ANNÉE 2021

LECTURE DIRIGÉE DE RECHERCHE

Jérémy BETTINGER & Florent CORNIQUEL

INTRODUCTION AUX CHAÎNES DE MARKOV ET MARCHE ALÉATOIRE SUR $\mathbb{Z}^d$.

Encadré par **Rémi MOREAU**

Table des matières

1 Généralités sur les chaînes de Markov

Dans tout ce qui suit, $(X_n)_{n\in\mathbb{N}}$ désigne une suite de variables aléatoires à valeurs dans E, l'ensemble des états d'un processus décrit par les X_n, que l'on assimilera à une partie de $\mathbb{N}$.

1.1 Définitions et premières propriétés

Définition 1.1

Une suite $(X_n)_{n\in\mathbb{N}}$ est une *chaîne de Markov* si pour tout $n \in \mathbb{N}$ et toute famille $(i_0, \ldots, i_{n-1}, i, j)$ d'éléments de E telle que la probabilité $\mathbb{P}(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)$ est strictement positive, on a :

$$\mathbb{P}(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j \mid X_n = i). \tag{1}$$

Autrement dit, si on considère l'évolution temporelle du processus en question, l'état du système à l'instant $(n+1)$ dépend exclusivement de son état à l'instant n et pas de ses états antérieurs. On parle de processus sans mémoire, ou non héréditaire.

Définition 1.2

Une chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ est dite *homogène* si la probabilité apparaissant dans la définition précédente ne dépend pas de n. Pour i et j deux entiers naturels, on appelle alors *probabilité de passage* de l'état i à l'état j en une transition la quantité :

$$p_{i,j} := \mathbb{P}(X_{n+1} = j \mid X_n = i).$$

Définition 1.3

On appelle *matrice de transition* (ou de passage) la matrice dont le coefficient (i, j) est $p_{i,j}$:

$$\mathcal{P} = \begin{pmatrix} p_{0,0} & p_{0,1} & p_{0,2} & \cdots \\ p_{1,0} & p_{1,1} & p_{1,2} & \\ p_{2,0} & p_{2,1} & p_{2,2} & \\ \vdots & & & \ddots \end{pmatrix}.$$

Par soucis de lisibilité, on notera souvent $\mathcal{P} = (p_{i,j})_{(i,j)\in E^2}$. C'est une matrice qui peut être finie ou dénombrable selon que l'ensemble des états E est lui-même fini ou dénombrable.

Proposition 1.1

Une matrice de transition $\mathcal{P} = (p_{i,j})_{(i,j)\in E^2}$ vérifie les propriétés suivantes :

1. *Pour tout couple $(i, j) \in E^2$, on a $p_{i,j} \geqslant 0$.*
2. *Pour tout $i \in E$, on a $\sum_{j\in E} p_{i,j} = 1$.*

Démonstration. Les nombres $p_{i,j}$ dont des probabilités donc sont positifs. Ensuite, il suffit de remarquer qu'à $i \in E$ fixé, l'application $B \mapsto \sum_{j\in E} p_{i,j}$ est une mesure de probabilité sur E. □

Remarque 1. Une matrice qui vérifie ces deux propriétés est appelée matrice stochastique.

Proposition 1.2

Soit $\mathcal{P}$ une matrice de transition. Alors :

1. *$\mathcal{P}$ admet 1 pour valeur propre.*
2. *Un vecteur propre associé est le vecteur V ayant toutes ses composantes égales à 1.*

Démonstration. Il suffit de vérifier que le produit de $\mathcal{P}$ et du vecteur V avec toutes ses composantes égales à 1 redonne V. Cela résulte directement de la Proposition 1.1. □

Définition 1.4

À toute matrice de transition on peut associer un *graphe*. Les sommets de ce graphe sont les différents états de la chaîne. Il y a une flèche étiquetée $p_{i,j}$ entre le sommet numéroté i et celui numéroté j dès que la probabilité de transition entre ces deux états est strictement positive i.e. $p_{i,j} > 0$.

Remarque 2. Dans le cas d'un ensemble d'états fini, la représentation de la matrice par un graphe est particulièrement utile et parlante.

1.2 Exemples de chaînes de Markov

Il y a des modèles classiques de chaînes de Markov homogènes que l'on rencontre assez régulièrement et avec lequels il est bon de se familiariser dès le début. En voici quelques-uns.

1.2.1 La chaîne à deux états

On considère ici ce qui est vraisemblablement parmi les exemples de base de chaînes de Markov. Il s'agit de la chaîne à deux états qui comme son nom l'indique est associée à l'évolution d'un système pouvant prendre atteindre exactement deux états. Sa matrice de transition est de la forme suivante :

$$\mathcal{P} = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}.$$

Pour exclure les cas triviaux, on considère α et β à valeurs dans $]0,1[$. Le graphe associé est très simple :

$$1-\alpha \circlearrowleft \overset{1}{\bullet} \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \overset{2}{\bullet} \circlearrowright 1-\beta.$$

1.2.2 Le modèle de diffusion d'Ehrenfest

On dispose ici de deux urnes A et B qui contiennent, à elles deux, a boules numérotées de 1 à a. À chaque instant, on choisit un nombre aléatoirement entre 1 et a avec une probabilité de $1/a$. On change alors d'urne la boule dont le numéro a été tiré. L'ensemble des états est ici $E = \{0, \dots, a\}$ et pour $j \in E$, on dit que le processus est dans l'état j lorsque l'urne A contient j boules. Cette convention étant prise, on note que lorsque le processus est dans l'état 0 (i.e. l'urne A est vide) ou dans l'état a (i.e. l'urne B est vide) la probabilité de passer respectivement dans les états 1 et $(a-1)$ sont de 1. La matrice de transition prend la forme suivante :

$$\mathcal{P} = \begin{pmatrix} 0 & 1 & 0 & & & (0) \\ 1/a & 0 & (a-1)/a & 0 & & \\ 0 & 2/a & 0 & (a-2)/a & 0 & \\ & & & & & \\ & \ddots & \ddots & \ddots & \ddots & 0 \\ & & & & 0 & 1/a \\ (0) & & & 0 & 1 & 0 \end{pmatrix}$$

Le graphe associé est de la forme suivante :

$$\overset{0}{\bullet} \underset{1/a}{\overset{1}{\rightleftarrows}} \overset{1}{\bullet} \underset{2/a}{\overset{(a-1)/a}{\rightleftarrows}} \overset{2}{\bullet} \underset{3/a}{\overset{(a-2)/a}{\rightleftarrows}} \cdots \underset{(a-2)/a}{\overset{3/a}{\rightleftarrows}} \overset{a-2}{\bullet} \underset{(a-1)/a}{\overset{2/a}{\rightleftarrows}} \overset{a-1}{\bullet} \underset{1}{\overset{1/a}{\rightleftarrows}} \overset{a}{\bullet}.$$

Cela étant clair, si on introduit maintenant la suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ où X_n désigne le nombre de boules dans l'urne A à l'instant n, et qu'on prend $X_0 = 0$ alors le processus décrit par le suite $(X_n)_{n\in\mathbb{N}}$ peut être la diffusion d'un gaz.

1.2.3 Le modèle de la ruine du joueur

Un joueur A affronte un joueur B dans une succession de parties d'un jeu du type "pile ou face" toutes indépendantes. À eux deux ils représentent une fortune totale de a euros. On prend pour convention que le joueur A gagne un euros qui lui est donné par le joueur B quand il gagne, et qu'il donne un euro à ce dernier quand il perd. On se donne p un réel tel que $0 < p < 1$ et on prend pour convention que le joueur A gagne une partie avec une probabilité p, et qu'il perd avec une probabilité $q = 1-p$. Le jeu prend fin dès que l'un des deux joueurs est ruiné. Pour tout $n \geqslant 0$ on note X_n la fortune du joueur A à l'issue de la n-ième partie. La suite $(X_n)_{n\in\mathbb{N}}$ est une chaîne de Markov homogène dans le temps caractérisée par un ensemble des états possibles $E = \{0, ..., a\}$, par la matrice de transition :

$$\mathcal{P} = \begin{pmatrix} 1 & 0 & & & & (0) \\ q & 0 & p & & & \\ 0 & q & 0 & p & & \\ & & \ddots & & \ddots & \\ & & & q & 0 & p \\ (0) & & & & 0 & 1 \end{pmatrix},$$

et par le graphe associé

$$\circlearrowleft \overset{0}{\bullet} \xleftarrow[q]{} \overset{1}{\bullet} \underset{q}{\overset{p}{\rightleftarrows}} \overset{2}{\bullet} \underset{q}{\overset{p}{\rightleftarrows}} \cdots \underset{q}{\overset{p}{\rightleftarrows}} \overset{a-2}{\bullet} \underset{q}{\overset{p}{\rightleftarrows}} \overset{a-1}{\bullet} \xrightarrow{p} \overset{a}{\bullet} \circlearrowleft .$$

On note qu'une fois entré dans l'état 0 i.e. lorsque A est ruiné, ou dans l'état a i.e. lorsque B est ruiné, le système n'évolue plus. Ces états sont dits absorbants.

1.3 La relation de Chapman-Kolmogorov

On appelle généralement ainsi la propriété qui relie, pour une chaîne de Markov homogène, les probabilités de transition en n étapes aux probabilités de transition en une seule étape. Dans la suite on conserve les notations du paragraphe précédent. La suite $(X_n)_n$ désigne une chaîne de Markov homogène dans le temps, caractérisée par un ensemble d'états E et une matrice de transition $\mathcal{P} = (p_{i,j})_{(i,j)\in E^2}$.

Définition 1.5

Pour $n \in \mathbb{N}$ et $i, j \in E$, on désigne par $p_{i,j}^{(n)}$ la probabilité partant de l'état i à l'instant 0 d'atteindre l'état j au temps n. Plus formellement, on pose

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i).$$

On pose également $\mathcal{P}^{(n)} = (p_{i,j}^{(n)})_{(i,j)\in E^2}$.

Théorème 1.1 (*Chapman-Kolmogorov*)

Pour tout $n \in \mathbb{N}$, la matrice de transition en n étapes est la puissance n-ieme de la matrice de transition en une étape i.e.

$$\mathcal{P}^{(n)} = (\mathcal{P})^n.$$

Démonstration. On prouve le résultat par récurrence sur la valeur de n. Les cas $n = 0$ et $n = 1$ sont des cas triviaux. Supposons donc maintenant $n \geqslant 2$ et supposons la propriété vraie au rang $n-1$.
Ainsi, $\mathcal{P}^n = \mathcal{P}^{n-1}\mathcal{P} = \mathcal{P}^{(n-1)}\mathcal{P}^{(1)}$. Dès lors, notons $p_{i,j}^n$ le coefficient en place (i, j) de la matrice $\mathcal{P}^{(n)}$. On a :

$$\begin{aligned} p_{i,j}^n &= \sum_{k\in E} p_{i,k}^{(n-1)} p_{k,j} = \sum_{k\in E} \mathbb{P}(X_{n-1=k \mid X_0=i})\mathbb{P}(X_1 = j \mid X_0 = k) \\ &= \sum_{k\in E} \mathbb{P}(X_{n-1=k \mid X_0=i})\mathbb{P}(X_n = j \mid X_{n-1} = k) \ \text{ par homogénéité de la chaîne.} \end{aligned}$$

Considérons maintenant l'événement $A(i_1, ..., i_{n-2}) := (X_1 = 1, ..., X_{n-2} = i_{n-2})$.
La chaîne étant de Markov, l'évolution du système ne dépend que de son état au temps précédent, donc :

$$\mathbb{P}(X_n = j \mid X_{n-1} = k, A(i_1, ..., A_{n-2}), X_0 = i) = \mathbb{P}(X_n = j \mid X_{n-1} = k).$$

D'autre part, les $A(i_1, ..., i_{n-2})$ forment un système complet d'événement, donc par la formule des probabilités totales, il vient que :

$$\mathbb{P}(X_n = j \mid X_{n-1} = k) = \sum_{i_1,...,i_{n-2}\in E} \mathbb{P}(X_{n-1} = k, A(i_1, ..., i_{n-2}) \mid X_0 = i).$$

Par suite, en remplaçant dans la somme initale, on obtient :

$$\begin{aligned} p_{i,j}^n &= \sum_{k\in E} \sum_{i_1,...,i_{n-2}} \mathbb{P}(X_{n-1} = k, A(i_1, ..., i_{n-2}) \mid X_0 = i)\mathbb{P}(X_n = j \mid X_{n-1} = k) \\ &= \sum_{k\in E} \sum_{i_1,...,i_{n-2}} \mathbb{P}(X_{n-1} = k, A(i_1, ..., i_{n-2}) \mid X_0 = i)\mathbb{P}(X_n = j \mid X_{n-1} = k, A(i_1, ..., i_{n-2}), X_0 = i) \\ &= \sum_{k\in E} \sum_{i_1,...,i_{n-2}} \mathbb{P}(X_n = j, A(i_1, ..., i_{n-2}) \mid X_0 = i) \ \text{ (probabilité conditionnelle)} \\ &= \mathbb{P}(X_n = j \mid X_0 = i) = p_{i,j}^{(n)}. \end{aligned}$$

□

Corollaire 1.1

Pour tout entier naturel, la matrice $\mathcal{P}^{(n)}$ est stochastique.

