Temps de mélange pour les transpositions adjacentes

L'objectif de ce texte est d'étudier le temps de mélange d'une chaîne de Markov dont l'espace des états est l'ensemble $\mathfrak{S}(N)$ des permutations de taille $N \geq 2$. La chaîne en question correspond à l'expérience aléatoire suivante. On considère une rangée de livres dans une bibliothèque : les livres sont numérotés de 1 à N, et initialement ordonnés de gauche à droite dans l'ordre $123\ldots N$. Le lecteur vient régulièrement choisir deux livres adjacents de cette bibliothèque, en positions i et $i+1$ (avec $i \in [\![1, N]\!]$) ; il les replace ensuite dans la bibliothèque en échangeant leurs deux positions, c'est-à-dire en replaçant le livre qui était en position i à la position $i+1$, et *vice et versa*. On convient que si $i = N$, alors le lecteur ne vient prendre que le livre en position N, et il le replace au même endroit.

On suppose que les choix des entiers $i = i_n$ sont indépendants à chaque étape $n \in \mathbb{N}_*$, et uniformes dans $[\![1, N]\!]$. On note $\sigma_n(k)$ le numéro du livre placé en position k après n étapes ; chaque fonction $\sigma_n : [\![1, N]\!] \to [\![1, N]\!]$ est une bijection dans $\mathfrak{S}(N)$, et on a par hypothèse $\sigma_0 = \mathrm{id}_{[\![1,N]\!]}$. La suite de permutations aléatoires $(\sigma_n)_{n\in\mathbb{N}}$ est une chaîne de Markov, et on va montrer que sa loi converge vers la mesure uniforme sur les permutations, en un temps au moins de l'ordre de $O(N^3 \log N)$.

1 Propriétés de la chaîne de Markov

Pour $i \in [\![1, N-1]\!]$, on note τ_i la transposition $(i, i+1)$:

$$\tau_i(j) = \begin{cases} i+1 & \text{si } j = i, \\ i & \text{si } j = i+1, \\ j & \text{sinon.} \end{cases}$$

1. Montrer que $(\sigma_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur $\mathfrak{S}(N)$, dont la matrice de transition est :

$$P(\sigma, \sigma') = \begin{cases} \frac{1}{N} & \text{si } \sigma' = \sigma, \\ \frac{1}{N} & \text{si } \sigma' = \sigma \circ \tau_i \text{ pour un } i \in [\![1, N-1]\!], \\ 0 & \text{sinon.} \end{cases}$$

2. On rappelle que le groupe symétrique $\mathfrak{S}(N)$ est engendré par les transpositions (i, j) avec $1 \leq i < j \leq N$: (i, j) échange i et j et laisse tous les autres entiers $k \in [\![1, N]\!] \setminus \{i, j\}$ invariants. Pour tout couple (i, j) avec $i < j$, donner une expression de cette transposition comme produit de transpositions adjacentes τ_a.

3. Montrer que la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$ est irréductible sur $\mathfrak{S}(N)$ et apériodique. Montrer aussi que la matrice P est bistochastique. En déduire que la mesure invariante de la chaîne est la mesure uniforme donnée par

$$\pi(\sigma) = \frac{1}{N!}$$

pour toute permutation $\sigma \in \mathfrak{S}(N)$.

4. On pose $\pi_n(\sigma) = \mathbb{P}[\sigma_n = \sigma]$. Donner pour tout $\sigma \in \mathfrak{S}(N)$ la valeur de $\lim_{n\to\infty} \pi_n(\sigma)$.

5. Programmer avec `Python/SageMath` la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$. On pourra stocker les permutations σ_n sous forme de listes `Python` $[\sigma_n(1), \ldots, \sigma_n(N)]$, ou utiliser la classe `Permutation` de `SageMath`. Illustrer le résultat de la question précédente, par exemple avec des permutations de taille $n = 5$.

