Processus de Markov inhomogènes finis

Laurent Miclo

Avertissement : Les notes qui suivent constituent une version (trop ?) détaillée du premier chapitre d'un cours de DEA, et elles ont justement été rédigées pour pouvoir être enseignées superficiellement. Il s'agit d'une introduction aux processus de Markov finis inhomogènes, dans l'esprit des livres de Blumenthal et Getoor [1], d'Ethier et Kurtz [4] et d'Iosifescu et Tăutu [7], adaptés à la situation particulière des espaces d'états finis.

1 Généralités sur la structure des processus de Markov finis

Nous allons dans ce chapitre introduire et analyser les processus de Markov avec lesquels on travaillera ensuite. Ces objets seront alors définis à partir d'une loi initiale m_0 et d'une famille continue $(L_t)_{t\geq 0}$ de générateurs, et notre but ici est donc de bien comprendre ce qui se cache derrière la formule incantatoire « Soit $(X_t)_{t\geq 0}$ associé à m_0 et $(L_t)_{t\geq 0}$ ».

1.1 Propriété de Markov : introduction des cadres probabilistes

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité quelconque et S un ensemble fini, que l'on munit de sa tribu usuelle $\mathcal{S}$, qui est celle formée par toutes les sous-parties de S.

Un processus stochastique défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ à valeurs dans S désignera la donnée, pour tout $t \in \mathbb{R}_+$, d'une application mesurable

$$X_t \, : \, (\Omega, \mathcal{F}) \to (S, \mathcal{S})$$

Le mot « processus » fera ici référence au fait que « l'ensemble des temps » est la demi-droite positive. On rencontrera aussi ultérieurement des processus stochastiques à valeurs dans d'autres espaces mesurables que $(S, \mathcal{S})$, notamment dans la droite réelle munie de sa tribu borélienne $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Un tel processus $X = (X_t)_{t\geq 0}$ engendre naturellement une filtration (de manière générale ce terme désigne une famille croissante de sous-tribus de $\mathcal{F}$) indexée par $\mathbb{R}_+$, $(\mathcal{F}_t)_{t\geq 0}$, définie par

$$\forall \, t \geq 0, \quad \mathcal{F}_t = \sigma(X_s \, ; \, 0 \leq s \leq t)$$

($\mathcal{F}_t$ représente donc l'ensemble des événements relatifs au processus qui sont antérieurs à l'instant t).

Pour $t \geq 0$, on pose également $\mathcal{F}^t = \sigma(X_s \, ; \, s \geq t)$, la tribu des événements ne faisant intervenir que les positions de la trajectoire à des temps postérieurs à l'instant t.

Définition 1

Le processus X sera dit *de Markov* (sous-entendu sous la probabilité $\mathbb{P}$), si à tout instant, la trajectoire passée et la trajectoire future du processus sont indépendantes connaissant la position actuelle, c'est-à-dire plus rigoureusement, si pour tout $t \geq 0$, on a l'indépendance de $\mathcal{F}_t$ et de $\mathcal{F}^t$ conditionnellement à $\sigma(X_t)$:

$$\mathcal{F}_t \underset{\sigma(X_t)}{\perp\!\!\!\perp} \mathcal{F}^t$$

Faisons quelques rappels sur l'indépendance conditionnelle : la dernière expression ci-dessus signifie que pour toutes applications $F, G : \Omega \to \mathbb{R}_+$ qui sont respectivement $\mathcal{F}_t$ et $\mathcal{F}^t$ mesurables, on a $\mathbb{P}$-p.s.,

$$\mathbb{E}[FG|X_t] = \mathbb{E}[F|X_t]\mathbb{E}[G|X_t]$$

Remarquons qu'il s'agit là d'espérances conditionnelles relativement « simples », car $\sigma(X_t)$ est une tribu finie : elle est engendrée par la partition de Ω formée par les ensembles $\{X_t = x\}$, pour $x \in S$. Ainsi par exemple on peut prendre pour tout $\omega \in \Omega$,

$$\mathbb{E}[F|X_t](\omega) = \sum_{x \in S} \frac{\mathbb{E}[F\mathbb{1}_{\{x\}}(X_t)]}{\mathbb{P}[X_t = x]} \mathbb{1}_{\{x\}}(X_t(\omega))$$

avec la convention que $0.\infty = 0$, que l'on supposera toujours implicite par la suite.

Enfin retenons que $\mathcal{F}_t \underset{\sigma(X_t)}{\perp\!\!\!\perp} \mathcal{F}^t$ équivaut aussi au fait que pour toute application $F : \Omega \to \mathbb{R}_+$ qui est $\mathcal{F}^t$ mesurable, on doit avoir $\mathbb{P}$-p.s.,

$$\mathbb{E}[F|\mathcal{F}_t] = \mathbb{E}[F|X_t]$$

ce qui revient à dire qu'à l'instant $t \geq 0$, l'évolution future ne dépend du passé que par l'intermédiaire du présent.

De manière plus pratique, un processus X est de Markov si et seulement si pour tout $t \geq 0$, on a

$$\forall\, n \in \mathbb{N},\ \forall\, 0 \leq t_0 < t_1 < \cdots < t_n = t < t_{n+1},\ \forall\, x_0, \cdots, x_{n+1} \in S,$$

$$\mathbb{P}[X_{t_{n+1}} = x_{n+1}|X_{t_n} = x_n, \cdots, X_{t_0} = x_0] = \mathbb{P}[X_{t_{n+1}} = x_{n+1}|X_{t_n} = x_n]$$

dès que $\mathbb{P}[X_{t_n} = x_n, \cdots, X_{t_0} = x_0] > 0$.

Exercice (sur les classes monotones et l'espérance conditionnelle) : vérifier cette affirmation.

Pour $s \geq 0$, introduisons $S_s = \{x \in S : \mathbb{P}[X_s = x] > 0\}$, puis posons pour $0 \leq s \leq t$, $x \in S_s$ et $y \in S_t$,

$$(1) \qquad p_{s,t}(x,y) = \mathbb{P}[X_t = y|X_s = x] = \frac{\mathbb{P}[\{X_t = y\} \cap \{X_s = x\}]}{\mathbb{P}[X_s = x]}$$

la probabilité d'être en y au temps t sachant que l'on est en x au temps s.

Il apparaît alors que pour tous $0 \leq s \leq t \leq u$,

$$\begin{array}{rrcl} \forall\, x, y \in S_s, & p_{s,s}(x,y) & = & \delta_x(y) \\ \forall\, x \in S_s,\ \forall\, y \in S_u, & p_{s,u}(x,y) & = & \displaystyle\sum_{z \in S_t} p_{s,t}(x,z)p_{t,u}(z,y) \end{array}$$

En effet, la première relation est immédiate, et pour $x \in S_s$ et $y \in S_u$, on a

$$\begin{aligned}
I\!\!P[X_s = x] p_{s,u}(x,y) &= I\!\!P[X_u = y,\, X_s = x] \\
&= I\!\!E[I\!\!P[X_u = y|\mathcal{F}_t] \mathbb{1}_{\{x\}}(X_s)] \\
&= I\!\!E[I\!\!P[X_u = y|X_t] \mathbb{1}_{\{x\}}(X_s)] \\
&= \sum_{z \in S} I\!\!E[I\!\!P[X_u = y|X_t = z] \mathbb{1}_{\{z\}}(X_t) \mathbb{1}_{\{x\}}(X_s)] \\
&= \sum_{z \in S_t} I\!\!E[\mathbb{1}_{\{z\}}(X_t) \mathbb{1}_{\{x\}}(X_s)] p_{t,u}(z,y) \\
&= I\!\!P[X_s = x] \sum_{z \in S_t} p_{s,t}(x,z) p_{t,u}(z,y)
\end{aligned}$$

d'où la seconde relation.

Il est naturel de considérer $p_{s,t} = (p_{s,t}(x,y))_{(x,y) \in S_s \times S_t}$ comme une matrice rectangulaire dont les colonnes et les lignes sont respectivement indicées par S_s et S_t. Il s'agit d'une matrice de probabilités de transition de S_s à S_t (on parle aussi parfois de matrice markovienne, ou plus improprement de matrice stochastique), au sens où pour tout $x \in S_s$,

$$\forall\, y \in S_y,\ p_{s,t}(x,y) \geq 0 \quad \text{et} \quad \sum_{y \in S_t} p_{s,t}(x,y) = 1$$

(c'est-à-dire que $p_{s,t}(x,\cdot)$ définit une probabilité sur S_t).

Les relations précédentes peuvent alors se réécrire sous forme matricielle :

$$(2) \qquad \left\{ \begin{array}{rrcl} \forall\, s \geq 0, & p_{s,s} &=& \mathrm{id}_{S_s} \\ \forall\, 0 \leq s \leq t \leq u, & p_{s,u} &=& p_{s,t} p_{t,u} \end{array} \right.$$

où id_{S_s} représente la matrice identité sur S_s.

Notons m_0 la loi de X_0. On a alors pour tout $n \in I\!\!N$, tous $0 = t_0 < t_1 < \cdots < t_n$ et tous $A_0, A_1, \cdots, A_n \subset S$,

$$(3) \quad \begin{aligned} & I\!\!P[X_{t_0} \in A_0, \cdots, X_{t_n} \in A_n] \\ &= \sum_{(x_0,\cdots,x_n) \in S_{t_0} \times \cdots \times S_{t_n}} m_0(x_0) \mathbb{1}_{A_0}(x_0) p_{t_0,t_1}(x_0,x_1) \mathbb{1}_{A_1}(x_1) \cdots p_{t_{n-1},t_n}(x_{n-1},x_n) \mathbb{1}_{A_n}(x_n) \end{aligned}$$

En effet, ceci se démontre par récurrence sur n, en utilisant la formule

$$\begin{aligned}
& I\!\!P[X_{t_0} \in A_0, \cdots, X_{t_n} \in A_n] \\
&= I\!\!E[\mathbb{1}_{A_0}(X_0) \cdots \mathbb{1}_{A_{n-1}}(X_{t_{n-1}}) I\!\!E[\mathbb{1}_{S_{t_n} \cap A_n}(X_{t_n}) | X_{t_{n-1}}, \cdots, X_0]] \\
&= I\!\!E[\mathbb{1}_{A_0}(X_0) \cdots \mathbb{1}_{S_{t_{n-1}} \cap A_{n-1}}(X_{t_{n-1}}) \sum_{x_n \in S_{t_n}} p_{t_{n-1},t_n}(X_{t_{n-1}},x_n) \mathbb{1}_{A_n}(x_n)] \\
&= \sum_{x_{n-1} \in S_{t_{n-1}}, x_n \in S_{t_n}} I\!\!E[\mathbb{1}_{A_0}(X_0) \cdots \mathbb{1}_{\{x_{n-1}\}}(X_{t_{n-1}})] \mathbb{1}_{A_{n-1}}(x_{n-1}) p_{t_{n-1},t_n}(x_{n-1},x_n) \mathbb{1}_{A_n}(x_n)
\end{aligned}$$

Souvent on préférera remplacer les sommes par des intégrales, car ces dernières se généralisent à d'autres espaces que l'ensemble fini S, ce qui permet d'unifier les notations, ainsi le membre de droite de (3) peut s'écrire

$$\int_{S_{t_0} \times S_{t_1} \times \cdots \times S_{t_n}} \mathbb{1}_{A_0}(x_0) \mathbb{1}_{A_1}(x_1) \cdots \mathbb{1}_{A_n}(x_n)\, m_0(dx_0) p_{t_0,t_1}(x_0,dx_1) \cdots p_{t_{n-1},t_n}(x_{n-1},dx_n)$$

On peut considérer $X = (X_t)_{t\geq 0}$ comme une v.a. à valeurs dans l'espace mesurable $(S^{I\!R_+}, \mathcal{S}^{\otimes I\!R_+})$ (la structure produit de cet espace assure que $X : \Omega \to S^{I\!R_+}$ est bien mesurable pour tout processus stochastique X). Notons Q l'image de $I\!P$ par X, qui est appelée la loi de X (sous-entendu sur $(S^{I\!R_+}, \mathcal{S}^{\otimes I\!R_+})$).

Soit également $\mathcal{A}$ l'ensemble des sous-ensembles de $S^{I\!R_+}$ de la forme

$$\{a = (a_t)_{t\geq 0} \,:\, \forall\ 0 \leq i \leq n,\, a_{t_i} \in A_i\}$$

quand $n \in I\!N$, $0 = t_0 < t_1 < \cdots < t_n$ et $A_0, A_1, \cdots, A_n \subset S$ varient. Il est clair que $\mathcal{A}$ est stable par intersection et engendre $\mathcal{S}^{\otimes I\!R_+}$, ainsi (par le théorème des classes monotones), la restriction de Q à $\mathcal{A}$ détermine Q, ce qui montre par le biais de (3) que si deux processus de Markov $X^{(1)}$ et $X^{(2)}$ (éventuellement définis sur des espaces de probabilité différents) satisfont, avec des notations évidentes,

$$\begin{array}{rl}
\forall\ s \geq 0, & S_s \stackrel{\text{déf.}}{=} S_s^{(1)} = S_s^{(2)} \\
\forall\ x \in S_0, & m_0^{(1)}(x) = m_0^{(2)}(x) \\
\forall\ 0 \leq s \leq t,\, \forall\ x \in S_s,\, \forall\ y \in S_t & p_{s,t}^{(1)}(x,y) = p_{s,t}^{(2)}(x,y)
\end{array}$$

alors $Q^{(1)} = Q^{(2)}$, c'est-à-dire que $X^{(1)}$ et $X^{(2)}$ ont même loi. On aura remarqué qu'il s'agit là d'une condition aussi nécessaire, et que les premières égalités seront satisfaites si $m_0^{(1)} = m_0^{(2)}$, ce qui implique que $S_0^{(1)} = S_0^{(2)}$, et si pour tout $t \geq 0$, tout $x \in S_0$ et tout $y \in S$, $p_{0,t}^{(1)}(x,y) = p_{0,t}^{(2)}(x,y)$, ces quantités pouvant se définir comme précédemment, sans les restrictions $y \in S_t^{(1)}$ et $y \in S_t^{(2)}$.

Pour $t \geq 0$, soit $Y_t : S^{I\!R_+} \ni \omega = (\omega_s)_{s\geq 0} \mapsto \omega_t \in S$, l'application coordonnée au temps t. Il apparaît facilement que le processus $(Y_t)_{t\geq 0}$, défini sur l'espace de probabilité $(S^{I\!R_+}, \mathcal{S}^{\otimes I\!R_+}, Q)$, a même loi que X et notamment est donc aussi markovien, car on aura remarqué que la propriété de Markov présentée ne dépend en fait que de la loi du processus.

Souvent dans l'étude d'un processus, seule sa loi intervient (typiquement quand on cherche à évaluer la probabilité d'événements ne dépendant que de la trajectoire de ce processus, i.e. appartenant à $\mathcal{F}_\infty = \sigma(X_t\,;\, t \geq 0)$), et dans ce cas on peut donc se ramener à supposer que $(\Omega, \mathcal{F}, I\!P) = (S^{I\!R_+}, \mathcal{S}^{\otimes I\!R_+}, Q)$, et $(X_t)_{t\geq 0} = (Y_t)_{t\geq 0}$. On dit alors que l'on s'est placé sur l'espace canonique des trajectoires, l'avantage étant qu'il est plus facile à manier car il est explicite.

On pourrait considérer X comme étant ($I\!P$-p.s.) à valeurs dans l'espace $\prod_{t\geq 0} S_t$, mais ce point de vue n'est pas très avantageux. Au contraire, il est plutôt commode de disposer de quantités $p_{s,t}(x,y)$ définies pour tous $0 \leq s \leq t$ et tous $x, y \in S$, de sorte que $p_{s,t}$ soit une matrice carrée indicée par $S \times S$. On pourrait convenir de prendre par exemple $p_{s,t}(x,y) = \delta_x(y)$ si $I\!P[X_s = x] = 0$. Cependant avec cette définition, la seconde relation de (2) risque de ne pas être satisfaite en général. Nous allons donc procéder à un autre choix, qui restera d'une certaine façon arbitraire, car les situations où $I\!P[X_s = x] = 0$ n'auraient pas besoin d'être traitées si l'on voulait garder le point de vue d'un processus X sous une seule probabilité $I\!P$ (mais ceci ne sera pas le cas!). Puisque pour tout $s \geq 0$, on a $S_s \neq \emptyset$, choisissons un élément $x_s \in S_s$, de manière à avoir pour tout $s \geq 0$, $I\!P[X_s = x_s] > 0$. Si $s \geq 0$ et $x \in S \setminus S_s$, on convient alors de prendre pour tout $t \geq s$ et tout $y \in S$,

$$p_{s,t}(x,y) = \begin{cases} p_{s,t}(x_s, y) = I\!P[\{X_t = y\} \cap \{X_s = x_s\}]/I\!P[X_s = x_s] & \text{, si } t > s \\ \delta_x(y) & \text{, si } t = s \end{cases}$$

Comme annoncé, ceci admet l'avantage que les matrices $p_{s,t}$, $0 \leq s \leq t$, sont markoviennes sur S et que

$$(4) \qquad \begin{cases} \forall\ s \geq 0, & p_{s,s} = \mathrm{id} \\ \forall\ 0 \leq s \leq t \leq u, & p_{s,u} = p_{s,t}p_{t,u} \end{cases}$$

où id représente la matrice identité (sur S).

Il apparaît également que pour tout $n \in \mathbb{N}$, tous $0 = t_0 < t_1 < \cdots < t_n$ et tous $A_0, A_1, \cdots, A_n \subset S$,

$$\begin{aligned}&\mathbb{P}[X_{t_0} \in A_0, \cdots, X_{t_n} \in A_n] \\ &\quad = \int_{S^{n+1}} \mathbf{1}_{A_0}(x_0)\mathbf{1}_{A_1}(x_1)\cdots\mathbf{1}_{A_n}(x_n)\, m_0(dx_0)p_{t_0,t_1}(x_0,dx_1)\cdots p_{t_{n-1},t_n}(x_{n-1},dx_n)\end{aligned}$$

De manière générale, toute famille $(p_{s,t})_{0\leq s\leq t}$ de matrices de transition qui satisfait les deux conditions (4) ci-dessus est appelée semi-groupe inhomogène markovien (sur S). On considèrera plus en détail cette propriété dans la section 3.

Une autre particularité de l'espace mesurable $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes\mathbb{R}_+})$ est que souvent quand on cherche à construire un processus de Markov sur S à partir d'un semi-groupe inhomogène markovien donné, on le fait directement sur cet espace, en utilisant le résultat suivant :

Théorème 2

Soit $(q_{s,t})_{0\leq s\leq t}$ un semi-groupe inhomogène markovien sur S et $x_0 \in S$. Il existe une unique probabilité $\mathbb{P}_{0,x_0}$ sur $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes\mathbb{R}_+})$ pour laquelle $(Y_t)_{t\geq 0}$ est un processus de Markov vérifiant pour tous $0 \leq s \leq t$ et tous $x, y \in S$ tels que $\mathbb{P}_{0,x_0}[Y_s = x] > 0$,

$$\mathbb{P}_{0,x_0}[Y_t = y | Y_s = x] = q_{s,t}(x,y)$$

et

$$\mathbb{P}_{0,x_0}[Y_0 = x_0] = 1$$

(on dit alors que $\mathbb{P}_{0,x_0}$ (ou Y sous $\mathbb{P}_{0,x_0}$) admet δ_{x_0} pour loi initiale).

L'unicité découle d'une discussion précédente et nous admettrons ici le résultat d'existence (la démonstration est basée sur le théorème des limites projectives de Daniell-Kolmogorov, et s'étend donc au cas où $(S, \mathcal{S})$ est remplacé par exemple par un espace topologique σ-compact, cf. le théorème 2.11 p. 17 de [1] ou le théorème 2.3.3 p. 197 de [7]).

Soit m_0 une probabilité (« initiale ») sur S, et notons $\mathbb{P}_{0,m_0}$ la loi définie sur $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes\mathbb{R}_+})$ par

$$\forall\ A \in \mathcal{S}^{\otimes\mathbb{R}_+}, \qquad \mathbb{P}_{0,m_0}[A] = \sum_{x\in S} m_0(x)\mathbb{P}_{0,x}[A]$$

On vérifie sans difficulté que l'on a encore pour tous $0 \leq s \leq t$ et tous $x, y \in S$ tels que $\mathbb{P}_{0,m_0}[Y_s = x] > 0$,

$$\mathbb{P}_{0,m_0}[Y_t = y | Y_s = x] = q_{s,t}(x,y)$$

que la loi de X_0 sous $\mathbb{P}_{0,m_0}$ est m_0, et que $\mathbb{P}_{0,m_0}$ est l'unique loi sur $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes\mathbb{R}_+})$ satisfaisant ces deux propriétés.

Ainsi, si l'on utilise le théorème 2 avec $(q_{s,t})_{0\leq s\leq t} = (p_{s,t})_{0\leq s\leq t}$, un semi-groupe inhomogène markovien associé à X comme précédemment, on obtient que $Q = \mathbb{P}_{0,m_0}$.

Plus généralement, soit $u \geq 0$, et appliquons le théorème précédent avec le semi-groupe inhomogène markovien $(q_{s,t})_{0\leq s\leq t} = (p_{u+s,u+t})_{0\leq s\leq t}$, pour obtenir pour tout $x \in S$ une probabilité sur $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes\mathbb{R}_+})$ que l'on notera désormais $\mathbb{P}_{u,x}$. Celle-ci étant de loi initiale δ_x, on a

$$\forall\ u \geq 0, \forall\ x \in S, \qquad \mathbb{P}_{u,x}[Y_0 = x] = 1 \tag{5}$$

On désignera par $\mathbb{E}_{u,x}$ l'espérance relative à cette probabilité $\mathbb{P}_{u,x}$ (si m est une probabilité sur S, on pose également $\mathbb{E}_{u,m}[\,\cdot\,] = \sum_{x\in S} m(x)\mathbb{E}_{u,x}[\,\cdot\,]$).

Pour donner une autre formulation de la propriété de Markov pour $Y = (Y_t)_{t\geq 0}$, introduisons les opérateurs de translations (dans le temps) $(\theta_t)_{t\geq 0}$ définis par

$$\forall\, t \in \mathbb{R}_+, \qquad \begin{array}{rccl} \theta_t \,:\, & S^{\mathbb{R}_+} & \to & S^{\mathbb{R}_+} \\ & \omega & \mapsto & \theta_t(\omega) \end{array}$$

où $\theta_t(\omega)$ est l'application $\mathbb{R}_+ \ni s \mapsto \omega(t+s) \in S$.

Avec cette notation, on a pour tous $t, s \geq 0$, $Y_s \circ \theta_t = Y_{t+s}$, ce qui montre notamment que θ_t est mesurable.

Enfin soit $(\mathcal{G}_t)_{t\geq 0}$ la filtration engendrée par Y (pour tout $t \geq 0$, on pourrait donc identifier naturellement $\mathcal{G}_t$ avec $\mathcal{S}^{\otimes[0,t]}$).

Exercice : Vérifier que la propriété de Markov (pour Y sous $Q = \mathbb{P}_{0,m_0}$) implique que pour toute application mesurable $F : S^{\mathbb{R}_+} \to \mathbb{R}_+$, on a $\mathbb{P}_{0,m_0}$-p.s.

$$\mathbb{E}_{0,m_0}[F \circ \theta_t | \mathcal{G}_t] = \mathbb{E}_{t,Y_t}[F] \tag{6}$$

Ainsi il apparaît, puisqu'alors $\mathbb{E}_{0,m_0}[F \circ \theta_t | \mathcal{G}_t] = \mathbb{E}_{0,m_0}[F \circ \theta_t | Y_t]$, que $\mathbb{P}_{t,x}$ est la loi de la trajectoire après l'instant $t \geq 0$, que l'on conditionne à repartir de $x \in S$ (ou de manière équivalente, après que l'on ait conditionné la position Y_t au temps t à être en x), ceci si $\mathbb{P}_{0,m_0}[Y_t = x] > 0$.

