Why Statistical Physics.

In this class we will study the physics of macroscopic systems. That is, the physics of systems with a large number of equivalent particles (or sub-systems). There are two key concepts here:

1. Large numbers: Typically the # of particles, $N$, is extremely large, $\sim 10^{23}$, as in bulk solids, liquids, gases, stars, etc.

2. Equivalent: Unlike biological systems, in physics we deal with large collections of objects which for all purposes are strictly identical. Reference to Thermodynamics.

As you know, at the microscopic level, the laws of physics are governed by Quantum Mechanics and at high energies, Relativity plays a dominant role. At length scales large compared to the de Broglie wavelength, physics is described by the description is -- Classical Physics.

In classical physics we describe the state of a system by giving, at a given time $t$, the coordinates and velocities of all the particles. Since each particle
require 6 numbers (3 coordinates and 3 components of velocity) to describe the state. That is $6 \times 10^{23}$ numbers! Suppose we know these numbers at a given time, to describe the system we need to solve Newton's equations for some length of time. Suppose we know these 6N numbers with some precision given by some uncertainty $\Delta x$ for the coordinates and $\Delta v$ for the velocities. As time goes by the particle's coordinates and velocities will change. The problem is that for generic mechanical systems, the uncertainty also grows and, in fact, they grow exponentially fast. Thus, after some time, which depends on the system, the uncertainty will become large $\Delta x \times \Delta v$ and the system will have forgotten the initial state.

For a generic mechanical system, there are seven conserved quantities: Energy (1), Momentum (3) and Angular Momentum (3). These are simple systems (i.e., two-body problem with central force) for which these conserved quantities are enough to
describe the system. In some exceptional cases more conservation laws are found and, very rarely, there are enough conservation laws to specify a full Pfaffian (integrable system).

Thus, for $N > 1$, even though at the microscopical level we may know everything at the macroscopic level the situation is quite different. We must conclude that even a perfect knowledge of the state of the system at the microscopical level is not sufficient to determine the state of a macroscopic system.

We can put these statements more formally. Let us define the state of the system by a point in the $6N$ dimensional space of coordinates and velocities ("phase space"). As the system evolves, this point will trace a trajectory in phase space.

Let us suppose we choose a coordinate system in which the total momentum $\vec{P}$ and angular momentum $\vec{L}$ are zero.

Then the trajectory is characterized by the total energy $E$. 
However, there are a large number of trajectories in phase space (i.e., histories of the system) for a given total energy, momentum, and angular momentum. Suppose now that we start the system a large number of times, always with the same values of \( E, \vec{P}, \) and \( \vec{L} \). Obviously, the system will be in one of the many possible histories compatible with these very few conditions. Alternatively, we can imagine a collection, or ensemble, of systems all with the same value of \( E, \vec{P}, \) and \( \vec{L} \). If the system is very complex so that it forgets about the detailed initial conditions very rapidly, it is natural to assume that each one of these trajectories (or histories) is just as likely to happen and in very many trials (or for a large ensemble) all possible trajectories are equally likely. In other words, we assume that very large systems with very complex dynamics behave effectively as if they were governed by simple statistical or probabilistic laws. This assumption that all possible states compatible
with the requirements of $E$, $P$ and $I$ fixed, are equally likely is known as the Postulate of Equal a Priori Probabilities. It was first stated by Ludwig Boltzmann (1870). The ensemble of systems governed by this probability law is known as the Microcanonical Ensemble.

Another way to think about this problem is to use the Axiom of Ergodicity of a complex mechanical system which states that if we wait long enough, the system will get arbitrarily close to any state in phase space. If we pick a small volume in phase space around a given state, the fraction of time at the system spends in that volume (where $T$ is the total time evolution of the system), follows that of the particular state (with a given energy).

Here too, the fraction $\frac{1}{T}$ can be regarded as the probability of that state.
with the requirement of $E$, $F$, and $I$, are equally likely is known as the Postulate of Equal a Priori Probabilities and it was first stated by Ludwig Boltzmann. The ensemble of systems governed by this probability law is known as the Microcanonical Ensemble.

Can Quantum Mechanics help to overcome this situation? The answer is actually no. QM in fact implies that even at the microscopic level the state of a physical system is always incomplete. The state of the system is determined by the wave function $\Psi$ whose $|\Psi|^2$ is the probability density to find the system in a given state. Moreover, only certain possible wave functions are allowed (each for a given allowed state with a fixed energy). But the state (or its wave function) is completely determined. However, for a large system this is no longer possible. The level spacing (in general) becomes very small and, once again, we can only give the probability of a state with given energy. In QM, this state is said to be a mixed
State since the wave function is only state since only the probability of the state is known.

Simple Example of Statistical Physics: Elementary Kinetic Theory

Pressure: \( \frac{F}{A} = P \)

Here \( F \) is the force that we must apply on the piston to keep the gas from expanding.

Thus from a macroscopic standpoint (thermodynamics), the pressure \( P \) is a parameter that characterizes the state of the gas. Another variable is the volume \( V \) occupied by the gas.

Microscopically, the gas has \( N \) atoms (or molecules) \((i=1, \ldots, N)\) each with mass \( m \) (monatomic gas) and each atom has velocity \( \mathbf{v}_i \) at time \( t \).