2 La méthode de Wilson

La suite du problème est consacrée à l'étude de la vitesse de convergence des lois π_n vers leur limite. Si μ et ν sont deux probabilités sur un ensemble mesurable $(\mathfrak{X}, \mathscr{F})$, on rappelle que leur distance en variation totale est

$$\|\mu - \nu\| = \sup_{A\in\mathscr{F}} |\mu(A) - \nu(A)|.$$

6. Soit μ et ν deux mesures de probabilité sur un ensemble mesurable $\mathfrak{X}$, et $f : \mathfrak{X} \to \mathfrak{Y}$ une application. On note $f_*\mu$ et $f_*\nu$ les mesures images de μ et ν par f. Montrer que

$$\|f_*\mu - f_*\nu\| \leq \|\mu - \nu\|.$$

7. Soit $f : \mathfrak{X} \to \mathbb{R}$ une application mesurable telle que

$$\int_{\mathfrak{X}} (f(x))^2 \, \mu(dx) < +\infty \qquad ; \qquad \int_{\mathfrak{X}} (f(x))^2 \, \nu(dx) < +\infty.$$

   Ceci permet de considérer l'espérance et la variance de f sous la loi μ et sous la loi ν. On note

$$m = \min(\mathbb{E}_\mu(f), \mathbb{E}_\nu(f));$$
$$\sigma^2 = \max(\mathrm{Var}_\mu(f), \mathrm{Var}_\nu(f)),$$

   et on suppose que

$$|\mathbb{E}_\mu[f] - \mathbb{E}_\nu[f]| \geq r\,\sigma$$

   pour une certaine constante $r > 0$. Utiliser la question précédente pour montrer que sous cette hypothèse,

$$\|\mu - \nu\| \geq 1 - \frac{8}{r^2}.$$

   Étant donnée la partie mesurable $(m + \frac{r\sigma}{2}, +\infty) = B \subset \mathbb{R}$, on pourra utiliser l'inégalité de Chebyshev pour contrôler $(f_*\mu)(B)$ et $(f_*\nu)(B)$.

Soit P une matrice stochastique irréductible apériodique sur un ensemble fini $\mathfrak{X}$, et π la mesure invariante de P, qu'on suppose réversible. On rappelle que dans ce cas les valeurs propres de P sont des nombres réels dans l'intervalle $(-1, 1]$. On considère une fonction non nulle $f : \mathfrak{X} \to \mathbb{R}$ que l'on représente par un vecteur colonne $(f(x))_{x\in\mathfrak{X}}$, et qui vérifie les deux hypothèses suivantes :

(a) La fonction f est un vecteur propre à droite pour P et pour une valeur propre λ avec $\frac{1}{2} < \lambda < 1$:

$$Pf = \lambda\, f.$$

(b) Pour tout $x \in \mathfrak{X}$, la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ issue de x et de matrice de transition P vérifie

$$\mathbb{E}_x[(f(X_1) - f(x))^2] \leq R$$

pour une certaine constante R.

On fixe $x \in \mathfrak{X}$ point de départ de la chaîne $(X_n)_{n\in\mathbb{N}}$, et on note $D_n = f(X_{n+1}) - f(X_n)$.

8. Montrer que $\mathbb{E}_x[D_n \mid X_n = y] = (\lambda - 1)\, f(y)$ et que $\mathbb{E}_x[(D_n)^2 \mid X_n = y] \leq R$. En déduire que

$$\mathbb{E}_x[(f(X_{n+1}))^2] \leq R + (2\lambda - 1)\, \mathbb{E}_x[(f(X_n))^2].$$

9. Montrer que, pour tout entier $n \in \mathbb{N}$,

$$\mathrm{Var}_x(f(X_n)) \leq \frac{R}{2(1-\lambda)}.$$

Montrer que $\mathrm{Var}_\pi(f)$ vérifie la même inégalité.

10. On note π_n la loi de X_n sous $\mathbb{P}_x$. Calculer $\mathbb{E}_{\pi_n}[f]$ et $\mathbb{E}_\pi[f]$. En utilisant les résultats des questions précédentes, montrer que

$$\|\pi_n - \pi\| \geq 1 - \frac{4R}{(1-\lambda)\lambda^{2n}\, |f(x)|^2}.$$

3 Spectre de la chaîne des transpositions adjacentes

On considère de nouveau la chaîne de Markov $(\sigma_n)_{n\in\mathbb{N}}$ de la première partie du problème.

