**Statistical Physics—Section 2: Assignment of Probabilities and Formulation of Statistical Mechanics**

*2. 1. Assignment of Probability*

> **Idea** $\{p\}_r$ should be such as to maximise $S$ subjects to *constraints* imposed by the available information

i.e. $\{p\}_r$ should *only* reflect the information available and are otherwise unbiased. In this way it is a *rational* assignment of probability

**Constraints**: These typically take the form of expectation values of some observables

**Procedure**: Suppose we have two known constraints

$$\langle y \rangle = \sum_i p_i y_i \tag{1}$$

$$\langle z \rangle = \sum_i p_i z_i \tag{2}$$

and of course we should not forget the constraint

$$1 = \sum_i p_i \ . \tag{3}$$

We wish to maximise the missing information, or Gibbs entropy,

$$S = -k \sum_i p_i \ln p_i \tag{4}$$

subject to these constraints.

The maximization can be done using Lagrange's method of undetermined multipliers. To extremize a function $f(x_i)$ of a list of variables $x_i$, subject to the constraint that another function $g(x_i)$ has a prescribed value $g_0$, the procedure is to construct

$$h(x_i) = f(x_i) - \lambda g(x_i) \tag{5}$$

and then extremize $h$ by requiring

$$\frac{\partial h}{\partial x_i} = 0 \tag{6}$$

for all the variables $x_i$ in the list. (In the present case the variables $x_i$ are, of course, the probabilities $p_i$ for the various microstates.) The 'undetermined' multiplier $\lambda$ is then chosen so as to make sure the constrained quantity has the desired value. The label 'undetermined' is because we often don't have to work out the particular value for lambda.

If there are several constraints, each is implemented with its own Lagrange multiplier.

To justify the procedure consider the condition for $f$ to be extremised

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i = 0 \tag{7}$$

but because of the constraint the set $\mathrm{d}x_i$ are not independent i.e. $g$ must be kept constant:

$$dg = \sum_i \frac{\partial g}{\partial x_i} dx_i = 0 \tag{8}$$

However, if we subtract some particular multiple $\lambda$ of (8) from (7) we end up with independent conditions

$$\frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} = 0 \tag{9}$$

to be satisfied for each $i$, which yields (6).

**Mathematical Aside:** What's going on is perhaps easiest to see geometrically. Recall that $\nabla g$ is in the direction of maximal increase of $g$ or equivalently is orthogonal to the level surfaces of $g$. Thus under the constraint that $g$ is constant we must have $\underline{\mathrm{d}r} \cdot \nabla g = 0$. However if $f$ is to be extremised we must have $\mathrm{d}f = \underline{\mathrm{d}r} \cdot \nabla f = 0$ from which we deduce that $\nabla g$ and $\nabla f$ are parallel which yields (9). **End of aside**

*Example 1*: First consider the case where there are no given constraints other than

$$\sum_{i=1}^{r} p_i = 1$$

Then

$$h = -k \sum_i p_i \ln p_i - \lambda \sum_i p_i$$

Then our conditions (6) become

$$\frac{\partial h}{\partial p_i} = -k\left[\ln p_i + 1\right] - \lambda = 0 \Rightarrow p_i = \exp[-1 - \lambda/k]$$

making $p_i$ constant for all $i$. But the normalisation constraint requires

$$\sum_{i=1}^{r} p_i = r \exp[-1 - \lambda/k] = 1 \qquad \Rightarrow p_i = \frac{1}{r}$$

*Example 2:* Now apply the method of Lagrange Multipliers to the three constraint example (1,2,3). We require that for all states $i$,

$$\frac{\partial}{\partial p_i}\left[-k \sum_j p_j \ln p_j - \lambda_1 \sum_j p_j - \lambda_y \sum_j y_j p_j - \lambda_z \sum_j z_j p_j\right] = 0$$

