WCGMS BCMB — BIOCHEMISTRY FALL 2021

STATISTICAL MECHANICS AND THERMODYNAMICS RELEVANT TO DRUG DISCOVERY

John Chodera[*]

Street-fighting statistical mechanics for biochemists

———

CONTENTS

For more visit http://choderalab.org.
[*]john.chodera@choderalab.org

Recommended reading

Molecular Driving Forces *2ed*, Dill and Bromberg
Chapter 2 ("Extremum Principles Predict Equilibria") and 3 ("Heat, Work, and Energy")
[PDF (click "Preview this Book")]
Chapter 5 ("Entropy and the Boltzmann Law") [PDF of 1ed version as Chapter 6]
Chapter 4 is optional mathematics refresher [PDF (click "Preview this Book")]

Optional reading and supportive texts

Statistical Physics of Biomolecules: An Introduction, Daniel Zuckerman
An excellent gentle introduction to most of the important concepts of statistical mechanics
relevant to biophysics, biochemistry, and drug discovery. [Amazon]
Street Fighting Mathematics, Sanjoy Mahajan
Possibly the most useful book you will read in your scientific career: How to solve order-of-
magnitude problems. [PDF] [Amazon]
Statistical Mechanics: Theory and Molecular Simulation, Mark E. Tuckerman
Mark Tuckerman teaches an excellent course at NYU for those looking to delve deeply into
the topic. [Amazon]
Introduction to Modern Statistical Mechanics, David Chandler
Known as the "little green book", this is the canonical introductory graduate statistical
mechanics text. [Amazon]
Monte Carlo Strategies in Scientific Computing, Jun S. Liu
A useful resource for those who wish to delve more into the mathematics behind the
algorithms we use for simulating biomolecular systems or doing Bayesian inference using
Markov chain Monte Carlo. [Amazon]

2 WHAT IS STATISTICAL MECHANICS AND WHY IS IT USEFUL?

The bad news is that it's impossible to learn all of statistical mechanics in two lectures. The good news is that you don't have to. A great deal of statistical mechanics and thermodynamics comes from some very simple ideas, and everything can be derived (with a bit of straightforward mathematics) from a few simple principles. While we focus on a few simple analytically soluble illustrative examples that can help build intuition, pretty much everything useful you might want to do with statistical mechanics requires some hard work with computer simulations or fancy algorithms to address questions in a quantitatively satisfying way, and even then, the quality of your results will be highly dependent on the accuracy of the model you use to represent physical systems.

In these two lectures, we'll focus on two goals:

- In the **first lecture**, we'll try to arm you with the fundamental ideas you will need to see if an idea is even remotely consistent with what we know about how physical systems *must* behave at the microscopic and macroscopic levels. We want to equip you with the tools to quickly weed out or identify Bad Ideas (TM) and hopefully, even some tools that help you create some simple useful models that help you understand systems at the microscopic level to explain the weird thing you're seeing in your experiments.

- In the **second lecture**, we want to give you the idea for how systems behave at the microscopic level that is relevant to biological macromolecules that carry out biological processes—where *real* biology happens. It's *very* different than our everyday experience in the macroscopic world. As a result, most of the movies that you've seen of how microscopic processes might work at the single-molecule level are just *wrong*—complexes moving together to associate with seemingly deterministic intent to associate, motor proteins chugging along at constant velocity, ribosomes happily translating away without ever stuttering, ATP synthases happily generating ATP without ever going backwards and hydrolyzing ATP—all of this is *wrong* in a way that can lead to intuitively satisfying but highly unpredictive mental models. Action at the microscopic level is not just corrupted by statistical motion—wiggling and jiggling—it is *dominated* by it. Consider a pool table: At the macroscopic scale, when you hit a ball with the cue ball, the balls beautifully follow Newton's laws, reflecting off the bumpers in highly predictable ways; at the microscopic scale, a pool table the size of the ribosome would see the pool balls all go in and out of all the pockets *all on their own*, and if you hit a ball with the cue ball, it wouldn't get very far before its trajectory would completely deviate from where you had intended to hit it. It's just *different*, and accounting for these differences is critical to building models that are *predictive*—the essence of science. We'll try to give you a sense of this stochasticity in the second lecture.

2.1   *Statistical mechanics*

Consider a single particle moving in a vacuum, or in something pretty close to a vacuum that contains very few other particles:
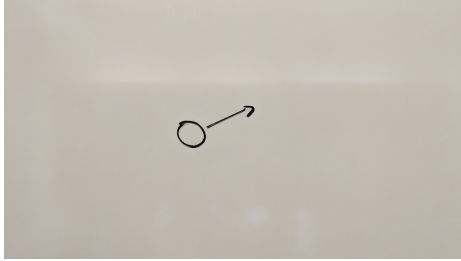
Maybe it's the earth moving in the gravitational pull of the sun, or an ionized protein moving in an electric field in a mass spectrometer. Over a broad and relevant range of length- and time-scales, it's easy to describe the motion of the object in these cases as they follow Newton's laws of motion to very good approximation: We can easily predict where the object will be some time t after we initially observe it if we know some basic properties of the system, like masses, velocities, positions, and the quantitative nature of interactions between the objects. In fact, we can do this so well that we can figure out how to toss a tiny spacecraft between two planets:
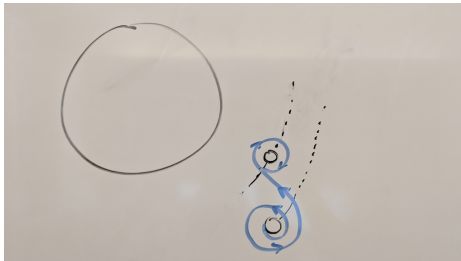
Now consider a crowded environment inside a cell or test tube:
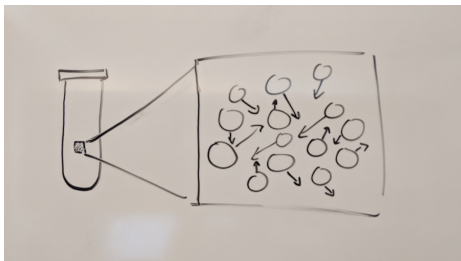
Consider this beautiful illustration of the interior of an *E. coli* by David S. Goodsell of the Scripps Research Institute:
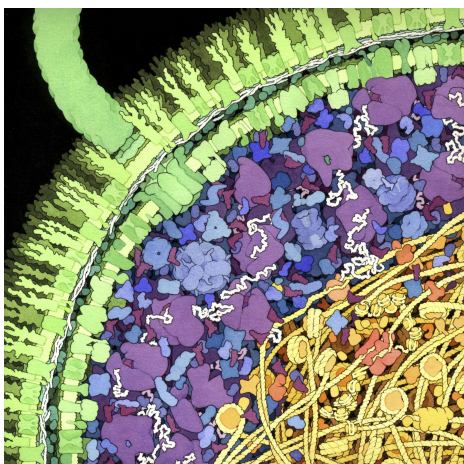
It might seem hopeless to make concrete predictions about the behavior of such a complex system. There's so much going on in this system—even if there are many identical copies of the same protein behaving in different ways, interacting with other molecules, jostled around by the solvent, that it's impossible to predict what any particle will do from Newton's laws unless we know *everything* about all of the biomolecules and solvent molecules in the system—their velocities, positions, and detailed interactions. It's almost information overload—even if we *had* all that information and could predict what would happen, what would we do with all of it? We need a simpler description of what is going on in this system that is actually *useful* to us. That's where statistical mechanics comes in.

Statistical mechanics is, unsurprisingly, just how it sounds: It is a set of tools for describing the mechanics of the system—that is, what is going on with the dynamical motion or configurations of a bunch of particles—in a statistical way so that we have some hope of describing the behavior of the system in a simple and useful way. We strive to come up with the simplest predictive statistical model of what we're interested in. For example, we might characterize the complex dynamics of many copies of a protein molecule in a test tube with a simple folding and unfolding rate that allows us to describe the complex dynamics of individual proteins or collections of them in a simple conceptual—but still predictive way. Often, there are many such models at different resolutions with different degrees of predictive power, such as a multistate model of protein folding that may also include a transiently populated on- or off-pathway intermediate state necessary to describe a higher-resolution spectroscopic experiment. At its core, statistical mechanics is about saying something simple and predictive about the average behavior of molecules or collections of molecules in a manner that exploits the fact that these molecules behave statistically.

2.2   *Thermodynamics*

Thermodynamics is a dinosaur. It is a complex, internally consistent set of relations that describe some properties of the macroscopic world in terms of self-consistently defined quantities that make no reference to the microscopic nature of matter; concepts like entropy are purely convenient constructs that do not make reference to the microscopic disorder of systems. While useful a century ago in the design of efficient steam engines, we find little use for it today beyond a few basic laws that can help act as a bullshit detector: If some proposed mechanism violates the laws of thermodynamics, it's probably incorrect. Most results we need in thermodynamics can be derived from statistical mechanics, but we'll review a few useful results from thermodynamics that will help you weed out bad ideas.