The velocities are in general very different, but there will be an average speed and the speed of most atoms will be close to the average speed.
Let's assume for simplicity that the atoms do not collide with each other but only with the rigid walls of the container (this is what we call an Ideal Gas). Between two collisions with the wall (or the piston) the velocity of each atom remains unchanged. Upon a collision we find

\[ \vec{v}_c = \vec{v} \]

\[ \vec{v}_p = \vec{v} \]

\[
\begin{align*}
\vec{v}' &= \vec{v} \\
\vec{v}'_x &= \vec{v}_x \\
\vec{v}'_y &= -\vec{a}_g \\
\vec{v}'_z &= \vec{v}_z
\end{align*}
\]

Change in momentum:

\[ \Delta \vec{p} = m \vec{v} - m \vec{v}' = m (\vec{v} - \vec{v}') \]

\[ = 2m \vec{a}_g \hat{y} \]

(Here \( \hat{y} \) is the unit vector in the y direction.)

**Force:** rate of change of momentum: \( \frac{d\Delta \vec{p}}{dt} \) (Newton's)

There is a force exerted against the wall (and piston) by the atoms that collide with the wall.

How many collisions in \( dt \)? Only by those atoms...
that can hit the wall between $t$ and $t + \Delta t$.

Say that an atom $A$ is at a distance $d$ from the wall and has velocity $v_y = v$. If $d < v \Delta t$ then the atom will hit the wall in the interval $\Delta t$ and it won’t if $d > v \Delta t$.

How many atoms hit the wall?

There are $N$ atoms in $V$. Thus the density is

$$\rho = \frac{N}{V} \quad (# \text{ atoms per unit volume}).$$

The piston has area $A$. The atoms that hit the piston in $\Delta t$ occupy a volume $A v_y \Delta t$ and there are

$$\rho \Delta t A v_y \Delta t$$

atoms in that volume (total).

The change in momentum is

$$\Delta P = (\rho \Delta t A v_y) 2m v_y \hat{y}$$

$$\Rightarrow \quad \dot{P} = \frac{\Delta P}{\Delta t} = \hat{y} \rho \Delta t A 2m v_y^2$$

Pressure:

$$P = 2m v_y^2 \rho$$
But all atoms have $\neq$ velocity

$$\Rightarrow \rho = 2\rho m \langle v_j^2 \rangle^{\frac{1}{2}}$$

(since, on average, $\frac{1}{2}$ of the atoms move forward and $\frac{1}{2}$ backwards)

$$\langle v_j^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} v_i^2$$

$$\Rightarrow \rho = \rho m \langle v_j^2 \rangle$$

But the gas is isotropic (i.e. no preferred direction)

$$\Rightarrow \langle v_j^2 \rangle = \langle v_x^2 \rangle = \langle v_z^2 \rangle = \frac{1}{3} \langle \overline{v}^2 \rangle$$

$$\Rightarrow \rho = \frac{1}{3} \rho m \langle \overline{v}^2 \rangle$$

$$\langle \overline{v}^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle$$

$$\Rightarrow \rho = \frac{2}{3} \rho \langle \frac{1}{2} m \overline{v}^2 \rangle$$

We recognize $\frac{1}{2} m \overline{v}^2$ as the kinetic energy of each atom $\Rightarrow$

$$p = \frac{2}{3} \rho \langle K \rangle \frac{N}{V}$$

For this case (no potential energy), the total energy $U = NK + V = NK$

$$\Rightarrow \frac{2}{3} \rho \frac{U}{V} \Rightarrow \boxed{PV = \frac{2}{3} U} \quad \text{(Boyle)}$$
In your undergraduate class you showed...
What did we do: 1) We made a microscopic model of the
gas whose behavior we predicted using Classical Mechanics
(Newton's laws). We calculated changes in momenta
due to collisions, estimated the number of collisions and
the total momentum change. 2) We used statistical arguments
to relate the pressure to the total energy.
The result is a relationship between the pressure $P$,
the volume $V$ and the total energy $U$, which
we derived from a microscopic kinetic consideration.
The relation

$$pV = \frac{2}{3} U$$

(1)
is a thermodynamic relation which connects the
thermodynamic variables $p, V, \text{and } U$ for a gas in
equilibrium (i.e. the averages do not change with
time). This equation states that all
thermodynamic variables can be independent of
the system is in equilibrium. This is also
known as an equation of state.
Eqn. (1) relates microscopic variables without reference to
a microscopic parameter.
Statistical Ensembles

More formally, let us consider a macroscopic body or system of bodies, and assume that this system is closed: i.e., it does not interact with any other bodies. A small part of the system, small but large enough to be considered macroscopic, can be regarded as being separate from the rest. Such part will be called a subsystem. However, a subsystem by definition is not closed since it interacts with other subsystems in complex ways. We will represent a subsystem by the coordinate and momentum of its center of mass. Let \( V \) be the volume of phase space occupied by the subsystem.

