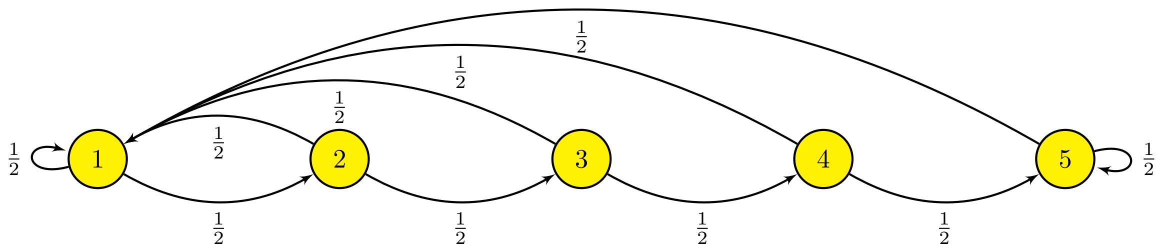
M2 – Temps de mélange – Corrigé 2020

Problème 1

1. Voici le graphe induit par la chaîne dans le cas $n = 5$:



2. Le noyau P est irreductible (comme le montre par exemple le chemin $1 \to 2 \to \cdots \to n \to 1$). Il y a donc une unique loi invariante π. L'équation $\pi = \pi P$ s'écrit ici

$$\begin{aligned}
\pi(1) &= \frac{\pi(1)}{2} + \cdots + \frac{\pi(n)}{2}; \\
\pi(k) &= \frac{\pi(k-1)}{2} \text{ pour tout } k \in \{2, \ldots, n-1\}; \\
\pi(n) &= \frac{\pi(n-1)}{2} + \frac{\pi(n)}{2}.
\end{aligned}$$

Par ailleurs, on doit avoir $\pi(1) + \cdots + \pi(n) = 1$ (on cherche une distribution de probabilité). Ce système se résout immédiatement, et l'on trouve :

$$\pi(k) = \begin{cases} \frac{1}{2^k} & \text{si } k \in \{1, \ldots, n-1\}; \\ \frac{1}{2^{n-1}} & \text{si } k = n. \end{cases}$$

3. Il faut traverser $n-1$ arêtes pour aller de 1 à n donc $\textsc{diam}(P) = n-1$. Attention cependant : la minoration $t_{\text{MIX}}(P, \varepsilon) \geq \textsc{diam}(P)/2$ vue en cours ne s'applique pas ici, car le support de P n'est pas symétrique ! Pour ce qui est du degré, il y a exactement deux destinations possibles depui chaque état, donc $\Delta(P) = 2$. Comme $\max(\pi) = 1/2$, on en déduit que

$$t_{\text{MIX}}(P, \varepsilon) \geq \left\lceil \frac{\log(2 - 2\varepsilon)}{\log 2} \right\rceil = 1,$$

pour tout $\varepsilon \in (0, 1)$, ce qui n'est pas très informatif...