Démonstration. En effet, pour tout $i \in E$, on a :

$$\sum_{j\in E} p_{i,j}^{(n)} = \sum_{j\in E} \mathbb{P}(X_n = j \mid X_0 = i) = 1$$

C'est une conséquence de la formule des probabilités totales. On remarque dans le même temps que pour un entier $n \geqslant 1$ fixé, la suite de variables aléatoires $(X_{mn})_{m\in\mathbb{N}}$ est encore une chaîne de Markov avec le même ensemble d'état, et dont la matrice de transition est $\mathcal{P}^n$. □

Corollaire 1.2

Pour tout couple $(i,j) \in E^2$ et $(m,n) \in \mathbb{N}^2$, on a l'identité suivante :

$$\mathbb{P}(X_{m+n} = j \mid X_0 = i) = \sum_{k\in E} \mathbb{P}(X_m = k \mid X_0 = i)\mathbb{P}(X_n = j \mid X_0 = k).$$

Ce qui peut s'écrire plus succintement comme :

$$p_{i,j}^{(m+n)} = \sum_{k\in E} p_{i,k}^{(m)} p_{k,j}^{(n)}. \tag{2}$$

Démonstration. Cela résulte directement de l'associativité du produit matriciel, dans la mesure où

$$\mathcal{P}^{(m+n)} = \mathcal{P}^{m+n} = \mathcal{P}^m\mathcal{P}^n = \mathcal{P}^{(m)}\mathcal{P}^{(n)}.$$

□

Proposition 1.3

Soient $n \geqslant 0$ et $r \geqslant 1$ deux entiers. Alors

$$\mathbb{P}(X_{n+1} = j_{n+1}, \dots, X_{n+r} = j_{n+r} \mid X_0 = i_0, \dots, X_n = i_n) = p_{i_n,j_{n+1}} p_{j_{n+1},j_{n+2}} \cdots p_{j_{n+r-1},j_{n+r}}. \tag{3}$$

Démonstration. On procède par récurrence sur r. L'initialisation est évidente, parce que le cas $r = 1$ correspond à (1) c'est-à-dire la définton d'une chaîne de Markov. Supposons que la relation (3) soit vérifiée pour $r-1$ avec $r \geqslant 2$. On introduit les événements suivants :

$$A_1 := \{X_0 = i_0, \dots, X_n = i_n\} \;\; ; \;\; A_2 := \{X_{n+1} = j_{n+1}, \dots, X_{n+r-1} = j_{n+r-1}\} \;\; ; \;\; A_3 := \{X_{n+r} = j_{n+r}\}$$

La relation (1) nous assure que $\mathbb{P}(A_3 \mid A_1, A_2) = p_{j_{n+r-1},j_{n+r}}$ et d'autre part

$$\mathbb{P}(A_2 \mid A_1) = p_{i_n,j_{n+1}} p_{j_{n+1},j_{n+2}} \cdots p_{j_{n+r-2},j_{n+r-1}} \quad \text{par hypothèse de récurrence.}$$

On peut alors conclure en utilisant la relation $\mathbb{P}(A_2, A_3 \mid A_1) = \mathbb{P}(A_3 \mid A_1, A_2)\mathbb{P}(A_2 \mid A_1)$. □

1.4 Mesure de probabilité d'une chaîne de Markov

On prend toujours en considération $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov homogène, prenant ses valeurs dans E fini ou dénombrable, et admettant pour matrice de transition $\mathcal{P} = (p_{i,j})_{(i,j)\in E^2}$. On se donne également une mesure de probabilité μ sur E. En d'autres termes, on se donne une famille $(\mu_i)_{i\in E}$ de nombres réels positifs tels que $\sum_{i\in E} \mu_i = 1$. Le but de ce paragraphe est de montrer que la matrice de transition $\mathcal{P}$ et la mesure μ permettent de définir une mesure de probabilité $\mathbb{P}$ sur l'espace probabilisable $(\Omega, \mathcal{A})$ sur lequel sont définies les variables X_n.

1.4.1 Mesure de probabilité et loi initiale

Définition 1.6

On dit que $\mathbb{P}$ est la *mesure de probabilité* de cette chaîne si les conditions suivantes sont vérifiées :

1. pour tout état i, on a $\mathbb{P}(X_0 = i) = \mu_i$. On dit alors que μ est la loi initiale de la chaîne.
2. pour tout couple d'états (i,j), on a $\mathbb{P}(X_1 = j, X_0 = i) = p_{i,j}$.

Dans ces conditions, la formule (3) donnée dans la proposition 1.3 assure pour $n = 0$ et $r \geqslant 1$ que :

$$\mathbb{P}(X_0 = i_0, \dots, X_r = i_r) = \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1, \dots, X_r = i_r \mid X_0 = i_0) = \mu_{i_0} p_{i_0,i_1} \cdots p_{i_{r-1},i_r}. \tag{4}$$

1.4.2 Processus canonique

Considérons l'ensemble $E^{\mathbb{N}}$ des suites à valeurs dans E un ensemble dénombrable. Pour un $n \in \mathbb{N}$, on définit la n-ième projection comme étant l'application X_n qui envoie un élément $\omega = (i_k)_{k\in\mathbb{N}}$ sur $X_n(\omega) := i_n$. On peut alors décrire une chaîne de Markov de mesure de probabilité P^μ en prenant pour Ω l'ensemble $E^{\mathbb{N}}$, pour suite de variables aléatoires la suite $(X_n)_{n\in\mathbb{N}}$ la suite des projections définies au-dessus, et pour tribu $\mathcal{A}$ la tribu engendrée par les variables X_n c'est-à-dire la tribu engendrée par les événements de la forme $(X_n = i)$, où $i \in E$ et $n \in \mathbb{N}$. On peut alors prouver qu'il existe une unique loi de probabilité sur $(\Omega, \mathcal{A})$ qui vérifie les relations écrites dans la définition 4. Le processus ainsi construit est appelé processus canonique.

1.4.3 Formule de conditionnement

Dans tout ce qui suit, on désigne par A un événement appartenant à la tribu $\mathcal{A}_n$ engendrée par les projections $X_0, \dots, X_n$. La proposition suivante assure que A est bien un événement.

Proposition 1.4

Soient $I = \{i_1, \dots, i_k \mid i_1 < \cdots < i_k\}$ un sous-ensemble fini de $\mathbb{N}$ et $(Y_i)_{i\in I}$ une famille de variables aléatoires à valeurs entières. Soit aussi $\mathcal{C}$ la classe des événements de la forme $\{Y_{i_1} = b_1, \dots, Y_{i_k} = b_k\}$ où les b_i sont des entiers. Alors les éléments de la tribu $\sigma((Y_i)_{i\in I})$ engendrée par les $(Y_i)_{i\in I}$ sont des événements.

Démonstration. Pour prouver ce résultat, il s'agit de se rendre compte que la tribu $\sigma((Y_i)_{i\in I})$ est en fait engendrée par les événements de la forme $(Y_i = b)$ où $i \in I$ et $b \in \mathbb{N}$, donc aussi par les éléments de $\mathcal{C}$. De fait, il est clair qu'un élément de $\mathcal{C}$ s'écrit comme intersection d'événements de la forme $(Y_i = b)$. D'autre part, les éléments de $\mathcal{C}$ forment une partition dénombrable de Ω donc un élément de $\sigma((Y_i)_{i\in I})$ de la forme $(Y_i = b)$ s'écrit comme une réunion dénombrable d'éléments de $\mathcal{C}$. Ce dernier point permet alors d'affirmer que $\overline{\mathcal{C}}$ l'ensemble des réunions dénombrables d'éléments de $\mathcal{C}$ est une tribu. Étant incluse dans $\sigma((Y_i)_{i\in I})$, on en déduit le résultat souhaité. □

Proposition 1.5

Soient $n \in \mathbb{N}$ et $A \in \mathcal{A}_n$. Si $\mathbb{P}(A, X_n = i) > 0$, alors :

$$\mathbb{P}(X_{n+1} = j \mid A,\ X_n = i) = \mathbb{P}(X_{n+1} = j \mid X_n = i) = p_{i,j}. \tag{5}$$

Démonstration. On considère donnés les éléments i, j de E qui apparaissent dans la proposition. Pour $n = 0$, l'événement A est nécessairement $(X_0 = i)$ pour qu'on ait bien $\mathbb{P}(A, X_0 = i) > 0$, et dans ce cas on a simplement $\mathbb{P}(X_1 = j \mid A, X_0 = i) = p_{i,j}$. On va dire qu'un événement A de la tribu $\mathcal{A}_n$ vérifie la propriété $(\mathcal{H})$ s'il vérifie d'une part $\mathbb{P}(A, X_n = i) > 0$ et d'autre part la relation (5). Considérons une suite $(A_k)_{k\in\mathbb{N}}$ d'événements disjoints deux à deux de la tribu $\mathcal{A}_n$ qui vérifient la propriété $(\mathcal{H})$. Alors on peut établir que :

$$\begin{aligned}
\mathbb{P}(X_{n+1} = j, \bigcup_k A_k, X_n = i) &= \sum_k \mathbb{P}(X_{n+1}, A_k, X_n = i) \\
&= \sum_k \mathbb{P}(X_{n+1} = j \mid A_k, X_n = i)\mathbb{P}(A, X_n = i) \\
&= p_{i,j} \sum_k A_k \mathbb{P}(A_k, X_n = i) = p_{i,j}\mathbb{P}(\bigcup_k A_k, X_n = i)
\end{aligned}$$

avec pour argument d'une part la σ-additivité de $\mathbb{P}$, et d'autre part le fait que chaque A_k vérifie $(\mathcal{H})$. Dès lors, vu que $\mathbb{P}(\bigcup_k A_k, X_n = i) > 0$, la réunion $\sum_k A_k$ vérifie encore la propriété $(\mathcal{H})$. Il suffit d'utiliser la formule de probabilité conditionnelle et de divisier par la probabilité précédente pour le voir. Maintenant, on sait par la relation fondamentale définissant les chaînes de Markov, à savoir la relation (1), on sait que chaque événement de la forme $(X_0 = i_0, \dots, X_n = i_n)$ qui vérifie $\mathbb{P}(X_0 = i_0, \dots, X_n = i_n, X_n = i) > 0$ (ce qui impose notamment $i = i_n$ vérifie la relation $(\mathcal{H})$. Dès lors, vu qu'un événement de $\mathcal{A}_n$ vérifiant $\mathbb{P}(A, X_n = i) > 0$ est une réunion dénombrable d'événements de cette forme qui sont deux à deux disjoints, on en déduit qu'il vérifie également la propriété $(\mathcal{H})$ par les calculs menés précédemment. □

On a naturellement envie de prolonger les propriétés comme celle de Markov (1) à des événements de la tribu $\mathcal{A}_n$ comme on l'a fait dans la relation 5. Il faut pour cela que le présent soit de la forme $(X_n = i)$ avec i dans l'ensemble E des états. S'il est de la forme $(X_n \in B)$ avec B qui n'est pas un singleton, alors la formule n'est plus vraie. On peut regarder la marche aléatoire sur $\mathbb{Z}^2$ pour s'en convaincre. À noter que l'on peut aussi prolonger la formule (5) pour une suite d'événements ponctuels tous postérieurs au temps $(n + 1)$. C'est ce que décrit la proposition suivante.

Proposition 1.6

Soient $n \geqslant 0$ et $r \geqslant 1$ deux entiers et A un événement de la tribu $\mathcal{A}_n$. Si $\mathbb{P}(A, X_n = i) > 0$ alors :

$$\mathbb{P}(X_{n+1} = j_{n+1}, \dots, X_{n+r} = j_{n+r} \mid A, X_n = i_n) = p_{i_n, j_{n+1}} p_{j_n, j_{n+1}} \dots p_{j_{n+r-1}, j_{n+r}}. \tag{6}$$

Démonstration. La preuve est analogue à celle faite pour montrer la relation (5) mais en utilisant la relation (3). □

On peut alors établir en sommant sur toutes les valeurs possibles de $j_{n+1}, \dots, j_{n+r-1}$ que :

$$\mathbb{P}(X_{n+r} = j \mid A, X_n = i) = p_{i,j}^{(r)}.$$

En prenant $A = \Omega$, on en déduit que

$$\mathbb{P}(X_{n+r} = j \mid X_n = i) = p_{i,j}^{(r)}. \tag{7}$$

C'est une remarque importante, car cela montre que l'homogénéité supposée pour une transition du système faite en une étape reste vraie pour une transition en r étapes.

2 Propriétés sur les chaînes de Markov

2.1 Classification des états et décomposition en classes

Dans ce qui suit, on va préciser la description d'une chaîne de Markov en introduisant la notion de classe. Avant cela, introduisons un peu de formalisme. On a toujours en contexte une chaîne de Markov, ainsi que la donnée de son ensemble E de valeurs et $\mathcal{P}$ sa matrice de transition.

