Processus aléatoires Thomas Budzinski
ENS Paris, 2016-2017 Bureau V2
thomas.budzinski@ens.fr

TD 8 : Encore des martingales
Corrigé

Lundi 21 Novembre

Exercice 1 (Modèle de Wright-Fisher)
Un gène a deux allèles α et β. On considère une population qui se renouvelle entièrement à chaque génération, le nombre d'individus restant fixe et égal à $n \geq 2$. On suppose que pour tout $k \geq 0$, à la génération $k+1$, chacun des n individus choisit son parent uniformément parmi les n individus de la génération k, indépendamment les uns des autres. On note X_k le nombre d'individus portant l'allèle α à la génération k, et $\mathscr{F}_k = \sigma(X_0, \dots, X_k)$. On suppose $X_0 = a \in [\![1, n-1]\!]$.

1. Quelle est la loi de X_{k+1} conditionnellement à $\mathscr{F}_k$? En déduire que X est une martingale pour $(\mathscr{F}_k)_{k \geq 0}$.
2. Montrer que X converge p.s. vers une variable X_∞, et donner sa loi.
   On pose $\tau = \inf\{k | X_k = X_\infty\}$.
3. Calculer $\mathbb{E}\left[X_{k+1}(n - X_{k+1}) | \mathscr{F}_k\right]$ et trouver une martingale.
4. En déduire un encadrement de $\mathbb{E}[\tau]$. On pourra par exemple montrer

$$\mathbb{E}[\tau] = O(n \ln n).$$

<u>*Solution de l'exercice 1*</u>

1. Conditionnellement à $\mathscr{F}_k$, chaque individu de la génération porte l'allèle α avec probabilité $\frac{X_k}{n}$, indépendamment les uns des autres, donc X_{k+1} suit une loi binomiale de paramètres n et $\frac{X_k}{n}$. On en déduit $\mathbb{E}[X_{k+1}|\mathscr{F}_k] = n \times \frac{X_k}{n} = X_k$. De plus, le processus X est bien intégrable et adapté, donc X est bien une martingale.
2. La martingale X est bornée par n, donc elle converge p.s. et dans L^1 vers une variable X_∞. On va maintenant montrer que $X_\infty \in \{0, n\}$ p.s. Soit $j \in [\![1, n-1]\!]$. Pour tout k, on pose

$$q = \mathbb{P}\left(X_{k+1} = j | X_k = j\right) = \binom{n}{j} \left(\frac{j}{n}\right)^j \left(1 - \frac{j}{n}\right)^{n-j} < 1.$$

   On a pour tout $\ell \geq 0$:

$$\mathbb{P}\left(X_k = X_{k+1} = \dots = X_{k+\ell} = j\right) = \mathbb{P}\left(X_k = j\right) \prod_{i=0}^{\ell-1} \mathbb{P}\left(X_{k+i+1=j} | X_{k+i=j}\right) \leq q^\ell \xrightarrow[\ell \to +\infty]{} 0.$$

   On en déduit $\mathbb{P}\left(\forall i \geq k, X_i = j\right) = 0$ pour tous $j \in [\![1, n-1]\!]$ et $k \geq 0$, donc $\mathbb{P}\left(X_\infty = j\right) = 0$ (comme X est à valeurs entières et converge, elle est constante à partir d'un certain rang). On a donc $X_\infty \in \{0, n\}$ p.s. De plus, par convergence L^1 on a $\mathbb{E}\left[X_\infty\right] = \mathbb{E}[X_0] = a$, donc $\mathbb{P}\left(X_\infty = 0\right) = 1 - \frac{a}{n}$ et $\mathbb{P}\left(X_\infty = n\right) = \frac{a}{n}$.

