Corrigé des exercices sur la physique statistique

1 Entropie de Shannon

On considère des expériences aléatoires avec n résultats $a_1, \cdots, a_n$. On cherche à associer à chaque n et à chaque mesure de probabilité $(p_1, \cdots, p_n)$ sur $\{a_1, \cdots, a_n\}$ une quantité qui mesure l'*incertitude sur le résultat d'une expérience en connaissant juste les probabilités* $p_1, \cdots, p_n$ de chaque événement. On notera, pour tout n :

$$\begin{array}{cccc} H_n : & \mathcal{M}[n] & \longrightarrow & \mathbb{R} \\ & (p_1, \cdots, p_n) & \longmapsto & H(p_1, \cdots, p_n) \end{array}$$

la fonction qui, à chaque $p_1, \cdots, p_n$, associe cette quantité, où $\mathcal{M}[n]$ est l'ensemble des probabilités sur $\{a_1, \cdots, a_n\}$. Suivant Claude Shannon, on postule les axiomes suivants sur la famille de fonctions $(H_n)_{n\in\mathbb{N}}$.
(A1) Dans le cas *équiprobable* où $p_1 = \cdots = p_n = 1/n$, on note

$$f(n) := H_n\left(\frac{1}{n}, \cdots, \frac{1}{n}\right)$$

et on définit ainsi une fonction $f : \mathbb{N}^* \longrightarrow \mathbb{R}$. On fait alors l'hypothèse que **la fonction f est croissante** (plus on a de choix, plus l'incertitude est grande).

(A2) Si on réalise deux expériences *indépendantes* : la première avec n_1 résultats équiprobables et la deuxième avec n_2 résultats équiprobables, alors **l'incertitude associée à l'expérience qui associe les deux** (et qui comporte alors $n_1 n_2$ résultats équiprobables) **est la somme des incertitudes** :

$$f(n_1 n_2) = f(n_1) + f(n_2).$$

(A3) Pour une expérience aboutissant à $n = n_1 + n_2$ résultats, si on regroupe les n_1 premiers résultats dans $A = \{a_1, \cdots, a_{n_1}\}$ et les n_2 suivants dans $B = \{a_{n_1+1}, \cdots, a_n\}$, alors **l'incertitude totale $H_n(p_1, \cdots, p_n)$ est la somme de l'incertitude de savoir si le résultat sera dans A ou dans B** (c'est à dire $H_2(p_A, p_B)$, où $p_A = p_1 + \cdots + p_{n_1}$ et $p_B = p_{n_1+1} + \cdots + p_n$) **et de la moyenne de l'incertitude sur le résultat dans $\{a_1, \cdots, a_{n_1}\}$ sachant que le résultat y est et de l'incertitude sur le résultat dans $\{a_{n_1+1}, \cdots, a_n\}$ sachant que le résultat y est**. Autrement dit :

$$H_n(p_1, \cdots, p_n) = H_2(p_A, p_B) + p_A H_{n_1}\left(\frac{p_1}{p_A}, \cdots, \frac{p_{n_1}}{p_A}\right) + p_B H_{n_2}\left(\frac{p_{n_1+1}}{p_B}, \cdots, \frac{p_n}{p_B}\right).$$

(A4) Enfin la fonction $p \longmapsto H_2(p, 1-p)$ est **continue**.

Le but de cet exercice est de montrer que, si les axiomes (A1), (A2), (A3) et (A4) sont satisfaits, alors il existe une constante C telle que $\forall n \in \mathbb{N}^*$, $\forall (p_1, \cdots, p_n) \in \mathcal{M}[n]$,

$$H_n(p_1, \cdots, p_n) = -C \sum_{i=1}^{n} p_i \log p_i. \tag{1}$$

1) Montrer que, pour tout $n \in \mathbb{N}^*$ et pour tout $s \in \mathbb{N}$, on a

$$kf(n) \leq sf(2) \leq (k+1)f(n),$$

où k est l'unique entier tel que $n^k \leq 2^s < n^{k+1}$. En déduire que

$$f(n) = \frac{f(2)}{\log 2} \log n.$$

Réponse — Soit $n \in \mathbb{N}^*$ et $s \in \mathbb{N}^*$. Considérons $k := \sup\{k \in \mathbb{N} |\ n^k \leq 2^s\}$. Alors $n^k \leq 2^s$ et $2^s < n^{k+1}$. D'après (A2) on a : $f(2^s) = sf(2)$, $f(n^k) = kf(n)$ et $f(n^{k+1}) = (k+1)f(n)$. Donc, comme f est croissante (A1),

$$n^k \leq 2^s < n^{k+1} \quad \Longrightarrow \quad kf(n) \leq sf(2) < (k+1)f(n) \quad \Longleftrightarrow \quad \frac{k}{s} \leq \frac{f(2)}{f(n)} < \frac{k+1}{s}.$$