After some long time \( T \), the trajectory of the subsystem in phase space will pass many times through each such possible volume. The postulate of equal a-priori probabilities is equivalent to the statement that the fraction of time a subsystem
speeds in some volume $d\mathbf{q}d\mathbf{p}$ is the same for all such volumes accessible for a certain total energy, total momentum and total angular momentum. Thus we can define the probability of the state as the fraction of time the system spends there.

$$W = \lim_{T \to \infty} \frac{\Delta t}{T}$$

This limit is constant in time.

$$\Rightarrow \Delta W = f(q_i, p_i) \frac{d\mathbf{p}d\mathbf{q}}{\text{volume of phase space}}$$

$f(q, p)$ is the statistical distribution function. It is normalized

$$\int d\mathbf{p}d\mathbf{q} \; f(q, p) = 1$$

Notice that since this distribution function is time-independent, it is independent of the initial condition! Thus is an equilibrium state.

If we know $f(q, p)$ we can compute
average of any physical observable \( f(\theta, p) \)

\[
<f> = \int d\theta dp \ f(\theta, p) \ \mathcal{L}(\theta, p)
\]

From the assumption that the distribution is

true independent (and from the definition of

the distribution) that \(<f>\) is also equal to

it follows

\[
<f> = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(t)
\]

Thus a complex deterministic system behaves

(under these hypotheses) as a probabilistic system.

We have assumed above that the subsystems

are not closed but interact weakly with each

other. By weakly we mean that the

interaction are of sufficiently short range

that their effects grow with the surface

of the subsystem and not its volume.

From this condition it follows that if \( A \) and \( B \)

are two subsystems they can be regarded as

being statistically independent. Hence, the
Joint probability distribution factors

\[ P_{AB} = P_A P_B \]

In practice, this condition is much too restrictive but what will matter is that if a system is composed of \( N \) subsystems \( \Rightarrow \)

\[ \ln P \leq \sum_{i=1}^{N} \ln p_i \leq N \max_i (\ln p_i) \]

\[ \Rightarrow \ln P \text{ is an extensive quantity as } N \to \infty \]

For statistically independent quantities, the following results follow. Let \( f \) be a physical additive quantity and \( p_i \) its value for a subsystem \( \Rightarrow \)

\[ \sum_{i=1}^{N} f_i = N \langle f \rangle \]

Likewise \( \Delta f_c = f - \langle f \rangle \)

\[ \langle (\Delta f)^2 \rangle = \langle \left( \sum_i \Delta f_i \right)^2 \rangle = \]

\[ = \sum_{i,j} \langle \Delta f_i \Delta f_j \rangle = \sum_i \langle \Delta f_i \rangle^2 + \sum_{i \neq j} \langle \Delta f_i \Delta f_j \rangle \]

\[ = \sum_i \langle \Delta f_i \rangle^2 \]

\[ \frac{1}{\langle (\Delta f)^2 \rangle} \sim \frac{1}{\sqrt{N}} \]
In the theory of probability, it is proven that the sum of weakly correlated random variables obey the central limit theorem, which implies that the conditions stated above are true.

**Liouville's Theorem**

Let us discuss the general properties of the distribution function \( f(q, p) \). We can draw the phase space evolution of the system by saying that if we observe it at times \( t_1, \ldots, t_n \), at each time the subsystem will be found at a set of points \( P_1, \ldots, P_n \) of phase space. Alternatively, \( P_1, \ldots, P_n \) can be regarded as the places in phase space at which an ensemble of identically prepared but independent subsystems are found at the same time \( t \). Then, the distribution function will tell us what is the probability to find a subsystem of the ensemble at a particular
state \( \mathbf{P} = (q_0, p_0, \ldots) \) of phase space. Each subsystem is then one point in phase space, which is a 2N-dimensional Euclidean space (for N coordinates). We can think of this as a gas of particles (each particle being a subsystem). Since the \# of subsystems is fixed (and unaltered) the density \( \rho \) obeys a continuity equation:

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0
\]

\( \mathbf{v} \) : "velocities"

Steady state: \( \frac{\partial \rho}{\partial t} = 0 \)

\[
\rightarrow \quad \nabla \cdot (\rho \mathbf{v}) = 0 \Rightarrow \sum_{i=1}^{N} \left[ \frac{\partial}{\partial q_i} \left( \rho \frac{dq_i}{dt} \right) + \frac{\partial}{\partial p_i} \left( \rho \frac{dp_i}{dt} \right) \right] = 0
\]

\[
\sum_{i=1}^{N} \left( \frac{\partial \rho}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial \rho}{\partial p_i} \frac{dp_i}{dt} \right) + \rho \sum_{i=1}^{N} \left[ \frac{\partial}{\partial q_i} \left( \frac{dq_i}{dt} \right) + \frac{\partial}{\partial p_i} \left( \frac{dp_i}{dt} \right) \right] = 0
\]

Hamilton's Eqs.

