be sufficient to derive the results necessary for our study. In general, extending the theory to other ensembles is straightforward and follows a similar pattern as used in our study of the microcanonical and canonical systems.

22 The Microcanonical Ensemble

22.1 Introduction and Definition

The microcanonical ensemble is conceptually the simplest of all thermodynamic ensembles, though that doesn't always make it either the most realistic or the easiest to work with. Our task in microcanonical systems is to assign probabilities to all possible microstates μ based on the constraints that the number of particles, the total volume, and the total energy of each microstate must remain fixed at some predefined set of values, N, V, and E.

It's pretty clear what states are *possible* in such a system: any microstate μ is possible as long as it contains the right number of particles in the correct volume and its total energy ε_μ is exactly equal to E. What about the microstate probabilities?

Intuitively, the answer may be clear already: If we're to assume no extra information than the state constraints, then there can be no reason to choose one microstate as more probable than another. But it's instructive to see that exactly the same result is determined by our information-theoretic approach. The maximum entropy method says that the correct set of probabilities is the one that maximizes the information entropy of Eq. (250) subject to the relevant constraints. In the microcanonical system, we can apply the necessary thermodynamic constraints just by restricting the sum over microstates μ to extend only over those states with the correct values of N, V, and E. But there is one other constraint that must be applied during the optimization so as to make sure the probabilities are statistically meaningful: the sum of all probabilities must equal 1. To implement this constraint, we'll need to grab another tool from our calculus mathematical toolbox: the method of Lagrange Multipliers.

22.2 Lagrange Multipliers

You probably learned about Lagrange multipliers in your primary calculus sequence, but it's okay if you don't remember them. The intuition behind them is a bit abstract, and they're often covered only very briefly. Unfortunately, we don't have time to give a thorough treatment here either, but we'll at least describe the basic mathematical method and give a qualitative explanation for why it works.

The method of **Lagrange Multipliers** is a method for *maximizing* or *minimizing* a given function subject to constraints. The mathematical basis of the method is the observation that: *the extremal points (maxima or minima) of a function $f(\boldsymbol{x})$, subject to a series of constraints $c_1(\boldsymbol{x}) = 0$, $c_2(\boldsymbol{x}) = 0$, ..., $c_K(\boldsymbol{x}) = 0$, are exactly the extremal points (maxima or minima) of the function*

$$g(\boldsymbol{x}) = f(\boldsymbol{x}) + \lambda_1 c_1(\boldsymbol{x}) + ... + \lambda_K c_K(\boldsymbol{x}) \tag{255}$$

where the quantities λ_1, ..., λ_K are constant factors termed Lagrange multipliers whose values must be determined by requiring that the extremal points satisfy the original constraints $c_1(\boldsymbol{x}) = 0$, ..., $c_K(\boldsymbol{x}) = 0$.

This theorem is something of a mouthful, so let's break it down more carefully. First, let's start with the notation. The bold script notation $\boldsymbol{x}$ here means that the functions $f(\boldsymbol{x})$ and $c_i(\boldsymbol{x})$ depend on several variables x_1, x_2, ... that we haven't bothered to specify explicitly. The bold script symbol $\boldsymbol{x}$ is just a short-hand notation for all of these coordinate collectively.

Second, what exactly does it mean to refer to an equation like "$c_1(\boldsymbol{x}) = 0$" as a constraint? The idea here is that, most generally, a constraint is just some mathematical relationship between the coordinates x_1, x_2, etc. For example, we could require that

$$x_1 = x_2$$

or that

$$x_1 + x_2 = 1$$

or that

$$\sum_i x_i = 0.$$

The point is that we can always *write* such a relationship in the form $c(\boldsymbol{x}) = 0$. In the three examples above, the last one is already in this form, while the others can be brought into it as

$$x_1 - x_2 = 0$$

or

$$x_1 + x_2 - 1 = 0,$$

respectively. Thus we can regard the constraint as an equation $c(\boldsymbol{x}) = 0$, as long as we choose $c(\boldsymbol{x})$ appropriately.

Finally, the theorem above asserts that the extremal points of any function $f(\boldsymbol{x})$ subject to a given set of such constraints *are the same as* the extremal points of the function $g(\boldsymbol{x})$ defined in the theorem. Thus by finding the extrema of $g(\boldsymbol{x})$, we can also find the extrema of $f(\boldsymbol{x})$. This result is not at all obvious, and we don't have the space here to explain it in detail. Qualitatively, though, the idea behind this is that in constrained optimization problems, the condition for extremization is that, at the extremal points, the differential of $f(\boldsymbol{x})$ must be zero for any small displacement *consistent with the specified constraints.* This contrasts with the situation in *unconstrained* optimization, where *df* must be zero for *any* small coordinate displacement. The extra terms $\lambda_1 c_1(\boldsymbol{x}) + \lambda_2 c_2(\boldsymbol{x}) + ...$ ensure that the displacements considered in the optimization are constructed so as to preserve the constraints.

The catch, of course, is that the factors λ_i (the "Lagrange multipliers") used to define $g(\boldsymbol{x})$ are *not known in advance.* They must be determined in each case by the original set of constraints imposed on the problem. For example, suppose $\boldsymbol{x}$ contains N coordinates $x_1, ..., x_N$, and there are K constraints $c_1(\boldsymbol{x}), ..., c_K(\boldsymbol{x})$ on the system. Extremization of $g(\boldsymbol{x})$ then implies that

$$\left(\frac{\partial g}{\partial x_i}\right)_{x_{j\neq i}} = 0, \tag{256}$$

for each of the N coordinates x_i. Each such equation contains (some of) the K variables $\lambda_1, ..., \lambda_K$ – a system of N variables in K unknowns. So long as $K \leq N$, this system determines a unique solution to the problem. If $K > N$, there are more constraints than coordinates, and (unless some of the constraints are equivalent) the problem has no solution – the original optimization question is ill-posed.

22.3 Application to the Microcanonical Ensemble

The method will likely become clearer when we apply it to our microcanonical optimization problem. Here we want to maximize the Shannon information entropy [Eq. (250)] subject to the constraint that the probabilities p_μ sum to unity, i.e.,

$$\sum_\mu p_\mu = 1. \tag{257}$$

Writing this in the form $c(\boldsymbol{p}) = 0$, we define the extremization function

$$g(\boldsymbol{p}) = -\sum_\mu p_\mu \ln p_\mu + \lambda \left(1 - \sum_\mu p_\mu\right). \tag{258}$$

According to the Lagrange multipliers method outlined above, the maximum entropy distribution can be identified by maximizing $g(\boldsymbol{p})$. For each microstate ν, this gives an equation

$$0 = \left(\frac{\partial g}{p_\nu}\right)_{p_{\mu\neq\nu}} = \frac{\partial s}{p_\nu} + \lambda \frac{\partial}{\partial p_\nu}\left(1 - \sum_\mu p_\mu\right) \tag{259}$$

$$= -\ln p_\nu - 1 - \lambda. \tag{260}$$

Rearranging the expression, we obtain

$$p_\nu = e^{-(\lambda+1)}. \tag{261}$$

Since the right-hand-side is independent of ν, this means that *all microstates in the microcanonical ensemble are equally likely.* The normalization factor λ must be determined by the original constraint that $\sum_\mu p_\mu = 1$ or

$$1 = \sum_\mu p_\mu = \sum_\mu e^{-(\lambda+1)} = N_{\text{states}} e^{-(\lambda+1)}, \tag{262}$$

where N_{states} is the total number of microstates consistent with the thermodynamic constraints (constant energy, volume, and number of particles). The multiplier λ must thus satisfy the equation

$$e^{-(\lambda+1)} = \frac{1}{N_{\text{states}}}, \tag{263}$$

so that we find

$$p_\mu = \frac{1}{N_{\text{states}}}. \tag{264}$$

This result simply expresses formally what we already expect intuitively: if we have no information upon which to distinguish between two different states μ and ν, then they should be assigned the same probability, i.e., $p_\mu = p_\nu$.

22.4 Entropy and Entropy

From this result, we can readily calculate the information entropy of the system:

$$s(\{p_\mu\}) = -\sum_{\mu=1}^{N_{\text{states}}} p_\mu \ln p_\mu \tag{265}$$

$$= -\sum_{\mu=1}^{N_{\text{states}}} \frac{1}{N_{\text{states}}} \ln \frac{1}{N_{\text{states}}} \tag{266}$$

$$= \ln N_{\text{states}}. \tag{267}$$

Referring back to Eq. (91), we see that this is *almost* exactly the Boltzmann entropy formula we introduced near the start of the course. The only differences are that

1. Boltzmann's formula is for the *thermodynamic* entropy, while our result is for the *information* entropy, and
2. Boltzmann's formula includes a prefactor of $k_{\text{B}} = \frac{R}{N_{\text{A}}}$, where R is the ideal gas constant and N_{A} is Avogadro's number.

These observations suggest the *postulate* that the information entropy is, in fact, directly proportional to the thermodynamic entropy:

Entropy Postulate: The thermodynamic entropy S for a macroscopic system can be calculated from the appropriate thermodynamic ensemble as the average

$$S = k_{\mathrm{B}} s(\{p_\mu\}) = -k_{\mathrm{B}} \sum_\mu p_\mu \ln p_\mu. \tag{268}$$

In fact, Boltzmann's entropy formula is simply his own statement of this postulate, applied to the microcanonical ensemble.

This postulate is supported by the fact that both the information entropy and the thermodynamic entropy are maximized at equilibrium. It should be emphasized, however, that ultimately this statement is a *postulate* because it cannot (in general) be derived from a more fundamental theory. In principle, the entropy of a real system could be calculated directly from a theoretical model using Eq. (105) or (106), but in practice such calculations are far too difficult to be completed. Instead, we *hypothesize* that the thermodynamic and information entropies are directly proportional to each other and then *test* this hypothesis by comparing the resulting statistical predictions against known thermodynamic relationships. As it turns out, the hypothesis *satisfies all known thermodynamic relationships*, so we regard it as correct to our best understanding.

23 The Canonical Ensemble

23.1 Definition

A similar approach can be applied to maximize the entropy for systems in the *canonical ensemble*, i.e., with N, V, and T held constant. The first two of these requirements are the same as in the microcanonical ensemble, and can be directly accounted for in our calculations by considering only those microstates with the correct volume and number of particles.

The last requirement (constant T) is a little trickier. In our development of thermodynamics, we *defined* the temperature in terms of the properties of ideal gases. Things are a bit more complex from a statistical perspective since we must calculate the properties of the ideal gas statistically *before* we can use its properties to define a temperature scale!

The practical resolution to this difficulty is to use a different condition to characterize the canonical ensemble: that the *average* energy be fixed at some specified value U. This is a constraint that we can easily incorporate into our maximum entropy formalism. And, once we've completed our calculations, we can show that the average-energy requirement is *equivalent* to fixing the temperature on the ideal-gas scale. Although this approach is a little round-about, it does turn out to be logically consistent, and it makes the math much easier than if we tried to use fixed temperature as our starting point for entropy maximization.

23.2 Maximizing the Entropy

With this introduction, our problem reduces to maximizing the information entropy $s(\{p_\mu\})$ for a system of microstates at fixed V and N subject to the constraints

- $\sum_\mu p_\mu = 1$
- $\sum_\mu p_\mu \varepsilon_\mu = U$.

The first constraint says that the sum of the probabilities must be one (a requirement for the set $\{p_\mu\}$ to be a well-defined probability distribution), while the second is exactly the condition that the average energy must equal the thermodynamic energy U. Here ε_μ is the total energy of the microstate μ, i.e., the value of the *total system Hamiltonian* evaluated at the coordinates and momenta specified for all particles by the microstate μ.

According to the method of Lagrange Multipliers, this means we must maximize the function

$$g(\{p_\mu\}) = -\sum_\mu p_\mu \ln p_\mu + \lambda \left(1 - \sum_\mu p_\mu\right) + \beta \left(U - \sum_\mu p_\mu \varepsilon_\mu\right). \tag{269}$$

Here λ is the Lagrange multiplier associated with the normalization condition $\sum_\mu p_\mu = 1$, and β is the independent Lagrange multiplier associated with the constraint $\sum_\mu p_\mu \varepsilon_\mu = U$. Calculating the partial derivative with respect to the probability p_ν of an arbitrary microstate ν, we obtain

$$\frac{\partial g}{\partial p_\nu} = -\ln p_\nu - 1 - \lambda - \beta \varepsilon_\nu = 0 \tag{270}$$

or after rearranging terms and exponentiating both sides

$$p_\nu = e^{-(\lambda+1)} e^{-\beta \varepsilon_\nu}. \tag{271}$$

The first of these factors has exactly the same form as we saw previously in the microcanonical optimization; it represents an overall normalization constant that ensures that the probabilities sum to unity. The second term is different: this term varies by state and indicates that *the probability of a microstate in the canonical ensemble is an exponential function of its energy.*

23.3 Canonical Normalization Constraint: The Partition Function

The exact values of the constants λ and β must be determined by making sure that they satisfy the normalization and average-energy constraints listed above. The normalization constraint is satisfied if

$$p_\mu = \frac{e^{-\beta\varepsilon_\mu}}{Z}, \tag{272}$$

where

$$Z \equiv \sum_\mu e^{-\beta\varepsilon_\mu} \tag{273}$$

is termed the *partition function.* We'll comment more carefully on what this function represents physically once we've established the value of β. For now, notice just that it is a normalization constant required to ensure that the probabilities p_μ sum to unity.

Since the probability of each microstate is exponentially weighted, some states are more likely to be occupied than others; in particular, if $\beta > 0$ (which we will see shortly is true), then low-energy states (where ε_μ is small) will be much more likely to be occupied than high-energy states. Intuitively, you might suspect that the parameter β should be related to the *temperature* of the system. Then our probability formula would say that many states are accessible to the system at high temperatures but only a few are accessible at low temperatures. But to develop this idea fully, we'll have to identify exactly what β represents thermodynamically.

To do that, it's useful to note first that many key thermodynamic quantities can be calculated as partition function derivatives. For example, the average energy can be calculated as

$$\begin{aligned}
\sum_\mu p_\mu\varepsilon_\mu &= \frac{\sum_\mu \varepsilon_\mu e^{-\beta\varepsilon_\mu}}{Z} && (274)\\
&= -\frac{\sum_\mu \left(\frac{\partial}{\partial\beta} e^{-\beta\varepsilon_\mu}\right)}{Z} = -\frac{\frac{\partial Z}{\partial\beta}}{Z} && (275)\\
&= -\frac{\partial \ln Z}{\partial\beta}. && (276)
\end{aligned}$$

Since the average energy is constrained to match the thermodynamic U, this implies that

$$U = -\frac{\partial \ln Z}{\partial\beta}. \tag{277}$$

We'll use this relation in the next section when we impose the constraint of constant energy.

23.4 The Inverse Temperature

Our last fundamental task in developing the canonical ensemble is to answer the question: How is β related to real thermodynamic parameters? Intuitively, we already surmised that β might be related to the thermodynamic temperature T. To flesh this idea out fully, let's see if we can derive an explicit expression for T in terms of statistically calculable quantities.

Recall from our first-law differential relation that, when volume and number of particles remain fixed, the entropy, energy, and temperature are connected by the relation

$$dU = TdS \qquad \text{Constant } V,\, N. \tag{278}$$

This means that, in statistical applications, we can identify the thermodynamic temperature simply by checking how U varies with small variations in S.

Since we're in the canonical ensemble, as long as we keep N and V fixed, the set of microstates μ that together produce the ensemble remain fixed. Only their relative probabilities p_μ are free to vary. This means that any differential displacements in the energy dU or entropy dS must result solely from differential changes to the probabilities dp_μ. With N and V fixed, the only parameter left to vary in the probabilities p_μ is the Lagrange multiplier β. So all displacements in dU and dS (in the Canonical ensemble) must result from displacements $d\beta$.

Now, note that the entropy S can be related to the energy U by

$$\begin{aligned} S &= -k_\mathrm{B}\sum_\mu p_\mu \ln p_\mu = -k_\mathrm{B}\sum_\mu p_\mu \ln\left(\frac{e^{-\beta\varepsilon_\mu}}{Z}\right) && (279)\\ &= k_\mathrm{B}\sum_\mu p_\mu\left(\beta\varepsilon_\mu + \ln Z\right) && (280)\\ &= k_\mathrm{B}\beta U + k_\mathrm{B}\ln Z. && (281)\end{aligned}$$

This means that

$$\begin{aligned} dS &= -k_\mathrm{B}\left(Ud\beta + \beta dU + \frac{dZ}{Z}\right) && (282)\\ &= -k_\mathrm{B}\left(Ud\beta + \beta dU + d\ln Z\right) && (283)\\ &= -k_\mathrm{B}\left(Ud\beta + \beta dU + \frac{\partial \ln Z}{\partial\beta}d\beta\right) && (284)\\ &= -k_\mathrm{B}\left(Ud\beta + \beta dU - Ud\beta\right) && (285)\\ &= k_\mathrm{B}\beta dU && (286)\end{aligned}$$

or, after rearranging terms,

$$dU = \frac{1}{k_\mathrm{B}\beta}dS. \tag{287}$$

Comparing Eqs. [(278)] and [(287)], we see that, for our statistical model to be consistent with thermodynamics we must have

The Inverse Temperature

$$\beta \equiv \frac{1}{k_\mathrm{B}T} \tag{288}$$

for any system in the canonical ensemble.

With this result, we can rewrite our probability and partition function formulas as

$$p_\mu = \frac{e^{-\frac{\varepsilon_\mu}{k_\mathrm{B}T}}}{Z}, \tag{289}$$

where the partition function is

$$Z \equiv \sum_\mu e^{-\frac{\varepsilon_\mu}{k_\mathrm{B}T}}. \tag{290}$$

In this form we can see more clearly that *the probability p_μ for a microstate μ in the canonical ensemble decreases exponentially with the ratio of its energy to the "thermal energy" k_BT.*

As we've commented previously, the term "thermal energy" can be applied to many different quantities, so one should be a bit cautious in using it. In the present context, though, it seems particularly appropriate: the quantity $k_\mathrm{B}T$ determines *which microstates are thermally populated at temperature* T. If $\varepsilon_\mu \ll k_\mathrm{B}T$, then the population will involve the exponent of a very small negative number, so that the $e^{-\frac{\varepsilon_\mu}{k_\mathrm{B}T}}$ term approaches unity. On the other hand, if $\varepsilon_\mu \gg k_\mathrm{B}T$, then the population will involve the exponent of a very large negative number so that $e^{-\frac{\varepsilon_\mu}{k_\mathrm{B}T}}$ approaches zero. The quantity $k_\mathrm{B}T$ can thus be thought of as the *amount of energy available from the thermal environment* that can be used to populate the system microstates, i.e., the "thermal energy".

