Analyse spectrale du modèle des urnes d'Ehrenfest

Dans tout le devoir, les coordonnées d'un N-uplet sont numérotées avec des exposants : $x = (x^{(1)}, x^{(2)}, \ldots, x^{(N)})$. Ceci permet de considérer sans ambiguïté une suite de N-uplets $(x_n)_{n\geq 0}$: alors, la i-ième coordonnée du n-terme de la suite est notée $x_n^{(i)}$.

L'objectif de ce devoir est de comparer la vitesse du convergence du modèle des urnes d'Ehrenfest à celle de la marche aléatoire sur l'hypercube, qui a été étudiée en cours. Dans tout ce qui suit, $N \geq 1$ est un entier fixé, et on considère les deux espaces d'états suivants :

— l'*espace microscopique* $\mathfrak{X}_{\mathrm{micro}} = \{+1, -1\}^N$, dont les états sont les N-uplets de signes $x = (x^{(1)}, \ldots, x^{(N)})$ avec chaque $x_i \in \{+1, -1\}$.

— l'*espace macroscopique* $\mathfrak{X}_{\mathrm{macro}} = [\![0, N]\!]$, dont les états sont les entiers k compris entre 0 et N.

Supposons données N particules numérotées $1, 2, \ldots, N$ et placées dans une boîte B, que l'on sépare en deux parties B_+ et B_-. Alors, l'appartenance de chaque particule à l'une des deux parties peut être encodée par un élément de $\mathfrak{X}_{\mathrm{micro}}$: par exemple, si $N = 5$, alors l'état $(+1, -1, +1, +1, -1) \in \mathfrak{X}_{\mathrm{micro}}$ indique que les particules 1, 3 et 4 sont dans B_+ et que les particules 2 et 5 sont dans B_-. On récupère à partir de $x = (x^{(1)}, x^{(2)}, \ldots, x^{(N)}) \in \mathfrak{X}_{\mathrm{micro}}$ le nombre de particules dans B_+ en considérant

$$k = S(x) = \frac{N}{2} + \frac{1}{2} \sum_{i=1}^{N} x^{(i)}.$$

Ce nombre de particules est un état de l'espace macroscopique $\mathfrak{X}_{\mathrm{macro}}$.

Une chaîne de Markov naturelle sur $\mathfrak{X}_{\mathrm{micro}}$ consiste à choisir aléatoirement une particule à chaque étape et à la déplacer de B_+ vers B_- ou de B_- vers B_+. Pour obtenir une chaîne de Markov apériodique, il sera utile de considérer une version paresseuse de ce processus, où l'on autorise la particule choisie à rester dans sa moitié de boîte avec probabilité $\frac{1}{2}$. Ainsi, la *marche aléatoire sur l'hypercube* est la chaîne de Markov sur $\mathfrak{X}_{\mathrm{micro}}$ dont la matrice de transition est :

$$P(x,y) = \begin{cases} \frac{1}{2} & \text{si } x = y, \\ \frac{1}{2N} & \text{si } x \text{ et } y \text{ pour seule coordonnée différente } x^{(i)} = -y^{(i)} \text{ pour un } i \in [\![1, N]\!], \\ 0 & \text{si } x \text{ et } y \text{ ont au moins deux coordonnées différentes.} \end{cases}$$

Dans le second cas donnant une probabilité de transition $\frac{1}{2N}$, on suppose donc qu'il existe un indice $i \in [\![1, N]\!]$ tel que $x^{(i)} = -y^{(i)}$ et $x^{(j)} = y^{(j)}$ pour tout $j \neq i$. Il y a N choix possibles pour l'indice i, de sorte que la somme des probabilités de transition est bien $\frac{1}{2} + N\,\frac{1}{2N} = \frac{1}{2} + \frac{1}{2} = 1$. Par rapport à la version vue en cours, notons qu'on a par exemple $P(x,x) = \frac{1}{2}$ au lieu de $P(x,x) = \frac{1}{N+1}$; ceci modifiera les valeurs propres et le temps de mélange, mais pas la méthode générale d'analyse. Dans tout ce qui suit, $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov de matrice P sur l'espace microscopique $\mathfrak{X}_{\mathrm{micro}}$. La première partie du problème rappelle l'analyse spectrale de cette chaîne, et la seconde partie introduit la chaîne image $(K_n)_{n\in\mathbb{N}}$ sur $\mathfrak{X}_{\mathrm{macro}}$, définie par $K_n = S(X_n)$. La dernière partie est consacrée à la comparaison des vitesses de convergence des deux chaînes.

1 Analyse spectrale de la chaîne microscopique

(Q1) Montrer que la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ est irréductible et apériodique sur l'espace $\mathfrak{X}_{\mathrm{micro}}$, et que sa mesure invariante est la mesure uniforme :

$$\pi(x) = \frac{1}{2^N} \quad \text{pour tout } x \in \mathfrak{X}_{\mathrm{micro}}.$$

Notons π_n la loi marginale de X_n : $\pi_n(x) = \mathbb{P}[X_n = x]$. Ces lois sont des vecteurs de $\mathbb{R}^{\mathfrak{X}_{\text{micro}}}$ qui vérifient la relation $\pi_n = \pi_0\, P^n$. Le théorème de convergence vers la loi stationnaire assure que $\lim_{n\to\infty} \pi_n = \pi$ (en tant que vecteurs dans $\mathbb{R}^{\mathfrak{X}_{\text{micro}}}$). On équipe $\mathbb{R}^{\mathfrak{X}_{\text{micro}}}$ du produit scalaire :

$$\langle \nu_1 \mid \nu_2 \rangle = \frac{1}{2^N} \sum_{x \in \mathfrak{X}_{\text{micro}}} \nu_1(x)\, \nu_2(x).$$

Autrement dit, $\mathbb{R}^{\mathfrak{X}_{\text{micro}}}$ est muni de la structure d'espace de Hilbert $\mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)$.