Définition 2.1

Soient i, j deux états. On dit que l'état j est *accessible* depuis l'état i, ou est conséquent de l'état i, s'il existe un entier $n \geqslant 0$ tel que $p_{i,j}^{(n)} > 0$. On note dans ce cas $i \to j$.

Proposition 2.1

La relation d'accessibilité entre deux états est réflexive et transitive.

Démonstration. Tout d'abord, on a toujours $p_{i,i}^{(0)} = \mathbb{P}(X_0 = i \mid X_0 = i) = 1$ donc par définition, $i \to i$. Supposons maitenant que pour $i, l, j \in E$, on ait $i \to l$ et $l \to j$. Alors il existe deux entiers $m, n > 0$ tels que $p_{i,l}^{(m)} > 0$ et $p_{l,j}^{(n)} > 0$. Dès lors, la formule (2) nous assure que

$$p_{i,j}^{(m+n)} = \sum_{k \in E} p_{i,k}^{(m)} p_{k,j}^{(n)} \geqslant p_{i,l}^{(m)} p_{l,j}^{(n)} > 0$$

prouvant alors que $i \to j$. □

Proposition 2.2

Soient i, j deux états. Les assertions suivantes sont équivalentes :

1. *L'état j est accessible depuis l'état i i.e. $i \to j$.*
2. *Le processus partant de i atteint j avec une probabilité strictement positive.*

Démonstration. Le sens 2. ⇒ 1. découle directement de la définition de la relation d'accessibilité. Prouvons que 1. ⇒ 2. par contraposition. Supposons donc que l'état j ne soit pas accessible depuis l'état i i.e. pour tout $n \geqslant 0$ on a $p_{i,j}^{(n)} = 0$. Soit A l'événement "le processus ne passe par l'état j". Alors la sous-additivité de $\mathbb{P}$ assure que :

$$\mathbb{P}(A \mid X_0 = i) = \mathbb{P}(\sum_{n \geqslant 0} X_n = j \mid X_0 = i) \leqslant \sum_{n \geqslant 0} \mathbb{P}(X_n = j \mid X_0 = i) = \sum_{n \geqslant 0} p_{i,j}^{(n)} = 0.$$

Donc il apparaît que $\mathbb{P}(A \mid X_0 = i) = 0$, ce qui conclut. □

Définition 2.2

Soient i, j deux états. On dit que i et j *communiquent* et on écrit $i \leftrightarrow j$ lorsque $i \to j$ et $j \to i$.

Proposition 2.3

La relation de communication entre les états définit une relation d'équivalence.

Démonstration. Les propriétés de transitivité et de réflexivité sont encore vérifiées, car vérifiées par la relation d'accessibilité définie plus tôt. De plus, la définition même de la relation de communication donne la symétrie. □

Remarque 3. En raison du fait que pour $i \in E$, $p_{i,i}^{(0)} = 1$, tout état communique avec lui même. On dit que l'état i est de retour lorsqu'il existe un entier $n > 0$ tel que $p_{i,i}^{(n)} > 0$. Il existe aussi des états tels que la probabilité précédente est nulle, et ce quelque soit l'entier $n > 0$. Ces états sont appelés états de non-retour.
C'est le cas par exemple dans la chaîne de Markov à deux états telle que $E = \{0, 1\}$ et $\mathcal{P} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$.
Ici l'état 1 est un état de non-retour. On rencontre aussi assez régulièrement des états dits absorbants. Ce sont les états i tels que pour tout $n \in \mathbb{N}$, $p_{i,i}^{(n)} = 1$. Dans l'exemple précédent, l'état 0 est absorbant.

D'autre part, le fait que la relation de communication soit une relation d'équivalence fournit naturellement une partition de l'ensemble E des états possibles de la chaîne de Markov selon les différentes classes d'équivalence. Il est à noter qu'elles sont disjointes en tant que classes d'équivalance, mais qu'il est possible qu'un état appartenant à l'une d'entre elles soit accessible depuis un état appartenant à une autre classe (au sens de la relation d'accessibilité définie précédemment). Par exemple, dans l'hypothèse où E se partitionne en deux classes C_1 et C_2, alors il est possible de passer, disons, de C_1 à C_2 mais dans ce cas, il est impossible de passer de C_2 à C_1 (sinon on aurait deux éléments communiquant au sein de deux classes disjointes ce qui est exclu).

Définition 2.3

On appelle *classes indécomposables* les classes d'équivalence de E pour la relation de communication. S'il n'y a qu'une unique classe d'équivalence, i.e. si tous les états communiquent entre eux, la chaîne est dite *irréductible*.

Exemple 1. Considérons la chaîne de Markov où l'ensemble des états est $E = \{0, 1, 2\}$, la matrice de transition est

$$\mathcal{P} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{pmatrix} \quad \text{et le graphe est} \quad 1/2 \circlearrowright \underset{0}{\bullet} \underset{1/2}{\overset{1/2}{\rightleftarrows}} \underset{1}{\overset{\overset{1/4}{\circlearrowleft}}{\bullet}} \underset{1/3}{\overset{1/4}{\rightleftarrows}} \underset{2}{\bullet} \circlearrowleft 2/3.$$

C'est une chaîne irréductible, car tous les états communiquent. Même si 0 et 2 ne sont pas directement reliés (i.e. $p_{0,2} = p_{2,0} = 0$) on peut atteindre l'état 2 depuis 0 en passant par 1 avec une probabilité strictement positive.

2.2 États récurrents et états transients

Commençons par énoncer une définition dans un contexte plus général que celui des chaînes de Markov.

Définition 2.4

Soit $(Y_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Soit également T une variable aléatoire définie sur ce même espace. On dit que T est un *temps d'arrêt* du processus $(Y_n)_{n\in\mathbb{N}}$ si les deux conditions suivantes sont vérifiées :

1. La variable aléatoire T est à valeurs dans $\mathbb{N} \cup \{+\infty\}$.
2. Pour tout $n \geqslant 0$, l'événement $\{T = n\}$ appartient à la tribu $\mathcal{A}_n = \sigma(Y_0, \dots, Y_n)$.

La condition 2 postule qu'au temps n, on ne peut décider de la réalisation ou non de l'événement $(T = n)$ qu'à partir de l'information fournie par le vecteur aléatoire $(Y_0, \dots, Y_n)$. Pour le dire autrement, l'événement $(T = n)$ ne dépend pas de l'évolution du processus postérieurement au temps n. On peut maintenant se replacer dans notre contexte de chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ dont l'ensemble des états accessibles est E avec pour matrice de transition $\mathcal{P}$.

Définition 2.5

Pour tout état $j \in E$, on désigne par T_j le *temps d'atteinte* de l'état j à partir de l'instant 1 de la chaîne $(X_n)_{n\in\mathbb{N}}$. Autrement dit,

$$T_j := \inf\{n \geqslant 1,\ X_n = j\}.$$

On définit également par récurrence des variables T_j^n par :

$$T_j^n := \inf\{n \geqslant T_j^{n-1},\ X_n = j$$

avec pour convention que $T_j^1 = T_j$. Ce sont les les temps de n-ième atteinte de l'état j.

Remarque 4. Ce temps d'attente est un temps d'arrêt de la chaîne au sens de la définition 2.4 où pour tout $n \geqslant 0$, l'évènement $(T_j = n) = (X_0 \neq j, \dots, X_{n-1} \neq j, X_n = j)$ ne dépend que du vecteur aléatoire $(X_0, \dots, X_n)$.

Cela nous permet maintenant de considérer un conditionnement selon le temps d'attente.

Définition 2.6

Pour tout couple d'états $(i, j) \in E^2$, et tout $n \geqslant 1$, on pose :

$$f_{i,j}^{(n)} := \mathbb{P}(T_j = n \mid X_0 = i)$$

Ainsi, $f_{i,j}^{(n)}$ pour $n \geqslant 1$ représente la probabilité que le processus, partant de l'état i, atteigne l'état j pour la première fois au temps n. Pour tout couple d'états $(i, j) \in E^2$, on pose par convention $f_{i,j}^{(0)} = 0$.

Théorème 2.1

Pour tout entier $n \geqslant 0$, on a l'identité :

$$p_{i,j}^{(n)} = \sum_{k=0}^{n} f_{i,j}^{(k)} p_{j,j}^{(n-k)}. \tag{8}$$

Puisque $p_{i,j}^{(0)} = \delta_{i,j}$ où $\delta_{i,j}$ désigne le symbole de Kronecker, la formule reste vraie pour $n = 0$ et $i \neq j$.

Démonstration. Une première approche intuitive à la preuve de ce résultat est de se dire que le processus passe de l'étape i à l'étape j en n étapes si et seulement si il passe de i à j en k étapes avec $0 \leqslant k \leqslant n$, et qu'il passe ensuite de j à j en $(n-k)$ étapes. Chacun de ces chemins pour des valeurs k distinctes sont disjoints, et la probabilité à k fixé d'avoir un chemin est de $f_{i,j}^{(k)} p_{i,j}^{(n-k)}$.
Plus formellement, prouvons cela par le calcul. Tout d'abord, notons que pour $n \geqslant 1$, on a :

$$(X_n = j) = \bigcup_{k=1}^{n-1} (T_j = k, X_n = j) \cup (T_j = n).$$

Ces événements étant tous disjoints, ils le sont aussi avec un conditionnement, et il vient pour $(i, j) \in E^2$:

$$\begin{aligned}
\mathbb{P}(X_n = j \mid X_0 = i) &= \sum_{k=1}^{n-1} \mathbb{P}(T_j = k, X_n = j \mid X_0 = i) + \mathbb{P}(T_j = n \mid X_0 = i) \\
&= \sum_{k=1}^{n-1} \mathbb{P}(T_j = k \mid X_0 = i) \mathbb{P}(X_n = j \mid T_j = k \mid X_0 = i) + f_{i,j}^{(n)}.
\end{aligned}$$

Or pour $1 \leqslant k \leqslant n-1$, l'événement $(T_j = k, X_0 = i)$ est de la forme $(A_k, X_0 = i)$ avec $A_k \in \mathcal{A}_{k-1} \subset \mathcal{A}_k$.
Dès lors, la formule (6) assure que :

$$\mathbb{P}(X_n = j \mid T_J = k, X_0 = i) = \mathbb{P}(X_n = j \mid A_k, X_k = j) = p_{j,j}^{(n-k)}.$$

D'où :

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i) = \sum_{k=1}^{n-1} f_{i,j}^{(k)} p_{j,j}^{(n-k)} + f_{i,j}^{(n)} = \sum_{k=0} f_{i,j}^{(k)} p_{j,j}^{(n-k)}.$$

□

Remarque 5. Cette identité permet de calculer les $f_{i,j}^{(n)}$ par récurrence à partir des $p_{i,j}^{(n)}$:

$$f_{i,j}^{(1)} = p_{i,j}^{(1)} \quad ; \quad f_{i,j}^{(n)} = p_{i,j}^{(n)} - \sum_{k=1}^{n-1} f_{i,j}^{(k)} p_{j,j}^{(n-k)} \quad , \forall n \geqslant 2.$$

Définition 2.7

Pour $(i, j) \in E^2$, on définit :

$$f_{i,j} := \mathbb{P}(T_j < +\infty \mid X_0 = i) = \sum_{n=1}^{+\infty} f_{i,j}^{(n)}. \tag{9}$$

C'est la probabilité pour que le processus partant de i passe au moins une fois par l'état j au cours de son évolution. Dans le cas où $i = j$, la quantité $f_{j,j}$ désigne alors la probabilité que le système partant de l'état j repasse au moins une fois par ce même état.

Définition 2.8

On dit que l'état j est *récurrent* lorsque $f_{j,j} = 1$. On dit qu'il est *transient* ou transitoire lorsque $f_{j,j} < 1$.

Théorème 2.2

L'état j est récurrent si et seulement si

$$\sum_{n=0}^{+\infty} p_{j,j}^{(n)} = +\infty.$$

Il est transient sinon.

Démonstration. Pour prouver ce résultat, nous allons nous appuyer sur des fonctions génératrices, ainsi que sur le lemme d'Abel sous la forme suivante que nous ne prouverons pas ici :

Lemme 2.1 (*Admis*)

Soient $(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ et $\alpha \in \mathbb{R}$.