De la même manière, on prouve plus généralement que pour tous $s, t \geq 0$, tout $x \in S$ et toute application mesurable $F : S^{\mathbb{R}_+} \to \mathbb{R}_+$,

$$\mathbb{E}_{s,x}[F \circ \theta_t | \mathcal{G}_t] = \mathbb{E}_{s+t,Y_t}[F], \qquad \mathbb{P}_{s,x}\text{-p.s.} \tag{7}$$

ce qui assure que le processus Y est aussi markovien (au sens de la définition 1) sous chacunes des lois $\mathbb{P}_{t,x}$, pour $t \geq 0$ et $x \in S$.

Ainsi à notre processus $(X_t)_{t\geq 0}$, on a associé une loi initiale m_0 et un sextuplet

$$(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes \mathbb{R}_+}, (\mathcal{G}_t)_{t\geq 0}, (\theta_t)_{t\geq 0}, (Y_t)_{t\geq 0}, (\mathbb{P}_{t,x})_{t\geq 0, x\in S}) \tag{8}$$

où la famille des probabilités $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ satisfait les conditions (5) et (7). Comme on peut reconstruire Q à partir de m_0 et de $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$, il suffit de s'intéresser à cette dernière famille. Rappelons toutefois que le semi-groupe inhomogène markovien $(p_{s,t})_{0\leq s\leq t}$ que l'on a construit à partir du processus de Markov X, et par suite la famille $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ que l'on en déduit, n'ont en général rien de canoniques (sauf par exemple si pour tout $s \geq 0$, la loi de X_s charge tous les points de S). Plus généralement, tout sextuplet de la forme (8) satisfaisant (5) et (7), et tel que pour tout $t \geq 0$ et toute application mesurable $F : S^{\mathbb{R}_+} \to \mathbb{R}_+$,

$$\mathbb{E}[F \circ \theta_t \circ X | \mathcal{F}_t] = \mathbb{E}_{t,X_t}[F], \qquad \mathbb{P}\text{-p.s.} \tag{9}$$

(ce qui n'est qu'une réécriture de (6)), sera dit associé à X.

Mais souvent on part directement de la donnée d'un tel objet (8), satisfaisant les conditions (5) et (7), et ce sextuplet est alors appelé un processus de Markov, car il est plus commode de travailler avec la famille de probabilités $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$, qu'avec une seule loi Q, cela évite notamment d'avoir à considérer maladroitement, comme précédemment, les situations où $\mathbb{P}(X_t = x) = Q(Y_t = x) = 0$. De plus, ceci n'est pas une exigence démesurée, car en général il n'est pas plus difficile de construire un processus de Markov sous une probabilité $\mathbb{P}$ que toute une famille $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ sur $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes \mathbb{R}_+})$ satisfaisant (5) et (7), par exemple si on utilise le théorème 2. Et puis surtout, quand on remplace $(S, \mathcal{S})$ par d'autres espaces (non dénombrables), un objet de la forme (8) devient difficilement contournable si l'on veut travailler correctement (cf. par exemple les définitions données pour le cas homogène p. 20 de [1] ou p. 8 de [9]), car il permet d'avoir des processus partant de n'importe quel point à n'importe quel instant dont les évolutions sont « compatibles », au sens où (7) est satisfait.

Définition 3

Dans le cadre donné par (8), on dit que le processus de Markov est *homogène (dans le temps)*, si pour tout $t \geq 0$ et tout $x \in S$,

$$\mathbb{P}_{t,x} = \mathbb{P}_{0,x}$$

Dans le contexte de la définition 1 d'un processus de Markov sous une loi $\mathbb{P}$, la notion d'homogénéité s'énonce plutôt : pour tous $0 \leq s \leq t$ et tous $x \in S_s$, $y \in S$, $p_{s,t}(x,y)$ ne dépend que de $t-s$, de x et de y.

Remarque : Si X est homogène au sens de la seconde acceptation ci-dessus, reprenons (1) et posons, pour $0 \leq s \leq t$ et $x, y \in S$ tels que $\mathbb{P}[X_s = x] = 0$,

$$p_{s,t}(x,y) \quad = \quad \begin{cases} \delta_x(y) & \text{, si pour tout } u \geq 0,\ \mathbb{P}[X_u = x] = 0 \\ p_{u,u+t-s}(x,y) & \text{, s'il existe un } u \geq 0 \text{ tel que } \mathbb{P}[X_u = x] > 0 \end{cases}$$

(l'hypothèse d'homogénéité assurant que cette dernière expression ne dépend pas du choix d'un tel u). La famille $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ que l'on en déduit est alors homogène. Ceci suggère que l'on aurait dû prendre comme définition équivalente dans le second cas : un processus de Markov X sous une loi $\mathbb{P}$ est homogène s'il existe une famille homogène $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ qui lui est associée.

Pour finir cette section, faisons quelques rappels sur les temps d'arrêt, ce sera aussi l'occasion de manipuler le cadre markovien donné par (8), pour essayer à nouveau de nous convaincre qu'il est relativement naturel.

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ un espace filtré, c'est-à-dire un espace mesuré muni d'une filtration. Une variable aléatoire $\tau \; : \; (\Omega, \mathcal{F}) \to \overline{\mathbb{R}}_+ \stackrel{\text{déf.}}{=} \mathbb{R}_+ \sqcup \{+\infty\}$ est appelée un temps d'arrêt, si pour tout $t \geq 0$, l'événement $\{\tau \leq t\}$ appartient à $\mathcal{F}_t$ (il s'agit donc d'une notion ne faisant intervenir aucune probabilité, d'ailleurs la donnée d'une filtration est équivalente à la donnée de tous les temps d'arrêt associés, et leur ensemble peut être compris comme une « bonne » description de la filtration ...).

Très heuristiquement, ceci signifie que pour savoir si τ prend la valeur t, il suffit de regarder les événements antérieurs à t, dont l'ensemble est représenté par $\mathcal{F}_t$. Cette phrase peut être rendue tout à fait rigoureuse dans le cas où τ ne prend qu'un nombre dénombrable de valeurs.

Exercice : soit τ une variable aléatoire définie sur Ω et à valeurs dans un ensemble dénombrable $D \subset \overline{\mathbb{R}}_+$, montrer que τ est un temps d'arrêt si et seulement si pour tout $t \in D$, $\{\tau = t\} \in \mathcal{F}_t$.

L'exemple le plus simple d'un temps d'arrêt est évidemment celui d'une variable aléatoire constante, i.e. prenant toujours une valeur $t \geq 0$ donnée, mais l'avantage des temps d'arrêt est qu'ils permettent d'étendre des propriétés vraies pour des temps fixes à des temps aléatoires.

Les temps d'arrêt admettent certaines propriétés de stabilité, ainsi on vérifie que si τ_1 et τ_2 sont deux temps d'arrêt, alors $\tau_1 \wedge \tau_2$ est un temps d'arrêt, et que si $(\tau_n)_{n\geq 0}$ est une suite croissante de temps d'arrêt, alors $\lim_{n\to\infty} \tau_n$ est aussi un temps d'arrêt. De plus, tout temps d'arrêt τ est limite d'une suite décroissante de temps d'arrêt ne prenant qu'un nombre dénombrable de valeurs.

En effet, il suffit de poser pour tout $n \in \mathbb{N}$ et tout $\omega \in \Omega$,

$$\tau_n(\omega) \quad = \quad \sum_{k\in\mathbb{N}} 2^{-n}(k+1)\mathbf{1}_{\{2^{-n}k \leq \tau(\omega) < 2^{-n}(k+1)\}} + \infty \mathbf{1}_{\{\tau(\omega)=\infty\}}$$

pour obtenir une telle suite décroissante $(\tau_n)_{n\geq 0}$ de temps d'arrêt.

Exercice : vérifier ces affirmations.

Soit τ un temps d'arrêt, on lui associe une tribu $\mathcal{F}_\tau$ définie de la manière suivante : pour tout $A \in \mathcal{F}_\infty \stackrel{\text{déf.}}{=} \vee_{t\geq 0} \mathcal{F}_t \subset \mathcal{F}$,

$$A \in \mathcal{F}_\tau \iff \forall\, t \geq 0, \qquad A \cap \{\tau \leq t\} \in \mathcal{F}_t$$

Il apparaît sans difficulté que $\mathcal{F}_\tau$ est effectivement une tribu. D'une certaine manière, elle représente l'ensemble des événements antérieurs au temps aléatoire τ.

Remarquons que le temps d'arrêt τ est $\mathcal{F}_\tau$-mesurable et que si τ_1 et τ_2 sont deux temps d'arrêt tels que $\tau_1 \leq \tau_2$, alors $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

Si τ est un temps d'arrêt ne prenant qu'un nombre dénombrable de valeurs, disons $\{t_i\,;\, i \in \mathbb{N}\} \subset \overline{\mathbb{R}}_+$ (la suite $(t_i)_{i\in\mathbb{N}}$ n'est pas nécessairement croissante, mais on suppose que tous les t_i, $i \in \mathbb{N}$, sont distincts), on vérifie que $\mathcal{F}_\tau$ est l'ensemble des événements de la forme

$$\bigsqcup_{i\in\mathbb{N}} A_i \cap \{\tau = t_i\}$$

où pour tout $i \in \mathbb{N}$, $A_i \in \mathcal{F}_{t_i}$.

Donnons-nous maintenant un processus de Markov sous la forme (8), ce qui nous amène à considérer l'espace filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}) = (S^{\mathbb{R}_+}, \mathcal{S}^{\otimes \mathbb{R}_+}, (\mathcal{G}_t)_{t\geq 0})$.

Soit τ un temps d'arrêt fini (i.e. à valeurs dans $\mathbb{R}_+$) ne prenant qu'un nombre dénombrable de valeurs, que l'on note à nouveau $\{t_i\,;\, i \in \mathbb{N}\} \subset \mathbb{R}_+$.

On définit alors une application $\theta_\tau : S^{\mathbb{R}_+} \to S^{\mathbb{R}_+}$, en posant $\theta_\tau(\omega) = \theta_{t_i}(\omega)$ si $\omega \in S^{\mathbb{R}_+}$ est tel que $\tau(\omega) = t_i$. Constatons que cette application θ_τ est mesurable relativement aux deux tribus $\mathcal{S}^{\otimes \mathbb{R}_+}$. En effet, de par la structure produit à l'arrivée, il suffit de voir que pour tout $t \geq 0$, $\omega \mapsto Y_t(\theta_\tau(\omega))$ est $\mathcal{S}^{\otimes \mathbb{R}_+}$-mesurable. Mais ceci découle du fait que pour tout $x \in S$,

$$\{Y_t \circ \theta_\tau = x\} \;=\; \bigsqcup_{i\in\mathbb{N}} \{Y_{t+t_i} = x\} \cap \{\tau = t_i\} \in \mathcal{S}^{\otimes \mathbb{R}_+}$$

En posant $Y_\tau(\omega) = Y_{\tau(\omega)}(\omega)$, on montre de même que cette application $Y_\tau : (S^{\mathbb{R}_+}, \mathcal{S}^{\otimes \mathbb{R}_+}) \to (S, \mathcal{S})$ est mesurable, du fait de la dénombrabilité de l'image de τ.

Proposition 4

Soit τ un temps d'arrêt fini ne prenant qu'un nombre dénombrable de valeurs. Soit $F : S^{\mathbb{R}_+} \to \mathbb{R}_+$ mesurable et $s \geq 0$. L'application $S^{\mathbb{R}_+} \ni \omega \mapsto \mathbb{E}_{s+\tau(\omega), Y_\tau(\omega)}[F]$ est $\mathcal{F}_\tau$-mesurable, et on a plus précisément, pour tout $x \in S$, $\mathbb{P}_{s,x}$-p.s.,

$$\mathbb{E}_{s,x}[F \circ \theta_\tau | \mathcal{F}_\tau] = \mathbb{E}_{s+\tau, Y_\tau}[F]$$

Preuve :

Pour la première affirmation, notons que par définition,

$$\mathbb{E}_{s+\tau, Y_\tau}[F] \;=\; \sum_{x\in S} \sum_{i\in\mathbb{N}} \mathbb{E}_{s+t_i, x}[F] \mathbf{1}_{\{\tau = t_i,\, Y_{t_i} = x\}}$$

et il suffit donc de voir que pour tout $i \in \mathbb{N}$ et tout $x \in S$, l'événement $\{\tau = t_i,\, Y_{t_i} = x\}$ appartient à $\mathcal{F}_\tau$. Soit donc $t \geq 0$, on a

$$\{\tau = t_i,\, Y_{t_i} = x\} \cap \{\tau \leq t\} \;=\; \begin{cases} \{\tau = t_i,\, Y_{t_i} = x\} \in \mathcal{F}_{t_i} \subset \mathcal{F}_t & \text{, si } t \geq t_i \\ \emptyset \in \mathcal{F}_t & \text{, si } t < t_i \end{cases}$$

d'où le résultat annoncé.

Pour le second point de la proposition, soit $A \in \mathcal{F}_\tau$, il nous faut et suffit de voir que

$$\mathbb{E}_{s,x}[F \circ \theta_\tau \mathbb{1}_A] = \mathbb{E}_{s,x}[\mathbb{E}_{s+\tau,Y_\tau}[F]\mathbb{1}_A]$$

or on a vu que l'on pouvait écrire A sous la forme $\sqcup_{i\in\mathbb{N}} A_i \cap \{\tau = t_i\}$, avec $A_i \in \mathcal{F}_{t_i}$ pour $i \in \mathbb{N}$, ainsi par σ-additivité, il suffit de vérifier que pour tout $i \in \mathbb{N}$,

$$\mathbb{E}_{s,x}[F \circ \theta_\tau \mathbb{1}_{A_i\cap\{\tau=t_i\}}] = \mathbb{E}_{s,x}[\mathbb{E}_{s+t_i,Y_{t_i}}[F]\mathbb{1}_{A_i\cap\{\tau=t_i\}}]$$

Cependant en utilisant que $A_i \cap \{\tau = t_i\} \in \mathcal{F}_{t_i}$, on calcule que

$$\begin{aligned}\mathbb{E}_{s,x}[F \circ \theta_{t_i} \mathbb{1}_{A_i\cap\{\tau=t_i\}}] &= \mathbb{E}_{s,x}[\mathbb{E}_{s,x}[F \circ \theta_{t_i}|\mathcal{F}_{t_i}]\mathbb{1}_{A_i\cap\{\tau=t_i\}}] \\ &= \mathbb{E}_{s,x}[\mathbb{E}_{s+t_i,Y_{t_i}}[F]\mathbb{1}_{A_i\cap\{\tau=t_i\}}]\end{aligned}$$

□

<u>Exercice :</u> soit τ un temps d'arrêt ne prenant qu'un nombre dénombrable de valeurs mais non nécessairement fini, on définit alors θ_τ comme précédemment, mais seulement sur $\{\omega \in S^{\mathbb{R}_+} : \tau(\omega) < +\infty\} \in \mathcal{D}_\tau$. On munit cet ensemble de la tribu trace de $\mathcal{D}$, les arguments ci-dessus montrent alors que l'application $\theta_\tau : \{\tau < +\infty\} \to S^{\mathbb{R}_+}$ reste mesurable. Prouver que pour tout $s \geq 0$, tout $x \in S$ et toute application $F : S^{\mathbb{R}_+} \to \mathbb{R}_+$ mesurable,

$$\mathbb{E}_{s,x}[\mathbb{1}_{\{\tau<+\infty\}} F \circ \theta_\tau|\mathcal{F}_\tau] = \mathbb{1}_{\{\tau<+\infty\}}\mathbb{E}_{s+\tau,Y_\tau}[F] \tag{10}$$

On voudrait prolonger ce résultat à tous les temps d'arrêt, mais dans le cadre actuel ceci est pratiquement impossible, ne serait ce que parce que $\theta_\tau : \{\tau < +\infty\} \to S^{\mathbb{R}_+}$ n'est plus en général nécessairement mesurable. Nous allons voir dans la section suivante une manière de procéder, en imposant certaines restrictions.

1.2 Régularité des trajectoires et propriété de Markov forte

L'espace $(S^{\mathbb{R}_+}, \mathcal{S}^{\otimes\mathbb{R}_+})$ considéré dans la section précédente est encore trop « gros » dans la majorité des applications pour être bien manipulé, et la propriété de Markov que nous avons définie ne suffit pas à assurer que les trajectoires du processus se comportent agréablement, notamment il peut y avoir $\mathbb{P}$-p.s. une infinité de sauts dans tout intervalle ouvert de temps (considérer des exemples où tous les X_t, $t \geq 0$, sont indépendants). Pour éviter ceci on va imposer une première condition de régularité.

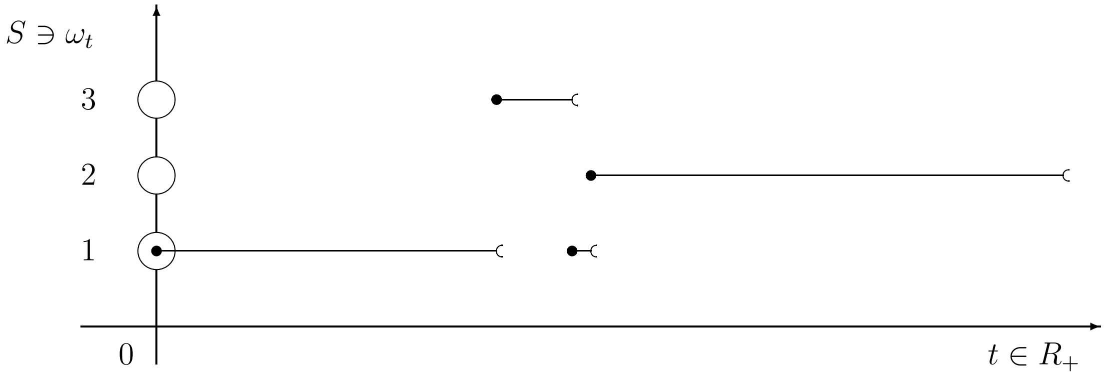
L'espace des états S étant muni de sa topologie discrète, on notera $D(\mathbb{R}_+, S)$ l'ensemble des applications $\omega : \mathbb{R}_+ \to S$ qui sont càdlàgs, i.e. qui en tous points de $\mathbb{R}_+$ sont continues à droite et admettent une limite à gauche (sauf évidemment en 0). Vu que S est fini, ceci signifie simplement que pour tout $t \geq 0$, il existe $\epsilon > 0$ tel que

$$\begin{aligned}\forall\ (t-\epsilon)_+ \leq s < t, \qquad \omega_s &= \omega_{t-\epsilon} \\ \forall\ t \leq s \leq t+\epsilon, \qquad \omega_s &= \omega_t\end{aligned}$$

On notera pour $\omega \in D(\mathbb{R}_+, S)$ et $t > 0$, ω_{t-} la limite de ω à gauche de t, en convenant que $\omega_{0-} = \omega_0$. On dira que ω admet un saut en t si $\omega_t \neq \omega_{t-}$.

<u>Exercice (de compacité) :</u> vérifier que pour tout $A > 0$, $\omega \in D(\mathbb{R}_+, S)$ n'admet qu'un nombre fini de sauts dans $[0, A]$.

Une telle fonction ω a donc shématiquement un aspect en palier comme ci-dessous (avec $S = \{1, 2, 3\}$) :



Pour $t \geq 0$, notons $X_t : D(I\!R_+, S) \to S$, la $t^{\text{ième}}$ application coordonnée canonique. $D(I\!R_+, S)$ peut être vu comme un sous-ensemble de $S^{I\!R_+}$ et X_t est donc la restriction à $D(I\!R_+, S)$ de Y_t, cette dernière notation correspondant à celle de la section précédente.

Exercice* : montrer que, malheureusement, dès que $\mathrm{card}(S) \geq 2$, $D(I\!R_+, S) \notin \mathcal{S}^{\otimes I\!R_+}$.

On posera aussi $\mathcal{D} = \sigma(X_t\,;\, t \geq 0)$, puis pour $t \geq 0$, $\mathcal{D}_t = \sigma(X_s\,;\, 0 \leq s \leq t)$. On reconsidère également les opérateurs de translations $\theta_t : (D(I\!R_+, S), \mathcal{D}) \to (D(I\!R_+, S), \mathcal{D})$, $t \geq 0$, qui sont définis par

$$\forall\, \omega \in D(I\!R_+, S),\ \forall\, s \geq 0, \qquad X_s(\theta_t(\omega)) = X_{t+s}(\omega)$$

ce qui montre qu'ils sont mesurables.

Une propriété de mesurabilité plus importante de l'espace $D(I\!R_+, S)$ est la suivante :

Lemme 5

Pour tout $t \geq 0$, l'application

$$G_t : [0, t] \times D(I\!R_+, S) \ni (s, \omega) \mapsto X_s(\omega) \in S$$

est $\mathcal{B}([0, t]) \otimes \mathcal{D}_t$-mesurable.

Preuve :

Si $t = 0$, l'affirmation est triviale. Sinon pour $n \in I\!N$, soit l'application

$$G_{t,n} : [0, t] \times D(I\!R_+, S) \ni (s, \omega) \mapsto X_{2^{-n}t\lceil 2^n \frac{s}{t}\rceil}(\omega) \in S$$

où pour tout $u \in I\!R_+$, $\lceil u \rceil = \inf\{n \in I\!N : n \geq u\}$, le plus petit entier plus grand que u. Cette application est $\mathcal{B}([0, t]) \otimes \mathcal{D}_t$-mesurable, car pour tout $x \in S$, on a

$$\begin{aligned} &\{(s, \omega) \in I\!R_+ \times D(I\!R_+, S) : G_{t,n}(s, \omega) = x\} \\ &= \{0\} \times \{\omega \in D(I\!R_+, S) : X_0(\omega) = x\} \bigsqcup \\ &\quad \bigsqcup_{0 \leq k < 2^n}]2^{-n}kt, 2^{-n}(k+1)t] \times \{\omega \in D(I\!R_+, S) : X_{2^{-n}(k+1)t}(\omega) = x\} \in \mathcal{B}([0, t]) \otimes \mathcal{D}_t \end{aligned}$$

Cependant la continuité à droite de X implique que pour tout $(t, \omega) \in I\!R_+ \times D(I\!R_+, S)$,

$$\lim_{n \to \infty} G_n(t, \omega) \;=\; G(t, \omega)$$

d'où en fin de compte la mesurabilité recherchée.

$\square$

De la même manière, on se convainc que l'application

$$G\,:\,\mathbb{R}_+\times D(\mathbb{R}_+,S)\ni(s,\omega)\mapsto X_s(\omega)\in S$$

est $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{D}$-mesurable.

Exercice* (continuité à droite de la filtration) : Pour $t\geq 0$, notons $\mathcal{D}_{t+}=\cap_{h>0}\mathcal{D}_{t+h}$. Montrer que $\mathcal{D}_{t+}=\mathcal{D}_t$.

Remarque (et indication de preuve) : ci-dessus, aucune probabilité n'est nécessaire, ceci vient du fait que S est discret et de la continuité à droite de X, qui imposent qu'infinitésimalement après l'instant $t\geq 0$, le processus n'a qu'un seul comportement possible (rester au même point), c'est-à-dire qu'il n'y a pas de possibilité d'un nouvel événement (i.e. d'une nouvelle information). Cela n'est plus nécessairement le cas si l'on remplace $(S,\mathcal{S})$ par des espaces plus généraux (par exemple si $(S,\mathcal{S})=(\mathbb{R}_+,\mathcal{B}(\mathbb{R}_+))$, immédiatement après l'instant t, la trajectoire peut commencer par croître ou décroître ...), typiquement il faut alors faire intervenir une « bonne » probabilité $\mathbb{P}$ et compléter les tribus par les ensembles négligeables (rappelons que dans notre contexte, ceux-ci sont les $A\subset D(\mathbb{R}_+,S)$ pour lesquels il existe $A'\in\mathcal{D}$ avec $A\subset A'$ et $\mathbb{P}(A')=0$) pour obtenir une continuité à droite de la filtration complétée.

