

**DEA de Probabilités et Applications**

**2003-2004**

**Chaînes de Markov, Processus de Poisson et Applications**

**Jean Jacod**

# Chapitre 1

## Chaînes de Markov

### 1.1 Introduction

L'idée des chaînes de Markov est très simple: il s'agit d'une suite  $(X_n)_{n \in \mathbb{N}}$  de variables aléatoires à valeurs dans un espace mesurable  $(E, \mathcal{E})$ , définies sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ , et telle que pour tout  $n$  on ait la propriété suivante:

$$\left. \begin{array}{l} \text{Conditionnellement à la valeur de } X_n, \text{ les variables } (X_0, \dots, X_{n-1}) \\ \text{d'une part, } (X_{n+1}, X_{n+2}, \dots) \text{ d'autre part, sont indépendantes.} \end{array} \right\} \quad (1.1.1)$$

Ce modèle permet de rendre compte d'un très grand nombre de situations concrètes.

Mathématiquement, la propriété précédente s'exprime ainsi: soit  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$  ( $\sigma(\cdot)$  désigne la tribu "engendrée par..."). On doit alors avoir pour tout  $n \geq 0$ :

$$\mathbb{E}(g(X_{n+1}, X_{n+2}, \dots) | \mathcal{F}_n) = \mathbb{E}(g(X_{n+1}, X_{n+2}, \dots) | \sigma(X_n)) \quad (1.1.2)$$

pour toute fonction  $g$  mesurable bornée (ou positive) sur  $E^\infty$ . Ci-dessus,  $\mathbb{E}(\cdot | \mathcal{F}_n)$  désigne l'espérance conditionnelle, et il faut évidemment définir la mesurabilité d'une fonction définie sur  $E^\infty$ . On verra que cela se ramène à la propriété apparemment plus faible suivante: pour tout  $n \geq 0$  on a

$$\mathbb{E}(f(X_{n+1}) | \mathcal{F}_n) = \mathbb{E}(f(X_{n+1}) | \sigma(X_n)) \quad (1.1.3)$$

pour toute fonction  $f$  mesurable bornée (ou positive) sur  $E$ . Ainsi, en itérant (1.1.3), on voit que la construction d'une chaîne de Markov se fait en deux étapes:

- on se donne la loi de la "variable initiale"  $X_0$ ,
- pour chaque  $n \geq 0$  on se donne un mécanisme décrivant  $X_{n+1}$ , ou plutôt sa loi, en fonction de  $X_n$ .

On voit donc que l'étude mathématique des chaînes de Markov nécessite un certain nombre de préliminaires que nous introduisons dans le paragraphe suivant.

## 1.2 Préliminaires

### 1.2.1 Espérances conditionnelles et classes monotones

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé, et  $\mathcal{G}$  une sous-tribu de  $\mathcal{F}$ . Rappelons d'abord la

**Définition 1.2.1** On appelle *espérance conditionnelle* de la variable aléatoire réelle  $Z$  (supposée positive, resp. intégrable) si  $\mathcal{G}$  (ou, "par rapport à  $\mathcal{G}$ "), et on note  $\mathbb{E}(Z|\mathcal{G})$ , toute variable aléatoire  $U$  telle que:

1.  $U$  est  $\mathcal{G}$ -mesurable.
2.  $\mathbb{E}(ZV) = \mathbb{E}(UV)$  pour toute variable aléatoire  $V$  positive (resp. bornée) et  $\mathcal{G}$ -mesurable.

L'espérance conditionnelle existe et est positive (resp. intégrable) si  $Z$  est positive (resp. intégrable). Elle est "unique à un ensemble de mesure nulle près": si  $U$  est une version de l'espérance conditionnelle, la variable  $U'$  en est une autre si et seulement si  $U = U'$   $\mathbb{P}$ -p.s.

Il y a essentiellement deux façons d'introduire l'espérance conditionnelle:

1. Le théorème de Radon-Nikodym: si  $Z$  est une variable aléatoire réelle positive, on définit la mesure (positive)  $\mu(A) = \mathbb{E}(Z1_A)$ . La restriction  $\mu_{\mathcal{G}}$  de la mesure  $\mu$  à la sous-tribu  $\mathcal{G}$  est absolument continue par rapport à  $\mathbb{P}_{\mathcal{G}}$  et l'espérance conditionnelle est définie comme la densité de  $\mu_{\mathcal{G}}$  par rapport à  $\mathbb{P}_{\mathcal{G}}$ . Pour  $Z$  de signe quelconque, on procède par différence.
2. Le théorème de projection dans les espaces de Hilbert: si  $F = L^2(\Omega, \mathcal{F}, \mathbb{P})$  et  $G = L^2(\Omega, \mathcal{G}, \mathbb{P})$  alors l'espérance conditionnelle de  $Z \in F$  est définie comme sa projection orthogonale sur  $G$ . Cette définition est ensuite étendue aux variables positives ou intégrables.