23.5 The Partition Function as Number of Thermally Accessible States

In our formulas above, we introduced the partition function Z as simply a normalization constant used to ensure that the sum of all microstate probabilities is exactly unity. A more interesting interpretation is obtained if we note that exactly the same probabilities p_μ are obtained if we multiply the top and bottom of Eq. (289) by an overall constant. In particular, suppose in a particular system that ε_min is the *lowest* energy of all microstates μ, i.e.

$$\varepsilon_\mathrm{min} = \min_\mu \varepsilon_\mu. \tag{291}$$

We can then recast Eqs. (289) and (290) as

$$p_\mu = \frac{e^{-\beta(\varepsilon_\mu - \varepsilon_\mathrm{min})}}{Z'} \tag{292}$$

with

$$Z' = \sum_\mu e^{-\beta(\varepsilon_\mu - \varepsilon_\mathrm{min})}. \tag{293}$$

Effectively, what we've done is to shift the system energy scale, defining a new "zero" for energy but leaving the *relative* probabilities of each state unchanged. Since these equations are obtained by simply multiplying through the equation for p_μ by the constant $e^{+\beta\varepsilon_\mathrm{min}}$, the actual value of the *probabilities* is left unchanged.

The partition function Z', however, now takes on a new meaning: it still serves as a normalization constant but it also *counts the number of microstates that are thermally accessible to the system.* For example, suppose the system has a unique "ground state", i.e., that there's only one state that has energy $\varepsilon_\mu = \varepsilon_\mathrm{min}$. Without losing generality, let's refer to this microstate as $\mu = 0$. Then for this particular microstate,

$$p_0 = \frac{e^{-\beta(\varepsilon_\mathrm{min} - \varepsilon_\mathrm{min})}}{Z'} = \frac{1}{Z'}. \tag{294}$$

In the low-temperature limit $T \to 0$ (which is equivalent to $\beta \to \infty$), Z' approaches 1 since the probabilities for all states *other than* $\mu = 0$ approach zero. More generally, if there are N_deg "ground states" with energy $\varepsilon_\mu = \varepsilon_\mathrm{min}$, then the partition function Z' approaches N_deg in the low-temperature limit. Conversely, suppose the system possesses N_states microstates in total. In the *high-temperature* limit $T \to \infty$, which is equivalent to $\beta \to 0$, the exponent $\beta\left(\varepsilon_\mu - \varepsilon_\mathrm{min}\right)$ approaches zero for *all* microstates, so that the partition function Z' approaches N_states.

Thus in both the high- and low-temperature limits, the modified partition function Z' reduces to just the number of microstates the system can occupy. At intermediate temperatures, Z' interpolates between these two limits, transitioning smoothly from N_deg to N_state and "counting" the number of thermally accessible states at each temperature.

24 Lattice Polymers

24.1 The Protein Folding Problem

To see how these rather abstract formulas apply to real-life systems, we'll explore next a simple model for protein folding. Protein folding is an incredibly complex process that in many ways defies simplistic explanations. Nevertheless, a few characteristic "driving forces" in protein folding can be identified:

- The **decrease in protein entropy** associated with folding into a single, well-defined conformation typically *opposes* folding. This observation offers a simple (in fact overly simplistic) explanation for why proteins tend to unfold at high temperatures. However, it should be born in mind that even a "folded" protein structure actually possesses some degree of conformational freedom since both the backbone and (especially) side chains are still able to flex and rearrange themselves locally, even in the folded state. This residual freedom helps somewhat to offset the entropy loss associated with folding.

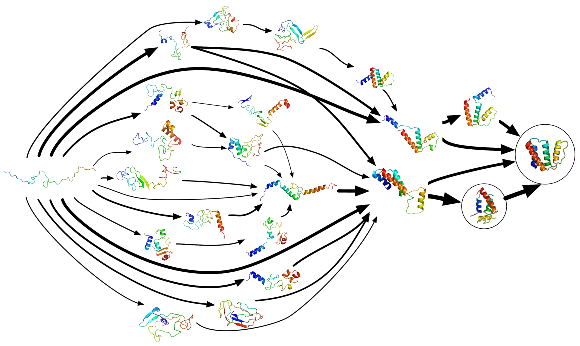
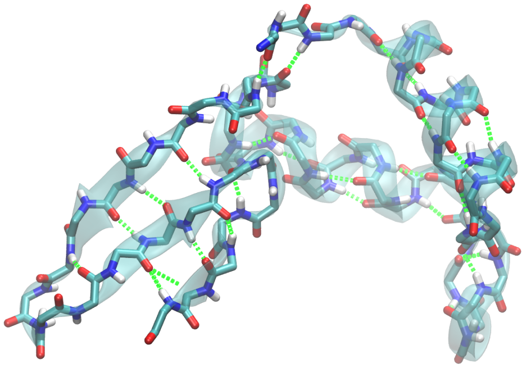
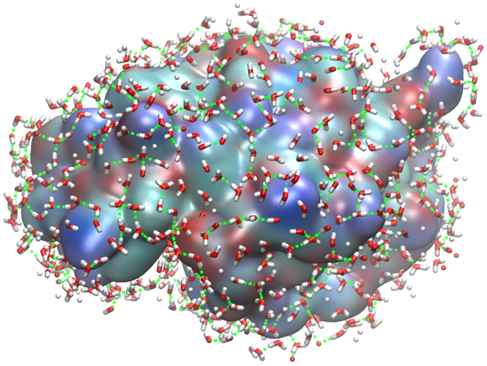
  

  Image by Vincent Voelz, online at `https://commons.wikimedia.org/wiki/File:ACBP_MSM_from_Folding@home.tiff`. Related to publication Voelz et al. *J. Am. Chem. Soc.* 2012, 134, 12565-12577; `http://dx.doi.org/10.1021/ja302528z`

- **Hydrogen-bonding interactions** within the protein backbone and between polar sidechains stabilize the folded structure. In many cases, this contributes to a *negative enthalpy change* associated with protein folding. However, protein-folding also *decreases* the number of hydrogen-bonding interactions that the protein can form with the solvent (usually water), which to some extent counteracts the stabilization from protein-protein interactions.

- Protein folding usually **increases the entropy of the solvent** (usually water) since the folded state of the protein is usually more compact than the unfolded state; this leaves water molecules that would otherwise be occupied with protein-water interactions free to diffuse through the entire solvent volume, increasing the solvent entropy.



As you may have noticed, the various "driving forces" for protein folding are very much in competition with each other. Protein folding *decreases* the entropy of the protein but *increases* the entropy of the solvent; it *increases* the number of energetically favorable protein-protein hydrogen bonds but *decreases* the number of protein-water hydrogen bonds. And the effects we've described are only the tip of the iceberg! Other key factors that are even harder to quantify include the *flexibility* of the folded structure (which can substantially increase the entropy of the folded state) and, most importantly, the *specific interactions* enabled by the unique sequence of amino acids that compose the protein chain. Examples of sequence-specific features include salt bridges formed between Arg and Lys residues and the steric "kinks" introduced into the backbone by the presence of a Pro residue.

The complexity and interconnectedness of these effects leads to what is known as the *protein-folding problem*:

The **Protein-Folding Problem** is a broad question in biophysical chemistry that asks *how the folded state of a protein can be predicted from its amino acid sequence* and *what mechanisms guide proteins to fold to their native structure.*

In the early days of biophysics, it was expected that relatively simple rules should be available to predict folded protein structures from the amino acid sequence, similar to the way a few simple hydrogen-bonding and base-stacking patterns account for the three-dimensional crystal structure of double-stranded DNA. As time passed, however, and as more crystal structures became available, it became gradually clearer that the protein-folding problem is much more complex than the parallel problem for DNA.

24.2 Lattice Proteins

In recent years, the *quality* of our computational models and the *quantity* of available computing power have together produced significant advances against this challenge. In what is becoming an impressive variety of cases, now, native protein structures *can* be predicted from only the amino acid sequence, and novel protein structures can even be designed from scratch. Nonetheless, the governing principles behind the problem remain complex since there are a huge number of degrees of freedom that must be simultaneously considered in analyzing the sequence of any given protein.

For exactly this reason, protein folding is very much a statistical problem and is ripe for analysis in terms of the canonical ensemble we've just described.

Since the full protein folding problem is much too complex for us to treat here, we'll consider a simplified problem know as the *lattice protein* model.

A **Lattice Protein** model is a simplified description of the protein-folding problem, where the protein chain is modeled by a sequence of "beads" (representing amino acids) on a 2D or 3D grid (the "lattice").

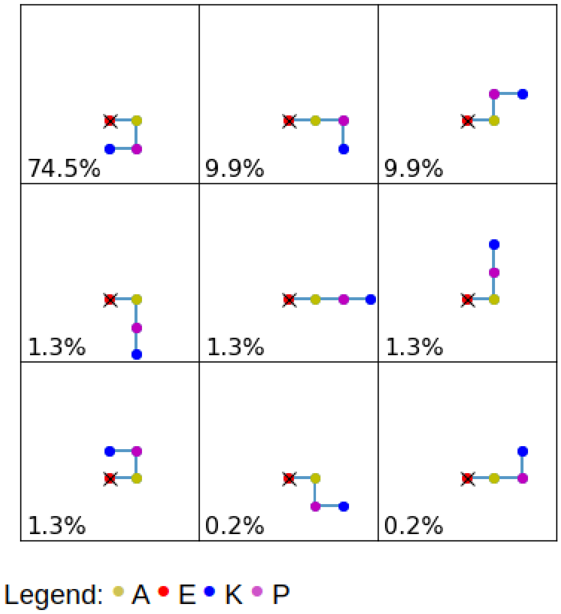
The use of a discrete lattice in this model (rather than the full, continuous, three-dimensional space available to real protein sequences) dramatically simplifies its analysis and in many cases allows us to enumerate conformational microstates explicitly.

To keep things simple, we'll think about protein conformations in only **two-dimensions** (2D), and we'll consider peptides made of only four amino acids with (dramatically!) simplified physical properties. Given a peptide sequence, we'll enumerate (computationally) all the possible conformations μ of the peptide on our 2D lattice, assign each conformation an energy ε_μ based on the amino-acid interactions in that structure, and then assign temperature-dependent probabilities in the canonical ensemble. Real proteins are, of course, much more complicated than the 2D models we'll study here, but this simplified model will at least give us a small taste of the kind of energetic and entropic forces the drive proteins to fold into stable, well-defined three-dimensional structures.

Specifically, we'll think about peptide sequences composed of four idealized amino acids:

- **Alanine** (Ala / A) residues will be treated as purely hydrophobic: A favorable energy shift $-\varepsilon_\mathrm{o}$ will be assigned for each Ala-Ala contact in our peptide structure.
- **Lysine** (Lys / K) is a positively charged amino acid. We assign an unfavorable energy offset $+\varepsilon_\mathrm{o}$ for each Lys-Lys contact and a favorable offset $-\varepsilon_\mathrm{o}$ for every Lys-Glu contact (see next item).
- **Glutamate** (Glu / E) is a negatively charged amino acid. We assign an unfavorable energy offset $+\varepsilon_\mathrm{o}$ for each Glu-Glu contact and a favorable offset $-\varepsilon_\mathrm{o}$ for every Lys-Glu contact.
- **Proline** (Pro / P) is a sterically constrained amino acid due to the fact that its alkyl sidechain is chemically bonded to the amine nitrogen, forming a five-membered ring. This leads to somewhat complicated steric constraints on the roles Pro can serve in protein structures, but we'll simplify the situation in our 2D model to assign
  - A favorable energy shift of $-\varepsilon_\mathrm{o}$ when the peptide chain makes a right-hand turn at a Pro residue
  - An unfavorable shift of $+\varepsilon_\mathrm{o}$ when the peptide chain makes a left-hand turn at a Pro residue
  - No offset when the chain is straight at a Pro residue.

Consider, for example, the nine possible combinations of the tetra-peptide EAPK sketched in the diagram below. Here each dot in the diagram represents an amino acid, and the thin blue lines indicate the amide bonds connecting them.



The conformer in the top left corner has an assigned energy of

$$\varepsilon_\mu = -2\varepsilon_\text{o} \tag{295}$$

since it is stabilized by both a Glu-Lys contact (salt bridge) and the favorable right-hand turn of the Pro residue. The next two conformers (top-center and top-right) are stabilized by favorable Pro conformations but not by any electrostatic contacts and hence are assigned slightly higher conformation energies of $\varepsilon_\mu = -\varepsilon_\text{o}$. The three conformations in the middle row have energies of $\varepsilon_\mu = 0$ since there are no electrostatic or hydrophobic contacts, and the Pro residue is in the neutral (neither favored nor disfavored) straight conformation. The bottom-left conformer is also assigned $\varepsilon_\mu = 0$ since the favorable electrostatic interaction is counter-acted by an unfavorable Pro conformation. Finally, the bottom-center and bottom-right conformations are assigned the unstable value $\varepsilon_\mu = +\varepsilon_\text{o}$ due to the disfavored left-turn Pro conformation and the lack of any stabilizing electrostatics.

From these conformation energies, we can calculate conformation *probabilities* for our peptide using the Canonical ensemble expressions of Eqs. (289) and (290). From the nine conformations in the figure above, we can calculate the partition function as

$$Z' = \sum_{\mu=1}^{9} e^{-\frac{(\varepsilon_\mu - \varepsilon_\text{min})}{k_\text{B}T}} \tag{296}$$

$$= e^0 + 2e^{-\beta\varepsilon_\text{o}} + 4e^{-2\beta\varepsilon_\text{o}} + 2e^{-3\beta\varepsilon_\text{o}}. \tag{297}$$

Now suppose for concreteness that

$$\varepsilon_\text{o} = 2 \cdot k_\text{B} \cdot 300\,\text{K}. \tag{298}$$

This gives a partition function at 300 K of

$$Z(300\,\text{K}) = 1 + 2e^{-2} + 4e^{-4} + 2e^{-6} \approx 1.35. \tag{299}$$

Intuitively, this says that the system is *mostly* restricted to the lowest-energy conformation, with a small population in the higher-energy conformations. The populations can be calculated quantitatively using Eq. (292); for example, the population of the lowest-energy conformer is simply

$$p_0 = \frac{1}{Z'} \approx 74\%. \tag{300}$$

The remaining populations at $T = 300\,\mathrm{K}$ are reported in the lower-right corner of each figure.

24.3 LatticeProtein Simulation App

To get a more "hands-on" feel for these lattice polymer simulations, go to the LatticeProtein app at `https://nanohub.org/tools/LatticeProtein`. (NanoHUB is a science computing site administered in part by Purdue, in collaboration with other research universities; you can create a free account using your Purdue login credentials.) There you can design peptides with between 1 and 15 amino acids, chosen from the four options (A, E, K, and P) that we've described above. You'll use this app to complete your homework assignment.

25 Protein Folding: Kinetics

25.1 Introduction to Kinetics

The LatticeProtein app we introduced last lecture helps to illustrate the driving forces that determine the *first* half of the protein folding problem: How does sequence determine the final structure of a folded protein? Even if we could answer that question exactly, however, there remains a difficult "second half" to the problem: How do proteins sort through an astronimical number of possible conformations to locate the "correct" folded structure *in a reasonable amount of time*?

This question of *how fast* a protein is able to transition between unfolded and folded states is our first introduction to the field of *kinetics* which studies *how fast different physical and chemical processes occur.* In contrast to thermodynamics, which focuses on the physical properties of *equilibrium systems*, kinetics asks *how quickly* the system reaches equilibrium. Such questions will be the focus of our next several lectures.

Before we dig into the topic, though, let's stop to define the scope of our discussion. "Kinetics" is a vast field, encompassing everything from enzyme ligand-binding problems to carbohydrate metabolism and even the rate at which chlorophyll molecules absorb electromagnetic energy from the sun. Since we can't possibly cover *all* of these topics in this course, we will instead simply cover a smattering of (hopefully!) interesting examples, introducing key concepts along the way. In today's lecture, we'll get a glimpse at how macroscopic protein-folding properties can help us to infer the basic mechanisms by which proteins fold at the microscale.

25.2 Levinthal's Paradox

The basic kinetic problem we're interested today is: *How does a protein manage to fold to the "correct" three-dimensional structure so quickly?* Experimentally, it's known that proteins can be "denatured" by either raising the temperature or adding certain "chaotropic" (literally "chaos-inducing") agents like urea or sodium dodecyl sulfate. In these denatured states, the native secondary structure of the protein is lost, and the protein is able to access an astronomical number of different conformations (the quoted paragraph below gives an estimate). In our lattice polymer simulations, this effect is visible via the observation that the partition function Z' increases with increasing temperature.

All of these experimental observations are consistent with expectations from statistical mechanics: The protein occupies a well-defined most-stable state at low temperatures but samples a much larger number of conformations at high temperatures. What is *not* necessarily expected is that, when the temperature is again lowered, the protein is able very quickly (typically in much less than a second) to "re-discover" the correct folded structure. The significance of this experimental finding is made clear when one considers just how many possible states an unfolded protein has available to it. If these states were sampled randomly, it would take an astronomically long time for the protein to fold. This contrast between real and naively expected protein-folding rates is often termed the "Levinthal Paradox" after Cyrus Levinthal who first formulated it precisely.

A concise summary of the paradox is provided by a recent protein-folding review:

> How long does it take for a protein to fold up into its native structure? In a standard illustration of the Levinthal paradox, each bond connecting amino acids can have several (e.g., three) possible states, so that a protein of, say, 101 amino acids could exist in $3^{100} = 5 \cdot 10^{47}$ configurations. Even if the protein is able to sample new configurations at the rate of 10^{13} per second, or $3 \cdot 10^{20}$ per year, it will take 10^{27} years to try them all. Levinthal concluded that random searches are not an effective way of finding the correct state of a folded protein. Nevertheless, proteins do fold, and in a time scale of seconds or less. This is the paradox.

[Quote from: Zwanzig R, Szabo A, Bagchi B. "Levinthal's paradox." *Proc Natl Acad Sci U S A*. 1992;89(1):20-22. `doi:10.1073/pnas.89.1.20`. Available online at `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC48166/`]

In other words, Levinthal's paradox points out that, if proteins *really* folded by sampling *all possible conformations* until they found the "correct" lowest-energy state, then it would take them an impossibly long time to fold!

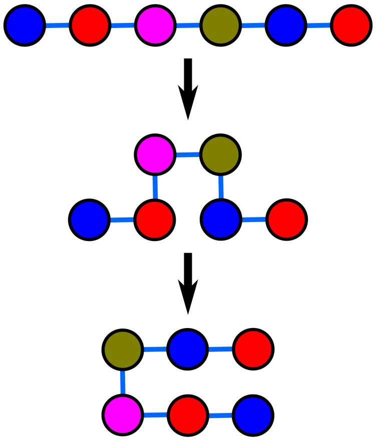
25.3 Resolution of the "Paradox"

How is this paradox resolved? It might be a stretch to say that the answer is fully known, but certainly we can outline some general principles.[9]

First, note that the problem outlined above – proteins in real life fold much faster than might be expected based on simple back-of-the-envelope calculations – isn't really a proper paradox (in the sense of a "self-contradiction"): it's just a sign that some of our back-of-the-envelope calculation made an invalid assumption somewhere along the line! In other words, real proteins must *not* actually sample *all* possible configurations on their way to the folded state. This was exactly the point Cyrus Levinthal was trying to make in his original paper.[10]

Levinthals argument thus shows that, however it is that proteins fold, they must *not* simply sample all possible configurations until they find the right one. From here, two possible explanations can be suggested:

- **Folding Pathways:** Levinthal's proposed solution to this problem was to propose that proteins fold by way of *specific mechanistic pathways*, i.e., that every time the protein fold it follows a single (or at least a small number) of well-defined paths from one unfolded state to another unfolded state until it reaches the final conformation. A hypothetical example is illustrated in the image below where a fully extended peptide chain first forms a β-turn triggered by a proline residue and stabilized by a Glu-Lys contact, which in turn brings the final Glu/Lys pair into close enough contact to "close the loop" to the final state. At the time of Levinthals' proposal, however, there was no way to verify this model, either experimentally or theoretically. Thus for many years it remained simply an untested hypothesis.