Par analogie avec ce que l'on a vu dans la section précédente, on convient que :

Définition 6

Un processus de Markov à trajectoires régulières est la donnée d'un sextuplet

$$(D(\mathbb{R}_+,S),\mathcal{D},(\mathcal{D}_t)_{t\geq 0},(\theta_t)_{t\geq 0},(X_t)_{t\geq 0},(\mathbb{P}_{t,x})_{t\geq 0,x\in S})$$

où la famille des probabilités $(\mathbb{P}_{t,x})_{t\geq 0,x\in S}$ sur $(D(\mathbb{R}_+,S),\mathcal{D})$ satisfait

$$\forall\ t\geq 0,\ \forall\ x\in S,\qquad \mathbb{P}_{t,x}[X_0=x]=1$$

et pour tous $s,t\geq 0$, tout $x\in S$ et toute application mesurable $F\,:\,D(\mathbb{R}_+,S)\to\mathbb{R}_+$,

$$\mathbb{E}_{s,x}[F\circ\theta_t|\mathcal{D}_t]=\mathbb{E}_{s+t,X_t}[F]\qquad \mathbb{P}_{s,x}\text{-p.s.}$$

On dit aussi dans ce cas que l'on s'est placé sur l'espace canonique $D(\mathbb{R}_+,S)$ des trajectoires régulières. D'ailleurs puisqu'ici la partie $(D(\mathbb{R}_+,S),\mathcal{D},(\mathcal{D}_t)_{t\geq 0},(\theta_t)_{t\geq 0},(X_t)_{t\geq 0})$ est canonique, pour simplifier on parlera souvent du processus de Markov à trajectoires régulières donné par $(\mathbb{P}_{t,x})_{t\geq 0,x\in S}$.

Remarquons comme précédemment, que $X=(X_t)_{t\geq 0}$ est markovien au sens de la définition 1 sous chacune des lois $\mathbb{P}_{t,x}$.

Remarque : soit Y un processus de Markov sur un espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$, on dit que Y est $\mathbb{P}$-p.s. càdlàg, s'il existe $N\in\mathcal{F}$ tel que $\mathbb{P}(N)=0$ et tel que pour tout $\omega\notin N$, la trajectoire $Y(\omega)\,:\,\mathbb{R}_+\ni t\mapsto Y_t(\omega)\in S$ appartienne à $D(\mathbb{R}_+,S)$. Cependant cette définition n'est pas entièrement compatible avec celle donnée plus haut, car il n'existe pas nécessairement un processus de Markov à trajectoires régulières $(\mathbb{P}_{t,x})_{t\geq 0,x\in S}$ qui lui soit associé (au sens donné par (9), pour tout $t\geq 0$, avec Y remplaçant X et $\mathcal{F}_t=\sigma(Y_s\,:\,0\leq s\leq t)$).

Exercice** : trouver un contre-exemple pour l'affirmation précédente (on en donnera un à la fin de la section 5). Par contre, vérifier que ce problème ne se pose pas si Y est de plus supposé homogène.

Nous allons maintenant vérifier qu'un tel processus satisfait la propriété de Markov forte, c'est-à-dire que l'équivalent de (10) est vrai pour tout temps d'arrêt relatif à l'espace filtré $(D(I\!\!R_+,S),\mathcal{D},(\mathcal{D}_t)_{t\geq 0})$.

On associe à un tel temps d'arrêt τ l'ensemble $A_\tau=\{\omega\in D(I\!\!R_+,S)\,:\,\tau(\omega)<+\infty\}\in\mathcal{D}_\tau$ que l'on munit de la tribu $\mathcal{A}_\tau=\{B\in\mathcal{D}\,:\,B\subset A_\tau\}$ qui est la trace de $\mathcal{D}$ sur A_τ. Commençons par voir que l'application $\theta_\tau\,:\,(A_\tau,\mathcal{A}_\tau)\to(D(I\!\!R_+,S),\mathcal{D})$ (qui évidemment est définie par $\forall\,\omega\in A_\tau,\ \forall\,t\geq 0,\ X_t(\theta_\tau(\omega))=X_{t+\tau(\omega)}(\omega)$) est mesurable. Par définition de $\mathcal{D}$, il suffit de voir que pour tout $t\geq 0$, l'application

$$A_\tau\ni\omega\mapsto X_t(\theta_\tau(\omega))\in S$$

est $\mathcal{A}_\tau$-mesurable. Cependant notons que cette fonction peut s'écrire comme la composition de l'application G définie après le lemme 5 avec

$$(A_\tau,\mathcal{A}_\tau)\ni\omega\mapsto(t+\tau(\omega),\omega)\in(I\!\!R_+\times D(I\!\!R_+,S),\mathcal{B}(I\!\!R_+)\otimes\mathcal{D})$$

or ces deux dernières sont mesurables, ce qui prouve le résultat annoncé.

On pose aussi pour $\omega\in A_\tau$, $X_\tau(\omega)=X_{\tau(\omega)}(\omega)$. Ce qui précède montre alors que $X_\tau\,:\,(A_\tau,\mathcal{A}_\tau,\mathcal{D})\to S$ est mesurable, mais en fait on a un peu mieux : elle est $\widetilde{\mathcal{A}}_\tau$-mesurable, où $\widetilde{\mathcal{A}}_\tau=\{B\in\mathcal{D}_\tau\,:\,B\subset A_\tau\}$ est la trace de $\mathcal{D}_\tau$ sur A_τ.

En effet, soit $t\geq 0$ et $x\in S$, il faut vérifier que

$$\{\omega\in A_\tau\,:\,G(\tau(\omega),\omega)=x\}\cap\{\tau\leq t\}\in\mathcal{D}_t$$

i.e.

$$\{\omega\in D(I\!\!R_+,S)\,:\,G_t(t\wedge\tau(\omega),\omega)=x\}\cap\{\tau\leq t\}\in\mathcal{D}_t$$

Du fait que $t\wedge\tau$ est un temps d'arrêt plus petit que le temps d'arrêt constant t, il apparaît que $t\wedge\tau$, qui est $\mathcal{D}_{t\wedge\tau}$-mesurable, est aussi $\mathcal{D}_t$-mesurable. Ainsi par composition de G_t avec l'application

$$(D(I\!\!R_+,S),\mathcal{D}_t)\ni\omega\mapsto(t\wedge\tau(\omega),\omega)\in([0,t]\times D(I\!\!R_+,S),\mathcal{B}([0,t])\otimes\mathcal{D}_t)$$

on fait apparaître la $\mathcal{D}_t$-mesurabilité de $D(I\!\!R_+,S)\ni\omega\mapsto G_t(t\wedge\tau(\omega),\omega)$, dont découle l'affirmation précédente.

Proposition 7 (propriété de Markov forte)

Soit τ un temps d'arrêt sur $(D(I\!\!R_+,S),\mathcal{D},(\mathcal{D}_t)_{t\geq 0})$. Soit $F\,:\,D(I\!\!R_+,S)\to I\!\!R_+$ mesurable et $s\geq 0$. L'application $D(I\!\!R_+,S)\ni\omega\mapsto I\!\!E_{s+\tau(\omega),X_\tau(\omega)}[F]$ est $\mathcal{D}_\tau$-mesurable, et on a plus précisément, pour tout $x\in S$, $I\!\!P_{s,x}$-p.s.,

$$I\!\!E_{s,x}[\mathbb{1}_{\{\tau<+\infty\}}F\circ\theta_\tau|\mathcal{D}_\tau]=\mathbb{1}_{\{\tau<+\infty\}}I\!\!E_{s+\tau,X_\tau}[F]$$

Preuve :

Par un argument usuel de classe monotone, il suffit de prouver ces affirmations dans le cas où F est de la forme,

$$F(\omega)\quad=\quad\prod_{0\leq i\leq n}\mathbb{1}_{\{x_i\}}(X_{t_i}(\omega))$$

avec $n\in I\!\!N$, $0\leq t_0<t_1<\cdots<t_n$ et $x_0,\cdots,x_n\in S$.

On peut de plus clairement se ramener au cas où $t_0 = 0$. Donnons-nous donc une telle application F jusqu'à la fin de cette démonstration.

Soit $x \in S$ fixé, commençons par vérifier que $\mathbb{R}_+ \ni t \mapsto \mathbb{E}_{t,x}[F]$ est continue à droite. Pour ceci rappelons que

$$\begin{aligned}\mathbb{E}_{t,x}[F] &= \mathbb{E}_{t,x}\left[\prod_{0\leq i\leq n} \mathbf{1}_{\{x_i\}}(X_{t_i})\right] \\ &= \mathbf{1}_{\{x_0\}}(x)\prod_{0\leq i<n} \mathbb{P}_{t+t_i,x_i}(X_{t_{i+1}-t_i} = x_{i+1})\end{aligned}$$

où il apparaît qu'il suffit de montrer que pour tout $s > 0$ et tous $y, z \in S$ fixés, l'application $\mathbb{R}_+ \ni t \mapsto \mathbb{P}_{t,y}(X_s = z)$ est continue à droite.

Cependant, par convergence dominée et continuité à droite de X, pour $t \geq 0$ fixé, l'application $\mathbb{R}_+ \ni u \mapsto \mathbb{P}_{t,y}[X_u = y]$ est aussi continue à droite, ainsi puisque $\mathbb{P}_{t,y}[X_0 = y] = 1$, pour $h > 0$ assez petit on aura $\mathbb{P}_{t,y}[X_h = y] > 0$. Supposons de plus que $h < s$, on a alors par la propriété de Markov,

$$\begin{aligned}\mathbb{P}_{t+h,y}[X_s = z] &= \mathbb{P}_{t,y}[X_{s+h} = z|X_h = y] \\ &= \frac{\mathbb{P}_{t,y}[X_{s+h} = z,\, X_h = y]}{\mathbb{P}_{t,y}[X_h = y]}\end{aligned}$$

et comme précédemment, il est clair que le numérateur et le dénominateur sont continus à droite en h, d'où la convergence voulue de cette expression vers $\mathbb{P}_{t,y}[X_s = z]$ quand h tend vers 0_+.

On en déduit que

$$\mathbb{R}_+ \times S \ni (t, x) \mapsto \mathbb{E}_{t,x}[F]$$

est $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{S}$-mesurable. En effet, puisque S est discret, il suffit que pour tout $x \in S$ fixé, $\mathbb{R}_+ \ni t \mapsto \mathbb{E}_{t,x}[F]$ soit mesurable, ce qui découle classiquement de la continuité à droite.

<u>Exercice :</u> soit $f : \mathbb{R}_+ \to \mathbb{R}$, une application continue à droite, montrer que f est mesurable.

La $\mathcal{D}_\tau$-mesurabilité de $D(\mathbb{R}_+, S) \ni \omega \mapsto \mathbf{1}_{\{\tau<+\infty\}}\mathbb{E}_{\tau(\omega),X_\tau(\omega)}[F]$ en découle par composition, car on a déjà vu avant l'énoncé de la proposition que $A_\tau \ni \omega \mapsto (\tau(\omega), X_\tau(\omega))$ était $\widetilde{\mathcal{A}}_\tau$-mesurable, puis on réutilise que $A_\tau \in \mathcal{D}_\tau$.

Pour vérifier la seconde affirmation de la proposition, soit $A \in \mathcal{D}_\tau$, il nous suffit de voir que

$$\mathbb{E}_{x,t}[F \circ \theta_\tau \mathbf{1}_{\{\tau<+\infty\}\cap A}] = \mathbb{E}_{x,t}[\mathbb{E}_{t+\tau,X_\tau}[F]\mathbf{1}_{\{\tau<+\infty\}\cap A}]$$

Pour ceci, soit $(\tau_n)_{n\in\mathbb{N}}$ une suite décroissante de temps d'arrêt ne prenant qu'un nombre dénombrable de valeurs, telle que pour tout $\omega \in D(\mathbb{R}_+, S)$, $\lim_{n\to\infty} \tau_n(\omega) = \tau(\omega)$ et telle que pour tout $n \in \mathbb{N}$, $\{\tau_n = +\infty\} = \{\tau = +\infty\}$. Il est clair, d'après les arguments donnés dans la preuve de la proposition 4, que pour tout $n \in \mathbb{N}$,

$$\mathbb{E}_{s,x}[\mathbf{1}_{\{\tau<+\infty\}}F \circ \theta_{\tau_n}|\mathcal{D}_{\tau_n}] = \mathbf{1}_{\{\tau<+\infty\}}\mathbb{E}_{s+\tau_n,X_{\tau_n}}[F]$$

D'autre part, par continuité à droite et convergence monotone, on a

$$\mathbb{E}_{x,t}[F \circ \theta_\tau \mathbf{1}_{\{\tau<+\infty\}\cap A}] = \lim_{n\to\infty} \mathbb{E}_{x,t}[F \circ \theta_{\tau_n} \mathbf{1}_{\{\tau<+\infty\}\cap A}]$$

or en utilisant que $A \in \mathcal{D}_\tau \subset \mathcal{D}_{\tau_n}$, on obtient pour tout $n \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}_{x,t}[F \circ \theta_{\tau_n} \mathbf{1}_{\{\tau<+\infty\}\cap A}] &= \mathbb{E}_{x,t}[\mathbb{E}_{x,t}[\mathbf{1}_{\{\tau<+\infty\}} F \circ \theta_{\tau_n} | \mathcal{D}_{\tau_n}] \mathbf{1}_A] \\ &= \mathbb{E}_{x,t}[\mathbb{E}_{s+\tau_n, X_{\tau_n}}[F] \mathbf{1}_{\{\tau<+\infty\}\cap A}] \end{aligned}$$

et ce dernier terme admet pour limite pour n grand, $\mathbb{E}_{x,t}[\mathbb{E}_{s+\tau,X_\tau}[F]\mathbf{1}_A]$, ceci par convergence dominée, du fait que sur $\{\tau < +\infty\}$,

$$\lim_{n\to\infty} \mathbb{E}_{s+\tau_n,X_{\tau_n}}[F] = \mathbb{E}_{s+\tau,X_\tau}[F]$$

à nouveau par continuité à droite.

$\square$

Remarque : en fait, l'existence des limites à gauche n'a pas été utilisée dans la preuve ci-dessus, ni pour le lemme 5, et on aurait donc pu se placer sur l'espace des trajectoires continues à droite tout en gardant la validité de la propriété de Markov forte.

Pour terminer cette section, voyons comment le contexte du temps discret peut s'injecter dans le cadre actuel.

Soit $(p_n)_{n\in\mathbb{N}}$ une suite de noyaux de probabilités de transition, à l'aide de la version en temps discret de la proposition 2 (qui d'ailleurs est alors plus simple, car elle se généralise à des espaces d'états mesurables quelconques, sans hypothèses topologiques, cf. par exemple la proposition V.1.1 p. 154 de [10]), on construit comme dans la première section (en définissant pour $m < n \in \mathbb{N}$, $p_{m,n} = \prod_{m\leq j<n} p_j$) un sextuplet

$$(S^{\mathbb{N}}, \mathcal{S}^{\otimes\mathbb{N}}, (\mathcal{G}_n)_{n\in\mathbb{N}}, (\theta_n)_{n\in\mathbb{N}}, (Y_n)_{n\in\mathbb{N}}, (Q_{n,x})_{n\in\mathbb{N},x\in S})$$

tel que pour tous $m, n \in \mathbb{N}$, tous $x, y \in S$ et toute application mesurable $F : S^{\mathbb{R}_+} \to \mathbb{R}_+$,

$$\begin{aligned} Q_{m,x}[Y_0 = x] &= 1 \\ \mathbb{E}_{Q_{m,x}}[F \circ \theta_n | \mathcal{G}_n] &= \mathbb{E}_{Q_{m+n,Y_n}}[F], \qquad Q_{m,x}\text{-p.s.} \\ p_m(x,y) &= \mathbb{P}_{m,x}[X_1 = y] \end{aligned}$$

Pour $0 \leq s < 1$, soit $\Psi_s : (S^{\mathbb{N}}, \mathcal{S}^{\otimes\mathbb{N}}) \to (D(\mathbb{R}_+, S), \mathcal{D})$, l'application qui à une suite $y = (y_n)_{n\in\mathbb{N}}$ d'éléments de S associe la fonction $\Psi_s(y) \in D(\mathbb{R}_+, S)$ déterminée par

$$\forall\, u \geq 0, \qquad X_u(\Psi_s(y)) = Y_{\lfloor s+u \rfloor}$$

($\lfloor \cdot \rfloor$ désignant la traditionnelle partie entière).

L'application Ψ_s est manifestement mesurable.

Soit $t \in \mathbb{R}_+$ et $x \in S$, on note $\mathbb{P}_{t,x}$ l'image de $Q_{\lfloor t \rfloor,x}$ par $\Psi_{t-\lfloor t \rfloor}$. On vérifie sans difficulté que la famille $(\mathbb{P}_{t,x})_{t\geq 0,x\in S}$ définie un processus de Markov à trajectoires régulières au sens de la définition 6.

Mais ce processus correspond plus à un déguisement d'une chaîne de Markov qu'à l'image que l'on pourrait se faire d'un « véritable » processus à temps continu, qui d'une certaine manière « évoluerait régulièrement en moyenne ». Pour éliminer ce type d'exemples, on va introduire une autre condition, faisant intervenir la « structure différentielle » de $\mathbb{R}_+$ sur le semi-groupe.

1.3 Semi-groupes, générateurs et mesures invariantes

On suppose désormais que l'on dispose d'un processus de Markov à trajectoires régulières.

Soit $F(S)$ l'ensemble des fonctions réelles définies sur S, l'avantage d'avoir pris S fini étant que ces fonctions sont toutes continues bornées. Pour $0 \leq s \leq t$, soit $P_{s,t}$ l'opérateur agissant sur $F(S)$ par

$$\forall\ f \in F(S), \forall\ x \in E, \qquad P_{s,t}(f)(x) \quad = \quad \mathbb{E}_{s,x}[f(X_{t-s})]$$

Puisque $F(S)$ est un espace vectoriel de dimension finie (= le cardinal de S), on identifie souvent l'opérateur $P_{s,t}$ avec la matrice $p_{s,t}$ qui le représente dans la base $(\mathbb{1}_{\{x\}})_{x\in S}$, et qui est donc définie par

$$\forall\ x, y \in S, \qquad p_{s,t}(x,y) = \mathbb{P}_{s,x}[X_{t-s} = y]$$

Le fait que cette matrice soit markovienne se traduit sur $P_{s,t}$ par

$$\begin{array}{rrcl} \forall\ f \in F(S), & f \geq 0 & \Longrightarrow & P_{s,t}(f) \geq 0 \\ & P_{s,t}(\mathbb{1}) & = & \mathbb{1} \end{array}$$

où $\mathbb{1} \in F(S)$ est la fonction prenant toujours la valeur 1.

Notons de plus que pour $0 \leq s \leq t \leq u$, on vérifie immédiatement que

$$P_{s,u} \quad = \quad P_{s,t} \circ P_{t,u} \tag{11}$$

on dit alors que la famille $(P_{s,t})_{0\leq s\leq t}$ forme un semi-groupe inhomogène markovien sur $F(S)$ (ou sur S par abus de language, et pour simplifier, on parlera aussi tout simplement de semi-groupe).

La régularité des trajectoires de X implique par convergence dominée que pour tout $t \geq 0$ et tout $f \in F(S)$, $\mathbb{R}_+ \ni h \mapsto P_{t,t+h}(f) \in F(S)$ est càdlàg, c'est-à-dire que l'application $\mathbb{R}_+ \ni h \mapsto P_{t,t+h}$, à valeurs dans l'espace des opérateurs sur $F(S)$, est càdlàg. Nous allons maintenant demander plus. Si elle existe, on note L_t^+ la limite suivante

$$L_t^+ \quad = \quad \lim_{h\to 0+} \frac{P_{t,t+h} - \mathrm{Id}}{h}$$

(Id représentant l'opérateur identité).

De même pour $t > 0$, si la limite suivante existe, on pose

$$L_t^- \quad = \quad \lim_{h\to 0+} \frac{P_{t-h,t} - \mathrm{Id}}{h}$$

Ces dérivées à droite et à gauche sont évidemment comprises comme des opérateurs sur $F(S)$, et il apparaît facilement que ce sont en fait des générateurs, au sens suivant : un opérateur L sur $F(S)$ est un générateur, si en notant l la matrice associée (dans la base $(\mathbb{1}_{\{x\}})_{x\in S}$), on a

$$\begin{array}{rrcl} \forall\ x \neq y \in S, & l(x,y) & \geq & 0 \\ \forall\ x \in S, & \sum_{y\in S} l(x,y) & = & 0 \end{array}$$

Définition 8

> Soit $t > 0$, on dit que semi-groupe inhomogène markovien $(P_{s,t})_{0\leq s\leq t}$ est *dérivable en* t, si L_t^+ et L_t^- existent et s'il sont égaux. On note alors L_t leur valeur commune, qui est appelé *le générateur au temps* t. Si $t = 0$, on convient que $L_0 = L_0^+$.
>
> La famille $(P_{s,t})_{0\leq s\leq t}$ est dite *de classe* C^1, si $\mathbb{R}_+ \ni t \mapsto L_t$ est une application continue.

La propriété de semi-groupe (11) permet de voir qu'une famille $(P_{s,t})_{0\leq s\leq t}$ de classe C^1 est effectivement continûment dérivable en chacune des deux variables temporelles.

Proposition 9

Soit $(P_{s,t})_{0\leq s\leq t}$ un semi-groupe de classe C^1, alors pour tout $s\geq 0$, l'application

$$[s,+\infty[\,\ni t\mapsto P_{s,t}$$

est dérivable de dérivée $\dfrac{d}{dt}P_{s,t}=P_{s,t}L_t$.

De plus, pour tous $0\leq s\leq t$, l'opérateur $P_{s,t}$ est inversible.

De manière similaire, pour $t\geq 0$ fixé, l'application

$$[0,t]\ni s\mapsto P_{s,t}$$

est dérivable de dérivée $\frac{d}{ds}P_{s,t}=-L_sP_{s,t}$.

Preuve :

Soit $0\leq s<t$ et $h>0$, en écrivant

$$\frac{P_{s,t+h}-P_{s,t}}{h} \quad=\quad P_{s,t}\frac{P_{t,t+h}-\mathrm{Id}}{h}$$

on obtient que $[s,+\infty[\,\ni t\mapsto P_{s,t}$ est dérivable à droite, de dérivée $\dfrac{d^+}{dt}P_{s,t}=P_{s,t}L_t$.

Par ailleurs, pour $t>0$, notons que

$$\lim_{h\to 0_+}P_{t-h,t}=\mathrm{Id}$$

ainsi du moins pour $h>0$ assez petit, $P_{t-h,t}$ sera inversible. De plus, toujours de par l'existence de L_t^-, il apparaît que $P_{t-h,t}^{-1}$ est dérivable à droite de 0 en h, et sa dérivée vaut $-P_{t,t}^{-1}L_t^-P_{t,t}^{-1}=-L_t$. Ceci permet de voir que

$$\begin{aligned}\lim_{h\to 0_+}\frac{P_{s,t}-P_{s,t-h}}{h} &= P_{s,t}\lim_{h\to 0_+}\frac{\mathrm{Id}-P_{t-h,t}^{-1}}{h}\\ &= P_{s,t}L_t\end{aligned}$$

ce qui termine de montrer la première affirmation de la proposition.

En se remémorant le calcul classique de l'étude de l'évolution du wronskien en théorie des équations différentielles linéaires ordinaires, on en déduit que pour tout $t\geq s$,

$$\frac{d}{dt}\det(P_{s,t}) \quad=\quad \mathrm{tr}(L_t)\det(P_{s,t})$$

<u>Exercice :</u> retrouver cette égalité.