Let's consider a molecular system with an instantaneous **microscopic configuration** x, which describes the precise positions of all the atoms in the system at an instant in time. Because each atom has three degrees of freedom, $x \in \mathfrak{R}^{3N}$, where N is the number of atoms. x might also denote some other aspects of the microscopic configuration, such as the size of the box containing the molecules if it can fluctuate or its volume, $V(x)$.

A molecular system has a **potential energy** $U(x)$ that is a function of the instantaneous microscopic configuration x of all the atoms. The energy provides the driving force for atomic motion—if we start with a high energy configuration on a steep potential energy surface, the system has a tendency to roll downhill toward lower-energy configurations, but we'll see how temperature influences this in a moment. In classical statistical mechanics, the potential energy is only really defined up to an arbitrary additive constant—we can set the zero value to be at anything we want, kind of like taring a balance. One convenient zero is to say the potential energy is zero when to particles are infinitely far apart, but that's just one choice—others may be more convenient, but should never affect the results of anything in classical statistical mechanics. (Quantum statistical mechanics is different, but we'll skip that today.)

For example, noble gas atoms (and larger molecules like methane) have attractive dispersion interactions (or London dispersion forces) but repel each other when you get them so close that their electrons start to exclude each other via the Pauli exclusion principle (electronic repulsion). We call these *van der Waals* interactions.
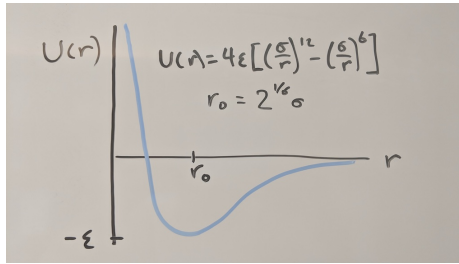
A simple model of van der Waals interactions has been used since the earliest days of minicomputers: the Lennard-Jones 12-6 model, which uses just two parameters—well depth ϵ and LJ radius ϵ—to describe the potential dependence on the interatomic separation r: The repulsive wall when two atoms get too close to each other should really be more like an exponential due to Pauli's exclusion principle, but it was easier to square the r^6 on early minicomputers to approximate an exponential overlap, so we seem to be stuck with that today in many computer models.

If we want to talk about the system's kinetic motion, the system also has velocities $v \in \mathfrak{R}^{3N}$, and kinetic energy $\mathcal{K}(v)$ defined by

$$\mathcal{K}(v) = \frac{1}{2} \sum_{n=1}^{N} m_n \|v_n\|^2 \tag{1}$$

The **total energy** $E(x, v)$ of an isolated system—one that is not in contact with some other system that allows it to exchange energy—is *constant* E_0, and cannot change.

$$E(x, v) = U(x) + \mathcal{K}(v) \tag{2}$$

You might sometimes see the total energy written as the Hamiltonian $H(x, v)$—it's identical to the total energy when you're talking about positions x and velocities v, but is a more general construct that allows you to talk about systems with other coordinate systems like angles, torsions, etc.
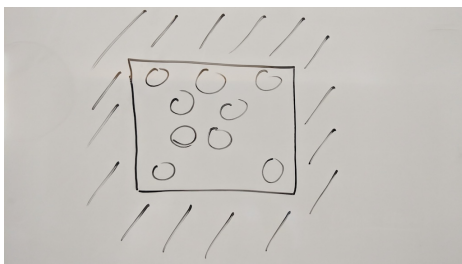
If we *do* bring the system into contact with another system where it can exchange energy, then the total energy can fluctuate, and the quantity of energy that flows in or out (for example, from collisions or interactions with other particles outside the system) is called *heat*.
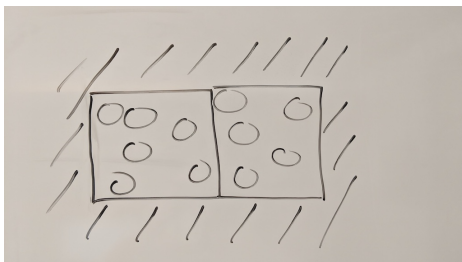

There's one more way we can change the total energy of the system: By doing **work** on the system (or allowing the system to do work on us). For example, we may pump some energy into the system with a laser, or push some particles around using an optical trap. Or we might let the particles bouncing against the wall of a container to do some work on us.

While heat and work are sometimes confusing quantities, they're just convenient labels for energy that is exchanged with either the environment (heat) or with something we are intentionally doing to the system (work), and will be useful for bookkeeping later on.

4 EQUILIBRIUM PROBABILITY DISTRIBUTIONS

Much of the time, we're interested in things we can say about the behavior of the system at equilibrium, which is often how we study biomolecules *in vitro*. Much of life is decidedly *not* at equilibrium, since the constant consumption of ATP converts chemical energy into work capable of driving chemical and biological processes such as reactions, motor proteins, and transport—often simply by providing a way to *suppress* reversibility and rectify thermal motion to drive things in a direction compatible with continued life.

We consider the case where a subsystem of interest—such as the region around a biological macromolecule in the context of part of a test tube or cellular environment. We can consider the region we study as a subsystem being in contact with some sort of macroscopic environment held at a fixed absolute temperature T (generally written in Kelvin, and always an *absolute* temperature where $T = 0$ denotes absolute zero—never Celsius or Fahrenheit, which use relative scales where $T = 0$ is not absolute zero!). We're interested in what we can say about the behavior of the macromolecule at equilibrium. Surprisingly, we can often say a great deal about the system with just a small amount of information.
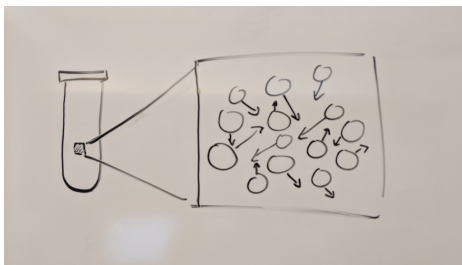
4.1 *The Boltzmann law*

The most important law governing the behavior of microscopic systems at equilibrium is the Boltzmann law. We'll start with a specific form of this law for the case of a subsystem with a fixed number N of particles at thermal equilibrium at temperature T:

$$\pi(x, v) \propto e^{-\beta E(x,v)} \tag{3}$$

where $\pi(x, v)$ is the probability of observing the system in a particular configuration x and with an instantaneous velocity v, $\beta \equiv (k_B T)^{-1}$ is called the **inverse temperature** and has dimensions of inverse energy, and $k_B \approx 1.381 \times 10^{-23}$ J/K is the Boltzmann constant. To make things simple, at room temperature (298 K), thermal energy $k_B T \approx 0.6$ kcal/mol. Because the argument of an exponential must be dimensionless, we note that the product of β (with dimensions of inverse energy) and E (with dimensions of energy) produces a unitless quantity that we can safely use as an exponent in Eq. 3.

Importantly, we've written the probability density $\pi(x, v)$ in terms of positions x and velocities v. If instead we used a different set of variables—such as internal degrees of freedom (angles, torsions) for a molecule in gas phase—we have to account for the change of variables and include a Jacobian term, and things can get messy. Just remember that if we're writing the density in terms of positions and velocities, we don't have to worry about adding in extra Jacobian terms.

The Boltzmann law can actually be derived in a variety of different ways—all of which are too complex to get in here—and is a simple consequence of counting of energy levels or energy densities in systems that can exchange energy with other systems to reach thermal equilibrium. The most interesting derivations are found in *Molecular Driving Forces* by Dill and Bromberg, which also show how the *principle of maximum entropy* can be used to derive this surprisingly powerful yet simple law. This form of the law—where we're describing a system where the number of particles N, volume V, and temperature T are held fixed—is said to describe the *canonical ensemble*, and the probability law is referred to as the *canonical distribution* or *canonical probability density*.

The first thing we notice if we substitute in the expression for the total energy $E(x, v)$ above (Eq. 2) is that we can factor it into potential and kinetic parts:

$$
\begin{aligned}
\pi(x, v) \quad &\propto \quad e^{-\beta E(x,v)} & (4) \\
&\propto \quad e^{-\beta[U(x)+\mathcal{K}(v)]} & (5) \\
&\propto \quad e^{-\beta U(x)} e^{-\beta \mathcal{K}(v)} & (6) \\
&\propto \quad \pi_x(x) \cdot \pi_v(v) & (7)
\end{aligned}
$$

Furthermore, we can work out exactly what the kinetic energy distribution is without knowing anything other than the atomic masses:

$$\pi_v(v) \quad \propto \quad e^{-\beta \mathcal{K}(v)} \tag{8}$$

$$\propto \quad \exp\left[-\frac{1}{2} \sum_{n=1}^{N} m_n \|v_n\|^2 \right] \tag{9}$$

That means that we can work out a huge number of kinetic equilibrium properties for our protein or small molecule without knowing anything more than the atomic masses!

Most of the time, however, we're interested in configurational properties, since the kinetic component is independent of the system and its interactions. To make any progress in configurational properties, we generally need to have a model of the system's interactions so we can write down a potential function $U(x)$ that we can do some computations with.

