Part 2: From Probabilities to Stat Mech

Probabilities
- Simple probabilities
- averages
- Composite probabilities
- Independent events
- independent events
- simple and complex
- lots of counting

Statistical Mechanics
- Ising
- Hamiltonian description
- Averages from derivatives
- statistical thermo
- quantum
- Pauli matrices
- one and many
- Structural Invariance
- Intensive and Extensive
- many Ising variables (defining entropy)
- many Ising variables (counting)

Gaussian
- probability Distributions
- Statistical Variables
- Integrals and Probabilities
- complete the square
- Statistical Distributions
- Many variables
- Gaussian averages (usual way)
- rapid variation
- Approximate Gaussian integrals

Quantum
- Harmonic oscillator
- average energy

Calculation of Averages and Fluctuations: in general
- sums and averages with distribution functions
- more sums and averages
- homework
Simple probabilities
six sides, numbers 1-6

mutually exclusive events described by $\alpha=1,2,3,\ldots$

number of times $\alpha$ turns up $= N_{\alpha}$; total number of events $N$

\[ N = \sum_{\alpha} N_{\alpha} \]

dice: probability of getting a side with number $\alpha$ is $\rho_{\alpha}$

\[ \rho_{\alpha} = \frac{N_{\alpha}}{N} \]

total probability $= 1$ --> \[ \sum_{\alpha} \rho_{\alpha} = 1 \]

relative probability: relative chance that $\alpha$ will turn up $= r_{\alpha}$, e.g. fair dice have $r_{\alpha} =$ constant

normalize (= fix up size): from $r$ to $\rho$, first find $z = \sum_{\alpha} r_{\alpha}$

then \[ \rho_{\alpha} = \frac{r_{\alpha}}{z} \]

Example: cubic dice 6 sides: fair dice --> all probabilities are equal --> $r_{\alpha} = 1$ --> $z = 6$ --> $\rho_{\alpha} = 1/6$ for all values of $\alpha$
Averages six sides, numbers 1-6

\[ \rho_{\alpha} = \frac{1}{6} \]

average number on a throw = \[ \langle \alpha \rangle = \left( \sum_{\alpha} \alpha N_{\alpha} \right) / N \]

average number on a throw = \[ \langle \alpha \rangle = \sum_{\alpha} \rho_{\alpha} \alpha = 3.5 \]

general rule: To calculate the average of any function \( f(\alpha) \) that gives the probability that what will come out will be \( \alpha \), you use the formula

\[ < f(\alpha) > = \sum_{\alpha} f(\alpha) \rho_{\alpha} \]

examples:

\[ <1> = \sum_{\alpha} \rho_{\alpha} = 1 \]
\[ <\alpha> = \sum_{\alpha} \rho_{\alpha} \alpha = 3.5 \quad \text{average } \alpha \]
\[ <\alpha^2> = \sum_{\alpha} \rho_{\alpha} \alpha^2 = ? \quad \text{average of } \alpha^2 \]

\[ <\alpha^2> - <\alpha>^2 = <(\alpha - <\alpha>)^2> = \sum_{\alpha} \rho_{\alpha} (\alpha - <\alpha>)^2 = ? \quad \text{variance of } \alpha \]
Composite Probabilities

α and β are two different kinds of events

α might describe the temperature on January 1, ρα computed as Nα / N

β might describe the precipitation on December 31, with probabilities ρβ

Both kinds of events are complete \( \sum_{\alpha} \rho_{\alpha} = 1 \) \( \sum_{\beta} \rho'_{\beta} = 1 \)

The prime indicates that the two probabilities are quite different from one another.

Let \( \rho_{\alpha,\beta} \) be the probability that both will happen. The technical term for this is a joint probability. The joint probability satisfies \( \sum_{\alpha,\beta} \rho_{\alpha,\beta} = 1 \)

\( \rho(\alpha|\beta) \) is the probability that event α occurs if that we know that event β has or will occur. This quantity is called a conditional probability. It obeys \( \rho(\alpha|\beta) = \rho_{\alpha,\beta} / \rho'_{\beta} \)

Something must happen, implies that \( \sum_{\alpha} \rho(\alpha|\beta) = 1 \)
Independent Events

Physically two events are independent if the outcome of one does not affect the outcome of the other. It is a mutual relation, if \( \alpha \) is independent of \( \beta \) then \( \beta \) is independent of \( \alpha \).

This can then be stated in terms of conditional probabilities. If \( \rho(\alpha|\beta) \) does not depend upon \( \beta \) then we say \( \alpha \) and \( \beta \) are statistically independent. After a little algebraic manipulation, it follows that the joint probability \( \rho_{\alpha,\beta} \) obeys

\[
\rho_{\alpha,\beta} = \rho_\alpha \rho'_\beta
\]

Equivalently, two events are statistically independent, if the number of times both show up is expressed in terms of the number of times each one individually shows up as

\[
N_{\alpha,\beta} = N_\alpha N'_\beta / N
\]

This can be generalized to the statement that a series of \( m \) different events are statistically independent if the joint probabilities of the outcomes of all these events is simply the product of all the \( m \) individual probabilities.

The word uncorrelated is also used to describe statistically independent quantities.
Simple and Complex

definition:  **simple outcome**: can happen only one way: like 2 coming up when a die is thrown.

definition:  **complex outcome**: can happen several ways: like 7 coming up when two dice are thrown. One should calculate probability of complex outcome as a sum of probabilities of simple outcomes.

If the simple outcomes are equally likely, probability of complex outcome is the number of different simple outcomes times the probability of a single simple outcome. There is lots of counting in statistical mechanics. The number of ways that something can happen is often denoted by the symbol $W$. Entropy is given by

$$S = k \ln W,$$

where $k = k_B$ is Boltzmann’s constant. This equation is on Boltzmann’s tombstone. He committed suicide.