This gives

$$-k(\ln p_i + 1) - \lambda_1 - \lambda_y y_i - \lambda_z z_i = 0$$

or, rearranging (check you can do this!),

$$p_i = \exp(-1 - \lambda_1/k) \exp(-[\lambda_y y_i + \lambda_z z_i]/k) \; .$$

The first factor, $\exp(-1-\lambda_1/k)$ is a constant (independent of $i$) and is there to ensure normalization:

$$1 = \sum_i p_i = \exp(-1-\lambda_1/k)\sum_i \exp\left[-\left(\frac{\lambda_y y_i + \lambda_z z_i}{k}\right)\right]$$

thus

$$p_i = \frac{1}{Z}\exp(-[\lambda_y y_i + \lambda_z z_i]/k) \quad \text{where} \quad Z = \sum_i \exp(-[\lambda_y y_i + \lambda_z z_i]/k)$$

*2. 2. Formulation of Statistical Mechanics*

Recall that the most detailed description possible of the assembly is the microstate. Most generally, embracing quantum mechanics, the microstate $i$ is defined to be the ith solution of the Schrodinger equation (S.E.) for the assembly. Microstate $i$ has energy $E_i$ — note that $E_i(\{X\})$ where $\{X\}$ are extensive thermodynamic co-ordinates such as volume $V$, or number of particles etc.

The macroscopic information which specifies the equilibrium state is the (expectation) values of extensive observables e.g. the internal energy, defined as

$$\overline{E} = \sum_i p_i E_i$$

**The different ensembles of Statistical Mechanics**

At this stage the word 'ensemble' is synonymous with (probability) distribution. Later we shall we will look at the idea of an 'ensemble' more closely.

**Microcanonical Ensemble**: Here the assembly is isolated and only fixed energy states are accessible. So the microstates all have the same energy and otherwise we have no information about the assembly. This corresponds to example 1 and

$$p_i = \frac{1}{\Omega} \tag{10}$$

where $\Omega$ is the number of available microstates. Thus, maximising $S$ in the microcanonical ensemble recovers the 'Principle of equal a priori probabilities' (PEAPP).

**Canonical Ensemble**: Here the assembly may explore states of different $E_i$ i.e. we can think of the assembly being connected to some heat reservoir to allow this. The observable $\overline{E}$ specifies the equilibrium state. Maximising $S$ subject to this constraint (plus probability normalisation)gives

$$\frac{\partial}{\partial p_i}\left[S - \lambda_E \sum_j p_j E_j - \lambda_1 \sum_j p_j\right] = 0$$

which leads to

$$\boxed{p_i = \frac{1}{Z_c}\exp(-\lambda_E E_i/k) \quad \text{where} \quad Z_c = \sum_i \exp(-\lambda_E E_i/k)} \tag{11}$$

$Z_c$ is the canonical partition function. As we shall see when we identify the Lagrange multiplier $\lambda_E$, the canonical distribution is just the Boltzmann distribution.

**Grand Canonical Ensemble**: Here the assembly may explore states of different energy and states with different particle number $N$. We refer to this as an *open* assembly. Since $N$ is an extensive thermodynamic variable it actually affects the solutions of S.E therefore we specify our states $i, N$ with energy $E_{i,N}$ meaning that it is the ith solution of S.E. for an assembly with $N$ particles. To allow both energy and particle number to fluctuate we can think of the assembly being connected to both heat and particle reservoirs. The observables $\overline{E}$ and $\overline{N}$ specify the equilibrium state. Maximising $S$ subject to these constraint (plus probability normalisation) follows the worked example 2

$$\frac{\partial}{\partial p_{i,N}}\left[S - \lambda_E \sum_{i,N} p_{i,N} E_{i,N} - \lambda_N \sum_{i,N} p_{i,N} N - \lambda_1 \sum_{i,N} p_{i,N}\right] = 0$$

Which leads to

$$\boxed{p_{i,N} = \frac{1}{\mathcal{Z}_{gc}} \exp(-\lambda_E E_{i,N}/k - \lambda_N N/k) \quad \text{where} \quad \mathcal{Z}_{gc} = \sum_{i,N} \exp(-\lambda_E E_{i,N}/k - \lambda_N N/k)} \tag{12}$$

$\mathcal{Z}_{gc}$ is the grand canonical partition function.

Thus we have derived the three main ensembles (or distributions) of statistical mechanics. They should look familiar except for the presence of the Lagrange multipliers. In order to identify these we must recap some thermodynamics.