11. Fixons $a \in [\![1, N]\!]$. Montrer que $((\sigma_n)^{-1}(a))_{n\in\mathbb{N}}$ est une chaîne de Markov sur $[\![1, N]\!]$ de matrice de transition

$$Q(k,l) = \begin{cases} \frac{1}{N} & \text{si } l = k-1, \\ \frac{1}{N} & \text{si } l = k+1, \\ 1 - \frac{1}{N} & \text{si } l = k \in \{1, N\}, \\ 1 - \frac{2}{N} & \text{si } l = k \in [\![2, N-1]\!]. \end{cases}$$

Que représente concrètement la valeur $\sigma_n^{-1}(a)$?

12. Montrer que $f(k) = \cos(\frac{(2k-1)\pi}{2N})$ est une fonction propre de Q, de valeur propre

$$\lambda = 1 + \frac{2}{N}\left(\cos\frac{\pi}{N} - 1\right).$$

Montrer que pour $N \geq 4$, cette valeur propre vérifie

$$1 - \frac{\pi^2}{N^3} < \lambda < 1 - \frac{\pi^2}{N^3}\left(1 - \frac{1}{N^2}\right)$$

et satisfait aux hypothèses de la méthode de Wilson.

13. Soit $F : \mathfrak{S}(N) \to \mathbb{R}$ la fonction définie par :

$$F(\sigma) = \sum_{a=1}^{N} f(a)\, f(\sigma^{-1}(a)).$$

Montrer que F est un vecteur propre de la matrice P pour la valeur propre λ de la question précédente. Montrer qu'on a aussi

$$F(\mathrm{id}_{[\![1,N]\!]}) = \frac{N}{2}.$$

14. Montrer que pour toute permutation σ,

$$\mathbb{E}_\sigma[(F(\sigma_1) - F(\sigma))^2] \leq \frac{4\pi^2}{N^2}.$$

On pourra remarquer que f est $\frac{\pi}{N}$-Lipschitzienne, et que si $(\sigma_n)_{n\in\mathbb{N}}$ est la chaîne de Markov issue de $\sigma_0 = \sigma$, alors dans $F(\sigma_1)$, au plus deux termes diffèrent de ceux de la somme $F(\sigma)$.

15. En appliquant l'inégalité de la question 10. avec $N \geq 4$, montrer que

$$\|\pi_n - \pi\| \geq 1 - \frac{K}{N(\lambda)^{2n}}.$$

pour une certaine constante $K > 0$. En déduire que pour tout $\varepsilon > 0$, il existe une constante réelle C_ε telle que

$$(\|\pi_n - \pi\| \leq \varepsilon) \quad \Rightarrow \quad \left(n \geq \frac{N^3(\log N + C_\varepsilon)}{2\pi^2}\right).$$

Ainsi, il faut un temps au moins égal à $\frac{N^3 \log N(1+o(1))}{2\pi^2}$ pour mélanger correctement les livres. Il a été démontré en 2013 par Lacoin que le résultat ci-dessus est précis, au sens suivant : si $n \geq (1+\eta)\,\frac{N^3 \log N}{2\pi^2}$ pour $\eta > 0$, alors réciproquement $\|\pi_n - \pi\|$ est petit.

Corrigé

1. Si σ est la permutation désignant les numéros des livres dans la bibliothèque, alors la suite
$$(\sigma(1), \sigma(2), \ldots, \sigma(N))$$
est la suite de ces numéros lorsqu'on consulte les livres de gauche à droite. L'échange des livres en position i et $i+1$ donne la suite
$$\begin{aligned}&(\sigma(1), \ldots, \sigma(i-1), \sigma(i+1), \sigma(i), \sigma(i+2), \ldots, \sigma(N))\\ &= (\sigma(\tau_i(1)), \sigma(\tau_i(i-1)), \sigma(\tau_i(i)), \sigma(\tau_i(i+1)), \sigma(\tau_i(i+2)), \ldots, \sigma(\tau_i(N))),\end{aligned}$$
donc est encodée par la permutation $\sigma \circ \tau_i$. On peut alors écrire $\sigma_{n+1} = g(\sigma_n, i_{n+1})$, où $(i_n)_{n\in\mathbb{N}^*}$ est une suite de variables i.i.d. choisies uniformément dans $[\![1, N]\!]$, et $g : \mathfrak{S}(N) \times [\![1, N]\!] \to \mathfrak{S}(N)$ est la fonction donnée par :
$$g(\sigma, i) = \begin{cases} \sigma \circ \tau_i & \text{si } i \in [\![1, N-1]\!], \\ \sigma & \text{si } i = N. \end{cases}$$
Par le théorème de représentation, $(\sigma_n)_{n\in\mathbb{N}}$ est bien une chaîne de Markov, de matrice de transition
$$P(\sigma, \sigma') = \mathbb{P}[g(\sigma, i) = \sigma'] = \begin{cases} \frac{1}{N} & \text{si } \sigma' = \sigma \circ \tau_i \text{ pour un } i \in [\![1, N-1]\!], \\ \frac{1}{N} & \text{si } \sigma' = \sigma, \\ 0 & \text{sinon.} \end{cases}$$