- what is minimum value of $S$ ?

- can you think of a way of getting sub-minimum values of $S$ ?
Lots of Counting in Statistical Theory

More on Independence: Given two kinds of events, A and B, statistically independent of one another. A can happen in \( N_A \) different ways, B in \( N_B \) ways. The total number of ways A and B can happen is \( N = N_A N_B \). Of course if we are considering a whole string of independent things that can happen A,B,C,D,… The total number of ways they can happen is \( N = N_A N_B N_C N_D \)....

Problem 1: Given \( N \) balls numbered 1, 2, 3, ..., \( N \). How many different ways can we arrange them in a row? Answer: Well we can pick the first ball in the row in \( N \) ways, the second in \( N-1 \), the third in \( N-2 \), ..., leaving one last ball that we can only stick on the end. The total number of different ways is then \( N(N-1)(N-2)\ldots(3)(2)(1) = N! \). \( N! \) is called \( N \) factorial.

About factorials: Mathematicians use the gamma function with \( \Gamma(N)=(N-1)! \).

The factorial function grows quite large as \( N \) grow large. To estimate the size of \( N! \) we use

\[
\ln N! = \sum_{n=0}^{N} \ln n \approx \int_0^{N} \ln n \, dn = N \ln N - N + \text{const}
\]

To next order \( \ln N! \approx = N \ln N - N + \frac{1}{2} \ln (2\pi) \) Sterling approximation

Move balls from random to ordered arrangement

\[
\begin{align*}
\begin{array}{c}
7 \\
3 \\
2 \\
\end{array}
&
\begin{array}{c}
1 \\
5 \\
6 \\
\end{array}
&
\begin{array}{c}
4 \\
3 \\
2 \\
\end{array}
\end{align*}
\]

\[
\begin{align*}
\begin{array}{c}
3 \\
2 \\
\end{array}
&
\begin{array}{c}
5 \\
6 \\
\end{array}
&
\begin{array}{c}
1 \\
7 \\
4 \\
\end{array}
\end{align*}
\]
problem 2: Given N balls and two bins. How many different different ways can we throw the balls into the bins?
Answer: Each ball can be thrown into one of the two bins. This event is repeated N times. We then have a number of events equal to $2^N$. 

problem 3: Given N balls and two bins. We want to put m balls into the first bin and N-m into the second. We do not care about the order in which the balls were added to a given bin. We count two ways as different only if balls with different numbers on them appear in a given bin. How many ways can we do add these balls?

Answer: Number the balls. Throw them one at a time into the bins, with the first m balls being added to bin 1 and the remainder to bin 2. There are N! different ways of doing this. However this overcounts the number of different end results. There are m! different orderings of the throws that leave us with exactly the same balls in bin 1. Similarly there are (N-m)! different equivalent orderings of throws into the second bin. There are thus a total of m! (N-m)! different ordering which will leave the same balls in the same bins. We only count as different those throwing events which leave us with a different outcome for the numbered balls. Therefore the number of outcomes is

\[ \frac{N!}{(m! (N-m)!)} \]

problem 4: What is the probability of the outcome (m,N-m) of problem 3 among all the different possible outcomes.

Answer: The probability, \( \rho_m \), is the total number of outcomes from problem 3 divided by the total number of events from problem 2. Therefore the probability is

\[ \rho_m = \frac{N!}{(2^N m! (N-m)!)} \]
Description of any system described by a Hamiltonian, $H$ or $H(\cdot)$: The dot is a placeholder for the variables that describe the internal state of the system. If you wait long enough the system will come to statistical equilibrium and then the relative probability of finding it in any configuration described by the dot is

$$r = \exp[-\beta H(\cdot)]$$

where $\beta^{-1}$ is the Boltzmann constant times the temperature, $k_B T$ or just $kT$. Then, of course the probability of finding a configuration is

$$\rho(\cdot) = \frac{\exp[-\beta H(\cdot)]}{Z(\beta)}$$

where $Z$ is the sum over all configurations of the relative probability. We call that sum a trace so we have

$$Z=\text{trace } \exp[-\beta H(\cdot)]$$

and the average of anything $X(\cdot)$, which depends upon the configurations of the system is

$$\langle X(\cdot) \rangle = \text{trace } \{\exp[-\beta H(\cdot)] \ X(\cdot)\}/Z$$

That’s all there is to statistical mechanics.

Quantum statistical mechanics is defined by a diagonal sum over all states.

$$\text{trace } X = \sum_\alpha <\alpha|X|\alpha>$$

Classical statistical mechanics is defined by starting from Hamiltonian mechanics, in which there are lots of coordinates $p_\alpha$ and $q_\alpha$ and then integrations over phase space: $d\gamma = \prod_\alpha dp_\alpha \ dq_\alpha$ and trace means $\int d\gamma$. 
A start: The Simplest Statistical Mechanics System

The Ising* system has as its basic variable a spin, with z-component, $\sigma_z$ that takes on the values $\pm 1$. This situation combines the simplicity of commuting variables, as in classical mechanics, with the simplicity of discrete variables, i.e. ones which take on only a few values, as in some quantum cases.

We shall use the abbreviation, $\sigma$ for this spin component.

The behavior of a physical system is described by its Hamiltonian, which not only gives the equilibrium statistical distributions, but also the time-dependence of the system. If we put this spin in a magnetic field in the z-direction it has a Hamiltonian $H = -\mu B_z \sigma_z$, with $\mu$ being the magnetic moment and $B_z$ being the component of the magnetic field in the $z$-direction.