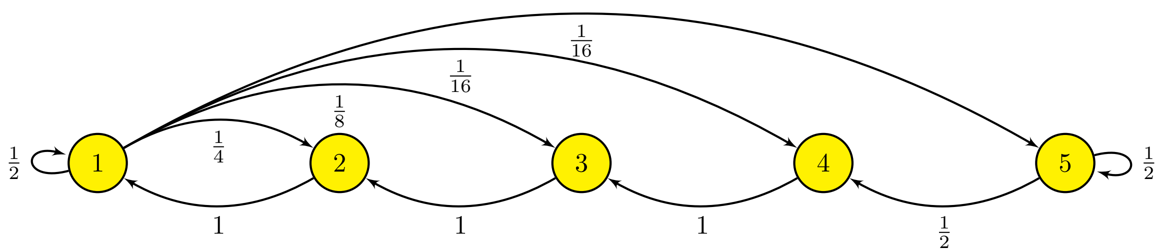
4. Considérons une partie non-vide $A \subseteq \{1, \ldots, n\}$ avec $\pi(A) \le \frac{1}{2}$. On distingue deux cas. Si A contient l'état 1, alors on a forcément $A = \{1\}$ car $\pi(1) = 1/2$. On a alors
$$\Phi(A) \quad = \quad \frac{\vec{\pi}(A \times A^c)}{\pi(A)} \quad = \quad \frac{\pi(1)P(1,2)}{\pi(1)} \quad = \quad P(1,2) \quad = \quad \frac{1}{2}.$$
Si au contraire A ne contient pas 1, alors le fait que $P(x,1) = 1/2$ pour tout x entraîne :
$$\vec{\pi}(A \times A^c) \quad = \quad \sum_{x \in A} \sum_{y \in A^c} \pi(x)P(x,y) \;\; \ge \;\; \sum_{x \in A} \pi(x)P(x,1) \quad = \quad \frac{\pi(A)}{2},$$
c'est-à-dire $\Phi(A) \ge 1/2$. On déduit de ces deux calculs que la conductance de la chaîne est
$$\Phi_\star(P) \quad = \quad \frac{1}{2}.$$
L'inégalité de Cheeger entraîne que $1/8 \le \gamma_\star(P) \le 1$, mais on ne peut pas en tirer une majoration de $t_{\text{MIX}}(P, \varepsilon)$ car le noyau P n'est pas paresseux ($P(2,2) = 0$). On peut en revanche toujours invoquer la minoration
$$t_{\text{MIX}}(P, \varepsilon) \quad \ge \quad \left\lceil \frac{1 - 2\varepsilon}{2\Phi_\star(P)} \right\rceil \;\; = \;\; 1,$$
pour $\varepsilon < \frac{1}{2}$. Hélas, cela n'est guère plus informatif que le résultat obtenu à la question 3.

5. La manière la plus simple de construire une chaîne de noyau P issue d'un état $x \in \{1, \ldots, n\}$ consiste à poser $X_0 = x$ puis inductivement, pour tout $t \ge 1$,
$$X_t \quad = \quad \begin{cases} \min(X_{t-1} + 1, n) & \text{si } B_t = 1 \\ 1 & \text{si } B_t = 0. \end{cases}$$
où $(B_t)_{t \ge 1}$ est une suite de variables de Bernoulli indépendantes de paramètre $1/2$. Pour coupler deux trajectoires $(X_t)_{t \ge 0}$ et $(Y_t)_{t \ge 0}$ issus d'états différents, on peut alors simplement utiliser les mêmes variables de mises à jour $(B_t)_{t \ge 1}$ dans les deux cas. Formellement, cela correspond au noyau de couplage
$$Q\left((x,y),(x',y')\right) \quad = \quad \begin{cases} \frac{1}{2} & \text{si } (x',y') = (1,1) \text{ ou si } (x',y') = (\max(x+1,n), \max(y+1,n)) \\ 0 & \text{sinon} \end{cases}$$
Pour que la coalescence se produise, sous ce couplage, il suffit que les chaînes soient envoyées sur l'état 1, c'est-à-dire que la variable de Bernoulli utilisée pour la mise à jour soit égale à 0. Ainsi, avec les notations du cours, on a
$$\mathcal{D}(t) \;\; \le \;\; \max_{x,y} \mathbb{P}_{x,y}(T > t) \quad \le \quad \frac{1}{2^t},$$
pour tout $t \ge 0$. Pour rendre le membre droit plus petit qu'une précision donnée $\varepsilon \in (0,1)$, il suffit de prendre $t \ge \log_2(1/\varepsilon)$. En conclusion, on obtient bien
$$t_{\text{MIX}}(P, \varepsilon) \quad \le \quad \left\lceil \log_2\left(\frac{1}{\varepsilon}\right) \right\rceil.$$