2. Un cycle $(a, a+1, a+2, \ldots, b)$ s'écrit comme le produit $(a, a+1)(a+1, a+2)\cdots(b-1, b) = \tau_a\tau_{a+1}\cdots\tau_{b-1}$. Par ailleurs, on a pour toute permutation σ et toute transposition (i, j) la relation de commutation
$$\sigma \circ (i, i') \circ \sigma^{-1} = (\sigma(i), \sigma(i')).$$
Si l'on applique cette relation avec $\sigma = (i+1, i+2, \ldots, j)^{-1} = (j, j-1, \ldots, i+1)$ et $i' = i+1$, on voit que pour $i < j$,
$$\begin{aligned}(i, j) &= (j, j-1, \ldots, i+1)(i, i+1)(i+1, i+2, \ldots, j)\\ &= \tau_{j-1}\tau_{j-2}\cdots\tau_{i+1}\,\tau_i\,\tau_{i+1}\tau_{i+2}\cdots\tau_{j-1}.\end{aligned}$$

3. Comme $P(\sigma, \sigma) = \frac{1}{N} > 0$ pour toute permutation σ, la chaîne est apériodique. D'après la question précédente, les transpositions adjacentes engendrent dans $\mathfrak{S}(N)$ toutes les transpositions, donc toutes les permutations :
$$\mathfrak{S}(N) = \langle (i, j) \,|\, 1 \leq i < j \leq N \rangle = \langle \tau_i \,|\, 1 \leq i \leq N-1 \rangle.$$
Par conséquent, étant données deux permutations $\sigma \neq \rho$ dans $\mathfrak{S}(N)$, on peut décomposer $\sigma^{-1}\rho$ comme un produit de transpositions adjacentes :
$$\sigma^{-1}\rho = \tau_{a_1}\tau_{a_2}\cdots\tau_{a_\ell}.$$
Notons que cette décomposition n'est pas unique, même si l'on demande en plus qu'elle soit de longueur minimale : on a par exemple $(i, i+2) = \tau_i\tau_{i+1}\tau_i = \tau_{i+1}\tau_i\tau_{i+1}$. La décomposition précédente implique que
$$P^\ell(\sigma, \tau) \geq \mathbb{P}[i_1 = a_1, \ldots, i_\ell = a_\ell] = \frac{1}{N^\ell} > 0,$$

donc la chaîne de Markov est irréductible.

Pour le caractère bistochastique, notons qu'on a même une matrice P qui est symétrique : comme $(\tau_i)^2 = \mathrm{id}_{[1,N]}$, si $\sigma' = \sigma\tau_i$, alors $\sigma = \sigma'\tau_i$. Ainsi, $P(\sigma, \sigma') = P(\sigma', \sigma)$ pour toutes permutations $\sigma, \sigma' \in \mathfrak{S}(N)$. Ceci implique que la mesure invariante est la mesure uniforme.

4. Par le théorème ergodique, la chaîne étant irréductible apériodique, la loi marginale π_n converge vers l'unique mesure de probabilité stationnaire. Ainsi,

$$\lim_{n\to\infty} \pi_n(\sigma) = \pi(\sigma) = \frac{1}{N!}$$

pour toute permutation $\sigma \in \mathfrak{S}(N)$.