1. *Si la serie de terme général α_n converge et a pour somme α alors $\lim_{s\to 1^-} \sum_{n=0}^{+\infty} \alpha_n s^n = \alpha$.*
2. *Si les α_n sont positifs et si $\lim_{s\to 1^-} \sum_{n=0}^{+\infty} \alpha_n s^n = \alpha \in \overline{\mathbb{R}}$ alors la série de terme général α_n converge et a pour somme α.*

Introduisons donc les fonctions génératrices suivantes :

$$P_{i,j}(s) := \sum_{n=0}^{+\infty} p_{i,j}^{(n)} s^n \quad \text{et} \quad F_{i,j}(s) := \sum_{n=0}^{+\infty} f_{i,j}^{(n)} s^n.$$

La relation récurrente (8) donnée dans le théorème 2.1 permet d'établir la relation fonctionnelle suivante :

$$P_{i,j}(s) = \delta_{i,j} + F_{i,j}(s) P_{j,j}(s).$$

Le théorème se prouve alors de la manière suivante : supposons j récurrent i.e. $\sum_{n\geqslant 0} f_{j,j}^{(n)}$ est de somme 1.
Le premier point du lemme d'Abel assure alors que

$$\lim_{s\to 1^-} \sum_{n=0}^{+\infty} f_{j,j}^{(n)} s^n = \lim_{s\to 1^-} F_{j,j}(s) = 1$$

ce qui grâce à la relation fonctionnelle précédente donne que :

$$\lim_{s\to 1^-} P_{j,j}(s) = \lim_{s\to 1^-} \frac{1}{1 - F_{j,j}(s)} = +\infty.$$

La deuxième partie du lemme d'Abel permet alors de conclure que la série $\sum_{n\geqslant 0} p_{j,j}^{(n)}$ diverge. De la même façon, si j est supposé transient, alors $\sum_{n\geqslant 0} f_{j,j}^{(n)}$ est de somme strictement plus petite que 1 et le même procédé que ci-dessus permet de prouver que

$$\lim_{s\to 1^-} P_{j,j}(s) < +\infty$$

et donc par le second point du lemme d'Abel, que la série $\sum_{n\geqslant 0} p_{j,j}^{(n)}$ est convergente. □

Proposition 2.4

Tout état de non-retour est transient. Tout état absorbant est récurrent.

Démonstration. Pour un état j de non-retour, on a $p_{j,j}^{(n)} = \delta_{n,0}$ donc il est évident que la série de terme général $p_{j,j}^{(n)}$ est convergente. Par le théorème 2.2, j est donc un état transient. Pour un état j absorbant, tous les termes $p_{j,j}^{(n)}$ valent 1, donc la série de terme général $p_{j,j}^{(n)}$ est grossièrement divergente, et l'état j est donc récurrent. □

Proposition 2.5

La récurrence est une propriété de classe i.e. si $i \leftrightarrow j$ alors i est récurrent si et seulement si j est récurrent.

Démonstration. Étant donné que $i \leftrightarrow j$, on a l'existence de deux entiers n_1, n_2 tels que $p_{i,j}^{(n_1)} > 0$ et $p_{j,i}^{(n_2)} > 0$. On a alors grâce à la formule (2) :

$$\sum_{n=0}^{+\infty} p_{j,j}^{(n)} \leqslant \sum_{n=0}^{+\infty} p_{j,j}^{(n+n_2+n_1)} \leqslant \sum_{n=0}^{+\infty} p_{j,i}^{(n_2)} p_{i,i}^{(n)} p_{i,j}^{(n_1)} = p_{j,i}^{(n_2)} p_{i,j}^{(n_1)} \sum_{n=0}^{+\infty} p_{i,i}^{(n)}.$$

On peut de la même manière établir que :

$$\sum_{n=0}^{+\infty} p_{i,i}^{(n)} \leqslant p_{j,i}^{(n_2)} p_{i,j}^{(n_1)} \sum_{n=0}^{+\infty} p_{j,j}^{(n)}.$$

On voit alors par le critère de majoration que $\sum_{n\geqslant 0} p_{i,i}^{(n)}$ diverge si et seulement si $\sum_{n\geqslant 0} p_{j,j}^{(n)}$ converge i.e. par le théorème 2.2, i est récurrent si et seulement si j l'est. □

Remarque 6. Cela légitime le fait de parler de classes récurrentes et de classes transientes. En fait, on pourrait même affaiblir les hypothèses. Si j est seulement accessible depuis i et que i est récurrent alors j l'est aussi.

Proposition 2.6

Si une chaîne de Markov a un nombre fini d'états, alors au moins l'un d'entre eux est récurrent.

Démonstration. Supposons que la chaîne n'ait qu'un nombre fini N d'états. On sait que pour tout $n \in \mathbb{N}$, pour tout $i \in E$ (qui est donc de cardinal N), alors $\sum_{j\in E} p_{i,j}^{(n)} = 1$, donc pour tout $n \in \mathbb{N}$:

$$\sum_{1\leqslant i,j\leqslant N} p_{i,j}^{(n)} = N.$$

Dès lors, il vient que :

$$\sum_{n=0}^{+\infty} \sum_{1\leqslant i,j\leqslant N} p_{i,j}^{(n)} = +\infty.$$

Par le critère de récurrence du théorème 2.2, on en déduit qu'au moins un état est récurrent. □

Remarque 7. À noter que ce résultat ne tient plus si la chaîne possède une infinité d'états. La chaîne qui a pour ensemble des états possibles $E = \mathbb{Z}$ et pour matrice de transition $\mathcal{P} = (p_{i,j})_{(i,j)\in\mathbb{Z}^2}$ où $p_{i,i+1} = 1$ et $p_{i,j} = 0$ dès que $j \neq i+1$ ne possède que des états transients.

2.3 Propriété de Markov forte

On a prouvé plus haut la formule (6) qui assure que pour $A \in \mathcal{A}_n$ tel que $\mathbb{P}(A, X_n = i_n) > 0$, on a :

$$\mathbb{P}(X_{n+1} = j_{n+1}, \dots, X_{n+r} = j_{n+r} \mid A, X_n = i_n) = p_{i_n,j_{n+1}} p_{j_n,j_{n+1}} \dots p_{j_{n+r-1},j_{n+r}}. \tag{10}$$

Cela traduit à la fois les caractères markovien, et homogène de la chaîne. L'idée ici va être d'étendre cette formule en ne considérant plus un instant présent $(X_n = i_n)$ mais un instant aléatoirement fixé $(X_T = i)$ où T désigne un temps d'arrêt de la chaîne. Avant cela, introduisons une nouvelle tribu.

Définition 2.9

On note $\mathcal{A}_T$ l'ensemble constitué des événements $A \in \mathcal{A}$ vérifiant que pour tout $n \in \mathbb{N}$, $A \cap (T = n) \in \mathcal{A}_n$.

Proposition 2.7

L'ensemble $\mathcal{A}_T$ décrit ci-dessus est une tribu.

Démonstration. Tout d'abord, on a que $\Omega \cap (T = n) = (T = n) \in \mathcal{A}_n$ car T est un temps d'arrêt. Ainsi, $\Omega \in \mathcal{A}_T$. Ensuite, soit $(A_m)_{m\in\mathbb{N}}$ une suite d'événements de $\mathcal{A}_T$, on a donc $\bigcup_{m\in\mathbb{N}} A_m \cap (T = n) = \bigcup_{m\in\mathbb{N}} \Big(A_m \cap (T = n)\Big)$ qui une union d'événements de $\mathcal{A}$, ainsi c'est un événement de cette dernière i.e. $\bigcup_{m\in\mathbb{N}} A_m \in \mathcal{A}_T$. Pour $A \in \mathcal{A}_T$, $A^c \cap (T = n) = (T = n)\backslash(A \cap (T = n))$ qui est dans $\mathcal{A}$ par intersection. Finalement $A^c \in \mathcal{A}_T$ ce qui conclut. □

Théorème 2.3

Soit T un temps d'arrêt de la chaîne de Markov $(X_n)_{n\in\mathbb{N}}$ prenant ses valeurs dans E, et ayant pour matrice de transition $\mathcal{P} = (p_{i,j})_{(i,j)\in E^2}$. Soit $A \in \mathcal{A}_T$ tel que $\mathbb{P}(T < +\infty, A, X_T = i) > 0$. Alors :

$$\mathbb{P}(X_{T+1} = j_1, \dots, X_{T+r} = j_r \mid A, X_T = i) = p_{i,j_1} p_{j_1,j_2} \dots p_{j_{r-1},j_r}. \tag{11}$$

Démonstration. Posons $A' := (T < +\infty, A, X_T = i)$, $B_n := (X_{n+1} = j_1, \dots, X_{n+r} = j_r)$ pour $n \geqslant 0$ puis $B := (X_{T+1} = j_1, \dots, X_{T+r} = j_r)$. Nous allons prouver que $\mathbb{P}(A', B) = p_{i,j_1} p_{j_1,j_2} \dots p_{j_{r-1},j_r} \mathbb{P}(A')$. D'abord,

$$\mathbb{P}(A', B) = \sum_{n=0}^{+\infty} \mathbb{P}(T = n, A, X_T = i, B) = \sum_{n=0}^{+\infty} \mathbb{P}(B_n \mid T = n, A, X_n = i) \mathbb{P}(T = n, A, X_n = i).$$

En fait, il n'apprait dans la somme ci-dessus que des termes tels que $\mathbb{P}(T = n, A, X_n = i)$ pour lesquels on peut directement utiliser la relation (10) donnée en introduction de cette section. On a donc :

$$\mathbb{P}(B_n \mid T = n, A, X_n = i) = \mathbb{P}(B_n \mid X_n = i) = p_{i,j_1} p_{j_1,j_2} \dots p_{j_{r-1},j_r}.$$

On peut donc conclure :

$$\mathbb{P}(A', B) = p_{i,j_1} p_{j_1,j_2} \dots p_{j_{r-1},j_r} \sum_{n=0}^{+\infty} \mathbb{P}(T = n, A, X_n = i) = p_{i,j_1} p_{j_1,j_2} \dots p_{j_{r-1},j_r} \mathbb{P}(T < +\infty, A, X_T = i)$$

□

Introduisons maintenant quelques nouvelles variables aléatoires. Tout d'abord le nombre de fois où le processus passe par l'état j entre les temps 1 et n :

$$N_n(j) := \sum_{k=1}^{n} \mathbb{1}_{(X_k = j)}.$$

Ensuite la fréquence de passage en j entre les temps 1 et n, $F_n(j) := \frac{1}{n} N_n(j)$.

Théorème 2.4

Dans une chaîne de Markov homogène, il n'y a que deux sortes d'états possibles : les états récurrents et les états transients. De plus,

1. *Soit j un état transient et $a = f_{j,j}$ la probabilité d'au moins un retour de la chaîne dans l'état j. Alors $a < 1$ et le nombre aléatoire N_j de retours de la chaîne dans l'état j suit la loi géométrique de paramètre $1 - a$:*

$$\mathbb{P}(N_j = k) = (1 - a) a^{k-1}.$$

2. *Si l'état j est récurrent, alors la suite $(X_n)_{n \in \mathbb{N}}$ partant de j revient presque sûrement une infinité de fois à son état initial :*

$$\mathbb{P}(N_j = +\infty) = 1.$$

Démonstration. Pour le premier point, il s'agit de voir que $(N_j = k) = (T_j < +\infty, \dots, T_j^{k-1} < +\infty, T_j^k = +\infty)$, i.e. $\mathbb{P}(N_j = k) = a^{k-1}(1 - a)$ par indépendance. Le second point repose sur le fait que lorsque j est récurrent, les variables $(T_j^n)_{n \geqslant 1}$ sont toutes finies presque sûrement, de sorte que la chaîne repasse par l'état j une infinité de fois avec une probabilité 1. □

Proposition 2.8 (*Admis*)

Soit j un état récurrent et $k \neq j$ tel que $i \to k$. Alors $k \to j$ de sorte que k est aussi récurrent, et dans la même classe que j. En particulier, une chaîne de Markov ne peut pas passer d'un état récurrent à un état transient.

Proposition 2.9 (*Admis*)

Notons Rec l'ensemble des états récurrents dans une chaîne de Markov. Alors, avec une probabilité de 1, la chaîne se trouve dans Rec au bout d'un temps fini.

2.4 Périodicité

On va ici étudier dans quelles conditions le temps qui sépare deux retours au même état est, ou n'est pas un multiple d'un temps minimum. On introduit pour cela la notion de période.

Définition 2.10

Soit j un état récurrent. On appelle *période* de j et on note $d(j)$ le pgcd de tous les entiers $n \geqslant 1$ tels que $p_{j,j}^{(n)} > 0$. Si $d(j) = d \geqslant 2$, on dit que j est *périodique* de période d. Si $d = 1$, on dit que j est *apériodique*. Enfin, si j est un état de non-retour i.e. $p_{i,j}^{(n)} = 0$ pour tout $n \in \mathbb{N}$, on pose $d(j) = +\infty$.

Théorème 2.5

La periodicité est une propriété de classe. Si deux états i et j communiquent, et que i est périodique de période d alors j est aussi périodique de période d.