3. On sait que l'espérance et la variance d'une variable de loi binomiale de paramètres n et p valent respectivement pn et $p(1-p)n$. On obtient donc

$$\begin{aligned}\mathbb{E}[X_{k+1}(n-X_{k+1})|\mathscr{F}_k] &= n\mathbb{E}[X_{k+1}|\mathscr{F}_k] - \mathbb{E}[X_{k+1}|\mathscr{F}_k]^2 - \mathrm{Var}(X_{k+1}|\mathscr{F}_k) \\ &= nX_k - X_k^2 - \frac{1}{n}X_k(n-X_k) = \frac{n-1}{n}X_k(n-X_k).\end{aligned}$$

On en déduit que $\left(\left(\frac{n}{n-1}\right)^k X_k(n-X_k)\right)_{k\geq 0}$ est une martingale, qu'on note M.

4. Soit $k \geq 0$. Si $\tau > k$, alors $1 \leq X_k \leq n-1$ donc $X_k(n-X_k) \geq n-1$ donc

$$\begin{aligned}\mathbb{P}(\tau \geq k+1) &= \mathbb{P}(X_k(n-X_k) \geq n-1) \\ &= \mathbb{P}\left(M_k \geq (n-1)\left(\frac{n}{n-1}\right)^k\right) \\ &\leq \frac{1}{n-1}\left(\frac{n-1}{n}\right)^k \mathbb{E}[M_k] \\ &= \frac{1}{n-1}\left(\frac{n-1}{n}\right)^k \mathbb{E}[M_0] \\ &\leq \frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k.\end{aligned}$$

On a donc $\mathbb{P}(\tau \geq k+1) \leq 1 \wedge \frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k$. On a $\frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k = 1$ pour $k \approx n\ln n$. On écrit donc

$$\begin{aligned}\mathbb{E}[\tau] &= \sum_{k\geq 0}\mathbb{P}(\tau \geq k+1) \\ &\leq \sum_{k=0}^{\lfloor n\ln n\rfloor} 1 + \sum_{k=1+\lfloor n\ln n\rfloor}^{+\infty} \frac{n^2}{4(n-1)}\left(1-\frac{1}{n}\right)^k \\ &\leq n\ln n + 1 + \frac{n^2}{4(n-1)}n\left(1-\frac{1}{n}\right)^{n\ln n} \\ &\leq n\ln n + 1 + \frac{n^3}{4(n-1)}\exp\left(-\frac{1}{n}n\ln n\right) \\ &= n\ln n + O(n).\end{aligned}$$

Par ailleurs, pour tout $k \geq 0$ on a $X_k(n-X_k) \leq \mathbb{1}_{\tau>k}\frac{n^2}{4}$ donc

$$a(n-a) = \mathbb{E}[M_k] \leq \frac{n^2}{4}\left(\frac{n}{n-1}\right)^k \mathbb{P}(\tau > k).$$

On en déduit

$$\mathbb{E}[\tau] = \sum_{k\geq 0}\mathbb{P}(\tau \geq k+1) \geq \frac{4}{n^2}a(n-a)\sum_{k\geq 0}\left(1-\frac{1}{n}\right)^k = \frac{4a(n-a)}{n}.$$

Par exemple, on prenant $a = \frac{n}{2}$, on obtient $\mathbb{E}[\tau] \geq (1+o(1))n$.

Remarque On peut montrer qu'en faisant tendre a et n vers $+\infty$ avec $\frac{a}{n} \to x$, on a

$$\mathbb{E}[\tau] \sim -2(x\ln x + (1-x)\ln(1-x))n.$$

Voir par exemple ce lien (pages 4 et 5) pour une explication heuristique.

Exercice 2 (Une preuve de la loi forte des grands nombres)
Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes. On suppose que $\mathbb{E}[Z_n] = 0$ pour tout $n \geq 1$ et $\sum_{n\geq 1} \frac{\mathrm{Var}(Z_n)}{n^2} < +\infty$. On pose, pour $n \geq 1$,

$$M_n = \sum_{j=1}^{n} \frac{Z_j}{j} \qquad \text{et} \qquad S_n = \sum_{j=1}^{n} Z_j\,.$$

1. Montrer que M_n converge p.s. quand n tend vers $+\infty$.
2. En exprimant S en fonction de M, en déduire $\frac{S_n}{n} \xrightarrow[n\to+\infty]{p.s.} 0$.
3. Soit X une variable aléatoire telle que $\mathbb{E}[X] = 0$. On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires i.i.d. de même loi que X. Pour tout $n \geq 1$, on pose $Y_n = X_n \mathbb{1}_{|X_n|\leq n}$. Montrer que :
(i) $\mathbb{E}[Y_n] \xrightarrow[n\to\infty]{} \mathbb{E}[X]$,
(ii) $\mathbb{P}\left(\exists n \geq 1, \forall j \geq n, X_j = Y_j\right) = 1$,
(iii) $\sum_{n\geq 1} \frac{\mathrm{Var}(Y_n)}{n^2} < \infty$.
4. En déduire la loi forte des grands nombres.