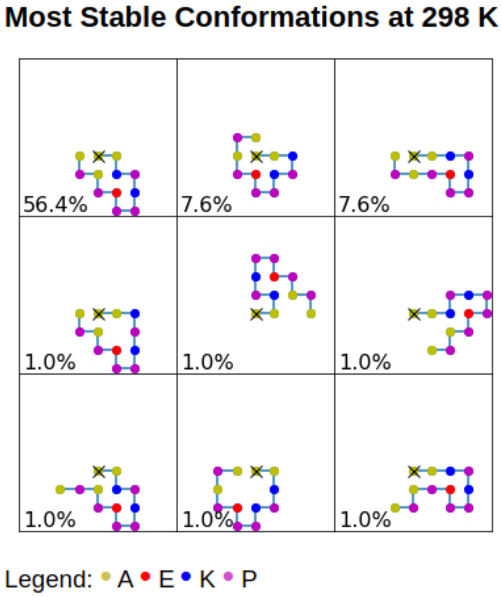


---

[9] For a more complete discussion, see `http://ww2.chemistry.gatech.edu/~lw26/course_Information/6572/papers/karplus_1997.pdf`

[10] In fact, Levinthal was arguing that proteins might not even make it to the lowest-energy conformation at all; they might remain kinetically trapped at a local energy minimum somewhere along the way. For Levinthal's argument see `https://www.cc.gatech.edu/~turk/bio_sim/articles/proteins_levinthal_1969.pdf`

- **Folding Funnel:** More recently, based in part on ultrafast kinetics measurements of real protein-folding events, it has been recognized that even without unique folding pathways, global driving forces can push an unfolded peptide rapidly toward *many different* possible pathways that all lead eventually to the folded state. Perhaps the most famous model for such process is the *hydrophobic collapse* model that suggests the key driving force for many proteins to fold is that hydrophobic residues like Ala "collapse" quickly to form a hydrophobic core (or "molten globule") that, although not actually part of the native structure, forces many residues into close enough contact that they rapidly "find" their bonding partners. For example, consider the lattice peptide populations depicted below for the peptide AAKPKPPEPAPA. Although there is only one unique lowest-energy state, there are a variety of *similar* states that are stabilized by both polar and hydrophobic interactions. An initial "collapse" event in which the terminal Ala residues on each end come into contact with each other can push the peptide rapidly toward the final folded state by keeping the remaining residues *close enough* to their final conformations that they can find their ideal binding partners.



The situation for any *real* protein is, of course, much more complex than can be described by simple lattice models and very likely contains features similar to each of these models, and perhaps many others. Modern computational power and improved molecular dynamics force fields provide a (hopefully!) more realistic picture of how real proteins fold. The video linked below shows one such example, one of the earliest explicit molecular dynamics simulations of a folding pathway for a real peptide: `https://www.youtube.com/watch?v=gFcp2Xpd29I`

26 Chemical Kinetics

26.1 Reaction Rates

The last section introduce the science of kinetics with the particular example of protein folding. But the question of "how fast" a process occurs can be applied to any chemical reaction and, indeed, even to many non-chemical processes. Although we'll have time in this course for only a cursory introduction to the topic, we can at least set the stage for future study by introducing some basic vocabulary.

To begin, we need to clarify is what *exactly* we mean when we ask "how fast" a reaction occurs. This may sound trivial, but there are actually several different ways we define reaction "speed". For example, when we ask "how fast" a reaction occurs, we could mean: "How many moles of reactant are consumed in each second?" But by this definition, just doubling the size of our reactive solution also doubles "how fast" the reaction occurs. A craft brewer, for example, can (in principle) produce beer twice as fast in a 400 L fermentation tank compared to a 200 L tank.

But this scaling of the reaction "speed" with system size has nothing to do with the chemistry taking place inside, and it's not usually what we care about in chemical kinetics. In comparing fermentation rates between differently sized vats, for example, the craft brewer will be interested in the rate at which alcohol *concentration* changes, not the overall rate at which alcohol is produced (which depends on the size of the vat). For the same reason, in chemical kinetics, the term "reaction rate" is reserved for the rate at which the *concentration* of a given species changes with time. It's also convenient to divide the measured rate by the stoichiometric coefficient of the measured species, so that *all* reactants and products produce the same reaction rate. Formally, we can summarize this definition as follows:

The **reaction rate** v for a generic chemical reaction

$$\nu_1 \mathrm{X}_1 + \nu_2 \mathrm{X}_2 + ... + \nu_{N_\mathrm{r}} \mathrm{X}_{N_\mathrm{r}} \longrightarrow \nu_{N_\mathrm{r}+1} \mathrm{X}_{N_\mathrm{r}+1} + \nu_{N_\mathrm{r}+2} \mathrm{X}_{N_\mathrm{r}+2} + ... + \nu_{N_\mathrm{r}+N_\mathrm{p}} \mathrm{X}_{N_\mathrm{r}+N_\mathrm{p}}$$

is defined as

$$v = \pm \frac{1}{\nu_i} \frac{d[X_i]}{dt} \tag{301}$$

where $[X_i]$ refers to the *molar concentration* of species X_i; the "+" sign applies if species i is a product, and the "-" sign applies if the species i is a reactant.

With this clarification, we can state precisely that the central question in chemical kinetics is to determine *how the rate v of a chemical reaction varies with parameters like temperature, pressure, and (most importantly) the concentration of each species in the system.* In this course, we'll primarily focus on reaction kinetics in solution, where *concentration* and *temperature* are usually the primary driving forces. Pressure plays a much greater role in gas-phase kinetics, but we won't worry too much about it in this course.

26.2 Rate Laws and Elementary Reactions

Information about reaction rates are conveniently summarized in *rate laws*, an equation of the form

$$v = f\left([X_1], [X_2], ..., [X_{N_\mathrm{r}+N_\mathrm{p}}]; T\right), \tag{302}$$

where $f(...)$ denotes some mathematical function of the concentrations $[X_1]$, ..., $[X_{N_\mathrm{r}+N_\mathrm{p}}]$ of the various reactants and products in the system; the rate law also depends parametrically on the temperature T, although our focus is usually on the species concentrations. In general, these functions can be monstrously complex and often can't be determined exactly.

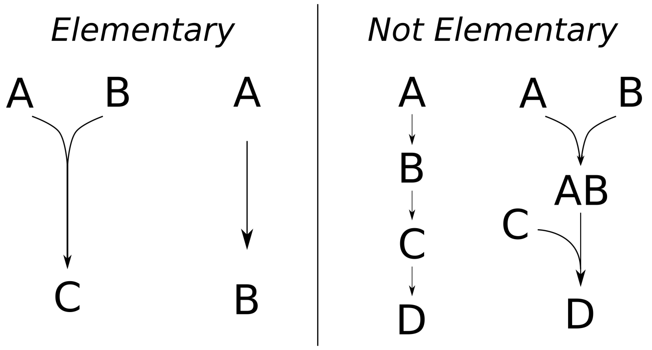
However, there some simple (but important!) cases cases where explicit expressions can be obtained. In this subsection, we'll discuss *elementary reactions*, for which rate laws can be written down directly in terms of empirical "rate constants" – temperature-dependent parameters that affect the reaction rate. In the next subsection, we'll discuss in more detail where those rate constants come from, and especially how they vary with temperature.

To begin, a definition:

An **elementary reaction** is one that takes place *in a single step* at the molecular level. Specifically, this means that

- The reaction happens *only* when all reactants encounter each other simultaneously to form a reactive complex.
- Once the reaction begins, it goes directly to completion, without the formation of multiple transition states.

For illustration, the nearby figure gives two examples of elementary reactions (left panel) and two examples of reactions that can*not* be considered elementary. The $A+B \to C$ reaction on the far left is elementary since the reaction proceeds to completion as soon as the two reactants collide. The $A \to B$ reaction is elementary since it involves only a single species A converting spontaneously to another species B. In contrast the reaction $A \to B \to C \to D$ is *not* elementary because it involves multiple intermediate steps. Similarly, the multi-step $A + B + C \to D$ reaction is *not* elementary because the A and B species must react to form a complex *before* they combine with reactant C to drive the reaction to completion.



For elementary reactions, it's easy to write down reaction rate laws. Because elementary reactions occur whenever (and *only* when) all the reactants encounter each other, the rate of the process is proportional to the rate at which such encounters occur. In both gas and solution phases, where the various reactants are distributed more or less randomly in space, this "encounter probability" is directly proportional to the *concentration* of each reactant. If a particular reactant appears more than once in the reactive complex, its concentration will occur more than once in the rate law. In general, an elementary reaction of the form

$$\nu_1 X_1 + \nu_2 X_2 + ... + \nu_{N_\mathrm{r}} X_{N_\mathrm{r}} \to \text{(products)} \tag{303}$$

will have a reaction rate

$$v = k[X_1]^{\nu_1}[X_2]^{\nu_2}...[X_{N_\mathrm{r}}]^{\nu_{N_\mathrm{r}}}, \tag{304}$$

where k is a numerical value called a *rate constant.*

26.3 Reaction Order

In Eq. (304), it is notable that the reaction rate v scales with the concentration of each species $[X_i]$ raised to the power ν_i of its stoichiometric coefficient. This exponent is referred to as the **order** of the reaction with respect to species X_i. The **overall order** of the reaction is the sum of all the stoichiometric coefficients of the reactants, i.e., $\nu_1 + \nu_2 + ... + \nu_{\mathrm{N}_r}$.

In the case that a reaction involves only a single species, the rate law becomes particularly simple and can be integrated directly. For illustration, consider the following examples:

- **First Order**: $A \rightarrow B$. The rate law for such reactions reads

  $$v = -\frac{d[A]}{dt} = k[A],$$

  which can be integrated to give

  $$[A](t) = [A]_0 e^{-kt}$$

- **Second Order**: $2A \rightarrow B$. The rate law reads

  $$v = -\frac{1}{2}\frac{d[A]}{dt} = k[A]^2,$$

  which can be integrated to give

  $$[A](t) = \frac{[A]_0}{1 + 2k[A]_0 t}.$$

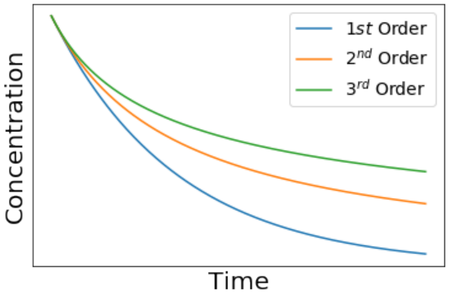
- **Third Order**: $3A \rightarrow B$. The rate law reads

  $$v = -\frac{1}{3}\frac{d[A]}{dt} = k[A]^3,$$

  which can be integrated to give

  $$[A](t) = \frac{[A]_0}{\sqrt{1 + 6k[A]_0^2 t}}.$$

For illustration, the figure below presents a plot of concentration ($[A](t)$) versus time for first-, second-, and third-order reactions, with parameters chosen so that the initial slope is the same. Note that in all three cases, the *rate* of the reaction (i.e., the slope of the concentration curve) eventually plateaus, approaching zero as the reactant A is consumed. Given the same initial rate, however, the 1$^{\mathrm{st}}$-order reaction goes to completion most quickly, while the 3$^{\mathrm{rd}}$-order reaction approaches completion most slowly. This is due to the fact that the higher reaction orders are *more sensitive* to the concentration of the reactant A and, thus, their reaction rates approach zero more quickly as the substrate is consumed.



For non-elementary reactions, rate laws become much more complicated and it is not always possible to assign a specific "order" of the reaction with respect to each species. However, when the rate of a reaction *does* scale in a simple way (i.e., with a fixed exponent) with respect to the concentration of a given species, we often say that the reaction is "first order" or "second order" (depending on the exponent), even if the reaction isn't elementary. Just keep in mind that for non-elementary reactions such distinctions may hold only under limited circumstances: it could be possible for a reaction to be first-order in reactant A under one set of conditions and second-order in A under other conditions. Such rate-law dependence usually indicates that the reaction is not, in fact, an elementary reaction.

26.4 Reaction Rates and Chemical Equilibrium

So far, we've treated chemical reactions as though they only go in one direction, e.g., $A \rightarrow B$. But all reactions are (to some extent) *reversible*: if $A \rightarrow B$ is possible physically, then so is the reverse reaction $B \rightarrow A$. How does this reversibility affect reaction rates?

The short answer is that it's complicated. The longer answer is that it's simple to write down formal rate expressions but more difficult to solve them. To illustrate, let's work through the simple $A \rightleftharpoons B$ reaction already mentioned. If this is an elementary reaction, then the $A \rightarrow B$ direction proceeds according to the simple rate law

$$v_{\mathrm{f}} = k_{\mathrm{f}}[A], \tag{305}$$

just as if the reverse reaction weren't at play at all. In this expression, the subscript "f" indicates that this is for the *forward* reaction. Conversely, the rate of the reverse reaction $B \rightarrow$ proceeds according to the simple rate law

$$v_{\mathrm{r}} = k_{\mathrm{r}}[B], \tag{306}$$

where the subscript "r" indicates the reverse reaction. Now comes the complication: The rate at which $[A]$ changes with time depends on *both* the *consumption* of A by the forward reaction and the *production* of A by the reverse reaction. Mathematically:

$$\frac{d[A]}{dt} = v_{\mathrm{r}} - v_{\mathrm{f}} = k_{\mathrm{r}}[B] - k_{\mathrm{f}}[A]. \tag{307}$$

The minus sign here on the v_{f} term indicates that species A is *consumed* by the forward reaction. A similar equation holds for the concentration of B:

$$\frac{d[B]}{dt} = v_{\mathrm{f}} - v_{\mathrm{r}} = k_{\mathrm{f}}[A] - k_{\mathrm{r}}[B]. \tag{308}$$

Such *partial differential equations* are generally much more difficult to solve than the simple rate laws we encountered earlier. In many multi-species cases, we must resort to making approximations to obtain simplified results, such as the Michaelis–Menten equation of enzyme kinetics.

However, even in complex systems, rate equations like this can often provide simple insight into the *equilibrium* state of chemical reactions. Recall that chemical equilibrium is defined by the fact that, on average, the reaction has stopped proceeding in any direction. Thus the overall rate of change of concentration of *any* reactant or product must be zero. And this criterion alone is sufficient to solve a set of rate equations to obtain expressions for the equilibrium concentration of each species in the system.

For example, in our $A \rightleftharpoons B$ case, setting $\frac{d[A]}{dt} = 0$ gives

$$k_{\mathrm{r}}[B]_{\mathrm{eq}} - k_{\mathrm{f}}[A]_{\mathrm{eq}} = 0 \tag{309}$$

or after some rearranging:

$$\frac{[B]_{\mathrm{eq}}}{[A]_{\mathrm{eq}}} = \frac{k_{\mathrm{f}}}{k_{\mathrm{r}}}. \tag{310}$$

But the factor $\frac{[B]_{\text{eq}}}{[A]_{\text{eq}}}$ here should look familiar: this is exactly the *equilibrium constant* K_{eq} for the reaction $A \rightleftharpoons B$! For a one-to-one elementary chemical reaction, we thus obtain the simple relationship

$$K_{\text{eq}} = \frac{k_{\text{f}}}{k_{\text{r}}}. \tag{311}$$

This simple example illustrates a more general relationship between rate constants and equilibrium constants. In general, a complete set of *rate constants* for even a complex reaction network is sufficient to determine the corresponding equilibrium constants. This result rests on the fact that the equilibrium constant reflects the point at which the rates of the forward and reverse reactions are balanced: thus knowledge of both forward and backward reaction rates is sufficient to determine the equilibrium constant.

Note, however, that the converse is not generally true: Knowing the equilibrium constant does not generally allow us to determine rate constants. For example, if K_{eq} in Eq. (311) is close to one, it means that k_{f} and k_{r} are approximately equal. But this doesn't tell us anything about the *magnitude* of either constant: it could be that k_{f} and k_{r} are either both very large or both very small. Insight into the actual values of the rate constants determines kinetic measurements that directly follow reaction concentrations as a function of time.

27 Big Ideas in Statistical Mechanics and Kinetics

- The **goal** of statistical mechanics is to find a reliable way to predict **macroscopic** thermodynamic properties as *average values* over **microscopic** physical models.
- The **central problem** in statistical mechanics is to figure out *how to assign probabilities* p_μ *to the various microstates* μ *that a system might occupy.*
- Our approach to this problem is based on the **Maximum Entropy** hypothesis that we should always get the right values if we *assume as little information as possible* about what microstates the system "lives in".
  - Why do we do this? Because *experimentally* we find that thermodynamic states are always defined by *only a few macroscopic parameters*, e.g., N, V, and T. If only a few parameters are needed to define the system state, then we *shouldn't need information* on any other parameters to assign the probabilities. Assuming more information would *add* constraints that aren't there in reality!
  - How do we do this? We use the **Shannon information entropy** $s(\{p_\mu\})$ as a quantitative measure of how much uncertainty there is in the system. *Maximizing* s means *minimizing* the assumed information!
- The method of **Lagrange Multipliers** is the technical tool we need to maximize the uncertainty subject to the experimental constraints
  - **Key Idea**: maximizing a function $f(\boldsymbol{x})$ subject to constraints $c_1(\boldsymbol{x}) = 0$, ..., $c_K(\boldsymbol{x}) = 0$ is equivalent to maximizing the auxiliary function

    $$g(\boldsymbol{x}) = f(\boldsymbol{x}) + \lambda_1 c_1(\boldsymbol{x}) + ... + \lambda_K c_K(\boldsymbol{x}),$$

    where $\lambda_1, ..., \lambda_K$ are unknown constants that must be determined from the constraints.
  - **Key Idea**: We maximize the function $g(\boldsymbol{x})$ by requiring that

    $$\left(\frac{\partial g}{\partial x_i}\right)_{x_j \neq x_i} = 0,$$

    i.e., that taking a step in any direction doesn't change the value of $g(\boldsymbol{x})$.
- In statistical mechanics, we consider two "flavors" of **Thermodynamic Ensemble**:
  - The **Microcanonical Ensemble** has fixed N, V, and U. This corresponds to an *isolated system.*
    - We optimize over the probabilities p_μ for all microstates μ that have *exactly* the right values for N, V, and U, subject to the constraint

      $$\sum_\mu p_\mu = 1.$$

    - The outcome is that **the probabilities are equal** for all states that have the right energy, volume, and number of particles.
  - The **Canonical ensemble** has fixed N and V, and T. This corresponds to a *constant-volume, constant-temperature* system.