5. Le programme suivant (dans `SageMath`) calcule la permutation aléatoire $\sigma_n \in \mathfrak{S}([\![1, N]\!])$:

```
def RandomPerm(N,n):
    res = Permutation([1..N])
    for a in range(n):
        i = randint(1,N)
        if i == N:
            pass
        else:
            res *= Permutation((i,i+1))
    return res
```

Pour illustrer le théorème limite, on va calculer avec $N = 5$ la probabilité empirique d'avoir $\sigma_n = 52413$ avec $n = 100$, et 10000 essais :

```
sigma = Permutation([5,2,4,1,3])
count = 0
for i in range(10000):
    if RandomPerm(5,100) == sigma:
            count += 1
RR(count/10000)
```

On obtient sur un essai la valeur 0.0085, très proche de la probabilité uniforme $\frac{1}{120} \simeq 0.008333$.

6. Si $B \subset \mathfrak{Y}$ est mesurable, on a

$$|f_*\mu(B) - f_*\nu(B)| = |\mu(f^{-1}(B)) - \nu(f^{-1}(B))| = |\mu(A) - \nu(A)|$$

avec $A = f^{-1}(B) \subset \mathfrak{X}$. Le terme de droite est majoré par $\|\mu - \nu\|$, donc en passant à la borne supérieure sur les parties mesurables $B \subset \mathfrak{Y}$, on en déduit que $\|f_*\mu - f_*\nu\| \leq \|\mu - \nu\|$.

7. On va évaluer $f_*\mu(B)$ et $f_*\nu(B)$, en supposant par exemple que $m = \mathbb{E}_\mu(f)$ (l'autre cas est symétrique). Par l'inégalité de Bienaymé–Chebyshev,

$$f_*\mu(B) = \mu\left(\left\{x \,|\, f(x) \geq \mathbb{E}_\mu[f] + \frac{r\sigma}{2}\right\}\right) \leq \frac{4\,\mathrm{Var}_\mu(f)}{r^2\sigma^2} \leq \frac{4}{r^2}.$$

De même,

$$\begin{aligned}
f_*\nu(B) &= \nu\Big(\Big\{x \,|\, f(x) \geq \mathbb{E}_\mu[f] + \frac{r\sigma}{2}\Big\}\Big) \geq \nu\Big(\Big\{x \,|\, f(x) \geq \mathbb{E}_\nu[f] - \frac{r\sigma}{2}\Big\}\Big) \\
&\geq 1 - \nu\Big(\Big\{x \,|\, f(x) \leq \mathbb{E}_\nu[f] - \frac{r\sigma}{2}\Big\}\Big) \\
&\geq 1 - \frac{4\,\mathrm{Var}_\nu(f)}{r^2\sigma^2} \geq 1 - \frac{4}{r^2}.
\end{aligned}$$

On a donc $\|\mu - \nu\| \geq \|f_*\mu - f_*\nu\| \geq |f_*\mu(B) - f_*\nu(B)| \geq 1 - \frac{8}{r^2}$.

8. La loi de X_{n+1} sachant $X_n = y$ est $P(y,\cdot)$, donc

$$\mathbb{E}_x[D_n \,|\, X_n = y] = \mathbb{E}[f(X_{n+1}) \,|\, X_n = y] - f(y) = (Pf)(y) - f(y) = (\lambda - 1)\, f(y)$$

puisque $Pf = \lambda\, f$. Par ailleurs, par la propriété de Markov simple appliquée au temps n, la loi de X_{n+1} sachant $X_n = y$ est la loi de X_1 sous $\mathbb{E}_y$, donc

$$\mathbb{E}_x[(D_n)^2 \,|\, X_n = y] = \mathbb{E}_y[(f(X_1) - f(y))^2] \leq R \text{ par hypothèse.}$$