Il apparaît ainsi que

$$\begin{aligned}\det(P_{s,t}) &= \det(P_{s,s})\exp(\int_s^t\mathrm{tr}(L_u)\,du)\\ &= \exp(\int_s^t\mathrm{tr}(L_u)\,du)\end{aligned}$$

Cette quantité restant strictement positive pour tous $0\leq s\leq t$, il en découle que $P_{s,t}$ est toujours inversible.

On en déduit aisément la dernière affirmation de la proposition, en écrivant que pour $0 \leq s \leq t$, $P_{s,t} = P_{0,s}^{-1} P_{0,t}$.

$\square$

Réciproquement, remarquons que si l'on se donne une famille continue $(L_t)_{t\geq 0}$ de générateurs, en résolvant (ce qui ne présente aucun problème) pour tout $s \geq 0$ les équations

$$(12) \qquad \left\{ \begin{array}{rcll} P_{s,s} & = & \mathrm{Id} & \\ \dfrac{d}{dt} P_{s,t} & = & P_{s,t} L_t & ;\, t \geq s \end{array} \right.$$

on construit un semi-groupe inhomogène markovien $(P_{s,t})_{0\leq s\leq t}$, qui admet effectivement $(L_t)_{t\geq 0}$ pour famille de générateurs. A l'aide du théorème 2, on peut lui associer un unique processus de Markov de la forme (8), par contre il n'est pas clair a priori que l'on puisse lui faire correspondre un processus de Markov à trajectoires régulières (ce qui est facile à voir c'est que s'il en existe un, il est nécessairement unique), ce qui revient à faire « porter » les lois $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ de (8), par $D(\mathbb{R}_+, S)$. On verra comment apporter une réponse positive à cette question dans la dernière section de ce chapitre, en construisant directement et explicitement la famille de lois $(\mathbb{P}_{t,x})_{t\geq 0,\, x\in S}$ sur $(D(\mathbb{R}_+, S), \mathcal{D})$.

Remarquons que dans le cas où l'on part d'un processus de Markov à trajectoires régulières qui est homogène, le semi-groupe associé vérifie pour tous $0 \leq s \leq t$,

$$P_{s,t} \quad = \quad P_{0,t-s}$$

et on convient désormais de noter $P_t \stackrel{\text{déf.}}{=} P_{0,t}$ dans cette situation. Ce semi-groupe sera donc de classe C^1 si et seulement si $\mathbb{R}_+ \ni t \mapsto P_t$ est dérivable à droite de 0. Cette dérivée L est appelée le générateur du processus de Markov homogène à trajectoires régulières considéré, c'est aussi le générateur à tout instant du semi-groupe et on a pour tout $t \geq 0$,

$$\frac{d}{dt} P_t = L P_t = P_t L$$

ce que l'on résout immédiatement pour obtenir

$$\forall\, t \geq 0, \qquad P_t = \exp(tL)$$

Par contre, en général dans le cas inhomogène, on ne peut pas se contenter de ne considérer que les dérivées à droite L_t^+, comme le montre l'exemple construit à partir d'une chaîne de Markov à la fin de la section précédente, qui vérifie $L_t^+ = 0$ pour tout $t \geq 0$. Par contre, on peut prouver que si pour tout $t > 0$, la dérivée à gauche L_t^- existe et si l'application $\mathbb{R}_+^* \ni t \mapsto L_t^-$ est continue et admet une limite en 0_+, alors le semi-groupe est de classe C^1.

Pour finir cette section, on va introduire quelques notions associés à un générateur L donné. L'objet le plus important est certainement le suivant :

Définition 10

Une probabilité μ sur S est dite invariante par L, si

$$\forall\, f \in F(S), \qquad \mu(L(f)) = 0$$

On dit qu'elle est réversible si elle vérifie de plus

$$\forall\, f, g \in F(S), \qquad \mu(gL(f)) = \mu(fL(g))$$

i.e. si L est symétrique dans $\mathbb{L}^2(\mu)$.

On aura noté que la réversibilité implique l'invariance, car il suffit de prendre $g = \mathbb{1}$ dans l'égalité ci-dessus.

On convient désormais d'identifier L avec sa matrice $(L(x,y))_{x,y\in S}$. La probabilité μ est alors invariante (respectivement réversible) si et seulement si

$$\forall\, x \in S, \qquad \sum_{y\in S\setminus\{x\}} \mu(x)L(x,y) \;=\; \sum_{y\in S\setminus\{x\}} \mu(y)L(y,x)$$

(resp. si pour tous $x \neq y \in S$, $\mu(x)L(x,y) = \mu(y)L(y,x)$).

Nous allons maintenant étudier plus précisément ces mesures.

Une suite finie $(x_i)_{0\leq i\leq n}$ d'éléments de S est appelée un chemin relatif à L (ou un L-chemin), si pour tout $0 \leq i \leq n-1$, $L(x_i, x_{i+1}) > 0$. Un tel chemin est dit partir de x_0 et aller en x_N et sa longueur est N. Les chemins que nous considérerons seront toujours de plus injectifs, c'est-à-dire que pour tous $0 \leq i \neq j \leq N$, on a $x_i \neq x_j$. Pour $x, y \in S$, on notera $\mathcal{C}_{x,y}$ l'ensemble des chemins allant de x à y (le générateur L étant sous-entendu).

Un ensemble $A \subset S$ est appelé une classe de récurrence (pour L), si pour tout $x \in A$, on a

$$\{z \in S \,:\, \mathcal{C}_{x,z} \neq \emptyset\} = A$$

et on dit que L est irréductible si S est une classe de récurrence.

Notons $E = \{\{x,y\}\,;\, x \neq y \in S\}$ l'ensemble des arêtes sur S, non orientées et où l'on ne considère pas les auto-boucles $\{x,x\}$, pour $x \in S$. Un graphe sur S est la donnée d'un sous-ensemble G de E, et celui-ci est appelé un arbre, s'il est connexe (tous points $x \neq y \in S$ peuvent être rejoints en utilisant des arêtes de G) et sans boucles (pour tous $x \neq y \in S$, il n'existe qu'un ensemble d'arêtes de G permettant de les relier). On note $\mathcal{T}(S)$ l'ensemble des arbres sur S. Un arbre pointé est la donnée d'un arbre $T \in \mathcal{T}(S)$ et d'un point particulier $x \in S$, qui joue le rôle de racine. Soit $y \in S \setminus \{x\}$, il existe alors un unique $z \in S$ tel que $\{y,z\}$ soit l'une des arêtes de T qui permettent de relier y à x, ce point z est alors appelé le père de y, et on note $z = p_{(T,x)}(y)$. L'ensemble des couples $(y, p_{(T,x)}(y))$, $y \neq x$, qui se construisent de cette manière à partir d'un arbre pointé (T,x), est souvent appelé un x-graphe, suivant la terminologie introduite dans [5], et détermine entièrement l'arbre pointé. On convient de poser $p_{(T,x)}(x) = x$.

Au générateur L et au point $x \in S$, on peut alors associer la quantité suivante

$$M(L,x) \;=\; \sum_{T\in\mathcal{T}(S)} \prod_{y\in S\setminus\{x\}} L(y, p_{(T,x)}(y))$$

qui sera strictement positive si et seulement si pour tout $y \in S$, on a $\mathcal{C}_{x,y} \neq \emptyset$.

Proposition 11

Si L est irréductible, il existe une unique probabilité invariante μ, et elle est donnée par

$$\forall\, x \in S, \qquad \mu(x) = \frac{M(L,x)}{\sum_{y\in S} M(L,y)}$$

Preuve :

Vérifions tout d'abord que la probabilité μ définie comme dans la proposition est invariante. Il suffit de voir que pour tout $x \in S$,

$$(13) \qquad \sum_{y\in S\setminus\{x\}} M(L,x)L(x,y) \;=\; \sum_{y\in S\setminus\{x\}} M(L,y)L(y,x)$$

Pour ceci, x étant fixé, soient $T \in \mathcal{T}(S)$ et $y_0 \neq x$. On note pour $n \in \mathbb{N}$, $p^n_{(T,x)}$ la composée $n^{\text{ième}}$ de $p_{(T,x)}$. Considérons $m = \min\{n \in \mathbb{N} \,/\, p^n_{(T,x)}(y_0) = x\} \in \mathbb{N}^*$ (ce nombre est parfois appelé la hauteur de y_0 dans l'arbre pointé (T, x)), et posons $z_0 = p^{m-1}_{(T,x)}(y_0) \in S \setminus \{x\}$. On appelle ensuite $\widetilde{T}$ l'arbre que l'on obtient à partir de T en enlevant l'arête $\{z_0, x\}$ et en rajoutant $\{x, y_0\}$.

On a

$$\prod_{y\in S\setminus\{x\}} L(y, p_{(T,x)}(y))L(x, y_0) \quad = \quad \prod_{y\in S\setminus\{z_0\}} L(y, p_{(\widetilde{T},z_0)}(y))L(z_0, x) \tag{14}$$

Cependant il est clair que l'application

$$\mathcal{T}(S) \times S \setminus \{x\} \ni (T, y_0) \mapsto (\widetilde{T}, z_0) \in \mathcal{T}(S) \times S \setminus \{x\}$$

est une bijection, ainsi en sommant les égalités (14) pour $(T, y_0) \in \mathcal{T}(S) \times S \setminus \{x\}$, on obtient (13).

Pour prouver l'unicité, soit ν une autre probabilité invariante, et notons pour $x \in S$, $f(x) = \nu(x)/\mu(x)$ (on aura remarqué que μ charge tous les points de S). On a alors

$$\mu(fL(f)) \quad = \quad \nu(L(f)) = 0$$

et on calcule que

$$\begin{aligned}
\mu(fL(f)) &= \sum_{x,y\in S} \mu(x)f(x)L(x, y)f(y) \\
&= \frac{1}{2} \sum_{x,y\in S} (\mu(x)L(x, y) + \mu(y)L(y, x))f(x)f(y)
\end{aligned}$$

Or l'invariance de μ implique que pour tout $x \in S$ fixé,

$$\sum_{y\in S} \mu(y)L(y, x) \quad = \quad \sum_{y\in S} \mu(x)L(x, y) = 0$$

ce qui permet de voir que

$$\sum_{y\in S}(\mu(x)L(x, y) + \mu(y)L(y, x))f(y) \quad = \quad \sum_{y\in S}(\mu(x)L(x, y) + \mu(y)L(y, x))(f(y) - f(x))$$

et on en déduit que

$$\begin{aligned}
&\sum_{x,y\in S} (\mu(x)L(x, y) + \mu(y)L(y, x))f(x)f(y) \\
&\quad = \sum_{x,y\in S} (\mu(x)L(x, y) + \mu(y)L(y, x))f(x)(f(y) - f(x)) \\
&\quad = -\sum_{x,y\in S} (\mu(x)L(x, y) + \mu(y)L(y, x))f(y)(f(y) - f(x)) \\
&\quad = -\frac{1}{2} \sum_{x\neq y\in S} (\mu(x)L(x, y) + \mu(y)L(y, x))(f(y) - f(x))^2
\end{aligned}$$

et cette expression ne peut être nulle que si $f(x) = f(y)$ pour tous $x \neq y$ tel que $L(x, y)$ ou $L(y, x)$ est non nul, ce qui par irréductibilité implique que f est constant sur S.

Reste alors à utiliser que $\mu(f) = \nu(\mathbb{1}) = 1$ pour conclure que $f = \mathbb{1}$, c'est-à-dire que $\nu = \mu$.

□

Exercice : dans le cas d'un générateur quelconque L, montrer qu'il existe toujours au moins une classe de récurrence. Notons $R_1, \cdots, R_n$ les classes de récurrence, qui sont nécessairement disjointes, et pour chaque $1 \leq i \leq n$, soit L_i l'opérateur sur $F(R_i)$ dont la matrice est $(L(x,y))_{x,y \in R_i}$. Montrer que L_i est un générateur irréductible sur $F(R_i)$. Il admet donc d'après la proposition 10 une unique probabilité invariante μ_i sur R_i. On la prolonge en une probabilité sur S en posant $\mu_i(x) = 0$ pour tout $x \in S \setminus R_i$. Montrer que l'ensemble des probabilités invariantes pour L est alors

$$\{\mu = \sum_{1 \leq i \leq n} a_i \mu_i \,/\, \forall\ 1 \leq i \leq n,\ a_i \geq 0 \text{ et } \sum_{1 \leq i \leq n} a_i = 1\}$$

Notamment il apparaît ainsi que L est irréductible si et seulement s'il admet une unique probabilité invariante chargeant tous les points de S.

Le terme « invariant » provient du fait suivant : soit $(\mathbb{P}_{t,x})_{t \geq 0, x \in S}$ un processus de Markov à trajectoires régulières homogène admettant un générateur L, et soit μ une probabilité quelconque sur S. On note $\mathbb{P}_{0,\mu}$, la probabilité définie sur $(D(\mathbb{R}_+, S), \mathcal{D})$ par

$$\forall\ A \in \mathcal{D}, \quad \mathbb{P}_{0,\mu}(A) = \sum_{x \in S} \mu(x) \mathbb{P}_{0,x}(A)$$

notamment la loi initiale de Y sous $\mathbb{P}_{0,\mu}$ est μ.

Alors μ est invariante pour L si et seulement si pour tout $t \geq 0$, μ est la loi de X_t.

En effet, soit $f \in F(S)$, on a

$$\mathbb{E}_{0,\mu}[f(X_t)] \quad = \quad \mu(P_t(f))$$

où $(P_t)_{t \geq 0}$ est le semi-groupe homogène associé à $(\mathbb{P}_{t,x})_{t \geq 0, x \in S}$, ce qui montre que cette expression est dérivable en $t \geq 0$ et que sa dérivée vaut

$$\mu(P_t(L(f))) \quad = \quad \mu(L(P_t(f))) = 0$$

Il en découle que

$$\mathbb{E}_{0,\mu}[f(X_t)] \quad = \quad \mathbb{E}_{0,\mu}[f(X_0)] = \mu(f)$$

qui est ce qu'il fallait démontrer pour l'implication directe.

Réciproquement, si pour tout $t \geq 0$, X_t a pour loi μ, alors pour tout $f \in F(S)$,

$$\mu(L(f)) \quad = \quad \lim_{t \to 0_+} \mu\left(\frac{P_t(f) - f}{t}\right) = 0$$

ce qui prouve que μ est invariante pour L.

La terminologie « réversible » quant à elle provient de la caractérisation suivante : intéressons-nous de nouveau à un processus de Markov à trajectoires régulières homogène $(\mathbb{P}_{t,x})_{t \geq 0, x \in S}$ de générateur L, et soit μ une probabilité invariante associée. Soit $D(\mathbb{R}, S)$ l'ensemble des applications de $\mathbb{R}$ dans S qui sont càdlàgs, que l'on munit des applications coordonnées canoniques $(X_t)_{t \in \mathbb{R}}$ et de la tribu qu'elles engendrent. Du fait que μ est invariante, on peut montrer par une variante du théorème 2 (cf. par exemple le théorème p. 154 de [3], en y prenant toutes les « lois d'entrée » égales à μ) que l'on peut prolonger la probabilité $\mathbb{P}_{0,\mu}$ en une probabilité $\mathbb{P}_\mu$ sur $D(\mathbb{R}, S)$ telle que pour tout $t \in \mathbb{R}$, X_t soit de loi μ et telle que pour tous $s \leq t \in \mathbb{R}$ et tout $f \in F(S)$

$$\mathbb{E}_\mu[f(X_t) | \sigma(X_u\,;\ -\infty < u \leq s)] \quad = \quad P_{t-s}(f)(X_s)$$

Considérons X comme une variable aléatoire à valeurs dans $S^{\mathbb{R}}$, et soit $\widehat{X}$ la variable aléatoire également à valeurs dans $S^{\mathbb{R}}$ définie par $\forall\ \omega \in D(\mathbb{R}, S)$, $\widehat{X}(\omega) = (X_{-t}(\omega))_{t\geq 0}$. On munit évidemment $S^{\mathbb{R}}$ de la tribu $\mathcal{S}^{\otimes \mathbb{R}}$, de sorte que les deux applications X et $\widehat{X}$ soient mesurables, et comme il est d'usage, notons $\mathbb{P}_\mu \circ X^{-1}$ (respectivement $\mathbb{P}_\mu \circ \widehat{X}^{-1}$) la loi image de $\mathbb{P}_\mu$ par X (resp. par $\widehat{X}$).

On vérifie alors sans difficulté que μ est réversible si et seulement si $\mathbb{P}_\mu \circ X^{-1} = \mathbb{P}_\mu \circ \widehat{X}^{-1}$, ce qui revient à dire que la loi image de $\mathbb{P}_\mu$ par X reste la même par retournement du temps.

1.4 Problèmes de martingales

On considère un processus de Markov $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ régulier, ce qui désormais signifiera qu'il est à trajectoires régulières et que son semi-groupe est de classe C^1, dont la famille de générateurs est $(L_t)_{t\geq 0}$. Un des objets principaux qui vont nous intéresser ici est le processus stochastique à valeurs réelles défini par

$$\forall\ t\geq 0, \qquad M_t^{(f)} = f(X_t) - f(X_0) - \int_0^t L_u(f)(X_u)\, du$$

où $f \in F(S)$ est une fonction donnée.

Soit m_0 une probabilité sur S, rappelons que l'on note $\mathbb{P}_{0,m_0} = \sum_{x\in S} m_0(x)\mathbb{P}_{0,x}$, qui s'appelle parfois la probabilité associée à la famille $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$ au temps 0 de loi initiale m_0 (on aurait dit au temps $s \geq 0$, si l'on avait considéré $\mathbb{P}_{s,m_0} = \sum_{x\in S} m_0(x)\mathbb{P}_{s,x}$).

Proposition 12

> Sous $\mathbb{P}_{0,m_0}$, le processus $(M_t^{(f)})_{t\geq 0}$ est une martingale par rapport à la filtration $(\mathcal{D}_t)_{t\geq 0}$

Preuve :

Il s'agit de montrer que pour tous $0 \leq s \leq t$, $\mathbb{P}_{0,m_0}$-p.s.,

$$\mathbb{E}_{0,m_0}[M_t^{(f)}|\mathcal{D}_s] \ = \ M_s^{(f)}$$

Pour tout $s \geq 0$, on remarque comme dans la preuve du lemme 5, que l'application $[0,s] \times D(\mathbb{R}_+, S) \ni (u,\omega) \mapsto L_u(f)(X_u(\omega))$ est $\mathcal{B}([0,s]) \otimes \mathcal{D}_s$-mesurable. Ainsi par la partie mesurabilité du théorème de Fubini, $M_s^{(f)}$ est effectivement $\mathcal{D}_s$-mesurable. Par ailleurs, pour tout $t \geq 0$ fixé, $M_t^{(f)}$ est clairement une variable aléatoire réelle bornée, c'est-à-dire intégrable par rapport à n'importe quelle probabilité sur $(D(\mathbb{R}_+, S), \mathcal{D})$. Il suffit donc de voir que

$$\mathbb{E}_{0,m_0}\left[M_t^{(f)} - M_s^{(f)}\middle|\,\mathcal{D}_s\right] \ = \ 0$$

ce qui se traduit par

$$\mathbb{E}_{0,m_0}\left[f(X_t) - f(X_s) - \int_s^t L_u f(X_u)\, du\middle|\,\mathcal{D}_s\right] \ = \ 0$$

or l'intégrant ci-dessus s'écrit aussi

$$\left(f(X_{t-s}) - f(X_0) - \int_s^t L_u f(X_{u-s})\, du\right) \circ \theta_s$$

et par la propriété de Markov on est donc ramené à voir que

$$\mathbb{E}_{s,X_s}\left[f(X_{t-s}) - f(X_0) - \int_s^t L_u f(X_{u-s})\, du\right] \ = \ 0$$

Il suffit donc de se persuader que pour tout $x \in S$ et tout $0 \leq s \leq t$,

$$\mathbb{E}_{s,x}\left[f(X_{t-s}) - f(X_0) - \int_s^t L_u f(X_{u-s})\, du\right] = 0$$

or le membre de gauche s'écrit aussi par le théorème de Fubini,

$$P_{s,t}(f)(x) - f(x) - \int_s^t P_{s,u}(L_u f)(x)\, du$$

expression qui est dérivable en $t \geq 0$, de dérivée $P_{s,t}(L_t f)(x) - P_{s,t}(L_t f)(x) = 0$, et puisqu'en $t = s$ elle est nulle, on conclut à la validité de la proposition.

□

Réciproquement, soit m_0 une probabilité sur S, qui sera vue comme initiale, et soit $(L_t)_{t\geq 0}$ une famille continue de générateurs. Une probabilité $\mathbb{P}$ sur $(D(\mathbb{R}_+, S), \mathcal{D})$ telle que X_0 soit de loi m_0 sous $\mathbb{P}$ et telle que pour tout $f \in F(S)$, $(M_t^{(f)})_{t\geq 0}$ soit une martingale sous $\mathbb{P}$ par rapport à la filtration $(\mathcal{D}_t)_{t\geq 0}$, est appelée une solution du problème de martingales associé à m_0 et à $(L_t)_{t\geq 0}$.

Soit $(P_{s,t})_{0\leq s\leq t}$ le semi-groupe que l'on obtient en résolvant les équations (12), et supposons, comme dans la discussion qui a suivi la preuve de la proposition 9, que l'on puisse lui associer un processus de Markov à trajectoires régulières $(\mathbb{P}_{t,x})_{t\geq 0, x\in S}$. On a alors

Proposition 13

Avec les notations introduites ci-dessus, $\mathbb{P}_{0,m_0}$ est l'unique solution du problème de martingales associé à m_0 et à $(L_t)_{t\geq 0}$.

Preuve :

Intéressons-nous à l'unicité, en considérant une solution $\mathbb{P}$. Commençons par montrer que X est markovien sous $\mathbb{P}$, et pour ceci, soit $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \cdots < t_n < t_{n+1}$ et $x_0, \cdots, x_{n+1} \in S$ tels que $\mathbb{P}[X_{t_0} = x_0, \cdots, X_{t_n} = x_n] > 0$. On a déjà vu dans la première section qu'il nous suffisait de vérifier que

$$\mathbb{P}[X_{t_{n+1}} = x_{n+1}|A] = \mathbb{P}[X_{t_{n+1}} = x_{n+1}|X_n = x_n]$$

où $A = \{X_{t_0} = x_0, \cdots, X_{t_n} = x_n\}$.

Pour $t \geq t_n$ et $x \in S$, posons

$$Q_{A,t}(x) = \mathbb{P}[X_t = x|A]$$

Le fait que $(M_t^{(\mathbf{1}_{\{x\}})})_{t\geq 0}$ est une martingale et que $A \in \mathcal{D}_{t_n}$ permet d'obtenir que

$$\mathbb{E}\left[\mathbf{1}_{\{x\}}(X_t) - \mathbf{1}_{\{x\}}(X_{t_n}) - \int_{t_n}^t L_u(\mathbf{1}_{\{x\}})(X_u)\, du \middle| A\right] = 0$$

ce qui s'écrit aussi

$$Q_{A,t}(x) - Q_{A,t_n}(x) - \int_{t_n}^t \sum_{y\in S} Q_{A,u}(y) L_u(\mathbf{1}_{\{x\}})(y)\, du = 0$$

Considérons $Q_{A,t}$ comme un vecteur ligne $(Q_{A,t}(x))_{x\in S}$ et rappelons que pour tous $x, y \in S$, $L_u(\mathbf{1}_{\{x\}})(y) = L_u(x, y)$, de sorte que matriciellement,

$$Q_{A,t} = Q_{A,t_n} + \int_{t_n}^t Q_{A,u} L_u\, du$$

ce qui montre que le vecteur ligne $F_t \stackrel{\text{déf.}}{=} Q_{A,t}$ est solution du système différentiel

$$\left\{\begin{array}{rcll} \frac{d}{dt}F_t & = & F_t L_t & ; t \geq t_n \\ F_{t_n} & = & Q_{A,t_n} & \end{array}\right.$$

Ce système admet une unique solution, et pour tout $t \geq t_n$, F_t s'exprime comme une fonction de t et de « la condition initiale » Q_{A,t_n}. Cependant on aura remarqué que pour tout $x \in S$,

$$Q_{A,t_n}(x) \;=\; I\!\!P[X_{t_n} = x|A] = \delta_{x_n}(x) = I\!\!P[X_{t_n} = x|X_n = x_n]$$

il apparaît ainsi que $Q_{A,t}$ ne dépend pas de $x_0, \cdots, x_{n-1}$, et il suffit alors de sommer, en les $x_0, \cdots, x_{n-1} \in S$ tels que $I\!\!P[X_{t_0} = x_0, \cdots, X_{t_{n-1}} = x_{n-1}] > 0$, les égalités

$$I\!\!P[X_{t_0} = x_0, \cdots, X_{t_n} = x_n, X_t = x] \;=\; Q_{A,t}(x) I\!\!P[X_{t_0} = x_0, \cdots, X_{t_n} = x_n]$$

pour se rendre compte que $Q_{A,t}(x) = I\!\!P[X_t = x|X_{t_n} = x_n]$, ce qu'il fallait démontrer.

