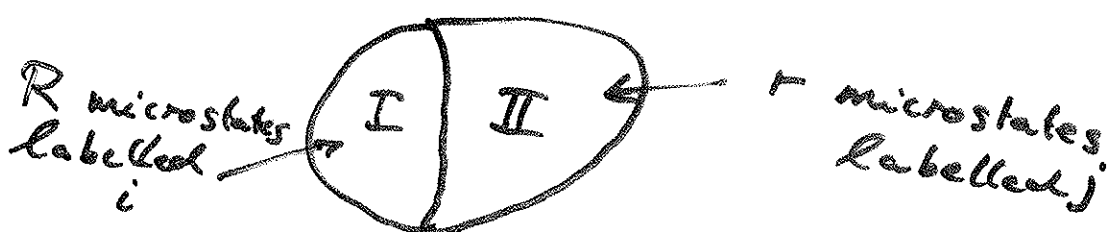


Proof that Shannon information

$$I[p] = c \sum_{i=1}^W p_i \ln p_i$$

is the unique solution of the  
Shannon axioms

( $c > 0$  arbitrary constant)



Consider (first) a special example,  
the uniform distribution

$$I, II$$

$$P_{i,j} = \frac{1}{N}$$

$$I, II$$

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

↑  
joint prob.  
↑  
conditional prob.  
of event  $j$  given  
that  $i$  has happened

$$P_i^I = \frac{1}{R}$$

$$P^{II}(j|i) = \frac{1}{r}$$

$$R \cdot r = N$$

Assertion (1) : Information measure  $I$  is a function of the number of states only

$$\Rightarrow I(\{p_{i,j}^{I,II}\}) = f(N)$$

$$I(\{p_i^I\}) = f(R)$$

$$I(\{p^{II}(j|i)\}) = f(r)$$

Assertion (4) yields

$$f(Rr) = f(R) + f(r)$$

$$\text{since } \sum p_i^I = 1$$

$$\Rightarrow f(R) = -c \ln R$$

where  $c$  is some constant

Assertion (2) and (3) yield

$$I\left(\frac{1}{R}, \dots, \frac{1}{R}\right) = I\left(\frac{1}{R}, \dots, \frac{1}{R}, 0\right)$$

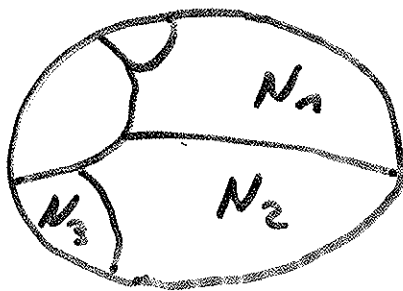
$$\geq I\left(\frac{1}{R+1}, \dots, \frac{1}{R+1}\right)$$

$$f(R) \geq f(R+1) \quad -??-$$

i.e.  $f$  is monotonically decreasing function of  $R$

Thus  $c > 0$

Now produce non-uniform distributions  $P_i^I$  in the following way:



Divide sample set containing  $N$  uniformly distributed events into  $R$  subsets containing  $N_i$  elements

$$P_i^I = \frac{N_i}{N}$$

$P(j|i)$  is still uniform  
 $\uparrow$   
 prob. to observe event  $j$  given that we are in subset  $i$

Axiom (4)

- 23 -

$$f(N) = I(\{p_i^I\}) + \sum_i p_i^I f(N_i)$$

$$\Leftrightarrow -c \ln N = I(\{p_i^I\}) - c \sum_i p_i^I \ln N_i$$

$$\Leftrightarrow +c \sum_i p_i^I \ln N_i - c \ln N \underbrace{\sum_i p_i^I}_1 = I(\{p_i^I\})$$

$\Leftrightarrow$

$$c \sum_i p_i^I \ln \underbrace{\frac{N_i}{N}}_{p_i^I} = I(\{p_i^I\})$$

pos. constant  $c$  can be fixed  
by definition. q.e.d.

Often

$$c = 1$$

$$c = k$$

Boltzmann const.

-64-

Remark : Axiom 4 can be written as :

$$S^{I, II} = S^I + \underset{\substack{\uparrow \\ \text{conditional entropy}}}{S^{II/I}}$$

If you generalize Axiom 4 as follows

$$S_q^{I, II} = S^I + S^{II/I} - (q-1) S_q^I \cdot S_q^{II/I}$$

then you end up uniquely with the Tsallis entropies.

We now assume validity of the (normal-type) Ghinchim-Axioms, we are thus lead to ordinary stat. mech. based on the Shannon entropy

$$S = -k \sum p_i \ln p_i$$

with  $p_i = \frac{1}{Z} e^{-\beta E_i}$  (canonical ensemble)

$$Z(\beta) = \sum_i e^{-\beta E_i} \quad \text{partition function}$$

$$\beta = \frac{1}{kT}$$

We already showed that if we define  $F := -\frac{1}{\beta} \log Z(\beta)$  then

$$F = U - TS \quad (*)$$

$\uparrow$  mean energy       $\uparrow$  entropy

Useful in the following:

Define  $k=1$

and

$$\begin{aligned} \psi(\beta) &:= -\log Z(\beta) \\ &= \beta F(\beta) \\ &= -\log \sum_i e^{-\beta E_i} \end{aligned}$$