6. En utilisant la question 2 et la définition du noyau adjoint, on trouve

$$P^\star(x,y) \quad = \quad \frac{\pi(y)P(y,x)}{\pi(x)} \quad = \quad \begin{cases} 1 & \text{si } x \neq n \text{ et } y = x-1; \\ \frac{1}{2} & \text{si } x = n \text{ et } y \in \{n-1, n\}; \\ \frac{1}{2^y} & \text{si } x = 1 \text{ et } y \neq n; \\ \frac{1}{2^{n-1}} & \text{si } x = 1 \text{ et } y = n. \end{cases}$$

Dans le cas particulier $n = 5$, le graphe induit est le suivant :



7. Par définition de la distance à l'équilibre $\mathfrak{D}_{P_\star}(t)$, on a pour tous $t \geq 0$, $x \in S$ et $A \subseteq S$,

$$\mathfrak{D}_{P^\star}(t) \quad \geq \quad \pi(A) - P_\star^t(x, A).$$

En prenant $t = n-2$, $x = n$ et $A = \{1\}$, on obtient $\mathfrak{D}_{P^\star}(n-2) \geq \frac{1}{2}$, ce qui montre que $t_{\text{MIX}}(P^\star, \varepsilon) \geq n-1$ dès que $\varepsilon < 1/2$.

8. Plus généralement, fixons $k \in \{1, \ldots, n-1\}$ et prenons $t = n-1-k$, $x = n$ et $A = \{1, \ldots, k\}$ dans l'inégalité ci-dessus. On obtient alors

$$\mathfrak{D}_{P^\star}(n-1-k) \quad \geq \quad \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^k} \;\; = \;\; 1 - \frac{1}{2^k}.$$

Si $k = 1 + \lfloor \log_2(\frac{1}{1-\varepsilon}) \rfloor$, le membre droit est strictement supérieur à ε, donc $t_{\text{MIX}}(P^\star, \varepsilon) \geq n-k$.

9. En remarquant que $P(x,1) = \frac{1}{2} = \pi(1)$ pour tout $x \in S$, on a

$$P^\star(1,x) \quad = \quad \frac{\pi(x)P(x,1)}{\pi(1)} \quad = \quad \pi(x).$$

Autrement dit, pour la chaîne de noyau $P^\star$ issue de 1, la loi au temps 1 est exactement π.

10. Nous allons utiliser la question précédente pour montrer que le temps de mélange est au plus le temps d'atteinte de l'état 1, plus 1. Considérons une chaîne $(X_t)_{t\geq 0}$ de noyau $P^\star$, et notons T_1 le temps d'atteinte de l'état 1. Pour tout état initial $x \in S$ et tout ensemble $A \subseteq S$, on a

$$\begin{aligned} \mathbb{P}_x(X_t \in A, T_1 < t) \quad &= \quad \sum_{s=0}^{t-1} \mathbb{P}_x(T_1 = s)\mathbb{P}_1(X_{t-s} \in A) \\ &= \quad \mathbb{P}_x(T_1 < t)\pi(A), \end{aligned}$$

où la première égalité découle de la propriété de Markov, et la seconde de la question 9. Comme on a par ailleurs $\mathbb{P}_x(X_t \in A, T_1 < t) \le \mathbb{P}_x(X_t \in A) = P_\star^t(x, A)$, on conclut que

$$\pi(A) - P_\star^t(x, A) \quad \le \quad \pi(A) - \pi(A)\mathbb{P}_x(T_1 < t) \quad = \quad \pi(A)\mathbb{P}_x(T_1 \ge t).$$