Comme ceci est vrai pour tout y, la variable aléatoire $\mathbb{E}_x[(D_n)^2 \,|\, X_n]$ est elle aussi toujours plus petite que R. On calcule alors $\mathbb{E}_x[(f(X_{n+1}))^2]$ en conditionnant par rapport à la valeur de X_n :

$$\begin{aligned}
\mathbb{E}_x[(f(X_{n+1}))^2] &= \mathbb{E}_x[\mathbb{E}_x[(D_n + f(X_n))^2 \,|\, X_n]] \\
&= \mathbb{E}_x\big[\mathbb{E}_x[(D_n)^2 \,|\, X_n] + 2\,(f(X_n))\,\mathbb{E}_x[D_n \,|\, X_n] + (f(X_n))^2\big] \\
&\leq R + (2(\lambda - 1) + 1)\,\mathbb{E}_x[(f(X_n))^2].
\end{aligned}$$

9. En itérant l'inégalité précédente, on obtient :

$$\begin{aligned}
\mathbb{E}_x[(f(X_n))^2] &\leq R + (2\lambda - 1)\,\mathbb{E}_x[(f(X_{n-1}))^2] \leq R(1 + (2\lambda - 1)) + (2\lambda - 1)^2\,\mathbb{E}_x[(f(X_{n-2}))^2] \\
&\leq R(1 + (2\lambda - 1) + \cdots + (2\lambda - 1)^{n-1}) + (2\lambda - 1)^n\,(f(x))^2 \\
&\leq R\,\frac{1 - (2\lambda - 1)^n}{2(1 - \lambda)} + (2\lambda - 1)^n\,(f(x))^2 \leq \frac{R}{2(1 - \lambda)} + (2\lambda - 1)^n\,(f(x))^2.
\end{aligned}$$

Par ailleurs, $\mathbb{E}_x[f(X_n)] = (P^n f)(x) = \lambda^n f(x)$, donc :

$$\mathrm{Var}_x(f(X_n)) \leq \frac{R}{2(1 - \lambda)} + ((2\lambda - 1)^n - \lambda^{2n})\,(f(x))^2.$$

Or, $\lambda^2 \geq 2\lambda - 1$, donc le dernier terme est négatif. On conclut que

$$\mathrm{Var}_x(f(X_n)) = \mathrm{Var}_{\pi_n}(f) \leq \frac{R}{2(1 - \lambda)}.$$

Cette inégalité passe clairement à la limite $n \to \infty$, car $\mu \mapsto \mathbb{E}_\mu[f]$ ou $\mu \mapsto \mathrm{Var}_\mu[f]$ sont des fonctions continues des coefficients $\mu(x)$ de la probabilité μ. Elle est donc encore valable pour la mesure stationnaire π.

10. On a déjà calculé $\mathbb{E}_{\pi_n}[f] = \lambda^n\, f(x)$, et par passage à la limite, $\mathbb{E}_\pi[f] = 0$ puisque $\lambda < 1$. Les hypothèses de la question 7. s'appliquent donc avec

$$|\mathbb{E}_{\pi_n}[f] - \mathbb{E}_\pi[f]| = \lambda^n |f(x)| \geq r\,\sigma$$

avec $r = \sqrt{\frac{2(1-\lambda)}{R}}\,\lambda^n\,|f(x)|$. Ainsi,

$$\|\pi_n - \pi\| \geq 1 - \frac{4R}{(1 - \lambda)\lambda^{2n}\,|f(x)|^2}.$$

11. Notons $x_n = (\sigma_n)^{-1}(a)$; c'est la position du livre numéroté a au temps n. Avec la même suite de variables aléatoires indépendantes uniformes $(i_n)_{n\in\mathbb{N}^*}$ que dans les questions précédentes, notons que
$$x_{n+1} = (\sigma_{n+1})^{-1}(a) = (\tau_{i_{n+1}} \circ (\sigma_n)^{-1})(a) = \tau_{i_{n+1}}(x_n).$$
Il existe donc une fonction mesurable $g : [\![1,N]\!]^2 \to [\![1,N]\!]$ telle que $x_{n+1} = g(x_n, i_{n+1})$ avec une suite de variables i.i.d. $(i_n)_{n\in\mathbb{N}^*}$. Ceci prouve que $(x_n)_{n\in\mathbb{N}}$ est une chaîne de Markov, de matrice de transition
$$Q(k,l) = \mathbb{P}[\tau_i(k) = l],$$
avec par convention $\tau_N = \mathrm{id}_{[\![1,N]\!]}$. L'analyse des différents cas donne la valeur des coefficients de cette matrice.

