Corrigé du texte sur les probabilités de retour

1. Comme G est un graphe connexe, si v et w sont deux sommets de G, alors soit $v = w$, auquel cas $Q(v,w) \geq \frac{1}{2} > 0$; soit il existe un chemin $v_0 = v, v_1, \ldots, v_r = w$ avec chaque $\{v_{i-1}, v_i\}$ arête de G. Dans ce cas,

$$Q^r(v,w) \geq \frac{1}{2^r}\, P(v_0,v_1)\, P(v_1,v_2) \,\cdots\, P(v_{r-1},v_r) \geq \frac{1}{(4|E|)^r}\, \deg v_0\, \deg v_1 \,\cdots\, \deg v_{r-1} > 0.$$

La chaîne de Markov de matrice de transition Q est donc irréductible, et comme par ailleurs $Q(v,v) > 0$ pour tout $v \in V$, elle est apériodique. Elle admet donc une unique mesure de probabilité invariante. C'est la mesure $\pi(v) = \frac{\deg v}{2|E|}$. En effet, cette quantité est bien une mesure de probabilité sur V, car par le lemme des poignées de mains,

$$\sum_{v \in V} \pi(v) = \frac{1}{2|E|} \sum_{v \in V} \deg v = \frac{2|E|}{2|E|} = 1.$$

De plus, π est réversible par rapport à P, car

$$\pi(v)\, P(v,w) = \frac{\deg v}{2|E|}\, \frac{1_{v\sim w}}{\deg v} = \frac{1_{v\sim w}}{2|E|} = \pi(w)\, P(w,v).$$

On en déduit que π est aussi réversible par rapport à Q :

$$\pi(v)\, Q(v,w) = \frac{\pi(v)\, P(v,w) + \pi(v)\, 1_{v=w}}{2} = \frac{\pi(w)\, P(w,v) + \pi(w)\, 1_{v=w}}{2} = \pi(w)\, Q(w,v).$$

La réversibilité implique l'invariance, donc π est bien la mesure stationnaire de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ de matrice Q. Par le théorème ergodique, $(X_n)_{n\in\mathbb{N}}$ converge en loi vers π pour toute mesure initiale sur X_0, donc, en particulier, si $X_0 = v$ presque sûrement, alors

$$Q^n(v,w) = \lim_{n\to\infty} \mathbb{P}[X_n = w] = \pi(w).$$

2. Pour encoder le graphe G, on peut utiliser un dictionnaire dont les clés sont les sommets $v \in V$, et dont les entrées sont les ensembles de voisins, donnés sous forme de listes. On pose donc :

```
G = {(0,0):[(0,1),(1,0)],
        (0,1):[(0,0),(0,2),(1,1)],
        (0,2):[(0,1),(1,2)],
        (1,0):[(0,0),(1,1),(2,0)],
        (1,1):[(0,1),(1,0),(1,2),(2,1)],
        (1,2):[(0,2),(1,1),(2,2)],
        (2,0):[(1,0),(2,1)],
        (2,1):[(1,1),(2,0),(2,2)],
        (2,2):[(1,2),(2,1)] }

V = G.keys()

def degr(v):
        return len(G[v])
```

La mesure invariante π et la matrice de transition Q sont données par les programmes :

```
def Q(v,w):
        if v==w:
                return float(1)/2
        else:
                if w in G[v]:
                        return float(1)/(2*degr(v))
                else:
                        return float(0)

def pi(v):
        return float(degr(v))/24
```

et on obtient les valeurs suivantes pour π :

$$
\begin{array}{ccccc}
\frac{1}{12} & \text{———} & \frac{1}{8} & \text{———} & \frac{1}{12} \\
| & & | & & | \\
\frac{1}{8} & \text{———} & \frac{1}{6} & \text{———} & \frac{1}{8} \\
| & & | & & | \\
\frac{1}{12} & \text{———} & \frac{1}{8} & \text{———} & \frac{1}{12}
\end{array}
$$