Notons $S_0 = \{x \in S \,:\, m_0(x) > 0\}$, ensemble qui est bien déterminé par m_0, et posons pour $x \in S_0$, $y \in S$ et $t \geq 0$, $Q_{0,t}(x,y) = I\!\!P[X_t = y|X_0 = x]$, d'après ce qui précède (avec $A = \{X_0 = x\}$), on a

$$\left\{\begin{array}{rcll} \frac{d}{dt}Q_{0,t}(x,y) & = & \sum_{z\in S} Q_{0,t}(x,z)L_t(z,y) & ; t \geq 0 \\ Q_{0,0}(x,y) & = & \delta_y(x) & \end{array}\right.$$

ce qui montre que $(Q_{0,t}(x,y))_{x\in S_0, y\in S}$ est entièrement déterminé par $(L_u)_{u\geq 0}$. Il en est de même pour $S_s = \{x \in S \,/\, I\!\!P[X_s = x] > 0\}$, puisque $I\!\!P[X_s = x] = \sum_{y\in S_0} m_0(y)Q_{0,s}(y,x)$, pour $s \geq 0$ et $x \in S$. Or si on pose pour $x \in S_s$ et $y \in S$, $Q_{s,t}(x,y) = I\!\!P[X_t = y|X_s = x]$, les calculs ci-dessus montrent également que

$$\left\{\begin{array}{rcll} \frac{d}{dt}Q_{s,t}(x,y) & = & \sum_{z\in S} Q_{s,t}(x,z)L_t(z,y) & ; t \geq s \\ Q_{s,s}(x,y) & = & \delta_y(x) & \end{array}\right.$$

de sorte que pour $0 \leq s \leq t$, $x \in S_s$ et $y \in S$, $Q_{s,t}(x,y)$ est aussi uniquement déterminé par $(L_u)_{u\geq 0}$ (plus précisément, $Q_{s,t}$ est la sous-matrice $S_s \times S$ de la solution $P_{s,t}$ de (12)). Cependant par la propriété de Markov, pour tous $n \in I\!\!N$, $0 = t_0 < t_1 < \cdots < t_n$ et $A_0, \cdots, A_n \subset S$, on a

$$\begin{array}{l} I\!\!P[X_0 \in A_0, \cdots, X_{t_n} \in A_n] \\ \quad = \displaystyle\sum_{x_0\in S_0, \cdots, x_n \in S_{t_n}} m_0(x_0)Q_{0,t_1}(x_0,x_1)\cdots Q_{t_{n-1},t_n}(x_{n-1},x_n)\mathbb{I}_{A_0}(x_0)\cdots\mathbb{I}_{A_n}(x_n) \end{array}$$

or comme on l'a déjà indiqué, ces marginales fini-dimensionnelles déterminent $I\!\!P$, d'où l'unicité annoncée.

Le fait que $I\!\!P_{0,m_0}$ soit effectivement une solution du problème de martingales associé à m_0 et à $(L_t)_{t\geq 0}$ découle immédiatement de la proposition 12.

$\square$

Plus généralement, soit $f \,:\, I\!\!R_+ \times S \to I\!\!R$ une application telle que pour tout $x \in S$ fixé, $I\!\!R_+ \ni t \mapsto f(t,x)$ soit de classe C^1. On notera $\partial_t f(t,x)$ la dérivée en t et $C^1(I\!\!R_+ \times S)$ désignera l'ensemble de telles fonctions f. Alors en posant

$$M_t^{(f)} \;=\; f(t,X_t) - f(0,X_0) - \int_0^t \partial_u f(u,X_u)\,du - \int_0^t L_u(f(u,\cdot))(X_u)\,du$$

on montre comme dans la preuve de la proposition 12, que $(M_t^{(f)})_{t\geq 0}$ est une $(\mathcal{D}_t)_{t\geq 0}$-martingale.

Le fait de disposer de martingales est important : cela permet de retranscire pour les processus de Markov les théorèmes bien connus qu'elles satisfont. Ainsi dans le cas homogène et irréductible, on peut obtenir, sous des renormalisations adéquates, des lois des grands nombres, des théorèmes de la limite centrale ou des lois du logarithme itéré pour les fonctionnelles « additives » du processus de Markov à trajectoires régulières X de la forme

$$\forall\ t \geq 0, \qquad F_t \stackrel{\text{déf.}}{=} \int_0^t f(X_s)\,ds$$

(l'additivité signifie que pour tous $t, s \geq 0$, $F_{t+s} = F_t + F_s \circ \theta_t$), historiquement, il semblerait d'ailleurs que ces relations, pour des processus de Markov généraux, aient été une motivation importante de l'étude des martingales.

Nous ne les présenterons pas maintenant, car on reviendra en détail dans un chapitre ultérieur sur des généralisations de tels résultats à certains processus de Markov inhomogènes réguliers.

Mais pour exploiter les martingales, il faut souvent disposer de renseignements sur leur « crochet oblique ». Pour donner une définition relativement générale de ce terme, tout en restant dans notre cadre, il nous faut considérer une nouvelle tribu $\mathcal{P}$ sur $\mathbb{R}_+ \times D(\mathbb{R}_+, S)$: soit $\mathcal{C}$ l'ensemble des applications $Z : \mathbb{R}_+ \times D(\mathbb{R}_+, S) \to \mathbb{R}$ telles que pour tout $t \geq 0$ fixé, $Z_t : D(\mathbb{R}_+, S) \ni \omega \mapsto Z(t,\omega)$ est $\mathcal{D}_t$-mesurable, et telles que pour tout $\omega \in D(\mathbb{R}_+, S)$ fixé, la trajectoire $\mathbb{R}_+ \ni t \mapsto Z(t,\omega) \in \mathbb{R}$ est continue. On appelle tribu prévisible sur $\mathbb{R}_+ \times D(\mathbb{R}_+, S)$, la tribu $\sigma(Z\,;\,Z \in \mathcal{C})$ engendrée par les applications de la forme précédente. Un processus réel $Z = (Z_t)_{t\geq 0}$ défini sur $(D(\mathbb{R}_+, S), \mathcal{D})$ est alors qualifié de prévisible, si l'application $Z : \mathbb{R}_+ \times D(\mathbb{R}_+, S) \ni (t,\omega) \mapsto Z_t(\omega)$ est $\mathcal{P}$-mesurable. Par ailleurs, si pour tout $\omega \in D(\mathbb{R}_+, S)$, la trajectoire $\mathbb{R}_+ \ni t \mapsto Z_t(\omega)$ est à variations localement finies, le processus Z est dit à variations finies, ce sera notamment le cas si pour tout $\omega \in D(\mathbb{R}_+, S)$ fixé, la trajectoire associée est absolument continue.

Soit $\mathbb{P}$ une probabilité sur $(D(\mathbb{R}_+, S), \mathcal{D})$ et $M = (M_t)_{t\geq 0}$ une $(\mathcal{D}_t)_{t\geq 0}$-martingale de carré intégrable (i.e. pour tout $t \geq 0$, $\mathbb{E}[M_t^2] < +\infty$), il existe alors un unique processus prévisible à variations finies, $Z = (Z_t)_{t\geq 0}$, tel que le processus $(M_t^2 - Z_t)_{t\geq 0}$ soit une $(\mathcal{D}_t)_{t\geq 0}$-martingale nulle au temps 0, et Z est traditionnellement appelé le crochet oblique de M.

Avant d'expliciter leur forme pour les martingales $M^{(f)}$, avec $f \in F(S)$, introduisons la notion suivante :

Soit L un générateur, on définit son carré du champs $\Gamma : F(S) \times F(S) \to F(S)$ en posant

$$\forall\ f, g \in F(S), \forall\ x \in S, \qquad \Gamma(f,g)(x) = L(fg)(x) - f(x)L(g)(x) - g(x)L(f)(x)$$

Considérons à nouveau une application $f \in C^1(\mathbb{R}_+ \times S)$, et un processus de Markov régulier X, dont la famille de générateurs est $(L_t)_{t\geq 0}$. On note pour tout $t \geq 0$,

$$\langle M^{(f)} \rangle_t = \int_0^t \Gamma_u(f(u,\,\cdot\,), f(u,\,\cdot\,))(X_u)\,du$$

où pour tout $u \geq 0$, Γ_u est le carré du champs associé à L_u.

Proposition 14

Pour toute probabilité initiale m_0, sous $\mathbb{P}_{0,m_0}$ le processus $((M_t^{(f)})^2 - \langle M^{(f)} \rangle_t)_{t\geq 0}$ est une $(\mathcal{D}_t)_{t\geq 0}$-martingale.

Preuve :

La fonction $f \in C^1(\mathbb{R}_+ \times S)$ étant fixée, on note pour $t \geq 0$, $F_t \stackrel{\text{déf.}}{=} \partial_t f(t, X_t) + L_t(f(t,\,\cdot\,))(X_t)$, puis on définit un processus réel à trajectoires continues $N = (N_t)_{t\geq 0}$ en posant

$$\forall\, t \geq 0, \qquad N_t \;=\; \int_0^t F_u\, du$$

On calcule alors que pour $t \geq 0$,

$$\begin{aligned}(f(0, X_0) + M_t^{(f)})^2 &= (f(t, X_t) - N_t)^2 \\ &= f^2(t, X_t) + N_t^2 - 2f(t, X_t)N_t\end{aligned}$$

Cependant, en introduisant la martingale $M^{(f^2)}$ associée à f^2, on a

$$f^2(t, X_t) \;=\; f^2(0, X_0) + 2\int_0^t f(u, X_u)\partial_u f(u, X_u)\, du + \int_0^t L_u(f^2(u,\,\cdot\,))(X_u)\, du + M_t^{(f^2)}$$

Par ailleurs, en utilisant les formules d'intégration par parties pour les intégrales de Stieljes et par continuité de N, on obtient, d'une part

$$N_t^2 \;=\; 2\int_0^t N_u F_u\, du$$

et d'autre part,

$$\begin{aligned}f(t, X_t)N_t &= \int_0^t N_u\, dM_u^{(f)} + \int_0^t N_u F_u\, du + \int_0^t f(u, X_{u-})F_u\, du \\ &= \int_0^t N_u\, dM_u^{(f)} + \int_0^t N_u F_u\, du + \int_0^t f(u, X_u)F_u\, du\end{aligned}$$

où on a évidemment pu remplacer dans la dernière intégrale X_{u-} par X_u, car l'ensemble des instants de sauts de X est négligeable pour la mesure de Lebesgue.

En regroupant toutes ces expressions, on fait apparaître que pour tout $t \geq 0$,

$$(M_t^{(f)})^2 \;=\; M_t + \langle M^{(f)}\rangle_t$$

avec

$$M_t \;=\; -2f(0, X_0)M_t^{(f)} + M_t^{(f^2)} + \int_0^t N_u\, dM_u^{(f)}$$

Or d'après la théorie élémentaire des intégrales stochastiques, du fait que le processus N est prévisible, on sait que $(\int_0^t N_u\, dM_u^{(f)})_{t\geq 0}$ est une martingale, d'où en fin de compte le résultat annoncé.

□

Il est manifeste que le processus $(\langle M^{(f)}\rangle_t)_{t\geq 0}$ est prévisible et à variations finies, ainsi la proposition 14 montre que $\langle M^{(f)}\rangle$ est bien le crochet de la martingale $M^{(f)}$.

Notons toutefois qu'il existe d'autres processus à variations finies tels que la propriété de la proposition 14 soit satisfaite : tout naturellement posons pour $t > 0$,

$$\begin{aligned}M_{t-}^{(f)} &= \lim_{s\to t-} M_s^{(f)} \\ &= f(X_{t-}) - f(X_0) - \int_0^t L_u(f)(X_u)\, du\end{aligned}$$

puis

$$[M^{(f)}]_t \quad = \quad \sum_{0<s\leq t} (M_s^{(f)} - M_{s-}^{(f)})^2$$

(il s'agit du fameux « crochet droit » associé à la martingale $M^{(f)}$, on renvoit, pour la définition générale de ce terme à tout livre traitant de la théorie des martingales à sauts en temps continu, par exemple [2] ou [8]), on peut alors vérifier que sous $\mathbb{P}_{0,m_0}$ le processus $((M_t^{(f)})^2 - [M^{(f)}]_t)_{t\geq 0}$ est également une $(\mathcal{D}_t)_{t\geq 0}$-martingale.

Pour finir cette section, nous allons présenter des formules de changement de loi sur $(D(\mathbb{R}_+, S), \mathcal{D})$ du type de celles données par Girsanov pour certaines diffusions sur des espaces euclidiens. Elles permettraient de fournir une preuve alternative à l'existence d'une solution au problème de martingales dans la proposition 13, et on reviendra succinctement sur cette possibilité à la fin de la section suivante. Toutefois, à part pour cette remarque, on ne se servira pas des résultats ci-dessous par la suite, et comme ils feront appels à des notions de calcul stochastique (élémentaire, mais que nous ne justifierons pas), nous conseillons au lecteur étudiant de passer directement à la section suivante.

Notons $\check{S} = \{(x,y) \in S \times S \,:\, x \neq y\}$, et pour $(x,y) \in \check{S}$, soit $M^{(x,y)}$ la martingale sous $\mathbb{P}_{0,m_0}$ définie par

$$\forall\, t \geq 0, \qquad M_t^{(x,y)} \quad = \quad \int_0^t \mathbb{1}_{\{x\}}(X_{s-})\, dM_s^{(\mathbb{1}_{\{y\}})}$$

Par définition, on calcule que

$$\begin{aligned} M_t^{(x,y)} &= \int_0^t \mathbb{1}_{\{x\}}(X_{s-})\, d\left(\mathbb{1}_{\{y\}}(X_s) - \mathbb{1}_{\{y\}}(X_0) - \int_0^s L_u(\mathbb{1}_{\{y\}})(X_u)\, du\right) \\ &= \int_0^t \mathbb{1}_{\{x\}}(X_{s-})\, d\mathbb{1}_{\{y\}}(X_s) - \int_0^t \mathbb{1}_{\{x\}}(X_{s-}) L_s(\mathbb{1}_{\{y\}})(X_s)\, ds \\ &= \sum_{0<s\leq t} \mathbb{1}_{\{x\}}(X_{s-})\left(\mathbb{1}_{\{y\}}(X_s) - \mathbb{1}_{\{y\}}(X_{s-})\right) - \int_0^t \mathbb{1}_{\{x\}}(X_s) L_s(\mathbb{1}_{\{y\}})(X_s)\, ds \end{aligned}$$

Notons que l'expression $\mathbb{1}_{\{x\}}(X_{s-})\left(\mathbb{1}_{\{y\}}(X_s) - \mathbb{1}_{\{y\}}(X_{s-})\right)$ est non nulle (et vaut alors 1) si et seulement si au temps $s \geq 0$, le processus de Markov X saute de x à y. Ceci nous amène à poser pour $t \geq 0$ et $(x,y) \in \check{S}$, $N_t^{(x,y)}$ le nombre de sauts de x à y dans l'intervalle $[0,t]$, car on peut alors écrire

$$\begin{aligned} \forall\, t \geq 0, \qquad M_t^{(x,y)} &= \sum_{0<s\leq t} \mathbb{1}_{\{x\}}(X_{s-})\mathbb{1}_{\{y\}}(X_s) - \int_0^t \mathbb{1}_{\{x\}}(X_s) L_s(\mathbb{1}_{\{y\}})(X_s)\, ds \\ &= N_t^{(x,y)} - \int_0^t L_s(x,y)\mathbb{1}_{\{x\}}(X_s)\, ds \end{aligned}$$

où il apparaît que $M^{(x,y)}$ est le processus du nombre de sauts de x à y compensé par un processus prévisible de manière à obtenir une martingale.

La famille de martingales $(M^{(x,y)})_{(x,y)\in\check{S}}$ est donc relativement naturelle, notamment si on lui ajoute X_0, elle permet de reconstituer X puisqu'elle donne les différents temps de sauts et les transitions effectuées. Plus précisément, un saut de X de x à y au temps $s \geq 0$ équivaut à un saut de $M^{(x,y)}$ (de hauteur 1) au temps s, et il est donc clair que pour tout $t \geq 0$,

$$\mathcal{D}_t \quad = \quad \sigma(X_0) \vee \sigma(M_s^{(x,y)} \,:\, 0 < s \leq t,\, (x,y) \in \check{S})$$

De plus toutes les martingales que l'on a rencontré jusqu'à présent peuvent s'écrire comme intégrales stochastiques par rapport à la famille $(M^{(x,y)})_{(x,y)\in\check{S}}$ (plus généralement, on peut d'ailleurs montrer que toutes les $(\mathcal{D}_t)_{t\geq 0}$-martingales $\mathbb{L}^2$ sous $\mathbb{P}_{0,m_0}$ s'obtiennent de cette manière, mais ceci n'admet pas beaucoup d'intérêt ici !).

Ainsi, si $f\in C^1(\mathbb{R}_+\times S)$, on a en intégrant par parties,

$$\begin{aligned}
&f(t,X_t)\\
&\quad= \sum_{x\in S} f(t,x)\mathbb{1}_{\{x\}}(X_t)\\
&\quad= \sum_{x\in S}\left(f(0,x)\mathbb{1}_{\{x\}}(X_0)+\int_0^t \mathbb{1}_{\{x\}}(X_{s-})\partial_s f(s,x)\,ds+\int_0^t f(s-,x)(L_s(\mathbb{1}_{\{x\}})(X_s)\,ds+dM_s^{(\mathbb{1}_{\{x\}})})\right)\\
&\quad= f(0,X_0)+\int_0^t\partial_s f(s,X_s)\,ds+\int_0^t L_s(f(s,\cdot))(X_s)\,ds+\sum_{x\in S}\int_0^t f(s,x)\,dM_s^{(\mathbb{1}_{\{x\}})}
\end{aligned}$$

ce qui permet de constater que

$$\begin{aligned}
M_t^{(f)} &= \sum_{x\in S}\int_0^t f(s,x)\,dM_s^{(\mathbb{1}_{\{x\}})}\\
&= \sum_{x\in S}\sum_{y\in S}\int_0^t f(s,x)\mathbb{1}_{\{y\}}(X_{s-})\,dM_s^{(\mathbb{1}_{\{x\}})}\\
&= \sum_{x\in S}\sum_{y\in S\setminus\{x\}}\int_0^t f(s,x)\mathbb{1}_{\{y\}}(X_{s-})\,dM_s^{(\mathbb{1}_{\{x\}})}+\sum_{x\in S}\int_0^t f(s,x)\mathbb{1}_{\{x\}}(X_{s-})\,dM_s^{(\mathbb{1}_{\{x\}})}\\
&= \sum_{(x,y)\in\check{S}}\int_0^t f(s,x)\,dM_s^{(y,x)}-\sum_{x\in S}\sum_{y\in S\setminus\{x\}}\int_0^t f(s,x)\mathbb{1}_{\{x\}}(X_{s-})\,dM_s^{(\mathbb{1}_{\{y\}})}\\
&= \sum_{(x,y)\in\check{S}}\int_0^t f(s,x)\,dM_s^{(y,x)}-\sum_{(x,y)\in\check{S}}\int_0^t f(s,x)\,dM_s^{(x,y)}\\
&= \sum_{(x,y)\in\check{S}}\int_0^t (f(s,y)-f(s,x))\,dM_s^{(x,y)}
\end{aligned}$$

où pour la quatrième égalité, on a utilisé que

$$\sum_{y\in S} M^{(\mathbb{1}_{\{y\}})} = M^{(\mathbb{1})}=0$$

Notamment, si pour $x_0\in S$ fixé, on considère $f:\mathbb{R}_+\times S\ni(s,x)\mapsto\mathbb{1}_{\{x_0\}}(x)$, on fait apparaître que

$$M^{(\mathbb{1}_{\{x_0\}})} = \sum_{x\neq x_0}M^{(x,x_0)}-\sum_{x\neq x_0}M^{(x_0,x)}$$

Par ailleurs, notons que les martingales $M^{(x,y)}$, $(x,y)\in\check{S}$, sont orthogonales, au sens où pour tous $(x,y)\neq(x',y')\in\check{S}$, $\langle M^{(x,y)},M^{(x',y')}\rangle=0$.

En effet, on calcule que pour $t\geq 0$,

$$\begin{aligned}
\langle M^{(x,y)},M^{(x',y')}\rangle_t &= \int_0^t \mathbb{1}_{\{x\}}(X_{s-})\mathbb{1}_{\{x'\}}(X_{s-})\,d\langle M^{(\mathbb{1}_{\{y\}})},M^{(\mathbb{1}_{\{y'\}})}\rangle_s\\
&= \delta_x(x')\int_0^t\mathbb{1}_{\{x\}}(X_s)\Gamma_s(\mathbb{1}_{\{y\}},\mathbb{1}_{\{y'\}})(X_s)\,ds\\
&= \delta_x(x')\delta_y(y')\int_0^t L_s(x,y)\mathbb{1}_{\{x\}}(X_s)\,ds
\end{aligned}$$

(plus rapidement, on pouvait noter que pour $(x,y) \neq (x',y') \in \check{S}$, les ensembles des temps de sauts des martingales totalement discontinues $M^{(x,y)}$ et $M^{(x',y')}$ sont disjoints, ce qui assure que $[M^{(x,y)}, M^{(x',y')}] = 0$, et par suite que $\langle M^{(x,y)}, M^{(x',y')}\rangle = 0$).

Nous allons maintenant considérer les intégrales stochastiques simples, au sens où l'on n'intègre par rapport aux $M^{(x,y)}$, $(x,y) \in \check{S}$, que des processus déterministes continus, qui généralisent les $M^{(f)}$, pour $f \in F(S)$.

Pour tout $(x,y) \in \check{S}$, donnons-nous $V(x,y) : \mathbb{R}_+ \ni s \mapsto V_s(x,y) \in [-1,+\infty[$ une application continue et définissons une martingale M en posant

$$\forall\ t \geq 0, \qquad M_t \;=\; \sum_{(x,y)\in\check{S}} \int_0^t V_s(x,y)\, dM_s^{(x,y)}$$

Rappelons que l'exponentielle stochastique (de Doléans-Dade) associée est la martingale $\mathcal{E}(M)$ définie par

$$\forall\ t \geq 0, \qquad \mathcal{E}(M)_t \;=\; \exp(M_t) \prod_{0<s\leq t} (1 + \triangle M_s) \exp(-\triangle M_s)$$

où pour tout $s > 0$, $\triangle M_s = M_s - M_{s-}$ (on aura noté que la partie martingale continue de M est nulle, car M est à variations finies).