\[
\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}
\]
\[
\Rightarrow \frac{\partial}{\partial \mathbf{q}_i} \left( \frac{d\mathbf{q}_i}{dt} \right) = \frac{\partial}{\partial \mathbf{q}_i} \frac{\partial^2 H}{\partial \mathbf{p}_i \partial \mathbf{q}_i} = \frac{\partial^2 H}{\partial \mathbf{q}_i \partial \mathbf{p}_i} \\
\frac{\partial}{\partial \mathbf{p}_i} \left( \frac{d\mathbf{p}_i}{dt} \right) = -\frac{\partial}{\partial \mathbf{p}_i} \frac{\partial H}{\partial \mathbf{q}_i} = -\frac{\partial^2 H}{\partial \mathbf{q}_i \partial \mathbf{p}_i} \\
\Rightarrow \frac{d\mathbf{s}}{dt} = \sum_{i=1}^{n} \left( \frac{\partial}{\partial \mathbf{q}_i} \frac{d\mathbf{q}_i}{dt} + \frac{\partial}{\partial \mathbf{p}_i} \frac{d\mathbf{p}_i}{dt} \right) = 0
\]

Lioville's Theorem

\Rightarrow \mathbf{s} \text{ is constant along a classical trajectory.}

of the subsystems.

\[\text{A consequence of Liouville's Theorem is that } \mathbf{s} \text{ depends only on the constants of motion. In addition, we saw before that } \mathbf{s} \text{ must be an additive function, } \Rightarrow \text{ only additive integrals of motion are allowed.}\]

In a generic mechanical system, there are only additive integrals of motion (there are exceptional systems that have more): the total
Energy \( E \), the total linear momentum \( \vec{P} \) and the total angular momentum \( \vec{M} \). For a subsystem \( \alpha \)

\[ \log \alpha = \alpha \left[ g, P \right] + \tilde{g} \cdot \vec{P} \left[ g, P \right] + \tilde{g} \cdot \vec{M} \left[ g, P \right] \]

\[ \text{independent of the subsystem!} \]

\[ \text{Normalization constant} \]

If we go to the center of mass frame \((\vec{P} = 0)\) and the system is not rotating \((\vec{M} = 0)\)

\[ \log \alpha \text{ depends only on the energy, } E. \]

This is almost (but not quite) correct. The manifold of points accessible to any subsystem is the phase space restricted by the condition that \( E, \vec{P} \) and \( \vec{M} \) are fixed. This is an \( N-1 \)-dimensional space. \( \Rightarrow \) if \( \int \int \ldots \delta (E - E_0) \delta (P - P_0) \delta (M - M_0) \neq 0 \Rightarrow \)

\( \gamma \) must diverge on the allowed manifold and vanish outside.

\[ \gamma = \text{constant} \times \delta (E - E_0) \delta (P - P_0) \delta (M - M_0) \]

This is the microcanonical Ensemble.
Thus, the Postulate of equal a-priori probability implies that in equilibrium, the system is a member of the microcanonical ensemble whose density distribution function satisfies \( \Delta \ll E \):

\[
P(p, q) = \begin{cases} \text{const.} & \text{if } E < H(q, p) < E + \Delta \\ 0 & \text{otherwise} \end{cases}
\]

All members of the ensemble have the same \( N, V \). Let \( \Gamma(E) \) be the volume of phase space occupied by the microcanonical ensemble, i.e.,

\[
\Gamma(E) = \int_{E < H(q, p) < E + \Delta} d^3q \, d^3p = \text{H of states in the energy shell}
\]

and let \( \Sigma(E) \) be

\[
\Sigma(E) = \int_{H(q, p) < E} d^3q \, d^3p
\]
\[ \Gamma(E) = \Sigma(E + \Delta) - \Sigma(E) \]

If \( \Delta \ll E \Rightarrow \Gamma(E) = \omega(E) \Delta + O(\Delta^2) \)

where \( \omega(E) \) is the density of thermodynamic states.

\[ \omega(E) = \frac{\partial \Sigma(E)}{\partial E} \]

Let us consider the quantity

\[ S(E,V) = k \log \Gamma(E) \]

which we will show behaves like the thermodynamic entropy. Let us prove 2 properties:

(a) \( S(E,V) \) is extensive.

(b) \( S(E,V) \) satisfies the properties of the entropy as required by the Second Law of Thermodynamics.

(c) Extensivity: let us divide the system into two subsystems with \( N_1 \) particles and volume \( V_1 \) and \( N_2, V_2 \) respectively. We will assume that the total energy (Hamiltonian) is

\[ H(p,q) = H_1(q_1, p_1) + H_2(q_2, p_2) \]

We have ignored the interaction term. In practice...
our conclusion will continue to hold provided
the interaction do not grow with the volume
of the system. We are also assuming that
the particles are distinguishable. We will see
that this will lead to problems which we will deal
with later on. (Gibbs Paradox)
Thus if we have two separate subsystems we
can write

$$S_1(E_1, V_1) = k \log \Gamma_1(E_1)$$

$$S_2(E_2, V_2) = k \log \Gamma_2(E_2)$$

For the composite system, the total energy
will lie between $E$ and $E + 2\Delta$, i.e.,