On simule la marche aléatoire avec le programme suivant (on utilise le minimum de fonctions prédéfinies) :

```
import random
alea = random.random

def MarcheAleatoire(n,v):
        pos = v
        res = [pos]
        for i in range(0,n):
                if alea()<0.5:
                        voisins = G[pos]
                        k = len(voisins)
                        pos = voisins[int(k*alea())]
                res.append(pos)
        return res
```

Par exemple, `MarcheAleatoire(30,(0,0))` renvoie la liste :

```
[(0, 0),(0, 0),(0, 0),(0, 0),(0, 0),(1, 0),(1, 0),(1, 1),(1, 1),(1, 1),(0, 1),
 (0, 1),(0, 1),(0, 0),(0, 0),(0, 0),(1, 0),(1, 0),(0, 0),(0, 1),(0, 2),(1, 2),
 (2, 2),(2, 2),(2, 2),(2, 2),(2, 1),(2, 2),(2, 2),(2, 2),(1, 2)]
```

Pour vérifier la convergence en loi vers π, on a besoin d'un N-échantillon de la variable X_n, avec n suffisamment grand. On peut par exemple utiliser :

```
def MesureEmpirique(N,n,v):
        res = {}
```

```
        for w in V:
            res[w] = float(0)
        for i in range(0,N):
            pos = MarcheAleatoire(n,v)[n]
            res[pos] += float(1)/N
        return res
```

Avec $v = (0, 0)$, $n = 1000$ et $N = 10000$, on obtient

```
{(0, 0): 0.07980000000000126,
 (0, 1): 0.1282000000000022,
 (0, 2): 0.08420000000000138,
 (1, 0): 0.1273000000000023,
 (1, 1): 0.1680999999999978,
 (1, 2): 0.12000000000000241,
 (2, 0): 0.08250000000000134,
 (2, 1): 0.12410000000000253,
 (2, 2): 0.08580000000000143}
```

avec des valeurs très proches de $\frac{1}{6} = 0.16666$ pour le centre du carré, $\frac{1}{8} = 0.125$ pour les milieux des côtés du carré, et $\frac{1}{12} = 0.083333$ pour les coins du carré.

3. Fixons S et $x \in (0, 1]$. Si $w \in f(S, x)$, alors par définition $x\,\pi(w) \leq \sum_{v\in S} \pi(v)\,Q(v, w)$. Supposons par l'absurde que w n'est pas dans
$$S_+ = S \cup \{\text{voisins d'éléments de } S\}.$$
Alors, $Q(v, w) = 0$ pour tout $v \in S$, donc le terme de droite de l'inégalité précédente vaut 0. C'est impossible, car $x\,\pi(w) > 0$. Donc, $f(S, x) \subset S_+$.

4. On vérifie d'abord que les états $\emptyset$ et V sont absorbants.
   - Pour l'état vide, on a pour tout $x \in (0, 1]$ $f(\emptyset, x) \subset \emptyset_+ = \emptyset$, donc $f(\emptyset, x) = \emptyset$ pour tout $x \in (0, 1]$. Ainsi, si $S_n = \emptyset$, alors $S_{n+1} = \emptyset$ presque sûrement, ce qui implique que $\emptyset$ est absorbant.
   - Pour l'état plein, on a $\sum_{v\in V} \pi(v)Q(v, w) = (\pi Q)(w) = \pi(w) \geq x\,\pi(w)$ pour tout $x \in [0, 1]$ et tout $w \in V$, donc $f(V, x) = V$ pour tout $x \in [0, 1]$. Ceci implique comme précédemment que V est absorbant.