L'intérêt de $\mathcal{E}(M)$ est qu'elle satisfait l'équation différentielle stochastique

$$d\mathcal{E}(M)_t \;=\; \mathcal{E}(M)_{t-}\, dM_t$$

Soit $T \geq 0$ fixé. Du fait que $V(x,y)$ prend ses valeurs dans $[-1,+\infty[$, on s'aperçoit que $\mathcal{E}(M)_T \geq 0$, et puisque $\mathcal{E}(M)$ est une martingale issue de 1, on a de plus $\mathbb{E}_{0,m_0}[\mathcal{E}(M)_T] = 1$, ce qui nous amène à poser $\mathbb{P}^{(T)}$ la probabilité définie sur $(D(\mathbb{R}_+,S), \mathcal{D}_T)$ qui admet $\mathcal{E}(M)_T$ pour densité par rapport à la restriction de $\mathbb{P}_{0,m_0}$ sur $\mathcal{D}_T$. De par la propriété de martingale de $\mathcal{E}(M)$, elle vérifie pour toute application $F : D(\mathbb{R}_+,S) \to \mathbb{R}_+$ qui est $\mathcal{D}_t$-mesurable, avec $0 \leq t \leq T$,

$$\mathbb{E}^{(T)}[F] = \mathbb{E}_{0,m_0}[\mathcal{E}(M)_T F] \;=\; \mathbb{E}_{0,m_0}[\mathcal{E}(M)_t F]$$

c'est-à-dire que la restriction de $\mathbb{P}^{(T)}$ à $\mathcal{D}_t$, $0 \leq t \leq T$, n'est autre que $\mathbb{P}^{(t)}$ (relations de compatibilité).

Proposition 15

> La probabilité $\mathbb{P}^{(T)}$ est solution du problème de martingales associé à m_0 et à $(R_t)_{0\leq t\leq T}$, où pour tout $0 \leq t \leq T$, R_t est le générateur défini par
>
> $$\forall\ (x,y) \in \check{S}, \qquad R_t(x,y) \;=\; L_t(x,y)(1 + V_t(x,y))$$

Preuve :

L'énoncé de la proposition signifie tout naturellement, d'une part que X_0 est de loi m_0 sous $\mathbb{P}^{(T)}$, ce qui est trivial, et d'autre part que pour tout $f \in F(S)$, le processus $(\widetilde{M}^{(f)})_{0\leq t\leq T}$ est une $(\mathcal{D}_t)_{0\leq t\leq T}$-martingale sous $\mathbb{P}^{(T)}$, où on a posé

$$\forall\ 0 \leq t \leq T, \qquad \widetilde{M}_t^{(f)} \;=\; f(X_t) - f(X_0) - \int_0^t R_s(f)(X_s)\, ds$$

Soit donc $f \in F(S)$ fixé, en notation différentielle et par définition du crochet droit, on a

$$\begin{aligned}
d(\mathcal{E}(M)_t f(X_t)) &= f(X_{t-})\,d\mathcal{E}(M)_t + \mathcal{E}(M)_{t-}(L_t(f)(X_t)\,dt + dM_t^{(f)}) + d[\mathcal{E}(M), f(X_\cdot)]_t \\
&= \mathcal{E}(M)_{t-} f(X_{t-}) dM_t + \mathcal{E}(M)_{t-} dM_t^{(f)} + d([\mathcal{E}(M), f(X_\cdot)]_t - \langle \mathcal{E}(M), f(X_\cdot)\rangle_t) \\
&\quad + \mathcal{E}(M)_t L_t(f)(X_t)\,dt + d\langle \mathcal{E}(M), f(X_\cdot)\rangle_t \\
&= dZ_t + \mathcal{E}(M)_t L_t(f)(X_t)\,dt + d\langle \mathcal{E}(M), f(X_\cdot)\rangle_t
\end{aligned}$$

où Z est une martingale sous $\mathbb{P}_{0,m_0}$. Or on calcule que

$$\begin{aligned}
d\langle \mathcal{E}(M), f(X_\cdot)\rangle_t &= d\langle M^{(f)}, \mathcal{E}(M)\rangle_t \\
&= \sum_{(x,y)\in\check{S}} \mathcal{E}(M)_{t-} V_t(x,y)\mathbb{1}_{\{x\}}(X_{t-})\,d\langle M^{(f)}, M^{(y)}\rangle_t \\
&= \sum_{(x,y)\in\check{S}} \mathcal{E}(M)_{t-} V_t(x,y)\mathbb{1}_{\{x\}}(X_{t-})\Gamma_t(f, \mathbb{1}_{\{y\}})(X_t)\,dt \\
&= \mathcal{E}(M)_t \sum_{(x,y)\in\check{S}} V_t(x,y)\mathbb{1}_{\{x\}}(X_t)\Gamma_t(f, \mathbb{1}_{\{y\}})(X_t)\,dt
\end{aligned}$$

puis on remarque que pour tout $x \in S$ fixé,

$$\begin{aligned}
&L_t(f)(x) + \sum_{y\neq x} V_t(x,y)\Gamma_t(f, \mathbb{1}_{\{y\}})(x) \\
&= \sum_{z\in S} L_t(x,z)(f(z) - f(x))(1 + \sum_{y\neq x} V_t(x,y)(\mathbb{1}_{\{y\}}(z) - \mathbb{1}_{\{y\}}(x))) \\
&= \sum_{z\in S} L_t(x,z)(f(z) - f(x))(1 + V_t(x,z)) \\
&= R_t(f)(x)
\end{aligned}$$

de sorte qu'en résumé, on a montré que

$$d(f(X_t)\mathcal{E}(M)_t) = dZ_t + R_t(f)(X_t)\,dt$$

Pour obtenir le résultat escompté, il reste à noter que pour tous $0 \le s \le t \le T$ et toute application $G : D(\mathbb{R}_+, S) \to \mathbb{R}$ bornée et $\mathcal{D}_s$-mesurable,

$$\begin{aligned}
&\mathbb{E}^{(T)}[(\widetilde{M}_t^{(f)} - \widetilde{M}_s^{(f)})G] \\
&= \mathbb{E}^{(T)}[(f(X_t) - f(X_s) - \int_s^t R_u(f)(X_u)\,du)G] \\
&= \mathbb{E}_{0,m_0}[\mathcal{E}(M)_t f(X_t)G] - \mathbb{E}_{0,m_0}[\mathcal{E}(M)_s f(X_s)G] - \int_s^t \mathbb{E}_{0,m_0}[\mathcal{E}(M)_u R_u(f)(X_u)G]\,du \\
&= \mathbb{E}_{0,m_0}[(\mathcal{E}(M)_t f(X_t) + \mathcal{E}(M)_s f(X_s) - \int_s^t \mathcal{E}(M)_u R_u(f)(X_u)\,du)G] \\
&= \mathbb{E}_{0,m_0}[(Z_t - Z_s)G] \\
&= 0
\end{aligned}$$

ce qui se traduit par

$$\mathbb{E}^{(T)}[\widetilde{M}_t^{(f)} - \widetilde{M}_s^{(f)} | \mathcal{D}_s] = 0$$

$\square$

Nous allons maintenant transformer l'expression de $\mathcal{E}(M)$ pour lui donner un aspect plus « habituel ».

Lemme 16

Pour tout $x \in S$ et tout $s \geq 0$, posons

$$\begin{aligned} H_s(x) &= -\sum_{y\neq x} L_s(x,y)V_s(x,y) = R_s(x,x) - L_s(x,x) \\ G_s(x) &= \sum_{y\neq x} L_s(x,y)(\ln(1+V_s(x,y)) - V_s(x,y)) \end{aligned}$$

On a alors pour tout $t \geq 0$

$$\mathcal{E}(M)_t = \exp\left(\sum_{(x,y)\in\check{S}} \int_0^t \ln(1+V_s(x,y))\, dN_s^{(x,y)} + \int_0^t H_s(X_s)\, ds\right)$$

(on aura remarqué que la première intégrale vaut $-\infty$ s'il existe $0 < s \leq t$ et $(x,y) \in \check{S}$ tels que $V_s(x,y) = -1$ et $\triangle M_s^{(x,y)} \neq 0$), et si de plus on suppose par exemple que $\sum_{(x,y)\in\check{S}} \int_0^t L_s(x,y) \ln(1+V_s(x,y))\, ds > -\infty$, on peut aussi écrire

$$\mathcal{E}(M)_t = \exp\left(\sum_{(x,y)\in\check{S}} \int_0^t \ln(1+V_s(x,y))\, dM_s^{(x,y)} + \int_0^t G_s(X_s)\, ds\right)$$

Preuve :

Pour tout $s > 0$, on a

$$\begin{aligned} \triangle M_s &= \sum_{(x,y)\in\check{S}} V_s(x,y)\triangle M_s^{(x,y)} \\ &= \sum_{(x,y)\in\check{S}} V_s(x,y)\triangle N_s^{(x,y)} \end{aligned}$$

de sorte que

$$\begin{aligned} (1+\triangle M_s)\exp(-\triangle M_s) &= \prod_{(x,y)\in\check{S}} (1+V_s(x,y)\triangle N_s^{(x,y)})\exp(-V_s(x,y)\triangle N_s^{(x,y)}) \\ &= \prod_{(x,y)\in\check{S}} \exp((\ln(1+V_s(x,y)) - V_s(x,y))\triangle N_s^{(x,y)}) \end{aligned}$$

On en déduit que

$$\begin{aligned} \prod_{0<s\leq t} (1+\triangle M_s)\exp(-\triangle M_s) &= \prod_{(x,y)\in\check{S}} \prod_{0<s\leq t} \exp((\ln(1+V_s(x,y)) - V_s(x,y))\triangle N_s^{(x,y)}) \\ &= \exp\left(\sum_{(x,y)\in\check{S}} \int_0^t \ln(1+V_s(x,y)) - V_s(x,y)\, dN_s^{(x,y)}\right) \end{aligned}$$

et les résultats annoncés en découlent immédiatement, via le fait que

$$\forall\, t \geq 0, \qquad M_t = \sum_{(x,y)\in\check{S}} \int_0^t V_s(x,y)\, dN_s^{(x,y)} + \int_0^t H_s(X_s)\, ds$$

$\square$

Remarque : soit $f \in C^1(\mathbb{R}_+ \times S)$ et considérons la situation où pour tout $(x,y) \in \check{S}$ et tout $s \geq 0$,

$$1 + V_s(x,y) \;=\; \exp(f(s,y) - f(s,x))$$

Il n'est pas difficile de montrer que l'on peut alors écrire pour tout $t \geq 0$,

$$\begin{aligned} M_t &= \int_0^t \exp(-f(s,X_{s-}))\, dM_s^{(e^f)} \\ \mathcal{E}(M)_t &= \exp\left(f(t,X_t) - f(0,X_0) - \int_0^t \exp(-f(s,X_s))(\partial_s + L_s)(\exp(f(s,\cdot))(X_s)\, ds\right) \end{aligned}$$

et on retrouve le fait bien connu que le processus défini par le membre de droite est une martingale strictement positive (cf. par exemple le lemme 3.2 p. 174 de [4], où l'on considère évidemment le processus de Markov homogène $((t,X_t))_{t\geq 0}$, parfois dit de temps-espace, dont le générateur $\bar{L}$ agit notamment sur les fonctions $f \in C^1(\mathbb{R}_+ \times S)$ par la formule : pour tout $(t,x) \in \mathbb{R}_+ \times S$, $\bar{L}(f)(t,x) = \partial_t f(t,x) + L_t(f(t,\cdot))(x))$. Néanmoins les densités de cette forme ne permettent pas d'effectuer tous les changements de loi précédents : ainsi par exemple si L_s, pour un $s \geq 0$, est réversible par rapport à une probabilité μ_s, la transformation ci-dessus nous donne un générateur R_s réversible par rapport à la loi ν_s définie par

$$\nu_s(x) \;=\; \frac{\exp(-2f(s,x))}{\sum_{y\in S}\exp(-2f(s,y))\mu_s(y)}\mu_s(x)$$

Soient μ la probabilité équidistribuée sur S et Λ le générateur défini par

$$\forall\ (x,y) \in \check{S}, \qquad \Lambda(x,y) \;=\; 1$$

Notons $\mathbb{P}$ la solution du problème homogène de martingales de loi initiale μ et de générateur Λ.

Soit toujours $T \geq 0$ fixé, un exemple d'application des calculs précédents est de fournir, du moins sur $\mathcal{D}_T$, i.e. sur l'intervalle de temps $[0,T]$, la solution $\mathbb{P}_{0,m_0}^{(T)}$ du problème de martingales associé à une loi initiale m_0 et à une famille continue de générateurs $(L_t)_{0\leq t\leq T}$, d'une manière relativement explicite à l'aide de $\mathbb{P}$:

$$\frac{d\mathbb{P}_{0,m_0}^{(T)}}{d\mathbb{P}_{|\mathcal{D}_T}} \;=\; \frac{dm_0}{d\mu}(X_0)\exp\left(\sum_{(x,y)\in\check{S}}\int_0^T \ln(L_s(x,y))\, dN_s^{(x,y)} + \int_0^T \mathrm{card}(S) - 1 + L_s(X_s,X_s)\, ds\right)$$

(on aura aussi noté qu'ici, on a pour tout $(x,y) \in \check{S}$ et tout $s \geq 0$, $M_s^{(x,y)} = N_s^{(x,y)} - \int_0^s \mathbf{1}_{\{x\}}(X_u)\, du$).

Par contre, cette absolue continuité ne peut pas se prolonger sur tout $\mathcal{D}$, car en général $\mathbb{P}_{0,m_0}$ et $\mathbb{P}$ diffèrent grandement sur la tribu de queue

$$\mathcal{T} \stackrel{\text{déf.}}{=} \bigcap_{t\geq 0} \sigma(X_s\,;\, t \geq s)$$

Par exemple, dans le cas homogène où pour tout $t \geq 0$, $L_t = L_0$, la tribu $\mathcal{T}$ est $\mathbb{P}_{0,m_0}$-atomique et admet exactement R $\mathbb{P}_{0,m_0}$- atome(s), où R est le nombre de classe(s) de récurrence de L_0. Ainsi $\mathcal{T}$ est $\mathbb{P}$-triviale et il n'en est pas nécessairement de même sous $\mathbb{P}_{0,m_0}$, ce qui suffit à interdire l'absolue continuité de $\mathbb{P}_{0,m_0}$ par rapport à $\mathbb{P}$.

1.5 Description des temps et positions de sauts

Nous allons ici décrire d'une manière probabiliste plus élémentaire le processus de Markov X sous la probabilité $\mathbb{P}_{0,m_0}$ solution du problème de martingales associé à une loi initiale m_0 et à une famille de générateurs $(L_t)_{t\geq 0}$, et l'analyse ainsi effectuée permettra réciproquement de donner un moyen de construire cette solution $\mathbb{P}_{0,m_0}$.

Revenons sur l'espace canonique des trajectoires càdlàgs et soit $\omega \in D(\mathbb{R}_+, S)$. La suite $(\tau_n(\omega))_{n\geq 0}$ des temps de sauts de ω est définie par récurrence en posant $\tau_0(\omega) = 0$, puis pour tout $n \in \mathbb{N}$,

$$\tau_{n+1}(\omega) = \begin{cases} \inf\{t > \tau_n(\omega) \,:\, X_t(\omega) \neq X_{t-}(\omega)\} \leq +\infty & , \text{si } \tau_n(\omega) < +\infty \\ +\infty & , \text{sinon} \end{cases}$$

Remarquons que pour tout $n \in \mathbb{N}$, $\tau_{n+1}(\omega) > \tau_n(\omega)$, sauf si $\tau_n(\omega) = +\infty$, et que l'on a

$$\lim_{n\to+\infty} \tau_n(\omega) \;=\; +\infty$$

du fait que sur tout intervalle borné de $\mathbb{R}_+$, ω n'admet qu'un nombre fini de sauts.

Pour ne pas avoir de problème de définition pour la suite $(Y_n(\omega))_{n\geq 0}$ des positions de sauts, on rajoute un élément à S qui ne lui appartenait pas, disons $\triangle$, qui est parfois appelé point cimetière, et on note

$$S_\triangle \;=\; S \sqcup \{\triangle\}$$

que l'on munit de sa tribu usuelle désignée par $\mathcal{S}_\triangle$.

On pose alors pour tout $n \geq 0$,

$$Y_n(\omega) \;=\; \begin{cases} X_{\tau_n(\omega)}(\omega) & , \text{si } \tau_n(\omega) < +\infty \\ \triangle & , \text{sinon} \end{cases}$$

puis $Z_n(\omega) = (\tau_n(\omega), Y_n(\omega))$. Il est clair que ω est entièrement déterminé par la suite $(Z_n(\omega))_{n\geq 0}$, car pour tout $t \geq 0$,

$$X_t(\omega) \;=\; Y_n, \qquad \text{si } \tau_n(\omega) \leq t < \tau_{n+1}(\omega), \text{ avec } n \in \mathbb{N}$$

Commençons par donner quelques propriétés de mesurabilité de ces objets.

Lemme 17

Pour tout $n \in \mathbb{N}$, τ_n est un temps d'arrêt et $Z_n \,:\, D(\mathbb{R}_+, S) \to (\overline{\mathbb{R}} \times S_\triangle, \mathcal{B}(\overline{\mathbb{R}}) \otimes \mathcal{S}_\triangle)$ est $\mathcal{D}_{\tau_n}$-mesurable.

Preuve :

L'affirmation étant triviale si $n = 0$, soit $n \in \mathbb{N}^*$ fixé. Pour tout $t \geq 0$, on a par continuité à droite,

$$\begin{aligned}
&\{\tau_n \leq t\} \\
&\quad = \{ \text{ il y a au moins } n \text{ sauts dans } [0,t] \ \} \\
&\quad = \{\exists\ 0 < t_1 < t_2 < \cdots < t_n \leq t,\ \exists\ x_0,\, x_1, \cdots, x_n \in S, \text{ avec } x_0 \neq x_1,\, x_1 \neq x_2, \cdots, \\
&\qquad\ x_{n-1} \neq x_n : X_0 = x_0,\, X_{t_1} = x_1, \cdots, X_{t_n} = x_n\} \\
&\quad = \{\exists\ 0 < t_1 < t_2 < \cdots < t_n \leq t,\ \exists\ x_0,\, x_1, \cdots, x_n \in S \,:\, X_0 = x_0 \text{ et} \\
&\qquad\ \forall\ 1 \leq i \leq n,\, t_i \in \mathbb{Q},\, x_{i-1} \neq x_i,\, X_{t_i} = x_i\} \\
&\quad = \bigcup_{t_1,\cdots,t_n \in \mathbb{Q}\cap]0,t]\,:\,0<t_1<t_2<\cdots<t_n} \ \bigcup_{x_0,x_1,\cdots,x_n \in S\,:\,x_0\neq x_1,\cdots,x_{n-1}\neq x_n} \{X_0 = x_0,\, X_{t_1} = x_1, \cdots, X_{t_n} = x_n\}
\end{aligned}$$

qui appartient bien à $\mathcal{D}_t$, ce qu'il fallait démontrer.

Pour la $\mathcal{D}_{\tau_n}$-mesurabilité de Z_n, on a déjà constaté que τ_n est $\mathcal{D}_{\tau_n}$-mesurable et que la restriction de Y_n à $\{\tau_n < +\infty\}$ était mesurable par rapport à la trace de $\mathcal{D}_{\tau_n}$ sur cet ensemble. Il suffit donc de se rappeler que $\{\tau_n < +\infty\} \in \mathcal{D}_{\tau_n}$ et que Y_n est constant sur le complémentaire de cet ensemble, pour se rendre compte du résultat annoncé.

□

La chaîne stochastique $(Z_n)_{n\geq 0}$ à valeurs dans $\overline{\mathbb{R}}_+ \times S_\triangle$ et définie sur $D(\mathbb{R}_+, S)$ est donc adaptée à la filtration en temps discret $(\mathcal{D}_{\tau_n})_{n\in\mathbb{N}}$. Nous allons montrer qu'elle est markovienne homogène sous $\mathbb{P}_{0,m_0}$.

Commençons par présenter quel sera son noyau markovien de transition.

Pour $t \in \mathbb{R}_+$ et $x \in S$, notons $\nu_{t,x}$ la loi sur $\overline{\mathbb{R}}_+$ définie par

$$\forall\ s \geq 0, \qquad \nu_{t,x}(]s, +\infty]) \;\; = \;\; \exp\left(\int_0^s L_{t+u}(x,x)\, du\right)$$

La restriction de $\nu_{t,x}$ à $\mathbb{R}_+$ admet donc la densité

$$\mathbb{R}_+ \ni s \;\; \mapsto \;\; -L_{t+s}(x,x) \exp\left(\int_0^s L_{t+u}(x,x)\, du\right)$$

par rapport à la mesure de Lebesgue, et $\nu_{t,x}$ dispose éventuellement en plus d'une masse $\exp\left(\int_0^{+\infty} L_{t+u}(x,x)\, du\right)$ en $+\infty$, ce qui en résumé s'écrit aussi

$$\nu_{t,x}(ds) \;\; = \;\; -L_{t+s}(x,x) \exp\left(\int_0^s L_{t+u}(x,x)\, du\right) \mathbb{1}_{\mathbb{R}_+}(s)\, ds + \exp\left(\int_0^{+\infty} L_{t+u}(x,x)\, du\right) \delta_{+\infty}(ds)$$

Pour $x \in S$, notons $A(x)$ le fermé $\{t \in \mathbb{R}_+ \;:\; L_t(x,x) = 0\}$, puis pour $t \notin A(x)$, on définit une probabilité $q_{t,x}$ sur S en posant

$$\forall\ y \in S, \qquad q_{t,x}(y) = \begin{cases} L_t(x,y)/|L_t(x,x)| & \text{, si } y \neq x \\ 0 & \text{, sinon} \end{cases}$$

On aura remarqué que $\nu_{t,x}(\mathbb{R}_+ \cap (A(x) - t)) = 0$, ce qui permet d'introduire, pour tout $t \geq 0$ et tout $x \in S$, la probabilité $Q((t,x), \cdot\,)$ sur $(\overline{\mathbb{R}}_+ \times S_\triangle, \mathcal{B}(\overline{\mathbb{R}}_+) \otimes \mathcal{S}_\triangle)$ donnée par

$\forall\ B \in \mathcal{B}(\overline{\mathbb{R}}_+) \otimes \mathcal{S}_\triangle$,

$$\begin{aligned} Q((t,x), B) \;\; &= \;\; \sum_{y\in S} \int_{\mathbb{R}_+} \nu_{t,x}(ds) \mathbb{1}_{A(x)}(t+s) q_{t+s,x}(y) \mathbb{1}_B(t+s, y) + \nu_{t,x}(+\infty) \mathbb{1}_B(+\infty, \triangle) \\ &= \;\; \sum_{y\in S} \int_{\mathbb{R}_+} \mathbb{1}_B(t+s, y) L_{t+s}(x,y) \exp\left(\int_0^s L_{t+u}(x,x)\, du\right) ds \\ &\quad\; + \exp\left(\int_0^{+\infty} L_{t+u}(x,x)\, du\right) \delta_{(+\infty,\triangle)}(B) \end{aligned}$$

On étend cette définition pour tout $(t,x) \in \overline{\mathbb{R}}_+ \times S_\triangle$, en posant de plus pour tout $t \in \overline{\mathbb{R}}_+$ et tout $x \in S_\triangle$,

$$Q((+\infty, x), \cdot\,) = \delta_{(+\infty,\triangle)}(\,\cdot\,) = Q((t, \triangle), \cdot\,)$$

Notons $(E, \mathcal{E})$ l'espace mesurable $(\overline{\mathbb{R}}_+ \times S_\triangle, \mathcal{B}(\overline{\mathbb{R}}_+) \otimes \mathcal{S}_\triangle)$, on vérifie que Q est un noyau de probabilités de transitions sur $(E, \mathcal{E})$, au sens où pour tout $x \in E$, $Q(x, \cdot\,)$ est une probabilité sur $(E, \mathcal{E})$ et où pour tout $A \in \mathcal{E}$, l'application $E \ni x \mapsto Q(x, A)$ est $\mathcal{E}$-mesurable.