Un résultat très utile pour la caractérisation des espérances conditionnelles est fourni par le théorème des classes monotones:

**Théorème 1.2.2** Soit  $\mathcal{H}$  un espace vectoriel de fonctions réelles bornées définies sur  $\Omega$  et  $\mathcal{C}$  un ensemble de parties de  $\Omega$  stable par intersection finie. On suppose que

- (i)  $1 \in \mathcal{H}$ ,
- (ii) si  $f_n \in \mathcal{H}$  et si  $0 \leq f_n \uparrow f$  et  $f$  est bornée, alors  $f \in \mathcal{H}$ .
- (iii) pour tout  $A \in \mathcal{C}$ ,  $1_A \in \mathcal{H}$ .

Alors  $\mathcal{H}$  contient toutes les fonctions  $\sigma(\mathcal{C})$ -mesurables bornées.

A titre d'application, on en déduit facilement le résultat suivant: soit  $X$  une variable définie sur  $(\Omega, \mathcal{F}, \mathbb{P})$  et à valeurs dans un espace mesurable  $(E, \mathcal{E})$ . On suppose que la sous-tribu  $\mathcal{G}$  est engendrée par une famille dénombrable  $(Y_n)$  de variables à valeurs dans un espace mesurable  $(Y, \mathcal{Y})$  et soit  $f$  une fonction réelle  $\mathcal{E}$ -mesurable, positive (resp. bornée). Pour tester qu'une variable  $U$ ,  $\mathcal{G}$ -mesurable positive (resp. bornée) est bien une version

de l'espérance conditionnelle de  $f(X)$  quand  $\mathcal{G}$  il suffit de vérifier que:

$$\mathbb{E}(f(X) \prod_{k=1}^n f_k(Y_k)) = \mathbb{E}(U \prod_{k=1}^n f_k(Y_k)) \quad (1.2.1)$$

pour tout  $n \geq 1$  et tout système  $f_1, f_2, \dots, f_n$  de fonctions  $\mathcal{Y}$ -mesurables, positives (resp. bornées).

## 1.2.2 Espérances conditionnelles et lois conditionnelles

Soit  $X$  une variable définie sur  $(\Omega, \mathcal{F}, \mathbb{P})$ , à valeurs dans un espace mesurable  $(E, \mathcal{E})$ , et  $\mathcal{G}$  une sous-tribu de  $\mathcal{F}$ . On définit la "loi conditionnelle de  $X$  si  $\mathcal{G}$ " par la formule:

$$\mathbb{P}_X(B|\mathcal{G}) = \mathbb{E}(1_B \circ X|\mathcal{G}), \quad B \in \mathcal{E}.$$

Bien sûr, pour  $B$  fixé cette formule ne définit  $\mathbb{P}_X(B)$  que comme une classe d'équivalence de variables aléatoires, donc le théorème de convergence monotone ne fournit qu'une  $\sigma$ -additivité  $\mathbb{P}$ -presque sûre. Il est important de savoir s'il est possible de trouver pour chaque  $B \in \mathcal{E}$  une version de cette classe d'équivalence, telle que pour chaque  $\omega$  la fonction  $B \mapsto \mathbb{P}_X(B|\mathcal{G})(\omega)$  soit une "vraie" probabilité sur  $(E, \mathcal{E})$ .

De manière plus précise introduisons la notion suivante, qui sera fondamentale pour les chaînes de Markov:

**Définition 1.2.3** Si  $(E, \mathcal{E})$  et  $(F, \mathcal{F})$  sont deux espaces mesurables, on appelle *mesure de transition* de  $(E, \mathcal{E})$  dans  $(F, \mathcal{F})$  toute famille  $Q = (Q(x, B) : x \in E, B \in \mathcal{F})$  de nombres dans  $[0, \infty]$ , qui vérifie:

1.  $x \mapsto Q(x, A)$  est  $\mathcal{E}$ -mesurable pour tout  $A \in \mathcal{F}$ ;
2.  $A \mapsto Q(x, A)$  est une mesure sur  $(F, \mathcal{F})$ , pour tout  $x \in E$ .