<u>Solution de l'exercice 2</u>

1. Comme les X_i sont d'espérance nulle, on vérifie facilement que M est une martingale. De plus, par indépendance des X_i, pour tout n on a

$$\mathbb{E}[M_n^2] = \mathrm{Var}(M_n) = \sum_{i=j}^{n} \mathrm{Var}\left(\frac{Z_j}{j}\right) = \sum_{j=1}^{n} \frac{\mathrm{Var}(Z_j)}{j^2},$$

qui est bornée par l'hypothèse sur les variances. La martingale M est donc bornée dans L^2, donc elle converge p.s. vers une variable M_∞.
2. Pour tout $n \geq 1$, on a

$$\frac{S_n}{n} = \frac{1}{n}\sum_{j=1}^{n} j(M_j - M_{j-1}) = \frac{1}{n}\left(\sum_{j=1}^{n} jM_j - \sum_{j=1}^{n-1}(j+1)M_j\right) = M_n - \frac{1}{n}\sum_{j=1}^{n-1} M_j.$$

D'après le lemme de Cesaro, le membre de droite tend vers 0, d'où $\frac{S_n}{n} \to 0$ p.s.
3. (i) On a $\mathbb{E}[Y_n] = \mathbb{E}[X\mathbb{1}_{|X|\leq n}]$, avec $X\mathbb{1}_{|X|\leq n} \to X$ p.s. et $|X\mathbb{1}_{|X|\leq n}| \leq |X|$, donc $\mathbb{E}[Y_n] \to \mathbb{E}[X] = 0$ par convergence dominée.
(ii) Pour tout $j \geq 1$, on a $\mathbb{P}(X_j \neq Y_j) = \mathbb{P}(|X_j| > j) = \mathbb{P}(|X| > j)$ donc

$$\sum_{j=1}^{\infty} \mathbb{P}(X_j \neq Y_j) = \sum_{j=1}^{\infty} \mathbb{P}(|X| > j) \leq \int_0^\infty \mathbb{P}(|X| > t)\,\mathrm{d}t = \mathbb{E}[|X|] < +\infty,$$

donc d'après le lemme de Borel-Cantelli, p.s., $X_j = Y_j$ pour j assez grand.
(iii) Enfin, on a

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\mathrm{Var}(Y_n)}{n^2} \leq \sum_{n=1}^{\infty} \frac{\mathbb{E}[Y_n^2]}{n^2} &= \sum_{n=1}^{\infty} \mathbb{E}\left[\frac{|X|^2}{n^2}\mathbb{1}_{|X|\leq n}\right] = \mathbb{E}\left[|X|^2 \sum_{n\geq 1\vee|X|} \frac{1}{n^2}\right] \\ &\leq \mathbb{E}\left[|X|^2 \frac{c}{|X|}\right] = c\mathbb{E}[|X|] < +\infty,\end{aligned}$$

en utilisant à la fin $\sum_{n\geq 1\vee a} = O\left(\frac{1}{a}\right)$.

