Appendix C extra - Concepts of statistical physics

A conventional introductory course in physics concentrates on one-body motion: the change in the position **x** and momentum **p** of a single object subject to a force **F**. Newton's second law tells us how momentum evolves through

$$\mathbf{F} = d\mathbf{p} / dt, \tag{Cx1.1}$$

where vector quantities are denoted in **boldface**. The change in position can be found by integrating the velocity $\mathbf{v} = \mathbf{p}/m$, where m is the object's mass. We use momentum, rather than velocity, to characterize an object's motion in order to make a smoother connection with quantum mechanics and relativity; for the situations of interest here, the only difference between velocity and momentum is the multiplicative scalar m. Although Eq. (Cx1.1) refers to translational motion, the rotation of a rigid body can be described by means of a similar equation for the change of its angular momentum in response to an applied torque.

In this course, we deal with a collection of interacting objects, not just one in isolation. The objects exchange energy with each other and their environment and are usually not in their lowest energy state (which might be a regular crystal, for example). Further, materials of biological interest are generally flexible - not rigid with a fixed shape. How do we describe such interacting many-body systems?

One approach is just to invoke as many copies of Eq. (Cx1.1), and its rotational cousin, as are needed to account for all the positions, orientations, momenta *etc.* of the objects in the system, and then use a computer to numerically integrate their equations of motion. This may be overkill if we are more interested in the macroscopic properties of a collection of objects (for example, the volume of a gas) than the detailed motion of the individual components (such as the molecules making up the gas). An alternative, then, is to describe the collective behavior of a system, for which statistical mechanics provides an appropriate formalism.

Description of a state

Our starting point in statistical mechanics is the definition of a state of the system. For an ensemble of N identical point-like objects, a specific *microstate* is a set of values for the $6N$ variables

positions: $\mathbf{x}_1 \ldots \mathbf{x}_N$
momenta: $\mathbf{p}_1 \ldots \mathbf{p}_N$

where we have neglected other characteristics such as spin or composition. The position labels $\mathbf{x}_1 \ldots \mathbf{x}_N$ need not be cartesian coordinates - they could be polar angles *etc.* To accommodate an angular representation of the object locations, the label $\mathbf{x}_i$ is often replaced by $\mathbf{q}_i$, which are the coordinates of object i. Note that the number of independent labels needed to describe the location of an object does not change with the representation: in three dimensions, each object requires:

three cartesian coordinates x_i, y_i, z_i

or

three polar coordinates r_i, θ_i, ϕ_i (θ, ϕ are angles)

The collection of coordinates q are called the degrees of freedom f of the system. Here, there are $3N$ coordinates so $f = 3N$. The coordinates q and momenta p of the objects form the *phase space* of the system.

Positions and momenta don't tell the whole story. Let's consider as an example the cell's energy currency ATP (adenosine triphosphate). As a molecule, ATP can provide energy in a biological process through its hydrolysis to ADP (adenosine diphosphate). Thus, to describe the energy available to the cell through this reaction, one would need an inventory of how many ATP molecules there are, and how many ADP. That is, a collection of these molecules would need another set of labels beyond q_i and p_i.

Other labels may be quantum mechanical in origin. Most readers are already familiar with the concept of a quantum number. For example, the energy of a one-dimensional harmonic oscillator is given by

$$E = (n + 1/2)\, \omega h\, /\, 2 \ . \qquad\qquad\qquad\qquad (Cx1.2)$$

where ω is the angular frequency of the oscillator and h is Planck's constant. The energy of the oscillator is specified by the quantum number n which is permitted to have the integer values 0 to   . For a collection of quantum oscillators, the number of degrees of freedom is the set of quantum numbers required to specify the state. For example:

|		
1 harmonic oscillator	$E = (n + 1/2)\, \omega h\, /\, 2$	$f = 1$
N weakly coupled oscillators	$E = (n_1 + n_2 \ldots + N/2)\, \omega h\, /\, 2$	$f = N$

Summary   A microscopic state or *microstate* of the system is a specified state of
• phase space coordinates
• quantum numbers
• other labels such as spin or chemical composition.
Each microstate has a unique energy. However, the reverse is not true: more than one microstate may have the same energy. The number of microstates with the same specified energy is referred to as the degeneracy.