   En particulier, ces deux états sont récurrents. Montrons maintenant que tout état $S \neq \emptyset$ est connecté par la chaîne des ensembles évoluants à la partie pleine V. Il suffit de montrer que si $S \neq \emptyset$, alors il y a une probabilité de transition positive de S vers l'ensemble S_+ défini à la question précédente. Comme G est un graphe connexe, ceci impliquera qu'avec une probabilité positive, on peut en un nombre fini d'étapes recouvrir tout V en suivant la chaîne des ensembles évoluants issue de S. Nous allons montrer que si x est assez petit et $w \in S_+$, alors $\sum_{v\in S} \pi(v)\,Q(v, w) \geq x\,\pi(w)$. Comme S_+ est un ensemble fini, on pourra alors choisir un $\varepsilon > 0$ commun à tous les $w \in S_+$ et tel que $w \in f(S, x)$ pour tout $x \in (0, \varepsilon)$ et tout $w \in S_+$. Combiné au résultat de la question précédente, ceci impliquera $f(S, x) = S_+$ pour tout $x \in (0, \varepsilon)$.

   Si $w \in S$, alors $\sum_{v\in S} \pi(v)\,Q(v, w) \geq \pi(w)\,Q(w, w) = \frac{\pi(w)}{2}$, donc il suffit de prendre $x \leq \frac{1}{2}$. Supposons maintenant $w \in S_+ \setminus S$; on choisit $v_0 \in S$ tel que v_0 et w soit voisins. On a alors
$$\sum_{v\in S} \pi(v)\,Q(v, w) = \sum_{v\in S} \pi(v)\,\frac{P(v, w)}{2} \geq \frac{\pi(v_0)\,P(v_0, w)}{2} = \frac{\pi(w)\,P(w, v_0)}{2} \geq \frac{\pi(w)}{2\deg G},$$

donc il suffit de prendre $x \leq \frac{1}{2 \deg G}$. On conclut que le résultat précédemment annoncé est bien vrai (avec $\varepsilon = \frac{1}{2 \deg G}$). Il implique que tous les états différents de $\emptyset$ et V sont transients.

5. Dans tout ce qui suit, les parties $S \subset V$ sont données sous forme de liste. Il suffit essentiellement de programmer la fonction $f(S, x)$:

```
def f(S,x):
        res = []
        for w in V:
                if x*pi(w) <= sum([pi(v)*Q(v,w) for v in S]):
                        res.append(w)
        return res
```

Alors, la chaîne de Markov des ensembles évoluants est donnée par :

```
def EnsemblesEvoluants(n,S):
        pos = S
        res = [pos]
        for i in range(0,n):
                pos = f(pos,alea())
                res.append(pos)
        return res
```

Par exemple, `EnsemblesEvoluants(20,[(1, 1), (1, 0), (0, 1)])` donne :

```
[[(1, 1), (1, 0), (0, 1)],
 [(0, 1), (1, 0), (1, 1)],
 [(0, 1), (0, 0), (1, 0), (1, 1)],
 [(0, 1), (0, 0), (1, 0), (1, 1)],
 [(0, 1), (0, 0), (1, 0), (1, 1)],
 [(0, 1), (0, 0), (2, 0), (1, 0), (0, 2), (1, 1)],
 [(0, 1), (0, 0), (1, 0)],
 [(0, 1), (0, 0), (2, 0), (1, 0), (0, 2), (1, 1)],
 [(0, 1), (0, 0), (2, 0), (1, 0), (0, 2), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)],
 [(0, 1), (1, 2), (0, 0), (2, 0), (1, 0), (2, 2), (0, 2), (2, 1), (1, 1)]]
```

qui est stationnaire à l'ensemble plein. Pour obtenir S_∞, on peut modifier le programme précédent et l'arrêter dès qu'on atteint $\emptyset$ ou V :

```
def Sinfinity(S):
        pos = S
```

```
        while not (len(pos) in [0,9]):
                pos = f(pos,alea())
        return pos
```

Avec $S = [(1,1),(1,0),(0,1)]$ qui a mesure $\pi(S) = \frac{5}{12} = 0.41666$, on obtient sur $N = 1000$ essais 418 fois l'ensemble plein, ce qui confirme empiriquement le résultat annoncé dans le texte pour la distribution de S_∞.