4. On se place dans le cadre de la question 3. Pour tout $n \geq 1$, posons $Z_n = Y_n - \mathbb{E}[Y_n]$. On a $\mathrm{Var}(Z_n) = \mathrm{Var}(Y_n)$ pour tout $n \geq 0$ donc $(Z_n)_{n\geq 1}$ vérifie les hypothèses des questions 1 et 2, donc

$$\frac{1}{n}\sum_{j=1}^{n}(Y_j - \mathbb{E}[Y_j]) \xrightarrow[n\to\infty]{\text{p.s.}} 0\,.$$

Or $\mathbb{E}[Y_n] \xrightarrow[n\to\infty]{} \mathbb{E}[X] = 0$, donc d'après le lemme de Cesaro on a

$$\frac{1}{n}\sum_{j=1}^{n}\mathbb{E}[Y_j] \xrightarrow[n\to\infty]{} 0, \qquad \text{d'où} \qquad \frac{1}{n}\sum_{j=1}^{n} Y_j \xrightarrow[n\to\infty]{\text{p.s.}} 0.$$

Enfin, p.s., il existe $J \geq 1$ tel que $Y_j = X_j$ pour tout $j \geq J$. On a alors

$$\frac{1}{n}\sum_{j=1}^{n} X_j = \frac{1}{n}\sum_{j=1}^{J-1} X_j + \frac{1}{n}\sum_{j\geq J} Y_j \xrightarrow[n\to\infty]{} 0,$$

ce qui achève la démonstration.

Exercice 3 (Théorème de Rademacher)
Le but de cet exercice est de montrer par une approche probabiliste que toute fonction lipschitzienne est primitive d'une fonction mesurable bornée. Soient X une variable aléatoire de loi uniforme sur $[0,1]$ et $f : [0,1] \longrightarrow \mathbb{R}$ une fonction lipschitzienne de constante de Lipschitz $L > 0$. Pour tout $n \geq 0$, on pose

$$X_n = \lfloor 2^n X \rfloor 2^{-n} \qquad \text{et} \qquad Z_n = 2^n\left(f\left(X_n + 2^{-n}\right) - f(X_n)\right).$$

1. Montrer les égalités de tribus suivantes :

$$\sigma(X_0, X_1, \ldots, X_n) = \sigma(X_n) \qquad \text{et} \qquad \bigcap_{n\geq 0}\sigma(X_n, X_{n+1}, \ldots) = \sigma(X)\,.$$

2. Déterminer $\mathbb{E}[h(X_{n+1})|X_n]$ pour toute fonction $h : [0,1] \to \mathbb{R}$ mesurable continue. En déduire que $(Z_n)_{n\geq 0}$ est une $\mathscr{F}_n$-martingale bornée (où $\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n)$ pour tout $n \geq 0$).
3. Montrer qu'il existe une variable aléatoire Z, limite p.s. et dans L^1 de $(Z_n)_{n\geq 0}$, puis qu'il existe une fonction $g : [0,1] \longrightarrow \mathbb{R}$ mesurable bornée telle que $Z = g(X)$.
4. Calculer $\mathbb{E}[h(X)|X_n]$ pour toute fonction $h : [0,1] \to \mathbb{R}$ mesurable bornée. En déduire que p.s. :

$$Z_n = 2^n \int_{X_n}^{X_n+2^{-n}} g(u)\mathrm{d}u\,.$$

5. Conclure que pour tout $x \in [0,1]$, $f(x) = f(0) + \int_0^x g(u)\mathrm{d}u$.

<u>*Solution de l'exercice 3*</u>

1. On remarque que, pour $0 \leq k \leq n$, $X_k = 2^{-k}\lfloor 2^k X_n \rfloor$. On peut l'écrire proprement, ou faire un dessin pour s'en convaincre... Ainsi, pour $0 \leq k \leq n$, X_k est $\sigma(X_n)$-mesurable. On en déduit que $\sigma(X_0, X_1, \ldots, X_n) = \sigma(X_n)$.
   De plus, pour tout $n \geq 0$, par définition de X_n, on sait que X_n est $\sigma(X)$-mesurable. Ainsi, on a l'inclusion
   $$\bigcap_{n\geq 0}\sigma(X_n, X_{n+1}, \ldots) \subset \sigma(X)\,.$$
   Enfin, X_n converge p.s. vers X quand n tend vers l'infini, donc X est $\sigma(X_n, X_{n+1}, \ldots)$-mesurable pour tout $n \geq 0$. Ainsi, on obtient l'inclusion réciproque
   $$\sigma(X) \subset \bigcap_{n\geq 0}\sigma(X_n, X_{n+1}, \ldots).$$