Si de plus  $Q(x, F) = 1$  pour tout  $x$ , on dit que  $Q$  est une *probabilité de transition*. On écrit aussi  $Q$  ainsi:  $Q(x, dy)$ .

On peut montrer (c'est le théorème de Jirina) que, dès que la tribu  $\mathcal{E}$  est séparable (= engendrée par une *algèbre* dénombrable, comme le sont les tribus boréliennes de  $\mathbb{R}$  ou de  $\mathbb{R}^n$ ), il existe une probabilité de transition  $Q$  de  $(\Omega, \mathcal{G})$  dans  $(E, \mathcal{E})$ , telle que pour tout  $B \in \mathcal{E}$ , la variable aléatoire  $\omega \mapsto Q(\omega, B)$  soit une version de l'espérance conditionnelle  $\mathbb{P}_X(B|\mathcal{G})(\omega)$ . On écrit souvent  $\mathbb{P}_{X|\mathcal{G}}$  au lieu de  $Q$ , et on a donc pour toute fonction  $\mathcal{E}$ -mesurable  $f$ , positive ou bornée:

$$\mathbb{E}(f(X)|\mathcal{G}) = \int_E f(x) \mathbb{P}_{X|\mathcal{G}}(\cdot, dx) \quad (1.2.2)$$

(cette égalité est vraie si  $f = 1_B$  pour  $B \in \mathcal{E}$  par définition même; par linéarité elle est vraie si  $f$  est mesurable étagée positive; par limite monotone elle est vraie si  $f$  est mesurable positive; par différence elle est vraie si  $f$  est mesurable bornée).

La probabilité de transition  $\mathbb{P}_{X|\mathcal{G}}$  s'appelle *une version régulière de la loi conditionnelle*. Il est important de noter qu'elle n'est pas définie de manière unique: si par exemple

$A$  est un ensemble  $\mathcal{G}$ -mesurable négligeable, si on pose  $\mathbb{P}'_{X/\mathcal{G}}(\omega, dx) = \mathbb{P}_{X/\mathcal{G}}(\omega, dx)$  pour  $\omega \notin A$  et  $\mathbb{P}'_{X/\mathcal{G}}(\omega, dx) = \mu(dx)$  pour  $\omega \in A$  (où  $\mu$  est une probabilité arbitraire sur  $E$ ), on obtient une autre version régulière de la loi conditionnelle.

### 1.2.3 Opérations sur les mesures de transition

Dans (1.2.2) on a vu apparaitre de manière détournée une première opération sur les mesures de transition. Plus généralement, si  $Q(x, dy)$  est une mesure de transition de  $(E, \mathcal{E})$  dans  $(F, \mathcal{F})$ , on pose

$$Qf(x) = \int_F Q(x, dy)f(y), \quad \forall x \in E \quad (1.2.3)$$

pour toute fonction  $\mathcal{F}$ -mesurable  $f$  sur  $F$  telle que cette formule ait un sens pour tout  $x \in E$  (par exemple si  $f \geq 0$ , ou si  $f$  est bornée quand  $Q$  est une probabilité de transition). Noter que la fonction  $Qf$  est  $\mathcal{E}$ -mesurable: c'est évident si  $f = 1_B$ , par linéarité c'est vrai si  $f$  est positive mesurable étagée, par limite monotone c'est vrai si  $f$  est mesurable positive, et par différence c'est vraie si  $f$  est mesurable et si l'intégrale (1.2.3) existe pour tout  $x$ .

On a une opération "duale": si  $\mu$  est une mesure sur  $(E, \mathcal{E})$ , on pose

$$\mu Q(B) = \int_E \mu(dx)Q(x, B), \quad \forall B \in \mathcal{F}, \quad (1.2.4)$$

ce qui définit (d'après la linéarité de l'intégrale et le théorème de limite monotone) une nouvelle mesure  $\mu Q$  sur  $(F, \mathcal{F})$ . Si  $\varepsilon_x$  désigne la masse de Dirac en  $x$ , on a  $\varepsilon_x Q = Q(x, \cdot)$ .

