What is the disorder of the two state system, \( S(p_1, p_2) \)?

Consider \( N \) copies of the two state system.
By additivity of \( S \) we want the disorder of this joint system of \( N \) copies to be

\[
(*) \quad S = NS(p_1, p_2)
\]

Now in any given sample of the \( N \) copy system, 
\( M \) of the systems will be in state \( \uparrow \), while 
\( N - M \) are in state \( \downarrow \). The prob for this will be given by the binomial distribution

\[
P_M = \frac{N!}{M!(N-M)!} \quad p_1^M \quad p_2^{N-M} \quad \text{Prob} [M \text{ in state } \uparrow ]
\]

For \( N \) large, this probability is very strongly peaked about the average \( M = Np_1 \). We have

\[
\text{average \# systems in state } \uparrow = \langle n_\uparrow \rangle = Np_1
\]

\[
\text{standard deviation of \# in state } \uparrow = \sqrt{\langle n_\uparrow^2 \rangle - \langle n_\uparrow \rangle^2} = \sqrt{Np_1p_2}
\]

So relative width of distribution \( \frac{\sqrt{\langle n_\uparrow^2 \rangle - \langle n_\uparrow \rangle^2}}{\langle n_\uparrow \rangle} \sim \frac{1}{\sqrt{N}} \)

\( \Rightarrow 0 \) as \( N \to \infty \).

\( \Rightarrow \) as \( N \) gets large we almost always find the system
of \( N \) copies with \( Np_1 \) in state \( \uparrow \) and \( Np_2 \) in state \( \downarrow \).

How many ways are there to choose which of the \( N \)
two level sub-systems are in state \( \uparrow \)?
There are \( \frac{N!}{(Np_1)! (Np_2)!} \) ways \( (Np_2 = N(1-p)) \) each of these ways are \textit{equally likely}!

\( \Rightarrow \) the entropy of the \( N \) copy system is

\[
S = k \ln \left[ \frac{N!}{(Np_1)! (Np_2)!} \right] \log \text{ of # equally likely states!}
\]

\[
= k \left[ \ln N! - \ln (Np_1)! - \ln (Np_2)! \right]
\]

\[\text{use Stirling formula}\]

\[
= k \left[ N \ln N - N - Np_1 \ln Np_1 + Np_1 - Np_2 \ln Np_2 + Np_2 \right]
\]

\[\text{use } Np_1 + Np_2 = N \text{ as } p_1 + p_2 = 1\]

\[
= k N \left[ -p_1 \ln p_1 - p_2 \ln p_2 \right]
\]

\( \Rightarrow \)

\[
S = k N \left[ -p_1 \ln p_1 - p_2 \ln p_2 \right] \text{ since } p_1 + p_2 = 1
\]

But by (\( \ast \)) we expect \( S = N S(p_1, p_2) \)

\[\Rightarrow S(p_1, p_2) = -k \left[ p_1 \ln p_1 + p_2 \ln p_2 \right] \]

Similarly, if our system had \( m \) possible states, with probabilities \( p_1, p_2, \ldots, p_m \), and we took \( N \) copies of the \( m \) level system, the joint system would have \( Np_1 \) subsystems in state 1, \( Np_2 \) in state 2, \ldots, \( Np_m \) in state \( m \). The number of \textit{equally likely} ways to divide the \( N \) subsystems this way is

\[
\frac{N!}{(Np_1)! (Np_2)! \cdots (Np_m)!}
\]
And so a similar line of argument results in:

\[
S(p_1, \ldots, p_m) = -k \left[ p_1 \ln p_1 + p_2 \ln p_2 + \cdots + p_m \ln p_m \right]
\]

\[
S(\varepsilon p_i) = -k \frac{\varepsilon}{2} p_i \ln p_i
\]

This defines our measure of the disorder of the prob distribution \( p_i \). We see it agrees with what we found for the entropy in both canonical and microcanonical ensembles.