Statistical Mechanics is defined by giving the relative weight of the different configurations. Here the configurations are defined by the two values of $\sigma$ and the relative weight is

$$r(\sigma) = \exp[-H/(kB)T] = \exp[\mu B_z \sigma/(kB)T]$$

We describe this by using the abbreviation, $h$, for the parameters in the probability

$$r(\sigma) = \exp(h \sigma) \quad \text{with} \quad h = \mu B_z / (kB)T$$

and with $\sigma$ having the two values $+1$ and $-1$.

* The Ising model was invented by the physicist Wilhelm Lenz (1920) who gave it as a problem to his student Ernst Ising after whom it is named. You can find more in Wikipedia.
Ising System, continued

Here the configurations are defined by the two values of $\sigma$ and the relative weight is

$$r(\sigma) = \exp(h \sigma) \quad \text{with} \quad h = \mu B_z / (k_B T) \quad \text{and with} \quad \sigma \ \text{having the two values} \ +1 \ \text{and} \ -1.$$  

Statistical Mechanics is defined by a probability. The probability is

$$\rho(\sigma) = (1/z) \ r(\sigma) \quad \text{and} \quad z = r(1) + r(-1) = \exp(h) + \exp(-h) = 2 \cosh h$$

**normalization:** total probability $= 1 = \rho(1) + \rho(-1) = (1/z) \ \exp(h) + (1/z) \ \exp(-h)$

**average:** $X = \langle X \rangle = \sum_\alpha \rho(\alpha) \ X_\alpha$

therefore $\langle \sigma \rangle = \rho(1)1 + \rho(-1)(-1) = 1/(2 \cosh h) \ \{\exp(h) - \exp(-h)\}$

$= (2 \sinh h) / (2 \cosh h) = \tanh h$

---

* The Ising model was invented by the physicist Wilhelm Lenz (1920) who gave it as a problem to his student Ernst Ising after whom it is named. You can find more in Wikipedia.
Averages from Derivatives: A calculational method

Use Ising Hamiltonian given by* $-\beta H = h \sigma$

$$z = \sum_\sigma \exp(h \sigma) = 2 \cosh h$$

partition function

$$d(\ln z) / dh = \sum_\sigma \sigma \exp(h \sigma) / z = \langle \sigma \rangle = \tanh h$$

average

$$d^2(\ln z) / (dh)^2 = \sum_\sigma (\sigma - \langle \sigma \rangle)^2 \exp(h \sigma) / z = \langle (\sigma - \langle \sigma \rangle)^2 \rangle$$

$$= 1 - \langle \sigma \rangle^2 = 1 - (\tanh h)^2$$

variance

All derivatives of the log of the partition function are thermodynamic functions of some kinds. As I shall say below, we expect simple behavior from the log of $Z$ but not $Z$ itself. The derivatives described above are respectively called the magnetization, $M = \langle \sigma \rangle$ and the magnetic susceptibility, $\chi = dM/dh$. The analogous first derivative with respect to $\beta$ is minus the energy. The next derivative with respect to $\beta$ is proportional to the specific heat, or heat capacity, another traditional thermodynamic quantity. The derivative of the logarithm of the partition function with respect to volume is proportional to the pressure, i.e $\partial \ln Z / \partial \Omega = \beta P$.

Result: variance $= \langle (\sigma - \langle \sigma \rangle)^2 \rangle = \langle \sigma \rangle (1 - \langle \sigma \rangle)$

* $\beta = 1/(k_B T)$, $h = \mu B_z / (k_B T)$
A little statistical thermodynamics

\[ z = \text{trace} \exp[-\beta H] \] is the partition function for a single particle. We write this as \[ z = \exp[-\beta f] \] where \( f \) is the free energy for a single particle. Then \[ -\beta f = s/k_B - \beta \langle H \rangle \] where \( s \) is the entropy for a single particle. This means, roughly speaking, that \( \exp[s/k_B] \) is the number of terms which count in the statistical sum for the particle.

Those of you who are in the know will understand that these concepts will very directly generalize to systems with many particles, only the notation will change so that the quantities we have just defined will morph into one involving capital letters: \( Z,F,S \) for the partition function, free energy, and entropy of the entire system.

Also \( d \ln Z = -E d\beta + \beta p d\Omega \) with \( \Omega \) being the volume.

For the Ising system just defined
\[ z = 2 \cosh h \; ; \; \; \; -\beta \langle H \rangle = h \tanh h \; \; \; \; -\beta f = \ln (2 \cosh h) \]
and \( s/k_B = \ln (2 \cosh h) - h \tanh h \)

Notice that all of these quantities are even under the interchange \( h \) goes into \( -h \).

For \( h = 0 \), \( s/k_B \approx \ln 2 \), indicating that two configurations contribute to the free energy in this situation, i.e. both \( \sigma = 1 \) and \( \sigma = -1 \) contribute

For \( h = +\infty \), \( s/k_B \approx \ln (e^h - h) = \ln 1 = 0 \), indicating that one configuration contributes to the free energy in this situation, i.e. for \( h \) large and positive \( \sigma = 1 \) contributes but not \( \sigma = -1 \).
Quantum Statistical Mechanics: Hamiltonian Description

The first idea in statistical mechanics is that the relative probability of a configuration labeled by $\alpha$, is given in terms of the Hamiltonian $H$ or equivalently the energy $E(\alpha)$ in that the relative probability of such a configuration is $\exp(-H/T)=\exp[-E(\alpha)/T]$ where $T$ is the temperature measured in energy units. (This is usually written $k_B T$ where $k_B$ is the Boltzmann constant.) The general formulation of statistical mechanics is then that

\[
\rho_\alpha = \frac{e^{-E(\alpha)/T}}{Z} \quad \text{with} \quad Z = \sum_\alpha e^{-E(\alpha)/T}
\]

In the example we have already given, a single atom with spin $\frac{1}{2}$ in a magnetic field has a Hamiltonian which can be written as $H= -B \sigma_z$, with $B$ the $z$ component of the magnetic field and $\sigma_z$ being the $z$ component of the spin. In this model, this component takes on the values $\sigma=\pm 1$. (Most often, one puts subscripts $z$ on the spin and magnetic field.)