    - ∗ We optimize over the probabilities p_μ for all microstates μ that have *exactly* the right values for N and V subject to the constraints

      $$\sum_\mu p_\mu = 1$$

      and

      $$\sum_\mu \varepsilon_\mu p_\mu = U.$$

    - ∗ This leads to probabilities

      $$p_\mu = \frac{e^{-\beta\varepsilon_\mu}}{Z}, \tag{312}$$

      where

      $$Z \equiv \sum_\mu e^{-\beta\varepsilon_\mu} \tag{313}$$

      is the **partition function** and

      $$\beta = \frac{1}{k_\mathrm{B}T}. \tag{314}$$

    - ∗ The partition function Z is a *normalization constant* that ensures that the probabilities p_μ sum to one.
    - ∗ If we *shift the energy scale* so that the lowest-energy microstate has energy $\varepsilon_\mu = 0$, then the modified partition function $Z'(T)$ **counts the number of microstates accessible** to the system at a given temperature.

- The **protein folding problem asks**
  - How can we predict the three-dimensional structure of a protein from its sequence?
  - How do real proteins *locate* the "correct" folded structure from the astronomically large number of different possible conformations?

- **Lattice Proteins** (or lattice polymers) are simplified models that let us test and understand how different forces affect protein folding.

- Effects that tend to **drive protein folding forward** (i.e., favor folding) include
  - Formation of intra-protein hydrogen bonds in the protein backbone which lower the energy of the system
  - The increase in the entropy of the surrounding water due to the reduced surface area of the protein (the "hydrophobic effect")

- Effects that tend to **oppose protein folding** include
  - The decrease in entropy of the protein due to the smaller number of possible conformations in the folded state
  - The loss of favorable hydrogen-bonding interactions between polar protein residues and the solvent, which raises the energy of the system

- The specific molecular properties of individual amino acids in the peptide chain controls the final conformation of the protein. Examples include:

  - Alanine (Ala / A) is hydrophobic
  - Lysine (Lys / K) holds a positive charge
  - Glutamate (Glu / E) is negatively charged
  - Proline (Pro / P) is sterically constrained and likes to form “kinks” in the chain

- **Levinthal’s Paradox** notes that it would take proteins billions of years to fold if they sampled all the possible conformations available to them

- Two possible resolutions to Levinthal’s paradox include

  - **Folding Pathways**: The protein may be driven by favorable interactions to follow a *specific sequence of conformations* to quickly move from the unfolded state to the folded state
  - **Folding Funnel**: There may be *global driving forces* like “hydrophobic collapse” that keep the protein *close enough* to the native structure that it eventually locates the correct fold. (I.e., there may be many different “folding pathways”, but they all bring the structure close enough to the native conformation to find the correct fold quickly.)

- **Kinetics** is the study of *how fast a process happens* or *how quickly it reaches equilibrium.*

- The **rate** of a chemical reaction

  $$\nu_1 \mathrm{X}_1 + \nu_2 \mathrm{X}_2 + ... + \nu_{N_\mathrm{r}} \mathrm{X}_{N_\mathrm{r}} \longrightarrow \nu_{N_\mathrm{r}+1} \mathrm{X}_{N_\mathrm{r}+1} + \nu_{N_\mathrm{r}+2} \mathrm{X}_{N_\mathrm{r}+2} + ... + \nu_{N_\mathrm{r}+N_\mathrm{p}} \mathrm{X}_{N_\mathrm{r}+N_\mathrm{p}}$$

  is a measure of how quickly the reaction proceeds in the forward direction:

  $$v = \pm \frac{1}{\nu_i} \frac{d[X_i]}{dt},$$

  with the “$-$” sign applicable if species X_i is a reactant (indicating that it is consumed by the reaction) and the “+” sign applicable if species X_i is a product (indicating that it is produced by the reaction). Note that the rate is scaled by the stoichiometric coefficient ν_i so that the rate is independent of which reactant or product we choose to measure it.

- The **units** for a reaction rate v are $\mathrm{M/sec} = \mathrm{moles/(liter \cdot s)}$

- A **rate law** is a mathematical relationship

  $$v = f\left([X_1], [X_2], ..., [X_{N_\mathrm{r}+N_\mathrm{p}}]; T\right)$$

  between a reaction rate v and the concentrations $[X_1]$, $[X_2]$, ... of the various chemical species involved in the reaction.

- A reaction is said to be of n^{th} **order** in species X if the reaction rate scales with concentration as $[X]^n$.

- An **elementary reaction** is one that takes place in a single step at the molecular level. Rate laws for elementary reactions take the simple form

  $$v = k[X_1]^{\nu_1}[X_2]^{\nu_2}...[X_{N_\mathrm{r}}]^{\nu_{N_\mathrm{r}}}.$$

  The coefficient k is called the **rate constant**.

- The **units** on the rate constant k depend on the reaction order. A first-order rate constant has units of s^{-1}.

- The **overall order** of an elementary reaction is the sum of the stoichiometric coefficients of its reactants $\nu_1 + \nu_2 + ... + \nu_{N_\mathrm{r}}$.
- **Rate constants** are denoted by lower-case k's, while **equilibrium constants** are denoted by upper-case K's.
- A **First Order** elementary reaction (e.g., $A \to B$) has a rate law of the form

  $$v = -\frac{d[A]}{dt} = k[A],$$

  which can be integrated to give

  $$[A](t) = [A]_0 e^{-kt}.$$

  In the absence of the reverse reaction, the reactant concentration decays **exponentially**.
- The rate of **higher order** chemical reactions depend *more sensitively* on reactant concentrations than do first-order reactions. (The rate slows down more rapidly as the reactants are consumed.)
- Real reactions are, to some extent, **reversible**. A reaction is at **equilibrium** when the rates of the forward and reverse reactions are equal.
- In a first-order reversible reaction $A \rightleftharpoons B$, setting the overall rate of A (or B) consumption to zero gives

  $$K_\mathrm{eq} = \frac{k_\mathrm{f}}{k_\mathrm{r}},$$

  where k_f and k_r are the respective rates for the forward and reverse reactions.
- More generally, equilibrium constants K_eq can be calculated if *all rate constants* are known.
- In contrast, rate constants *cannot* be calculated from equilibrium constants alone.
- **Brownian motion** is the random motion that microscopic particles go through in a fluid. At the macroscopic scale, Brownian motion leads to **diffusion**.
- The **random walk** is a simple probabilistic model for Brownian motion.
  - An **unbiased** random walk proceeds with equal probability in any direction.
  - A **biased** random walk prefers to move in one direction relative to others.
- Brownian motion provides a good model for the way many real reactions move along their **reaction coordinate**, an "abstract one-dimensional coordinate which represents progress along a reaction pathway." (`https://en.wikipedia.org/wiki/Reaction_coordinate`).
- The **Arrhenius Equation**

  $$k = Ae^{-\frac{E_a}{RT}}$$

  is an empirical relationship that describes the dependence of many rate constants k on temperature.
- The Arrhenius prefactor A indicates how quickly a reaction can "diffuse" or "randomly walk" along the reaction coordinate.
- The **activation energy** E_a in the Arrhenius equation represents an energy barrier the reaction must overcome to proceed to completion.

28 The Canonical Ideal Gas

28.1 Introduction

With the statistical tools we've developed, we're now ready to give a more complete account of the thermodynamic properties of ideal gases. Early in the course, we derived an expression for the energy of a monomeric ideal gas, and later related that energy law to the temperature. In this lecture, we'll repeat the derivation of the energy expression using our new statistical tools and then extend the treatment to consider a *diatomic* ideal gases like O_2 and N_2.

28.2 Monomeric Ideal Gas

Let's begin with a box of N monotomic ideal gas atoms contained inside a sealed box at fixed temperature, i.e., average energy. The total Hamiltonian $\mathcal{H}$ for the N gas particles is

$$\mathcal{H}\left(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{v}^{(x)},\boldsymbol{v}^{(y)},\boldsymbol{v}^{(z)}\right)=\sum_{i=1}^{N}\frac{m\left[\left(v_i^{(x)}\right)^2+\left(v_i^{(y)}\right)^2+\left(v_i^{(z)}\right)^2\right]}{2}, \tag{315}$$

where m is the molecular weight of the gas, x_i, y_i, and z_i are the three Cartesian coordinates of the i^{th} atom, and $v_i^{(x)}$, $v_i^{(y)}$, and $v_i^{(z)}$ are the corresponding velocities. Since the coordinates and velocities are *continuous* functions, the partition function is calculated as an integral

$$Z=\int_V d\boldsymbol{x}\int_{-\infty}^{\infty}dv_1^{(x)}...\int_{-\infty}^{\infty}dv_N^{(z)}e^{-\beta\mathcal{H}}, \tag{316}$$

where the integral $d\boldsymbol{x}$ indicates an integral of the spatial coordinates of each particle over the entire volume of the box:

$$\int_V d\boldsymbol{x}=V^N. \tag{317}$$

Since the exponent can be factored as

$$e^{-\beta\mathcal{H}}=\prod_{i=1}^{N}e^{-\frac{\beta m\left(v_i^{(x)}\right)^2}{2}}e^{-\frac{\beta m\left(v_i^{(y)}\right)^2}{2}}e^{-\frac{\beta m\left(v_i^{(z)}\right)^2}{2}}, \tag{318}$$

the remaining integrals can be calculated as

$$\int_{-\infty}^{\infty}dv\,e^{-\frac{\beta m v^2}{2}}=\sqrt{\frac{2\pi}{\beta m}}. \tag{319}$$

Here we've used the Gaussian integral identity:

Gaussian Integral

$$\int_{-\infty}^{\infty}dxe^{-\frac{x^2}{2\sigma^2}}=\sqrt{2\pi\sigma^2}, \tag{320}$$

by taking $\sigma^2=\frac{1}{\beta m}$. Since there are three of these integrals for each particle, the velocity integrals in total produce a factor of $\left(\frac{2\pi}{\beta m}\right)^{\frac{3N}{2}}$. The partition function is then

$$Z=V^N\left(\frac{2\pi}{\beta m}\right)^{\frac{3N}{2}}. \tag{321}$$

Now, in the canonical ensemble, the thermodynamic energy U must be equal to the average ensemble energy. Eq. (276) thus tells us that

$$\begin{aligned} U &= -\frac{\partial \ln Z}{\partial \beta} && (322) \\ &= -\frac{\partial}{\partial \beta} \ln \left[V^N \left(\frac{2\pi}{\beta m} \right)^{\frac{3N}{2}} \right] && (323) \\ &= -\frac{\partial}{\partial \beta} \left(N \ln V + \frac{3N}{2} \ln \frac{2\pi}{\beta m} \right) && (324) \\ &= \frac{3N}{2} \frac{\partial}{\partial \beta} \left(\ln \frac{\beta m}{2\pi} \right) && (325) \\ &= \frac{3N}{2} \frac{m}{2\pi} \frac{2\pi}{\beta m} && (326) \\ &= \frac{3N}{2\beta}. && (327) \end{aligned}$$

Recalling the relations

$$\begin{aligned} \beta &= \frac{1}{k_{\mathrm{B}} T} && (328) \\ R &= N_{\mathrm{A}} k_{\mathrm{B}} && (329) \\ n &= \frac{N}{N_{\mathrm{A}}}, && (330) \end{aligned}$$

we see that this is exactly our original ideal gas energy formula [Eq. (42)]:

$$U = \frac{3N}{2\beta} = \frac{3N k_{\mathrm{B}} T}{2} = \frac{3 \frac{N}{N_{\mathrm{A}}} (N_{\mathrm{A}} k_{\mathrm{B}}) T}{2} = \frac{3}{2} nRT. \tag{331}$$

So for the monatomic ideal gas, the maximum entropy approach gives exactly the same result as we obtained earlier by directly integrating Newton's equations. (Good news!)

28.3 Diatomic Ideal Gas

With our new statistical tools, however, we can readily solve problems that would be extremely difficult if we started from Newton's equations. As a useful example, let's see what happens if we repeat our calculation for a *diatomic* ideal gas, a good model for atmospheric gases like O_2 and N_2. To a good approximation, the Hamiltonian for a diatomic gas can be broken up into three separate terms

$$\mathcal{H} = \mathcal{H}_{\mathrm{trans}} + \mathcal{H}_{\mathrm{vib}} + \mathcal{H}_{\mathrm{rot}} \tag{332}$$

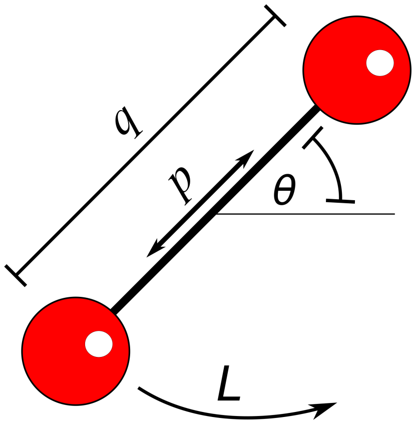
corresponding to the *translational* (trans) motion of the molecules through space, the *vibrational* (vib) motion of the bond length between the two atoms, and the *rotational* (rot) motion of the molecule around its center of mass. Mathematically, $\mathcal{H}_{\mathrm{trans}}$ is exactly the monomeric ideal gas Hamiltonian of Eq. (315); the only update is that m represents the mass of the whole molecule rather than that of a single atom. The vibrational Hamiltonian

$$\mathcal{H}_{\mathrm{vib}} = \sum_{i=1}^{N} \frac{\omega^2 q_i^2 + p_i^2}{2} \tag{333}$$

describes the expansion and compression of the bond length q_i between the two atoms of the i^{th} gas molecule, while the rotational Hamiltonian

$$\mathcal{H}_{\text{rot}} = \sum_{i=1}^{N} \frac{L_{A,i}^2}{2I} + \sum_{i=1}^{N} \frac{L_{B,i}^2}{2I}$$

accounts for the rotational energy associated with the angular momenta $L_{A,i}$ and $L_{B,i}$ of the i^{th} molecule rotating around its center of mass. Note that a diatomic molecule can rotate in *two distinct ways*, which is why there are two terms $L_{A,i}$ and $L_{B,i}$; For example, in the figure, the molecule is shown rotating in the plane of the page, but it might also rotate *out* of the plane of the page as well. (The quantity I is called the *moment of inertia* of the molecule, and depends on the molecular mass and bond length.)



You don't need to memorize the details of each of these terms. What you *should* remember is that

Key Points in Diatomic Ideal Gas Calculation

- There are three *independent* contributions to the diatomic ideal gas Hamiltonian, corresponding to *translational*, *vibrational*, and *rotational* motion.
- Each contribution to the Hamiltonian depends *quadratically* on the relevant coordinates and/or momenta. (Explicitly, $\mathcal{H}_{\text{trans}}$ is proportional to the sum of $\left(v_i^{(x)}\right)^2$, $\left(v_i^{(y)}\right)^2$, and $\left(v_i^{(z)}\right)^2$, $\mathcal{H}_{\text{vib}}$ is proportional to the sum of q_i^2 and p_i^2, and $\mathcal{H}_{rot}$ is proportional to the sum of $L_{i,A}^2$ and $L_{i,B}^2$.)
- In total, there are *seven* quadratic terms in the Hamiltonian.

This last two observations are important because they allow us to find the partition function Z for the diatomic gas almost without doing any new calculations at all. Because the translational, vibrational, and rotational contributions to the Hamiltonian are all independent, the total partition function can be factored as

$$Z_{\text{diatom}} = \int d\Gamma_{\text{trans}} \int d\Gamma_{\text{vib}} \int d\Gamma_{\text{rot}} e^{-\beta(\mathcal{H}_{\text{trans}}+\mathcal{H}_{\text{vib}}+\mathcal{H}_{\text{rot}})} \tag{334}$$

$$= \left(\int d\Gamma_{\text{trans}} e^{-\beta\mathcal{H}_{\text{trans}}}\right)\left(\int d\Gamma_{\text{vib}} e^{-\beta\mathcal{H}_{\text{vib}}}\right)\left(\int d\Gamma_{\text{rot}} e^{-\beta\mathcal{H}_{\text{rot}}}\right) \tag{335}$$

$$= Z_{\text{trans}} Z_{\text{vib}} Z_{\text{rot}}. \tag{336}$$

Here the notation $\int d\Gamma_{\text{trans}}$ is a short-hand for integrating over all the spatial coordinates and momenta of the gas particles, $\int d\Gamma_{\text{vib}}$ likewise indicates integration over the bond-stretch coordinates q_i and momenta p_i, and $\int d\Gamma_{\text{rot}}$ is an integral over the rotational angular momenta L_i and the corresponding *rotation angles* θ_i. On the second line, we've regrouped the exponents in each term into *separate* integrals, each of which we identify in the third line as a separate partition function Z_{trans}, Z_{vib}, or Z_{rot}. This regrouping is only possible because the translationa, vibrational, and rotational degrees of freedom are *independent* in our model.

From our earlier calculation, we already know that

$$Z_{\text{trans}} = V^N \left(\frac{2\pi}{\beta m}\right)^{\frac{3N}{2}} \tag{337}$$

since the translational part of the Hamiltonian is exactly the same as for the monatomic case (except that the mass is now a molecular mass instead of atomic mass). But we can in fact use this same result to write down almost immediately the remaining partition functions Z_{vib} and Z_{rot}. Looking back over our earlier derivation, we see that each Gaussian integral contributes a factor of $\sqrt{2\pi\sigma^2}$, where σ^2 is whatever factor is needed to put the corresponding term of the Hamiltonian in the form $\frac{x^2}{2\sigma^2}$. In particular, we can write

$$\sigma_q^2 = \frac{1}{\beta\omega^2} \tag{338}$$

$$\sigma_p^2 = \frac{1}{\beta} \tag{339}$$

$$\sigma_{L,A}^2 = \sigma_{L,B}^2 = \frac{I}{\beta}. \tag{340}$$

Since there are N such Gaussian integrals for each degree of freedom, the complete partition funtions are

$$Z_{\text{vib}} = \left(\frac{2\pi}{\beta^2\omega^2}\right)^{\frac{N}{2}} \tag{341}$$

$$Z_{\text{rot}} = \left(\frac{2\pi I}{\beta}\right)^{N}. \tag{342}$$

Now notice that the *natural log* of the partition function can be split up as

$$\begin{aligned}
\ln Z &= \ln\left(Z_{\text{trans}} Z_{\text{vib}} Z_{\text{rot}}\right) && (343)\\
&= \ln Z_{\text{trans}} + \ln Z_{\text{vib}} + \ln Z_{\text{rot}} && (344)\\
&= \ln\left(V^N \left(\frac{2\pi}{m}\right)^{\frac{3N}{2}}\right) - \ln\beta^{\frac{3N}{2}} + \ln\left(\frac{2\pi}{\omega^2}\right)^{\frac{N}{2}} - \ln\beta^N + \ln\left(\frac{2\pi I}{\beta}\right)^N - \ln\beta^N. && (345)
\end{aligned}$$

Here in each term we have factored out all the factors involving the inverse temperature $\beta^{\frac{N}{2}}$ – one for each quadratic term that appears in the Hamiltonian.