$$E < E_1 + E_2 < E + 2\Delta$$

```
What is the volume of the composite system in
phase space? Obviously it is $\Gamma_1(E_1) \Gamma_2(E_2)$

However, we will only specify the total energy $E$

$$\Rightarrow \Gamma(E) = \sum \Gamma_1(E_i) \Gamma_2(E - E_i)$$

Intervals

$$\Rightarrow$$

The entropy for a composite system will
\[ N = N_1 + N_2 \quad \text{and} \quad V = V_1 + V_2 \]

\[ S(E,V) = k \log \left( \sum_{i=1}^{E/\Delta} P_1(E_i) P_2(E-E_i) \right) \]

However, as \( N_1 \to \infty \) and \( N_2 \to \infty \) there is a configuration with \( E_1, E_2 \), which makes the probability maximal. This term (one among many positive terms) diminishes the sum in this limit.

Clearly, we can bound \( \Gamma(E) \) as follows:

\[ P_2(E_1) P_2(E_2) \leq \Gamma(E) \leq \frac{E}{\Delta} P_1(E_1) P_2(E_2) \]

\( \# \) of energy intervals

\( \Rightarrow \)

\[ k \log \left( P_1(E_1) P_2(E_2) \right) \leq S(E,V) \leq k \log \left( \frac{E}{\Delta} P_1(E_1) P_2(E_2) \right) + k \log \left( \frac{E}{\Delta} \right) \]

since \( P_1(E_1) \gg \frac{E}{\Delta} \) and \( P_2(E_2) \gg \frac{E}{\Delta} \)

\( \Rightarrow \)

\[ S(E,V) \to k \log P_1(E_1) + k \log P_2(E_2) \quad \text{as} \quad N \to \infty \]

\[ = S_1(E_1, V_1) + S_2(E_2, V_2) \]

\( \Rightarrow \) \( S \) is additive.
Moreover, what we just proved is that two systems weakly coupled to each other will have energies $E_1$ and $E_2$, $E = E_1 + E_2$, chosen in such a way that the probability is largest.

$$\Rightarrow \delta \left( \Gamma_1(E_1) \Gamma_2(E_2) \right) = 0 \quad \delta (E_1 + E_2) = 0$$

$$\frac{\partial \log \Gamma_1(E_1)}{\partial E_1} \bigg|_{E_1 = E_1} = \frac{\partial \log \Gamma_2(E_2)}{\partial E_2} \bigg|_{E_2 = E_2}$$

$$\Rightarrow \text{the configuration of maximum probability is such that the quantity}$$

$$\frac{\partial \log \Gamma(E)}{\partial E}$$

is the same for all systems in contact and in equilibrium.

We will define the absolute temperature $T$ s.t.

$$\frac{1}{T} = \frac{\partial S}{\partial E} \Rightarrow T_1 = T_2 \quad \text{This is the Zeroth law of thermodynamics}.$$
Thus in statistical mechanics there is a natural definition of entropy. Entropy is the logarithm of the number of accessible states. The principle of maximum probability is the principle of maximum entropy. This also implies that in a closed system the equilibrium state is the maximum entropy state. Entropy cannot decrease. This is a probabilistic law. This is the nature of the Second Law of Thermodynamics.

Properties of the Entropy

Consider two systems, I and II. Suppose we remove an amount of energy \( \Delta U \) from system I and that we add it to system II.

\[
\Delta E_I = - \Delta U \\
\Delta E_{II} = + \Delta U
\]

How does the entropy change?
Suppose that initially system I was at temperature \( T_1 \) and system II at \( T_2 \).

\[
\frac{1}{T_1} = \frac{\partial S_I}{\partial E_I} \quad \frac{1}{T_2} = \frac{\partial S_{II}}{\partial E_{II}}
\]

The entropy of the combined system is

\[
S = S_I + S_{II} = S(E, V) = S(E_I, E_{II})
\]

How does \( S \) change as \( \Delta E_I = -\Delta U \)

\[
\Delta E_{II} = +\Delta U
\]

\[
\Delta S = \frac{\partial S}{\partial E_I} \bigg|_{E_{II}} \Delta E_I + \frac{\partial S}{\partial E_{II}} \bigg|_{E_I} \Delta E_{II}
\]

\[
\Delta S = \frac{\partial S_I}{\partial E_I} \bigg|_{E_{II}} \Delta E_I + \frac{\partial S_{II}}{\partial E_{II}} \bigg|_{E_I} \Delta E_{II}
\]

Since \( \frac{\partial S_I}{\partial E_{II}} = \frac{\partial S_{II}}{\partial E_I} = 0 \)

\[
\Rightarrow \Delta S = \frac{1}{T_1} \Delta E_I + \frac{1}{T_2} \Delta E_{II} = \left( \frac{1}{T_1} - \frac{1}{T_2} \right) \Delta U
\]
\[ \Rightarrow \text{If } T_1 > T_2 \Rightarrow \Delta S > 0 \]
and \[ T_1 < T_2 \Rightarrow \Delta S < 0 \]

\[ \Rightarrow \text{if we transfer energy from a hotter body to a colder body, the entropy increases.} \]