The relative probabilities for positive and negative values of the spin are respectively $e^h$ and $e^{-h}$ with $h= B \mu / (kT)$ so that the partition function is $Z = e^h + e^{-h} = 2 \cosh h$. 
A Simple Example: Pauli Spin Matrices

In the real, three-dimensional, world spins are described by a three-component spin vector \((\tau_1, \tau_2, \tau_3)\) where the different components correspond to the different directions of a vector \(r = (x, y, z)\). These components are two by two matrices which obey the algebraic relations
\[
\tau_j \tau_k + \tau_k \tau_j = 2 \delta_{j,k} \mathbf{1} \quad \tau_1 \tau_2 \tau_3 = i \mathbf{1}
\]
Here \(\mathbf{1}\) is a two by two unit matrix.

Using these equations, we see that \((\tau_2)^2 = 1\) and \(\tau_1 \tau_2 = -\tau_2 \tau_1 = i \tau_3\)

Thus, for example the operator \(\tau \cdot B = \tau_1 B_x + \tau_2 B_y + \tau_3 B_z\), that is the dot product of a magnetic field vector and the spin vector has a square which we can calculate as
\[
(\tau \cdot B)^2 = B_x^2 + B_y^2 + B_z^2 = B^2
\]
The right hand side is just the squared magnitude of the vector \(B\).
Therefore the two eigenvalues are plus and minus the magnitude of \(B\).
\[
(\tau \cdot B)^{-} = B \quad \text{and} \quad -B
\]

Then if \(-\beta H = \beta \mu \tau \cdot B\), the partition function is \(z = 2 \cosh [\beta \mu (B_x^2 + B_y^2 + B_z^2)^{1/2}]\) For later reference we should note that when the Hamiltonian is a sum of non-commuting terms the partition function is **not** a product of the individual partition functions of the parts. This is in contrast to the case in which the parts commute.
Imagine a material with many atoms, each with its own spin. The system
has a Hamiltonian which is a sum of the Hamiltonian of the different atoms

\[ H = \sum_{\alpha=1}^{N} h\sigma_{\alpha} \]

and a probability distribution

\[ \rho = \exp(-\beta H) / Z = (1 / Z)\prod_{\alpha=1}^{N} \exp(h\sigma_{\alpha}) \]

which is a product of pieces which belong to the different atoms. The
different pieces are then statistically independent of one another. Note
that the partition function is

\[ Z = \prod_{\alpha=1}^{N} \sum_{\sigma^\alpha = \pm 1} \exp(h\sigma_{\alpha}) = (2 \cosh h)^N = z^N \]

so that the entire probability is a product of N pieces connected with the N atoms

\[ \rho\{\sigma\} = \prod_{\alpha} [\exp(h\sigma_{\alpha}) / z] \]

The appearance of a product structure depends only upon having a Hamiltonian
which is a sum of terms referring to individual parts of the system

Hamiltonian is sum \(<---\>\) stat mech probability is product \(<---\>\) statistical independence
Structural invariance

Note how the very same structure which applies to one atom, $\exp(-\beta H)/Z$, carries over equally to many atoms.

This structural invariance is characteristic of the mathematical basis of physical theories. Newton’s gravitational theory seemed natural because the same law which applied to one apple equally applies to an entire planet composed of apples.

This same thing works for electromagnetism. The same law which gives the force for a single electron also gives the force pattern produced outside a spherically symmetric object containing many charged particles.

A wave function is the same sort of thing for one electron or many. The structure of space and time has a similar invariance property. Remember that a journey of a thousand miles starts with but a single step. The similarity between a single step and a longer distance is a kind of structural invariance. This invariance of space is called a scale invariance. It is quite important in all theories of space and time.
Intensive and Extensive

Imagine a situation in which we have many independent identical subsystems, called iid by the mathematicians for independent and identically distributed. Then $Z(\beta) = z(\beta)^N$ as in equation ii.4. In this kind of system, the free energy, the entropy, and the average energy will be precisely $N$ times that for the single subsystem.

Quantities that are a sum or integral over the entire system are called extensive. One which define local properties or are ratios of extensive quantities are called intensive. The former are almost always linear in $N$, the latter are almost always independent of $N$.

The mean square fluctuation in energy is extensive. Why?? The root mean square fluctuation in energy, $\sqrt{<(H-<H>)>}$, is not extensive. Why?? In the limiting case in which the number of subsystems (particles) is very large, which is the relative size of the average energy and the root mean square fluctuation in the energy?

I have just told you the reason why classical thermodynamics does not consider fluctuations in things. Again, what is this reason?

How big is $N$, say for the particles in this room?