### *2. 3. Thermodynamics*

Let us review the combined 1st and 2nd laws. For a fluid or '$PVT$' system we have

$$\boxed{\mathrm{d}\overline{E} = T\mathrm{d}S - P\mathrm{d}V} \tag{13}$$

where $P$ is the pressure. You should memorise this.

More generally we can write

$$\mathrm{d}\overline{E} = T\mathrm{d}S + \sum_{\gamma} f_\gamma \mathrm{d}X_\gamma \tag{14}$$

where $f_\gamma$ is an applied force and $X_\gamma$ is a thermodynamic variable (sometimes referred to as a displacement) *conjugate* to that force. e.g. in the $PVT$ example the 'force' is $-P$ and the conjugate displacement is the volume. In a magnetic system the force would be the external magnetic field $\underline{H}$ and the displacement would be the magnetisation $\underline{M}$. (Beware that there are many different notations for applied field and magnetisation.) Note that the forces are **intensive** (i.e. don't depend on the size of the assembly) but the displacements $X_\gamma$ are **extensive** (i.e scale linearly with the size of the assembly).

Then one can generalise further to include the effect on the internal energy of changing the number of particles or indeed the number of each species of particle

$$\boxed{\mathrm{d}\overline{E} = T\mathrm{d}S + \sum_{\gamma} f_\gamma \mathrm{d}X_\gamma + \sum_{\alpha} \mu_\alpha \mathrm{d}\overline{N}_\alpha} \tag{15}$$

This is our most general form of the 1st/2nd law. It defines the **chemical potential** $\mu_\alpha$ as the change in internal energy by adding a particle of species $\alpha$ to the system. Clearly the internal energy of the system should change due to e.g. the kinetic energy of the added particle. The definition of the chemical potential may be written

$$\boxed{\mu_\alpha = \left(\frac{\partial \overline{E}}{\partial \overline{N}_\alpha}\right)_{S,\{X_\gamma\}} .}$$

It is the increase in internal energy due to adding a particle, with the other thermodynamic co-ordinates $S, \{X_\gamma\}$ held fixed.

## *2. 4. Thermodynamic Potentials*

From (15), the energy is a 'natural function' of $S$, the displacements $X_\gamma$ and the particle numbers $\overline{N}_\alpha$. In principle, we should obtain the equilibrium state at fixed $S$, $X_\gamma$, $\overline{N}_\alpha$ by minimising the energy with respect to other variables with these natural variables held fixed. However, no practical instrument exists for the measurement and control of entropy experimentally! Instead what is more convenient is to keep the temperature fixed. Thus we would like to change to a thermodynamic potential with natural variable $T$ rather than $S$.

This is easily done by defining

$$F = E - TS \qquad \text{definition of Helmholtz free energy} \tag{16}$$

(In some texts the Helmholtz free energy is denoted $A$.) For clarity consider a $PVT$ system and use (13) then

$$\mathrm{d}F = \mathrm{d}(E - TS) = \mathrm{d}E - T\mathrm{d}S - S\mathrm{d}T = -SdT - PdV$$

$$\Rightarrow \quad T, V \text{ are natural variables for } F$$

$$\Rightarrow S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \qquad P = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \tag{17}$$

In the last expressions we have added a reminder that the above equations were derived at fixed $N$. The fact that $T, V$ are natural variables for $F$ means that to obtain the thermodynamic state we should minimize $F$ at fixed $T, V$ (more on that later).

**Mathematical Aside**: Actually what we have done is known as a Legendre Transform. More generally consider a function $f(\{x_i\})$ $i = 1 \ldots k$ then

$$\mathrm{d}f = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}\mathrm{d}x_i = \sum_{i=1}^{k} u_i x_i$$

Then we can let $g = f - \sum_{i=r+1}^{k} u_i x_i$ and

$$\mathrm{d}g = \sum_{i=1}^{r} u_i \mathrm{d}x_i - \sum_{i=r+1}^{k} x_i \mathrm{d}u_i$$

and $g$ is a 'natural function' of $x_1 \ldots x_r, u_{r+1} \ldots u_k$ and is a *Legendre Transform* of $f$.