Let's consider the simplest tractable system with continuous coordinates: **A harmonic oscillator**. This might be a simple model of the motion of a colloidal bead confined to a laser trap, or a protein moving in a crystal at cryocooled temperatures, or the motion of an AFM tip used to probe microscopic matter: We can write the potential for a particle confined to a harmonic well:

$$U(x) = \frac{K}{2}x^2 \tag{10}$$

where K is the *force constant* or *spring constant* since the force is the negative partial derivative of the potential, and the K constant sits out front:

$$F(x) = -\frac{\partial}{\partial x}U(x) = -K\,x \tag{11}$$

The spring constant K has dimensions of energy over distance-squared, so that when multiplied by x^2, which has dimensions of distance-squared, it gives us an energy. (Dimensional analysis is perhaps the most useful tool in all of biology to keep you from making simple mistakes, so I highly encourage you check out the chapter in *Street Fighting Mathematics* on it.)

We can plug the potential $U(x)$ into the Boltzmann equation (Eq. 3) and subsequently know everything there is to know about the equilibrium static behavior of the harmonic oscillator:

$$\pi_x(x) \quad \propto \quad e^{-\beta U(x)} \tag{12}$$

$$\propto \quad e^{-\frac{\beta K}{2}x^2} \tag{13}$$

Here's an important trick: Whenever you see anything of the form

$$p(x) \propto e^{-\alpha x^2} \tag{14}$$

you should recognize you're dealing with a **Gaussian distribution** (or *normal density*). It's good to keep an eye out for these kinds of equations because you can exactly solve pretty much anything that features them. If you can remember that the Gaussian distribution has

the following form, you'll be in good shape:

$$p(x) \quad = \quad \frac{1}{(2\pi)^{d/2}\sigma^d} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{15}$$

where d is the dimension of the vector x, σ is the standard deviation, and μ is the mean. This is a normalized probability density, so
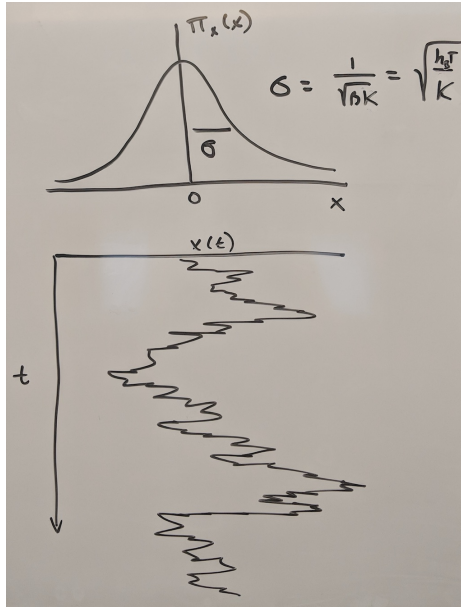
$$\int_{-\infty}^{+\infty} dx\, p(x) = 1 \tag{16}$$

Recognizing that it is a Gaussian means we can go further and write out an exact expression for the equilibrium behavior of our one-dimensional (d = 1) harmonic oscillator (Eq. 13).

$$\pi_x(x) \quad \propto \quad e^{-\beta U(x)} \tag{17}$$

$$\propto \quad e^{-\frac{\beta K}{2}x^2} \tag{18}$$

$$= \quad \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{19}$$

where $\mu = 0$ and $\sigma = (\beta K)^{-1/2} = \sqrt{k_B T/K}$. That means we can can plot the equilibrium distribution in positions:

We can also compute properties—such as dependence of the size of fluctuations on the temperature and spring constant, directly. The width of the distribution, σ, increases as the square-root of the temperature, and decreases as the square root of the spring constant. That means to confine our bead to half the fluctuation size in an optical laser trap, we'd have to quadruple the spring constant, or quadruple the laser power (assuming we're in the linear regime).

While it's hard to analytically solve many problems with continuous configuration spaces beyond this one—we'll need numerical methods and computers to study more complicated systems—we'll work some more problems a bit later using some simple *discrete*

systems that are still exactly solvable, and show how this can often be a useful enough model for a biological process that you can get a lot of mileage in understanding common mechanisms this way.

4.2 *Reduced potentials and general equilibrium densities*

It turns out the Boltzmann law for describing equilibrium distributions has an even simpler and more general form:

$$\pi(x) \propto e^{-u(x)} \tag{20}$$

where $u(x)$ is a dimensionless quantity called the *reduced potential* [1]. The reduced potential is an easy way to construct the equilibrium distribution for whatever **thermodynamic parameters** are held fixed in your circumstance. The general form of the reduced potential is

$$u(x) \equiv \beta \left[\sum_{i=1}^{N} \lambda_i U_i(x) + pV(x) + \sum_{m=1}^{M} \mu_m N_m(x) + \cdots \right] \tag{21}$$

This form captures the conjugate relationship between thermodynamic parameters—like inverse temperature β, external fields λ_i, pressure p, and chemical potentials μ_m for various components of the system—and their conjugate dynamical variables or properties that can fluctuate—such as potential energy components U_i, the volume of the system V, and the number of particles of a given type N_m. Using this general form, we can construct the equilibrium probability density for any situation of interest by just including the thermodynamic parameters that are held fixed during the experiment.

4.3 *Isothermal-isobaric ensemble*

In the simplest case, we may have a system with a fixed number of particles N where the box volume V is allowed to fluctuate at atmospheric external pressure p.
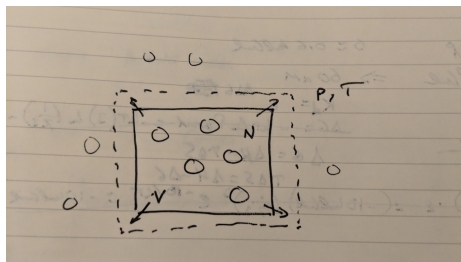
In this case, we form the reduced potential by including any terms where one of components of the term is allowed to fluctuate: This includes the inverse temperature β (where the potential $U(x)$ will fluctate) and the pressure p (where the volume $V(x)$ is allowed to fluctuate):

$$u(x) = \beta \left[U(x) + pV(x) \right] \tag{22}$$

4.4 *Ideal gas*

There's a simple illustrative case we can analytically solve: The ideal gas. In this case, the potential energy $U(x)$ is always zero, so we can eliminate the term that it appears in:

$$u(x) = \beta p V(x) \tag{23}$$

We are left with a probability density that can tell us everything about the system behaves

Suppose we have N particles in our ideal gas:

$$\pi(x) \propto e^{-\beta p V(x)} \tag{24}$$

we can make a transformation of variables and carefully integrate over all the configurations x where the volume $V(x)$ is equal to a desired volume v:

$$\pi(v) \quad \propto \quad \int dx \, e^{-\beta p V(x)} \, \delta(V(x) - v) \tag{25}$$

$$\propto \quad v^N e^{-\beta p v} \tag{26}$$

where the v^N factor is the Jacobian we pick up when transforming from the 3N-dimensional x to box volume v.

From this, we can compute the average volume of the ideal gas as a function of pressure and temperature!

$$\langle v \rangle_{\beta,p,N} = \int_0^\infty dv \, v \, \pi(v) = \frac{\int_0^\infty dv \, v \, v^N e^{-\beta p v}}{\int_0^\infty dv \, v^N e^{-\beta p v}} = \frac{N}{\beta p} = \frac{N k_B T}{p} \tag{27}$$

The $\langle A \rangle_{\beta,p,N}$ brackets mean the **expectation** or average value of a **mechanical observable** $A(x)$ with respect to the thermodynamic parameters in the lower-right corner—in this case, $\{\beta, p, N\}$. If we took many measurements of an ideal gas under these fixed thermodynamic conditions, we would see that the individual systems would have an array of different volumes—whose probability is described by $\pi(v)$—but the average would be $N k_B T/p$.

4.5 *Optical trap*

In a more complex example, suppose we have a laser optical trap, where one end of an assembly is anchored to a fixed wall and the other end is tethered to a polystyrene bead held in an optical laser trap, which acts pretty close to a harmonic potential on the bead center r:
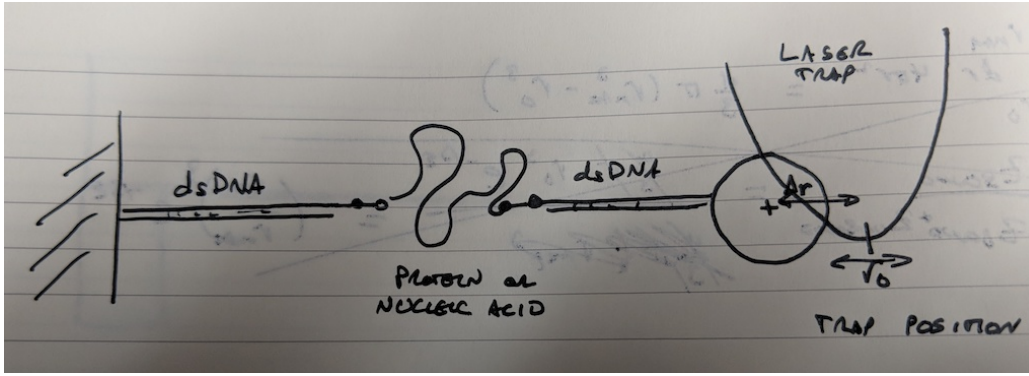
We can write the reduced potential for this system as

$$u(x; \beta, p, K, r_0) = \beta\left[U_0(x) + \frac{K}{2}(r - r_0)^2 + pV(x)\right] \tag{28}$$

Our instrument might allow us to modulate several thermodynamic parameters: the inverse temperature β, external pressure p, spring constant K (via laser power), and beam center r_0 (through a translation stage or mirrors). Some of these may be easier to experimentally control than others, but all of these will modulate the equilibrium distribution

$$\pi(x; \beta, K, r_0) \quad \propto \quad e^{-u(x;\beta,K,r_0)} \tag{29}$$

in a way that we can predict if we knew the potential energy function of the polymer being studied and surrounding solvent $U_0(x)$. Surprisingly, there are tricks we can play to *infer* the polymer's equilibrium properties by collecting data at many trap positions—a good example of this, where the free energy landscape of a dsDNA hairpin was inferred to sub-angstrom and sub-$k_B T$ resolution, can be found in [1].