Plus généralement, si  $Q$  est comme ci-dessus et si  $R$  est une mesure de transition de  $(F, \mathcal{F})$  dans un autre espace  $(G, \mathcal{G})$ , on pose

$$(QR)(x, C) = \int_F Q(x, dy)R(y, C), \quad \forall x \in E, \forall C \in \mathcal{G}. \quad (1.2.5)$$

Noter que  $(QR)(x, C) = Qf(x)$  si  $f(y) = R(y, C)$ , tandis que  $(QR)(x, C) = \mu R(C)$  si  $\mu(\cdot) = Q(x, \cdot)$ : par suite  $QR$  est une mesure de transition de  $(E, \mathcal{E})$  dans  $(G, \mathcal{G})$ .

Par itération, si  $(E_i, \mathcal{E}_i)$  est une suite d'espaces mesurables, si  $Q_i$  est une mesure de transition de  $(E_{i-1}, \mathcal{E}_{i-1})$  dans  $(E_i, \mathcal{E}_i)$ , si  $\mu_0$  est une mesure sur  $(E_0, \mathcal{E}_0)$ , et si  $f_i$  est une fonction mesurable sur  $(E_i, \mathcal{E}_i)$ , on a ainsi:

- $Q_1 Q_2 \dots Q_i$ : une mesure de transition de  $(E_0, \mathcal{E}_0)$  dans  $(E_i, \mathcal{E}_i)$ ;
- $\mu_0 Q_1 Q_2 \dots Q_i$ : une mesure sur  $(E_i, \mathcal{E}_i)$ ;
- $Q_1 Q_2 \dots Q_i f_i$ : une fonction mesurable sur  $(E_0, \mathcal{E}_0)$ ;
- $\mu_0 Q_1 Q_2 \dots Q_i f_i$ : un nombre !

Si on suppose de plus que  $\mu_0$  est une probabilité et que les  $Q_i$  sont des probabilités de transition, notre objectif maintenant est de leur associer une probabilité sur l'espace produit  $\prod E_i$ .

### 1.2.4 Probabilités sur un produit: le cas fini

Commençons par le cas d'un produit de deux espaces  $(E, \mathcal{E})$  et  $(F, \mathcal{F})$ . On pose  $G = E \times F$  et  $\mathcal{G} = \mathcal{E} \otimes \mathcal{F}$  (rappelons que  $\mathcal{G}$  est la plus petite tribu sur  $G$  contenant les pavés mesurables  $A \times B$ , i.e. tels que  $A \in \mathcal{E}$  et  $B \in \mathcal{F}$ ).

Rappelons l'essentiel du théorème de Fubini, dans le cas des probabilités: si  $\mu$  et  $\nu$  sont des probabilités sur les deux espaces ci-dessus,  $\eta = \mu \otimes \nu$  est l'unique mesure sur  $(G, \mathcal{G})$  telle que  $\eta(A \times B) = \mu(A)\nu(B)$  pour tous  $A \in \mathcal{E}$  et  $B \in \mathcal{F}$ , et on a pour toute fonction  $\mathcal{G}$ -mesurable  $f$ , bornée ou positive:

$$\int f d\eta = \int_E \mu(dx) \left( \int_F \nu(dy) f(x, y) \right).$$

Enfin  $\eta$  est la loi d'un couple de variables  $(X, Y)$  lorsque  $X$  et  $Y$  sont de lois respectives  $\mu$  et  $\nu$ , et sont *indépendantes*.

Ce qui suit est une extension du théorème de Fubini, avec essentiellement la même preuve, lorsqu'on veut construire la loi d'un couple de variables  $(X, Y)$  avec  $X$  de loi  $\mu$ , et la loi conditionnelle de  $Y$  sachant que  $X = x$  égale à  $Q(x, \cdot)$ .

**Théorème 1.2.4** *Soit  $\mu$  une probabilité sur  $(E, \mathcal{E})$  et  $Q$  une probabilité de transition de  $(E, \mathcal{E})$  dans  $(F, \mathcal{F})$ . Il existe une probabilité  $\eta$  et une seule sur le produit  $(G, \mathcal{G})$ , telle que*

$$\eta(A \times B) = \int \mu(dx) 1_A(x) Q(x, B), \quad \forall A \in \mathcal{E}, \forall B \in \mathcal{F}. \quad (1.2.6)$$