Comme $\mathfrak{D}_{P^\star}(t) = \max\{\pi(A) - P_\star^t(x, A) \colon x \in S, A \subseteq S\}$, on en déduit que

$$\mathfrak{D}_{P^\star}(t) \quad \le \quad \max_{x \in S} \mathbb{P}_x(T_1 \ge t) \quad = \quad \mathbb{P}_n(T_1 \ge t),$$

où l'on a remarqué que sous le noyau dual $P^\star$, le seul moyen d'atteindre 1 depuis x est de passer par $x-1, x_2 \ldots, 1$. Mais, partant de l'état n, la chaîne reste sur l'état n pendant un temps géométrique de paramètre $1/2$, puis descend de manière déterministe vers 1. Autrement dit, $T_1 = G + n - 1$ avec $G \sim \mathcal{G}(\frac{1}{2})$ et donc

$$\mathfrak{D}_{P^\star}(t) \quad \le \quad \mathbb{P}\left(G \ge t + 1 - n\right) \quad = \quad \frac{1}{2^{(t+1-n)_+}}.$$

Pour rendre le membre droit plus petit que $\varepsilon \in (0,1)$, il suffit de prendre $t \ge n - 1 + \log_2(\frac{1}{\varepsilon})$:

$$t_{\text{MIX}}(P^\star, \varepsilon) \quad \le \quad n - 1 + \left\lceil \log_2\left(\frac{1}{\varepsilon}\right) \right\rceil.$$

En combinant ce résultat avec celui de la question 8, on obtient bien que $t_{\text{MIX}}(P^\star, \varepsilon) = n + o(n)$ lorsque $n \to \infty$, indépendamment de la précision $\varepsilon \in (0,1)$ choisie (cutoff).

Problème 2

1. Pour chaque $0 \le s \le t$, on se donne un état $x(s) = (x_1(s), \ldots, x_n(s)) \in S^n$. On a alors

$$\begin{aligned} \mathbb{P}\left(X(0) = x(0), \ldots, X(t) = x(t)\right) \quad &= \quad \mathbf{1}_{(x(0)=x)} \prod_{s=1}^{t} P_n\left(x(s-1), x(s)\right) \\ &= \quad \prod_{i=1}^{n} \mathbf{1}_{(x_i(0)=x_i)} \prod_{s=1}^{t} P\left(x_i(s-1), x_i(s)\right). \end{aligned}$$

   Cette forme produit montre que les processus $\left(X_1(t) \colon t \ge 0\right), \ldots, \left(X_n(t) \colon t \ge 0\right)$ sont indépendants, et que ce sont des chaînes de Markov de noyau P issues des états $x_1, \ldots, x_n$.

2. Par l'indépendance ci-dessus, on a pour tout $(x, y) \in S^n \times S^n$,

$$P_n^t(x, y) \quad = \quad \prod_{i=1}^{n} P^t(x_i, y_i). \tag{1}$$

P étant ergodique, il existe un entier t tel que toutes les entrées de P^t sont strictement positives. La formule ci-dessus montre que pour ce même entier t, toutes les entrées de P_n^t sont strictement positives. Ainsi, P_n est ergodique, et il est clair que la loi produit

$$\pi_n(x) \quad := \quad \pi(x_1) \times \cdots \times \pi(x_n) \tag{2}$$

vérifie l'équation de réversibilité $\pi_n(x)P_n(x,y) = \pi_n(y)P_n(y,x)$.

(3a) D'après la question 1, les variables $X_1(t), \ldots, X_n(t)$ sont i.i.d. de loi $P^t(z,\cdot)$. Ainsi,

$$\begin{aligned}
\mathbb{E}\left[\phi\left(X(t)\right)\right] &= n(P^t f)(z) \;=\; n\lambda^t f(z) \;=\; n\lambda^t \\
\mathrm{Var}\left[\phi\left(X(t)\right)\right] &= n\mathrm{Var}\left(f(X_1(t))\right) \;\le\; \|f\|_\infty = 1.
\end{aligned}$$

On conclut en invoquant l'inégalité de Bienaymé Chebychev pour la variable $\phi\left(X(t)\right)$. L'hypothèse que t est pair assure que $\lambda^t > 0$.