12. On rappelle que $\cos a + \cos b = 2\,\cos(\frac{b-a}{2})\,\cos(\frac{a+b}{2})$. Si $k \notin \{1,N\}$, on
$$\begin{aligned}(Qf)(k) &= \sum_l Q(k,l)\,f(l)\\ &= \left(1-\frac{2}{N}\right)\,\cos\left(\frac{(2k-1)\pi}{2N}\right)+\frac{1}{N}\,\left(\cos\left(\frac{(2k-3)\pi}{2N}\right)+\cos\left(\frac{(2k+1)\pi}{2N}\right)\right)\\ &= \left(1+\frac{2}{N}\left(\cos\frac{\pi}{N}-1\right)\right)\,\cos\left(\frac{(2k-1)\pi}{2N}\right) = \left(1+\frac{2}{N}\left(\cos\frac{\pi}{N}-1\right)\right)\,f(k).\end{aligned}$$
On a la même identité pour $k=1$ et $k=N$, car pour tout angle θ, on a :
$$(\cos 2\theta - 1)(\cos\theta) = \frac{(\mathrm{e}^{2\mathrm{i}\theta}+\mathrm{e}^{-2\mathrm{i}\theta}-2)(\mathrm{e}^{\mathrm{i}\theta}+\mathrm{e}^{-\mathrm{i}\theta})}{4} = \frac{\mathrm{e}^{3\mathrm{i}\theta}+\mathrm{e}^{-3\mathrm{i}\theta}-\mathrm{e}^{\mathrm{i}\theta}-\mathrm{e}^{-\mathrm{i}\theta}}{4} = \frac{\cos 3\theta - \cos\theta}{2}.$$
Ceci implique avec $k=1$:
$$\begin{aligned}(Qf)(1) &= Q(1,1)\,f(1) + Q(1,2)\,f(2)\\ &= \left(1-\frac{1}{N}\right)\,\cos\frac{\pi}{2N} + \frac{1}{N}\,\cos\frac{3\pi}{2N} = \cos\frac{\pi}{2N} + \frac{2}{N}\left(\frac{\cos\frac{3\pi}{2N}-\cos\frac{\pi}{2N}}{2}\right)\\ &= \left(\cos\frac{\pi}{2N}\right)\left(1+\frac{2}{N}\left(\cos\frac{\pi}{N}-1\right)\right) = \left(1+\frac{2}{N}\left(\cos\frac{\pi}{N}-1\right)\right)\,f(1).\end{aligned}$$
Le calcul pour $k=N$ est similaire. Ainsi, f est vecteur propre de Q pour la valeur propre $\lambda = 1+\frac{2}{N}(\cos\frac{\pi}{N}-1)$. Pour $N \geq 4$, $\frac{\pi}{N} < 1$ et la série entière du cosinus est alternée, donc vérifie
$$-\frac{\pi^2}{2N^2} < \cos\frac{\pi}{N} - 1 < -\frac{\pi^2}{2N^2} + \frac{\pi^4}{24N^4} = -\frac{\pi^2}{2N^2}\left(1-\frac{\pi^2}{12N^2}\right) < -\frac{\pi^2}{2N^2}\left(1-\frac{1}{N^2}\right).$$
L'estimée de λ s'en déduit, et on a bien
$$1 > \lambda > 1 - \frac{\pi^2}{64} \geq 0.84 > \frac{1}{2}.$$

13. Pour $a \in [\![1,N]\!]$, notons $g_a : \sigma \in \mathfrak{S}(N) \mapsto f(\sigma^{-1}(a))$. Si $(\sigma_n)_{n\in\mathbb{N}}$ (respectivement, $(x_n)_{n\in\mathbb{N}}$) désigne la chaîne de Markov de matrice de transition P (respectivement, de matrice de transition Q), alors
$$\begin{aligned}(Pg_a)(\sigma) &= \mathbb{E}_\sigma[g_a(\sigma_1)] = \mathbb{E}_\sigma[f((\sigma_1)^{-1}(a))] = \mathbb{E}_{\sigma^{-1}(a)}[f(x_1)] = (Qf)(\sigma^{-1}(a))\\ &= \lambda\,f(\sigma^{-1}(a)) = \lambda\,g_a(\sigma).\end{aligned}$$