Proposition 18

Sous la solution $\mathbb{P}_{0,m_0}$ au problème de martingales associé à la loi m_0 et à la famille de générateurs $(L_t)_{t\geq 0}$, $(Z_n)_{n\in\mathbb{N}}$ est une chaîne de Markov sur $(E,\mathcal{E})$ dont le noyau de probabilités de transition est Q et dont la loi initiale est $\delta_0\otimes m_0$.

Preuve :

Le fait que $(Z_n)_{n\in\mathbb{N}}$ soit de loi initiale $\delta_0\otimes m_0$ est trivial, et d'ailleurs quitte à conditionner par la tribu initiale $\mathcal{D}_0$, on se ramène immédiatement au cas où il existe $x_0\in S$ tel que $m_0=\delta_{x_0}$.

La première étape de la démonstration consiste à calculer la loi de Z_1 sous $\mathbb{P}_{0,x_0}$:

Lemme 19

Sous $\mathbb{P}_{0,x_0}$, la loi du premier couple de saut (τ_1,Y_1) est $Q((0,x_0),\,\cdot\,)$.

Preuve du lemme :

Soit $y\in S\setminus\{x_0\}$, par définition, on a pour tout $T\geq 0$,

$$\mathbb{1}_{\{y\}}(X_T) \quad=\quad \mathbb{1}_{\{y\}}(X_0)+\int_0^T L_s(\mathbb{1}_{\{y\}})(X_s)\,ds+M_T^{(\mathbb{1}_{\{y\}})} \tag{15}$$

et pour $t\geq 0$ donné, considérons plutôt cette égalité avec $T=t\wedge\tau_1$.

Par le théorème d'arrêt, le processus $(M_{s\wedge\tau_1}^{(\mathbb{1}_{\{y\}})})_{s\geq 0}$ est une martingale, que l'on sait de plus issue de 0, il apparaît donc que

$$\mathbb{E}_{0,x_0}[M_{t\wedge\tau_1}^{(\mathbb{1}_{\{y\}})}] \quad=\quad 0$$

et par ailleurs,

$$\mathbb{E}_{0,x_0}[\mathbb{1}_{\{y\}}(X_0)] \quad=\quad \delta_{x_0}(y)=0$$

ainsi en prenant l'espérance dans (15), on obtient

$$\begin{aligned}
\mathbb{P}_{0,x_0}[Y_1=y,\tau_1\leq t] &= \mathbb{E}_{0,x_0}[\mathbb{1}_{\{y\}}(X_{t\wedge\tau_1})]\\
&= \mathbb{E}_{0,x_0}\left[\int_0^{t\wedge\tau_1} L_s(\mathbb{1}_{\{y\}})(X_s)\,ds\right]\\
&= \mathbb{E}_{0,x_0}\left[\int_0^{t\wedge\tau_1} L_s(\mathbb{1}_{\{y\}})(x_0)\,ds\right]\\
&= \mathbb{E}_{0,x_0}[G_y(t\wedge\tau_1)]
\end{aligned}$$

où on a posé pour tout $s\geq 0$,

$$G_y(s) \quad=\quad \int_0^s L_u(x_0,y)\,du$$

car on aura remarqué que pour tout $0\leq s<t\wedge\tau_1$, $X_s=x_0$.

Il est commode d'introduire aussi pour tout $s\geq 0$,

$$G(s) \quad=\quad \sum_{y\neq x_0} G_y(s)=-\int_0^s L_u(x_0,x_0)\,du$$

car en sommant sur $y\in S\setminus\{x_0\}$ les égalités

$$\mathbb{P}_{0,x_0}[Y_1=y,\tau_1\leq t] \quad=\quad \mathbb{E}_{0,x_0}[G_y(t\wedge\tau_1)] \tag{16}$$

on obtient une équation fermée satisfaite par la loi de τ_1 :

$$\begin{aligned} \mathbb{P}_{0,x_0}[\tau_1 \le t] &= \mathbb{E}_{0,x_0}[G(t\wedge\tau_1)] \\ &= G(t)\mathbb{P}_{0,x_0}[\tau_1 > t] + \mathbb{E}_{0,x_0}[G(\tau_1)\mathbb{1}_{\{\tau_1\le t\}}] \\ &= G(t)h(t) - \int_{[0,t]} G(s)\,dh(s) \end{aligned}$$

où $h : \mathbb{R}_+ \to [0,1]$ est l'application définie par

$$\forall\ s \ge 0, \qquad h(s) \;=\; \mathbb{P}_{0,x_0}[\tau_1 > s]$$

et où la dernière intégrale précédente est comprise au sens de Stieljes, h étant clairement à variations localement finies, car décroissante.

On a donc aboutit à

$$\forall\ t \ge 0, \qquad 1 - h(t) = G(t)h(t) - \int_{[0,t]} G(s)dh(s)$$

cependant en notant que G est également à variations localement finies et est continu, on a par la formule d'intégration par parties pour les intégrales de Stieljes,

$$\begin{aligned} G(t)h(t) &= G(0)h(0) + \int_{[0,t]} G(s-)\,dh(s) + \int_{[0,t]} h(s-)\,dG(s) + [G,h]_t \\ &= \int_{[0,t]} G(s)\,dh(s) + \int_{[0,t]} h(s-)\,dG(s) \end{aligned}$$

ce qui nous permet de réaliser que

$$-\int_{[0,t]} dh(s) = 1 - h(t) \;=\; \int_{[0,t]} h(s-)\,dG(s)$$

ce qui s'écrit aussi $h(0) = 1$ et $dh(t) = -h(t-)dG(t)$, équations qui se résolvent immédiatement en

$$\forall\ t \ge 0, \quad h(t) \;=\; \exp(-G(t))$$

c'est-à-dire que la loi de τ_1 est ν_{0,x_0}.

Pour trouver la loi du couple (τ_1, Y_1), revenons à l'équation (16), dont on intègre par parties le membre de droite comme ci-dessus, G_y remplaçant G, pour obtenir,

$$\begin{aligned} \mathbb{P}_{0,x_0}[Y_1 = y, \tau_1 \le t] &= \int_{[0,t]} h(s-)\,dG_y(s) \\ &= \int_{[0,t]} \exp\left(\int_0^s L_u(x_0,x_0)\,du\right) L_s(x_0,y)\,ds \end{aligned}$$

ce qui est le résultat annoncé, car rappelons que si $\tau_1 = +\infty$, alors $Y_1 = \triangle$.

$\square$

Considérons à nouveau une probabilité initiale quelconque m_0, le lemme ci-dessus peut s'interpréter de la manière suivante : pour toute application mesurable bornée $F : E \to \mathbb{R}$, $\mathbb{P}_{0,m_0}$-p.s.,

$$\mathbb{E}_{0,m_0}[F(Z_1)|Z_0] \;=\; Q(Z_0, F)$$

et la proposition sera prouvée si l'on montre plus généralement que pour tout $n \in \mathbb{N}$,

$$\mathbb{E}_{0,m_0}[F(Z_{n+1})|\mathcal{D}_{\tau_n}] \;=\; Q(Z_n, F)$$

car la chaîne $(Z_n)_{n\in\mathbb{N}}$ est adaptée à la filtration $(\mathcal{D}_{\tau_n})_{n\in\mathbb{N}}$.

Remarquons déjà que sur $\{\tau_n = +\infty\}$, $Z_{n+1} = (+\infty, \triangle)$, ce qui permet de voir qu'il suffit de prouver

$$\mathbf{1}_{\{\tau_n<+\infty\}}\mathbb{E}_{0,m_0}[F(Z_{n+1})|\mathcal{D}_{\tau_n}] \quad = \quad \mathbf{1}_{\{\tau_n<+\infty\}}Q(Z_n, F) \tag{17}$$

i.e.

$$\mathbb{E}_{0,m_0}[\mathbf{1}_{\{\tau_n<+\infty\}}F(Z_{n+1})|\mathcal{D}_{\tau_n}] \quad = \quad \mathbf{1}_{\{\tau_n<+\infty\}}Q(Z_n, F)$$

Cependant sur $\{\tau_n < +\infty\}$, on vérifie que

$$\begin{aligned} \tau_{n+1} &= \tau_n + \tau_1 \circ \theta_{\tau_n} \\ Y_{n+1} &= Y_1 \circ \theta_{\tau_n} \end{aligned}$$

ainsi la propriété de Markov forte implique que $\mathbb{P}_{0,m_0}$-p.s. en $\omega \in D(\mathbb{R}_+, S)$,

$$\mathbf{1}_{\{\tau_n(\omega)<+\infty\}}\mathbb{E}_{0,m_0}[F(\tau_n + \tau_1 \circ \theta_{\tau_n}, Y_1 \circ \theta_{\tau_n})|\mathcal{D}_{\tau_n}](\omega) \quad = \quad \mathbf{1}_{\{\tau_n(\omega)<+\infty\}}\mathbb{E}_{\tau_n(\omega),Y_n(\omega)}[F(\tau_n(\omega) + \tau_1, Y_1)]$$

Or la preuve du lemme 19 montre, à un décallage près dans le temps de la famille de générateurs, que pour tout $t \geq 0$ et tout $x \in S$,

$$\mathbb{E}_{t,x}[F(t + \tau_1, Y_1)] \quad = \quad Q((t, x), F)$$

d'où finalement (17).

□

Ces résultats suggèrent une méthode de construction de la solution $\mathbb{P}_{0,x_0}$, pour $x_0 \in S$, du problème de martingales associé à δ_{x_0} et à $(L_t)_{t\geq 0}$.

En effet, remarquons que le noyau markovien Q défini sur $(E, \mathcal{E})$ ci-dessus ne dépend que de la famille de générateurs $(L_t)_{t\geq 0}$, or comme on l'a déjà mentionné précédemment, pour tous $t \geq 0$ et $x \in S$, il est toujours possible de construire sur $(E^{\mathbb{N}}, \mathcal{E}^{\otimes\mathbb{N}})$ une probabilité $\widetilde{\mathbb{P}}_{t,x}$ telle que la chaîne canonique des coordonnées sur cet espace produit, $\widetilde{Z} = (\widetilde{Z}_n)_{n\in\mathbb{N}} = ((\widetilde{\tau}_n, \widetilde{Y}_n))_{n\in\mathbb{N}}$, soit markovienne homogène de noyau de probabilités de transition Q et de loi initiale $\delta_{(t,x)}$.

Lemme 20

Soient $t \geq 0$ et $x \in S$ fixés. Sous la probabilité $\widetilde{\mathbb{P}}_{t,x}$, on a p.s. que pour $n \in \mathbb{N}$, si $\widetilde{\tau}_n < +\infty$, alors $\widetilde{Y}_n \in S$ et $\widetilde{\tau}_{n+1} > \widetilde{\tau}_n$. De plus, toujours $\widetilde{\mathbb{P}}_{t,x}$-p.s., on est assuré de

$$\lim_{n\to+\infty} \widetilde{\tau}_n \quad = \quad +\infty$$

Preuve :

Commençons par remarquer que pour tous $s \geq 0$ et $y \in S$, $Q((s, y),]0, +\infty[\times\{\triangle\}) = 1$, ce qui permet d'obtenir par une récurrence sur n, en utilisant qu'initialement $\widetilde{Z}_0 = (t, x)$, $\widetilde{\mathbb{P}}_{t,x}$-p.s., que pour $n \in \mathbb{N}^*$,

$$\begin{aligned} \widetilde{\mathbb{P}}_{t,x}[\widetilde{Y}_n \notin S, \widetilde{\tau}_n < +\infty] &= \widetilde{\mathbb{P}}_{t,x}[\widetilde{Y}_n = \{\triangle\}, \widetilde{\tau}_n < +\infty, \widetilde{\tau}_{n-1} < +\infty, \widetilde{Y}_{n-1} \in S] \\ &= \widetilde{\mathbb{E}}_{t,x}[\mathbf{1}_{\{\widetilde{\tau}_{n-1}<+\infty,\, \widetilde{Y}_{n-1}\in S\}}Q((\widetilde{\tau}_{n-1}, \widetilde{Y}_{n-1}),]0, +\infty[\times\{\triangle\})] = 0 \end{aligned}$$

De même, puisque pour tous $s \geq 0$ et $y \in S$, $\nu_{s,y}(]0, +\infty]) = 1$, on réalise que

$$\widetilde{\mathbb{P}}_{t,x}[\widetilde{\tau}_{n+1} \leq \widetilde{\tau}_n, \widetilde{\tau}_n < +\infty] \quad = \quad \widetilde{\mathbb{E}}_{t,x}[\mathbf{1}_{\{\widetilde{\tau}_n<+\infty,\, \widetilde{Y}_n\in S\}}(1 - \nu_{\widetilde{\tau}_n,\widetilde{Y}_n}(]0, +\infty]))] = 0$$

et on en déduit la première affirmation, qui elle même implique déjà que $\widetilde{I\!P}_{t,x}$-p.s., $\lim_{n\to\infty}\widetilde{\tau}_n$ existe à valeurs dans $]0,+\infty]$.

Soit $M>0$, et notons $A=\max_{0\leq s\leq 2M,\, x\in S}|L_s(x,x)|<+\infty$, de sorte que pour tout $n\in I\!N$, on ait $\widetilde{I\!P}_{t,x}$-p.s.

$$\begin{aligned}\widetilde{I\!P}_{t,x}[\widetilde{\tau}_{n+1}-\widetilde{\tau}_n>M|\widetilde{Z}_n,\widetilde{Z}_{n-1},\cdots,\widetilde{Z}_0]\mathbb{1}_{\{\widetilde{\tau}_n\leq M\}} &= \mathbb{1}_{\{\widetilde{\tau}_n\leq M,\,\widetilde{Y}_n\in S\}}\nu_{\widetilde{\tau}_n,\widetilde{Y}_n}(]M,+\infty])\\ &= \mathbb{1}_{\{\widetilde{\tau}_n\leq M,\,\widetilde{Y}_n\in S\}}\exp\left(\int_0^M L_{\widetilde{\tau}_n+u}(\widetilde{Y}_n,\widetilde{Y}_n)\,du\right)\\ &\geq \mathbb{1}_{\{\widetilde{\tau}_n\leq M\}}\exp(-MA)\end{aligned}$$

ce qui par récurrence sur $n\in I\!N$, permet de constater que

$$\begin{aligned}\widetilde{I\!P}_{t,x}[\widetilde{\tau}_n\leq M] &\leq \widetilde{I\!P}_{t,x}[\widetilde{\tau}_n-\widetilde{\tau}_{n-1}\leq M,\widetilde{\tau}_{n-1}\leq M]\\ &\leq (1-\exp(-MA))\widetilde{I\!P}_{t,x}[\widetilde{\tau}_{n-1}\leq M]\\ &\leq (1-\exp(-MA))^n\end{aligned}$$

puis que

$$\widetilde{I\!P}_{t,x}[\lim_{n\to\infty}\widetilde{\tau}_n\leq M] = \lim_{n\to\infty}\widetilde{I\!P}_{t,x}[\widetilde{\tau}_n\leq M]=0$$

et il reste alors à faire tendre M vers $+\infty$ pour se convaincre que

$$\widetilde{I\!P}_{t,x}[\lim_{n\to\infty}\widetilde{\tau}_n<+\infty] = 0$$

□

Exercice : montrer de même que p.s. sous $\widetilde{I\!P}_{t,x}$, pour $n\in I\!N$, si $\widetilde{\tau}_n<+\infty$, alors $\widetilde{Y}_{n+1}\neq\widetilde{Y}_n$.

Pour $n\in I\!N$, notons

$$A_n = \{\omega\in E^{I\!N}\,:\,\tau_n(\omega)<+\infty,\,\widetilde{Y}_n(\omega)\in S,\,\widetilde{\tau}_{n+1}(\omega)>\widetilde{\tau}_n(\omega),\,\widetilde{Y}_{n+1}(\omega)\neq\widetilde{Y}_n(\omega)\}$$

puis posons

$$\begin{aligned}\widetilde{\Omega} &= \bigcap_{n\in I\!N}(A_n\sqcup\{\widetilde{\tau}_n=+\infty\})\\ \Omega &= \{\omega\in\widetilde{\Omega}\,:\,\lim_{n\to+\infty}\widetilde{\tau}_n(\omega)=+\infty\}\end{aligned}$$

On vérifie que l'ensemble Ω appartient à $\mathcal{E}^{\otimes I\!N}$ et on le munit de la tribu $\mathcal{F}$ trace de $\mathcal{E}^{\otimes I\!N}$.

Par abus de notation, $\widetilde{I\!P}_{t,x}$ désignera encore la restriction de $\widetilde{I\!P}_{t,x}$ à $\mathcal{F}$, qui est une probabilité puisque l'on vient de vérifier que $\widetilde{I\!P}_{t,x}(\Omega)=1$. Sur l'espace probabilisé $(\Omega,\mathcal{F},\widetilde{I\!P}_{t,x})$ on peut bien définir un processus stochastique $\widetilde{X}^{(t)}=(\widetilde{X}^{(t)}_s)_{s\geq 0}$ à valeurs dans S, en convenant que pour tout $s\geq 0$,

$$\forall\,\omega\in\Omega,\qquad \widetilde{X}^{(t)}_s = \widetilde{Y}_n,\qquad \text{si } \widetilde{\tau}_n-t\leq s<\widetilde{\tau}_{n+1}-t$$

Remarquons que $\widetilde{X}^{(t)}$ peut être considéré comme une variable aléatoire à valeurs dans $(D(I\!R_+,S),\mathcal{D})$, ce qui nous incite à noter $I\!P_{t,x}$ l'image de $\widetilde{I\!P}_{t,x}$ par $\widetilde{X}^{(t)}$.

Proposition 21

La famille $(I\!P_{t,x})_{t\geq 0,x\in S}$ introduite ci-dessus sur $(D(I\!R_+,S),\mathcal{D})$ est un processus de Markov à trajectoires régulières au sens de la définition 6.

Preuve :

La première condition de la définition 6 est clairement satisfaite et il suffit donc de vérifier que pour tous $s, t \geq 0$, tout $x \in S$ et toute application mesurable bornée $F : D(\mathbb{R}_+, S) \to \mathbb{R}$,

$$\mathbb{E}_{t,x}[F \circ \theta_s | \mathcal{D}_s] = \mathbb{E}_{t+s,X_s}[F] \qquad \mathbb{P}_{t,x}\text{-p.s.}$$

Cependant sur l'espace mesurable $(D(\mathbb{R}_+, S), \mathcal{D})$ reconsidérons la chaîne $(Z_n)_{n\geq 0} = (\tau_n, Y_n)_{n\geq 0}$ introduite au début de cette section, puis pour $s \geq 0$, posons

$$N_s \;=\; \inf\{n \in \mathbb{N} \,:\, \tau_n > s\} - 1$$

qui est le nombre de sauts intervenus avant le temps s (inclus).

Il apparaît sans trop de difficulté que

$$\mathcal{D}_s \;=\; \sigma(N_s,\, (Z_{N_s \wedge i})_{i\in\mathbb{N}})$$

Revenons sur l'espace probabilisé $(\Omega, \mathcal{F}, \widetilde{\mathbb{P}}_{t,x})$, et remarquons que par définition de Ω, on a pour tout $n \in \mathbb{N}$, $\widetilde{\tau}_n = t + \tau_n \circ \widetilde{X}^{(t)}$, $\widetilde{Y}_n = Y_n \circ \widetilde{X}^{(t)}$.

Ceci nous amène à introduire

$$\widetilde{N}_{t+s} \;=\; N_s \circ \widetilde{X}^{(t)} = \inf\{n \in \mathbb{N} \,:\, \widetilde{\tau}_n > t+s\} - 1 < +\infty$$

puis pour tout $n \geq 1$,

$$\widehat{Z}_n = (\widehat{\tau}_n, \widehat{Y}_n) \;\stackrel{\text{déf.}}{=}\; (\widetilde{\tau}_{\widetilde{N}_{t+s}+n}, \widetilde{Y}_{\widetilde{N}_{t+s}+n})$$

car cette petite discussion permet de se persuader qu'il suffit de vérifier que pour toute application $G : (E^{\mathbb{N}}, \mathcal{E}^{\otimes\mathbb{N}}) \to \mathbb{R}$ mesurable bornée,

$$\widetilde{\mathbb{E}}_{t,x}[G((\widehat{Z}_n)_{n\geq 1}) | \widetilde{N}_{t+s}, (\widetilde{Z}_{\widetilde{N}_{t+s}\wedge i})_{i\in\mathbb{N}}] \;=\; \widetilde{\mathbb{E}}_{t+s,\widetilde{Y}_{\widetilde{N}_{t+s}}}[G((\widetilde{Z}_n)_{n\geq 1})] \tag{18}$$

Pour $n \in \mathbb{N}$, soit $\mathcal{E}_n$ la tribu engendrée sur Ω par les coordonnées $\widetilde{Z}_j$, pour $0 \leq j \leq n$. La variable aléatoire $\widetilde{N}_{t+s} + 1$ à valeurs dans $\mathbb{N}$ est alors un temps d'arrêt relativement à la filtration discrète $(\mathcal{E}_n)_{n\in\mathbb{N}}$, et on lui associe une tribu $\mathcal{E}_{\widetilde{N}_{t+s}+1}$. Par la propriété de Markov forte appliquée à $\widetilde{Z}$ et au temps d'arrêt $\widetilde{N}_{t+s} + 1$, on montre facilement que $(\widehat{Z}_n)_{n\geq 1}$ est conditionnellement à $\mathcal{E}_{\widetilde{N}_{t+s}+1}$ sous $\widetilde{\mathbb{P}}_{t,x}$, une chaîne de Markov homogène de noyau de probabilités de transition Q. Il en est évidemment de même pour $(\widetilde{Z}_n)_{n\geq 1}$ sous toute loi $\widetilde{\mathbb{P}}_{t+s,y}$, pour $y \in S$, ainsi les égalités (18) seront vérifiées dès que l'on aura montré que pour toute application $G : (E, \mathcal{E}) \to \mathbb{R}$ mesurable bornée,

$$\widetilde{\mathbb{E}}_{t,x}[G(\widehat{Z}_1) | \widetilde{N}_{t+s}, (\widetilde{Z}_{\widetilde{N}_{t+s}\wedge i})_{i\in\mathbb{N}}] \;=\; \widetilde{\mathbb{E}}_{t+s,\widetilde{Y}_{\widetilde{N}_{t+s}}}[G(\widetilde{Z}_1)]$$

On reconnait que le membre de droite n'est autre que $Q((t+s, \widetilde{Y}_{\widetilde{N}_{t+s}}), G)$, et par un théorème de classe monotone, on se ramène à prouver ces relations pour G de la forme

$$\forall\ (v, z) \in E, \qquad G((v, z)) \;=\; \mathbb{1}_{]t+s+u,+\infty]}(v)\mathbb{1}_{\{y\}}(z)$$

où $u \geq 0$ et $y \in S$ sont fixés.