In terms of real MKS numbers, roughly what is the average kinetic energy of these particles and what is the size of its fluctuation?
Many Ising Variables: Defining Entropy

Consider the behavior of a large number \(N\) of Ising Variables in a magnetic field. The partition function is given by a Hamiltonian

\[-\beta H = \sum_{j=1}^{N} h\sigma_j\]

and then a partition function of the form.

\[Z = \prod_{j=1}^{N} \sum_{\sigma_j=\pm1} \exp(-\beta H) = \prod_{j=1}^{N} \sum_{\sigma_j=\pm1} \exp(h\sigma_j) = (2 \cosh h)^N\]

It follows at once that the average energy is

\[< -\beta H > = \sum_{j} h < \sigma_j > = h \frac{\partial \ln Z}{\partial h} = Nh \tanh h\]

Now the partition function is \(Z=\exp(-\beta F)\) which is defined to be \(\exp(S/k_B-\beta<H>)\) where \(S\) is the entropy and \(\exp(S/k_B)\) counts how many terms contribute to the free energy. Since we know the average energy, we can immediately calculate the entropy as \(S/k_B=N[ -(h \tanh h )+ \ln (2 \cosh h)]\).

The exponential of \(S/k_B\) counts the number of contributions to the energy.

When \(h\) goes to \(\pm\)infinity, \(\tanh h\) goes to \(\pm\)1 and \(\cosh h\) goes to \(e^{h/2}\). The net result is that in this situation \(S/k_B\) is zero so that precisely one term contributes.

In the opposite limit when \(h=0\), \(S/k_B\) is \(N \ln 2\) so that all of the \(2^N\) possible terms are contributing.
Many Ising Variables: Counting

\[-\beta H = \sum_{j=1}^{N} h \sigma_j\]

\[Z = \sum \exp(h \sum \sigma_j) = (2 \cosh h)^N = (e^{h} + e^{-h})^N = \Sigma_j \frac{N!}{j!(N-j)!} e^{jh} e^{-h(N-j)} \quad \text{(from binomial theorem)}\]

Number of times we have exactly j spins up = \(\frac{N!}{j!(N-j)!} e^{jh} e^{-h(N-j)}\)

If we were just throwing spins at random into spin up or spin down states, then we would have \(h=0\). In that case, number of times we have exactly j spins up = \(\frac{N!}{j!(N-j)!}\) just like the balls in problem 3.

We would have exactly \(Z = 2^N\) total possibilities, so the probability of having j balls is \(\frac{N!}{2^N j! (N-j)!}\).
Probability Distributions

So far we have talked about discrete outcomes. A die may take on one of six possible values. But measured things are often continuous. For example, in one dimension, the probability that a quantum particle will be found between $x$ and $x + dx$ is given in terms of the wave function, $|\psi(x)|^2 dx$. In this context, the squared wave function appears as a *probability density*. In general, we shall use the notation $\rho(x)$ for a probability density, saying that $\rho(x) \ dx$ is the probability for finding a particle between $x$ and $x + dx$. The general properties of such probability densities are simple. They are positive. Since the total probability of some $x$ must be equal to one they satisfy the normalization condition

$$\int_{-\infty}^{+\infty} \rho(x) \ dx = 1$$

For example, in classical statistical mechanics, the probability density for finding a particle with $x$-component of momentum equal to $p$ is

$$\left(\frac{2\pi\beta}{m}\right)^{1/2} \exp\left[-\beta \frac{p^2}{(2m)}\right]$$

This is called a *Gaussian* probability distribution, i.e. one that is based on $\exp(-x^2)$. Such distributions are very important in theoretical physics, mathematics, and statistics.
Gaussian Statistical Variables

A Gaussian random variable, $X$, is one which has a probability distribution which is the exponential of a quadratic in $X$.

$$\rho(x) = \left[\frac{\beta}{(2\pi)}\right]^{1/2} \exp[-\beta(x-\langle X \rangle)^2/2]$$

$1/\beta$ is the variance of this distribution.

The sum of two statistically independent Gaussian variables is also Gaussian. How does the variance add up?

A Gaussian variable is an extreme example of a structurally stable quantity.

Central Limit Theorem: Given a large number of random variables, all having variance less than some preset maximum. The sum of these variables is, to a good approximation, a Gaussian random variable, with variance and mean equal to the sum of those of the summed variables.

A Gaussian distribution has a lot of structurally invariant properties.
Gaussian integrals are of the form
\[ l = \int dx \, \exp(-ax^2 / 2 + bx + c) \]
with \( a, b, \) and \( c \) being real numbers, complex numbers, or matrices. They are very, very useful in all branches of theoretical physics.

We define the probability that the random variable \( X \) will take on the value between \( x \) and \( x + dx \) as \( \rho(X=x)dx \) or more simply as \( \rho(x)dx \)

There is a canonical* form for Gaussian probability distributions, namely
\[ \rho(X=x) = (\beta/2\pi)^{1/2} \exp[-\beta(x - <X>)^2 /2] \]
produced by “completing the square". Here \( X \) is the random variable \( 1/\beta \) is its variance and \(<X>\) is the average of this random variable.

\[ \rho(x) \sim \exp[-ax^2/2+bx +c]= \exp[-a(x-b/a)^2/2+b^2/(2a) +c] \]
so pick \( c=-b^2/(2a)+[\ln (\beta/2\pi)]/2 \) we now have the canonical form

with \(<X> = b/a \) and variance \( 1/a \)

* canonical  |ˈkanən̩ikəl|   adjective
1 according to or ordered by canon law: the canonical rites of the Roman Church.
2 ...... Mathematics of or relating to a general rule or standard formula.
Complete the square

For Gaussian probability distributions, there is a very important result:

\[ < \exp(iqX) >= \exp(iq < X >) \exp[-q^2 / (2\beta)] \]

Based on the canonical form \( \rho(X=x)= (\beta/2\pi)^{\frac{1}{2}} \exp[-\beta(x - <X>)^2 /2] \)

Notice how the \( \beta \) that appears in the numerator of the probability distribution reappears in the denominator of the average.