The reason this form is useful is because the energy U can now be calculated easily by dropping all terms independent of β:

$$\begin{aligned}
U = -\frac{\partial \ln Z}{\partial\beta} &= -\frac{\partial}{\partial\beta}\left(-\frac{3N}{2}\ln\beta - N\ln\beta - \frac{N}{2}\ln\beta\right) && (346)\\
&= \frac{7N}{2}\frac{d\ln\beta}{d\beta} = \frac{7N}{2\beta} = \frac{7Nk_{\text{B}}T}{2} = \frac{7}{2}nRT. && (347)
\end{aligned}$$

28.4 The Equipartition Principle

In addition to giving us a useful formula for the energy of a diatomic ideal gas, this formula illustrates a very general principle in classical mechanics: that energy is *partitioned equally* among different degrees of freedom in equilibrium systems. Formally, this result is known as the *equipartition principle*:

The Equipartition Principle: At equilibrium, energy is partitioned equally among all quadratic degrees of freedom in a system. Each quadratic degree of freedom "stores" an energy of $\frac{k_\mathrm{B}T}{2}$ at equilibrium.

The equipartition principle gives us a simple explanation for why some substances have larger specific heats than others: substances with many degrees of freedom have more room (i.e., degrees of freedom) in which to "store" energy, so it takes more energy to bring about the same change in temperature.

29 Something Rotten in Denmark: The Fall of Classical Mechanics

29.1 Heat Capacities of Real Gases

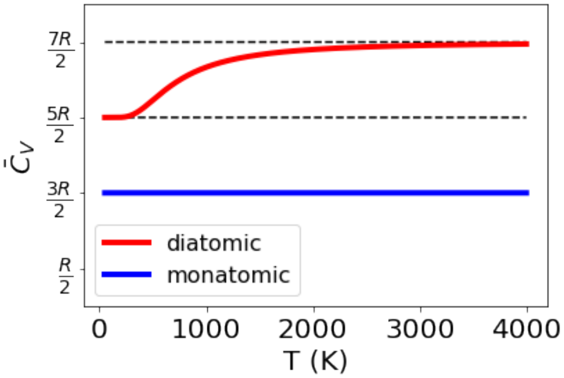
The last section closes our study of classical statistical mechanics. Conveniently, it also *opens* our study of quantum mechanics: The first hint that something was missing from classical Newtonian mechanics was that *it failed – even qualitatively – to accurately describe the heat capacities of real diatomic gases.* While the formulas worked perfectly well for *monatomic* gases like Ar, Kr, and Xe, strange behavior was observed experimentally for diatomic gases like O_2, N_2, and H_2.

To see the problem, note that the formulas we derived in the last section for the *energy* of ideal gases can also be used to calculate the heat capacities. In particular, a very brief calculation gives:

$$C_V \equiv \left.\frac{dQ}{dT}\right|_V = \left(\frac{\partial U}{\partial T}\right)_V = \begin{cases} \frac{3}{2}nR, & \text{monatomic} \\ \frac{7}{2}nR, & \text{diatomic} \end{cases} \tag{348}$$

Thus, according to our calculations, the heat capacities of both types of gas should be independent of temperature, with the value of the diatomic heat capacity greater than the monatomic capacity by a term $\frac{4nR}{2} = 2nR$.

Unfortunately, when the heat capacity of *real* gases are measured experimentally, the results look very different. The figure sketches the behavior of the specific heat capacities for real monatomic (blue line) and diatomic (red line) gases. The monatomic gas behaves exactly as expected: $\bar{C}_V$ is independent of T, with a constant value of $\frac{3R}{2}$. The specific heat capacity for the *diatomic* gase approaches the expected value of $\frac{7R}{2}$ at (very!) high temperatures (thousands of Kelvin), but is much closer to $\frac{5R}{2}$ at low temperatures.



Such findings were deeply worrying to the founders of statistical mechanics, particularly to James Clerk Maxwell who originally derived the classical results. At first, scientists worried that the entire atomistic picture of the world that statistical mechanics builds on might be in error. Soon, however, evidence begin to pile up that the *statistical* (and therefore atomistic) part of the theory was just fine; it was the *mechanics* (based on Newton's laws) that was failing.

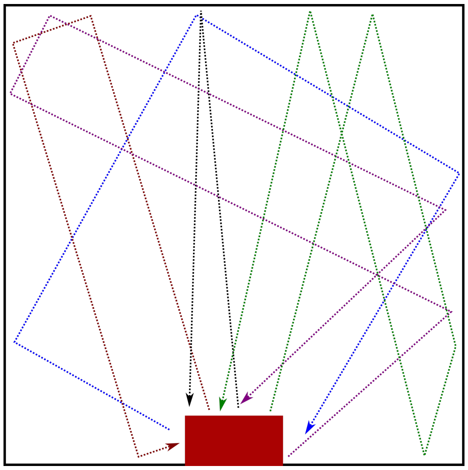
29.2 Blackbody Radiation

The most famous – and arguably the definitive – experimental finding that struck down Newtonian physics was that experimental data on *blackbody radiation* was in direct contradiction to classical theory.

Blackbody radiation refers the electromagnetic energy emitted by a material system that is *at equilibrium at a fixed temperature T.*

There are several ideas here that may be more or less unfamiliar, so let's unpack them carefully. *Electromagnetic radiation* is the energy carried by light waves.[11] *Emission* is the process of light radiating away from an object, carrying energy with it; for example, an incandescent light bulb *emits* radiant energy from a tungsten filament that is, in turn, heated by the electricity passing through it. More subtle – but critical for our discussion – is that the requirement that the emission occur *at equilibrium at temperature T* implies that the material is also *absorbing* light (i.e., accepting radiant energy) at the same rate that the energy is being emitted.

A simple way to think about this is to imagine an object inside of a box with mirror-coated walls, as pictured in the figure below. The object's thermal energy causes it to emit radiation (depicted as lines emanating outward from the object). If this were the only process taking place, the temperature of the object would gradually drop as it loses energy via emission. But because the walls are mirror-coated, the emitted light eventually "bounces" its way back to the object where it is absorbed again. In this way, the temperature stays fixed at the equilibrium value T.



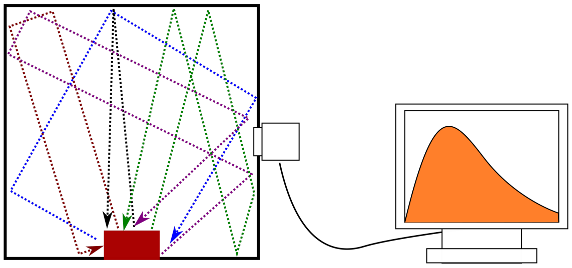
The question of interest to scientists around the turn of the 20th century was: What is the *spectrum* of blackbody radiation from an object at temperature T?

The **spectrum** $\rho(\nu)$ of a radiation source is a curve that plots *how much electromagnetic energy is emitted at each frequency* ν of light.

In terms of our "mirrored box" example, measuring the blackbody spectrum corresponds to poking a small hole in the side of the box, putting up a filter so that only light of a single frequency can come out at a time, and then measuring (for each frequency filter) how much radiant energy is emitted. In a modern-day experiment, the set-up might look something like what's shown in the figure below.

---

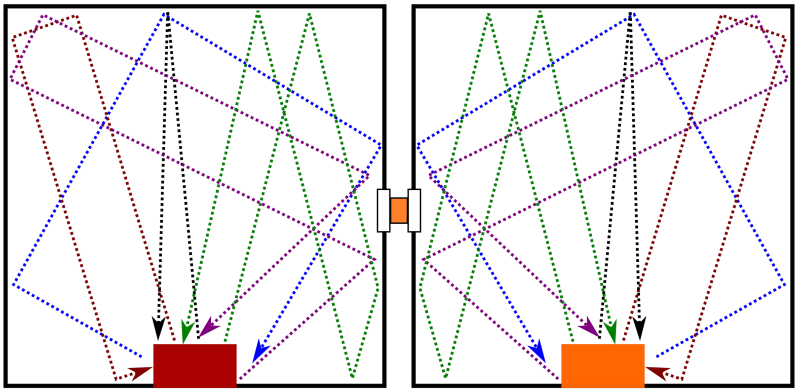
[11] Technically, "light" is usually reserved for radiation in the visible range of the spectrum, but the physics is identical for all electromagnetic radiation, from the microwave to the extreme ultraviolet.



29.3 Kirchhoff: Everything emits the same at equilibrium

Why were scientists so interested in this question? Well, partly because the process of absorption and emission of radiation wasn't very well understood yet. And partly because the blackbody radiation spectrum turns out to have a surprisingly general form. In fact, the physicist Gustav Kirchhoff demonstrated theoretically that *all objects in thermal equilibrium at a given temperature T must give rise to the same blackbody radiation spectrum.*

How can this be true? Kirchhoff arrived at this conclusion through a simple though experiment, a "proof by contradiction" that invokes our old friend, the Second Law of Thermodynamics. Imagine two mirrored boxes, each containing an object at thermal equilibrium at temperature T. Suppose that the blackbody radiation spectrum of the two objects are *not* identical. Specifically, let's suppose that the object on the right emits more orange-colored light. If we then poke a small hole in each box, place a small optical filter over the whole that allows passage only of orange light, and align the two holes so that orange light can pass between the two boxes, what will happen? Because the box on the right emits more orange light, there will be a *net transfer of energy* into the box on the left. Thus energy will flow *spontaneously* from one object at temperature T to another object at the same temperature! This flow of energy could, in turn, be used to drive a heat engine that would produce work without any net input of heat – a violation of the Second Law!

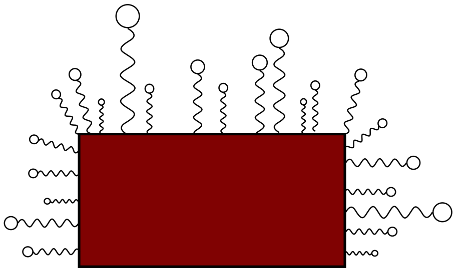


The only way to avoid this problem – and thus preserve the Second Law of Thermodynamics – is if *all objects exhibit the same blackbody spectrum at equilibrium at a given temperature.*

29.4 Theoreticians Everywhere: Let's pretend everything is harmonic!

Kirchhoff's proof was of great interest to physicists since it meant that if you could solve the blackbody radiation problem for *any* material system, you had solved it for *all* material systems. In fact, the material system doesn't even need to really exist: so long as you can solve the problem for some hypothetical material, the results should hold for real systems too.

As a general rule in physics, if one has a choice of what system to study, the easiest is almost *always* the harmonic oscillator. For that reason, theoretical physicists of all stripes began to imagine materials made *entirely* of harmonic oscillators. The exact details of what elements or chemicals could compose such systems didn't matter; what mattered is that one could *imagine* an object made up entirely of tiny harmonic oscillators (i.e., springs), oscillators with every possible oscillation frequency imaginable, from zero to infinity. With such a hypothetical material, the blackbody radiation problem could be solved!



And solved it was. Using classical statistical mechanics (in particular the Equipartition Theorem illustrated in the last lecture), along with basic electrodynamics, it was quickly shown that the blackbody spectrum for harmonic systems – and hence for *all* systems must have the form

$$\rho(\nu) = 2k_\mathrm{B}T\frac{\nu^2}{c^2}, \tag{349}$$

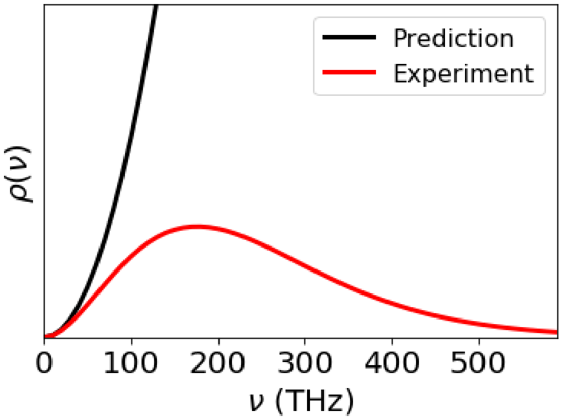
where $c \approx 2.9979 \cdot 10^8$ m/s is the speed of light.

29.5 Experimentalists Everywhere: It doesn't work!

The formula was a beautiful piece of theory work. Unfortunately, it suffered from two problems:

- It doesn't describe the blackbody spectrum of real materials, and
- If it did, we'd all be dead.

The first problem is illustrated in the figure below. Empirically, when the blackbody spectrum of a real system was measured, it was found to resemble the red curve in the figure. As a function of frequency ν, the spectrum first increases, then peaks (in this case in the infrared region near 200 THz) and begins to fall back toward zero at very high frequencies. The black curve is the theoretical prediction of classical mechanics: that the radiant energy should increase continuously with frequency (proportional to ν^2) without limit!



This predicted infinite increase in blackbody radiation at high frequencies became known as the *ultraviolate catastrophe*, and with good reason: If blackbody emitters actually emitted radiation this way, life on earth would be impossible! The sun, for example would constantly emit such a high intensity of X-rays (along with even higher-energy radiation) that living creatures would be scorched almost instantaneously.

At the turn of the 20th century, this failure seemed inexplicable. The arguments leading to came from the very bedrock of what was then known of physics: Newton's, equations, Maxwell's equations, and the Laws of Thermodynamics. Clearly *something* had to be wrong. But what?

29.6 Planck: It works if you quantize?

The first clue toward the answer came from a German physicist named Max Planck. Purely as a mathematical exercise, Planck pointed out that the correct form for the blackbody spectrum was obtained if one assumed that the energy states of the harmonic oscillators were *quantized*, i.e., that they could take on *only* those energy values

$$\varepsilon_n = nh\nu, \tag{350}$$

where ν was the frequency at which the oscillator vibrates, n was any non-negative integer (i.e., $n = 0, 1, 2, ...$), and h was a mathematical constant that now bears Planck's name:

Planck's Constant h has the numerical value

$$h \approx 6.626 \cdot 10^{-34}\,\mathrm{J\,s}. \tag{351}$$

Planck himself believed this result to be a coincidence; he spent much of the rest of his career searching for an explanation for why such quantization might occur or what alternative explanation could give the same result. To his surprise (and somewhat to his chagrin) he had started a revolution.

30 Quantization and the Breakdown of Equipartition

30.1 Equipartition and Blackbody Radiation

In the last lecture, we stated that Planck's *ad hoc* idea of "quantizing" harmonic oscillator energy levels avoids the ultraviolet catastrophe and results in a blackbody radiation spectrum in close agreement with experiment. The key to understanding *why* Planck's approach avoids disaster has to do with the equipartition principle we noted earlier in reference to the energy of the diatomic ideal gas: namely, that all quadratic degrees of freedom in a system should, at equilibrium, acquire an average energy of $\frac{k_{\mathrm{B}}T}{2}$.

To see how this works, let's pick apart the blackbody radiation formula [Eq. (349)] in more detail. First, let's rewrite the equation as

$$\rho(\nu) = k_{\mathrm{B}}T \cdot \frac{2\nu^2}{c^2}. \tag{352}$$

In this form, the factor of $k_{\mathrm{B}}T$ in the equation corresponds to the *average energy stored in a harmonic oscillator of frequency ν at equilibrium.* Recall that, according to the classical equipartition principle, all qudratic degrees of freedom should store an energy of $\frac{k_{\mathrm{B}}T}{2}$ at equilibrium. Since a harmonic oscillator has two quadratic coordinates in its Hamiltonian (the bond distance q and the momentum p), the total energy stored in each oscillator of frequency ν is just twice that value or

$$\langle E_\nu \rangle = k_{\mathrm{B}}T. \tag{353}$$

The second term in in Eq. (352), the factor of $\frac{2\nu^2}{c^2}$ comes from electrodynamics (i.e., the study of electromagnetic fields) and dictates how quickly radiation is emitted from a harmonic oscillator *per unit energy stored* in the oscillator. It is this term that produces the ν^2 divergence of the blackbody spectrum $\rho(\nu)$ as ν increases toward infinity.

30.2 Averting Disaster

The reason Planck's quantization "trick" avoids the ultraviolet catastrophe is that it replaces the equipartition principle with a new result, based on the assumption that the oscillator states are "quantized", i.e., allowed to exist only at specific energy levels, namely those given by Eq. (350).

To see how this helps, let's derive an expression for the average energy $\langle E_\nu \rangle$ for the energy of an oscillator with frequency ν that is allowed only to take quantized energy levels. Since we want to describe a blackbody object at a fixed temperature T, we work in the canonical ensemble. The probability of the n^{th} oscillator state is then

$$p_n = \frac{e^{-\beta n h\nu}}{Z}, \tag{354}$$

where Z is the partition function

$$Z = \sum_{n=0}^{\infty} e^{-\beta n h\nu}. \tag{355}$$

This complicated-looking infinite sum is surprisingly easy to evaluate. First note that it can be rewritten as

$$Z = \sum_{n=0}^{\infty} \left(e^{-\beta h\nu}\right)^n. \tag{356}$$

Then use the identity

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + ... = \frac{1}{1-x} \tag{357}$$

to obtain

$$Z = \frac{1}{1 - e^{-\beta h\nu}}. \tag{358}$$

Finally, use Eq. (277) to calculate the energy stored in each mode as

$$\begin{aligned}
\langle E_\nu \rangle = -\frac{\partial \ln Z}{\partial \beta} &= \frac{\partial}{\partial \beta} \ln\left(1 - e^{-\beta h\nu}\right) && (359)\\
&= \frac{h\nu e^{-\beta h\nu}}{1 - e^{-\beta h\nu}} && (360)\\
&= \frac{h\nu}{e^{\beta h\nu} - 1}. && (361)
\end{aligned}$$

Replacing the factor of $k_\mathrm{B}T$ in Eq. (352) with this new expression, we obtain

$$\rho(\nu) = \frac{h\nu e^{-\beta h\nu}}{1 - e^{-\beta h\nu}} \cdot \frac{2\nu^2}{c^2} = \frac{2h\nu^3}{\left(e^{\beta h\nu} - 1\right) c^2} \tag{362}$$

When ν is very small, the denominator here is approximately $\beta h\nu$, and the expression matches the original (classical) formula. But as ν increases, the exponent $e^{\beta h\nu}$ increases toward infinity *faster* than the factor of ν^3 in the numerator. Thus the radiation spectrum ρ initially increases as ν^2 for very small frequencies, but as $\nu \to \infty$, it drops back down toward zero, reproducing exactly the experimentally observed spectrum.

30.3 Energy-level Populations

To understand *why* quantization solves the blackbody problem, it may help to look at some examples involving the populations p_n of various harmonic oscillator energy levels. For this purpose, it's convenient to rewrite [Eq. (354)] as

$$p_n = \frac{e^{-\frac{n\bar{\nu}}{\frac{k_\mathrm{B}}{hc}T}}}{Z}, \tag{363}$$

where

$$\bar{\nu} = \frac{\nu}{c} \tag{364}$$

is the frequency in units of "wavenumbers" (cm^{-1}). Here c is the speed of light

$$c = 3.0 \cdot 10^{10} \text{ cm/s}, \tag{365}$$

and the quantity $\frac{k_\mathrm{B}}{hc}$ appearing in the p_n exponent has a value

$$\frac{k_\mathrm{B}}{hc} = \frac{1.381 \cdot 10^{-23}\mathrm{J/K}}{(6.626 \cdot 10^{-34}\mathrm{J \cdot s}) \cdot (3.0 \cdot 10^{10}\mathrm{cm/s})} \approx 0.7 \text{ cm}^{-1}/\mathrm{K}. \tag{366}$$

Thus the population of the n^{th} harmonic oscillator state can be written as

$$p_n \approx \frac{e^{-\frac{n\bar{\nu}/\mathrm{cm}^{-1}}{0.7 \cdot T/\mathrm{K}}}}{Z}. \tag{367}$$

The reason for all this unit manipulation is that it eliminates the awkwardness of working with miniscule quantities like 10^{-34}. For real molecules, vibrational frequencies typically span a range from around 100 cm^{-1} to around 3700 cm^{-1}. Since in most real-world applications, temperatures likewise range from around 200 K (the temperature of dry ice) to 6000 K (the temperature of the sun), the numerator and denominator of the quantity

$$\frac{n\bar{\nu}/\text{cm}^{-1}}{0.7\cdot T/\text{K}}. \tag{368}$$

take on a similar range of values, making it easy to do numerical calculations.