But now we will use it to derive the microcanonical and the canonical ensembles!

\( S \) above agrees with the desired properties (i) and (ii).

\( S = 0 \) if any \( p_i = 1 \) and all others are zero.

We soon see that \( S \) is max if all \( p_i \) are equal.

We can now use the above as our definition of entropy and define equilibrium as the prob distribution that maximizes \( S \), subject to whatever constraints may exist on the distribution. Each such constraint represents some "information" we have about the system.
microcanonical ensemble - each state $i$ has an energy $E_i$

We have $p_i = 0$ for $E_i \neq E$, $p_i > 0$ for $E_i = E$

Considering only those states $i$ with $E_i = E$, we now want to maximize $S$ over these non-zero $p_i$.

We want to maximize $S = -k \sum_i p_i \ln p_i$

subject to the constraint $\sum_i p_i = 1$ (normalization of probabilities)

Use method of Lagrange multipliers

$\Rightarrow$ maximize in an unconstrained way

$S + k \lambda \sum_i p_i$

where $\lambda$ is the Lagrange multiplier - we then determine the value of $\lambda$ by imposing the constraint.

So if there are $N$ states of energy $E$, the maximization gives

$0 = \frac{\partial}{\partial p_i} \left( S + k \lambda \sum_i p_i \right) = \frac{\partial}{\partial p_i} \left( -k \sum_j \left( p_i \ln p_i - \lambda p_i \right) \right)$

$\Rightarrow p_i \left( \frac{1}{p_i} \right) + \ln p_i - \lambda = 0$

$1 + \ln p_i - \lambda = 0$

$\Rightarrow p_i = e^{-1} \quad \text{same for all } i$
\[ \sum \phi_i e^{-\beta E_i} = 1 \quad \Rightarrow \quad \phi_i = \frac{e^{-\beta E_i}}{\sum e^{-\beta E_i}} \]

**Canonical Ensemble**

Now any \( E_i \) is allowed, but we have the constraint that the average energy \( \langle E \rangle \) is fixed \( \Rightarrow \sum \phi_i E_i = \langle E \rangle \)

\[ \sum \phi_i = 1. \]  

Thus the maximization requires two Lagrange multipliers.

\[ 0 = \frac{\partial}{\partial \phi_i} \left( -k \sum \left[ \phi_i \ln \phi_i - \alpha \phi_i + \beta \phi_i E_i \right] \right) \]

\[ \Rightarrow \quad 0 = 1 + \ln \phi_i - \alpha + \beta E_i \]

\[ \phi_i = e^{\lambda-1} e^{-\beta E_i} \]

Normalisation \( \Rightarrow \sum \phi_i = e^{\lambda-1} \sum e^{-\beta E_i} = 1 \)

\[ \Rightarrow \quad e^{\lambda-1} = \frac{1}{\sum e^{-\beta E_i}} \]

\[ \Rightarrow \quad \sqrt{\phi_i} = \frac{e^{-\beta E_i}}{\sum e^{-\beta E_i}} \]

Determine \( \beta \) by condition that

\[ \frac{\sum e^{-\beta E_i} E_i}{\sum e^{-\beta E_i}} = \langle E \rangle \]
If we interpret $\beta = \frac{1}{k_BT}$ we recover the canonical distribution!

More generally, if we had any quantity $X$ constrained, i.e. $X_i$ is value in state $i$, and average value $\langle X \rangle = \sum_i \pi_i X_i$ is fixed, then

$$\pi_i = \frac{e^{-\beta X_i}}{\sum_j e^{-\beta X_j}}$$
gives maximum $S$ consistent with the constraint.

$\beta$ determined by requiring $\frac{\sum_i X_i e^{-\beta X_i}}{\sum_j e^{-\beta X_j}} = \langle X \rangle$
gives the desired value of $\langle X \rangle$.