Accessible states and the fundamental postulate

Knowing the physics governing each object, we can determine the energy of a given microstate. Let's consider a grossly simplified model of RNA folding in one dimension. As an example, we take a short RNA molecule of just three bases. As you know, the ends of the string of bases are inequivalent, so there is a direction to the string. In one dimension, the allowed configurations look like:
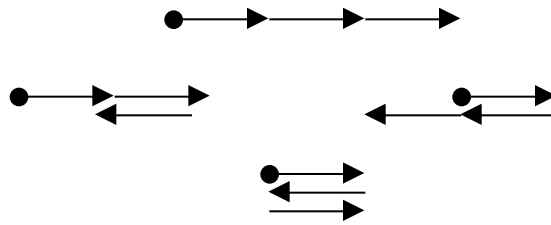
We have shown only those configurations where the first subunit of the string points to the right, starting from the disc: there is an equivalent set where the first unit points to the left. In this collection, there is one unfolded state, two states with a single fold, and one state with a double fold.   The 1:2:1 ratio of folded states is unchanged if we include the four "left pointing" configurations as well.

What is the energy of these states? Suppose that two subunits have a *lower* energy when they are bound: that is, it takes an amount of energy E_{bind} to undo the fold. Then the energies of the configurations are

$E = 0$

$E = -E_{bind}$

$E = -2E_{bind}$
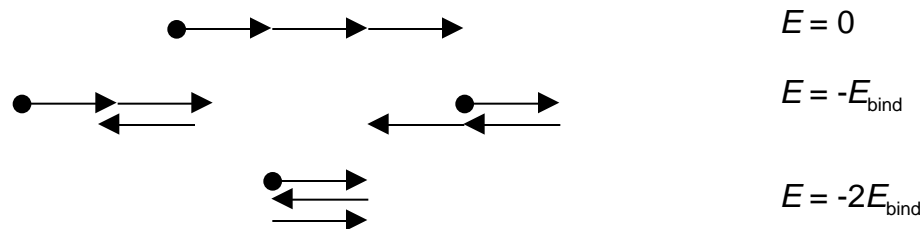
The minus sign shouldn't trouble us, it only indicates that energy is required to undo the fold. That is, the completely unfolded configuration has a higher energy than a state with one fold:

$$E_{unfold} - E_{1\text{-}fold} = 0 - (-E_{bind}) = +E_{bind} \qquad \text{(Cx1.3)}$$

Now let's think about the probability of finding a system in one if its microstates. Out of a (potentially large) set of states, there exist a set of *accessible states* of the system, specified by a global system characteristic such as the energy. In our 3-base RNA example:

configuration	energy	# of accessible states in energy range
no folds	0	1
one fold	$-E_{bind}$	2
two folds	$-2E_{bind}$	1

where the "energy range" is taken to be small compared to E_{bind}. In this example, the system in isolation can only have three values of the energy. More generally, say a gas of 10^{23} molecules, there may be a huge number of accessible states in a given energy range.

How do we know in which accessible state or states a system will be found? The fundamental postulate of statistical mechanics is:

An isolated system in equilibrium is equally likely to occupy any of its accessible states.

Note the choice of words!
1. *Isolated*: this means that there is no exchange of energy or objects with other systems.
2. *Equilibrium*: the system will evolve in time if it is not in equilibrium.

Example:
Consider a classical harmonic oscillator in one dimension. Does the fundamental postulate mean that the oscillator (or an ensemble of isolated oscillators, each with its own phase space) have equal likelihood of having any **spatial** position *x* allowed by conservation of energy? **NO**. Why?

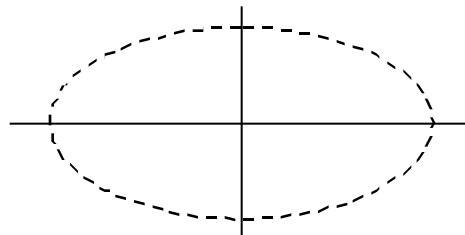The phase space of the oscillator is elliptical, owing to conservation of energy:

$$E = \frac{p^2}{2m} + \frac{kx^2}{2}$$ (Cx1.4)

or

$$1 = \frac{p^2}{2mE} + \frac{kx^2}{2E} = \left(\frac{p}{p_{max}}\right)^2 + \left(\frac{x}{x_{max}}\right)^2$$ (Cx1.5)
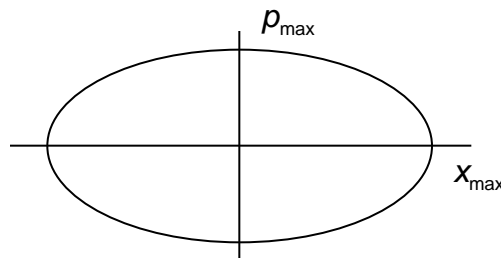