Donc, chaque fonction g_a est vecteur propre de P pour la valeur propre λ, et il en va de même pour $F = \sum_{a=1}^{N} f(a)\, g_a$. Calculons maintenant $S_N = F(\mathrm{id}_{[1,N]})$:

$$\begin{aligned} S_N &= \sum_{k=1}^{N} \cos^2\left(\frac{(2k-1)\pi}{2N}\right) = \sum_{k=1}^{N} \frac{1}{2} + \frac{1}{2}\cos\left(\frac{(2k-1)\pi}{N}\right) \\ &= \frac{N}{2} + \frac{1}{4}\,\mathrm{e}^{-\mathrm{i}\frac{2\pi}{N}} \sum_{k=1}^{N} \mathrm{e}^{\mathrm{i}\frac{2k\pi}{N}} + \frac{1}{4}\,\mathrm{e}^{\mathrm{i}\frac{2\pi}{N}} \sum_{k=1}^{N} \mathrm{e}^{-\mathrm{i}\frac{2k\pi}{N}}. \end{aligned}$$

Les sommes de racines N-ièmes de l'unité s'annulent, donc $S_N = \frac{N}{2}$.

14. Remarquons pour commencer que $F(\sigma) = \sum_{a=1}^{N} f(a)\, f(\sigma(a))$, en faisant le changement de variable $a = \sigma^{-1}(b)$. Alors, si $\rho = \sigma \circ \tau_i$, on a :

$$\begin{aligned} F(\sigma) - F(\rho) &= f(i)\, f(\sigma(i)) + f(i+1)\, f(\sigma(i+1)) - f(i)\, f(\sigma(i+1)) - f(i+1)\, f(\sigma(i)) \\ &= (f(i) - f(i+1))(f(\sigma(i)) - f(\sigma(i+1))). \end{aligned}$$

Comme $i \mapsto f(i)$ est $\frac{\pi}{N}$-Lipschitzienne, le premier terme est plus petit que $\frac{\pi}{N}$, tandis que le second peut être majoré par 2. Ainsi, pour tout $\rho = \sigma_1$ possible,

$$|F(\sigma) - F(\sigma_1)| \leq \frac{2\pi}{N}$$

d'où l'inégalité en élevant au carré et en prenant l'espérance.

15. On peut appliquer l'inégalité de la question 10. avec $x = \mathrm{id}_{[1,N]}$, $F(x) = \frac{N}{2}$, $R = \frac{4\pi^2}{N^2}$ et $\lambda = 1 + \frac{2}{N}(\cos\frac{\pi}{N} - 1)$:

$$\|\pi_n - \pi\| \geq 1 - \frac{32\pi^2}{N^3(1-\cos\frac{\pi}{N}))\,\lambda^{2n}} \geq 1 - \frac{32\pi^2}{N^3(1-\cos\frac{\pi}{N})\lambda^{2n}} \geq 1 - \frac{64}{N(1-\frac{1}{N^2})\,\lambda^{2n}},$$

d'où la première inégalité demandée avec $K = \frac{64}{1-\frac{1}{16}} \leq 69$. Si $\|\pi_n - \pi\| \leq \varepsilon$, alors :

$$\begin{aligned} \varepsilon &\geq 1 - \frac{K}{N\lambda^{2n}}; \\ \frac{K}{1-\varepsilon} &\geq N\lambda^{2n}; \\ \log\frac{K}{1-\varepsilon} &\geq \log N + 2n\log\lambda. \end{aligned}$$

Comme $\lambda - 1 = -\frac{\pi^2}{N^3} + O(\frac{1}{N^5})$, la même estimée vaut pour $\log\lambda$, et la dernière inégalité devient

$$\begin{aligned} n &\geq \frac{N^3}{2\pi^2}\left(\log N - \log\frac{K}{1-\varepsilon}\right)(1 + O(N^{-2})) = \frac{N^3}{\pi^2}\left(\log N - \log\frac{K}{1-\varepsilon} + O(N^{-1})\right) \\ &\geq \frac{N^3}{\pi^2}(\log N + C_\varepsilon) \end{aligned}$$

pour une certaine constante C_ε.