(Q2) Pour tout $\varepsilon = (\varepsilon^{(1)}, \ldots, \varepsilon^{(N)})$ dans $\{0,1\}^N$ et tout état microscopique x, on note $\varepsilon(x) = \prod_{i=1}^N (x^{(i)})^{\varepsilon^{(i)}}$; les éléments de $\{0,1\}^N$ peuvent donc être considérés comme des fonctions dans $\mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)$. Montrer que

$$\{\, \varepsilon \mid \varepsilon \in \{0,1\}^N \,\}$$

est une base orthonormale de $\mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)$, et que pour tout élément de cette base,

$$\varepsilon P = \left(1 - \frac{|\varepsilon|}{N}\right) \varepsilon, \quad \text{avec } |\varepsilon| = \sum_{i=1}^N \varepsilon^{(i)}.$$

(Q3) Dans tout ce qui suit, on suppose que N est pair et que $X_0 = (-1, -1, \ldots, -1, 1, 1, \ldots, 1)$: initialement, les $\frac{N}{2}$ premières particules sont presque sûrement dans B_- et les $\frac{N}{2}$ dernières particules sont dans B_+. Que vaut dans ce cas le produit scalaire $\langle \pi_0 \mid \varepsilon \rangle$ pour $\varepsilon \in \{0,1\}^N$ (vu comme une fonction de $\mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)$) ? Montrer que pour tout temps n, la loi marginale π_n s'écrit :

$$\pi_n = \frac{1}{2^N} \sum_{\varepsilon \in \{0,1\}^N} (-1)^{\left|\frac{\varepsilon}{2}\right|} \left(1 - \frac{|\varepsilon|}{N}\right)^n \varepsilon, \quad \text{avec } \left|\frac{\varepsilon}{2}\right| = \sum_{i=1}^{\frac{N}{2}} \varepsilon^{(i)}.$$

(Q4) Redémontrer à partir de la question précédente la convergence $\pi_n(x) \to \pi(x)$ pour tout $x \in \mathfrak{X}_{\text{micro}}$.

2 Chaîne image dans l'espace macroscopique

Pour tout état microscopique $x \in \mathfrak{X}_{\text{micro}}$, on note comme dans l'introduction $S(x) = \frac{N}{2} + \frac{1}{2}\sum_{i=1}^N x^{(i)}$ l'état macroscopique correspondant, qui est le nombre de particules dans la partie B_+ de la boîte. On a donc une application $S : \mathfrak{X}_{\text{micro}} \to \mathfrak{X}_{\text{macro}}$, et on s'intéresse au processus aléatoire $(K_n)_{n\in\mathbb{N}}$ sur l'espace macroscopique, défini par $K_n = S(X_n)$ avec $(X_n)_{n\in\mathbb{N}}$ chaîne de Markov de matrice P dans $\mathfrak{X}_{\text{micro}}$. En particulier, comme $X_0 = (-1, \ldots, -1, 1, \ldots, 1)$, $K_0 = \frac{N}{2}$ presque sûrement. Comme l'image d'une chaîne de Markov n'est pas en général une chaîne de Markov, il n'est pas clair que $(K_n)_{n\in\mathbb{N}}$ soit une chaîne de Markov ; les premières questions de cette partie traitent ce problème.

(Q5) Si $S(x) = k$, quels sont les entiers $l \in \mathfrak{X}_{\text{macro}}$ tels qu'existe $y \in \mathfrak{X}_{\text{micro}}$ avec $P(x,y) > 0$ et $S(y) = l$? En déduire que pour tout $(n+1)$-uplet d'entiers $(a_0, a_1, \ldots, a_n) \in (\mathfrak{X}_{\text{macro}})^{n+1}$, la probabilité

$$\mathbb{P}[K_0 = a_0,\, K_1 = a_1, \ldots, K_n = a_n]$$

est nulle s'il existe un indice $i \in [\![1, n]\!]$ tel que $|a_i - a_{i-1}| \geq 2$.

(Q6) On décompose

$$\mathbb{P}[K_0 = a_0,\, K_1 = a_1, \ldots, K_n = a_n] = \sum_{\substack{y_0, \ldots, y_n \in \mathfrak{X}_{\text{micro}} \\ \forall i,\, S(y_i) = a_i}} \mathbb{P}[X_0 = y_0,\, X_1 = y_1, \ldots, X_n = y_n].$$

En utilisant cette décomposition, montrer que cette probabilité s'écrit :

$$1_{(a_0=\frac{N}{2})}\, Q(a_0,a_1)\, Q(a_1,a_2)\cdots Q(a_{n-1},a_n),$$

avec

$$Q(k,k)=\frac{1}{2}\quad;\quad Q(k,k+1)=\frac{N-k}{2N}\quad;\quad Q(k,k-1)=\frac{k}{2N}.$$

On pourra procéder par récurrence sur n. En déduire que $(K_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur $\mathfrak{X}_{\mathrm{macro}}$ de matrice Q.