4.6 *Chemical potential*

The chemical potential μ is the driving force that controls different components that can flow into and out of the system if you consider the outside to be semi-permeable to some components. This might be protons—since proteins and small molecules do not have fixed protonation states—where the chemical potential is determined by the pH. Alternatively, these might be salt pairs, where the chemical potential is determined by the macroscopic salt concentration in the cell or cuvette [2].

5 PARTITION FUNCTIONS AND PHASE EQUILIBRIA

5.1 *Partition functions*

Up to now, we've written probability densities of microscopic configurations as a proportionality relation:

$$\pi(x) \propto e^{-u(x)} \tag{30}$$

From now on, we're going to explicitly write the normalizing constant that ensures these probability densities integrate to unity:

$$\int dx\, \pi(x) = 1 \tag{31}$$

The normalizing constant has a special name in statistical mechanics—the **partition function**:

$$\pi(x; \theta) = Z(\theta)^{-1} e^{-u(x;\theta)} \implies Z(\theta) \equiv \int dx\, e^{-u(x;\theta)} \tag{32}$$

If θ is some collection of thermodyanmic parameters that characterize our equilibrium thermodynamic state, then the partition function $Z(\theta)$ is a function of those parameters. $Z(\theta)$ has dimensions of length to the 3N power, since it reflects in integral of the hypervolume accessible to the system. For different thermodynamic ensembles characterized by different collections of thermodynamic variables, the partition function might have a different name attached—such as the *canonical partition function* if β, N, V are fixed; or the *grand canonical partition function* if the particle number can vary and μ is fixed, but they're all the same idea.

For the one-dimensional harmonic oscillator, where $u(x) = \frac{\beta K}{2} x^2$, we can easily compute the partition function (since it's a Gaussian!):

$$Z(\beta) = \int dx\, e^{-u(x;\beta)} = \int dx\, e^{-\frac{\beta K}{2} x^2} = \sqrt{2\pi} \sqrt{\frac{k_B T}{K}} \tag{33}$$

If we're talking about the probability density over velocities as well, there's a contribution to the partition function that comes from the normalizing constant for the velocity part of the equilibrium distribution too, but that can be analytically computed and only depends on the system masses:

$$Z_v(\theta) \equiv \int dv\, e^{-\frac{\beta}{2} \sum_{n=1}^{N} m_n \|v_n\|^2} = (2\pi)^{3N/2} \prod_{n=1}^{N} \left(\frac{k_B T}{m_n} \right)^{3/2} \tag{34}$$

5.2 *Partition coefficients and phase equilibria*

Not to be confused with partition functions, **partition coefficients** describe the partitioning of molecules between different environments.

For example, suppose we have a drug molecule and want to measure how much it partitions between water and octanol, which is often used as a surrogate for lipophilicity, or how much the drug likes to permeate cell membranes:
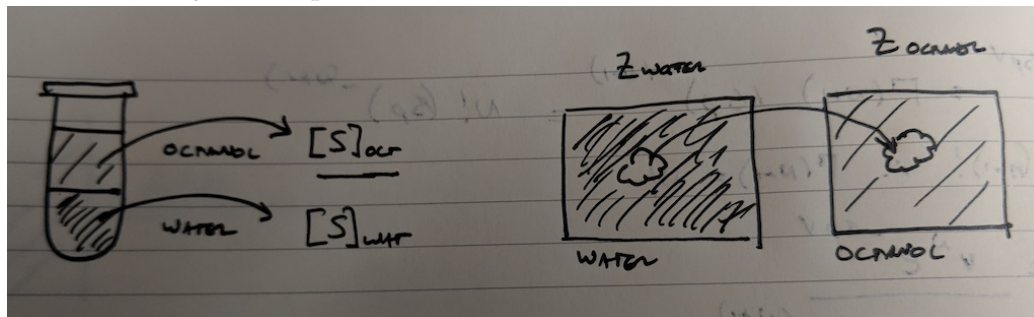
We have a few ways of measuring partitioning between these phases: The **partition coefficient** P (often appearing as $\log_{10} P$) gives the ratio of the neutral form of the compound between two phases:

$$\log_{10} P_{\text{oct/wat}} \equiv \log_{10} \frac{[S]_{\text{oct}}}{[S]_{\text{wat}}} \propto \frac{Z_{\text{octanol}}}{Z_{\text{water}}} \tag{35}$$

Surprisingly, these partition coefficients can also be computed from the ratio of partition coefficients for the small molecule in a solvent environment for each phase [3].

The concept of a **distribution coefficient** (D, often appearing as $\log_{10} D$) is related to the partition coefficient, and simply means that instead of using only the neutral form of the solute, we sum up the total concentration of all ionization states in each phase [4].

6 FREE ENERGY, ENTHALPY, AND ENTROPY

Partition functions are directly related to free energies. In fact, we *define* the **free energy** of a system in terms of partition functions:

$$F(\beta, \ldots) \equiv -\beta^{-1} \ln Z(\beta, \ldots) = -k_B T \ln \int dx\, e^{-\beta U(x)} \tag{36}$$

Defined in this manner, the free energy has dimensions of energy, and can only be compared at the same temperature T because of the $\beta^{-1} \equiv (k_B T)^{-1}$ that appears in front of $\ln Z$.

More generally, we can define a dimensionless free energy $f(\theta)$ we can use for any set of thermodynamic parameters, which can be thought of having units of thermal energy $(k_B T)$

$$f(\theta) \equiv -\ln Z(\theta) \tag{37}$$

We can convert back to dimensional quantities by just multiplying by $k_B T$ to add back in units of energy: $F(\theta) = k_B T\, f(\theta)$.

The **enthalpy** of the system is the *average energy* at a given thermodynamic state θ (again written as a dimensionless quantity):

$$h(\theta) \equiv \langle u \rangle_\theta = \int dx\, u(x)\, e^{-u(x;\theta)} \tag{38}$$

We can again convert back to dimensional quantities by multiplying by $k_B T$ to get the enthalpy with units of energy: $H(\theta) = k_B T\, h(\theta)$.

The **entropy** of the system can be computed from the relationship

$$f(\theta) = h(\theta) - s(\theta) \;\Rightarrow\; s(\theta) = h(\theta) - f(\theta) \tag{39}$$

When written in dimensional form, $S(\theta) = k_B\, s(\theta)$, since entropy is generally written in weird dimensions of energy divided by temperature (e.g. Joules/Kelvin) for obscure historical reasons.

6.1 *Simple model of ligand binding*

We can also write the free energy, enthalpy, and entropy *changes* for a process that involves changing the thermodynamic parameters (such as increasing the external pressure on a piston) or restricting parts of the system to specific regions of configuration space (such as a ligand binding a target receptor). While the free energy is a very useful quantity, since it tells us directly about the probability of finding the system somewhere (e.g. the fraction of time we will find the ligand bound), the enthalpy and entropy are only modestly useful, even though a great deal of noise has been made about them.

Consider a simple model of ligand binding in which two components have a distance-dependent potential where, if $r < r_0$, there is a favorable binding energy of $-\epsilon$, while for $r \geq r_0$, the interaction energy is zero.

$$U(r) = \begin{cases} -\epsilon & 0 \leq r < r_0 \\ 0 & r_0 \leq r < r_{max} \\ \infty & r \geq r_{rmax} \end{cases} \tag{40}$$
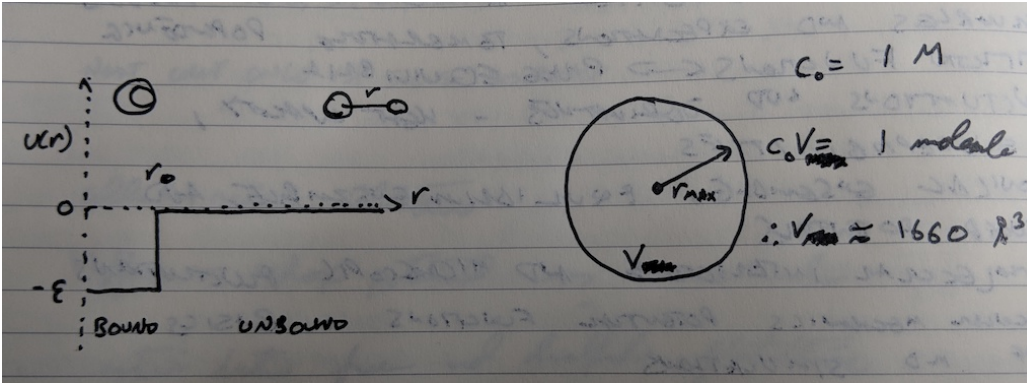
Recall that, for ligand binding, the dissociation constant K_d is given by

$$K_d \equiv \frac{[R][L]}{[RL]} \tag{41}$$

and that we can define a free energy ΔG for the binding process as

$$\Delta G \equiv -k_B T \ln \frac{K_d}{c_0} \tag{42}$$

where, since K_d has dimensions of concentration, we need to divide by some standard state concentration c_0 (generally taken to be $c_0 \equiv 1$ M) in order to be able to take the logarithm.