Démonstration. Supposons que $i \leftrightarrow j$. Alors il existe deux entiers n et m tels que $p_{i,j}^{(n)} > 0$ et $p_{j,i}^{(m)} > 0$. Supposons aussi i périodique de période $d(i) = d$. Alors il existe un entier $s \geqslant 1$ tel que $p_{i,i}^{(s)} > 0$. La formule (2) nous assure que :

$$p_{j,j}^{(n+m+s)} \geqslant p_{j,i}^{(m)} p_{i,i}^{(s)} p_{i,j}^{(n)} > 0.$$

De plus, $p_{i,i}^{(s)} > 0 \to p_{i,i}^{(2s)} > 0$, donc on peut ausssi obtenir que $p_{j,j}^{(n+m+2s)} > 0$.
Dès lors, la période $d(j)$ de j divise $m + 2s + n$ et $m + s + n$, donc leur différence i.e. divise s et en particulier $d(i)$. En suivant le même raisonnement, on peut aussi montrer que $d(i)$ divise $d(j)$, ce qui conclut. □

2.5 Loi stationnaire

Pour rappel, nous avons parlé dans les parties précédentes de temps d'attente de la chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ en introduisant les variables aléatoires

$$T_j = \inf\{n \geqslant 1,\ X_n = j\}.$$

On a posé également $f_{i,j}^{(n)} = \mathbb{P}(T_j = n \mid X_0 = i) = \mathbb{P}^i(T_j = n)$ en prenant pour loi initiale de la chaîne la loi singulière ε_i. L'espérance mathématique de la variable T_j pour la loi $\mathbb{P}^i$ est notée $M_{i,j} := \mathbb{E}[T_j]$. C'est un temps d'attente moyen pour atteindre j partant de i. En particulier, la quantité $M_{i,i}$ est appelée temps de retour moyen en i, et son inverse $1/M_{i,i}$ est une fréquence moyenne de retour en i. Si $M_{i,i} = +\infty$, on pose $1/M_{i,i} = 0$.

Définition 2.11

On dit que l'état i est *positif* lorsque $M_{i,i} < +\infty$. Lorsque $M_{i,i} = +\infty$, on dit alors que que l'état i est *nul*.

Théorème 2.6 (*Admis*)

Soit i un état.

1. *Si* $\limsup_{n \to +\infty} p_{i,i}^{(n)} > 0$ *alors l'état i est positif.*
2. *Si* $\lim_{n \to +\infty} p_{i,i}^{(n)} = 0$ *alors l'état i est nul.*

Remarque 8. Tout état transient est nul.

Théorème 2.7

Les propriétés de nullité et de positivité des états est une relation de classe pour la relation de communication.

Démonstration. C'est en fait une conséquence du critère de positivité donné dans le théorème 2.6. De fait, si j est un état nul communiquant avec un état i, , alors il existe deux entiers n_1 et n_2 tels que $p_{i,j}^{(n_1)} > 0$ et $p_{j,i}^{(n_2)} > 0$. Dès lors, pour tout $n \geqslant 0$, on a $p_{j,j}^{(n_2+n+n_1)} \geqslant p_{i,j}^{(n_1)} p_{j,i}^{(n_2)} p_{i,i}^{(n)}$, donc par nullité de j, $\lim_{n \to +\infty} p_{j,j}^{(n)} = 0$, donc $\lim_{n \to +\infty} p_{i,i}^{(n)} = 0$ et donc i est nul. Par contraposition, si i est supposé non nul, alors la relation ci-dessus donne que j l'est aussi ce qui conclut. □

Ce formalisme nous sera utile pour la suite avec la notion de loi stationnaire dont on va maitenant parler. On prend par commodité la convention de noter une loi de probabilité u sur l'ensemble dénombrable E comme un vecteur ligne $(u_0, u_1, \dots)$ où la somme des composantes u_i vaut 1, chaque u_i étant positif ou nul. On a toujours en contexte une chaîne de Markov prenant ses états dans E, et de matrice de transition $\mathcal{P}$.

Définition 2.12

Une loi de probabilité u sur E est dite *stationnaire* si $u = u.\mathcal{P}$, ou encore de manière équivalente si pour tout $j \in E$:

$$u_j = \sum_{i \in E} u_i p_{i,j}$$

Si on suppose qu'il existe une loi de probabilité stationnaire u on a que pour tout $n \in \mathbb{N}$, $u = u.\mathcal{P}^n$, de sorte que si on prend u comme loi initiale pour la chaîne de Markov $(X_n)_{n \in E}$ à valeurs dans E, de matrice de transition $\mathcal{P}$, alors pour tout $k \geqslant 0$, les vecteurs aléatoires $(X_0, \dots X_n)$ et $(X_k, \dots, X_{n+k})$ ont la même loi.

Remarque 9. À noter que l'existence d'une loi stationnaire n'est pas un fait. Par exemple, avec $E = \mathbb{N}$ et le graphe

$$0 \xrightarrow{1} 1 \xrightarrow{1} 2 \xrightarrow{1} \dots$$

le système $\sum_{i \in E} u_i p_{i,j} = u_j$ se réduit en fait à $u_0 = 0$ et pour tout $j \geqslant 1$, $u_j = u_{j-1}$ de sorte que les u_j sont tous nuls. On a donc pas une loi de probabilité. Par ailleurs, il peut aussi exister une infinité de lois de probabilité stationnaires. Considérons par exemple de nouveau le modèle de la ruine du joueur présenté plus haut, alors pour tout $\alpha \in]0,1[$, la loi $u = (\alpha, 0, \dots, 0, 1-\alpha)$ est une loi de probabilité stationnaire. La théorie concernant l'existence et l'unicité de lois de probabilité stationnaire est complète lorsque l'ensemble des états E est fini. Nous nous intéresserons dans notre cas à un ensemble E au plus dénombrable.

On peut alors énoncer un dernier résultat, que l'on appellera théorème ergodique.

Théorème 2.8 (*Ergodique*)

Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov irréductible. Alors :

1. *Pour presque tout $j \in E$,*
$$\frac{1}{n} N_n(j) \underset{n \to +\infty}{\overset{p.s}{\to}} \frac{1}{\mathbb{E}[T_j]}.$$
2. *Les assertions suivantes sont équivalentes :*
   (a) *Il existe une unique probabilité invariante.*
   (b) *La quantité $u = (u_i)_{i \in E}$ où $u_j = \mathbb{E}[T_j]^{-1}$ est la probabilité invariante.*
   (c) *Tous les états sont récurrents positifs.*
   (d) *Un état est récurrent positif.*

Démonstration. Le point 1 est en fait une extension du même théorème que l'on peut prouver dans le cas d'une chaîne de Markov dont l'ensemble des états est fini. Nous ne reviendrons pas sur ce point ici. Pour plus de précisions il est possible de se référer à BENAÏM et EL KAROUI 2004, p. 86.

Prouvons le second point. On va prouver $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$. Soit $u = (u_i)_{i \in E}$ une probabilité stationnaire et supposons que X_0 suive la loi u. Alors pour tout $n \geqslant 1$ c'est aussi le cas de X_n de sorte que :

$$\mathbb{E}\left[\frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{(X_k = j)}\right] = u_j.$$

Enfin, l'application du théorème de convergence dominée dans la relation du point 1 donne que $u_j = \frac{1}{\mathbb{E}[T_j]}$, et comme les u_i sont tous strictement positifs, il vient que tous les états sont récurrents positifs. On a donc prouvé que $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d)$. Supposons maintenant l'existence d'un état j récurrent positif et un état i quelconque de E. Alors les variables

$$\sum_{k=T_j^n+1}^{T_j^{n+1}} \mathbb{1}_{(X_k = i)}, \ \ n \geqslant 0$$

sont positives, et i.i.d. Par la loi des grands nombres, on a

$$\frac{1}{n}\sum_{s=1}^{T_j^n} \mathbb{1}_{(X_s = i)} \underset{n \to +\infty}{\to} \mathbb{E}\left[\sum_{s=1}^{T_j} \mathbb{1}_{(X_s = i)}\right].$$

Mais d'autre part,

$$\frac{1}{n}\sum_{s=1}^{T_j^n} \mathbb{1}_{(X_s = i)} = \frac{T_j^n}{n}\sum_{s=1}^{T_j} \frac{1}{T_j^n} \mathbb{1}_{(X_s = i)} \underset{n \to +\infty}{\to} \frac{\mathbb{E}[T_j]}{\mathbb{E}[T_i]}.$$

Or

$$\sum_{i \in E} \mathbb{E}\left[\sum_{s=1}^{T_j^n} \mathbb{1}_{(X_s = i)}\right] = \mathbb{E}[T_j]$$

de sorte que

$$\sum_{i \in E} \frac{1}{\mathbb{E}[T_i]} = 1.$$

Donc $u_i = \mathbb{E}[T_i]^{-1}$ définit bien une probabilité. Il reste encore à prouver que c'est une probabilité stationnaire.
Par application du théorème de convergence dominée à la relation du point 1, on obtient que :

$$\lim_{n\to+\infty} \sum_{k=1}^{n} p_{k,j}^{n} = u_j.$$

Dès lors en se donnant toujours un autre état i,

$$\sum_{j\in E} u_j p_{j,i} = \sum_{j\in E} \lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} p_{k,j}^{n} p_{j,i} \leqslant \lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} p_{k,j}^{n+1}.$$

Or on remarque par télescopage que :

$$\left|\frac{1}{n}\sum_{k=1}^{n} p_{k,j}^{n+1} - \frac{1}{n}\sum_{k=1}^{n} p_{k,j}^{n}\right| \leqslant \frac{2}{n}.$$

Par passage à la limite,

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} p_{k,j}^{n+1} = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} p_{k,j}^{n} = u_j.$$

Donc les inégalités $\sum_{j\in E} u_j p_{j,i} \leqslant u_i$ sont en fait toutes des égalités, ce qui conclut. □

3 Marche aléatoire

3.1 Outils d'analyse

3.1.1 Équivalent de Karamata

Théorème 3.1 (*Équivalent de Karamata*)

On considère une suite réelle $(a_k)_{k\in\mathbb{N}}$ positive telle que la série de terme général $a_k x^k$ converge absolument pour $x \in]-1,1[$. Ainsi on note $f(x)$ la somme de cette série et l'on suppose que

$$\lim_{x\to 1^-} \sqrt{1-x}\, f(x) = \sqrt{\pi}.$$

On a alors :

$$\sum_{k=0}^{n} a_k \underset{n\to+\infty}{\sim} 2\sqrt{n}.$$

Démonstration. ETAPE 1 : Commençons par montrer que :

$$\lim_{x\to 1^-} \sqrt{1-x} \sum_{k=0}^{+\infty} a_k x^{(p+1)k} = \sqrt{\frac{\pi}{p+1}}.$$

En effet, pour tout $x \in]-1,1[$ et $p \in \mathbb{N}$, on a $x^p \in]-1,1[$ et

$$\begin{aligned}
\sqrt{1-x} \sum_{k=0}^{+\infty} a_k x^{(p+1)k} &= \sqrt{1-x}\, f(x^{p+1}) \\
&= \frac{\sqrt{1-x}}{\sqrt{1-x^{p+1}}} \sqrt{1-x^{p+1}}\, f(x^{p+1}) \\
&= \frac{1}{\sqrt{1+x+\cdots+x^p}} \sqrt{1-x^{p+1}}\, f(x^{p+1}) \underset{x\to 1^-}{\longrightarrow} \sqrt{\frac{\pi}{p+1}}.
\end{aligned}$$

ETAPE 2 : ÉGALITÉ DE KARAMATA SUR LES POLYNÔMES

Remarquons ensuite que

$$\lim_{x\to 1^-} \sqrt{1-x} \sum_{k=0}^{+\infty} a_k x^{(p+1)k} = \int_0^{+\infty} \frac{e^{-(p+1)t}}{\sqrt{t}}\, dt.$$

En utilisant le changement de variable bijectif et $\mathcal{C}^1$ $t = x^2/(p+1)$ on se ramène à l'intégrale de Gauss pour en déduire que l'intégrale est égale à $\sqrt{\frac{\pi}{p+1}}$ ce qui conclut.
On étend ce résultat par linéarité aux polynômes :

$$\lim_{x\to 1^-} \sqrt{1-x} \sum_{k=0}^{+\infty} a_k x^k Q(x^k) = \int_0^{+\infty} \frac{e^{-t} Q(e^{-t})}{\sqrt{t}}\, dt.$$

ETAPE 3 : ÉGALITÉ DE KARAMATA

Soit h la fonction définie, pour tout $x \in [0,1]$, par :

$$h(x) = \begin{cases} 0 & \text{si } x \in [0, e^{-1}[\\ \frac{1}{x} & \text{si } x \in [e^{-1}, 1] \end{cases}$$

On admet l'égalité de Karamata :

$$\lim_{x \to 1^-} \sqrt{1-x} \sum_{k=0}^{\infty} a_k x^k h(x^k) = \int_0^{+\infty} \frac{e^{-t} h(e^{-t})}{\sqrt{t}} \, dt.$$

Elle se prouve en étendant le résultat du corollaire aux fonctions continues par morceaux en utilisant notamment le théorème de Stone-Weierstrass et le fait que les a_n sont positifs mais la preuve est trop technique pour rester dans le cadre de notre étude.