(Q7) On note ν la mesure image de la loi uniforme π par l'application S :

$$\nu(k)=\sum_{\substack{x\in\mathfrak{X}_{\mathrm{micro}}\\ S(x)=k}}\pi(x).$$

Calculer $\nu(k)$ pour tout $k\in[\![0,N]\!]$. Plus généralement, la fonction $S:\mathfrak{X}_{\mathrm{micro}}\to\mathfrak{X}_{\mathrm{macro}}$ induit une application linéaire

$$\begin{aligned}S_*:\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}},\pi)&\to\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}},\nu)\\ f&\mapsto\left(S_*f:k\mapsto\sum_{\substack{x\in\mathfrak{X}_{\mathrm{micro}}\\ S(x)=k}}f(x)\right).\end{aligned}$$

Par exemple, $\nu=S_*\pi$. On fait agir les deux matrices de transition P et Q à droite des vecteurs de $\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}},\pi)$ et $\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}},\nu)$. Montrer que $Q\circ S_*=S_*\circ P$.

(Q8) Montrer que la chaîne de Markov $(K_n)_{n\in\mathbb{N}}$ est irréductible et apériodique sur l'espace $\mathfrak{X}_{\mathrm{macro}}$, et que sa mesure de probabilité invariante est ν.

(Q9) Montrer que si f est vecteur propre pour P avec valeur propre λ, alors S_*f est vecteur propre pour Q avec valeur propre λ.

(Q10) Soit $k\in\mathfrak{X}_{\mathrm{macro}}$ et $\varepsilon\in\{0,1\}^N$ avec $|\varepsilon|=k$. Montrer que la fonction $F_k=S_*\varepsilon$ ne dépend que de k, et s'écrit :

$$F_k(m)=\sum_{j=0}^{k}(-1)^j\binom{k}{j}\binom{N-k}{N-m-j}\quad\text{pour tout } m\in\mathfrak{X}_{\mathrm{macro}}.$$

Montrer que $F_k\,Q=(1-\frac{k}{N})\,F_k$.

(Q11) On considère l'espace de Hilbert $\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}},\beta)$, avec $\frac{1}{\beta(m)}=\binom{N}{m}$. Montrer que $(F_0,F_1,\ldots,F_N)$ est une base orthogonale de $\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}},\beta)$, avec $\langle F_k\mid F_k\rangle=2^N\,\beta(k)$. On pourra commencer par établir l'identité suivante : pour tout y tel que $S(y)=m$,

$$\sum_{\substack{\varepsilon\in\{0,1\}^N\\ |\varepsilon|=k}}\varepsilon(y)=\frac{\binom{N}{k}}{\binom{N}{m}}\,F_k(m).$$

(Q12) On note ν_n la loi marginale de K_n : $\nu_n(k)=\mathbb{P}[K_n=k]$. Montrer que pour tout temps n, la loi marginale ν_n s'écrit :

$$\nu_n=\frac{1}{2^N}\sum_{k=0}^{N}1_{(k\text{ est pair})}\,(-1)^{\frac{k}{2}}\binom{\frac{N}{2}}{\frac{k}{2}}\left(1-\frac{k}{N}\right)^n F_k.$$

Il sera utile de calculer de deux façons le coefficient de x^k dans $(1-x)^{\frac{N}{2}}(1+x)^{\frac{N}{2}}$. Montrer aussi que $\lim_{n\to\infty}\nu_n=\nu$.

3 Comparaison des vitesses de convergence

Une bonne mesure de la vitesse de convergence pour la chaîne de Markov microscopique $(X_n)_{n\in\mathbb{N}}$ est la distance :

$$d_{\text{micro}}(\pi_n, \pi) = 2^N \|\pi_n - \pi\|_{\mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)} = \sqrt{2^N \sum_{x\in\mathfrak{X}_{\text{micro}}} |\pi_n(x) - \pi(x)|^2}.$$

(Q13) Calculer $d_{\text{micro}}(\pi_0, \pi)$. En remarquant que la loi stationnaire π est proportionnelle à l'une des fonctions $\varepsilon \in \{0,1\}^N$, et en utilisant la formule de la question (Q3) et l'orthonormalité des fonctions ε dans $\mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)$, montrer que :

$$d_{\text{micro}}(\pi_n, \pi) = \sqrt{\sum_{k=1}^{N} \binom{N}{k} \left(1 - \frac{k}{N}\right)^{2n}}.$$

De la même façon, une bonne mesure de la vitesse de convergence pour la chaîne de Markov macroscopique $(K_n)_{n\in\mathbb{N}}$ est la distance :

$$d_{\text{macro}}(\nu_n, \nu) = \sqrt{2^N}\, \|\nu_n - \nu\|_{\mathscr{L}^2(\mathfrak{X}_{\text{macro}}, \beta)} = \sqrt{\sum_{k\in\mathfrak{X}_{\text{macro}}} \frac{2^N}{\binom{N}{k}} |\nu_n(k) - \nu(k)|^2}.$$

(Q14) Montrer que les deux distances d_{micro} et d_{macro} s'écrivent :

$$(d_{\text{micro}}(\pi_n, \pi))^2 = \sum_{x\in\mathfrak{X}_{\text{micro}}} \frac{|\pi_n(x) - \pi(x)|^2}{\pi(x)}$$
$$(d_{\text{macro}}(\nu_n, \nu))^2 = \sum_{k\in\mathfrak{X}_{\text{macro}}} \frac{|\nu_n(k) - \nu(k)|^2}{\nu(k)}.$$

L'analogie entre les deux formules justifient leur utilisation et leur comparaison. Par ailleurs, on a vu en cours que ces distances $\mathscr{L}^2$ majoraient la distance $\mathscr{L}^1$ (distance en variation totale).

(Q15) En utilisant l'orthogonalité des fonctions F_k dans $\mathscr{L}^2(\mathfrak{X}_{\text{macro}}, \beta)$, montrer que :

$$d_{\text{macro}}(\nu_n, \nu) = \sqrt{\sum_{k=1}^{N} 1_{(k \text{ est pair})} \frac{\binom{\frac{N}{2}}{\frac{k}{2}} \binom{\frac{N}{2}}{\frac{k}{2}}}{\binom{N}{k}} \left(1 - \frac{k}{N}\right)^{2n}}.$$

(Q16) En déduire que $d_{\text{macro}}(\nu_n, \nu) \leq d_{\text{micro}}(\pi_n, \pi)$ pour tout n. Commenter cette inégalité.