2. Soit $h : [0,1] \longrightarrow \mathbb{R}$ une fonction mesurable continue. Alors h est bornée sur $[0,1]$ donc $h(X_n)$ est intégrable pour tout n. On a, pour $n \geq 0$ et $0 \leq k \leq 2^n - 1$,

$$\begin{aligned}
\mathbb{E}\left[h(X_{n+1})\mathbb{1}_{X_n=k/2^n}\right] &= \mathbb{E}\left[h(X_{n+1})\mathbb{1}_{X\in[k/2^n,(2k+1)/2^{n+1}[}\right] + \mathbb{E}\left[h(X_{n+1})\mathbb{1}_{X\in[(2k+1)/2^{n+1},(k+1)/2^n[}\right] \\
&= 2^{-(n+1)}\left(h\left(\frac{k}{2^n}\right) + h\left(\frac{2k+1}{2^{n+1}}\right)\right).
\end{aligned}$$

On en déduit

$$\mathbb{E}(h(X_{n+1}) \mid X_n) = \frac{h(X_n)}{2} + \frac{h\left(X_n + 2^{-(n+1)}\right)}{2}.$$

Pour tout $n \geq 0$, la variable Z_n est $\mathscr{F}_n$-mesurable, et $|Z_n| \leq L$ donc Z_n est intégrable. De plus,

$$\begin{aligned}
\mathbb{E}[Z_{n+1}|\mathscr{F}_n] &= 2^{n+1}\mathbb{E}\left[f(X_{n+1} + 2^{-(n+1)}) - f(X_{n+1}) \mid \mathscr{F}_n\right] \\
&= 2^{n+1}\mathbb{E}\left[f(X_{n+1} + 2^{-(n+1)}) - f(X_{n+1}) \mid X_n\right] \\
&= 2^n\left[f(X_n + 2^{-(n+1)}) - f(X_n) + f(X_n + 2^{-n}) - f(X_n + 2^{-(n+1)})\right] \\
&= Z_n\,,
\end{aligned}$$

en utilisant à la deuxième ligne la première égalité de tribus de la question 1. Donc $(Z_n)_{n\geq 0}$ est une $(\mathscr{F}_n)$-martingale bornée par L.

3. D'après la question 2, on sait que (Z_n) est une martingale bornée dans L^p pour tout $p > 0$, donc (Z_n) converge p.s. et dans L^1. On note Z sa limite. Pour tout $n \geq 0$, Z_n est mesurable par rapport à la tribu $\sigma(X_n, X_{n+1}, \ldots)$ donc Z est mesurable par rapport à la tribu $\bigcap_{n\geq 0}\sigma(X_n, X_{n+1}, \ldots)$. D'après la question 1, Z est ainsi $\sigma(X)$-mesurable. Il existe donc une fonction $g : [0,1] \longrightarrow \mathbb{R}$ borélienne telle que $Z = g(X)$. De plus, Z étant bornée par L, on peut choisir g bornée (en prenant remplaçant g par $g \wedge L$ par exemple).

4. Soit $h : [0,1] \longrightarrow \mathbb{R}$ une fonction mesurable bornée. La variable $h(X)$ est intégrable et on a, pour $n \geq 0$ et $0 \leq k \leq 2^n - 1$,

$$\mathbb{E}\left[h(X)\mathbb{1}_{X_n=k2^{-n}}\right] = \mathbb{E}\left[h(X)\mathbb{1}_{X\in[k2^{-n},(k+1)2^{-n}[}\right] = \int_{k2^{-n}}^{(k+1)2^{-n}} h(x)\mathrm{d}x.$$

On en déduit que

$$\mathbb{E}[h(X) \mid X_n] = 2^n \int_{X_n}^{X_n+2^{-n}} h(x)\mathrm{d}x.$$

La $(\mathscr{F}_n)$-martingale $(Z_n)_{n\geq 0}$ converge p.s. et dans L^1 vers Z, donc $Z_n = \mathbb{E}[Z|\mathscr{F}_n]$ pour tout $n \geq 0$. On a donc p.s.