The idea is that the function $g$ should contain the same information as $f$. To get a feeling for this consider a function of one variable $f(x)$ which is specified by the values over a set of points $x$. But we could also specify the function (up to some overall constant) by the values of the derivative $u(x) = \dfrac{\partial f}{\partial x}$ at the set of points i.e. the envelope of tangents to the function. Now these tangents are straight lines each with a slope $u$ and its intercept, $g$, say along the $y$ axis i.e. the point $f(x) = g + ux$. Now one can think of $g(u)$ as the intersect as a function of slope, then

$$g(u) = f(x) - ux \; . \tag{18}$$

$g(u)$ contains the same information as the original relation $f(x)$ and is the Legendre transformation. **End of aside**

We can continue this process to construct thermodynamic potentials which are natural functions of different sets of variables

$$\begin{array}{lllc}
(T,V,N) & F = \overline{E} - TS & & \text{Helmholtz Free Energy} \\
(S,P,N) & H = \overline{E} - (-PV) & = \overline{E} + PV & \text{Enthalpy} \\
(T,P,N) & G = F - (-PV) & = \overline{E} - TS + PV & \text{Gibbs Free Energy} \\
(T,V,\mu) & \Phi = F - N\mu & = \overline{E} - TS - N\mu & \text{Grand Potential}
\end{array}$$

All these thermodynamic potentials provide minimisation principles e.g. to find the equilibrium state at fixed $T, P, N$ we minimise the Gibbs free energy. Of particular significance in statistical mechanics are $F$, the Helmholtz Free Energy, and $\Phi$, the Grand Potential.

**Extremisation Principles:**

Recall that our original programme was to maximise $S$ subject to constraints. If the constraint is $\overline{E}$ (Canonical ensemble) then we ended up imposing

$$\frac{\partial}{\partial p_i}\left[S - \lambda_E \sum_j p_j E_j - \lambda_1 \sum_j p_j\right] = 0$$

which, if we think about it, is precisely the same as

1. minimising $F = \overline{E} - TS$ where $\lambda_E = 1/T$
2. minimising $\overline{E}$ subject to the constraint of constant $S$ (i.e. we would have a Lagrange multiplier $\lambda_S = 1/\lambda_E$)

Thus the different extremisation principles (maximisation of entropy, minimisation of free energy, minimisation of energy etc) are all intimately related and which one applies is determined by which ensemble one is working in, i.e. which constraints apply and which variables are held fixed.

**Gibbs-Duhem relation:** Going back to the energy we know that the energy should be *extensive* i.e. proportional to the size of the system. Similarly all of the natural variables of the energy are extensive. This implies that multiplying each extensive variable by a factor $b$ results in an increase of the energy by a factor $b$

$$bE(S,\{X_\gamma\},\{N_\alpha\}) = E(bS,\{bX_\gamma\},\{bN_\alpha\})$$

where $b$ is a positive parameter. Now we differentiate both sides with respect to $b$

$$\begin{aligned}\overline{E}(S,\{X_\gamma\},\{N_\alpha\}) &= \frac{\mathrm{d}}{\mathrm{d}b}E(bS,\{bX_\gamma\},\{bN_\alpha\}) \\ &= TS + \sum_\gamma f_\gamma X_\gamma + \sum_\alpha \mu_\alpha \overline{N}_\alpha\end{aligned}$$

where we have used (15). But then

$$\mathrm{d}\overline{E} = T\mathrm{d}S + S\mathrm{d}T + \sum_\gamma [f_\gamma \mathrm{d}X_\gamma + X_\gamma \mathrm{d}f_\gamma] + \sum_\alpha \left[\mu_\alpha \mathrm{d}\overline{N}_\alpha + \overline{N}_\alpha \mathrm{d}\mu_\alpha\right]$$

and subtracting (15) yields

$$0 = S\mathrm{d}T + \sum_\gamma X_\gamma \mathrm{d}f_\gamma + \sum_\alpha \overline{N}_\alpha \mathrm{d}\mu_\alpha \tag{19}$$

