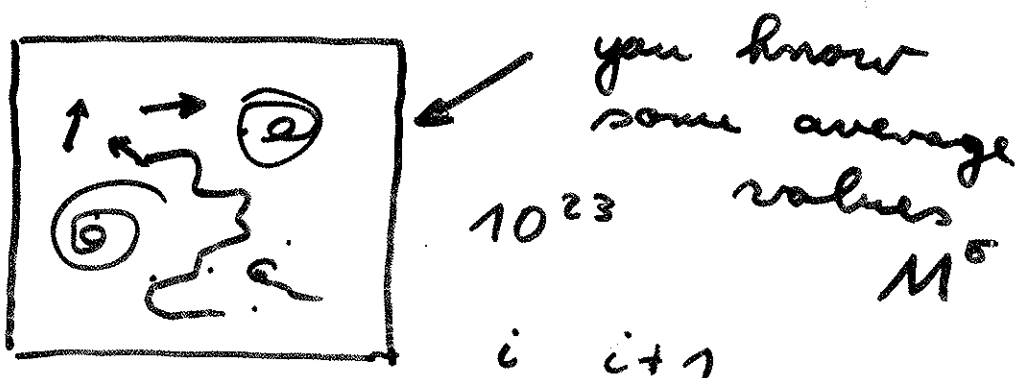


Statistical Mechanics by L. Bick



General question:

What type of probabilities of microstates i should we assume if we only know the quantities M^5 and nothing else?

Principle idea of SM

Define entropy function (= negative information measure).

Maximize entropy subject to given constraints.

Those probabilities that maximize the entropy are the relevant ones that correctly describe the system under consideration.

Information measures -2-

Consider events labelled by $i \in N$.
Probabilities for these events to happen are p_i .

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

Suppose event j occurs.

Suppose $p_j \approx 1$, whereas

$p_i \approx 0$ ($i \neq j$).

$\Rightarrow$ we learn very little from the actual occurrence of the event.

But if $p_i \approx \frac{1}{N}$ ($\forall i$) then we learn a lot from the actual occurrence of event j .

Information gain due to occurrence of an event is given by some function $h(p_j)$.

For example, choose $h(p_j) = \log p_j$.

In a long sequence of independent

trials, the average information gain is given by

$$I(\{p_i\}) = \sum_{i=1}^w p_i h(p_i)$$

The entropy S is then defined as

$$S = -I$$

The Shannon axioms

These general axioms describe what properties an information measure should have.

Axiom 1

$$I = I(p_1, p_2, \dots, p_w)$$

Information should only depend on frequencies of events and nothing else.

Axiom 2

$$I\left(\frac{1}{w}, \dots, \frac{1}{w}\right) \leq I(p_1, \dots, p_w)$$

-4-

The information measure I takes on a minimum for uniform distribution

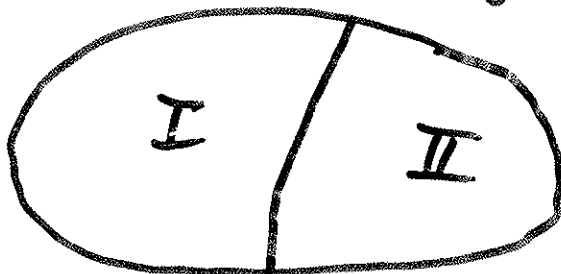
Axiom 3

$$I(P_1, \dots, P_N) = I(P_1, \dots, P_N, 0)$$

This means the information measure should not change if the sample set is enlarged by another event that has prob. 0.

Axiom 4

$$I(\{P_{ij}^{I, II}\}) = I(\{P_i^I\}) + \sum_i P_i^I I(\{P^{II}(j|i)\})$$



Meaning: The way the information is collected should be independent of how we divide the entire system into subsystems.

In general system I and II are not statistically independent.

But, in case they are, then

$$P_{ij}^{I,II} = P_i^I \cdot P_j^{II}$$

and thus

$$I(\{P_{ij}^{I,II}\}) = I(\{P_i^I\}) + I(\{P_j^{II}\})$$

since $P^{II}(j|i) = P_j^{II}$

Important result

The Shannon entropy

$$S := -K \sum_{i=1}^W P_i \ln P_i = -I$$

is the only solution to the Shannon axioms 1-4.

Proof

in one of the next lectures.

Often $K=1$

Other information measures

Rényi entropy

$$S_q^{(R)} = \frac{1}{q-1} \ln \sum_i p_i^q$$

$q \in \mathbb{R}$ a parameter

The summation is over all events with $p_i \neq 0$.

This satisfies axioms 1-3 but not 4.

For $q \rightarrow 1$ $S_q^{(R)}$ reduces to the Shannon entropy.