We can break the partition coefficient $Z(\beta, r_0, \epsilon)$ into two pieces—one for the bound

17

region, and one for the unbound region:

$$Z(\beta, r_0, \epsilon) = Z_{bound}(\beta, r_0, \epsilon) + Z_{unbound}(\beta, r_0, \epsilon) \tag{43}$$

$$Z_{bound}(\beta, r_0, \epsilon) = \int_0^{r_0} dr\, 4\pi r^2\, e^{\beta\epsilon} = \frac{4}{3}\pi r_0^3 e^{\beta\epsilon} \tag{44}$$

$$Z_{unbound}(\beta, r_0, \epsilon) = \int_{r_0}^{r_{max}} dr\, 4\pi r^2\, 1 = \frac{4}{3}\pi (r_{max}^3 - r_0^3) \tag{45}$$

The fraction of bound ligand can easily be expressed in terms of these partition functions:

$$P_{bound}(\beta, r_0, \epsilon) = \frac{Z_{bound}(\beta, r_0, \epsilon)}{Z_{bound}(\beta, r_0, \epsilon) + Z_{unbound}(\beta, r_0, \epsilon)} \tag{46}$$

$$= \frac{\frac{4}{3}\pi r_0^3 e^{\beta\epsilon}}{\frac{4}{3}\pi r_0^3 e^{\beta\epsilon} + \frac{4}{3}\pi(r_{max}^3 - r_0^3)} \tag{47}$$

$$= \frac{1}{1 + \left[\left(\frac{r_{max}}{r_0}\right)^3 - 1\right]e^{-\beta\epsilon}} \tag{48}$$

We can also compute the (dimensionless) free energy difference between the bound and unbound states:

$$\Delta f(\beta, r_0, \epsilon) \tag{49}$$

$$\equiv f_{bound}(\beta, r_0, \epsilon) - f_{unbound}(\beta, r_0, \epsilon) = -\ln Z_{bound}(\beta, r_0, \epsilon) + \ln Z_{unbound}(\beta, r_0, \epsilon) \tag{50}$$

$$= -\ln\frac{4}{3}\pi - 3\ln r_0 - \beta\epsilon + \ln\frac{4}{3}\pi + \ln(r_{max}^3 - r_0^3) \tag{51}$$

$$= -\beta\epsilon + \ln\left[\left(\frac{r_{max}}{r_0}\right)^3 - 1\right] = -\beta\epsilon + \ln c \tag{52}$$

where we have defined the constant $c = \left[\left(\frac{r_{max}}{r_0}\right)^3 - 1\right]$ since it doesn't depend on any thermodynamic parameters. We can go on to compute the entropy and enthalpy, and see how these behave as a function of β, r_0, and ϵ. The enthalpy of the whole system,

$$h(\beta, r_0, \epsilon) = \int_0^{\infty} dr\, 4\pi r^2\, u(r)\, \pi(r) = -\epsilon \cdot P_{bound}(\beta, r_0, \epsilon), \tag{53}$$

can be split into separate parts for bound and unbound, and we can compute the enthalpy

change upon binding.

$$h_{\text{bound}}(\beta, r_0, \epsilon) \quad = \quad \int_0^{r_0} dr\, 4\pi r^2\, u(r)\, \pi(r) = -\epsilon \cdot P_{\text{bound}}(\beta, r_0, \epsilon) \tag{54}$$

$$h_{\text{unbound}}(\beta, r_0, \epsilon) \quad = \quad \int_{r_0}^{\infty} dr\, 4\pi r^2\, u(r)\, \pi(r) = 0 \tag{55}$$

$$\Delta h_{\text{bound}}(\beta, r_0, \epsilon) \quad \equiv \quad h_{\text{bound}}(\beta, r_0, \epsilon) - h_{\text{unbound}}(\beta, r_0, \epsilon) \tag{56}$$

$$= \quad -\epsilon \cdot P_{\text{bound}}(\beta, r_0, \epsilon) \tag{57}$$

$$\tag{58}$$

Finally, the entropy change can be computed by noting from Eq. 39,

$$\Delta s \quad = \quad \Delta h - \Delta f \tag{59}$$

$$= \quad -\epsilon \cdot P_{\text{bound}}(\beta, r_0, \epsilon) + \beta\epsilon - \ln c \tag{60}$$

$$= \quad -\frac{\epsilon}{1 + \alpha e^{-\beta\epsilon}} + \beta\epsilon - \ln c \tag{61}$$

7 THE FUNDAMENTAL LAWS OF THERMODYNAMICS

As the poet Alan Ginsburg put it,

1. You can't win.

2. You can't break even.

3. You can't even get out of the game.

Surprisingly, this neatly sums up the laws of thermodynamics in a useful way. We'll explore these in more detail.

7.0.1 Zeroth law: Two systems in equilibrium with a third system are in equilibrium with each other

This is the thermodynamic equivalent of the transitive property of equality: If $a = b$ and $b = c$, then $a = c$. Similarly, if I allow two systems A and B to come into thermal equilibrium, and also B and C are in thermal equilibrium, then A and C must also be in thermal equilibrium. This is a good trick for finding flaws in designs for perpetual motion machines.

7.0.2 First law: Energy can neither be created nor destroyed.

If we consider a closed "universe" that contains a system of interest in contact with or isolated from some larger macroscopic environment that is big enough to be considered as having some fixed thermodynamic parameters (like temperature), energy has to go *somewhere*—it just can't be created or destroyed—at least not in classical statistical mechanics. The only way the total energy $E(x, v) = U(x) + \mathcal{K}(v)$ of our system can change is through **work** we do on the system (or it does on us), or through **heat** exchanged with the environment.
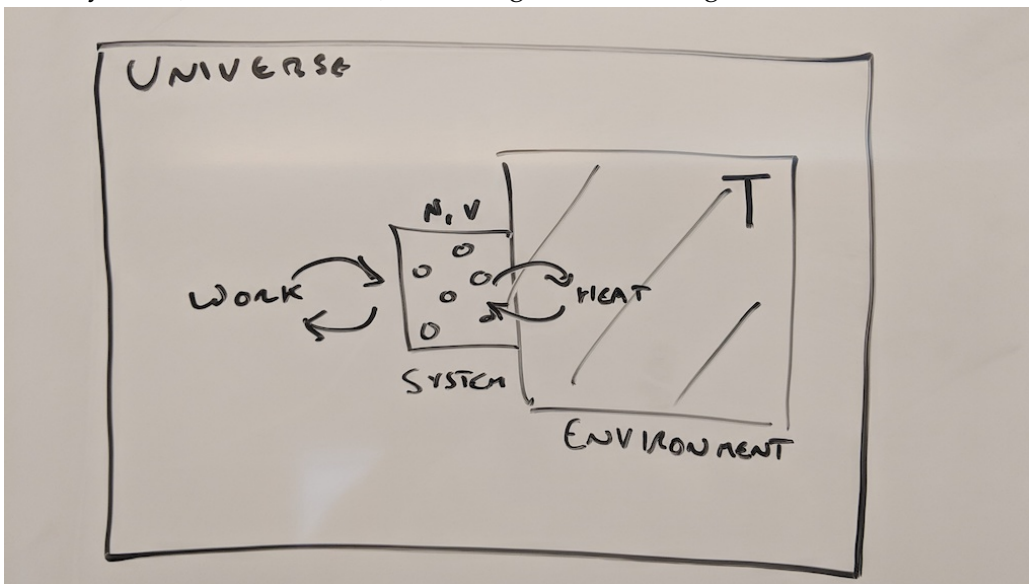
7.0.3 Second law: The entropy of an isolated system will tend to increase over time, approaching a maximum value at equilibrium

On average, when you prepare the system with some initial configuration and let it evolve in time, the entropy of the system will increase. That doesn't mean that it will *always* increase each time you try this experiment—some times, the entropy could surprisingly *decrease* as all the particles decide to go into one corner! But *on average*, if you do this many times or for many noninteracting copies of the same system, entropy will increase with time.
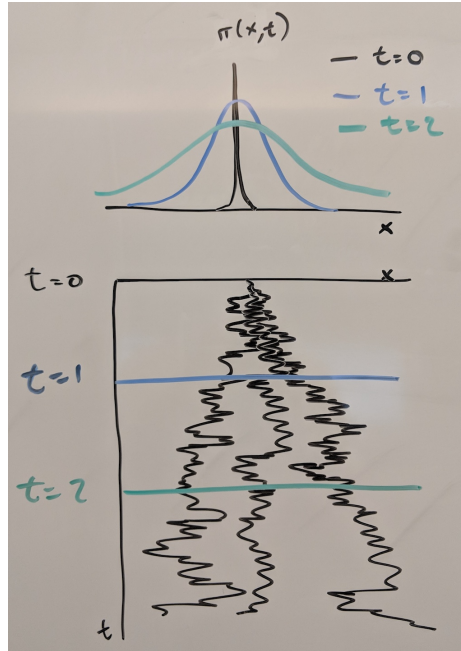
Surprisingly, only in the last couple of decades was a new relationship discovered that allows us to *exactly* relate systems driven out of equilibrium with true equilibrium properties in a remarkable way. Two mathematical theorems—called the Jarzynski equality and the Crooks fluctuation theorem, after their discoverers—allow us to say *exact* things about the distributions of system driven out of equilibrium; we'll examine this in more detail later (time permitting).