This is known as the Gibbs-Duhem relation and implies that the intensive variables $T, \{f_\gamma\}$ and $\{\mu_\alpha\}$ are not all independent. For example, in a $PVT$ system with one species, only two of $T, P, \mu$ are independent since (19) becomes

$$0 = S\mathrm{d}T - V\mathrm{d}P + \overline{N}\mathrm{d}\mu \tag{20}$$

i.e. a change in $T$, $P$ implies a specified change in $\mu$

*2. 5. Identification of Lagrange multipliers and the bridge equations*

**Canonical distribution**

Let us consider first the canonical ensemble. The idea is to work out from the form of the probability (11), an expression for $\mathrm{d}\overline{E}$ then compare with the first/second law of thermodynamics (14). The microscopic definition of the total energy of an assembly is

$$\overline{E} = \sum_i p_i E_i \tag{21}$$

But we know from quantum mechanics, that changing the volume of the 'box' for example must change the energy levels. It follows that a change $\mathrm{d}X_\gamma$ must give rise to a change in the mean energy. Therefore a change in the mean energy has two types of contribution, one from changes in $p_i$ and one from changes in $E_i$ due to $X_\gamma$:

$$\begin{aligned}\mathrm{d}\overline{E} &= \sum_i \frac{\partial \overline{E}}{\partial p_i}\,\mathrm{d}p_i + \sum_\gamma \frac{\partial \overline{E}}{\partial X_\gamma}\mathrm{d}X_\gamma \\ &= \sum_i E_i\,\mathrm{d}p_i + \sum_\gamma \frac{\partial \overline{E}}{\partial X_\gamma}\mathrm{d}X_\gamma\ . \end{aligned} \tag{22}$$

where the last line follows from (21). We now consider a small change in the entropy. Since $S$ is just a function of the probabilities $p_i$, a change in $S$ comes from a change in the probability distribution, and this may be expressed as

$$\begin{aligned} \mathrm{d}S &= \sum_i \frac{\partial S}{\partial p_i} \mathrm{d}p_i \\ &= -k \sum_i \left[\ln p_i + 1\right] \mathrm{d}p_i \\ &= -k \sum_i \ln p_i \, \mathrm{d}p_i \end{aligned} \tag{23}$$

where we have used the definition of the Gibbs entropy and the normalization condition in the form

$$\sum_i \mathrm{d}p_i = 0 \tag{24}$$

i.e. since the sum of all probabilities must always be 1 the sum of any changes in the $p_i$ must sum to zero. By substituting from (11) for $p_i$, and again using the condition $\sum_i \mathrm{d}p_i = 0$, we may further write our expression for the change in the entropy as

$$\mathrm{d}S = \lambda_E \sum_i E_i \mathrm{d}p_i,$$

Then with a little rearrangement (22) becomes

$$\mathrm{d}\overline{E} = \frac{\mathrm{d}S}{\lambda_E} + \sum_\gamma \frac{\partial \overline{E}}{\partial X_\gamma} \mathrm{d}X_\gamma \tag{25}$$

Comparison with the thermodynamic expression for the change in mean energy, as given by equation (14), then yields the Lagrange multiplier as

$$\lambda_E = 1/T, \tag{26}$$

along with an expression for the thermodynamic 'generalised forces'

$$f_\gamma = \frac{\partial \overline{E}}{\partial X_\gamma} \,. \tag{27}$$

For example the pressure is given by

$$-P = \frac{\partial \overline{E}}{\partial V} = \sum_i p_i \frac{\partial E_i}{\partial V} \,. \tag{28}$$

(The latter equation can be used to introduce the instantaneous pressure $P_i = -\dfrac{\partial E_i}{\partial V}$, such that the mean pressure takes the form

$$P = \sum_i p_i P_i,$$

as is used in Tutorial 2.7.)

Having identified the Lagrange multiplier we can now write the Canonical distribution in the familiar form,

$$\boxed{p_i = \frac{1}{Z_c}\exp(-\beta E_i) \qquad \text{where} \quad Z_c = \sum_i \exp(-\beta E_i)} \tag{29}$$

where we have used the convenient abbreviation

$$\boxed{\beta = \frac{1}{kT}\ .} \tag{30}$$

Substituting the above form for $p_i$ into equation (4) we obtain an expression for the entropy in terms of the partition function and the mean energy, thus:

$$S = -k\sum_i p_i\left[-\beta E_i - \ln Z_c\right] = \frac{\overline{E}}{T} + k\ln Z_c \tag{31}$$

Or, recalling the Helmholtz free energy $F = \overline{E} - TS$, we may rewrite the above equation as

$$\boxed{F = -kT\ln Z_c} \tag{32}$$

This remarkable result is often referred to as a 'bridge equation', as it provides a bridge between the microscopic and macroscopic descriptions of an assembly. Basically we get the helmholtz free energy, a potential we introduced in section  on macroscopic grounds, from the canonical partition function which comes from summing over microscopic states. The basic procedure of statistical physics is essentially to obtain an expression for the partition function from purely microscopic considerations, and then to use the bridge equation to obtain the thermodynamic free energy. Thus we come from the microscopic description to macroscopic properties.