We can use the definition

$$S = -k_B \sum_i \pi_i \ln \pi_i$$

more generally than for systems in equilibrium in the thermodynamic limit. It can be used just as well for systems of finite size, and for systems out of equilibrium.
Grand Canonical Ensemble

Consider a system of interest which is in contact with both a thermal and a particle reservoir.

System of interest $E, V, N$ — henceforth referred to as "the system".

Walls allow exchange of energy and particles.

Reservoir $E_R, V_R, N_R$.

One way such a situation may arise physically is if the "system of interest" is just a certain volume immersed in a much larger volume of the same "stuff", and the walls around the "system of interest" are just our mental constructs.

Gas in a box

Reservoir is the rest of the gas.

System of interest is some interior region of the gas. Dashed lines are mental constructs, not physical walls.

The energy $E$ and number of particles $N$ in the region of interest are not fixed but fluctuate as energy + particles flow between the region and the rest of the gas.
The reservoir is so large, that no matter how much energy or particles the system of interest transfers to it, its temperature $T_R$ and chemical potential $\mu_R$ do not change – this is what we mean by it being a reservoir.

We see this as we argued before. If heat $dQ = TdS$ is transferred to the reservoir then the change in $T_R$ is

$$\Delta T_R = \frac{2T_R}{\delta S_R} dS = \left(\frac{2^2 E_R}{\delta S_R^2}\right) ds \sim \frac{N}{N_R} T_R \quad \text{as} \quad E_R, S_R \sim N_R$$

$$ds \sim N \text{ at most}$$

So if $N \ll N_R$, $\Delta T_R \ll T_R$

Similarly, if $dN$ is transferred to the reservoir

$$\Delta N_R = \frac{2N_R}{\delta N_R} dN = \left(\frac{2^2 E_R}{\delta N_R^2}\right) dN \sim \frac{N}{N_R} \mu_R \quad \text{as} \quad E_R, N_R \sim N_R$$

$$dN \sim N \text{ at most}$$

So if $N \ll N_R$, $\Delta N_R \ll N_R$

So we regard $T_R$ and $\mu_R$ of the reservoir as fixed.

Now because the "system of interest" is in equilibrium with the reservoir, we have $T = T_R$ and $\mu = \mu_R$. 
Now \( N + N_R = N_T \) \((N+1)\), \( E + E_R = E_T \) \((N+1)\)

\( V, V_R \) are fixed

Similar to what we had for the canonical ensemble, the density of states for the total system of reservoir + system of interest is

\[
g_T(E_T, V, V_R, N_T) = \int \frac{dE}{A} \sum_N g(E, V, N) g_R(E_T - E, V_R, N_T - N)
\]

or for the number of states \( \Omega_T = g_A \) \((\Delta \equiv \text{small energy interval as before})\)

\[
\Omega_T(E_T, V, V_R, N_T) = \int \frac{dE}{A} \sum N \Omega(E, V, N) \Omega_R(E_T - E, V_R, N_T - N)
\]

\[
= \int \frac{dE}{A} \sum_N \Omega(E, V, N) e^{S_R(E_T - E, V_R, N_T - N)/k_B}
\]

Probability density for system to have \( E \) and \( N \) is proportional to the number of states that have the system with \( E \) and \( N \)

\[
p(E, N) \propto \Omega(E, V, N) e^{S_R(E_T - E, V_R, N_T - N)/k_B}
\]

except

\[
S_R(E_T - E, V_R, N_T - N) \approx S_R(E_T, V_R, N_T) + \frac{\partial S_R}{\partial E_T} (-E_T)
\]

\[
= S_R - \frac{E_T}{T} + \frac{\mu N}{T} + \left( \frac{\partial S_R}{\partial N} \right) (-N)
\]

\[
p(E, N) \propto \frac{\Omega(E, V, N) e^{-(E - \mu N)/k_B T}}{\sum_N \int \frac{dE}{A} \Omega(E, V, N) e^{-E/k_B T} e^{\mu N/k_B T}}
\]

Normalized