6. Pour toute partie S et tout w, on a

$$0 \leq \frac{\sum_{v\in S}\pi(v)\,Q(v,w)}{\pi(w)} \leq \frac{\sum_{v\in V}\pi(v)\,Q(v,w)}{\pi(w)} = \frac{(\pi Q)(w)}{\pi(w)} = \frac{\pi(w)}{\pi(w)} = 1$$

car π est Q-invariante.

7. Avec probabilité 1, S_n n'est plus dans $\mathfrak{P}(V) \setminus \{\emptyset, V\}$ pour n assez grand, car c'est un ensemble fini d'états transients. Or, les autres états sont absorbants, donc dès que S_n quitte cet ensemble, il devient constant à $S_\infty \in \{\emptyset, V\}$. Pour déterminer la loi de la variable S_∞, notons que comme $\pi(S_n) = M_n$ est une martingale bornée, elle converge p.s. et sa valeur limite $M_\infty = \pi(S_\infty)$ a la même espérance que toutes les variables M_n. Ainsi,

$$\mathbb{P}[S_\infty = V] = \mathbb{E}[\pi(S_\infty)] = \mathbb{E}[M_\infty] = \mathbb{E}[M_0] = \pi(S_0).$$

8. Comme $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ est un espace de Hilbert, toute forme linéaire continue ϕ sur cet espace s'écrit de manière unique $\phi(Y) = \mathbb{E}[TY]$ avec $T \in \mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$. C'est en particulier vrai pour $\phi(Y) = \mathbb{E}[XY]$, qui est linéaire continue puisque $|\phi(Y)| \leq \|X\|_{\mathrm{L}^2(\Omega,\mathcal{F},\mathbb{P})}\,\|Y\|_{\mathrm{L}^2(\Omega,\mathcal{G},\mathbb{P})}$. Si T est la fonction $\mathcal{G}$-mesurable de carré intégrable associée à cette forme ϕ, alors par définition de l'espérance conditionnelle, pour toute variable Y qui est $\mathcal{G}$-mesurable bornée, on a

$$\mathbb{E}[TY] = \mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[X \,|\, \mathcal{G}]\,Y].$$

En prenant $Y = \mathrm{sgn}(T - \mathbb{E}[X \,|\, \mathcal{G}])$, on en déduit que $\mathbb{E}[|T - \mathbb{E}[X \,|\, \mathcal{G}]|] = 0$, et donc que $\mathbb{E}[X \,|\, \mathcal{G}] = T$ est de carré intégrable. Alors, l'identité

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[X \,|\, \mathcal{G}]\,Y]$$

valable pour toute Y variable $\mathcal{G}$-mesurable bornée s'étend par densité à toutes les variables $Y \in \mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$. Cette identité caractérise la projection orthogonale de $\mathrm{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ vers $\mathrm{L}^2(\Omega, \mathcal{G}, \mathbb{P})$.

On calcule maintenant

$$\begin{aligned}\mathrm{var}(\mathbb{E}[X \,|\, \mathcal{G}]) &= \mathbb{E}[(\mathbb{E}[X \,|\, \mathcal{G}])^2] - \left(\mathbb{E}[\mathbb{E}[X \,|\, \mathcal{G}]]\right)^2 \\ &\leq \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathrm{var}(X)\end{aligned}$$

puisque la projection orthogonale sur un sous-espace réduit les distances, et puisque $\mathbb{E}[\mathbb{E}[X \,|\, \mathcal{G}]] = \mathbb{E}[X]$.