Cependant, il est clair que

$$\begin{aligned}
&\widetilde{I\!P}_{t,x}[\widehat{\tau}_1 > t+s+u,\, \widehat{Y}_1 = y | \widetilde{N}_{t+s}, (\widetilde{Z}_{\widetilde{N}_{t+s}\wedge i})_{i\in I\!N}] \\
&\quad = \sum_{n\in I\!N} \mathbf{1}_{\{\widetilde{N}_{t+s}=n\}} \widetilde{I\!P}_{t,x}[\widetilde{\tau}_{n+1} > t+s+u,\, \widetilde{Y}_{n+1} = y | \widetilde{N}_{t+s} = n, (\widetilde{Z}_i)_{0\le i\le n}]
\end{aligned}$$

et on est donc amené à estimer les termes de la somme de droite. Pour ceci, soit $n \in I\!N$ fixé, et $H : E^{n+1} \to I\!R$ une application mesurable bornée, on calcule que

$$\begin{aligned}
&\widetilde{I\!E}_{t,x}[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{N}_{t+s}=n}\mathbf{1}_{\{\widetilde{\tau}_{n+1}>t+s+u,\, \widetilde{Y}_{n+1}=y\}}] \\
&\quad = \widetilde{I\!E}_{t,x}[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{\tau}_n\le t+s}\mathbf{1}_{\{\widetilde{\tau}_{n+1}>t+s+u,\, \widetilde{Y}_{n+1}=y\}}] \\
&\quad = \widetilde{I\!E}_{t,x}[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{\tau}_n\le t+s}Q(\widetilde{Z}_n,]t+s+u, +\infty] \times \{y\})] \\
&\quad = \widetilde{I\!E}_{t,x}\left[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{\tau}_n\le t+s}Q(\widetilde{Z}_n,]t+s+u, +\infty] \times \{y\})\frac{\mathbf{1}_{\widetilde{\tau}_{n+1}>t+s}}{Q(\widetilde{Z}_n,]t+s, +\infty] \times S_\triangle)}\right] \\
&\quad = \widetilde{I\!E}_{t,x}[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{\tau}_n\le t+s}\mathbf{1}_{\widetilde{\tau}_{n+1}>t+s}Q((t+s, \widetilde{Y}_n),]t+s+u, +\infty] \times \{y\})]
\end{aligned}$$

car on aura remarqué que pour tous $0 \le v \le t+s$ et $y \ne z \in S$,

$$\begin{aligned}
&\frac{Q((v,z),]t+s+u, +\infty] \times \{y\})}{Q((v,z),]t+s, +\infty] \times S_\triangle)} \\
&\quad = \frac{Q((v,z),]t+s+u, +\infty[\times\{y\})}{\nu_{v,z}(]t+s-v, +\infty])} \\
&\quad = \frac{\int_{I\!R_+} \mathbf{1}_{]t+s+u,+\infty[}(v+w)L_{v+w}(z,y)\exp(\int_0^w L_{v+w'}(z,z)\,dw')\,dw}{\exp(\int_0^{t+s-v} L_{v+w'}(z,z)\,dw')} \\
&\quad = \int_{I\!R_+} \mathbf{1}_{]t+s+u,+\infty[}(v+w)L_{v+w}(z,y)\exp\left(\int_{t+s-v}^{w} L_{v+w'}(z,z)\,dw'\right)\,dw \\
&\quad = \int_{I\!R_+} \mathbf{1}_{]t+s+u,+\infty[}(t+s+\widetilde{w})L_{t+s+\widetilde{w}}(z,y)\exp\left(\int_0^{\widetilde{w}} L_{t+s+w'}(z,z)\,dw'\right)\,d\widetilde{w} \\
&\quad = Q((t+s,z),]t+s+u, +\infty] \times \{y\})
\end{aligned}$$

(on aura noté, avec les notations ci-dessus, que $w \ge t+s-v$). Par contre si $y = z$, tous les membres ci-dessus sont en fait nuls, et donc égaux.

En résumé, on a ainsi obtenu

$$\begin{aligned}
&\widetilde{I\!E}_{t,x}[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{N}_{t+s}=n}\mathbf{1}_{\{\widetilde{\tau}_{n+1}>t+s+u,\, \widetilde{Y}_{n+1}=y\}}] \\
&\quad = \widetilde{I\!E}_{t,x}[H(\widetilde{Z}_0, \cdots, \widetilde{Z}_n)\mathbf{1}_{\widetilde{N}_{t+s}=n}Q((t+s, \widetilde{Y}_n),]t+s+u, +\infty] \times \{y\})]
\end{aligned}$$

ce qui signifie que

$$\begin{aligned}
&\mathbf{1}_{\{\widetilde{N}_{t+s}=n\}}\widetilde{I\!P}_{t,x}[\widetilde{\tau}_{n+1} > t+s+u,\, \widetilde{Y}_{n+1} = y | \widetilde{N}_{t+s} = n, (\widetilde{Z}_i)_{0\le i\le n}] \\
&\qquad\qquad = \mathbf{1}_{\{\widetilde{N}_{t+s}=n\}}Q((t+s, \widetilde{Y}_n),]t+s+u, +\infty] \times \{y\})
\end{aligned}$$

et il reste donc à sommer ces égalités pour $n \in I\!N$ pour obtenir le résultat escompté.

$\square$

Pour se persuader que la probabilité $I\!P_{0,x_0}$ ainsi construite, pour $x_0 \in S$, est bien la solution au problème de martingales associé à δ_{x_0} et à $(L_t)_{t\ge0}$, il nous reste donc, par le biais de la proposition 12, à montrer la

Proposition 22

> Le processus de Markov à trajectoires régulières $(\mathbb{P}_{t,x})_{t\geq 0,x\in S}$ introduit avant le lemme 20 est de classe C^1 et $(L_t)_{t\geq 0}$ est sa famille continue de générateurs.

Preuve :

Soit $T>0$ fixé, on pose $A=\max_{0\leq s\leq T+1,\, y\in S}|L_s(y,y)|$. Pour tous $0\leq t\leq T$, $0\leq h\leq 1$ et $y\in S$,

$$\begin{aligned}\mathbb{P}_{t,y}[\tau_1\leq h] &= 1-\exp\left(\int_0^h L_{t+u}(y,y)\,du\right)\\ &\leq 1-\exp(-Ah)\leq Ah\end{aligned}$$

ce qui montre que pour $x\neq y\in S$, quand h tend vers 0_+,

$$\begin{aligned}|\mathbb{P}_{t,x}[X_h=y]-\mathbb{P}_{t,x}[\tau_1\leq h,Y_1=y]| &\leq \mathbb{P}_{t,x}[\tau_2\leq h]\\ &\leq \mathbb{E}_{t,x}[\mathbb{1}_{\tau_1\leq h}\mathbb{P}_{\tau_1,Y_1}[\tau_1\leq h]]\\ &\leq A^2h^2=\circ(h)\end{aligned}$$

Cependant, pour $h>0$,

$$\frac{1}{h}\mathbb{P}_{t,x}[\tau_1\leq h,Y_1=y] \;=\; \frac{1}{h}\int_0^h L_{t+u}(x,y)\exp\left(\int_0^u L_{t+v}(x,x)\,du\right)\,du$$

expression qui tend vers $L_t(x,y)$ pour h petit, ainsi on obtient bien en regroupant ces considérations que

$$L_t \;=\; \lim_{h\to 0_+}\frac{P_{t,t+h}-\mathrm{Id}}{h}$$

où on désigne comme d'habitude par $(P_{s,t})_{0\leq s\leq t}$ le semi-groupe inhomogène associé à la famille $(\mathbb{P}_{t,x})_{t\geq 0,x\in S}$.

De la même manière il apparaît que pour $0<t\leq T$,

$$L_t \;=\; \lim_{h\to 0_+}\frac{P_{t-h,t}-\mathrm{Id}}{h}$$

d'où en fin de compte l'affirmation de la proposition.

□

Les résultats de cette section permettent donc de simuler le processus X sous la loi $\mathbb{P}_{0,m_0}$, solution du problème de martingales associé à m_0 et à $(L_t)_{t\geq 0}$.

En effet, il suffit de savoir échantillonner les variables aléatoires $(\widetilde{Z}_n)_{n\geq 0}$ sous $\widetilde{\mathbb{P}}_{0,m_0}$, et pour ceci on procède de la manière suivante :

On commence par choisir suivant la loi m_0 sur S une position Y_0, puis on « tire » suivant la probabilité ν_{0,Y_0} une valeur τ_1 sur $\overline{\mathbb{R}}_+$.

• Si $\tau_1=+\infty$, on pose $Y_1=\triangle$, puis pour tout $i\geq 2$, $Z_i=(+\infty,\triangle)$, et on arrête la procédure.

• Si $\tau_1<+\infty$, presque sûrement $L_{\tau_1}(Y_0,Y_0)>0$, et on tire suivant la loi q_{τ_1,Y_0} une valeur Y_1 sur $S\setminus\{Y_0\}$.

On passe ensuite à la seconde étape : on tire une valeur T_2 sur $\overline{\mathbb{R}}_+$ suivant la probabilité ν_{τ_1,Y_1}, puis on pose $\tau_2=\tau_1+T_2$. On distingue alors à nouveau deux cas :

• Si $\tau_2=+\infty$, on pose $Y_2=\triangle$, puis pour tout $i\geq 3$, $Z_i=(+\infty,\triangle)$, et on arrête la procédure.

• Si $\tau_1 < +\infty$, presque sûrement $L_{\tau_2}(Y_1, Y_1) > 0$, et on tire suivant la loi q_{τ_2,Y_1} une valeur Y_2 sur $S \setminus \{Y_1\}$.

Vient après une troisième étape, et ainsi de suite ...

Remarques :

a) Il apparaît dans les preuves ci-dessus que la condition de continuité de la famille de générateurs est trop forte ; on peut obtenir des résultats similaires en supposant seulement les applications $\mathbb{R}_+ \ni s \mapsto L_s$ mesurables et localement dans l'espace $\mathbb{L}^1$ de la mesure de Lebesgue, ce qui signifie simplement que pour tout $t \geq 0$ et tout $x \in S$,

$$\int_0^t L_s(x,x)\, ds \quad > \quad -\infty$$

Dans la section 3, il faudrait alors considérer des semi-groupes absolument continus, séparément en chacune des deux variables temporelles, les différentiations devant être comprises p.s. au sens de Radon-Nikodym.

Pour ce genre de considérations, on renvoit à la revue effectuée dans la dernière section du livre d'Iosifescu [6].

b) La construction précédente est encore possible dans certaines situations particulières.

Ainsi par exemple, considérons $S = \{0, 1\}$ et la famille de générateurs $(L_t)_{t>0}$ donnée par

$$\forall\, t > 0, \qquad L_t \quad = \quad \begin{pmatrix} -1 & 1 \\ 1/t & -1/t \end{pmatrix}$$

On peut alors comme ci-dessus construire sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ un processus de Markov $\widetilde{X}$ correspondant en quelque sorte à la loi initiale δ_0 et à la famille $(L_t)_{t>0}$.

Mais on laisse le soin à l'éventuel lecteur de vérifier que l'on ne peut pas lui associer un processus de Markov à trajectoires régulières au sens de la définition 6, ce qui fournit un contre-exemple pour le quatrième exercice de la section 3.

c) Il existe une situation où la construction ci-dessus est particulièrement simple : il s'agit du cas homogène où pour tout $t \geq 0$, $L_t = L$, et où de plus il existe $a > 0$ tel que pour tout $x \in S$, $L(x,x) = -a$.

En effet, il apparaît alors, en notant pour tout $n \geq 1$, $T_n = \tau_n - \tau_{n-1}$, que les suites $(T_n)_{n\geq 1}$ et $(Y_n)_{n\geq 0}$ sont indépendantes, que $(T_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes et identiquement distribuée suivant la loi exponentielle de paramètre a (i.e. de moyenne a^{-1}), et que $(Y_n)_{n\geq 0}$ est une chaîne de Markov dont le noyau q de probabilités de transitions est défini par

$$\forall\, x, y \in S, \qquad q(x,y) = \begin{cases} L(x,y)/a & \text{, si } y \neq x \\ 0 & \text{, si } y = x \end{cases}$$

Ainsi notamment il est extrémement facile de construire le processus de Markov homogène associé à la probabilité μ et au générateur Λ introduits à la fin de la section précédente, car les conditions ci-dessus sont vérifiées avec $a = \mathrm{card}(S) - 1$.

On en déduit alors, via les densités que l'on y a décrites, la solution du problème de martingales associé à une loi m_0 et à une famille de générateurs $(L_t)_{t\geq 0}$ quelconques, du moins sur $\mathcal{D}_T$, pour tout $T > 0$, mais par une variante du théorème 2, il n'est pas très difficile de montrer que ces probabilités se prolongent en une loi $\mathbb{P}_{0,m_0}$ sur tout $(D(\mathbb{R}_+, S), \mathcal{D})$.

Exercice* : rédiger les détails de la suggestion précédente.

d) Soit L un générateur irréductible et $(L_t)_{t\geq 0}$ une famille continue de générateurs satisfaisants

$$\forall\, t \geq 0, \forall\, x \neq y \in S, \qquad L_t(x,y) = 0 \;\;\Rightarrow\;\; L(x,y) = 0$$

Soit m_0 une loi initiale et $\mathbb{P}_{0,m_0}$ la probabilité associée comme d'habitude à m_0 et à $(L_t)_{t\geq 0}$.

Pour $t > 0$, notons m_t la loi de X_t sous $\mathbb{P}_{0,m_0}$. La description précédente permet alors de réaliser que m_t, pour $t > 0$, charge tous les points de S : pour tout $x \in S$, $m_t(x) > 0$.

On pouvait aussi se rendre compte de ce phénomène en utilisant les résultats de densité de la fin de la section précédente (sous une forme contraposée) et le fait que pour le processus homogène de générateur L et de loi initiale m_0, ceci peut s'obtenir directement grâce à la formule explicite

$$\forall\, t > 0, \qquad m_t \;=\; m_0 \exp(tL) = m_0 \left(\sum_{n\in\mathbb{N}} \frac{t^n L^n}{n!} \right)$$

Pour terminer cette section, nous allons présenter une dernière propriété du processus de Markov X sous $\mathbb{P}_{0,m_0}$.

Proposition 23 (quasi-continuité à gauche)

> Soit $(S_n)_{n\geq 0}$ suite croissante de $(\mathcal{D}_t)_{t\geq 0}$-temps d'arrêt.
>
> Notons $S = \lim_{n\to\infty} S_n \leq +\infty$, qui est nécessairement aussi un temps d'arrêt. Alors $\mathbb{P}_{0,m_0}$-p.s., sur $\{S < +\infty\}$, on est assuré de
>
> $$\lim_{n\to\infty} X_{S_n} \;=\; X_S$$

Preuve :

On constate que l'événement $\{S < +\infty,\, \lim_{n\to\infty} X_{S_n} \neq X_S\}$ sera réalisé si et seulement s'il existe un $p \in \mathbb{N}^*$ tel que $S = \tau_p < +\infty$ et tel que pour tout $n \in \mathbb{N}$, $S_n < \tau_p$.

La proposition sera donc démontrée dès que l'on se sera assuré que pour tout $p \geq 1$,

$$\mathbb{P}_{0,m_0}[S = \tau_p < +\infty,\ \text{et}\ \forall\, n \in \mathbb{N},\ S_n < \tau_p] \;=\; 0$$

Par l'absurde, si ceci n'était pas réalisé, on pourrait trouver un $M > 0$ tel que

$$\lim_{\epsilon\to 0_+} \lim_{n\to\infty} \mathbb{P}_{0,m_0}[\tau_p - \epsilon \leq S_n < \tau_p \leq M] \;>\; 0$$

Or en conditionnant par $\mathcal{D}_{S_n}$ et en reprenant un calcul de la preuve de la proposition 22, on fait apparaître que pour $0 < \epsilon \leq 1$,

$$\begin{aligned}
\mathbb{P}_{0,m_0}[\tau_p - \epsilon \leq S_n < \tau_p \leq M] &\leq \mathbb{E}_{0,m_0}[\mathbb{1}_{\{S_n\leq M\}} \mathbb{P}_{S_n, X_{S_n}}[\tau_1 \leq \epsilon]] \\
&= A\epsilon
\end{aligned}$$

avec $A = \max_{0\leq s\leq M+1,\, y\in S} |L_s(y,y)|$, et on aboutit ainsi à une contradiction.

$\square$

Ce résultat admet pour conséquence que chacun des temps d'arrêt τ_i, pour $i \geq 1$, est totalement inaccessible, au sens où il n'existe pas de suite croissante $(S_n^{(i)})_{n\geq 0}$ de temps d'arrêt tel que

$$\mathbb{P}_{0,m_0}[\tau_i = \lim_{n\to\infty} S_n^{(i)} < +\infty \text{ et } \forall\, n \in \mathbb{N},\ S_n^{(i)} < \tau_i] > 0$$

1.6 Exemple : les algorithmes de recuit simulé généralisés

Pour revenir sur l'introduction de ce chapitre, une fois donnés une probabilité initiale m_0 et une famille continue de générateurs $(L_t)_{t\geq 0}$, le processus de Markov associé sera le processus des coordonnées canoniques $(X_t)_{t\geq 0}$ sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P}) \stackrel{\text{déf.}}{=} (D(\mathbb{R}_+, S), \mathcal{D}, \mathbb{P}_{0,m_0})$, où $\mathbb{P}_{0,m_0}$ est la solution du problème de martingales relatif à m_0 et à $(L_t)_{t\geq 0}$. Si le besoin se fait ressentir de devoir reconsidérer toute une famille de lois $(\mathbb{P}_{t,x})_{t\geq 0,\, x\in S}$ de semi-groupe de classe C^1 dont les générateurs sont donnés par $(L_s)_{s\geq 0}$, on désignera plutôt la probabilité $\mathbb{P}_{t,x}$, $t \geq 0$, $x \in S$, comme étant la solution du problème de martingales correspondant à δ_x et à $(L_{t+s})_{s\geq 0}$.

Nous sommes maintenant en mesure de décrire le type de processus de Markov inhomogènes réguliers qui nous préoccupera essentiellement par la suite : les algorithmes de recuit simulé généralisés (finis).

Pour ceux-ci, on dispose d'un générateur irréductible L sur S, qui est parfois appelé noyau d'intensités de transition a priori ou d'exploration, et d'une fonction de coût compatible V : il s'agit d'une application $V : S \times S \setminus \{(x,x)\,;\, x \in S\} \to \overline{\mathbb{R}}_+$ telle que

$$\forall\, x \neq y \in S, \qquad V(x,y) = +\infty \;\;\Leftrightarrow\;\; L(x,y) = 0$$

Pour $\beta \geq 0$, qui représente l'inverse de la température, on note ensuite L_β le générateur défini par

$$\forall\, x \neq y \in S, \qquad L_\beta(x,y) \;\;=\;\; \exp(-\beta V(x,y))L(x,y)$$

Soit $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ une application de classe C^1, figurant l'évolution temporelle de l'inverse de la température, qui satisfait $\lim_{t\to+\infty} \beta_t = +\infty$ (le système considéré aura donc tendance à se refroidir) et soit m_0 une probabilité initiale sur S.

L'algorithme de recuit simulé (généralisé) associé aux objets précédents désignera alors le processus de Markov $(X_t)_{t\geq 0}$ relatif à m_0 et à $(L_{\beta_t})_{t\geq 0}$.

Les cas classiques correspondent à ceux pour lesquels on fait les deux hypothèses supplémentaires suivantes :

- La probabilité invariante μ associée au générateur L est réversible.
- La fonction de coût V dérive d'une certaine « énergie » $U \in F(S)$, au sens où

$$\forall\, x \neq y, \qquad V(x,y) = \begin{cases} (U(y) - U(x))_+ & \text{, si } L(x,y) > 0 \\ +\infty & \text{, sinon} \end{cases}$$

avec la notation $a_+ = \max(0,a)$, pour tout $a \in \mathbb{R}$.

Dans cette situation, rappelons que la mesure de Gibbs μ_β associée à μ, U et à l'inverse de température $\beta \geq 0$, est la probabilité définie par

$$\forall\, x \in S, \qquad \mu_\beta(x) \;\;=\;\; Z_\beta^{-1} \exp(-\beta U(x))\mu(x)$$

où Z_β est la constante de renormalisation, parfois appelée fonction de partition en physique statistique,

$$Z_\beta \;\;=\;\; \sum_{y\in S} \exp(-\beta U(y))\mu(y)$$

Remarquons qu'elle est réversible pour le générateur L_β (et donc invariante), car on calcule que pour tous $x \neq y$,

$$\begin{aligned} \mu_\beta(x)L_\beta(x,y) &= Z_\beta^{-1} \exp(\beta((U(y) - U(x))_+ + U(x)))\mu(x)L(x,y) \\ &= Z_\beta^{-1} \exp(\beta(U(y) \vee U(x)))\mu(x)L(x,y) \\ &= \mu_\beta(y)L_\beta(y,x) \end{aligned}$$

Dans le cas généralisé, ce résultat admet une sorte d'extension, car on peut définir une énergie virtuelle de la manière suivante : tout d'abord notons que pour tout $\beta \geq 0$, le générateur L_β est irréductible, il admet donc une unique probabilité invariante, disons μ_β.

Reprenons les notations introduites avant la proposition 11. Si $(T, x) \in \mathcal{T}(S) \times S$ est un arbre pointé, on lui associe la quantité

$$V(T, x) \quad = \quad \sum_{y \in S \setminus \{x\}} V(y, p_{(T,x)}(y))$$

puis on définit les objets suivants

$$\begin{aligned}
U(x) &= \min_{T \in \mathcal{T}(S)} V(T, x) \\
N &= \{y \in S \,:\, U(y) = \min_S U\} \\
V_*(x) &= \{T \in \mathcal{T}(S) \,:\, V(T, x) = U(x)\} \\
r(x) &= \sum_{T \in V_*(x)} \prod_{y \in S \setminus \{x\}} L(y, p_{(T,x)}(y)) \\
\rho(x) &= \frac{r(x)}{\sum_{y \in N} r(y)} > 0
\end{aligned}$$

La proposition 11 permet alors immédiatement d'obtenir

Proposition 24

Pour tout $x \in S$, on a quand β tend vers $+\infty$,

$$\mu_\beta(x) \sim \rho(x) \exp(-\beta(U(x) - \min_S U))$$

L'application $U - \min_S U \in F(S)$ est pour cette raison appelée l'énergie virtuelle associée à la fonction de coût V.

Pour ce genre d'algorithme, les transitions de x à y, avec $V(x, y) > 0$, seront donc de plus en plus rares en temps grand, et ce d'autant plus que $V(x, y)$ est grand, elles étant d'ailleurs « impossibles » dans le cas extrême où $V(x, y) = +\infty$. Dans les situations classiques, le processus sera ainsi asymptotiquement proche d'un algorithme de descente de U (qui correspond au générateur L_∞, en général réversible, les classes de récurrence étant certaines classes d'équivalence des minima locaux), toutefois les perturbations devront être suffisantes pour permettre au processus d'échapper aux puits d'attraction des minima locaux non globaux de U et permettre une recherche des minima globaux, but pour lequel ces algorithmes stochastiques ont été initialement introduits, dans le cadre de problèmes d'optimisation combinatoire sur des espaces finis très grands (la forme généralisée est quant à elle apparue en parallélisant ce type de processus, pour chercher à en accélérer la vitesse de convergence).

Heuristiquement l'idée est que si l'évolution de la température est suffisament lente, la loi du processus à un temps t grand restera proche de la probabilité instantanée d'équilibre μ_{β_t}, et puisque celle-ci à tendance à se concentrer sur les minima globaux de U, on obtiendra effectivement une convergence vers ceux-ci en temps grand.

Ce type de comportement sera étudié en détail dans des chapitres ultérieurs.

Références

[1] R.M. Blumenthal and R.K. Getoor. *Markov Processes and Potential Theory.* Pure and Applied Mathematics, A Series of Monographs and Textbooks 29. Academic Press, New York, 1968.

[2] C. Dellacherie and P.A. Meyer. *Probabilités et potentiel ; théorie des martingales.* Hermann, 1980.

[3] C. Dellacherie and P.A. Meyer. *Probabilités et potentiel : chapitres XII à XVI ; théorie du potentiel associée à une résolvante, théorie des processus de Markov.* Hermann, 1987.

[4] S. Ethier and T. Kurtz. *Markov Processes, Characterization and Convergence.* Wiley series in probability and mathematical statistics. John Wiley and Sons, New York, 1986.

[5] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of Dynamical Systems.* A Series of Comprehensive Studies in Mathematics 260. Springer-Verlag, 1984.

[6] M. Iosifescu. *Finite Markov Processes and Their Applications.* John Wiley and Sons, 1980.

[7] M. Iosifescu and P. Tăutu. *Stochastic Processes and Applications in Biology and Medicine I : Theory.* Biomathematics 3. Springer-Verlag, Berlin, 1973.

[8] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes.* A Series of Comprehensive Studies in Mathematics 288. Springer-Verlag, 1987.

[9] T.M. Liggett. *Interacting Particle Systems.* A Series of Comprehensive Studies in Mathematics 276. Springer-Verlag, New York, 1985.

[10] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson, 1970.