$$Z_n = \mathbb{E}[g(X)|X_n] = 2^n \int_{X_n}^{X_n+2^{-n}} g(u)\mathrm{d}u.$$

5. D'après la question 4., pour tout $n \geq 0$,

$$f(X_n + 2^{-n}) - f(X_n) = \int_{X_n}^{X_n+2^{-n}} g(u)\mathrm{d}u \quad \text{p.s.}$$

Donc, pour tout $n \geq 0$ et pour tout $0 \leq k \leq 2^n - 1$,

$$f((k+1)2^{-n}) - f(k2^{-n}) = \int_{k2^{-n}}^{(k+1)2^{-n}} g(u)\mathrm{d}u$$

puis, en sommant, pour tout $0 \leq k \leq 2^n$,

$$f(k2^{-n}) = f(0) + \int_0^{k2^{-n}} g(u)\mathrm{d}u\,.$$

Ainsi, pour tout $x \in [0,1]$,

$$f\left(2^{-n}\lfloor 2^n x\rfloor\right) \;=\; f(0)+\int_0^{2^{-n}\lfloor 2^n x\rfloor} g(u)\mathrm{d}u$$

et en faisant tendre n vers l'infini, par continuité de f on obtient

$$f(x) \;=\; f(0)+\int_0^{x} g(u)\mathrm{d}u.$$

Exercice 4 (Vaisseau spatial perdu)
Le *Millenium Falcon* se trouve à une distance D_0 du Soleil mais ses commandes ne répondent plus : toutes les heures, Han Solo ne peut qu'entrer une distance r_n inférieure à la distance au Soleil dans l'ordinateur de bord, qui effectue alors un saut dans l'hyperespace de longueur r_n et de direction choisie uniformément dans la sphère S^2. On note D_n la distance du vaisseau au Soleil après n sauts et $\mathscr{F}_n$ la tribu engendrée par les n premiers sauts. Han Solo veut revenir dans le système solaire, c'est-à-dire à distance au plus d du soleil.

1. En utilisant des souvenirs de physique de prépa (théorème de Gauss), montrer que $\left(\frac{1}{D_n}\right)$ est une martingale.
2. En déduire que la probabilité que Han Solo revienne un jour dans le système solaire est inférieure ou égale à $\frac{d}{D_0}$.
3. A la place du pilote, feriez-vous plutôt de grands ou de petits sauts ?

<u>*Solution de l'exercice 4*</u> Toutes les justifications des interversions seront laissées en exercice.

1. Soit X_n la variable aléatoire à valeurs dans S^2 qui indique la direction du n-ième saut. Alors on veut montrer que $\mathbb{E}\left[\|S_n+r_{n+1}X_{n+1}\|^{-1}|\mathscr{F}_n\right] = \|S_n\|^{-1}$. Pour tout x dans $\mathbb{R}^3\backslash\{0\}$, on pose $f(x)=\|x\|^{-1}$. On veut donc montrer que pour tous $x\in\mathbb{R}^3$ et $r<\|x\|$:

   $$\frac{1}{4\pi}\int_{S^2} f(x+ru)^{-1}\mathrm{d}u = f(x).$$

   Il suffit pour cela de vérifier que la dérivée par rapport à r du membre de gauche est nulle (par convergence dominée, il tend bien vers 0 quand $r\to 0$). Notons que le membre de gauche a une discontinuité en $r=\|x\|$, c'est pourquoi on impose $r<\|x\|$. En intervertissant dérivée et intégrale puis en appliquant le théorème de Gauss, on obtient

   $$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}r}\int_{S^2} f(x+ru)^{-1}\mathrm{d}u &= \int_{S^2} r\,\nabla f(x+ru)\cdot u\,\mathrm{d}u\\ &= r^2\int_{B_1}\mathrm{div}(\nabla f(x+y))\,\mathrm{d}y\\ &= r^2\int_{B_1}\Delta f,\end{aligned}$$