Since the entropy is the log of the # of accessible states, the final state with more accessible states is more probable. \[ \Rightarrow \]

the physically allowed state is the one with largest probability \[ \Rightarrow \text{Second Law.} \]

This stops in the equilibrium state where \[ T_1 = T_2 \]

**Definition:**

Heat: is the spontaneous flow of energy between and two systems in thermal contact at different temperature.

In other words, it is the exchange of energy not due to the work of external forces. Let us denote by \( Q \) this exchange of energy and let \( W \) be the work done by external forces on system.
\[ I \left( W \text{ in System II} \right) = \]
\[ \Delta E_I = -Q + W_I \]
\[ \Delta E_{II} = +Q + W_{II} \]

In general
\[ \Delta E = Q + W \]

First Law of Thermodynamics

\( Q > 0 \) if heat flows into the system
\( W > 0 \) if work is done on the system.

Quantum Systems: The Density Matrix

In quantum mechanics, the physical states allowed to an isolated (finite) system are pure states.

Pur\( e \) states are linear superpositions of a set of basis states, \( \{ |\psi_n\rangle\} \). This set is linearly independent and complete, i.e.

\[ \hat{I} = \sum_n |\psi_n\rangle \langle \psi_n| \]

Projection operator on the state \( |\psi_n\rangle \)
Thus a pure state is

\[ |\Psi\rangle = \sum_{n} c_{n} |\Psi_{n}\rangle \]

\(c_{n} \in \mathbb{C}\)

Once this state is given at time \(t_{0}\), the ultimate evolution of the system is fully determined.

\[ i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = \hat{H} |\Psi(t)\rangle \quad \text{Schrödinger's Equation} \]

with \( |\Psi(t_{0})\rangle = |\Psi\rangle \) \(\Rightarrow\) phase coherence.

\(\Rightarrow\) the functions \(c_{n}(t)\) describe the evolution of the system (in the basis \( \langle |\Psi_{n}\rangle |\Psi\rangle \))

For a macroscopic system this is as hard to do as it is for a classical system.

In practice systems are never completely isolated.

The true "system" is the system + the "universe".

Let the Hilbert space of the combined system be spanned

states of the form \( |\Phi_{i}\rangle |\Theta_{j}\rangle \equiv |\Phi_{i},\Theta_{j}\rangle \)

\(\uparrow\) system \(\uparrow\) universe
\[ 14 > = \sum_{i,j} c_{i,j} \ket{\varphi_i} \ket{\Theta_j} \]

("Schmidt decomposition")

Let \( x \) be the "coordinates" of the system and \( \bar{x} \) be the "coordinates" of the rest \( \Rightarrow \)

\[ \varphi_i(x) = \langle x | \varphi_i \rangle \]
\[ \Theta_j(\bar{x}) = \langle \bar{x} | \Theta_j \rangle \]
\[ \Rightarrow \psi(x, \bar{x}) = \langle x, \bar{x} | \psi \rangle \]

\[ = \sum_{i,j} c_{i,j} \varphi_i(x) \Theta_j(\bar{x}) \]

\[ = \sum_i \left( \sum_j c_{i,j} \Theta_j(\bar{x}) \right) \varphi_i(x) \]

\[ c_{i}(\bar{x}) = \sum_j c_{i,j} \Theta_j(\bar{x}) \]

Let \( \hat{A} \) be a physical observable which acts only on the states of the system \( \psi(x, \bar{x}) \)

\[ \hat{A} = \sum_{i,i'} A_{i,i'} \ket{\varphi_i} \bra{\varphi_{i'}} \]
\[ = \sum_i \sum_j A_{i,j} \ket{\varphi_i} \bra{\Theta_j} \bra{\Theta_j} \]
\[ \text{such that} \quad \hat{A} \ket{\Theta_j} = \ket{\Theta_j} \]
\[ \text{completeness} \]
\[ \langle \Psi | \hat{A} | \Psi \rangle = \sum_{i,i'} c_{i,j}^{*} c_{i',j'} \langle \Psi_i \theta_j | \hat{A} | \Psi_{i'} \theta_{j'} \rangle \]

\[ = \sum_{i,i'} c_{i,j}^{*} c_{i',j'} A_{i'i'} \]

\[ = \sum_{i,i'} \left( \sum_{j} c_{i,j}^{*} c_{i',j'} \right) A_{i'i'} \]

Density matrix: \[ P_{i,i'} = \sum_{j} c_{i,j}^{*} c_{i',j'} = \langle \Psi_i | \hat{P} | \Psi_{i'} \rangle \]

\[ \langle \Psi | \hat{A} | \Psi \rangle = \sum_{i,i'} P_{i,i'} \hat{A}_{i'i'} = \text{tr} (\hat{P} \hat{A}) = \text{tr} \left( \hat{A} \hat{P} \right) \]

\[ s_{i,i'}^{*} = \sum_{j} c_{i,j}^{*} c_{i',j} = S_{i,i'} \]

\[ \Rightarrow \hat{P} = \hat{P}^{+} \quad \hat{P} \text{ is hermitian.} \]

And \[ \text{tr} \hat{P} = \sum_{i} s_{i,i} = \sum_{i} \sum_{j} c_{i,j}^{*} c_{i,j} \]