ETAPE 4 : CONCLUSION

Soit $n \in \mathbb{N}^*$, utilisons l'égalité de Karamata pour $x = e^{-\frac{1}{n}}$. Par définition de h, on a

$$\sqrt{1-e^{-1/n}} \sum_{k=0}^{+\infty} a_k e^{-k/n} h(e^{-k/n}) = \sqrt{1-e^{-1/n}} \sum_{k=0}^{n} a_k.$$

Quand $n \to +\infty$, $e^{-1/n} \to 1^-$ et l'égalité de Karamata donne

$$\lim_{n \to +\infty} \sqrt{1-e^{-1/n}} \sum_{k=0}^{n} a_k = \int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}} h(e^{-t}) \, dt = \int_0^1 \frac{dt}{\sqrt{t}} = \left[2\sqrt{t}\right]_0^1 = 2.$$

Donc

$$\sum_{k=0}^{n} a_k \underset{n \to +\infty}{\sim} \frac{2}{\sqrt{1-e^{-1/n}}} \underset{n \to +\infty}{\sim} 2\sqrt{n}.$$

□

3.1.2 Théorème Taubérien

Théorème 3.2 (*Taubérien*)

Soit $(a_n)_{n \in \mathbb{N}}$ une suite décroissante de réels positifs. Pour tout entier naturel n, on pose $S_n = \sum_{k=0}^{n} a_k$. Si

$$S_n \underset{n \to +\infty}{\sim} 2\sqrt{n},$$

Alors

$$a_n \underset{n \to +\infty}{\sim} \frac{1}{\sqrt{n}}.$$

On notera $[x]$ la partie entière d'un réel x.

Démonstration. ETAPE 1 : ENCADREMENT DE a_n

Tout d'abord on a pour $0 < \alpha < 1 < \beta$:

$$\frac{S_{[\beta n]} - S_n}{[\beta n] - n} \leqslant a_n \leqslant \frac{S_n - S_{[\alpha n]}}{n - [\alpha n]}.$$

En effet, comme $n \geq [\alpha n]$, on a

$$S_n - S_{[\alpha n]} = \sum_{k=[\alpha n]+1}^{n} a_k \geqslant (n - [\alpha n]) a_n$$

car la suite (a_k) est décroissante. On peut ensuite diviser par $n - [\alpha n] > 0$ pour obtenir

$$a_n \leqslant \frac{S_n - S_{[\alpha n]}}{n - [\alpha n]}.$$

De même pour l'autre inégalité.

ETAPE 2 : LIMITE DES BORNES ENCADRANTES
Pour $\gamma > 0$ différent de 1, on écrit que

$$\sqrt{n}\frac{S_n - S_{[\gamma n]}}{n - [\gamma n]} = \frac{1}{1 - \frac{[\gamma n]}{n}} \left(\frac{S_n}{\sqrt{n}} - \frac{S_{[\gamma n]}}{\sqrt{n}} \right).$$

Or

$$\lim_{n \to +\infty} \frac{[\gamma n]}{n} = \gamma.$$

Comme $[\gamma n] \to +\infty$, l'hypothèse sur la suite (S_n) est : $S_{[\gamma n]} \sim 2\sqrt{[\gamma n]}$ soit en divisant par $\sqrt{n}$:

$$\frac{S_{[\gamma n]}}{\sqrt{n}} \sim 2\sqrt{\frac{[\gamma n]}{n}}.$$

Et ainsi par la limite calculée ci-dessus

$$\lim_{n \to +\infty} \frac{S_{[\gamma n]}}{\sqrt{n}} = 2\sqrt{\gamma}.$$

Et alors

$$\sqrt{n}\frac{S_n - S_{[\gamma n]}}{n - [\gamma n]} \xrightarrow[n \to +\infty]{} \frac{2(1 - \sqrt{\gamma})}{1 - \gamma}.$$

On a alors montré que pour $0 < \alpha < 1 < \beta$:

$$\frac{2(1 - \sqrt{\beta})}{1 - \beta} \leqslant \lim_{n \to +\infty} \sqrt{n} a_n \leqslant \frac{2(1 - \sqrt{\alpha})}{1 - \alpha}.$$

ETAPE 3 : CONCLUSION
Or fixons $\varepsilon > 0$. Comme $\frac{2(1-\sqrt{x})}{1-x} = \frac{2}{1+\sqrt{x}}$ est de limite 1 en 1, il existe $0 < \alpha < 1 < \beta$ tels que

$$1 - \varepsilon \leqslant \frac{2(1 - \sqrt{\beta})}{1 - \beta} \text{ et } \frac{2(1 - \sqrt{\alpha})}{1 - \alpha} \leqslant 1 + \varepsilon.$$

On applique alors le résultat de l'étape 2 à ces α, β et on obtient un rang n_0 tel que :

$$\forall n \geqslant n_0, \ 1 - 2\varepsilon \leqslant \frac{2(1 - \sqrt{\beta})}{1 - \beta} - \varepsilon \leqslant \sqrt{n} a_n \leqslant \frac{2(1 - \sqrt{\alpha})}{1 - \alpha} + \varepsilon \leqslant 1 + 2\varepsilon.$$

Par définition des limites, on a donc $\sqrt{n} a_n \to 1$ et donc

$$a_n \underset{n \to +\infty}{\sim} \frac{1}{\sqrt{n}}.$$

□

3.2 Marche aléatoire en dimension 1

On considère $\Omega = \mathbb{Z}^{\mathbb{N}^*}$ l'ensemble des suites indexées par $\mathbb{N}^*$ à valeurs dans $\mathbb{Z}$.
On définit les applications coordonnées, pour tout $i \geq 1$,

$$X_i \ : \ \omega = (\omega_1, \omega_2, \dots) \in \Omega \mapsto \omega_i \in \mathbb{Z}.$$

On admet que l'on peut construire une tribu $\mathcal{B}$ et une mesure de probabilité $\mathbb{P}$ sur Ω, de sorte que les X_i soient des variables aléatoires, indépendantes et de loi de Rademacher :

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}.$$

On définit la suite de variables aléatoires $(S_n, n \geq 0)$ par :

$$S_0 = 0, \ S_n = \sum_{i=1}^{n} X_i.$$

On définit enfin la variable aléatoire T du plus petit retour en 0 par $T(\Omega) = \mathbb{N}^* \cup \{+\infty\}$:

$$T = \begin{cases} +\infty & \text{si } S_n \neq 0, \ \forall n \geq 1 \\ \inf\{n \geq 1, \ S_n = 0\} & \text{sinon} \end{cases}.$$

Pour tout entier naturel n, on note $E_n = \{T > n\}$, et pour $n \geq 1$ et $k \in \{0, \dots, n-1\}$,

$$A_k^n = \{S_k = 0\} \cap \bigcap_{i=k+1}^{n} \{S_i \neq 0\}.$$

Proposition 3.1

Soit $1 \leq k < n$.

1. *Pour tout* $(i_1, \dots, i_{n-k}) \in \{-1, 1\}^{n-k}$,
$$\mathbb{P}(X_{k+1} = i_1, \dots, X_n = i_{n-k}) = \mathbb{P}(X_1 = i_1, \dots, X_{n-k} = i_{n-k}).$$
2. *Pour tout* $(j_1, \dots, j_{n-k}) \in \mathbb{Z}^{n-k}$,
$$\mathbb{P}(S_{k+1} - S_k = j_1, \dots, S_n - S_k = j_{n-k}) = \mathbb{P}(S_1 = j_1, \dots, S_{n-k} = j_{n-k}).$$
3. *On a*
$$\mathbb{P}(A_k^n) = \mathbb{P}(S_k = 0)\mathbb{P}(E_{n-k}).$$
4. *Pour tout réel* x *de* $]0, 1[$*, on a l'égalité :*
$$\frac{1}{1-x} = \left(\sum_{n=0}^{\infty} \mathbb{P}(S_n = 0)x^n\right)\left(\sum_{n=0}^{\infty} \mathbb{P}(E_n)x^n\right).$$

Démonstration. 1. Comme les (X_i) sont indépendants et de même loi on a :

$$\mathbb{P}(X_{k+1} = i_1, \dots, X_n = i_{n-k}) = \prod_{j=k+1}^{n} \mathbb{P}(X_j = i_{j-k}) \prod_{j=1}^{n-k} \mathbb{P}(X_j = i_j) = \mathbb{P}(X_1 = i_1, \dots, X_{n-k} = i_{n-k}).$$

2. Par définition des S_n, on a
$$\begin{cases} S_{k+1} - S_k = j_1 \\ S_{k+2} - S_k = j_2 \\ \vdots \\ S_n - S_k = j_{n-k} \end{cases} \iff \begin{cases} X_{k+1} = j_1 \\ X_{k+1} + X_{k+2} = j_2 \\ \vdots \\ X_{k+1} + \dots + X_n = j_{n-k} \end{cases} \iff \begin{cases} X_{k+1} = j_1 \\ X_{k+2} = j_2 - j_1 \\ \vdots \\ X_n = j_{n-k} - j_{n-k-1} \end{cases}.$$
On peut alors utiliser le point 1 pour en déduire que
$$\mathbb{P}(S_{k+1} - S_k = j_1, \dots, S_n - S_k = j_{n-k}) = \mathbb{P}(X_1 = j_1, X_2 = j_2 - j_1, \dots, X_{n-k} = j_{n-k} - j_{n-k-1}).$$
On fait alors le trajet dans l'autre sens et on obtient alors le résultat.
3. On remarque que $A_k^n = (S_k = 0) \cap (S_{k+1} \neq 0) \cap \dots (S_n \neq 0)$ et donc
$$\mathbb{P}(A_k^n) = \mathbb{P}(S_k = 0)\mathbb{P}(S_{k+1} \neq 0, \dots, S_n \neq 0 \mid S_k = 0).$$
De façon évidente,
$$\mathbb{P}(S_{k+1} \neq 0, \dots, S_n \neq 0 \mid S_k = 0) = \mathbb{P}(S_{k+1} - S_k \neq 0, \dots, S_n - S_k \neq 0 \mid S_k = 0).$$
On remarque que les variables $S_{k+1} - S_k$, ..., $S_n - S_k$ ne dépendent que de $X_{k+1}, \dots, X_n$ et sont donc indépendantes de S_k qui ne dépend que de $X_1, \dots, X_k$ (et puisque les X_i sont, elles, indépendantes). Ainsi :
$$\mathbb{P}(S_{k+1} \neq 0, \dots, S_n \neq 0) = \mathbb{P}(S_{k+1} - S_k \neq 0, \dots, S_n - S_k \neq 0).$$
Remarquons que
$$\begin{aligned} \mathbb{P}(S_{k+1} - S_k \neq 0, \dots, S_n - S_k \neq 0) &= \mathbb{P}\left(\bigsqcup_{j_1, \dots, j_{n-k} \neq 0} (S_{k+1} - S_k = j_1, \dots, S_n - S_k = j_{n-k})\right) \\ &= \sum_{j_1, \dots, j_{n-k} \neq 0} \mathbb{P}\left(S_{k+1} - S_k = j_1, \dots, S_n - S_k = j_{n-k}\right) \\ &= \sum_{j_1, \dots, j_{n-k} \neq 0} \mathbb{P}\left(S_1 = j_1, \dots, S_{n-k} = j_{n-k}\right) \text{ par le point 2} \\ &= \mathbb{P}(S_1 \neq 0, \dots, S_{n-k} \neq 0) \\ &= \mathbb{P}(E_{n-k}). \end{aligned}$$
D'où le résultat.