The oscillator is equally likely to be found at any *phase space* point on the trajectory. For illustration, let the dots on the following drawing represent phase space points


The oscillator has more "dots" at large *x* than it does near *x* = 0, where it has its greatest momentum. Thus, the oscillator is more likely to be found at large *x* than near *x* = 0.

Ergodic assumption

There is a technical subtlety about the measurement of averages that should be mentioned at this point. Let's consider a group of hard spheres confined within a container. The spheres have no other interaction with each other beyond their hard-core potential: when they collide, they exchange energy and momentum according the usual conservation laws.
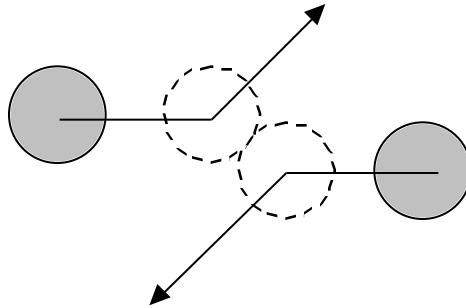
Suppose that we now want to find the probability of a sphere having a given speed *s*. Few spheres are moving either very slowly or very fast, so the probability distribution might look like:
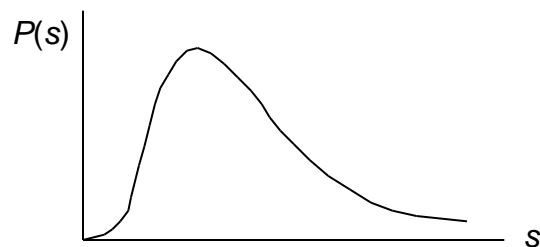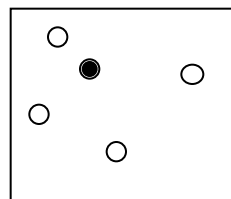
where *P*(*s*) *ds* is the probability of finding a sphere with speed between *s* and *s* + *ds* Strictly speaking, *P*(*s*) is a *probability density*: you multiply *P*(*s*) by the range *ds* to find the probability. Varying between 0 and 1, the probability has no units, whereas the probability density must have units of 1/*s*.

We can imagine two ways of measuring this distribution:
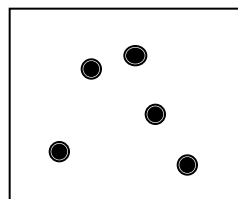
measure all spheres at once   make repeated measurements
                                              on a single sphere

In one case, all the speeds are measured at a single instant in time; in the other case, the motion of a single object is followed, and its changing speed is measured

repeatedly.   The **ergodic assumption** states that the results of these two measurements will be the same.

Probability

Now we start to put these ideas into a mathematical framework. Suppose that we have an energy range E to $E + \delta E$, which specifies the accessible states of the system:

$E + \delta E$

define Ω to be the number of states in this region

E

Let's measure the value of an observable y, which we say has discrete values y_1 y_2 ... y_k ... The number of accessible states with observable y having the value y_k we denote by $\Omega(E; y_k)$. By the fundamental postulate, the probability of observing the system with observable y_k is then

[*Prob. of y_k*] = [*# of accessible states with y_k*] / [*# accessible states in total*]

or

$$P(y_k) = \frac{\Omega(E; y_k)}{\Omega(E)} \qquad\qquad\qquad\qquad (Cx1.6)$$

where

$$\Omega(E) = \sum_k \Omega(E; y_k). \qquad\qquad\qquad\qquad (Cx1.7)$$

Eq. (Cx1.6) permits us to calculate the probability distribution from the number of accessible states.

Once the probability distribution has been determined, it can be used to calculate a variety of quantities that characterize the system as a whole. For example, the mean value of y is found from the weighted sum over the individual y_k's

$$\bar{y} = \sum_k y_k P(y_k) = \frac{\sum_k y_k \Omega(E; y_k)}{\Omega(E)} . \qquad\qquad\qquad\qquad (Cx1.8)$$