7.0.4 Third law: As temperature approaches absolute zero, the entropy of a system approaches a minimum

For classical mechanics, this minimum is zero entropy, provided this is done so slowly as to anneal the system into a unique solid crystalline ground state.

In real systems, it's sometimes very hard to cool them down slowly enough to reach a unique solid ground state, and instead we get an *amorphous solid*, or glass. This occurs when the barrier heights get so large compared to thermal energy ($k_B T$) that barrier crossing events become exponentially rare, and the system simply *can't* reach the global minimum ground state. For aqueous protein solutions, this happens around 200 K, where large-scale motions of proteins are highly suppressed and simple harmonic vibrations dominate [5]. Structural biologists exploit this property of proteins for both cryocooled X-ray spectroscopy and especially cryo-electron microscopy (cryo-EM): There, a tiny grid with samples suspended in thin liquid sheets in the holes of the grid is plunged into liquid ethane at < 100 K [6] to rapidly freeze out large-scale conformational motions and trap the biomolecule in conformations close to equilibrium at the initial temperature.

Quantum systems are weird, and can still have weird ground-state effects even at zero temperature, but since zero absolute temperature is not incredibly relevant for biology, we won't go into that.

7.1   *The sum of free energy differences around a thermodynamic cycle* must *equal zero.*

Free energy is a state function—no matter what sequence of thermodynamic states you go through and sum up free energy differences, you should arrive at the same free energy irrespective of which path you took. That also means that the free energy changes around a thermodynamic cycle should sum to zero. This can have very important consequences for protein conformational landscapes or ligand binding reactions!
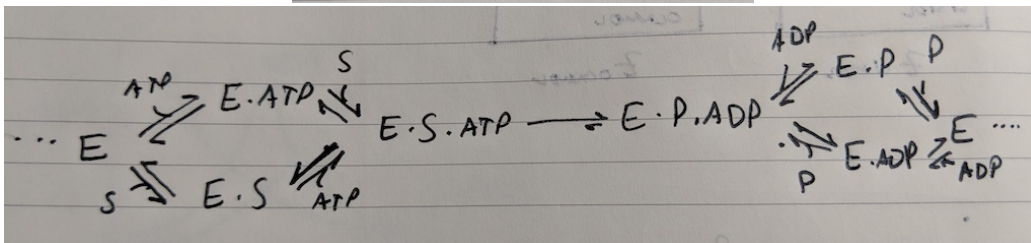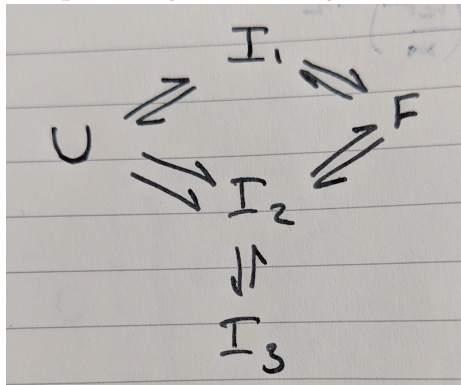
8   OBSERVABLES, EXPECTATIONS, AND SUSCEPTIBILITIES

### 8.1   *Expectations of observables*

Observables are functions of $A(x)$ that we might be able to measure experimentally. For example, we might be able to make a measurement of the total volume of a liquid as part of the estimation of its density, or perhaps the FRET efficiency between two fluorophores, or the fluorescence intensity of tryptophan emission. Generally, we can only measure equilibrium averages of these quantities over many molecules in a cuvette (unless we are doing single-molecule experiments, in which case we still often have to average over some time interval). The average of this observable over many copies or long times is called an **expectation**:

$$\langle A \rangle_\theta \equiv \int dx \, A(x) \, \pi(x; \theta) \tag{62}$$

These expectations will depend on the thermodynamic parameters θ imposed during the measurement.

Suppose we have something like a protein in a crystal. What's the average potential energy of the crystal? We can model this to a simple first approximation by assuming

that each atom in the crystal oscillates around a single stationary equilibrium point in an independent harmonic oscillator in three dimensions—this is called an **Einstein crystal**.
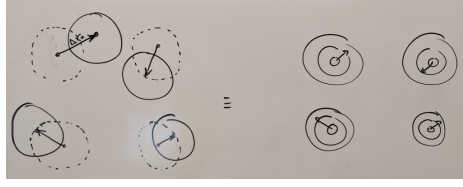
For example, we can compute the expectation of the system energy for an array of three-dimensional harmonic oscillators, where the total system potential energy might be described by

$$U(x) \;\equiv\; \sum_{n=1}^{N} \frac{K_n}{2}\|\Delta r_n\|^2 \tag{63}$$

The average unit-bearing potential energy can be computed

$$\langle U\rangle_\beta \;=\; \int_{-\infty}^{+\infty} dx\, U(x)\, Z^{-1}(\beta) e^{-\beta U(x)} \tag{64}$$

$$=\; \int_{-\infty}^{+\infty} dx\, \left[\sum_{n=1}^{N}\frac{K_n}{2}\|\Delta r_n\|^2\right] Z^{-1}(\beta) e^{-\beta \sum_{n'=1}^{N} \frac{K_{n'}}{2}\|\Delta r_{n'}\|^2} \tag{65}$$

$$=\; \int_{-\infty}^{+\infty} dx\, \left[\sum_{n=1}^{N}\frac{K_n}{2}\|\Delta r_n\|^2\right] \prod_{n'=1}^{N} Z_{n'}^{-1} e^{-\beta \frac{K_{n'}}{2}\|\Delta r_{n'}\|^2} \tag{66}$$

$$=\; \sum_{n=1}^{N}\frac{K_n}{2} \int_{-\infty}^{+\infty} dx\, \|\Delta r_n\|^2 \prod_{n'=1}^{N} Z_{n'}^{-1} e^{-\beta \frac{K_{n'}}{2}\|\Delta r_{n'}\|^2} \tag{67}$$

$$=\; \sum_{n=1}^{N}\frac{K_n}{2} \int_{-\infty}^{+\infty} dx\, \|\Delta r_n\|^2 \, Z_n^{-1} e^{-\beta \frac{K_n}{2}\|\Delta r_n\|^2} \tag{68}$$

$$=\; \sum_{n=1}^{N}\frac{K_n}{2} \int_{0}^{+\infty} dr_n\, 4\pi r_n^2\, r_n^2\, \frac{1}{(2\pi)^{3/2}\sigma^3} e^{-\frac{r_n^2}{2\sigma_n^2}} \tag{69}$$

$$=\; \sum_{n=1}^{N}\frac{K_n}{2} 4\pi \int_{0}^{+\infty} dr_n\, r_n^4\, \frac{1}{(2\pi)^{3/2}\sigma^3} e^{-\frac{r_n^2}{2\sigma_n^2}} \tag{70}$$

$$=\; \sum_{n=1}^{N}\frac{K_n}{2} 3\sigma^2 \tag{71}$$

$$=\; \sum_{n=1}^{N}\frac{3}{2}K\sigma^2 = \sum_{n=1}^{N}\frac{3}{2}k_B T = \frac{3}{2}Nk_B T \tag{72}$$

23

where we recall we write $\sigma^2 = (\beta K)^{-1}$ for simplicity.

In fact, it's a good rule of thumb to remember that, if there are a lot of independent harmonic oscillators, the average potential energy picks up roughly $k_B T/2$ per degree of freedom. In this case, there are N particles, and each has three degrees of freedom—the three coordinate axes x, y, and z.