(Q17) En s'inspirant des techniques de calcul vues en cours, montrer que si $k = \frac{1}{2} N(\log N + c)$, alors

$$(d_{\text{micro}}(\pi_n, \pi))^2 \leq \mathrm{e}^{\mathrm{e}^{-c}} - 1 = O(\mathrm{e}^{-c}),$$

et si $k = \frac{1}{4} Nc$, alors

$$(d_{\text{macro}}(\nu_n, \nu))^2 \leq \frac{1}{\sqrt{1 - \mathrm{e}^{-c}}} - 1 = O(\mathrm{e}^{-c}).$$

La dernière question est liée au fait que la chaîne microscopique $(X_n)_{n\in\mathbb{N}}$ admet une coupure avec temps de mélange $\frac{N \log N}{2}$, alors que la chaîne macroscopique $(K_n)_{n\in\mathbb{N}}$ converge tout de suite vers 0, en temps $O(N)$.

(Qbonus) Écrire des programmes qui simulent : les chaînes de Markov $(X_n)_{n\in\mathbb{N}}$ et $(K_n)_{n\in\mathbb{N}}$, les distances $d_{\text{micro}}(\pi_n, \pi)$ et $d_{\text{macro}}(\nu_n, \nu)$ (disons avec $N = 50$). Commenter les résultats de ces programmes.

Corrigé

(Q1) Les états que l'on peut atteindre en une transition à partir d'un état x sont :

- l'état x lui-même,
- et les N états $x\uparrow^{(i)}$ où l'on a changé seulement le signe de la i-ième coordonnée de x.

Si x et y sont deux états microscopiques arbitraires, alors en modifiant au fur et à mesure les coordonnées de x, on peut bien atteindre y, donc la chaîne de matrice P est irréductible. Par ailleurs, $P(x,x) = \frac{1}{2} > 0$ pour tout état x, donc la chaîne est apériodique. Finalement, comme $P(x,y) = P(y,x)$ pour tous états x et y, la matrice de transition P est symétrique et donc bistochastique, ce qui implique que la mesure uniforme est invariante :

$$(\pi P)(x) = \sum_{y \in \mathfrak{X}_{\mathrm{micro}}} \pi(y)\, P(y,x) = \frac{1}{2^N} \sum_{y \in \mathfrak{X}_{\mathrm{micro}}} P(x,y) = \frac{1}{2^N} = \pi(x) \quad \text{pour tout état } x \in \mathfrak{X}_{\mathrm{micro}}.$$

(Q2) Si ε et η sont deux éléments de $\{0,1\}^N$, alors

$$\langle \varepsilon \mid \eta \rangle = \frac{1}{2^N} \sum_{x \in \{+1,-1\}^N} \left(\prod_{i=1}^N (x^{(i)})^{\varepsilon^{(i)}+\eta^{(i)}} \right) = \prod_{i=1}^N \frac{1 + (-1)^{\varepsilon^{(i)}+\eta^{(i)}}}{2}.$$

Si $\varepsilon = \eta$, alors chaque terme du produit à droite vaut $\frac{1+1}{2} = 1$, donc $\langle \varepsilon \mid \varepsilon \rangle = 1$. Par contre, si $\varepsilon^{(i)} \neq \eta^{(i)}$ pour une coordonnée d'indice i, alors $\varepsilon^{(i)} + \eta^{(i)} = 1$, donc le terme correspondant dans le produit vaut $\frac{1-1}{2} = 0$, et $\langle \varepsilon \mid \eta \rangle = 0$. Les fonctions ε sont donc de norme 1 et deux-à-deux orthogonales : comme il y en a $2^N = \dim(\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}}, \pi))$, elles forment une base orthonormale de $\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}}, \pi)$.

Fixons maintenant $\varepsilon \in \{0,1\}^N$, et calculons εP :

$$\begin{aligned}(\varepsilon P)(x) &= \frac{1}{2}\,\varepsilon(x) + \frac{1}{2N} \sum_{i=1}^N \varepsilon(x\uparrow^{(i)}) = \varepsilon(x) \left(\frac{1}{2} + \frac{1}{2N} \sum_{i=1}^N (-1)^{\varepsilon^{(i)}} \right) \\ &= \varepsilon(x) \left(1 - \frac{1}{N} \sum_{i=1}^N \varepsilon^{(i)} \right) = \left(1 - \frac{|\varepsilon|}{N} \right) \varepsilon(x)\end{aligned}$$

En effet, si $h \in \{0,1\}$, alors $(-1)^h = 1 - 2h$. On a donc bien $\varepsilon P = (1 - \frac{|\varepsilon|}{N})\, \varepsilon$.

(Q3) Puisque les fonctions ε forment une base orthonormale de $\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}}, \pi)$ et sont des vecteurs propres pour l'action de la matrice P, on peut écrire :

$$\pi_n = \pi_0\, P^n = \sum_{\varepsilon \in \{0,1\}^N} \langle \pi_0 \mid \varepsilon \rangle\ (\varepsilon P^n) = \sum_{\varepsilon \in \{0,1\}^N} \langle \pi_0 \mid \varepsilon \rangle\ \left(1 - \frac{|\varepsilon|}{N} \right)^n \varepsilon.$$

Comme π_0 est le Dirac en $(-1, \ldots, -1, 1, \ldots, 1)$, on a

$$\langle \pi_0 \mid \varepsilon \rangle = \frac{1}{2^N} \prod_{i=1}^{\frac{N}{2}} (-1)^{\varepsilon^{(i)}} \prod_{i=\frac{N}{2}+1}^{N} 1^{\varepsilon^{(i)}} = \frac{(-1)^{|\frac{\varepsilon}{2}|}}{2^N}.$$

On en déduit la formule souhaitée pour la loi π_n.