Proof

$$q = 1 + \epsilon$$

ϵ small

$$S_{1+\epsilon}^{(R)} = \frac{1}{\epsilon} \log \sum_i p_i^{1+\epsilon}$$

$$\log \equiv \ln$$

$$= \frac{1}{\epsilon} \log \sum_i p_i \cdot \underbrace{p_i^\epsilon}_{e^{\epsilon \log(p_i)}}$$

$$\underbrace{e^{\epsilon \log(p_i)}}_{e^{\epsilon \log p_i} \approx 1 + \epsilon \log p_i}$$

$$e^{\epsilon \log p_i} \approx 1 + \epsilon \log p_i$$

$$\begin{aligned}
&= \frac{1}{\epsilon} \log \left(\sum_i P_i (1 + \epsilon \log P_i) + \dots \right) \\
&= \frac{1}{\epsilon} \log \left(1 + \epsilon \sum_i P_i \log P_i + \dots \right) \\
&= \frac{1}{\epsilon} \left(\epsilon \sum_i P_i \log P_i + \dots \right) \\
&= \sum_i P_i \log P_i
\end{aligned}$$

if $\epsilon \rightarrow 0$

■

The Tsallis entropies

$$S_q^{(T)} = \frac{1}{q-1} \left(1 - \sum_{i=1}^W P_i^q \right)$$

Exercise 1

Show that for $q \rightarrow 1$ the Tsallis entropy reduces to the Shannon entropy

Again these types of entropies satisfy versions 1-3 but not version 4.

Further entropies ⁻⁸⁻

Landsberg - Pedral entropy

$$S_q^{(L)} = \frac{1}{q-1} \left(\frac{1}{\sum_{i=1}^n p_i^q} - 1 \right)$$

Abbe entropy

$$S_q^{Abbe} = - \sum_i \frac{p_i^q - p_i^{q^{-1}}}{q - q^{-1}}$$

Jaynes' information theory gives a simple information-theoretic derivation of stat. mech.

Assume we know some mean values $M^{\bar{\sigma}}$, $\bar{\sigma} = 1, \dots, S$ of a complex system under consideration.

Suppose we measure information with some information measure

$$I(\{p_i\}) = T \ln T$$

What probabilities p_i for the microstates i of the system should we assume?

Our information should have a minimum given the constraints of system. Because otherwise we would have unjustified prejudices on the system, which we should have entered in terms of additional constraints M^0 in the first place.

$\Rightarrow I[p]$ must have minimum
 $\Rightarrow S[p] = -I[p]$ must have a maximum (under the given constraints M^0).

Constraints are given in form of the following conditions:

$$\sum_{i=1}^W p_i M_i^\sigma = M^\sigma \quad (\sigma = 1, \dots, S)$$

Example: mean energy of a system particles:

$$\sum_{i=1}^W p_i E_i = U$$

$\uparrow$ energy associated with microstate i $\uparrow$ mean energy

Normalization condition is also regarded as a constraint

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

To find the distributions that maximize the entropy under the given constraints we introduce a function $\Psi[p]$

$$\Psi[p] := I[p] + \sum_{\sigma} \beta_{\sigma} \left(\sum_i p_i M_i^{\sigma} \right)$$

$\uparrow$ Lagrange multiplier

Went to find minimum of
 $\Psi[p] = \min \text{ of } I[p] \text{ given the constraints}$

Minimum can be obtained by differentiating:

$$\frac{\partial}{\partial p_i} \Psi[p] = 0 \quad (i=1, \dots, W)$$

This means

$$\frac{\partial}{\partial p_i} I[p] + \sum_c \beta_c M_i^c = 0$$

Ordinary stat. mech.

$$I[p] = \sum_i p_i \log p_i$$

$$\frac{\partial}{\partial p_i} I[p] = \log p_i + 1$$

$$\log p_i + 1 + \alpha + \beta E_i = 0$$

(*)

reminder: for the canonical ensemble one has

$$M_i = E_i$$

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

energy of i-th microstate

$$\beta_1 \equiv \alpha$$

$$\beta_2 \equiv \beta$$

$$\sigma = 1, 2$$

meaning

$$\psi[p] = \sum p_i \log p_i + \alpha \sum p_i + \beta \sum p_i E_i$$

Solving (*) for p_i gives

$$p_i = \frac{1}{Z} e^{-\beta E_i}$$

Z : partition function

$$Z = \sum_i e^{-\beta E_i} = e^{1+\alpha}$$

Definition

The free energy F of any complex system with partition function $Z(\beta)$ is defined as

$$F := -\frac{1}{\beta} \log Z(\beta)$$

Theorem

For the canonical ensemble one has