\[ = \sum_{i,i',j} |c_{i,j}|^2 = |\langle \Psi | \Psi \rangle|^2 \]

\[ \text{I + } |\langle \Psi | \Psi \rangle| = 1 \Rightarrow \text{[tr} \hat{P} = 1] \]
let \( |w_i\rangle \) be a basis of eigenvectors of \( \hat{\mathcal{Q}} \)
\[
\hat{\mathcal{Q}} |w_i\rangle = w_i |w_i\rangle
\]
\[
\hat{\mathcal{Q}} = \sum_i w_i |w_i\rangle \langle w_i |
\]
\[
\text{tr} \hat{\mathcal{Q}} = \sum_i w_i = 1
\]
Consider \( \hat{A} = |i\rangle \langle i'| = \sum_j |j\rangle \langle j| |i\rangle \langle i'| 
\]
\[
\Rightarrow \langle \hat{A} \rangle = \langle \psi | \hat{A} | \psi \rangle = \sum_j | \langle \psi | j \rangle \langle j | \psi \rangle |^2 \geq 0
\]
But \( \langle \hat{A} \rangle = \text{tr}(\hat{A} \hat{\mathcal{Q}}) = \sum_i \langle j | i \rangle \langle i | \psi \rangle \langle \psi | j \rangle \langle j | i \rangle = \sum_i | i \rangle \langle i | \psi \rangle = \psi
\]
\[
\Rightarrow w_i \geq 0 \quad \text{and} \quad \sum_i w_i = 1 \Rightarrow \text{probabilities!}
\]

We can reformulate theorem as follows

\( 1 \) Any system is describable by a density matrix
\[
\hat{\mathcal{Q}} = \sum_i w_i |i\rangle \langle i|
\]
where \( \{ |i\rangle \} \) is a complete set of orthogonal vectors
\( 2 \) \( w_i \geq 0 \)
\( 3 \) \( \sum_i w_i = 1 \)
\[ \langle A \rangle = \text{tr}(\hat{A} \hat{A}^\dagger) = \sum_i w_i \langle i | \hat{A}^\dagger i \rangle \]

\[ \langle i | \hat{A}^\dagger i \rangle \] is the expectation value of \( \hat{A} \) in state \( i \).

\( w_i \) is the probability to find the system in state \( i \). We have an ensemble of the systems with probability distribution \( \{ w_i \} \).

We will call this a mixed state. A pure state is a state in which only one \( w_i \neq 0 \).

\[ \therefore w_i = 1 \]

\[ w_j = 0 \quad j \neq i \]

\[ \Rightarrow \langle A \rangle = \langle \hat{A} \hat{A}^\dagger \rangle \]

**Equation of Motion for \( \hat{a} \):**

\[ \hat{a}(t) = \sum_i w_i \left| \psi_i(t) \right\rangle \left\langle \psi_i(t) \right| \]

\[ \frac{\partial}{\partial t} \left| \psi(t) \right\rangle = \hat{A} \left| \psi(t) \right\rangle \]
\[
-i \frac{\hbar}{\partial t} \langle \psi(t) | = \langle \psi(t) | \hat{H} \\
\implies \frac{\partial \hat{N}}{\partial t} = \sum_i w_i \frac{1}{\sqrt{\hbar}} \langle \psi(t) | \hat{N} | \psi(t) \rangle - \sum_i w_i \frac{1}{\sqrt{\hbar}} \langle \psi(t) | \hat{N} | \psi(t) \rangle \\
i \frac{\hbar}{\partial t} \frac{\partial \hat{N}}{\partial t} = \hat{H} \hat{N} - \hat{N} \hat{H} \\
\implies \left[ i \frac{\hbar}{\partial t} \frac{\partial \hat{N}}{\partial t} = [\hat{H}, \hat{N}] \right]
\]

Quantum mechanical Liouville's theorem.

and

\[
\frac{d}{dt} \langle \hat{A} \rangle = \frac{d}{dt} \text{tr} (\hat{\rho} \hat{A}) = \text{tr} \left( \frac{d\hat{\rho}}{dt} \hat{A} \right) + \text{tr} \left( \hat{\rho} \frac{d\hat{A}}{dt} \right) \\
= \frac{1}{i\hbar} \text{tr} \left( [\hat{H}, \hat{\rho}] \hat{A} \right) + \text{tr} \left( \hat{\rho} \frac{d\hat{A}}{dt} \right) \\
\frac{d}{dt} \langle \hat{A} \rangle = \langle \frac{d\hat{A}}{dt} \rangle + \frac{1}{i\hbar} \text{tr} \left( [\hat{A}, \hat{H}] \hat{\rho} \right) 
\]
Microcanonical Ensemble:

\[ \rho_{\text{cc}} = w_i \delta (\epsilon_i - \epsilon) \quad (i.e. \hat{\rho} \text{ is diagonal}) \]

\[ w_i = \begin{cases} \text{const} & \text{for } E \leq E_i \leq E + \Delta \\ 0 & \text{otherwise} \end{cases} \]

i.e. all accessible states have the same probability. (where \( \hat{H} | E_n \rangle = E_n | E_n \rangle \))

\[ \Rightarrow \hat{\rho} = \sum \limits_{E \leq E_n \leq E + \Delta} | \psi_n \rangle \langle \psi_n | \quad \text{(unnormalized)} \]

\[ \text{Tr} \hat{\rho} = \sum \limits_n p_{nn} = \Gamma(E) \quad \text{is the \# of states with energy in the interval } [E, E + \Delta] \]

\[ \Rightarrow \text{Entropy} \]

\[ S(E, V) = k \log \Gamma(E) = \log \text{det} \hat{\rho} \]

\[ = k \log (\text{Tr} \hat{\rho}) \quad \text{for } \hat{\rho} \text{ unnormalized} \]