One can show this by a process of “completing the square”. For simplicity let us have a Gaussian distribution in \( \phi \), with \( < \phi >=0 \)

Specifically notice that an integral of the form

\[ z(a) = \int d\phi \exp[-\frac{1}{2} \phi \beta \phi + i \phi a] \]

can be calculated by replacing the integration variable by

\[ \psi = \phi - i \frac{a}{\beta} \]

to give an integration result. One can displace the \( \psi \)-contour to the real axis, to find

\[ z(a) = \int d\psi \exp[-\frac{1}{2} \psi \beta \psi - a^2/(2 \beta)] = (2\pi/\beta)^{\frac{1}{2}} \exp[-a^2/(2 \beta)]. \]

We can consider \( z(a)/z(0) = \exp[-a^2/\beta] \) to be \( < \exp[ i \phi a] > \) with a Hamiltonian \(-\beta H= -\frac{1}{2} \phi \beta \phi\). With this Hamiltonian

\[ i < \phi > = (\partial \ln z(a)/\partial a)|_{a=0} = 0 \]

\[ <\phi^2 > = -(\partial^2 \ln z(a)/(\partial a)^2)|_{a=0} = 1/\beta \]
Gaussian Distributions in Stat Mech

According to Ludwig Boltzmann (1844 – 1906) and James Clerk Maxwell (1831-1879) the probability distribution for a particle in a weakly interacting gas is given by defining the probability distribution for finding the particle with a particular momentum and position, as

\[ \rho(p, r) = \frac{1}{z} \exp(-\beta H) \]

\[ H = \left[ p_x^2 + p_y^2 + p_z^2 \right] / 2m + U(r) \]

Here, the potential holds the particles in a box of volume \( \Omega \), so that \( U \) is zero inside a box of this volume and infinite outside of it. As usual, we go after thermodynamic properties by calculating the partition function, \( z = \int dp dr \exp[-\beta H] \)

\[ z = \Omega \left[ \int dp \exp(-\beta p^2 / (2m)) \right]^3 = \Omega (2\pi m / \beta)^{3/2} \]

In the usual way, we find that the average energy is \( 3/(2\beta) = (3/2)kT \). The classical result is the average energy contains a term \( kT/2 \) for each quadratic degree of freedom. Thus a harmonic oscillator has \( <H>=kT \).

Hint for theorists: Calculations of \( Z \) (or of its quantum equivalent, the vacuum energy) are important. Once you can get this quantity, you are prepared to find out most other things about the system. Specifically \( \ln Z = -F/T \quad d(F/T) = - (P/T) d\Omega + <H>d\beta \) so knowing \( Z \) you can calculate average energy and pressure.

The relation in red gives a starting point for thermodynamics. It is quite important.
Many variables are as easy as one

Let $M$ be an $N$ by $N$ symmetric real matrix with $N$ positive real eigenvalues, $m_1, m_2, \ldots, m_N$ are the eigenvalues of this matrix. We can then easily calculate an integral involving many Gaussian variables by taking linear combinations of variables to diagonalize the matrix, $M$, giving

$$Z = \int d\phi_1 \ldots d\phi_N \exp\left[-\frac{1}{2} \sum_{i,j} \phi_i M_{i,j} \phi_j\right] = \left[(2\pi)^N / \det M\right]^{1/2}$$

The last equality follows from the fact that the determinant of $M$ is the product of its eigenvalues. More specifically, if $M_{i,j}$ is a diagonal matrix, $M_{i,j} = \delta_{i,j} m_i$ then it follows at once that $Z = \prod_i (2\pi/m_i)^{1/2}$ since the determinant of $M$ is the product of its eigenvalues. If $M$ is not diagonal we form linear combinations of the $\phi$'s that are eigenstates of the matrix, $M$. If we use those linear combinations as integration variables, then our formula for the integration immediately follows once again.
Gaussian Averages

The usual way

Let one particle be confined to a box of volume $\Omega$. Let $U(r)$ be zero inside the box and $+\infty$ outside. Then, in three dimensions

$$Z = \int d^3p \, d^3r \, \exp(-\beta[p^2/(2m)+U(r)]) = \Omega (2\pi m/\beta)^{3/2}$$

Let $\varepsilon = [p^2/(2m)+U(r)]$

$$\frac{\partial \ln Z}{\partial \beta} = -\frac{1}{Z} \int d^3p \, d^3r \, \varepsilon \, \exp(-\beta \varepsilon) = -\langle \varepsilon \rangle$$

$$\langle \varepsilon \rangle = 3/(2\beta) = (3/2)k_B T \quad (1/2 kT \text{ for every quadratic term in the Hamiltonian.})$$

$$\frac{\partial^2 \ln Z}{\partial \beta^2} = -\frac{\partial \langle \varepsilon \rangle}{\partial \beta} = ??? \quad (\text{The usual way})^2$$
Homework assignment is up on web site. Homework due in class on Wednesday. The
assignment sheet also gives schedule for next two weeks:

Monday                      Wednesday                Friday
1/10 class                   1/12 class                Arnab goes over homework
1/17 day off                 1/19 class                1/21 class
Rapidly Varying Gaussian random variable

Later on we shall make use of a time-dependent gaussian random variable, $\eta(t)$. In its usual use, $\eta(t)$ is a very rapidly varying quantity, with a time-integral which behaves like a Gaussian random variable. Specifically, it is defined to have two properties:

$$< \eta(t)> = 0$$

$$X(t) = \int_s^t du \eta(u)$$ is a Gaussian random variable with variance $\Gamma |s-t|$. Here $\Gamma$ defines the strength of the oscillating random variable.
The Quantum Harmonic Oscillator