   où B_1 est la boule de rayon 1 autour de l'origine dans $\mathbb{R}^3$. Un simple calcul (ou, à nouveau, des souvenirs de physique de prépa) montre que le Laplacien Δf est nul sur $\mathbb{R}^3\backslash\{0\}$, donc $\frac{1}{D_n}$ est bien une martingale.
2. Soit $T=\inf\{n|D_n\leq d\}$. Le temps T est un temps d'arrêt (éventuellement infini), donc on peut appliquer le théorème d'arrêt à $T\wedge t$ et à la martingale $\frac{1}{D_n}$:

   $$\mathbb{E}\left[\frac{1}{D_{T\wedge t}}\right]=\frac{1}{D_0}.$$

On en déduit

$$\begin{aligned}
\mathbb{P}(T \leq t) &= \mathbb{P}(D_{T\wedge t} \leq d) \\
&= \mathbb{P}\left(\frac{1}{D_{T\wedge t}} \geq \frac{1}{d}\right) \\
&\leq \left(\frac{1}{d}\right)^{-1} \mathbb{E}\left[\frac{1}{D_{T\wedge t}}\right] \\
&= \frac{d}{D_0}.
\end{aligned}$$

Ceci est valable pour tout $t > 0$ donc $\mathbb{P}(T < +\infty) \leq \frac{d}{D_0}$.

3. On veut que l'inégalité de la question précédente soit la plus serrée possible. Le seul endroit où on n'a pas égalité ci-dessus est dans l'inégalité de Markov (avant-dernière ligne du dernier calcul). Pour que l'inégalité de Markov soit serrée, il faut que $\frac{1}{D_{T\wedge t}}$ ne puisse pas être "beaucoup" plus grande que $\frac{1}{d}$. Il faut donc faire de petits sauts à l'approche du système solaire. On peut vérifier (exercice !) que pour tout $\varepsilon > 0$, si le saut à chaque étape n est inférieur ou égal à $D_n - d + \varepsilon$, alors on a $\mathbb{P}(T < +\infty) \geq \frac{d-\varepsilon}{D_0}$.

Remarque L'hypothèse "les sauts sont plus petits que la distance au Soleil" peut paraître arbitraire. En supprimant cette hypothèse, le processus $\left(\frac{1}{D_n}\right)_{n\geq 0}$ n'est plus forcément une martingale mais une *surmartingale*, c'est-à-dire que $E\left[\frac{1}{D_{n+1}}|\mathscr{F}_n\right] \leq \frac{1}{D_n}$. La raison est que, au sens des distributions, le Laplacien de $x \to \|x\|^{-1}$ sur $\mathbb{R}^3$ est (à une constante) multiplicative près) $-\delta_0$, donc est négatif. Le théorème d'arrêt pour les surmartingales donne

$$\mathbb{E}\left[\frac{1}{D_{T\wedge t}}\right] \leq \frac{1}{D_0}.$$

L'inégalité étant dans le bon sens, le résultat de la question 2 reste vrai. Cela montre que faire des sauts trop grands ne peut qu'aggraver la situation de notre vaisseau.

Exercice 5 (Propriété de Liouville)
Soit $G = (V, E)$ un graphe infini, connexe et localement fini (i.e. où chaque sommet n'a qu'un nombre fini de voisins) et $h : V \to \mathbb{R}$. On dit que h est *harmonique sur* G si pour tout $x \in V$, on a

$$h(x) = \frac{1}{\deg(x)} \sum_{y\sim x} h(y),$$

où la somme est effectuée sur les voisins y de x et où $\deg(x)$ est le nombre de ces voisins. On dit que G vérifie la *propriété de Liouville* si toute fonction bornée harmonique sur G est constante.

1. Montrer que si h est harmonique et (X_n) est une marche aléatoire simple sur G, alors $(h(X_n))_{n\geq 0}$ est une martingale.
2. Montrer que si la marche aléatoire simple sur G est récurrente (i.e. si elle visite presque sûrement tous les points une infinité de fois), alors G vérifie la propriété de Liouville.
3. Soient $x, y \in \mathbb{Z}^d$ tels que $\sum_{i=1}^d x_i \equiv \sum_{i=1}^d y_i \pmod 2$. Montrer qu'il existe (X_n) et (Y_n) deux marches aléatoires simples (non indépendantes !) issues respectivement de x et y telles que p.s., pour n assez grand, $X_n = Y_n$.
4. En déduire que $\mathbb{Z}^d$ vérifie la propriété de Liouville.
5. Donner un exemple de graphe (connexe, localement fini) ne vérifiant pas la propriété de Liouville.