### Grand Canonical distribution

We follow the same procedure used for the canonical distribution and write

$$\mathrm{d}\overline{E} = \sum_{i,N} E_{i,N}\mathrm{d}p_{i,N} + \sum_\gamma \frac{\partial\overline{E}}{\partial X_\gamma}\,\mathrm{d}X_\gamma\ . \tag{33}$$

One subtle point here is that changes in $\overline{N}$ lead to changes in $p_{i,N}$ but not to any changes in $E_{i,N}$ since these are calculated for the fixed allowed values of $N$. The change in entropy may be expressed as

$$\mathrm{d}S = \sum_{i,N}\frac{\partial S}{\partial p_{i,N}}\,\mathrm{d}p_{i,N} = -k\sum_{i,N}\ln p_{i,N}\,\mathrm{d}p_{i,N}, \tag{34}$$

where we have used the definition of the Gibbs entropy and the normalization condition in the form $\sum_{i,N}\mathrm{d}p_{i,N} = 0$. By substituting from (12) for $p_{i,N}$, and again using the condition $\sum_{i,N}\mathrm{d}p_{i,N} = 0$, we may further write our expression for the change in the entropy as

$$\mathrm{d}S = \sum_{i,N}\left[\lambda_E E_i + \lambda_N N\right]\,\mathrm{d}p_{i,N}\ . \tag{35}$$

Finally we consider

$$\mathrm{d}\overline{N} = \sum_{i,N} N\,\mathrm{d}p_{i,N}, \tag{36}$$

and equation (34) becomes

$$\mathrm{d}S = \lambda_E \sum_{i,N} E_{i,N}\,\mathrm{d}p_{i,N} + \lambda_N \mathrm{d}\overline{N}\ .$$

Then with a little rearrangement (33) becomes

$$\mathrm{d}\overline{E} = \frac{\mathrm{d}S}{\lambda_E} - \frac{\lambda_N}{\lambda_E}\mathrm{d}\overline{N} + \sum_{\gamma} \frac{\partial \overline{E}}{\partial X_\gamma}\mathrm{d}X_\gamma$$

We compare this to the version of the first/second law (15) for one species

$$\mathrm{d}\overline{E} = T\mathrm{d}S + \mu\mathrm{d}\overline{N} + \sum_{\gamma} f_\gamma \mathrm{d}X_\gamma$$

and identify

$$\lambda_E = \frac{1}{T} \qquad \lambda_N = -\lambda_E\mu = -\frac{\mu}{T} \tag{37}$$

Having identified the Lagrange multipliers we can now write the Grand Canonical distribution in the familiar form,

$$\boxed{p_{i,N} = \frac{1}{\mathcal{Z}_{gc}} \exp(-\beta(E_i - N\mu)) \quad \text{where} \quad \mathcal{Z}_{gc} = \sum_{i,N} \exp(-\beta(E_i - N\mu))} \tag{38}$$

where again we have used $\beta = 1/kT$

Substituting the above form for $p_{i,N}$ into equation (4) we obtain an expression for the entropy in terms of the partition function, the mean energy and mean particle number, thus:

$$S = -k\sum_{i,N} p_{i,N}\left[-\beta E_i + \beta\mu N - \ln \mathcal{Z}_{gc}\right] = \frac{\overline{E}}{T} - \frac{\mu\overline{N}}{T} + k\ln \mathcal{Z}_{gc} \tag{39}$$

Or, recalling the Grand potential $\Phi = \overline{E} - TS - \mu\overline{N}$, we may rewrite the above equation as

$$\boxed{\Phi = -kT\ln \mathcal{Z}_{gc}} \tag{40}$$

Again we have a remarkable 'bridge equation' connecting the microscopic and macroscopic descriptions of an open assembly.