The Canonical Ensemble (the Gibbs Ensemble)

Consider a physical system in equilibrium at temperature \( T \). What is the probability to find the system in a state with energy \( E \)?

Q1: How do we maintain a system at fixed temperature?

A: Let \( R \) be a very large system already in equilibrium at temperature \( T \). Let \( S \) be the system we are interested in.

R: reservoir or heat bath

Let us assume that the joint system \( R S \) is isolated. After \( R \) and \( S \) become in thermal contact they reach an equilibrium state with common temperature \( T' \). In this process an amount \( \Delta U \) of energy is exchanged between \( R \) and \( S \).

\[
\Delta T = T' - T = \frac{\Delta U}{C_m_{reservoir}}
\]

\( C_m_{reservoir} \) is the specific heat of the reservoir.

Since \( m_{reservoir} \to \infty \Rightarrow (\Delta T)_{reservoir} = 0 \Rightarrow T \to \text{const.} \)
Since $S'OR$ is isolated \[ E = U_R + U_S \]

In state $|n\rangle$, the system has energy $U_S = E_n$ and $U_R = E - E_n$

The probability of being in state $|n\rangle$ is proportional to the number of quantum states of the joint system when $S'$ is in $|n\rangle$. This equals to the number of quantum states of $R$ when $S'$ is in $|n\rangle$ times the number of quantum states of $S'$ when $S'$ is in $|n\rangle$ ($= 1$) (proportional to)

$\Rightarrow$ The probability is \[ \frac{1}{n} \] of the number of quantum states of $R$ when $S'$ is in state $|n\rangle$ = \[ \frac{1}{E_n} \]

Notice that $|\theta| \gg E_n$

Since $R$ is large \[ T \] does not change for small changes of the energy $E_n$

$\Rightarrow$ \[ \text{const.} = \frac{1}{kT} = \frac{\partial S_{\text{reservoir}}}{\partial E_R} \]

\[ \frac{\partial \log P_{\text{res}}}{\partial E} = \frac{1}{kT} = \log P_{\text{res}} = \frac{E_n}{kT} + \text{const.} \]
\[ P_{\text{res}}(E_{\text{res}}) = e^{\frac{E_{\text{res}}}{kT}} \times \text{const.} \]

\[ \Rightarrow \] Probability to find the system with energy \( E_n \) at temperature \( T \) is \( \propto P_{\text{res}}(E_{\text{res}}) \)

\[ P(E_n) = \text{const} \times e^{E_{\text{res}}/kT} = \text{const} \times e^{(E - E_n)/kT} \]

\[ \Rightarrow P(E_n) = \text{const}' \times e^{-E_n/kT} \]

Normalization

\[ \sum_n P(E_n) = 1 \]

\[ 1 = \text{const}' \times \sum_n e^{-E_n/kT} \]

Partition Function

\[ Z = \sum_n e^{-E_n/kT} \]

\[ P(E_n) = \frac{1}{Z} e^{-\frac{E_n}{kT}} \]

Gibbs Distribution

(Gibbs Ensemble) or Canonical Ensemble.
Entropy and the Canonical Ensemble

In the Microcanonical Ensemble the entropy is the logarithm of the # of available states at fixed energy. For the canonical ensemble the energy is not fixed but the probability to find the system at that energy is given by the Gibbs distribution. The entropy now is the expectation value of the log of the # of states at each energy:

\[ S = -k \sum_{n} P_n \ln P_n = -k \langle \ln P \rangle = -k \text{tr}(\hat{\rho} \ln \hat{\rho}) \]

where \( \hat{\rho} \) is the density matrix for the Gibbs distribution. Here I used \( \text{tr}(\hat{\rho}) = \sum_{n} P_n = 1 \).

Using this definition we get the following result:

\[ S = -k \sum_{n} P_n \ln P_n = -k \sum_{n} e^{-\frac{E_n}{kT}} \left[ -\frac{E_n}{kT} - \ln Z \right] \]

\[ = \frac{1}{T} \langle E \rangle + k \ln Z \]
If we now define the Free Energy \( F \)
\[
F = -kT \ln Z
\]
\[
\Rightarrow \frac{\partial F}{\partial T} = -S
\]
we will shortly see that this relation holds in thermodynamics.

Similarly, the internal energy \( U \) = \( \langle E \rangle \)
\[
U = \langle E \rangle = \frac{1}{Z} \sum_n E_n e^{-E_n/kT}
\]

If we now define \( \beta = \frac{1}{kT} \) \( \Rightarrow \beta F = -\ln Z \)
\[
\Rightarrow \langle E \rangle = \frac{\partial \ln Z}{\partial \beta} = \frac{\partial (\beta F)}{\partial \beta}
\]
Thus we can compute the internal energy and the entropy once we know the partition function.