As a simple example, C=O stretch vibrations typically have a frequency of around $\bar{\nu} = 1600\ \text{cm}^{-1}$. What is the relative population $\frac{p_1}{p_0}$ of the ground ($n = 0$) and first excited state ($n = 1$) at room temperature (roughly $T = 300$ K)? From Eq. (367), we find

$$\frac{p_1}{p_0} \approx \frac{e^{-1\cdot\frac{\bar{\nu}/\text{cm}^{-1}}{0.7\cdot T/\text{K}}}}{e^{-\frac{0\cdot\bar{\nu}/\text{cm}^{-1}}{0.7\cdot T/\text{K}}}} = e^{-\frac{1600}{0.7\cdot 300}} = e^{-\frac{1600}{210}} \approx 5\cdot 10^{-4}. \tag{369}$$

On the other hand, large scale deformation modes of proteins often have vibrational frequencies as low as 50 cm^{-1}. The relative population of the first vibrational excited state is dramatically higher

$$\frac{p_1}{p_0} \approx e^{-\frac{50}{210}} = 0.788. \tag{370}$$

These simple calculations reflect a general result of quantization: the thermal excited-state populations for *high-frequency* vibrations are (at finite temperatures) much smaller than the corresponding populations in *low-frequency* vibrations. Because of this "suppression" of thermal population in high-frequency vibrations, the average energy stored in high-frequency coordinates is much smaller at room temperature than the $\frac{k_\text{B}T}{2}$ that would be expected from the classical equipartition principle. In the blackbody problem, this suppression of high-frequency populations is indeed so rapid that it overcomes the quadratic increase $\frac{2\nu^2}{c^2}$ of the emission rate, causing the radiation spectrum $\rho(\nu)$ to tail off smoothly to zero at high frequency – exactly as is observed experimentally.

30.4 Ideal Gas Heat Capacities

A very similar calculation explains the anomalous heat capacity values calculated for diatomic ideal gases. Because the translational, rotational, and vibrational degrees of freedom in the diatomic gas are (approximately) independent of each other in the Hamiltonian, the total heat capacity can be written as a sum of three separate contributions:

$$\bar{C}_\text{V} = \bar{C}_\text{V}^{(\text{trans})} + \bar{C}_\text{V}^{(\text{rot})} + \bar{C}_\text{V}^{(\text{vib})}. \tag{371}$$

The classical calculation for the first two of these terms turns out to work just fine at room temperature. The problem is with the third term, which describes the heat capacity associated with the harmonic vibration of the diatomic molecular bond.

However, using the expression we just derived for the energy stored in a harmonic oscillator, we can easily calculate the correct "quantized" heat capacity associated with bond vibration. Since

$$C_\text{V} \equiv \left.\frac{dU}{dT}\right|_\text{V}, \tag{372}$$

the contribution $\bar{C}_{\mathrm{V}}^{(\mathrm{vib})}$ to the heat capacity from harmonic vibrational motion is just

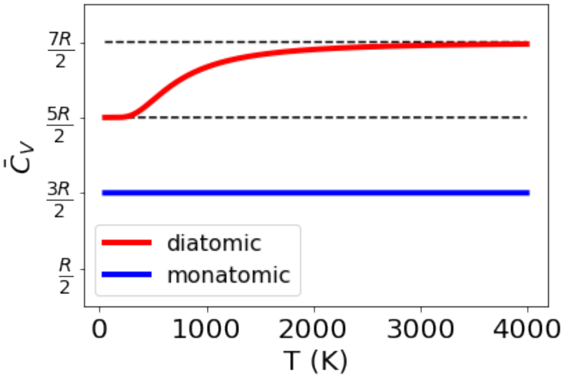
$$C_{\mathrm{V}}^{(\mathrm{vib})} = N\frac{d}{dT}\frac{h\nu}{e^{\frac{h\nu}{k_{\mathrm{B}}T}}-1} = -N\frac{h\nu}{k_{\mathrm{B}}T^2}e^{\frac{h\nu}{k_{\mathrm{B}}T}}\frac{h\nu}{\left(e^{\frac{h\nu}{k_{\mathrm{B}}T}}-1\right)^2} \tag{373}$$

$$= Nk_{\mathrm{B}}\frac{\left(\frac{h\nu}{k_{\mathrm{B}}T}\right)^2}{\left(e^{\frac{h\nu}{2k_{\mathrm{B}}T}}-e^{-\frac{h\nu}{2k_{\mathrm{B}}T}}\right)^2}. \tag{374}$$

This expression looks complicated, but it becomes simple in two different limits:

- When $T \to 0$, the denominator in this expression approaches infinity. As a result, $C_{\mathrm{V}}^{(\mathrm{vib})} \to 0$.
- When $T \to \infty$, the denominator approaches $\left(\frac{h\nu}{k_{\mathrm{B}}T}\right)^2$, i.e., the same value as the numerator. Thus, $C_{\mathrm{V}}^{(\mathrm{vib})} \to Nk_{\mathrm{B}} = nR$.

In light of these predictions, the observed behavior for the diatomic ideal gas begins to make sense. The translational and rotational degrees of freedom together contribute a $\frac{5nRT}{2}$ to the energy, corresponding to the value of $\bar{C}_{\mathrm{V}} = \frac{5R}{2}$ measured experimentally at room temperature. At very high temperatures (roughly when $0.7 \cdot T/\mathrm{K}$ approaches $\bar{\nu}/\mathrm{cm}^{-1}$), the vibrational degrees of freedom gradually become accessible and the heat capacity approaches the classical limit $\bar{C}_{\mathrm{V}} = \frac{5R}{2} + R = \frac{7R}{2}$.



31 Equations Gone Wild

31.1 Introduction

One of the things that makes quantum mechanics hard to wrap our minds around is that the equations describing it operate (pun intended) on very different terms than more familiar relations like Newton's equations or Beer's Law. In this section, we'll take a tour through some of these strange relations and describe what it is that makes them so strange – and so useful.

31.2 Plank's Quantization of the Harmonic Oscillator

Arguably the very first quantum-mechanical equation to be written down was Max Planck's relation

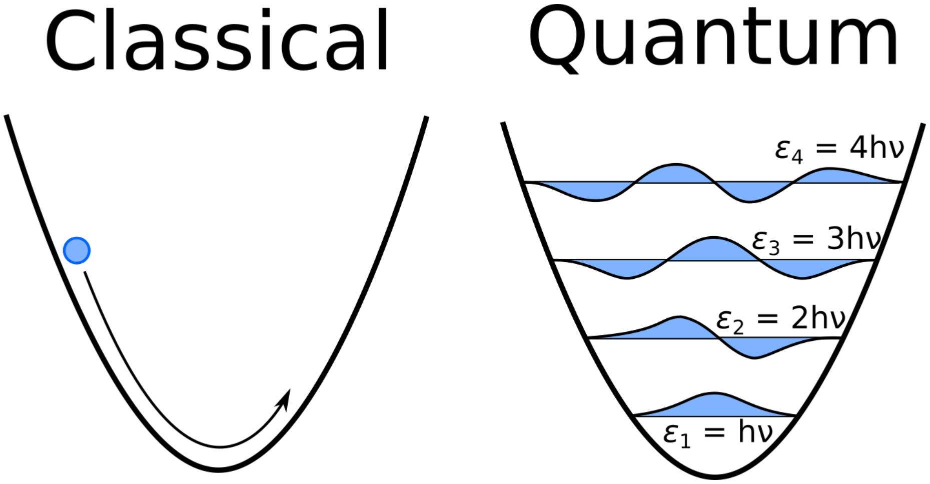
$$\varepsilon_n = nh\nu. \tag{375}$$

for the possible energies ε_n of the harmonic oscillator (basically a ball on a spring). Here

$$h \approx 6.626 \cdot 10^{-34}\mathrm{J} \cdot \mathrm{s} \tag{376}$$

is a universal constant known today as *Planck's Constant*, while ν is the oscillation frequency of the oscillator. The integer n is called a *quantum number* and can take on any non-negative value, enumerating the possible energy states of the oscillator.

The implications of this statement go far beyond what Planck himself envisioned, but the key point for today is that Planck realized that in order to match experimental data, *the energy of the harmonic oscillator had to be quantized*, i.e., it had to have only a discrete set of possibilities: $h\nu$, $2h\nu$, $3h\nu$, etc. This may not sound revolutionary, but at the time it was earth-shattering. In classical physics, the energy of a harmonic oscillator can take on a *continuous* range of values, depending on the position of the particle (i.e., the extension of the spring) and the speed at which the particle is moving. Planck found that a wide range of experimental observations could be explained only if, at the microscopic scale, harmonic oscillators took on discrete or *quantized* energies. This was the beginning of the quantum revolution.



31.3 Einstein's Frequency Relation

Einstein picked up on Planck's argument and pointed out that if the energy of *material systems* (like harmonic oscillators) is quantized, then we should logically also assume that the energy of the *electromagnetic field* is quantized. This led Einstein to a simple expression

$$E_{\text{photon}} = h\nu \tag{377}$$

for the energy of what we now call *photons* – small "particles" of light that transport electromagnetic energy through space. Again, this idea was revolutionary: since James Clerk Maxwell's authoritative treatment, light had previously been considered purely a wavelike phenomenon.[12] It was Einstein's demonstration that another puzzling experimental observation, the *photoelectric effect* could be explained by "quantizing" the electromagnetic field that led to his receiving the 1921 Nobel Prize in Physics.

31.4 The Bohr Frequency Relation

Building further on these ideas, Neils Bohr suggested that energy and frequency were related for *all* systems through the energy *difference* relationship

$$\varepsilon_m - \varepsilon_n = h\nu_{mn}. \tag{378}$$

Here ε_m and ε_n are the energies of two quantum states m and n, i.e., microstates for some material system. The quantity ν_{mn} is the *frequency of radiation emitted* by the system when it passes from energy state m to energy state n.

31.5 The Old Quantum Theory

Embedded in each of these relations is a pair of distinctly quantum-mechanical concepts:

- That there is a *direct, fundamental relationship* between the *energy* of a system and the *frequency* with which it moves, and
- That the energy states of a system may be *quantized*, i.e., limited to discrete values.

Together with a somewhat complex framework for determining the correct energy states for real systems, these hypotheses formed the basis of what became known as the *Old Quantum Theory.* In this theory, objects were generally regarded as still moving classically according to Newton's equations (at least in a limited sense), only with a limited set of "orbits" available to them. The most famous example of the use of this Old Quantum Theory was Niels Bohr's analysis of the hydrogen atom. Bohr's explanation of the electronic absorption spectrum of atomic hydrogen was the first great triumph of quantum theory, and hopes were high for some time afterward that the Old Quantum Theory would offer a simple, concrete description of *all* real systems.

Unfortunately, it turned out that the Hydrogen atom was as good as life got for the Old Quantum Theory. Even the next element on the periodic table – atomic Helium – turned out to be impossible to describe using the Planck-Bohr-Einstein prescription for energy quantization. A new theory was needed. And, with it, the equations would get only wilder.

31.6 The Canonical Commutation Relation

The first critical step away from the "Old Quantum Theory" was the recognition by Werner Heisenberg, Max Born, and Pascual Jordan that Bohr's frequency formula [Eq. (378)] appeared naturally from the *classical* equations of motion if one interpreted the usual position and momentum variables x and p not as simple *functions* but as *matrices.*

---

[12] It's worth noting, however, that both Newton and Lucretius (among many others) believed light was composed of particles. It was only relatively recently (the late 1800s) that light had come to be viewed definitively as a wave.

We don't have time to go here into the details of matrix algebra, but the key distinction of interest to us here is that, while matrices can be added, subtracted and multiplied just like numbers and functions, the multiplication is *non-commutative.* In simple terms, this means that the left-side product $\mathbf{AB}$ of two matrices $\mathbf{A}$ and $\mathbf{B}$ is *not* the same as the right-side product $\mathbf{BA}$. This is very different from the multiplication of two numbers a and b where we have the commutative relationship $ab = ba$.

Specifically, Heisenberg discovered that Bohr's frequency formula was recovered from the classical equation of motion (based on Newton's equations) if one replaced the classical position x and momentum $p = mv$ with infinite-dimensional matrices $\mathbf{x}$ and $\mathbf{p}$ that satisfied the *canonical commutation relation*

$$\mathbf{xp} - \mathbf{px} = i\frac{h}{2\pi}. \tag{379}$$

In a certain sense, this finding decisively "solved" the problem of quantum theory. If consistently applied, the Born-Jordan-Heisenberg "matrix mechanics" embodied in this formula turns out to provide an *exact* treatment of the quantum-mechanical properties of *all* material systems. The problem was that the method was impossibly opaque and extremely difficult to apply in practice to even simple systems. (It was an extremely difficult achievement to show that it did, in fact, give the correct fluorescence spectrum for both the Hydrogen and Helium atoms.) Fortunately, other great minds were at work on the same problem, and a simpler – but also exact – approach was soon to appear.

31.7 De Broglie's Wavelength Formula

While Heisenberg was puzzling over matrix products, a young French physicist named Louis de Broglie was thinking about Bohr's solution to the hydrogen-atom spectrum from a completely different perspective. Whereas Einstein had pointed out that a logically consistent theory should "quantize" the electromagnetic field (like the harmonic oscillators it interacts with), de Broglie began to see that one could achieve similar results by instead "wave-ifying" the matter. Despite Einstein's particle-like explanation of the behavior of light in experiments like the photoelectric effect, many classic experiments made clear that light nonetheless had many wave-like properties. de Broglie reasoned that if a classical wave phenomenon like *light* had particle-like properties, then it was perfectly reasonable that classical particles (like electrons) should also have *wavelike* properties.

In particular, de Broglie realized that Bohr's quantization of the hydrogen atom could be obtained (without any reference to matrices or noncommutative products!) by assuming that the electrons orbited the hydrogen nucleus as *waves* with wavelength

$$\lambda = \frac{h}{p} = \frac{h}{mv} \tag{380}$$

where $p = mv$ is the momentum of the electron. (Here m is the mass of the electron, and v is its velocity.)

Since de Broglie's formula predicts that the wavelength of an object is inversely proportional to its momentum (and therefore its mass), the de Broglie wavelength of macroscopic objects works out to be vanishingly small. For example, the wavelength of a baseball (roughly 150 grams) traveling at the modest speed of 40 miles per hour works out to be

$$\lambda = \frac{6.626 \cdot 10^{-34}\,\mathrm{kg\,m^2/s}}{40\mathrm{mph} \cdot 0.150\,\mathrm{kg}} \cdot \frac{\mathrm{mile}}{1609\mathrm{m}} \cdot \frac{3600\,\mathrm{s}}{\mathrm{hr}} \approx 2.5 \cdot 10^{-34}\mathrm{m}, \tag{381}$$

many orders of magnitude below the detection limit of any modern experimental equipment. For very small objects like electrons, however, the de Broglie wavelength turns out to be significant, on the order of the the diameter of the atoms binding them.

The key significance of de Broglie's observation is that it provided a *rationale*, however counterintuitive, for *why* quantization should occur: whereas at macroscopic scales, physics is dominated by *either* wave-like *or* particle-like behavior, at the atomic scale a *wave-particle duality* becomes visible: particles behave in some ways like waves and in others like particles. Quantization falls out as a necessary consequence of the periodicity of electronic "particle waves" in regular orbits around atomic nuclei.

Unfortunately, like the Einstein-Bohr "Old Quantum Theory", de Broglie's analysis turned out not to give a complete solution to the quantum mechanical problem. In particular, it was not at all obvious how it could be applied to more complex systems like the helium atom in order to give reasonable agreement with experiment. Fortunately, that development was waiting just around the corner.

31.8 The Time-Dependent Schrödinger Equation

Inspired by de Broglie's wave-particle hypothesis, a young Austrian physicist named Erwin Schrödinger began puzzling over how a general "wave theory" of quantum mechanics might be formulated. How exactly Schrödinger reasoned through this process is not entirely clear. Certainly he was inspired by the classical diffusion equation that describes the spread of the probability density for particles diffusing through a solvent, the ensemble equivalent of the random-walk process we discussed earlier in the semester. And he was clearly influenced by de Broglie's formulation of wave-particle duality.

But whatever his reasoning, the ultimate result is that he proposed a formula – now known as the *time-dependent Schrödinger equation*

$$\frac{ih}{2\pi}\frac{\partial\psi}{\partial t} = \hat{H}\psi \tag{382}$$

for the time-evolution of a new quantity he called the "wave function" for quantum mechanical systems.

This equation involves several new features that we haven't seen before. First, the wavefunction ψ itself was in Schrödinger's time a somewhat mysterious object and, to a considerable extent, remains mysterious to this day. The simplest interpretation of ψ is that, when given as a function of position x, it represents the *probability amplitude* for the position of a particle. Mathematically, this means that

$$\text{Probability (position} \approx x) = |\psi(x)|^2 = \psi^*(x)\psi(x), \tag{383}$$

i.e., the probability that a particle is at the location x is equal to the *squared absolute value* of the wavefunction. (Since the wavefunction is, in general, a complex creature, the squared absolute value $|\psi(x)|^2$ is obtained by multiplying $\psi(x)$ by its complex conjugate $\psi^*(x)$.) Note that this interpretation of ψ as a probability density already hints at the probabilistic aspect of modern quantum mechanics that Schrödinger himself came to dislike so strongly.

The second unfamiliar feature of Schrödinger's equation is the appearance of the Hamiltonian not as a classical function $H(x, p)$ but as an *operator* $\hat{H}$. Schrödinger was familiar with the mathematics of linear operators from his study of mathematical analysis and differential equations, but his introduction of these ideas to the treatment of quantum mechanics opened an entirely new vista toward understanding the developing theory. In analysis and quantum mechancs, an *operator* is a sort of "function of functions": Just like a function takes numbers as input and returns other numbers, an operator takes *functions* as inputs and returns *other funtions*. You're already very familiar with the *differential operator* $\frac{d}{dx}$ which takes a function $f(x)$ and returns its derivative $f'(x)$:

$$\frac{d}{dx}f(x) = \frac{df}{dx} = f'(x). \tag{384}$$

But in general *any* method for mapping functions to other functions can be used to define an operator. In our notes, we'll always denote such operators by "hats", e.g., $\hat{H}$ is an operator, whereas H would represent a number or a function.