4. On effectue le produit de Cauchy de ces deux séries absolument convergentes et on obtient pour $x \in]0,1[$:

$$\left(\sum_{n=0}^{\infty} \mathbb{P}(S_n = 0)x^n\right)\left(\sum_{n=0}^{\infty} \mathbb{P}(E_n)x^n\right) = \sum_{n=0}^{\infty} u_n x^n \text{ avec } u_n = \sum_{k=0}^{n} \mathbb{P}(S_k = 0)\mathbb{P}(E_{n-k})$$

Or, comme $S_0(\omega) = 0$, il existe un plus grand $k \in [\![0;n]\!]$ tel que $S_k(\omega) = 0$ et on a alors $\omega \in A_k^n$. La réunion des parties $A_0^n, \ldots, A_n^n$ est donc égale à Ω. Ces parties étant disjointes, on a donc d'après le point 3 :

$$1 = \sum_{k=0}^{n} \mathbb{P}(A_k^n) = \sum_{k=0}^{n} \mathbb{P}(S_k = 0)\mathbb{P}(E_{n-k}) = u_n.$$

Donc :

$$\left(\sum_{n=0}^{\infty} \mathbb{P}(S_n = 0)x^n\right)\left(\sum_{n=0}^{\infty} \mathbb{P}(E_n)x^n\right) = \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}.$$

□

Proposition 3.2

1. *On a le développement en série entière suivant :*

$$\forall x \in]-1,1[, \ \frac{1}{\sqrt{1-x}} = \sum_{k=0}^{+\infty} \frac{\binom{2k}{k}}{4^k} x^k.$$

2. *L'expression de la probabilité d'avoir $S_n = 0$ est :*

$$\forall p \in \mathbb{N}, \ \mathbb{P}(S_{2p+1} = 0) = 0 \quad \textit{et} \quad \mathbb{P}(S_{2p} = 0) = \frac{\binom{2p}{p}}{4^p}.$$

3. *Un équivalent de cette probabilité est :*

$$\mathbb{P}(S_{2n} = 0) \underset{n \to +\infty}{\sim} \sqrt{\frac{1}{\pi \, n}}.$$

4. *Pour tout $x \in]0,1[$, on a*

$$\sum_{n=0}^{+\infty} \mathbb{P}(E_n)x^n = \sqrt{\frac{1+x}{1-x}}.$$

5. *On a l'équivalent de $\mathbb{P}(E_n)$ suivant :*

$$\mathbb{P}(E_n) \sim \sqrt{\frac{2}{\pi n}}.$$

6. *On a $\mathbb{P}(T = +\infty) = 0$. La probabilité de ne jamais retourner en 0 est nulle.*
7. *Pour tout réel x de $]-1,1[$, on a l'égalité :*

$$\sum_{k=0}^{+\infty} \mathbb{P}(T = k)x^k = 1 - \sqrt{1-x^2}.$$

Démonstration. 1. D'après un développement en série entière usuel on a :

$$\forall x \in]-1,1[, \ (1+x)^\alpha = 1 + \sum_{k=1}^{+\infty} \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!} x^k$$

En choisissant $\alpha = -1/2$ et en substituant $-x$ à x, on a donc :

$$\forall x \in]-1,1[, \ \frac{1}{\sqrt{1-x}} = 1 + \sum_{k=1}^{\infty} a_k x^k \text{ avec } a_1 = \frac{1}{2}, \ \forall k \geqslant 1 \ : \ a_{k+1} = -\frac{-\frac{1}{2} - k}{k+1} a_k = \frac{2k+1}{2(k+1)} a_k$$

On montre par récurrence que

$$\forall k \geqslant 1, \ a_k = \frac{\binom{2k}{k}}{4^k}$$

La formule étant encore valable au rang $k = 0$ ($a_0 = 1$), on a ainsi

$$\forall x \in]-1,1[, \ \frac{1}{\sqrt{1-x}} = \sum_{k=0}^{+\infty} \frac{\binom{2k}{k}}{4^k} x^k.$$

2. La somme de n quantités valant 1 ou -1 ne peut donc être nulle que s'il y a autant de 1 que de -1, c'est à dire que si n est pair. On a donc
$$\forall n \in \mathbb{N},\ \mathbb{P}(S_{2n+1}=0)=0.$$
Supposons maintenant n pair et écrivons que $n=2p$. La valeur de $S_n(\omega)$ ne dépend que des valeurs de $\omega_1,\dots,\omega_{2p}$. Il y a $2^{2p}=4^p$ choix pour ces valeurs et on a
$$\mathbb{P}(S_{2p}=0)=\frac{\alpha_p}{4^p}$$
où α_p est le nombre de uplets $(\omega_1,\dots,\omega_{2p})$ contenant autant de 1 que de -1. Choisir un tel uplet, c'est choisir la position des 1 et il y a $\binom{2p}{p}$ tels choix. On a donc le résultat.

3. Par le point 2 et d'après un équivalent connu de Stirling on a :
$$\mathbb{P}(S_{2n}=0)=\frac{1}{4^n}\binom{2n}{n}\underset{n\to+\infty}{\sim}\frac{1}{4^n}\times\frac{4^n}{\sqrt{\pi}\sqrt{n}}=\sqrt{\frac{1}{\pi\, n}}.$$

4. On en déduit du point 1 et 2 que pour $x\in]-1,1[$,
$$\sum_{n=0}^{+\infty}\mathbb{P}(S_n=0)x^n=\sum_{p=0}^{+\infty}\frac{\binom{2p}{p}}{4^p}x^{2p}=\frac{1}{\sqrt{1-x^2}}$$
et donc que
$$\sum_{n=0}^{+\infty}\mathbb{P}(E_n)x^n=\frac{\sqrt{1-x^2}}{1-x}=\sqrt{\frac{1+x}{1-x}}.$$

5. Considérons la fonction f définie par
$$\forall x\in]-1,1[,\ f(x)=\sqrt{\frac{\pi}{2}}\sqrt{\frac{1+x}{1-x}}=\sqrt{\frac{\pi}{2}}\sum_{n=0}^{+\infty}\mathbb{P}(E_n)x^n\quad\text{par le point 4.}$$
Comme $\sqrt{1-x}f(x)\to\sqrt{\pi}$ quand $x\to1^-$, on peut utiliser le théorème de Karamata (3.1) pour obtenir :
$$\sum_{k=0}^{n}\sqrt{\frac{\pi}{2}}\mathbb{P}(E_n)\sim 2\sqrt{n}.$$
Par ailleurs $\mathbb{P}(T>n)=\mathbb{P}(E_n)$ est le terme général d'une suite décroissante ($(T>n)\subset(T>n-1)$) et alors par le théorème Taubérien (3.2) on a
$$\sqrt{\frac{\pi}{2}}\mathbb{P}(E_n)\sim\frac{1}{\sqrt{n}}.$$
On a donc
$$\mathbb{P}(E_n)\sim\sqrt{\frac{2}{\pi n}}.$$

6. On a
$$(T=+\infty)=\bigcap_{n\in\mathbb{N}}(T>n)=\bigcap_{n\in\mathbb{N}}E_n.$$
Par continuité décroissante des probabilités (appliqué avec la suite décroissante d'ensemble de terme général $E_0\cap\cdots\cap E_n=E_n$), on en déduit que
$$\mathbb{P}(T=+\infty)=\lim_{n\to+\infty}\mathbb{P}(E_n)=0.$$

7. On a $(T=n)=(T>n-1)\setminus(T>n)$ et comme $(T>n)\subset(T>n-1)$, on a alors l'expression : $\mathbb{P}(T=n)=\mathbb{P}(T>n-1)-\mathbb{P}(T>n)$ et ainsi
$$\begin{aligned}\sum_{k=0}^{n}\mathbb{P}(T=k)x^k &= \sum_{k=0}^{n}(\mathbb{P}(T>k-1)-\mathbb{P}(T>k))x^k\\ &= \mathbb{P}(T>-1)+x\sum_{k=1}^{n}\mathbb{P}(T>k-1)x^{k-1}-\sum_{k=0}^{n}\mathbb{P}(T>k)x^k\\ &= 1+x\sum_{k=0}^{n-1}\mathbb{P}(E_k)x^k-\sum_{k=0}^{n}\mathbb{P}(E_k)x^k.\end{aligned}$$

Pour $x \in]-1,1[$, les termes du membre de droite admettent une limite quand $n \to +\infty$ et, avec l'expression des sommes,

$$\sum_{k=0}^{+\infty} \mathbb{P}(T=k)x^k = 1 + x\sqrt{\frac{1+x}{1-x}} - 1\sqrt{\frac{1+x}{1-x}} = 1-(1-x)\sqrt{\frac{1+x}{1-x}} = 1-\sqrt{1-x^2} = 1-\sqrt{1-x^2}.$$

La relation reste vraie pour $x=1$ car $(T=n)_{n\in\mathbb{N}}$ est un système complet d'événements. □

Théorème 3.3

Pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(T=2n+1)=0 \quad \textit{et} \quad \mathbb{P}(T=2n) = \frac{1}{2n-1}\frac{\binom{2n}{n}}{4^n}.$$

Et alors :

$$\mathbb{P}(T=2n) \underset{n\to+\infty}{\sim} \frac{1}{2\sqrt{\pi}\, n^{3/2}}.$$

Démonstration. Première partie : Expression de la probabilité
Le fait que $\mathbb{P}(T=2n+1)=0$ est clair : on ne peut pas revenir en 0 avec un nombre impair d'itérations.
Montrons le deuxième point.

On commence par chercher le développement en série entière de $h : x \mapsto 1-\sqrt{1-x^2}$. Pour cela on remarque que pour $x \in]-1,1[$:

$$h'(x) = \frac{x}{\sqrt{1-x^2}} = x\sum_{k=0}^{\infty}\frac{\binom{2k}{k}}{4^k}x^{2k} = \sum_{k=0}^{\infty}\frac{\binom{2k}{k}}{4^k}x^{2k+1}.$$

Par théorème de primitivation des séries entières,

$$h(x) = h(0) + \sum_{k=0}^{\infty}\frac{\binom{2k}{k}}{4^k(2k+2)}x^{2k+2} = \sum_{j=1}^{\infty}\frac{\binom{2(j-1)}{j-1}}{4^{j-1}.2j}x^{2j}.$$

Or d'après la proposition on a aussi

$$1-\sqrt{1-x^2} = \sum_{k=0}^{+\infty}\mathbb{P}(T=k)x^k = \sum_{k=1}^{+\infty}\mathbb{P}(T=2k)x^{2k}. \quad (\mathbb{P}(T=0)=0)$$

Par unicité du DSE, on en déduit que

$$\mathbb{P}(T=2j) = \frac{\binom{2(j-1)}{j-1}}{4^{j-1}.2j}.$$

Il nous reste à remarquer que

$$\binom{2(j-1)}{j-1} = \frac{(2j-2)!}{((j-1)!)^2} = \frac{(2j)!}{(j!)^2}\frac{j^2}{2j(2j-1)} = \binom{2j}{j}\frac{j}{2(2j-1)}$$

pour en déduire que

$$\mathbb{P}(T=2j) = \binom{2j}{j}\frac{j}{2(2j-1)}\frac{1}{4^{j-1}.2j} = \frac{\binom{2j}{j}}{4^j(2j-1)}.$$

Ce qui démontre la première partie du théorème. □

Deuxième partie : Équivalent
La première partie du théorème fournit l'expression de la probabilité. L'équivalent se trouve grâce à la formumle de Stirling, on a alors :

$$\mathbb{P}(T=2n) \underset{n\to+\infty}{\sim} \frac{1}{2n-1}\frac{4^n}{\sqrt{\pi}\sqrt{n}\,4^n} \underset{n\to+\infty}{\sim} \frac{1}{2\sqrt{\pi}\,n^{3/2}}.$$

□

Lemme 3.1 (*Équivalent du reste d'une série de Riemann*)

Si $\alpha > 1$,

$$\sum_{k=n+1}^{+\infty}\frac{1}{k^\alpha} \underset{n\to+\infty}{\sim} \frac{1}{(\alpha-1)n^{\alpha-1}}.$$

Démonstration. Si $\alpha > 1$, on obtient par comparaison série intégrale que pour $n \geqslant 1$:

$$\frac{1}{(\alpha-1)(n+1)^{\alpha-1}} = \int_{n+1}^{+\infty} \frac{dt}{t^\alpha} \leqslant \sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \leqslant \int_{n}^{+\infty} \frac{dt}{t^\alpha} = \frac{1}{(\alpha-1)n^{\alpha-1}}.$$

On conclut en passant à la limite. □

Définition 3.1

Pour $i \in \mathbb{N}^*$, soit Y_i la variable de Bernoulli indicatrice de l'événement $(S_i \notin \{S_k,\ 0 \leqslant k \leqslant i-1\})$.
Notons pour $n \in \mathbb{N}$, N_n le cardinal du sous-ensemble de $\mathbb{Z}^d$ $\{S_k,\ k \in \{0,\dots,n\}\}$.
Le nombre de points de $\mathbb{Z}^d$ visités par la marche aléatoire $(S_n)_{n\in\mathbb{N}}$ après n pas est donc N_n.