### *2. 6. Ensembles*

In deriving our distributions we used the a priori definition of probability: we used the available information then constructed rational probabilities representing our degree of belief that the assembly would be in microstate $i$. In order to use the frequency definition of probability we can introduce the concept of an ensemble.

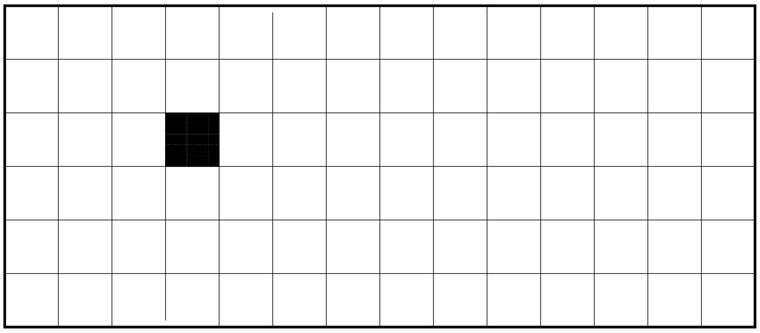

Figure 1: The idea of an ensemble: the 'megasystem' is divided up into a large number $M$ of assemblies e.g. the black square is one assembly

An ensemble is a very large number of copies, $M$, of the same assembly i.e. an assembly of assemblies if you like. Then the probability of microstate $i$ is given by the number of assemblies in microstate $i$ divided by $M$

$$p_i = \lim_{M\to\infty} \frac{m_i}{M}$$

Practically one can think of a block of material: each assembly being a small piece of the whole block. Even though each of the many pieces is small, they still contain Avogadro size numbers of atoms so we can have both the number $M$ of assemblies large and the number $N$ of particles in each assembly large

The megasystem is isolated from the rest of the universe so all microstates of the megasystem are equally likely.

In tutorial question 1.4. it is shown how using the Boltzmann definition of entropy for the whole ensemble yields the Gibbs definition of entropy for each assembly.

One constructs the Canonical ensemble *for the assemblies* by allowing them to exchange energy with each other. Therefore for each assembly the rest of the assemblies act as a heat reservoir.

Similarly one construct the Grand canonical ensemble for by allowing the assemblies to exchange energy and particles with one another. Thus for each assembly the others act as a heat and particle reservoir. This lends itself to naturally to considering a small portion of a fluid, say, where the particle number will clearly fluctuate due to particles moving in from neighbouring assemblies.

Actually, the ensemble construction we have just described defines the probability of a state of an assembly using the frequency definition and is the old-fashioned route. In our assignments of probabilities we have actually bypassed the need for this construction by using an a priori information definition of probability to obtain the probability distribution for an assembly. However the terminology 'ensemble' endures and we will continue to uses it as a synonym for distribution. The idea of a large number of copies of an assembly will prove useful later on when we consider dynamics.

## 2. 7. Summary

Let us summarise our achievements in this section.

- By maximising the missing information, or Gibbs entropy, subject to various constraints we have derived the three standard 'ensembles' (or distributions) of Statistical mechanics: Microcanonical, Canonical and Grand Canonical
- The Lagrange multipliers that enter into the maximisation procedure have been identified with intensive thermodynamics variables: temperature and chemical potential.
- In the tutorial you are invited to show how two systems in equilibrium under exchange of a certain quantity should share the same value of the Lagrange multiplier corresponding to that quantity. For example systems in equilibrium with respect to energy exchange should share the same value of $\lambda_E$ i.e. the same temperature. This recovers the $0^{\text{th}}$ Law of thermodynamics. Similarly systems in equilibrium with respect to particle exchange should share the same value of $\lambda_N$ i.e. the same chemical potential.

  Thus, remarkably, the abstract mathematical entities, Lagrange multipliers, become the concrete physical concepts enshrined in thermodynamics, such as temperature.
- We have arrived at bridge equations which allow us to express macroscopic potentials (e.g. free energies) in terms of the logarithms of the partition functions.

  Again remarkably we link abstract mathematical entities (partition functions) with established physical quantities (potentials and free energies).