The basic Hamiltonian is $p^2/(2m) + kq^2/2$ which transform into $\hbar \omega [N + \frac{1}{2}]$, where $N$ has the eigenvalues 0,1,2,..... The partition function is

$$z = \sum_{n=0,1,2,...} e^{-\beta \hbar \omega [n+\frac{1}{2}]} = e^{-\frac{1}{2} \beta \hbar \omega} [1 + e^{-\beta \hbar \omega} + e^{-2\beta \hbar \omega} + ...]$$

This geometric series sums to

$$z = e^{-\frac{1}{2} \beta \hbar \omega} [1 - e^{-\beta \hbar \omega}]$$

We can calculate the average energy by using:

$$\partial \ln z / \partial \beta = -<H> = -\hbar \omega <N+\frac{1}{2}>$$

The direct calculation gives:

$$\partial \ln z / \partial \beta = -\frac{1}{2} \hbar \omega - \hbar \omega e^{-\beta \hbar \omega} [1 - e^{-\beta \hbar \omega}]$$

We thus see that the average value of the quantum number is

$$<N> = 1/[ e^{\beta \hbar \omega} - 1]$$

and

$$<H> = \frac{1}{2} \hbar \omega + \hbar \omega <N>$$
Average Energy of the Quantum Harmonic Oscillator

\[ \langle N \rangle = \frac{1}{e^{\beta \hbar \omega} - 1} \]
and \[ \langle H \rangle = \frac{1}{2} \hbar \omega + \hbar \omega \langle N \rangle \]

**In the quantum limit**, \( \beta \hbar \omega \) is large. In that situation the average of \( \langle N \rangle \) is small indicating that most of the time, the oscillator is not excited at all. Thus for example for a single diatomic molecule, at room temperature it is unlikely that we will see much excitation of oscillatory motion. Hence such motion will contribute little to the specific heat. This is part of a general rule that excitation with \( \hbar \omega \gg k_B T \) are “frozen out” and play a tend to play a small role in the behavior of a quantum

**In the classical limit**, \( \beta \hbar \omega \) is small. Here with \( \hbar \omega \ll k_B T \) we expect to see the classical result that there is a contribution to the heat capacity, \( \partial \langle H \rangle / \partial T \), of \( k_B / 2 \) for each classical degree of freedom. In this case, sure enough we can expand the exponential as \( e^{\beta \hbar \omega} \approx 1 + \beta \hbar \omega \) so that \[ \langle N \rangle \approx \frac{1}{(\beta \hbar \omega)} \] Then the mode of oscillation has average energy \( \langle H \rangle = 1/\beta = k_B T \), as we might expect from a classical quadratic system with two degrees of freedom. Thus, the rotational degrees of freedom of a diatomic molecule contribute an amount \( k_B \) to the specific heat of a low density gas, because there are two observable modes of rotation of the molecule each with a quadratic kinetic contribution to the Hamiltonian.

This behavior of molecular specific heats confused Gibbs and Boltzmann quite considerably. It was not understood until the invention of quantum mechanics.

For an application see http://www.decompression.org/maiken/Molecules.htm
Calculation of Averages and Fluctuations: Reprise

The usual way: in general

Let’s start from a Hamiltonian on a Lattice

\[ W\{\sigma\} = -\beta H\{\sigma\} = \sum_r h_r \sigma_r + \sum_{<r,s>} K \sigma_r \sigma_s \]

This Hamiltonian defines what is called the Ising model. The first sum is a sum over all lattice sites, \( r \). The second sum is a sum over nearest neighbors. The field, \( h_r \), which depends upon \( r \), multiplies a spin variable which is different, of course, different on each site. The notation \( \{\sigma\} \) says that these things depend upon many spin variables. You can assume that the spin variables take on the value +1 or -1 if you want, but the argument below on this slide is very general and the result does not depend upon what \( \sigma \) might be. Start from

\[ Z = \text{Tr} \exp W\{\sigma\} \]

where Tr means a summation over all possible values of all the spin variables. It is a repeat of the argument that we have given before to say that

\[ \langle \sigma_r \rangle = \frac{\partial \ln Z}{\partial h_r} \]

The partial derivative means that we hold all other \( h \)’s constant. The second derivative is given by

\[ \langle \sigma_r \sigma_s \rangle - \langle \sigma_r \rangle \langle \sigma_s \rangle = \langle [\sigma_r - \langle \sigma_r \rangle][\sigma_s - \langle \sigma_s \rangle] \rangle = \frac{\partial^2 \ln Z}{\partial h_s \partial h_r} = \frac{\partial <\sigma_s>}{\partial h_r} \]
Sums and Averages via Distribution Functions

The probability distribution for a single particle in a weakly interacting gas as is given by

$$\rho(p, r) = (1/ z) \exp(-\beta H)$$

$$H = [p_x^2 + p_y^2 + p_z^2] / 2m + U(r)$$

Here, the potential holds the particles in a box of volume $\Omega$, so that $U$ is zero inside a box of this volume and infinite outside of it. The partition function, is

$$z = \Omega \left[ \int dp \exp(-\beta p^2 / (2m)) \right]^3 = \Omega (2\pi m / \beta)^{3/2}$$

The average of any function of $p$ and $r$ is given by

$$<g(p,r)> = \int dp \ dr \ \rho(p,r) \ g(p,r)$$

Since there are $N$ particles in the system $N \ dp \ dr \ \rho(p,r)$ is the number of particles which have position and momentum within $dp \ dr$ about the phase space point $p, r$. The quantity $N \ \rho(p,r)=f(p,r)$ is called the distribution function. The total amount of the quantity represented by $g(p,r)$ is given in terms of the distribution function as

$$\text{total amount of } g = \int dp \ dr \ f(p,r) \ g(p,r)$$

Example: We calculated the average energy $< p^2/(2m) > = 3 k T/2 = \int dp \ dr \ \rho(p,r) \ p^2/(2m)$

The total energy in the system is $\int dp \ dr \ f(p,r) \ p^2/(2m) = 3N \ k T/2$. 