Mais comme $n^k \leq 2^s < n^{k+1}$ implique pour des raisons analogues $\frac{k}{s} \leq \frac{\log 2}{\log n} < \frac{k+1}{s}$, on en déduit que

$$\left| \frac{f(2)}{f(n)} - \frac{\log 2}{\log n} \right| < \frac{1}{s}.$$

Comme comme cela est valable pour $s \in \mathbb{N}$, on en déduit le résultat demandé. Dans la suite nous poserons $C := \frac{f(2)}{\log 2}$, si bien que

$$f(n) = C \log n.$$

2) Montrer que, pour $p \in [0,1]$, on a

$$H_2(p, 1-p) = -C\left[p \log p + (1-p)\log(1-p)\right].$$

On pourra pour cela considérer d'abord les cas où $p = r/s \in \mathbb{Q}$.
Réponse — Supposons d'abord $p = r/s$, avec $r \in \mathbb{N}$ et $s \in \mathbb{N}^*$, tels que $r \leq s$. Considérons une expérience avec s résultats équiprobables et divisons les s résultats $\{a_1, \cdots, a_s\}$ en deux paquets $A := \{a_1, \cdots, a_r\}$ et $B := \{a_{r+1}, \cdots, a_s\}$. Nous appliquons alors (A3) :

$$H_s\left(\frac{1}{s}, \cdots, \frac{1}{s}\right) = H_2\left(\frac{r}{s}, \frac{s-r}{s}\right) + \frac{r}{s} H_r\left(\frac{1}{r}, \cdots, \frac{1}{r}\right) + \frac{s-r}{s} H_{s-r}\left(\frac{1}{s-r}, \cdots, \frac{1}{s-r}\right)$$

qui nous donne :

$$f(s) = H_2\left(\frac{r}{s}, \frac{s-r}{s}\right) + \frac{r}{s} f(r) + \frac{s-r}{s} f(s-r)$$

ou encore :

$$H_2\left(\frac{r}{s}, \frac{s-r}{s}\right) = f(s) - \frac{r}{s} f(r) - \frac{s-r}{s} f(s-r) = \frac{r}{s}(f(s) - f(r)) + \frac{s-r}{s}(f(s) - f(s-r)).$$

En tenant compte du résultat de la question précédente, on obtient :

$$H_2\left(\frac{r}{s}, \frac{s-r}{s}\right) = -C\frac{r}{s} \log \frac{r}{s} - C\frac{s-r}{s} \log \frac{s-r}{s},$$

ce qui donne la relation demandée pour $p = r/s$. On déduit de la condition (A4) que cela est vrai pour tout $p \in [0,1]$.
3) Conclure en montrant (1) par récurrence.
Réponse — Nous venons de démontrer (1) pour $n = 2$. Supposons que nous ayons montré (1) pour H_{n-1} et montrons que cette relation pour H_n. Nous divisons les résultats $\{a_1, \cdots, a_n\}$ d'une expérience en deux paquets $A = \{a_1\}$ et $B = \{a_2, \cdots, a_n\}$. En appliquant (A3), on obtient :

$$H_n(p_1, \cdots, p_n) = H_2(p_1, 1-p_1) + p_1 H_1(1) + (1-p_1) H_{n-1}\left(\frac{p_2}{1-p_1}, \cdots, \frac{p_n}{1-p_1}\right).$$

En utilisant $H_1(1) = f(1) = C$ et les hypothèses de récurrence $H_2(p_1, 1-p_1) = Cp_1 \log p_1 + C(1-p_1)\log(1-p_1)$ et

$$H_{n-1}\left(\frac{p_2}{1-p_1}, \cdots, \frac{p_n}{1-p_1}\right) = C\frac{p_2}{1-p_1} \log \frac{p_2}{1-p_1} + \cdots C\frac{p_n}{1-p_1} \log \frac{p_n}{1-p_1},$$

on en déduit le résultat.

2 Equirépartition de l'énergie

On considère un système de N oscillateurs harmoniques indépendants : l'état à un certain instant t de ce système est caractérisé par un vecteur $z(t) = (z_1(t), \cdots, z_N(t)) \in \mathbb{R}^N$ dont l'énergie est $H(z(t))$ où :

$$H(z) = \sum_{i=1}^{N} a_i \frac{z_i^2}{2}$$

et où les constantes a_i sont toutes strictement positives. On munit $\mathbb{R}^N$ de la mesure de probabilité

$$d\nu = \frac{1}{Z(\beta)} e^{-\beta H(z)} d\mu(z),$$

où $d\mu(z) = dz_1 \cdots dz_N$ et $Z(\beta)$ est tel que $\int_{\mathbb{R}^N} d\nu = 1$. Calculer, pour tout $i = 1, \cdots, N$, la valeur moyenne de l'énergie correspondant au degré de liberté numéro i, c'est à dire $\langle a_i z_i^2/2 \rangle_\nu$ et montrer que cette quantité ne dépend pas de a_i.

Réponse —Voir le cours « Introduction à la physique statistique et à la mécanique quantique » (sur la page web), pages 97 à 99.