What about the kinetic energy? Recall that the kinetic energy is just given in terms of the masses and velocities:

$$\mathcal{K}(v) = \frac{1}{2}\sum_{n=1}^{N} m_n \|v_n\|^2 \tag{73}$$

We can compute the average kinetic energy in much the same way:

$$\langle \mathcal{K} \rangle_\beta = \int_{-\infty}^{+\infty} dv\, \mathcal{K}(v)\, Z^{-1}(\beta) e^{-\beta \mathcal{K}(x)} \tag{74}$$

$$= \int_{-\infty}^{+\infty} dv \left[\sum_{n=1}^{N} \frac{m_n}{2}\|\Delta v_n\|^2 \right] Z^{-1}(\beta) e^{-\beta \sum_{n'=1}^{N} \frac{m_{n'}}{2}\|\Delta v_{n'}\|^2} \tag{75}$$

In fact, we realize we have *exactly* the same integral as before, with positions x replaced with velocities v and spring constants K_n replaced with masses m_n. We can then write down the result directly

$$\langle \mathcal{K} \rangle_\beta = \frac{3}{2}N k_B T \tag{76}$$

Surprisingly, the average kinetic and potential energies are equal!

8.2 *Fluctuation properties and susceptibilities*

We can also compute **fluctuation properties**, which measure how much some observable (like the energy) fluctuates about its average. There is a surprising relationship between fluctuation properties and **susceptibility properties**, which measure how much the average for that property changes as a thermodynamic parameter is varied. For example, consider the heat capacity, which measures how much heat (total energy change) is required to raise the temperature by one unit.

For the Einstein crystal, this is simply given by

$$\frac{d}{dT}\langle E \rangle_\beta = \frac{d}{dT}[3N k_B T] = 3N k_B \tag{77}$$

It's a constant for the Einstein crystal, which is always in a solid phase.

There's a surprising (but surprisingly useful!) relationship between the **susceptibility** of a property $\langle A \rangle_\beta$ to changes in temperature—that is, the measure of how much it changes when you change a thermodynamic parameter θ—and the magnitude of fluctuations in

that property at fixed θ. Let's take a look at this:

$$\frac{\partial}{\partial\theta}\langle A\rangle_\theta \;=\; \frac{\partial}{\partial\theta}\int dx\, A(x)\, Z(\theta)^{-1}\, e^{-u(x;\theta)} \tag{78}$$

$$=\; \int dx\, A(x)\, \frac{\partial}{\partial\theta}\Big[Z(\theta)^{-1}\, e^{-u(x;\theta)}\Big] \tag{79}$$

$$=\; \int dx\, A(x)\, \Big[Z(\theta)^{-1}\frac{\partial}{\partial\theta}e^{-u(x;\theta)} + \Big(\frac{\partial}{\partial\theta}Z(\theta)^{-1}\Big)e^{-u(x;\theta)}\Big] \tag{80}$$

Where we have simply used the chain rule for derivatives in the last line. Let's work out the partial derivatives individually. The first partial derivative simply requires we apply the chain rule again:

$$\frac{\partial}{\partial\theta}e^{-u(x;\theta)} \;=\; -\frac{\partial u}{\partial\theta}e^{-u(x;\theta)} \tag{81}$$

The second partial derivative requires we use the chain rule and substitute the definition of the partition function

$$\frac{\partial}{\partial\theta}Z(\theta)^{-1} \;=\; -Z(\theta)^{-2}\frac{\partial}{\partial\theta}Z(\theta) \tag{82}$$

$$=\; -Z(\theta)^{-2}\frac{\partial}{\partial\theta}\int dx\, e^{-u(x;\theta)} \tag{83}$$

$$=\; -Z(\theta)^{-2}\int dx\, \frac{\partial}{\partial\theta}e^{-u(x;\theta)} \tag{84}$$

$$=\; Z(\theta)^{-2}\int dx\, \frac{\partial u}{\partial\theta}e^{-u(x;\theta)} \tag{85}$$

$$=\; Z(\theta)^{-1}\Big\langle\frac{\partial u}{\partial\theta}\Big\rangle_\theta \tag{86}$$

Putting this all together, we obtain

$$\frac{\partial}{\partial\theta}\langle A\rangle_\theta \;=\; \int dx\, A(x)\, \Big[-Z(\theta)^{-1}\frac{\partial u}{\partial\theta}e^{-u(x;\theta)} + Z(\theta)^{-1}\Big\langle\frac{\partial u}{\partial\theta}\Big\rangle_\theta e^{-u(x;\theta)}\Big] \tag{87}$$

$$=\; -\Big[\Big\langle A\frac{\partial u}{\partial\theta}\Big\rangle_\theta - \langle A\rangle_\theta\Big\langle\frac{\partial u}{\partial\theta}\Big\rangle_\theta\Big] \tag{88}$$

$$=\; -\mathrm{cov}_\theta\Big(A, \frac{\partial u}{\partial\theta}\Big) \tag{89}$$

Now, consider what happens when we are interested in the potential energy component of the constant-volume heat capacity, where $A(x) = U(x)$, and $\theta = T$, given $u(x) \equiv \beta U(x)$ (and recalling $\beta \equiv (k_B T)^{-1}$ so that $\partial\beta/\partial T = -1/(k_B T^2)$):

$$\frac{\partial u}{\partial T} \;=\; \frac{\partial}{\partial T}[\beta U(x)] \tag{90}$$

$$=\; -\frac{1}{k_B T^2}U(x) \tag{91}$$

We therefore have

$$\frac{\partial}{\partial T} \langle u \rangle_\theta = -\left[\left\langle u \left(-\frac{1}{k_B T^2} \beta u \right) \right\rangle_\theta - \langle u \rangle_\beta \left\langle -\frac{1}{k_B T^2} u \right\rangle_\beta \right] \tag{92}$$

$$= \frac{1}{k_B T^2} \left[\langle u^2 \rangle_2 - \langle u \rangle_\beta^2 \right] = \frac{1}{k_B T^2} \mathrm{var}_\beta(u) \tag{93}$$

Surprisingly, the susceptibility of the average potential energy (the heat capacity) is related to the magnitude of fluctuations in the potential energy!

For the Einstein crystal, that means we can easily compute the magnitude of fluctuations in the potential energy:

$$\langle u^2 \rangle_2 - \langle u \rangle_\beta^2 = k_B T^2 \frac{\partial}{\partial T} \langle u \rangle_\theta = k_B T^2 \left[\frac{3}{2} N k_B \right] = \frac{3}{2} N (k_B T)^2 \tag{94}$$

In fact, if we have a generic thermodynamic state where we can write $u(x; \theta)$ as a linear sum $u(x; \theta) \equiv \sum_{n=1}^{N} \theta_i A_i(x)$, we have the special case that $\partial u / \partial \theta_i = A_i(x)$ and this relationship is

$$\frac{\partial}{\partial \theta_j} \langle A_i \rangle_\theta = -\mathrm{cov}_\theta \left(A_i, A_j \right) \tag{95}$$

which is related to the *thermodynamic metric tensor* that describes how much things change as the thermodynamic parameter θ_j is tweaked. This has important consequences for the design of optimal single-molecule or biomolecular simulation experiments [7].

8.3 *Phase transitions*

For real substances we deal with in the laboratory, we can typically access solid, liquid, and gas phases by putting in or taking out heat. At each phase transition boundary, putting in more heat will cause the system to stop increasing in temperature for a while as the system absorbs enough energy to reconfigure its internal structure to switch phases. For solid to liquid transitions, this quantity of heat is called the **enthalpy of fusion**, and is about 1.4 kcal/mol for water at 0 C and atmosphere pressure; for liquid to gas, this is called the **enthalpy of vaporization**, and is about 12 kcal/mol for water at 100 C. The heat capacity of water varies over the 0–100 C range, but is roughly 1 cal to raise 1 g of water by 1 K.

Different phases are accessible at equilibrium with different combinations of thermodynamic parameters:
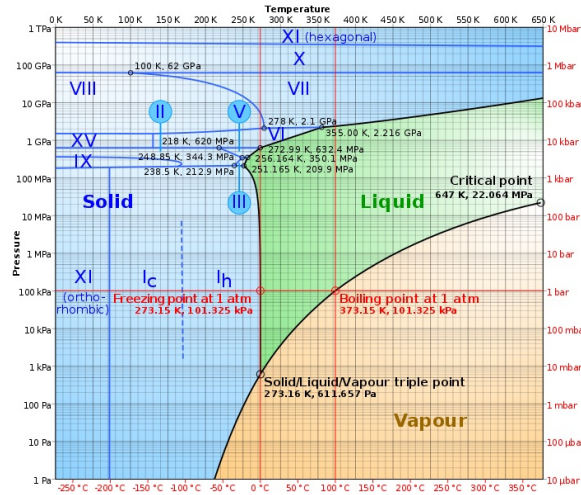
The phase diagram of water (CC BY-SA 3.0, made by Wikipedia user Cmglee) shows solid, liquid, and vapour (gas) phases.

9 A MICROSCOPIC VIEW OF ENTROPY

To better understand what the concept of *entropy* really means microscopically, consider this two-state one-dimensional system, which might be a model for ligand binding or protein folding:
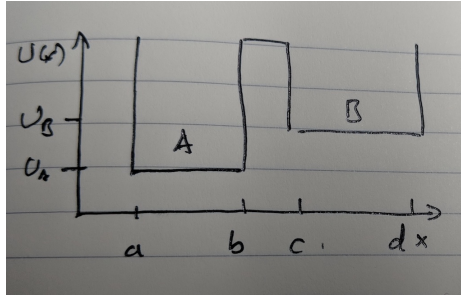
Let's compute the behavior of this system as a function of temperature.