(Q4) Remarquons que toutes les valeurs propres $(1 - \frac{|\varepsilon|}{N})$ sont comprises entre 0 et 1 ; de plus, le seul vecteur propre ε de valeur propre égale à 1 est $\varepsilon = (0, \ldots, 0)$, qui correspond à la fonction constante égale à 1. On a donc :

$$\pi_n(x) = \frac{1}{2^N} + \sum_{\varepsilon \neq (0,\ldots,0)} (-1)^{|\frac{\varepsilon}{2}|} \underbrace{\left(1 - \frac{|\varepsilon|}{N} \right)^n}_{\to_{n\to\infty} 0} \varepsilon(x)$$

et $\lim_{n\to\infty} \pi_n(x) = \frac{1}{2^N}$.

(Q5) Si x est un état microscopique avec k signes $+1$ et $N-k$ signes -1, et si $P(x,y)>0$, alors :

- soit $x=y$, et donc $S(y)=k$.
- soit y est obtenu en changeant l'un des signes $+1$ de x en un -1, donc $S(y)=S(x)-1=k-1$.
- soit y est obtenu en changeant l'un des signes -1 de x en un $+1$, donc $S(y)=S(x)+1=k+1$.

Les entiers l atteignables en un pas depuis k sont donc dans $\{k-1,k,k+1\}$. Ainsi, si $|a_i-a_{i-1}|\geq 2$, alors

$$\mathbb{P}[K_0=a_0,\,K_1=a_1,\ldots,K_n=a_n]\leq \mathbb{P}[S(X_{i-1})=a_{i-1} \text{ et } S(X_i)=a_i]=0.$$

La probabilité ci-dessus est donc nulle s'il existe un indice $i\in[\![1,n]\!]$ tel que $|a_i-a_{i-1}|\geq 2$.

(Q6) Fixons une suite $(a_0,a_1,\ldots,a_n)$ dans $\mathfrak{X}_{\mathrm{macro}}$, et démontrons la formule attendue par récurrence sur n. Le cas $n=0$ est bien vérifié :

$$\mathbb{P}[K_0=a_0]=1_{(a_0=\frac{N}{2})}$$

car $X_0=(-1,\ldots,-1,1,\ldots,1)$ presque sûrement, donc $K_0=\frac{N}{2}$ presque sûrement. Supposons la formule vérifiée au rang $n-1$. Si $(y_0,y_1,\ldots,y_n)$ est une suite d'états microscopiques tels que $S(y_i)=a_i$ pour tout i, alors

$$\mathbb{P}[X_0=y_0,X_1=y_1,\ldots,X_n=y_n]=\mathbb{P}[X_0=y_0,\ldots,X_{n-1}=y_{n-1}]\;P(y_{n-1},y_n)$$

car $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov. Par conséquent,

$$\mathbb{P}[K_0=a_0,\,K_1=a_1,\ldots,K_n=a_n]=\sum_{\substack{y_0,\ldots,y_{n-1}\\ \forall i,\,S(y_i)=a_i}}\mathbb{P}[X_0=y_0,\ldots,X_{n-1}=y_{n-1}]\left(\sum_{\substack{y_n\in\mathfrak{X}_{\mathrm{micro}}\\ S(y_n)=a_n}}P(y_{n-1},y_n)\right).$$

Montrons que $\alpha=\sum_{\substack{y_n\in\mathfrak{X}_{\mathrm{micro}}\\ S(y_n)=a_n}}P(y_{n-1},y_n)$ ne dépend que de a_{n-1} et a_n, et pas spécifiquement de y_{n-1} :

- Si $a_n=a_{n-1}$, alors il existe un unique état y_n avec $S(y_n)=a_n$ et $P(y_{n-1},y_n)>0$, à savoir $y_n=y_{n-1}$. La somme α vaut alors $\frac{1}{2}$.
- Si $a_n=a_{n-1}-1$, alors le nombre d'états y_n avec $S(y_n)=a_n$ et $P(y_{n-1},y_n)>0$ est a_{n-1} : ce sont les états obtenus à partir de y_n en changeant l'une des a_{n-1} coordonnées $+1$ en un -1. La somme α vaut alors $\frac{a_{n-1}}{2N}$.
- Si $a_n=a_{n-1}+1$, alors le nombre d'états y_n avec $S(y_n)=a_n$ et $P(y_{n-1},y_n)>0$ est $N-a_{n-1}$: ce sont les états obtenus à partir de y_n en changeant l'une des a_{n-1} coordonnées -1 en un $+1$. La somme α vaut alors $\frac{N-a_{n-1}}{2N}$.

Ainsi,

$$\begin{aligned}\mathbb{P}[K_0=a_0,\,K_1=a_1,\ldots,K_n=a_n]&=\sum_{\substack{y_0,\ldots,y_{n-1}\\ \forall i,\,S(y_i)=a_i}}\mathbb{P}[X_0=y_0,\ldots,X_{n-1}=y_{n-1}]\,Q(a_{n-1},a_n)\\&=\mathbb{P}[K_0=a_0,\ldots,K_{n-1}=a_{n-1}]\,Q(a_{n-1},a_n),\end{aligned}$$

et l'hypothèse de récurrence permet de conclure. Par définition, la formule obtenue implique que $(K_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur l'espace $\mathfrak{X}_{\mathrm{macro}}$ de matrice de transition Q.