Indication : Pour la question 3, commencer par le cas $d = 1$ puis essayer d'adapter à d quelconque.

Solution de l'exercice 5

1. Soit $\mathscr{F}_n = \sigma(X_0, \dots, X_n)$. Conditionnellement à $\mathscr{F}_n$, le sommet X_{n+1} est uniforme parmi les voisins de X_n, donc $\mathbb{E}\left[h(X_{n+1})|\mathscr{F}_n\right]$ est la moyenne de h sur les voisins de X_n, c'est-à-dire X_n car h est harmonique.
2. Soit h harmonique bornée sur G. Soient $x, y \in V$ et (X_n) une marche aléatoire simple issue de x. Alors $(h(X_n))_{n \geq 0}$ est une martingale, donc elle converge p.s. Or, elle visite une infinité de fois x et y, donc elle prend une infinité de fois les valeurs $h(x)$ et $h(y)$, donc $h(x) = h(y)$, et ce pour tous x et y. La fonction h est donc constante, donc G est Liouville.
3. Dans le cas $d = 1$, l'idée est de faire démarrer deux marches aléatoires indépendantes $\widetilde{X}$ et $\widetilde{Y}$ issues de x et y. Par récurrence de la marche aléatoire simple sur $\mathbb{Z}$ (plus la condition de parité), le temps $\tau_1 = \inf\{n | \widetilde{X}_n = \widetilde{Y}_n\}$ est fini p.s. On prend alors $X = \widetilde{X}$ ainsi que $Y_n = \widetilde{Y}_n$ pour $n \leq \tau_1$ et $Y_n = \widetilde{X}_n$ pour $n \geq \tau_1$.

   Dans le cas général, il faut "coupler les coordonnées une par une" : on démarre deux marches indépendantes de x et y, et on note τ_1 le premier temps où leurs coordonnées selon e_1 coïncident. À partir de τ_1, on applique la stratégie du cas $d = 1$ pour que les coordonnées selon e_1 restent les mêmes pour $n \geq \tau_1$. Puis on attend τ_2, le premier temps où les coordonnées selon e_2 coïncident, et ainsi de suite.

   Les détails sont laissés en exercice, vous pouvez aussi venir me voir au bureau V2.
4. Soit h harmonique bornée sur G et soient x, y, X et Y comme dans la question précédente. Alors $(h(X_n))_{n \geq 0}$ et $(h(Y_n))_{n \geq 0}$ sont deux martingales bornées, donc elles convergent p.s. et dans L^1 vers respectivement X_∞ et Y_∞. Comme $X_n = Y_n$ pour n assez grand, on a $X_\infty = Y_\infty$ p.s. Par convergence L^1, on peut donc écrire

   $$h(x) = \mathbb{E}\left[X_\infty\right] = \mathbb{E}\left[Y_\infty\right] = h(y),$$

   donc h est constante sur $\{x \in \mathbb{Z}^d | \sum_{i=1}^d x_i \equiv 0 \pmod 2\}$, et il est facile d'en conclure que h est constante sur V.
5. Considérer l'arbre binaire infini. La marche aléatoire est transiente (car la "hauteur" augmente de 1 avec proba $\frac{2}{3}$ et diminue de 1 avec proba $\frac{1}{3}$), donc à partir d'un certain rang elle reste soit dans la moitié gauche de l'arbre, soit dans la moitié droite. On peut alors poser

   $$h(x) = \mathbb{P}\left(\text{la marche aléatoire issue de } x \text{ finit dans la moitié gauche de l'arbre}\right).$$

   On vérifie facilement que h est harmonique et bornée (par 1). De plus, si x est "très haut" dans la moitié gauche de l'arbre, il a une très faible probabilité de redescendre jusqu'à la racine, donc $h(x)$ est proche de 1. Les détails sont laissés en exercice.