Monday, January 10, 2011
More sums and averages

The normalization condition for the probability is
\[ \int d\mathbf{p} \, d\mathbf{r} \, \rho(\mathbf{p}, \mathbf{r}) = <1> = 1 \]

The normalization for the distribution function is
\[ \int d\mathbf{p} \, d\mathbf{r} \, f(\mathbf{p}, \mathbf{r}) = N \]

The pressure, \( P \), is defined as the total momentum transfer to a wall per unit of area and unit of time. Call these \( dA \) and \( dt \). Since a low density gas is the same all over, the number hitting is the number within the distance \( p_x/m \, dt \) of the area, for \( p_x > 0 \), and hence the number within the volume \( p_x/m \, dt \, dA \) which is
\[ \int d\mathbf{p} \, f(\mathbf{p}, \mathbf{r}) \, p_x/m \, dt \, dA \] with the integral covering all \( \mathbf{p} \)'s with the condition that \( p_x > 0 \). If a particle hits the wall and bounces back it transfers momentum \( 2p_x \). do we need specular reflection condition??.

Therefore the total momentum transferred is
\[ \int d\mathbf{p} \, f(\mathbf{p}, \mathbf{r}) \, p_x/m \, dt \, dA \, 2p_x \] once again with the condition that \( p_x > 0 \). An integral over all momenta would give a result twice as large. In the end we get that the pressure is
\[ P = \int d\mathbf{p} \, f(\mathbf{p}, \mathbf{r}) \, p_x^2/m \]
which is then \( NkT/\Omega \) as we knew it would be.

The partition function is the sum over all variables of \( \exp(-\beta H) \). For large \( N \), it can be interpreted as \( W \exp(-\beta <H>) \), where \( W \) is the number of configuration which enter. Boltzmann got \( W \) in terms of the entropy as \( \ln W = S/k \). We put together previous results and find
\[ W \exp(-3N/2) = z^N = \Omega^N \, (2\pi m kT)^{3N/2} \] so that \( S/k_B = N \, [\ln \Omega + 3 \, (\ln (2\pi e m kT))/2 ] \)
Please look at web site
http://jfi.uchicago.edu/~leop/
for my course outline and my record of the first group of lectures

In addition do problems 2.0-2.9, due Wednesday January 12 in class:

**Homework**

2.0 For problem of distributing N balls in two bins, calculate for N=20 the probability of finding zero and ten balls in bin 1 via exact theory and via Sterling approximation. In which case is absolute error of Sterling approximation larger? In which case is relative error larger?
2.1 What is probability of 2 coming up when one die is thrown? What is the probability of getting a total of 2 when 2 dice are thrown? What is the probability of getting a total of seven when 2 dice are thrown.

2.2 Given two fair dice what is the average sum of the two numbers that turn up. What is the average product of these numbers.

2.3 Work out the value of the pressure for a classical relativistic gas with \( H = |p| c \). Do this both by using kinetic theory, as just above, and also by differentiating the partition function. Calculate the mean squared fluctuations in the pressure.

2.4 The Hamiltonian for \( N \) particles in a fluid is \( H = \sum \alpha \frac{p^2}{2m} + \sum_{\alpha<\beta} V(|r_\alpha - r_\beta|) \). If the interaction, \( V \), is weak we can assume that the particles move independently of one another, and have positions which are independent variables. What is the value of the pressure to first order in the interaction? How does this result relate to van der Waals’ approximate equation of state for a fluid?

2.5 The length of a random walk of \( N \) steps is of the form of a sum of \( N \) iid variables* \( \sigma_j \) each of which takes on the values 0 and \( \pm 1 \) with equal probability. What is the probability that a walker who takes 1000 steps will end up at a point at least 30 steps away from her starting point? (A relative accuracy of two decimal digits will suffice.)

2.6 Now we roll one hundred dice all at the same time. What is the average of the sum of the dice-values. What is the variance of this sum? How can one define a root mean square fluctuation in the sum of the dice values? How big is it?

2.7 Given two Gaussian random variables \( X \) and \( Y \), with averages \( <X> \) and \( <Y> \) and respective variances \( \sigma^2 \) and \( \sigma^2 \). What are the values of the average and variance of \( Z = X + Y \) of \( Z^2 \)?

2.8 Given two particles, called \( A \) and \( B \), which respectively have spin matrices \( (\tau_1^A, \tau_2^A, \tau_3^A) \) and \( (\tau_1^B, \tau_2^B, \tau_3^B) \). The matrices for the different particles commute. Because of a magnetic field in the \( z \) direction and an interaction between the particles they have a Hamiltonian \(-\beta H = hm + K (\tau_3^A \tau_3^B) \) where \( m = (\tau_3^A + \tau_3^B) \). Calculate the partition function and the average of \( m \).

2.9 Given two particles, called \( A \) and \( B \), which respectively have spin matrices \( (\tau_1^A, \tau_2^A, \tau_3^A) \) and \( (\tau_1^B, \tau_2^B, \tau_3^B) \). The matrices for the different particles commute. Because of a magnetic field in the \( z \) direction and an interaction between the particles they have a Hamiltonian \(-\beta H = hm + K (\tau_1^A \tau_1^B + \tau_2^A \tau_2^B) \) where \( m = (\tau_3^A + \tau_3^B) \). Calculate the partition function and the average of \( m \).

*The abbreviation iid is used in mathematical probability theory. It stands for independently an identically distributed. The different things are statistically independent and each one has the same probability distribution as the others.