(Q7) La mesure image $\nu=S_*\pi$ est la loi binomiale $\mathcal{B}(N,\frac{1}{2})$:

$$\nu(k)=\frac{1}{2^N}\;\mathrm{card}\big(\{\,x\in\mathfrak{X}_{\mathrm{micro}}\mid S(x)=k\,\}\big)=\frac{\binom{N}{k}}{2^N}.$$

Fixons une fonction $f \in \mathscr{L}^2(\mathfrak{X}_{\text{micro}}, \pi)$, et calculons $(S_* f)Q$. On a :

$$((S_*f)Q)(l) = \sum_{k=0}^{N} (S_*f)(k)\, Q(k,l) = \sum_{k=0}^{N} \sum_{\substack{x \in \mathfrak{X}_{\text{micro}} \\ S(x)=k}} f(x)\, Q(S(x),l) = \sum_{x \in \mathfrak{X}_{\text{micro}}} f(x)\, Q(S(x),l).$$

Or, d'après la question précédente, $Q(S(x),l) = \sum_{\substack{y \in \mathfrak{X}_{\text{micro}} \\ S(y)=l}} P(x,y)$, donc :

$$((S_*f)Q)(l) = \sum_{x \in \mathfrak{X}_{\text{micro}}} \sum_{\substack{y \in \mathfrak{X}_{\text{micro}} \\ S(y)=l}} f(x)P(x,y) = \sum_{\substack{y \in \mathfrak{X}_{\text{micro}} \\ S(y)=l}} (fP)(y) = (S_*(fP))(l).$$

Ainsi, $Q \circ S_* = S_* \circ P$.

(Q8) Comme $Q(k,k+1) > 0$, on peut passer en $l-k$ étapes de k à $l > k$ par des transitions de la matrice Q, et comme $Q(k,k-1)$, on peut de même passer en $k-l$ étapes de k à $l < k$ par des transitions de la matrice Q. Ainsi, la matrice est irréductible sur l'espace $\mathfrak{X}_{\text{macro}}$. Comme $Q(k,k) > 0$, elle est aussi apériodique. Finalement, puisque π est invariante par P, compte tenu de la question précédente,

$$\nu\, Q = (S_*\pi)\, Q = S_*(\pi P) = S_*\pi = \nu$$

et ν est invariante par Q.

(Q9) De la même façon, si f est un vecteur propre de P pour la valeur propre λ, alors

$$(S_*f)\, Q = S_*(fP) = \lambda\, (S_*f),$$

donc S_*f est un vecteur propre de Q pour la valeur propre λ.

(Q10) Fixons $\varepsilon \in \{0,1\}^N$ avec $|\varepsilon| = k$. On notera $I = \{i_1 < i_2 < \cdots < i_k\}$ l'ensemble des indices i tels que $\varepsilon^{(i)} = 1$, et $J = [\![1,N]\!] \setminus I$. Alors,

$$(S_*\varepsilon)(m) = \sum_{\substack{x \in \mathfrak{X}_{\text{micro}} \\ S(x)=m}} \varepsilon(x) = \sum_{\substack{x \in \mathfrak{X}_{\text{micro}} \\ S(x)=m}} \left(\prod_{i \in I} x^{(i)} \right).$$

Chaque produit $\prod_{i \in I} x^{(i)}$ met en jeu k signes $+1$ ou -1. Supposons donné x avec $S(x) = m$ et tel que l'ensemble des indices $i \in I$ avec $x^{(i)} = -1$ soit de cardinal j ; l'entier j appartient à $[\![0,k]\!]$. Alors, il y a $\binom{k}{j}$ possibilités pour ces indices, et les autres indices $i \in [\![1,N]\!]$ tels que $x^{(i)} = -1$ sont dans J ; il y en a $N-m-j$, et il y a $\binom{N-k}{N-m-j}$ possibilités pour ces autres indices. Ainsi,

$$\operatorname{card}\left(\left\{ x \in \mathfrak{X}_{\text{micro}} \;\middle|\; S(x) = m \text{ et il y a } j \text{ signes } -1 \text{ dans } \{x^{(i_1)}, \ldots, x^{(i_k)}\} \right\} \right) = \binom{k}{j} \binom{N-k}{N-m-j}.$$

On en déduit la formule

$$(S_*\varepsilon)(m) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \binom{N-k}{N-m-j}.$$

Le résultat ne dépend que de k ; notons F_k cette fonction sur $\mathfrak{X}_{\text{macro}}$. D'après la question précédente, avec $|\varepsilon| = k$, on a

$$F_k\, Q = (S_*\varepsilon)\, Q = S_*(\varepsilon\, P) = \left(1 - \frac{k}{N}\right) S_*\varepsilon = \left(1 - \frac{k}{N}\right) F_k.$$

(Q11) On constate pour commencer que si $S(y) = m$, alors pour toute fonction ε telle que $|\varepsilon| = k$, le nombre d'indices i tel que $y^{(i)} = -1$ et $\varepsilon^{(i)} = 1$ est un entier $j \in [\![0, k]\!]$. La fonction $\varepsilon(y)$ vaut alors $(-1)^j$, et si y et j sont fixés, alors ε est entièrement déterminé par l'un des $\binom{N-m}{j}$ choix pour les indices i tels que $y^{(i)} = -1$ et $\varepsilon^{(i)} = 1$, et par l'un des $\binom{m}{k-j}$ choix pour les indices i tels que $y^{(i)} = 1$ et $\varepsilon^{(i)} = 1$. On obtient donc :

$$\sum_{\substack{\varepsilon \in \{0,1\}^N \\ |\varepsilon|=k}} \varepsilon(y) = \sum_{j=0}^{k} (-1)^j \binom{N-m}{j} \binom{m}{k-j}.$$