Schrödinger's key insight was to propose that each physical *observable* – i.e., a quantity like position, momentum, or energy that can be "observed" for a system – should be assigned their own unique operator. The Schrödinger equation exhibits one such operator $\hat{H}$ representing the Hamiltonian of a quantum system, i.e., its energy. More generally, Schrödinger proposed that the correct operators for each "observable" could be obtained by using the definitions

$$\hat{x} = x \tag{385}$$

$$\hat{p} = -i\frac{h}{2\pi}\frac{\partial}{\partial x}. \tag{386}$$

The position operator $\hat{x}$ thus just corresponds to multiplying by the position variable x, while the momentum operator corresponds exactly to the differential operator discussed earlier, with the addition of the imaginary prefactor $-i\frac{h}{2\pi}$.

31.9 The Time-Independent Schrödinger Equation

If you find Schrödinger's wave equation to be difficult to wrap your head around, you're in good company. The physics community to whom he first presented it also found it hard to understand. Schrödinger's next step, however, quickly convinced even the staunchest skeptics of the value of his approach. Starting from his wave equation, Schrödinger showed that, if quantum mechanics was to give an experimentally consistent description of system dynamics, then the energy levels for a system must be determined by the *eigenvalues* of its Hamiltonian operator. Specifically, Schrödinger deduced from his time-dependent wave equation [Eq. (382)], that the n^{th} energy value ε_n of a system must satisfy what is now known as the *time-independent Schrödinger equation*:

$$\hat{H}\psi_n = \varepsilon_n\psi_n, \tag{387}$$

where ψ_n is the n^{th} *eigenfunction* of the Hamiltonian operator $\hat{H}$ and ε_n is the corresponding *eigenvalue.* In general, an *eigenfunction* of an operator is simply a function that (apart from multiplication by a constant) is "left the same" by application of that operator. Thus if ψ_n is an eigenfunction of $\hat{H}$, then, according to Eq. (387) it is "left the same" by the action of $\hat{H}$, apart from being multiplied by the eigenvalue ε_n. As Schrödinger knew from his study of differential equations, each operator comes with its own unique set of eignefunctions and eigenvalues. His theory now proposed that the possible *energy levels* of *all quantum systems* could be calculated by simply finding this unique set of eigenvalues for the corresponding Hamiltonian operator.

The great advantage of Schrödinger's formalism is that the mathematics behind it was already well-developed in the mathematical fields of analysis and differential relations and that it offered a much more concrete prescription for solving quantum mechanical problems than the infinite-dimensional matrices of Heisenberg's matrix mechanics. Equally important, Schrödinger was able to demonstrate explicitly that the two theories were *mathematically equivalent*: Heisenberg's matrix theory could be derived rigorously from his own "wave mechanics" and (with the benefit of hindsight) vice versa. Although there was much still to be understood in the *interpretation* of these equations, Schrödinger's theory provided an essentially complete mathematical framework within which quantum mechanical problems could be analyzed.

31.10 Where things get weird: The Measurement Postulate

In the early 1900's, Heisenberg's matrix mechanics, together with Schrödinger's more palatable wave mechanics, gave a remarkably complete description of known experiemental observations. The absorption spectra of not only the hydrogen and helium atoms but even small *molecules* could be described with exquisite precision. Even the statistics of bizarre new experiments where electrons apparently interfered with themselves or changed back and forth from wave-like to particle-like properties were readily derived mathematically. But one problem still remained – and in fact, still remains with us today: interpreting what quantum mechanics *means* about the nature of the universe, when and where particles exist, and what properties they have *before* a measurement is made.

It takes a fair bit of mathematics to explain the "strangeness" of quantum theory, but much of it comes from the last (and most-debated) hypothesis to be added to the theory: the *measurement postulate.*

The **measurement postulate** in quantum mechanics says that when a measurement is made on a quantum system, its wavefunction *collapses* instantaneously to match the value of the property that was just measured. In equations

$$\psi' = \hat{P}\psi \tag{388}$$

where ψ' is the wavefunction *after* the measurement, ψ is the wavefunction before the measurement, and $\hat{P}$ is a *projection operator* that "filters out" only those parts of ψ that are consistent with the experimental values just measured.

For the purposes of this course, you don't need to worry about exactly what a "projection operator" does or how it works. What you *do* need to know is that, according to the measurement postulate, *measurements inevitably and instantaneously change the state of quantum systems.* The state ψ' *after* the measurement is generally different from the state *before* the measurement.

To be sure, there are some classical analogs to this behavior. One way to measure exactly how much weight a rope could hold would be to keep adding very small weights one at a time until the rope broke. (This is actually very nearly the way many protein force-pulling measurements work!) This would give you a very precise measurement of the "maximum load" of the rope, but it would also destroy the rope in the process! The state of the rope is *changed* by the measurement process.

What's strange about the quantum measurement postulate is that it claims that the state of a quantum system is *necessarily* modified by measurement – regardless of how you do the measurement. In the rope analogy, you could probably devise other ways to make the measurement that keep your rope intact (e.g., test an identical rope or even just a single fiber from the rope of interest). But the measurement postulate says that it's *impossible* to devise a measurement method that doesn't perturb the system. Intuitively, this can be understood as a result of "quantization": becuase energy in quantum systems comes in discrete packets (like a photon), there's always some minimum perturbation to a system when you make a measurement on it. Classically, energy can come in infinitely small quantities (because it's continuous), so you can always imagine carrying out the measurement in such a way that the system is left effectively unperturbed. Quantization means that you can't do this for quantum systems: some amount of perturbation is inevitable.

31.11 The Uncertainty Principle

The measurement postulate might not sound like that big of a deal. So what if a system gets perturbed by measurement? In the real world, this is *usually* true anyway (even in classical mechanics), so what does it matter if it's *always* true in quantum mechanics?

As it turns out, it matters a great deal. One of the most immediate consequences is what is now known as Heisenberg's *uncertainty principle.* Heisenberg realized early on in his development of matrix mechanics that, because of the canonical commutation relation $\mathbf{xp}-\mathbf{px} = i\frac{h}{2\pi}$, a given system can have precisely defined values of *either x or p*, but never both at the same time. Mathematically, this is stated as

$$\sigma_x \sigma_p \geq \frac{h}{4\pi}, \tag{389}$$

where σ_x represents the *uncertainty* (i.e., standard deviation) of the position x, σ_p represents the uncertainty in the momentum p, and h is our old friend Planck's constant. Intuitively, this relationship says: There is a limit to how precisely we can know *both* the position *and* the momentum of a quantum system simultaneously; if we can predict the outcome of an x-measurement precisely (σ_x is small), the we must *not* be able to predict the outcome of p measurements precisely (σ_p is large); conversely, if we can predict the outcome of a p-measurement precisely (σ_p is small), the we must *not* be able to predict the outcome of x measurements precisely (σ_x is large).

What does this have to do with the measurement postulate? Well, initially it was hoped that one could overcome the uncertainty principle by making successive measurements: if we first measure x precisely and *then* measure p precisely, we'll know both of them to any precision we desire! Unfortunately, this scheme turns out not to work. As a result of the measurement postulate, it turns out that any measurement of x completely "scrambles" any knowledge of p; and any measurement of p "scrambles" our knowledge of x. As a result, successive measurements don't actually buy us anything: we *can* measure x first and then p, but once we measure p our previous measurement of x isn't valid any more – the particle is so perturbed by the measurement of its momentum that it's position will have changed completely (and unpredictably) by the time the p measurement is finished!

32 Quantum Mechanics: Fact or Fiction?

32.1 Introduction to Quantum Mechanics

The last few lectures have introduced us to the final and probably least-understood topic in our survey course: Quantum Mechanics. At a fundamental level, quantum mechanics is the field of physics that governs material systems *at very small energy and length scales.* When we're dealing with atoms (or sub-atomic particles), quantum mechanics is the only game in town. But, because *everything* is ultimately made of a very large number of very small particles, the rules of quantum mechanics ultimately govern *all material systems*, from diode lasers to the human body.

Unfortunately, because complex systems (such as individual cells, let alone higher organisms) involve so many different particles, we usually can't trace out all the ways that quantum mechanics manifests itself in real life. And in the short time we have left int his course, we won't have time even to give an in-depth treatment of the basic rules of quantum mechanics (a course in itself), let alone describing how those rules play out in living systems.

Instead, we'll offer a "birds-eye view" of *what's so distinctive* about quantum mechanics and how some of those distinctive characteristics play out in biological systems. Today we'll start off with a "fact or fiction" series dealing with some of the common conceptions and misconceptions surrounding quantum theory.

32.2 Quantum Mechanics is really hard.

Fact or Fiction? A bit of both:

- *Quantum mechanics is hard at the beginning,* because both the *mathematical tools* and the *philosophical concepts* involved are very different from what most of us encounter in daily life. As a result, quantum concepts are likely to feel abstract or exotic to newcomers, and the techniques are likely to feel awkward and difficult. This is explained nicely by an xkcd comic: `https://explainxkcd.com/wiki/index.php/1861:_Quantum`.

- *But quantum mechanics gets easier fast.* On the other hand, quantum mechanics gets easier much faster than many other technical subjects. Once you've got the core concepts and tools in place, you can cover a great deal of "ground" very quickly. In fact, there are quite a few problems in physics that are relatively easy to solve in quantum mechanics but very difficult in classical mechanics. This situation again varies dramatically from other physical fields like relativity or Newtonian dynamics. In both cases, a fairly simple set of equations is needed to *get started* in the subject, but the mathematical complexity balloons dramatically as you ask more difficult questions. Quantum mechanics is just the opposite: The initial required effort is rather large, but this "buys" results for a fairly large number of interesting problems.

32.3 Quantum Mechanics proves that life is random.

Fact or Fiction? Fiction.

One of the most frustrating aspects of quantum mechanics is that – although the equations work beautifully – it's fundamentally unclear *how the theory should be interpreted*, i.e., what it is exactly that those equations tell us about *how the universe actually is* or what even really *exists.*

The traditional description of quantum theory (the Copenhagen interpretation) says that quantum mechanics indicates a fundamental "randomness" to the universe. Heisenberg's famous *uncertainty principle* – which states that certain quantities (e.g., position and momentum) can never simultaneously be known with certainty – is perhaps the most commonly cited "proof" for this idea.

However, there are actually several different ways to interpret both the basic philosophy of quantum mechanics and, in particular, the uncertainty embedded in Heisenberg's principle. The most extreme opposite of the Copenhagen Interpretation is that of absolute determinism or *Superdeterminism* which suggests that all events in both the past and future are determined ahead of time – including the conscious choices that human beings make apparently of their own free will.

The take-home point is that the rules of quantum mechanics *do* point to a variety of uncertainty that is *not* present in classical physics, but whether or not this uncertainty reflects a fundamental "randomness" in the universe remains a philosophical question – much as it did *before* the advent of quantum theory. Compare, for example, the arguments of Lucretius (c. 99 BC to c. 55 BC) on atomic motion and free will [`http://gutenberg.org/files/785/785-h/785-h.htm`, Book II, "Atomic Motions"]:

The atoms, as their own weight bears them down
Plumb through the void, at scarce determined times,
In scarce determined places, from their course
Decline a little – call it, so to speak,
Mere changed trend. For were it not their wont
Thuswise to swerve, down would they fall, each one,
Like drops of rain, through the unbottomed void;
And then collisions ne'er could be nor blows
Among the primal elements; and thus
Nature would never have created aught.

(The underlining is mine.) Compare this with the perspective of Erwin Schrödinger (one of the founders of quantum theory):

> If we are going to stick to this damned quantum-jumping, then I regret that I ever had anything to do with quantum theory.

The point is that the choice of whether to embrace "randomness" as a fundamental characteristic of the universe has much less to do with the math than with philosophical predisposition.

32.4 Quantum Mechanics proves that there are an infinite number of parallel universes.

Fact or Fiction? Fiction.

However, while quantum mechanics doesn't *prove* anything about other universes, a variety of multi-universe theories have been *proposed* to resolve (in a certain sense) some of the interpretation problems that quantum mechanics otherwise leaves unresolved. The most famous is the *Many-Worlds* interpretation, proposed originally by Hugh Everett, which posits that each time a random event occurs in quantum mechanics the universe splits itself into two or more parallel tracks one corresponding to each possible outcome from the random event. Whether such "multiverse" theories actually improve the situation or not is something of an open (and mostly philosophical) question since no experiments have so far been proposed that could actually test them. While, depending on your perspective, quantum mechanics may indeed *suggest* a particular idea, there are very few cases in which the scientifically testable facets of quantum theory can really be said to "prove" something – outside of the rather technical scope of quantum mechanics itself.

32.5 Quantum Mechanics proves X about Y, where Y is a topic in philosophy, religion, art, etc.

Fact or Fiction? Almost always fiction.

As you're probably realizing by this point, quantum mechanics poses a rich array of conundrums, some of which are very much related to traditional philosophical and religious questions like free will and determinism. Combine this with the fact that very few people understand the math behind quantum theory in detail, and it can be very hard to resist the temptation to use it to defend or "prove" one philosophical perspective or another. Keep in mind, however, that most of these "proofs" depend on one particular interpretation of quantum theory (Copenhagen and many-worlds are probably the most popular)

32.6 Schrödinger's cat is both alive and dead.

Fact or Fiction? Depends on who you ask.

Along with his scientific and philosophical ally Albert Einstein, Erwin Schrödinger had a knack for putting a fine point on the conceptual difficulties associated with the quantum theory he helped to formulate. Few concepts in quantum theory are now as famous as his (purely hypothetical!) "cat experiment" in which a live cat is placed in a box equipped with a "diabolical device" that may or may not kill the cat, depending on the outcome of a measurement that – according to the rules of quantum mechanics – has a 50/50 chance of resulting in life or death for the cat. The point that Schrödinger wanted to make is that, if the rules of quantum mechanics are followed, one is led to the surprising conclusion that the cat is *neither alive nor dead* until the box is opened by a (presumably human) outside observer. Although not explicitly stated, Schrödinger appears to regard this situation as essentially nonsensical, and played a considerable role in his own distaste for the standard (Copenhagan) interpretation of quantum theory (that "damned quantum-jumping" again). Thus Schrödinger's own view seemed to be that the possibility of a cat being both dead and alive was obviously unreasonable.

In spite of such vigorous complaints from Schrödinger and Einstein, however, the Copenhagen interpretation was vigorously defended by Bohr and Heisenberg, who seemed to regard hesitation in embracing the "new quantum theory" as something of a philosophical weakness. Though they often disputed the details and philosophical implications of Schrödinger's and Einstein's arguments, they did not shy away from acknowledging that the theory was radical and could, ultimately, lead to bizarre outcomes like dead/alive cats.

On the other hand, the entire debate was viewed with some impatience by others on the scene, particularly Paul Dirac who regarded the whole thing as something of a waste of time. In Dirac's view, quantum mechanics had better-defined (more mathematically precise) problems that should be dealt with first. The philosophy could be understood later. This is perhaps the perspective that has become most widely adopted among physicists today, although the temptation to raise eyebrows by introducing undead cats into casual conversation or research headlines can be almost irresistible at times.

So, in abbreviated form, three answers to the "dead-and-alive cat" proposition:

- Bohr and Heisenberg: Fact!
- Schroödinger and Einstein: Fiction!
- Dirac: Can we focus on something important?

32.7 Einstein was wrong about Quantum Mechanics.

Fact or Fiction? A little of both:

- *Einstein was right* about the quantum behavior of light (think "photons"): he even got a Nobel prize for it. (Rather astonishingly, he never did get a Nobel prize for his formulation of the theory of relativity.)
- *Einstein was right* that quantum mechanics implied bizarre behavior ("spooky action at a distance") in experiments involving "entangled" particles. Apart from Schrödinger, nobody else seemed much to care about this at the time.
- *Einstein was wrong* in guessing that the predicted bizarre behavior wouldn't happen in the real world. Experiments showed only decades after his death that the experiments really are bizarre!
- *Einstein's obsession* with this bizarre behavior led after his death to the development of fields like quantum computing and quantum cryptography.

In short, although his "guess" about the experiment turned out to be wrong, his intuition about where the physics got interesting was exactly right and resulted in a completely new field of science and technology. If you count this as being "wrong", it's a great way to be wrong!

32.8 $E = mc^2$

Fact or Fiction? Fact. But wrong field – this equation belongs to Einstein's theory of relativity, though it's often confused with quantum mechanics.

32.9 Quantum Mechanics contradicts Relativity.

Fact or Fiction? Fact.

One of Einstein's chief complaints against quantum theory was that it (in a certain sense) contradicted his own theory of relativity. Most of the specific problems he proposed turned out to be solvable without major changes to quantum theory itself, but even today inconsistencies between the two theories reveal themselves in extreme cases. Relativity can be thought of as the physics of very large systems, while quantum mechanics describes very small ones. Interpolating between those two results turns out to be problematic, and physicists are still looking for the "unified field theory" that will bring them together – and perhaps resolve some of the philosophical conundrums raised by the two theories individually.

33 Quantum Mechanics in Biology

After all this discussion of quantum mechanics, you may be left wondering: What does all this have to do with biology or the life sciences?

It's a reasonable question. Overt "quantum effects" – like entanglement and tunneling – are usually visible only at the microscopic scale. The wet, complicated, noisy environment of most living cells obscures most of these effects to the point that they aren't recognizable any longer as something fundamentally quantum.

That's not to say, though, that quantum mechanics doesn't have important implications for biology. Today we'll look at a smattering of these, ranging from familiar and clear-cut examples like molecular bonding to the more obscure and tenuous such as proposals of "quantum consciousness".

33.1 The UV Catastrophe, Molecular bonding, and the Genetic Code

We've already covered one key result of quantum theory that's very much relevant to biology: The **avoidance of the ultraviolet (UV) catastrophe**. The classical blackbody radiation formula [Eq. (352)] predicts that hot objects like the sun should produce vast quantities of radiation in the x-ray, gamma-ray and even higher-frequency portions of the spectrum. As you might imagine, this wouldn't be a healthy situation for life on earth!

But, more fundamentally, the same "quantization" of molecular energy states that averts the UV catastrophe also plays a fundamental role in building up material systems at the atomic and molecular level. In a classical atom, electrons would orbit closer and closer to the nucleus, gradually losing all of their orbital energy as radiation in a continuous "death spiral" into the nucleus. It is only the quantization of electronic energies that forbids this behavior in quantum mechanics. Thus, at a foundational level, quantum mechanics is relevant to biology in the sense that it forms the basis for **atomic stability and the molecular bonding** that produces the water, proteins, nucleic acids, and carbohydrates that form the building blocks for life.