(\*)

$$\Rightarrow \psi = -S + \beta U$$

Lemma 1       $\frac{\partial \psi}{\partial \beta} = U$

Proof

$$\Psi(\beta) = -\log \sum_i e^{-\beta E_i}$$

$$\frac{\partial}{\partial \beta} \Psi = \frac{+ \sum_i E_i e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}$$

$$= \sum_i \underbrace{\frac{1}{Z} e^{-\beta E_i}}_{p_i} \cdot E_i$$

$$= U \quad \text{q.e.d.}$$

Lemma 2

$$\frac{\partial S}{\partial U} = \beta$$

(originally  $S = S(\{p_i\})$  was a function of all prob.  $p_i$ , but now (physically) we regard it as a function of just one variable,  $U = \frac{\partial \Psi}{\partial \beta}$  )

Proof

$$S = -\Psi(\beta) + \beta U(\beta)$$

$$\frac{\partial S}{\partial U} = \frac{\partial S}{\partial \beta} \frac{\partial \beta}{\partial U}$$

$$\begin{aligned}\frac{\partial S}{\partial \beta} &= - \underbrace{\frac{\partial}{\partial \beta} \psi(\beta)}_{U(\beta)} + U(\beta) + \beta \frac{\partial U}{\partial \beta} \\ &= \beta \frac{\partial U}{\partial \beta}\end{aligned}$$

$$\Rightarrow \frac{\partial S}{\partial U} = \beta \frac{\partial U}{\partial \beta} \cdot \frac{\partial \beta}{\partial U} = \beta \quad \text{q.e.d.}$$

In thermodynamics / statistical mechanics, there are usually many parameters  $\beta_\alpha$  (corresponding e.g. to inverse temperature, pressure, chemical potential of the system under consideration).

For all those parameters relations similar to Lemma 1 and 2 are valid.

Lemma 1 and 2 imply that  $\Psi$  and  $\Sigma$  are connected by a so-called Legendre transformation.

Consider generally a function  $f(\beta)$  and let

$$M := \frac{\partial f}{\partial \beta}$$

Define Legendre transform of  $f$  by

$$L(M) = -f(\beta) + \beta M$$

$$\begin{aligned} \frac{dL}{dM} &= - \underbrace{\frac{df}{d\beta}}_M \frac{d\beta}{dM} + M \frac{d\beta}{dM} + \beta \\ &= \beta \end{aligned}$$

In general

$$M^\sigma = \frac{\partial f}{\partial \beta_\sigma} \quad f = f(\beta_1, \beta_2, \dots)$$

$$L(M) + F(\beta) = \sum_\sigma \beta_\sigma M^\sigma$$

$$\frac{\partial L}{\partial M^\sigma} = \beta_\sigma$$

Need also to generalize the concept of Legendre transformation to functions that are not necessarily differentiable.

$$f(\beta) = \min_M (\beta M - L(M))$$

$$L(M) = \min_{\beta} (\beta M - f(\beta))$$

Let us denote by  $M(\beta)$  the particular value  $M$  where the minimum in the 1st eq. is attained.

Similarly, denote by  $\beta(M)$  the particular value  $\beta$  where min in 2nd eq. is attained.

$$f(\beta) = \beta M(\beta) - L(M(\beta))$$

$$L(M) = \beta(M) \cdot M - f(\beta(M))$$

If  $f$  and  $L$  are differentiable, then these minima are determined by

$$\frac{\partial}{\partial \beta} f(\beta) = M(\beta) + \beta \frac{\partial M}{\partial \beta} - \frac{\partial L}{\partial M} \frac{\partial M}{\partial \beta} = 0$$

$$\frac{d}{dM} L(M) = \beta(M) + M \frac{\partial \beta}{\partial M} - \frac{df}{d\beta} \frac{d\beta}{dM} = 0 \quad -30-$$

solved by

$$\frac{dL}{dM} = \beta$$

$$\frac{df}{d\beta} = M$$

$$\begin{aligned} dS &= \frac{\partial S}{\partial U} dU \\ &= \beta dU \end{aligned}$$

$$\Rightarrow dU = T dS$$

Generally, if you also allow for volume changes or particle number changes, you have

$$dU = T dS - p dV + \mu dN$$

$\uparrow$   
 pressure

$\uparrow$   
 volume

$\uparrow$   
 chem.  
potential

$\uparrow$   
 particle  
number

$\uparrow$   
Gibbs fundamental eq.

measured macroscopic quantities  
can be mean energy  $U$ , mean  
volume  $V$ , mean particle number  $N$

### canonical ensemble

$$P_i \sim e^{-\beta E_i}$$

$$\langle E_i \rangle = U$$

$$S = \beta(U - F)$$

↑  
Helmholtz free energy

### pressure ensemble

$$P_i \sim e^{-\beta E_i - \bar{p} V_i}$$

$$\langle V_i \rangle = V$$

$$\langle E_i \rangle = U$$

$$S = \beta(U + \bar{p}V - G)$$

↑  
Gibbs free energy

### grand canonical ensemble

$$P_i \sim e^{-\beta E_i - \bar{p} V_i + \mu N_i}$$

$$\langle N_i \rangle = N$$

$$\langle V_i \rangle = V$$

$$\langle E_i \rangle = U$$

$$S = \beta(U + \bar{p}V - \mu N - \Omega)$$

↑  
grand canonical free energy