Une petite manipulation avec les coefficients binomiaux montre que le terme de droite est égal à $\frac{\binom{N}{k}}{\binom{N}{m}}$ fois la somme considérée dans la question (Q10). Alors, si k et l sont fixés dans $[\![0, N]\!]$, calculons le produit scalaire $\langle F_k \mid F_l \rangle$ pour la mesure β sur $\mathfrak{X}_{\mathrm{macro}}$:

$$\begin{aligned}\langle F_k \mid F_l \rangle = \sum_{m=0}^{N} \frac{1}{\binom{N}{m}} F_k(m)\, F_l(m) &= \sum_{m=0}^{N} \frac{\binom{N}{m}}{\binom{N}{k}\binom{N}{l}} \left(\frac{\binom{N}{k}}{\binom{N}{m}} F_k(m) \right) \left(\frac{\binom{N}{l}}{\binom{N}{m}} F_l(m) \right) \\ &= \frac{1}{\binom{N}{k}\binom{N}{l}} \sum_{m=0}^{N} \binom{N}{m} \left(\sum_{\substack{\varepsilon \in \{0,1\}^N \\ |\varepsilon|=k}} \varepsilon(y_m) \right) \left(\sum_{\substack{\eta \in \{0,1\}^N \\ |\eta|=l}} \eta(y_m) \right),\end{aligned}$$

où pour chaque m, y_m est un état microscopique arbitraire de poids $S(y_m) = m$. Comme les sommes considérées ne dépendent que de m, on peut réécrire ce qui précède sous la forme :

$$\frac{1}{\binom{N}{k}\binom{N}{l}} \sum_{y \in \mathfrak{X}_{\mathrm{micro}}} \sum_{\substack{\varepsilon \in \{0,1\}^N \\ |\varepsilon|=k}} \sum_{\substack{\eta \in \{0,1\}^N \\ |\eta|=l}} \varepsilon(y)\, \eta(y)$$

puisqu'il y a $\binom{N}{m}$ états y_m avec $S(y_m) = m$. On conclut que :

$$\langle F_k \mid F_l \rangle_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}}, \beta)} = \frac{2^N}{\binom{N}{k}\binom{N}{l}} \sum_{\substack{\varepsilon, \eta \in \{0,1\}^N \\ |\varepsilon|=k,\, |\eta|=l}} \langle \varepsilon \mid \eta \rangle_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}}, \pi)} = \frac{2^N}{\binom{N}{k}} 1_{(k=l)}.$$

En effet, les produits scalaires $\langle \varepsilon \mid \eta \rangle_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}}, \pi)}$ sont tous nuls si $k \neq l$, et il y en a $\binom{N}{k}$ égaux à 1 si $k = l$.

(Q12) Il y a deux approches possibles au calcul de ν_n :

- soit décomposer ν_0 sur la base des fonctions F_k, puis utiliser la formule $F_k\, Q^n = (1 - \frac{k}{N})^n F_k$ pour obtenir $\nu_n = \nu_0\, Q^n$.
- soit utiliser la formule de la question (Q3), puis la relation $\nu_n = S_* \pi_n$.

Détaillons la première approche. Comme $(F_0, \ldots, F_N)$ est une base orthogonale de $\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}}, \beta)$, on peut écrire :

$$\nu_0 = \sum_{k=0}^{N} \frac{\langle \nu_0 \mid F_k \rangle_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}}, \beta)}}{\langle F_k \mid F_k \rangle_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}}, \beta)}} F_k = \frac{1}{2^N} \sum_{k=0}^{N} \binom{N}{k} \langle \nu_0 \mid F_k \rangle_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}}, \beta)} F_k = \frac{1}{2^N} \sum_{k=0}^{N} \frac{\binom{N}{k}}{\binom{N}{\frac{N}{2}}} F_k\left(\frac{N}{2}\right) F_k.$$

D'après la question (Q11), pour tout y avec $S(y) = \frac{N}{2}$, le coefficient devant F_k dans la somme à droite est égal à

$$\sum_{\substack{\varepsilon \in \{0,1\}^N \\ |\varepsilon|=k}} \varepsilon(y) = \sum_{j=0}^{k} (-1)^j \binom{\frac{N}{2}}{j} \binom{\frac{N}{2}}{k-j}.$$

Si k est impair, alors le j-ième terme de la somme se compense avec le $(k-j)$-ième, donc la somme vaut 0. Supposons maintenant k pair : alors, la somme ci-dessus est le coefficient de x^k dans $(1-x)^{\frac{N}{2}}(1+x)^{\frac{N}{2}} = (1-x^2)^{\frac{N}{2}}$, c'est-à-dire $(-1)^{\frac{k}{2}}\binom{\frac{N}{2}}{\frac{k}{2}}$. On obtient donc :

$$\nu_0 = \frac{1}{2^N}\sum_{k=0}^{N} 1_{(k \text{ est pair})}\,(-1)^{\frac{k}{2}}\binom{\frac{N}{2}}{\frac{k}{2}}\,F_k,$$

et la formule pour ν_n s'en déduit par application de Q^n. La convergence $\nu_n \to \nu$ s'en déduit : seul le terme $k=0$ subsiste par passage à la limite $n \to \infty$, et il vaut $\frac{1}{2^N}\,F_0 = \nu$.