(3b) Notons que $\pi f = 0$ car f est une fonction propre non-triviale. La forme produit (2) montre que $Y_1, \ldots, Y_n$ sont i.i.d., de loi π. On en déduit que

$$\begin{aligned}
\mathbb{E}\left[\phi\left(Y\right)\right] &= n\pi f \;=\; 0; \\
\mathrm{Var}\left[\phi\left(Y\right)\right] &= n\mathrm{Var} f \;\le\; \|f\|_\infty \;=\; 1.
\end{aligned}$$

On conclut en invoquant l'inégalité de Bienaymé Chebychev pour la variable $\phi\left(Y\right)$.

(3c) Posons $A = \left\{x \in S^n \colon \phi(x) \le \frac{n}{2}\lambda^t\right\}$. Les questions précédentes montrent que

$$\begin{aligned}
\mathfrak{D}_n(t) &\ge \pi_n(A) - P_n^t((z,\ldots,z),A) \\
&\ge 1 - \frac{8}{n\lambda^{2t}}.
\end{aligned}$$

Le membre droit tend vers 1 si $n \to \infty$ et $t \sim \alpha \ln n$ avec $\alpha < \left(\ln(\frac{1}{\lambda^2})\right)^{-1}$. Or $\left(\ln(\frac{1}{\lambda^2})\right)^{-1} = \frac{\tau}{2}$ si l'on choisit une valeur propre λ telle que $|\lambda| = \lambda_\star$. Notons que l'argument ne s'applique pas si $\lambda_\star = 0$, mais dans ce cas on a $\tau = 0$ et la minoration cherchée est triviale.

(4a) L'inégalité de Cauchy-Schwarz montre que

$$d_{\mathrm{TV}}(\mu,\pi) \quad = \quad \frac{1}{2}\sum_x \pi(x)\left|\frac{\mu(x)}{\pi(x)} - 1\right| \;\le\; \frac{1}{2}\sqrt{\chi(\mu,\pi)}.$$

(4b) En développant le carré dans la définition de χ, on voit que

$$1 + \chi(\mu,\pi) \quad = \quad \sum_x \frac{\left(\mu(x)\right)^2}{\pi(x)} \;=\; \mathbb{E}\left[\frac{\mu(X)}{\pi(X)}\right],$$

où X désigne une variable de loi μ. Appliquons cela à $\mu = P_n^t(x, \cdot)$ et $\pi = \pi_n$:

$$\begin{aligned}
1 + \chi\left(P_n^t(x, \cdot), \pi_n\right) &= \mathbb{E}\left[\frac{P_n^t(x, X)}{\pi_n(X)}\right] \\
&= \mathbb{E}\left[\prod_{i=1}^n \frac{P^t(x_i, X_i)}{\pi(X_i)}\right] \\
&= \prod_{i=1}^n \mathbb{E}\left[\frac{P^t(x_i, X_i)}{\pi(X_i)}\right] \\
&= \prod_{i=1}^n \left(1 + \chi(P^t(x_i, \cdot), \pi\right),
\end{aligned}$$

où l'on a utilisé les formes produit (1)-(2) et le fait que les variables $X_1, \ldots, X_n$ sont indépendantes et de lois respectives $P^t(x_1, \cdot), \ldots, P^t(x_n, \cdot)$ (question 1).

(4c) On a vu en cours que la décomposition spectrale d'un noyau auto-adjoint P entraîne

$$\chi\left(P^t(x, \cdot), \pi\right) \leq \frac{\lambda_\star^{2t}}{\pi_\star},$$

pour tout $x \in S$. Le résultat demandé découle alors de l'identité précédente.

3. On déduit des questions (4a) et (4c) que

$$\mathfrak{D}_n(t) \leq \frac{1}{2}\sqrt{\left(1 + \frac{\lambda_\star^{2t}}{\pi_\star}\right)^n - 1}.$$

Comme $(1 + o\left(\frac{1}{n}\right))^n \to 1$ lorsque $n \to \infty$, il suffit de prendre $t \sim \alpha \ln n$ avec $\alpha > \frac{\tau}{2}$, pour que $\mathfrak{D}_n(t) \to 0$, ce qui fournit la majoration demandée.