Proposition 3.3

Pour $i, n \in \mathbb{N}^*$:

$$\mathbb{P}(Y_i = 1) = \mathbb{P}(T > i) \quad \textit{et} \quad \mathbb{E}(N_n) = 1 + \sum_{i=1}^{n} \mathbb{P}(T > i).$$

Démonstration. Soit $i \in \mathbb{N}^*$. On remarque que : $(Y_i = 1) = \bigcap_{k=0}^{i-1} (S_i \neq Sk)$ donc

$$(Y_i = 0) = \bigcup_{k=0}^{i-1} (S_i = Sk) = \bigcup_{k=0}^{i-1} \left(\sum_{j=k+1}^{i} X_j = 0_d \right).$$

Alors

$$\mathbb{P}\left(Y_i = 0\right) = \mathbb{P}\left(\bigcup_{k=0}^{i-1} \left(\sum_{j=k+1}^{i} X_j = 0_d \right) \right) = \mathbb{P}\left(\bigcup_{\ell=1}^{i} \left(\sum_{j=1}^{\ell} X_j = 0_d \right) \right)$$

Or

$$\bigcup_{\ell=1}^{i} \left(\sum_{j=1}^{\ell} X_j = 0_d \right) = \bigcup_{\ell=1}^{i} (S_\ell = 0_d) = (T \leqslant i)$$

Ainsi

$$\mathbb{P}\left(Y_i = 0\right) = \mathbb{P}\left(T \leqslant i\right)$$

En passant aux événements contraires, on obtient :

$$\mathbb{P}(Y_i = 1) = \mathbb{P}(T > i).$$

On précise que $(T > i) = (T = +\infty) \bigcup (T \in \{k \in \mathbb{N}, k > i\})$.
Or par récurrence immédiate sur $n \in \mathbb{N}$, on montre que $N_n = 1 + \sum_{i=1}^{n} Y_i$.
Les Y_i suivant des lois de Bernoulli, on a donc existences des membres et les inégalités :

$$\mathbb{E}(N_n) = 1 + \sum_{i=1}^{n} \mathbb{E}\left(Y_i\right) = 1 + \sum_{i=1}^{n} \mathbb{P}\left(Y_i = 1\right).$$

On en déduit que, pour $n \in \mathbb{N}^*$:

$$\mathbb{E}(N_n) = 1 + \sum_{i=1}^{n} \mathbb{P}(T > i).$$

□

Proposition 3.4

On a l'équivalent suivant :

$$\mathbb{E}(N_n) \underset{n\to+\infty}{\sim} \frac{2\sqrt{2n}}{\sqrt{\pi}}.$$

Démonstration. On a pour $p \in \mathbb{N}$, $(T > 2p) = (T > 2p+1) = \bigsqcup_{k=p+1}^{+\infty} (T = 2k)$. Ainsi,

$$\mathbb{P}\left(T > 2p\right) = \mathbb{P}\left(T > 2p+1\right) = \sum_{k=p+1}^{+\infty} \mathbb{P}\left(T = 2k\right).$$

La série de terme général $\mathbb{P}(T = 2k)$ converge car c'est un $O\left(n^{-3/2}\right)$ d'après le théorème 3.3. Alors en utilisant la sommation des relation de comparaison et en utilisant le lemme 3.1 on a quand p tend vers $+\infty$:

$$\mathbb{P}\left(T > 2p\right) = \mathbb{P}\left(T > 2p+1\right) \sim \sum_{k=p+1}^{+\infty} \frac{1}{2\sqrt{\pi}k^{3/2}} \sim \frac{1}{(3/2-1)2\sqrt{\pi}p^{3/2-1}} \sim \frac{1}{\sqrt{\pi}p^{1/2}}.$$

Donc

$$\sqrt{\frac{\pi}{2}}(2p)^{1/2}\mathbb{P}\left(T > 2p\right) \xrightarrow[p\to+\infty]{} 1 \quad \text{et} \quad \sqrt{\frac{\pi}{2}}(2p+1)^{1/2}\mathbb{P}\left(T > 2p+1\right) \xrightarrow[p\to+\infty]{} 1.$$

Ainsi puisque les suites extraites convergent vers 1 alors $\frac{\sqrt{\pi}}{\sqrt{2}}i^{1/2}\mathbb{P}\left(T > i\right) \xrightarrow[i\to+\infty]{} 1$ donc

$$\mathbb{P}\left(T > i\right) \underset{i\to+\infty}{\sim} \frac{\sqrt{2}}{\sqrt{\pi}i^{1/2}}.$$

À nouveau en utilisant le lemme 3.1 et les sommations des relations de comparaison de séries divergentes, on a :

$$\sum_{i=1}^{n}\mathbb{P}\left(T > i\right) \underset{n\to+\infty}{\sim} \sum_{i=1}^{n}\frac{\sqrt{2}}{\sqrt{\pi}i^{1/2}} \underset{n\to+\infty}{\sim} \frac{\sqrt{2}n^{1-1/2}}{\sqrt{\pi}(1-1/2)}.$$

Or d'après la proposition 3.3, $\mathbb{E}(N_n) = 1 + \sum_{i=1}^{n}\mathbb{P}\left(T > i\right)$ et on en déduit donc l'équivalent :

$$\mathbb{E}(N_n) \underset{n\to+\infty}{\sim} \sum_{i=1}^{n}\mathbb{P}\left(T > i\right) \underset{n\to+\infty}{\sim} \frac{2\sqrt{2n}}{\sqrt{\pi}} = \sqrt{\frac{8n}{\pi}}.$$

□

3.3 Marche aléatoire en dimension 2

Proposition 3.5

Pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(S_{2n} = 0_2) = \left(\frac{\binom{2n}{n}}{4^n}\right)^2$$

Et alors

$$\mathbb{P}\left(S_{2n} = 0_2\right) \underset{n\to+\infty}{\sim} \frac{1}{\pi n}.$$

Démonstration (Rapide par dénombrement)**.**
Pour retomber sur $0_2 = (0,0)$, sur l'axe horizontal on doit être aller n fois sur $\mathbb{Z}^+$ et n fois sur $\mathbb{Z}^-$. De même sur l'axe vertical. Autrement dit, $Z = (X, Y) = (0,0)$ si et seulement si X et Y (indépendantes) arrivent sur 0.
Ainsi, $\mathbb{P}(Z = 0) = \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$. Alors X et Y effectuent une marche aléatoire sur $\mathbb{Z}$ (X sur l'axe horizontal, Y sur l'axe vertical). Alors on obtient pour $n \in \mathbb{N}$,

$$\mathbb{P}(S_{2n} = 0_2) = \frac{\binom{2n}{n}}{4^n} \times \frac{\binom{2n}{n}}{4^n}.$$

Pour l'équivalent c'est à nouveau la formule de Stirling. □

Remarque 10 (Équivalent en l'infini)**.** On admet les résultats suivants :

$$\mathbb{P}(T > n) \underset{n\to+\infty}{\sim} \frac{1}{\pi\ln(n)} \quad \text{et} \quad \mathbb{E}(N_n) \underset{n\to+\infty}{\sim} \frac{n}{\pi\ln(n)}.$$

3.4 Marche aléatoire en dimension supérieure

On se place dans le cas plus général d'une marche aléatoire sur $\mathbb{Z}^d$. On peut la modéliser par une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ définie par récurrence par $X_0 = 0_d$, et pour tout n, $X_{n+1} = S_n + \theta_{n+1}$ où les (θ_n) sont des variables indépendantes et identiquement distribuées (i.i.d) de loi uniforme sur $\{\pm e_1, \ldots, \pm e_d\}$. On définit de la sorte une chaîne de Markov, et nous allons voir que le formalisme markovien est particulièrement bien adapté à l'étude de ce problème. Nous allons ici prouver le théorème de Polya.

Théorème 3.4 (*De Polya*)

Pour $d = 1$ et $d = 2$, la marche aléatoire $(X_n)_{n\in\mathbb{N}}$ est récurrente. Pour $d \geqslant 3$, elle est transiente.

Démonstration. Les cas $d = 1$ et $d = 2$ ont été traités précédement sans avoir recours aux chaînes de Markov. En effet, nous avons pu établir dans le cadre de la marche aléatoire sur $\mathbb{Z}$ que pour tout $n \in \mathbb{N}$:

$$\mathbb{P}(X_{2n+1} = 0) = 0 \quad \text{et que} \quad \mathbb{P}(X_{2n} = 0) \sim \sqrt{\frac{1}{\pi n}}.$$

Dès lors, on obtient par comparaison la divergence de la série $\sum_{n \geqslant 0} \mathbb{P}(X_n = 0)$ ce qui conclut sur le caractère transient de la chaîne.

De même pour la marche aléatoire sur $\mathbb{Z}^2$, on a pu établir que pour tout $n \in \mathbb{N}$:

$$\mathbb{P}(S_{2n+1} = 0) = 0 \quad \text{et que} \quad \mathbb{P}(S_{2n} = 0) \sim \frac{1}{\pi n}.$$

De la même façon, cela nous donne la divergence de la série $\sum_{n \geqslant 0} \mathbb{P}(X_n = 0)$ et donc la récurrence de la chaîne. Plaçons nous maintenant dans le cas où $d \geqslant 3$. Rappelons que les θ_n sont i.i.d et notons φ leur fonction caractéristique $\varphi : \mathbb{R}^d \to \mathbb{C}$ telle que pour tout vecteur $t = (t_1, \ldots, t_n)$:

$$\varphi(t) = \mathbb{E}\left[\exp\left(i\langle t, \theta_1\rangle\right)\right] = \frac{1}{d}(\cos(t_1) + \cdots + \cos(t_d)).$$

ETAPE 1 : UNE PREMIÈRE IDENTITÉ On va prouver que :

$$\sum_{k=0}^{+\infty} \mathbb{P}(X_{2k} = 0_d) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{\mathrm{d}t}{1 - \varphi(t)^2}.$$

Soit $x = (x_1, \ldots, x_d) \in \mathbb{Z}^d$. On note :

$$I_d(x) := \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \exp\left(i\langle x, t\rangle\right) \mathrm{d}t.$$

On vérifie grâce au théorème de Fubini que $I_d(x) = \prod_{j=1}^d I_1(x_j)$. Il s'ensuit que $I_d(x) = 1$ si $x = 0_d$ et $I_d(x) = 0$ sinon. En effet, supposons par exemple que $x_1 \neq 0$, alors $I_1(x_1) = \sin(x_1\pi)/x_1\pi = 0$ car $x_i \in \mathbb{Z}$.
Soit $\varphi_{X_n} : \mathbb{R}^d \to \mathbb{C}$ la fonction caractéristique de X_n. Par le théorème de Fubini, et ce qui précède, on montre très simplement qu'alors :

$$\frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \varphi_{X_n} \mathrm{d}t = \mathbb{E}\left[I_d(X_n)\right] = \mathbb{P}(X_n = 0_d).$$

Ensuite, étant donné que $X_n = \theta_1 + \cdots + \theta_n$, et que par hypothèse les θ_i sont i.i.d., on en déduit alors que $\varphi_{X_n} = \prod_{j=1}^n \varphi_{\theta_j} = \varphi(t)^n$. Par une permutation série-intégrale légitime puisqu'on intègre sur le compact $[-\pi, \pi]^d$ une fonction continue, il vient :

$$\sum_{k=0}^{+\infty} \mathbb{P}(X_{2k} = 0) = 0_d = \frac{1}{(2\pi)^d} \sum_{k=0}^{+\infty} \int_{[-\pi,\pi]^d} \varphi(t)^{2k} \mathrm{d}t = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \sum_{k=0}^{+\infty} \varphi(t)^{2k} \mathrm{d}t.$$

Enfin, le fait que pour presque tout $t \in \mathbb{R}^d$, $|\varphi(t)^2| < 1$ permet de conclure que :

$$\sum_{k=0}^{+\infty} \mathbb{P}(X_{2k} = 0) = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{\mathrm{d}t}{1 - \varphi(t)^2}.$$

ETAPE 2 : FINITUDE DE L'INTÉGRALE EN QUESTION
Les points où la fonction $t \mapsto 1 - \varphi(t)^2$ s'annule sont 0_d et les $(\varepsilon_1\pi, \ldots, \varepsilon_d\pi)$ où les ε_i valent 1 ou -1.
Il s'agit donc de montrer que la fonction question admet des limites finies en chacun de ces points.
Un développement limté à l'ordre 2 en 0_d de φ donne que $\varphi(t) = 1 - \|t\|^2/2d + o(\|t\|^2)$ et alors $\frac{1}{1-\varphi(t)^2} \underset{t \to 0_d}{\sim} \frac{d}{\|t\|^2}$.
La fonction de l'équivalent est localement intégrable au voisinage de 0. Pour le voir, considérons une boule ouverte centrée en 0 de rayon $\varepsilon > 0$, et faisons un changement de variable sphérique $t = r\alpha$ où $r \geqslant 0$ et $\alpha \in \mathbb{S}^{d-1}$ la sphère unité de $\mathbb{R}^d$. Le changement de variable sphérique fait apparaitre une expression régulière en des fonctions cosinus et sinus. Cette expression étant continue sur le compact $\mathbb{S}^{d-1}$ elle est majorée par une constante C_1 qui ne dépend que de d. On note ensuite C_2 la constante valant le produit entre C_1 et l'intégrale sur la sphère $\mathbb{S}^{d-1}$. Alors :

$$\int_{B(0,\varepsilon)} \frac{1}{\|t\|^2} \mathrm{d}t \leqslant \int_0^\varepsilon \int_{\mathbb{S}^{d-1}} \frac{1}{r^2} C_1 r^{d-1} \mathrm{d}\alpha \mathrm{d}r \leqslant C_2 \int_0^\varepsilon r^{d-3} \mathrm{d}r.$$

Or lorsque $d \geqslant 3$ cette intégrale est convergente. Le raisonnement est le même pour les points $(\pi, \ldots, \pi)$ et $(-\pi, \ldots, -\pi)$. Finalement, la transience de la chaîne se déduit des deux étapes précédentes car alors :

$$\sum_{n=0}^{+\infty} \mathbb{P}(X_n = 0_d) = \sum_{k=0}^{+\infty} \mathbb{P}(X_{2k} = 0_d) < +\infty.$$

□

Références

[BE04] Michel BENAÏM et Nicole EL KAROUI. *Promenade aléatoire - Chaînes de Markov, algorithmes, martingales, arrêt optimal*. 2004, p. 41-93.

[FF04] Dominique FOATA et Aimé FUCHS. *Processus stochastiques - Processus de Poisson, chaînes de Markov et martingales*. Dunod, 2004.

[Min16] Concours MINES-PONTS. *Retours à l'origine d'une marche aléatoire*. 2016.