(Q13) Avec $x_0 = (-1,\dots,-1,1,\dots,1)$, on a

$$(d_{\mathrm{micro}}(\pi_0,\pi))^2 = 2^N\left(1-\frac{1}{2^N}\right)^2 + 2^N(2^N-1)\left(\frac{1}{2^N}\right)^2 = 2^N - 1.$$

En effet, le terme correspondant à x_0 dans la somme donne une contribution $2^N(1-\frac{1}{2^N})^2$, et les 2^N-1 autres termes donnent chacun une contribution $2^N(\frac{1}{2^N})^2$. Ainsi,

$$d_{\mathrm{micro}}(\pi_0,\pi) = \sqrt{2^N-1}.$$

Pour le calcul général, remarquons que la loi stationnaire π est à un coefficient multiplicatif près l'une des fonctions ε :

$2^N\pi$ est la fonction constante à 1, associée à $\varepsilon = (0,0,\dots,0)$.

En utilisant l'orthonormalité des fonctions ε dans $\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}},\pi)$, on obtient alors à partir de la question (Q3) :

$$\begin{aligned}(d_{\mathrm{micro}}(\pi_n,\pi))^2 &= \left(\left\|\sum_{\varepsilon\neq(0,\dots,0)} (-1)^{|\frac{\varepsilon}{2}|}\left(1-\frac{|\varepsilon|}{N}\right)^n\varepsilon\right\|_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{micro}},\pi)}\right)^2 \\ &= \sum_{\varepsilon\neq(0,\dots,0)}\left(1-\frac{|\varepsilon|}{N}\right)^{2n}.\end{aligned}$$

La somme porte sur toutes les fonctions $\varepsilon \neq (0,\dots,0)$, car on considère $\pi_n - \pi$ au lieu de π. Pour chaque poids $k \in [\![0,N]\!]$, il y a $\binom{N}{k}$ fonctions $\varepsilon \in \{0,1\}^N$ de poids k, donc la somme ci-dessus vaut $\sum_{k=1}^{N}\binom{N}{k}\left(1-\frac{k}{N}\right)^{2n}$.

(Q14) C'est immédiat compte tenu des formules $\pi(x) = \frac{1}{2^N}$ et $\nu(k) = \frac{\binom{N}{k}}{2^N}$.

(Q15) En utilisant l'orthogonalité des fonctions F_k dans $\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}},\beta)$, on obtient à partir de la question (Q12) :

$$\begin{aligned}(d_{\mathrm{macro}}(\nu_n,\nu))^2 &= \frac{1}{2^N}\left(\left\|\sum_{k=1}^{N} 1_{(k \text{ est pair})}\,(-1)^{\frac{k}{2}}\binom{\frac{N}{2}}{\frac{k}{2}}\left(1-\frac{k}{N}\right)^n F_k\right\|_{\mathscr{L}^2(\mathfrak{X}_{\mathrm{macro}},\beta)}\right)^2 \\ &= \sum_{k=1}^{N} 1_{(k \text{ est pair})}\,\frac{\binom{\frac{N}{2}}{\frac{k}{2}}\binom{\frac{N}{2}}{\frac{k}{2}}}{\binom{N}{k}}\left(1-\frac{k}{N}\right)^{2n}\end{aligned}$$

car $\langle F_k \mid F_k\rangle = \frac{2^N}{\binom{N}{k}}$.

(Q16) Pour tout entier pair $k \in [0, N]$, $\binom{\frac{N}{2}}{\frac{k}{2}} \leq \binom{N}{k}$, donc la somme donnant $(d_{\text{macro}}(\nu_n, \nu))^2$ est plus petite que $(d_{\text{micro}}(\pi_n, \pi))^2$ (il y a moins de termes, et chaque terme est plus petit). Ceci implique que la convergence en loi $K_n \rightharpoonup \nu$ est *plus rapide* que la convergence en loi $X_n \rightharpoonup \pi$. Intuitivement : le nombre de particules dans chaque moitié de boîte (état macroscopique) est plus rapidement proche de sa loi stationnaire que la distribution individuelle des particules dans chaque moitié de boîte (état microscopique).

(Q17) Pour la convergence en loi de $(X_n)_{n\in\mathbb{N}}$, on majore chaque coefficient $\binom{N}{k}$ par $\frac{N^k}{k!}$, et chaque puissance $(1-\frac{k}{N})^{2n}$ par $\mathrm{e}^{-\frac{2nk}{N}}$. Alors :

$$(d_{\text{micro}}(\pi_n, \pi))^2 \leq \sum_{k=1}^{N} \frac{1}{k!} \left(N\,\mathrm{e}^{-\frac{2n}{N}}\right)^k \leq \sum_{k=1}^{\infty} \frac{1}{k!} \left(\mathrm{e}^{-c}\right)^k = \mathrm{e}^{\mathrm{e}^{-c}} - 1$$

si $n = \frac{1}{2}\,N\,(\log N + c)$. Pour la convergence en loi de $(K_n)_{n\in\mathbb{N}}$, on utilise la majoration suivante :

$$\frac{\binom{\frac{N}{2}}{k}\binom{\frac{N}{2}}{k}}{\binom{N}{2k}} = \binom{2k}{k} \frac{(\frac{N}{2})^{\downarrow k}(\frac{N}{2})^{\downarrow k}}{N^{\downarrow 2k}} \leq \frac{1}{4^k}\binom{2k}{k}.$$

Alors, avec $n = \frac{1}{4}\,Nc$, on obtient :

$$(d_{\text{macro}}(\nu_n, \nu))^2 \leq \sum_{k=1}^{\infty} \binom{2k}{k} \left(\frac{\mathrm{e}^{-\frac{4n}{N}}}{4}\right)^k = \sum_{k=1}^{\infty} \binom{2k}{k} \left(\frac{\mathrm{e}^{-c}}{4}\right)^k = \frac{1}{\sqrt{1-\mathrm{e}^{-c}}} - 1.$$