First, let's assume the barrier is large, and compute the partition functions of the two states A and B:

$$Z_A(\beta) = \int_a^b dx \, e^{-\beta U_A} = e^{-\beta U_A} \int_a^b dx = e^{-\beta U_A} (b - a) \tag{96}$$
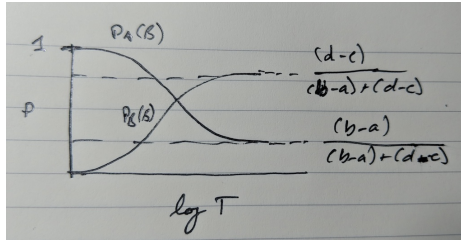
$$Z_B(\beta) = \int_c^d dx \, e^{-\beta U_B} = e^{-\beta U_B} \int_c^d dx = e^{-\beta U_B} (d - c) \tag{97}$$

We can use the partition functions to compute the probability of being in state A or B as a function of inverse temperature β:

$$p_A(\beta) = \frac{Z_A(\beta)}{Z_A(\beta) + Z_B(\beta)} = \frac{e^{-\beta U_A} (b - a)}{e^{-\beta U_A} (b - a) + e^{-\beta U_B} (d - c)} \tag{98}$$

$$p_B(\beta) = \frac{Z_A(\beta)}{Z_A(\beta) + Z_B(\beta)} = \frac{e^{-\beta U_B} (d - c)}{e^{-\beta U_A} (b - a) + e^{-\beta U_B} (d - c)} \tag{99}$$

27

If we plot these, it looks something like this:

As $T \to 0$, then the inverse temperature $\beta \to \infty$ and the lowest-energy state dominates the probability. On the other hand, as $T \to 0$, the exponential terms all become unity, and the probabilities of each state approach the proportion of their volumes in configuration space.

Now, let's compute the free energies of the two states:

$$F_A(\beta) \;=\; -k_B T \ln Z_A(\beta) = -k_B T \ln\left[e^{-\beta U_A}\,(b-a)\right] = U_A - T\,k_B \ln(b-a) \quad (100)$$

$$F_B(\beta) \;=\; -k_B T \ln Z_B(\beta) = -k_B T \ln\left[e^{-\beta U_B}\,(d-c)\right] = U_B - T\,k_B \ln(d-c) \quad (101)$$

Recall the relationship between free energy and enthalpy and entropy:

$$F = U - T\,S \tag{102}$$

You can see that the $H = U$ term can be identified with the *enthalpy* contribution for the state, while the $S = k_B \ln W$ terms (where W is the width of the state) can be identified with the entropy of the state. In fact, in a cemetery outside of Vienna, Ludwig Boltzmann has exactly this carved on his gravestone!

We can now compute the free energy difference between the states:

$$\Delta F(\beta) \;\equiv\; F_A(\beta) - F_B(\beta) = (U_A - U_B) - T\,k_B \ln \frac{b-a}{d-c} \tag{103}$$

where the enthalpy difference is $\Delta H = (U_A - U_B)$ and the entropy difference is $\Delta S = k_B \ln \frac{b-a}{d-c}$.

Microscopically, entropy is a logarithmic measure of the number of microscopic states the system has accessible to it. The logarithm is importance since taking two different systems that are completely independent as a collection *multiplies* the number of microstates accessible to them, but the entropy is *additive*. In this simple example, the entropy was a measure of the width (or 1D volume) of a flat basin. More generally, the entropy is computed

from a normalized probability density $p(x)$ as

$$S[p(x)] \equiv -\int dx\, p(x) \ln p(x) \qquad (104)$$

or for discrete systems with discrete probabilities p_i for $i \in \{1, 2, \ldots, N\}$,

$$S[p] \equiv -\sum i = 1^N p_i \ln p_i \qquad (105)$$

This microscopic thermodynamic definition of entropy is actually identical to the concept of information entropy!

9.1 Lattice model of protein folding

We can use this concept to understand the behavior of a simple model of protein folding where all the states can be exactly enumerated, so the complete behavior can be understood. Consider this example from Dill 2ed, Fig 8.6 (pg. 136):
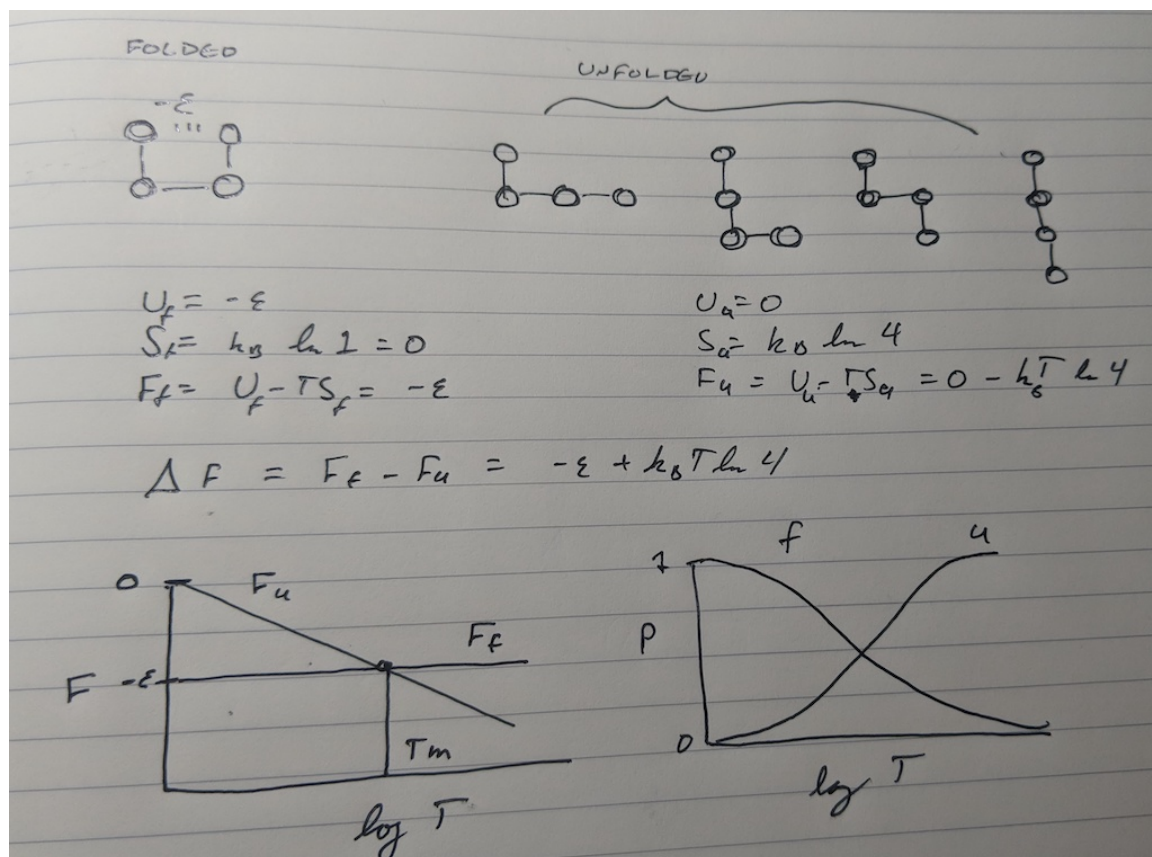
In this very simple lattice model, we can use this definition of entropy to compute the enthalpy and entropy separately, summing them to compute the free energy for folded and unfolded states. Generally, it's not possible to do this directly for complex continuous models, but fortunately, there are easier ways to compute free energy differences by estimating ratios of partition coefficients directly. We'll hear more about that shortly.

* * *

REFERENCES

[1] Michael R Shirts and John D Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of chemical physics*, 129(12):124105, 2008.

[2] Gregory A Ross, Ariën S Rustenburg, Patrick B Grinaway, Josh Fass, and John D Chodera. Biomolecular simulations under realistic macroscopic salt conditions. *The Journal of Physical Chemistry B*, 122(21):5466–5486, 2018.

[3] Caitlin C Bannan, Gaetano Calabró, Daisy Y Kyu, and David L Mobley. Calculating partition coefficients of small molecules in octanol/water and cyclohexane/water. *Journal of chemical theory and computation*, 12(8):4015–4024, 2016.

[4] Ariën S Rustenburg, Justin Dancer, Baiwei Lin, Jianwen A Feng, Daniel F Ortwine, David L Mobley, and John D Chodera. Measuring experimental cyclohexane-water distribution coefficients for the sampl5 challenge. *Journal of computer-aided molecular design*, 30(11):945–958, 2016.

[5] Dagmar Ringe and Gregory A Petsko. The 'glass transition' in protein dynamics: what it is, why it occurs, and how to exploit it. *Biophysical chemistry*, 105(2-3):667–680, 2003.

[6] Christopher J Russo, Steve Scotcher, and Martin Kyte. A precision cryostat design for manual and semi-automated cryo-plunge instruments. *Review of Scientific Instruments*, 87(11):114302, 2016.

[7] Gavin E Crooks. Measuring thermodynamic length. *Physical Review Letters*, 99(10):100602, 2007.