9. Pour la première identité à établir, on écrit :

$$\begin{aligned}\mathbb{E}_S\left[\pi(S_1)\,|\,U_0 \leq \frac{1}{2}\right] &= \sum_{w\in V} \pi(w)\,\mathbb{E}_S\left[1_{(w\in S_1)}\,|\,U_0 \leq \frac{1}{2}\right] \\ &= \sum_{w\in S}\pi(w) + \sum_{w\notin S}\left(2\sum_{v\in S}\pi(v)\,Q(v,w)\right) = \pi(S) + 2\sum_{\substack{v\in S\\ w\notin S}}\pi(v)\,Q(v,w).\end{aligned}$$

En effet, $w \in S_1$ si et seulement si $\sum_{v\in S}\pi(v)\,Q(v,w) \geq U_0\,\pi(w)$, et on s'est ramené aux identités précédant dans le texte celles que l'on souhaite établir. Pour la seconde identité, on a de même :

$$\begin{aligned}\mathbb{E}_S\left[\pi(S_1)\,|\,U_0 \geq \frac{1}{2}\right] &= \sum_{w\in V} \pi(w)\,\mathbb{E}_S\left[1_{(w\in S_1)}\,|\,U_0 \geq \frac{1}{2}\right] \\ &= \sum_{w\in S}\left(2\sum_{v\in S}\pi(v)\,Q(v,w) - \pi(w)\right) = 2\sum_{v,w\in S}\pi(v)\,Q(v,w) - \pi(S).\end{aligned}$$

On transforme la dernière quantité en rajoutant et en elevant la somme sur $w \notin S$:

$$\begin{aligned}2\sum_{v,w\in S}\pi(v)\,Q(v,w) - \pi(S) &= 2\sum_{v\in S}\pi(v) - 2\sum_{\substack{v\in S\\ w\notin S}}\pi(v)\,Q(v,w) - \pi(S) \\ &= \pi(S) - 2\sum_{\substack{v\in S\\ w\notin S}}\pi(v)\,Q(v,w).\end{aligned}$$

10. Comme $M_{n+1} = \pi(f(S_n,U_n))$, il suffit de montrer que $\pi(f(S,x)) \leq (1+\deg G)\pi(S)$ si $x \in (0,1]$, et par la propriété $f(S,x) \subset S_+$, il suffit d'évaluer $\pi(S_+)$. On a

$$\pi(S_+) = \pi(S) + \sum_{\substack{w\notin S\\ w\sim v \text{ avec } v\in S}} \pi(w) \leq \pi(S) + \sum_{v\in S}\sum_{w\sim v}\pi(w).$$

Si $w \sim v$, alors $P(w,v) = \frac{1}{\deg w}$ et $(\deg G)\,P(w,v) \geq 1$. Ainsi,

$$\begin{aligned}\pi(S_+) &\leq \pi(S) + (\deg G)\sum_{v\in S}\sum_{w\sim v}\pi(w)\,P(w,v) \\ &= \pi(S) + (\deg G)\sum_{v\in S}\sum_{w\sim v}\pi(v)\,P(v,w) = (1+\deg G)\pi(S)\end{aligned}$$

en utilisant la réversibilité de π par rapport à P.

11. Le programme suivant simule la variable τ_v :

```
def tau(v):
        pos = [v]
        count = 0
        while not len(pos) in [0,9]:
                pos = f(pos,alea())
                count += 1
        return count
```

Sur 100000 essais, la distribution de $\tau_{(1,1)}$ est la suivante :

```
{1: 50096,
 2: 24883,
 3: 12471,
 4: 6307,
 5: 3132,
 6: 1544,
 7: 809,
 8: 384,
 9: 178,
 10: 100,
 11: 48,
 12: 29,
 13: 13,
 14: 4,
 15: 2}
```

On semble ici très proche d'une distribution géométrique de paramètre $1/2$.

12. De façon générale, pour une chaîne de Markov irréductible apériodique sur un ensemble fini, on sait que la convergence vers la loi stationnaire se fait exponentiellement vite : il existe $\lambda < 1$ et une constante K dépendant de l'espace telle que $|Q^n(v,v) - \pi(v)| \leq K\,\lambda^n$. L'inégalité du texte n'est donc pas du tout optimale en n pour un graphe G fixé, mais son intérêt est quelle est valide avec une constante explicite, et qui est commune à tous les graphes si l'on se restreint aux graphes de degré maximal borné par une constante D.