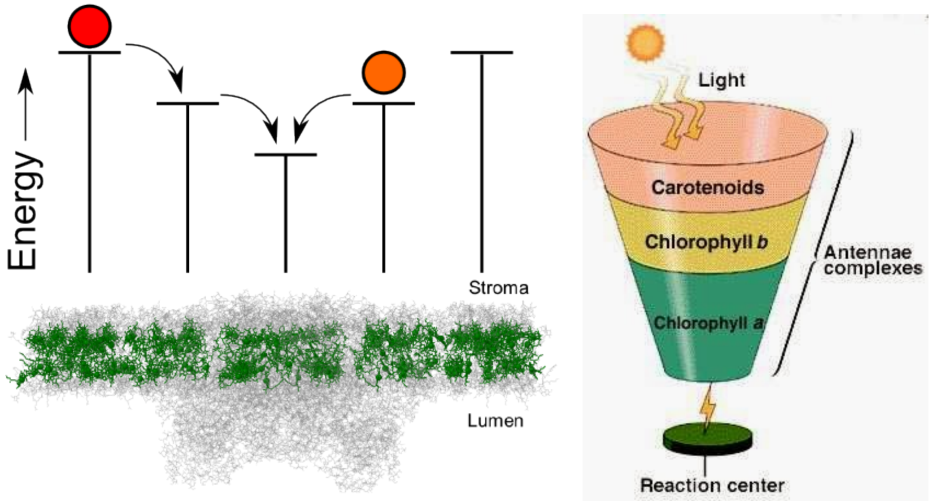
The significance of quantization is particularly apparent in the context of nucleic acids and the genetic code. As Erwin Schrödinger pointed out in his book *What Is Life?* – a mix of exploratory science and personal philosophy – the ability of organisms **to store and transmit genetic information** is dramatically aided by the fact that chemical bonds can be used to produce a *discrete chemical code* for the storage of information. Imagine if adenine and thymine weren't discrete chemical units – as they appear in DNA sequences – but could be continuously changed from one to the other. (This is indeed exactly what might be expected from a classical system in which electrons and atoms exist not in a limited set of quantized states but in a continuous range of classical configurations.) Such an arrangement would hopelessly complicate the storage of the information essential to biological replication and reproduction, or even to the consistent functioning of a individual protein or organism.

Most of the quantum effects just described involve processes that are fundamental to almost all processes in the physical world, not only biology. But the biological world has produced more distinctive quantum mechanical effects as well. The next three examples we'll consider fall into this categorie. Keep in mind, however, that the physics of life processes are often incompletely understood. The examples we'll consider range from well established (the photosynthetic energy funnel) to highly speculative (a hypothesized relationship between consciousness and quantum collapse).

33.2 The Photosynthetic Energy Funnel

One quantum-mechanical effect that's rather well-established the existence of the *photosynthetic energy funnel*: the spatial organization of photosynthetic light-harvesting proteins in a way that rapidly "funnels" solar excitation energy from a broad array of antenna proteins into the "reaction center" where photochemistry takes place. The illustration below shows one example, the Photosystem II lightharvesting supercomplex of green plants; in this system, pigments that absorb at short wavelengths (high frequency) such as carotenoids and Chlorophyll *b* molecules are situated far from the reaction center, while pigments that absorb at long wavelengths (low frequency) comprise the reaction center and its immediate environment.

If light-harvesting operated *classically* in this system, energy would diffuse through the complex in an unbiased "random walk" from complex to complex, seeking the even distribution of energy mandated by the classical equipartion principle. Because of the quantum-mechanical correlation between energy and frequency, however, energy absorbed by high-frequency pigments moves preferentially toward lower-energy sites, moving steadily from the periphery of the complex to the reaction center at the core where it can be converted to useful chemical energy. Thus violation of the classical equipartition principle is leveraged to **increase the efficiency of energy transfer** from protein to protein toward the photochemical reaction center.



Right-hand image from: `https://thebiologs.blogspot.com/2014/09/cape-2-photosynthesis-structure-of-leaf.html`

This well-established phenomenon in photosynthetic light harvesting should be distinguished from a variety of more recent claims regarding the role that might or might not be played by "quantum coherence" in photosynthetic function. (For a recent discussion see: `https://physicsworld.com/a/is-photosynthesis-quantum-ish/`.) These claims originated around 15 years ago in response to the result of new experiments using ultrafast (laser-based) spectroscopic experiments. Early interpretations of the data suggested that exotic quantum effects like entanglement might play a role in enhancing light-harvesting efficiency, but these ideas have been hotly debated and in some cases definitely debunked. Final answers are still a topic of current research.

33.3 Magnetoreception

Another phenomenon in which quantum mechanics is widely believed to be involved is the functioning of the "avian compass", specifically the ability of birds (and some other organisms) to navigate by *magnetoreception* the detection of the earth's magnetic field. Many migratory birds appear to use the earth's magnetic field as an internal compass during long flights and become confused and disoriented when placed inside of cages with changing magnetic fields. The mechanism by which this "magnetic sensing" occurs, however, is essentially unknown. Several possibilities have been suggested, ranging from light-active proteins in the bird's eyes to iron-containing particles in their beaks.

Although a complete understanding of the phenomenon must clearly wait for better biochemical data, the physical basis almost certainly involves in some form the response of unpaired electrons to the earth's magnetic field. At some level, this process clearly involves quantum-mechanical phenomena since the quantization of electron spin is itself a quantum-mechanical effect. Whether this spin quantization (which is the basis for even non-biological compasses) is the *only* quantum effect at work in this process or if there are more exotic forces at work is a topic of considerable debate. What *is* clear is that avian magnetoreception is an exciting open field of research!

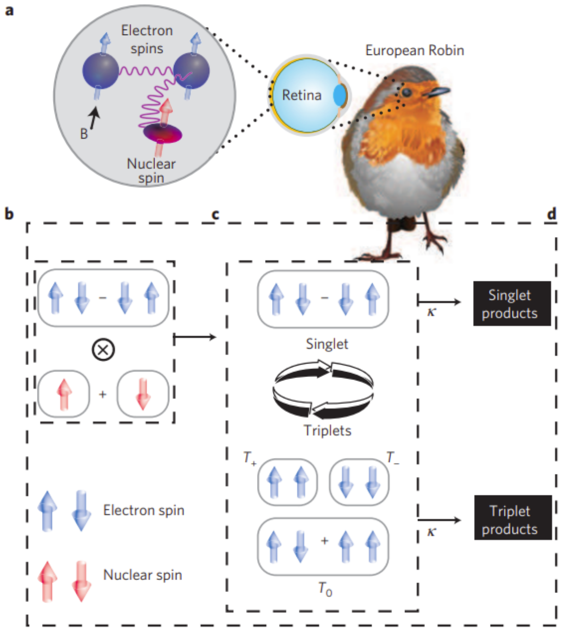


Image from:

33.4 Quantum Consciousness?

We'll close with a *highly speculative* proposal regarding the origins of consciousness in biological creatures such as human beings. This particular example is offered only to illustrate the range of ideas at play in the intersection between quantum mechanics and biology; it is **not in any way** intended as a personal endorsement of the idea. (In fact, I personally find the idea highly unlikely.)

The idea of a quantum-mechanical basis for human consciousness was popularized most emphatically by Roger Penrose, a British physicist and Nobel laureate who in his later career turned to thinking about biological problems. Based on abstract mathematical arguments involving Gödel's Incompleteness Theorems, Penrose suggested that any instrument capable of carrying out the neural processes common to the human brain must at some level rely on quantum-mechanical effects. An American researcher named Stuart Hameroff subsequently proposed to Penrose that cellular microtubules might be the physical origin of such a process, and the two of them together formulated a theory of "orchestrated objective reduction" that purports to describe how the collapse of quantum-mechanical wavefunction collapse in neuronal microtubular structures gives rise to human consciousness and decision-making. (For a more extensive discussion and references see here: `https://en.wikipedia.org/wiki/Orchestrated_objective_reduction`.)

34 Quantum Technologies

34.1 What's the big deal?

You've probably heard recently about one or more "quantum technologies" like quantum computing, quantum cryptography, or quantum sensing. Indeed, in the fall of 2019, Google announced to much fanfare that their quantum computing team had achieved "quantum supremacy" – a point where quantum computing can accomplish tasks that are beyond the power of any classical computer. (See `https://ai.googleblog.com/2019/10/quantum-supremacy-using-programmable.html`). Others have been more skeptical of Google's announcement (see here for IBM's response: `https://www.ibm.com/blogs/research/2019/10/on-quantum-supremacy/`). Nonetheless, the question remains: Why all the excitement about these "emerging" technologies, and what makes them "quantum"?

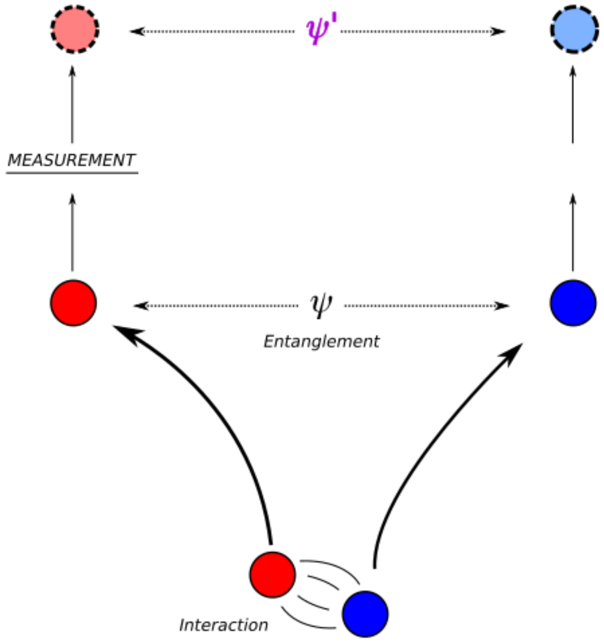
The name "quantum technology" serves as an umbrella term for a wide range of different devices, and it doesn't come with a single, precise, universally accepted definition. Indeed, as we've already seen, quantum mechanics is fundamentally at work in just about every technology you could imagine, since at the atomic scale the properties of *all* materials are governed by quantum mechanical principles. However, most devices commonly referred to as "quantum technologies" do share something in common: a reliance on the quantum mechanical phenomenon of *entanglement.* Thus we'll begin our discussion of quantum technologies with a brief look at entanglement itself.

34.2 Quantum Entanglement

The word "entanglement" was first used by Erwin Schrödinger to describe what he viewed as a deeply disturbing feature of quantum theory: the idea that (according to quantum mechanics) the physical state of two particles that had at any point in the past interacted with each other would continue to be defined by *a single wavefunction* – regardless of whether the particles continued to interact with each other or not. According to quantum mechanics, a two-particle system is described by a single wavefunction $\psi(x_1, y_1, z_1, x_2, y_2, z_2)$, where the coordinates x_1, y_1 and z_1 represent the first particle, and the coordinates x_1, y_2, and z_2 represent the second particle. At face value, this seems very similar to the situation in classical mechanics where (for example) we might use a probability density function to describe our knowledge about two classical particles, even if they weren't spatially close to each other.

What bothered Schrödinger about the situation in quantum mechanics is that – according to our old friend the Measurement Postulate – any measurement made on *one* of the two particles would *instantaneously change the state* of the *other* particle. Recall that the measurement postulate asserts that any measurement made on a quantum system necessarily perturbs the system; in the case of two entangled particles, this means that the state of particle #2 is *necessarily* disturbed when a measurement is made on particle #1 – regardless of how far apart the two particles are or whether or not they continue to interact with each other. Even worse, the Copenhagen interpretation of quantum mechanics indicates that the state of particle #2 changes *instantaneously* in response to measurements on particle #1. This apparent contradiction of the Theory of Relativity (which asserts that no signal can travel faster than the speed of light) was the feature that Einstein found most deeply troubling about the entire framework of quantum theory and that led him to reject the Copenhagen interpretation.

This entanglement "phenomenon" is illustrated schematically in the figure below. Two particles (red and blue) are initially close to each other in space and interact, causing their collective wavefunction ψ to become "entangled". Later, the particles are sent to different places where they no longer interact. Nevertheless, because the wavefunction is entangled, the act of measurement on the red particle *instantaneously* affects the state of the blue particle. Any subsequent measurements on the blue particle will reflect the new post-measurement wavefunction ψ' instead of the pre-measurement wavefunction ψ.



34.3 Quantum Computing

For many years, quantum entanglement was viewed as a purely abstract thought-experiment that didn't have much relevance to the real world. Indeed, it wasn't clear whether, if such measurements could actually be made, they would follow the rules that quantum mechanics predicted. It wasn't until after the deaths of both Schrödinger and Einstein that a British Physicist named John Stewart Bell outlined an experiment (and proved a related theorem) for how such measurements might actually be made. And it wasn't until some years after these experiments were actually done – confirming, contrary to Einstein's expectations, that entangled particles really *do* carry "spooky" correlations – that people begin to realize that these strange quantum mechanical effects might have some useful applications.

Probably the most famous (though still mostly hypothetical) application is *quantum computing*. Quantum computing takes advantage of the fact that a quantum particle can exist in a *superposition* of two or more states – just like Schrödinger's cat can be (in the Copenhagen view) both alive and dead at the same time. In classical computers, information is stored in a series of "bits": electronic markers that can take on the values 1 and 0. A 3-digit binary number on a classical computer could have the values

$$000 \text{ or } 001 \text{ or } 011 \text{ or } 101 \text{ or } 110 \text{ or } 111 \ .$$

At first glance, information storage in quantum computers works similarly: data is stored in "qubits" (quantum bits) that can take likewise on the values 1 or 0. For example, a qubit could correspond physically to a C=O stretch vibration, with "0" corresponding to the vibrational ground state and "1" corresponding to the first excited state. What makes quantum computing different (and, potentially, more powerful) than classical computing is that qubits can exist as a *superposition* of *all* possible states at the same time. Thus, the 3-digit qubit on a quantum computer could have the values

$$000 \textit{ and } 001 \textit{ and } 011 \textit{ and } 101 \textit{ and } 110 \textit{ and } 111$$

all at the same time! Now, suppose that the two computers were asked to do a minimization problem over these three-digit numbers. Because the bits in the classical computer can take on only one value at a time, the classical computer would need to check each of the 6 numbers in sequence. In contrast, because the qubits can take on multiple values simultaneously, a cleverly-designed quantum-computing algorithm could do the calculation on all six numbers *at the same time*, a six-fold speed-up relative to the classical case.

For such simple calculations, the quantum vs. classical speedup is rather modest. But what if the calculation were on a 100-digit number? A 100-digit binary number has approximately 10^{30} possible values. If a quantum computer could really do 10^{30} calculations in parallel, the speedup compared to the corresponding classical calculation would be astronomical!

This discussion is, of course, rather simplistic. For one thing, it's not at all obvious how to implement most practical calculations on a quantum computer in a way that leverages these potential advantages. For this reason, there's great interest right now in developing *quantum algorithms* that maximize the advantage offered by quantum computing, with the goal of solving problems that simply couldn't be solved using classical computers. This is what's meant by "quantum supremacy" – the idea that quantum computers can do things that classical computers fundamentally cannot.

The other reason why quantum computing hasn't yet changed our lives is that quantum computers simply don't exist yet – at least not on a scale needed to solve most practical problems. At the time of writing this text (Spring 2022), the biggest quantum computers available work with on the order of 100 qubits. This is big enough to do some interesting demonstrations, but not big enough to change the world. (For comparison, a classical computer like your laptop with 8 Gb RAM works with around 10^{10} bits.) Nonetheless, progress has ramped up dramatically over the last few years, and there's a strong expectation that we'll see real-world applications of quantum computing in the next decade.

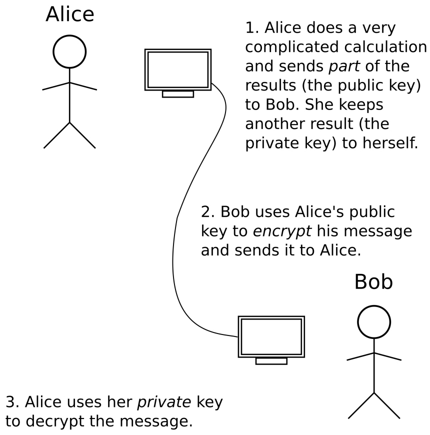
34.4 Quantum Cryptography

So what will we do with quantum computers once we have them? Ironically, the most obvious application available appears to be to destroy internet security. One of the few robust quantum algorithms that has an immense computational advantage over its classical competitors is *Shor's algorithm* for factoring large numbers. This may seem innocent enough. (What could be so malicious about breaking down a number into its prime factors?) But it has a rather surprising dark side: precisely *because* factorizing large numbers is very difficult on a classical computer, it has for decades served as the basis for most internet security protocols. A quantum computer capable of factoring large numbers millions of times faster than a classical computer could quickly break through the standard data encryption methods that protect both your personal credit card number and government military secrets alike.

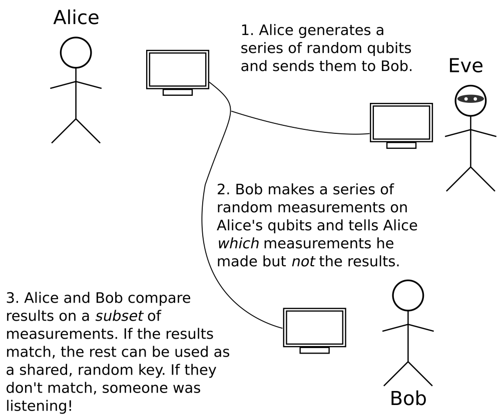
A schematic glimpse at how this works is provided in frame (a) of the figure below. Suppose Bob wants to send Alice a secret message. There are many standard *cryptographic ciphers* that Bob could use to encrypt his message (e.g., "padding" the message with irrelevant data to obscure the real content). But all these ciphers require the receiver of the message to have advance knowledge of a *key* that is used "unlock" the message, i.e., a way to filter out the nonsense information and restore the message to its original form. The challenge of getting the encryption key *securely* from sender to receiver is known as the *key distribution problem.*

Current key distribution protocols work something like this: The receiver (Alice) first does a very complicated calculation on her local computer. She then sends part of the results from that calculation – the "public key" to Bob. On Bob's local computer, he does a second calculation that uses the public key to "encrypt" his message; this encryption is very hard to reverse, without knowledge of the *other* half of Alice's original calculation, what's known as the *private key.* Even Bob can't decrypt the message once it's been encrypted! Only when Bob sends the encrypted message to Alice can she use her private key to descrypt and read the message.

Now, the "complicated calculation" that Alice did at the start of this process typically involves



(a) **Classical key distribution.**



(b) **Quantum key distribution.**

multiplying two large, randomly selected prime numbers. The reason the decryption is difficult to reverse is that (if you only know the public key) it's very difficult to guess *which* two prime numbers Alice started with. The danger posed by quantum computing to this system is that, using Shor's algorithm, a quantum computer could quickly factor the public key, determine the original prime numbers, recalculate Alice's private key, and then decrypt and read any messages Bob sends to Alice. With this in mind, it's not difficult to imagine that a malicious hacker armed with a working quantum computer could very quickly wreak havoc on our modern financial system!

Partly for this reason, interest has been ramping up recently in the idea of *quantum key distribution* – a solution to the key distribution problem that uses quantum entanglement and the measurement postulate to determine whether or not a communication line is secure. The basic idea behind this scheme is illustrated in frame (b) of the figure above. Briefly, Alice sends Bob a sequence of qubits, quantum particles in one of several possible states. Bob then makes a measurement on each particle, choosing *randomly* which measurements to make. Alice and Bob then choose a subset of the qubits for comparison. If Bob's measurement results match Alice's expectations, all is well, and the *rest* of the measurement results can be used as a shared random key that both Bob and Alice know and can use for encryption. The beauty of the scheme comes from the fact that if anyone else (say an "eavesdropper" named Eve) tried to monitor the original stream of qubits between Alice and Bob, she would (according to the measurement postulate) *necessarily* disrupt the corresponding quantum states. Alice and Bob would then be alerted of the interference when they compared Bob's measurement results with Alice's expectations.