*De plus, pour toute fonction  $f$  mesurable sur  $(G, \mathcal{G})$ , positive ou bornée, la fonction  $x \mapsto \int_F Q(x, dy) f(x, y)$  est  $\mathcal{E}$ -mesurable, et on a*

$$\int f d\eta = \int_E \mu(dx) \left( \int_F Q(x, dy) f(x, y) \right). \quad (1.2.7)$$

*Si enfin  $X$  et  $Y$  sont deux variables aléatoires à valeurs dans  $E$  et  $F$  respectivement, définies sur un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  arbitraire, et telles que la loi du couple  $(X, Y)$  soit  $\eta$ , alors on a:*

- la loi de  $X$  est  $\mu$ ,
  - $Q(X(\omega), \cdot)$  est une version régulière de  $\mathbb{P}_{Y/\sigma(X)}$ .
- $$\left. \vphantom{\begin{matrix} \bullet \\ \bullet \end{matrix}} \right\} \quad (1.2.8)$$

Noter à l'inverse qu'il existe toujours un couple  $(X, Y)$  vérifiant les dernières propriétés ci-dessus: il suffit de prendre  $\Omega = G$ ,  $\mathcal{A} = \mathcal{G}$ ,  $\mathbb{P} = \eta$ , et pour  $(X, Y)$  l'application identité de  $G$  dans lui-même.

**Preuve.** 1) Soit  $\mathcal{H}$  l'espace de toutes les fonctions  $f$  mesurables bornées sur  $(G, \mathcal{G})$  pour lesquelles  $x \mapsto \int_F Q(x, dy) f(x, y)$  soit  $\mathcal{E}$ -mesurable, et  $\mathcal{C}$  l'ensemble des pavés mesurables  $A \times B$ , qui est stable par intersection finie. Il est immédiat que  $\mathcal{H}$  satisfait les hypothèses du théorème 1.2.2, donc égale la classe de toutes les fonctions  $\mathcal{G}$ -mesurables bornées. Par

limite croissante, on en déduit aussi que  $x \mapsto \int_F Q(x, dy)f(x, y)$  est  $\mathcal{E}$ -mesurable pour toute  $f$  mesurable positive.

2) D'après ce qui précède, pour tout  $C \in \mathcal{G}$  on peut poser

$$\eta(C) = \int_E \mu(dx) \left( \int_F Q(x, dy)1_C(x, y) \right).$$

Par linéarité et limite monotone, on voit que cela définit une probabilité  $\eta$  sur  $(G, \mathcal{G})$ , qui vérifie évidemment (1.2.6). On a (1.2.7) lorsque  $f = 1_C$ . Par linéarité, limite monotone, puis différence, on en déduit (1.2.7) pour toute fonction  $f$  mesurable, positive ou bornée.

3) L'unicité de  $\eta$  vérifiant (1.2.6) découle du théorème de Fubini (via une nouvelle application du théorème des classes monotones).

4) Il reste à prouver (1.2.8). Le fait que  $\mathcal{L}(X) = \mu$  (où  $\mathcal{L}(X)$  désigne la loi de  $X$ ) découle de (1.2.6) appliqué avec  $B = F$ , donc  $Q(x, B) = 1$  pour tout  $x$ . En appliquant (1.2.7) à  $f(x, y) = g(x)1_B(y)$ , on obtient

$$\mathbb{E}(g(X)1_B(Y)) = \int_E \mu(dx)g(x)Q(x, B) = \mathbb{E}(g(X)Q(X, B)).$$

La seconde partie de (1.2.8) en découle immédiatement. □

Cette procédure se généralise aisément à un produit fini d'espaces. Nous nous contentons ci-dessous d'une version adaptée aux chaînes de Markov, mais il est possible de faire plus général.

La situation est la suivante: on a des espaces mesurables  $(E_i, \mathcal{E}_i)$ , des probabilités de transition  $Q_i$  de  $(E_{i-1}, \mathcal{E}_{i-1})$  dans  $(E_i, \mathcal{E}_i)$ , et une probabilité  $\mu_0$  sur  $(E_0, \mathcal{E}_0)$ . On pose  $F_N = E_0 \times \dots \times E_N$  et  $\mathcal{F}_N = \mathcal{E}_0 \otimes \dots \otimes \mathcal{E}_N$ . On a alors:

**Théorème 1.2.5** (1) Pour toute fonction  $f$  mesurable sur  $(F_N, \mathcal{F}_N)$ , positive ou bornée, les formules  $f_N = f$  et

$$f_n(x_0, \dots, x_n) = \int Q_{n+1}(x_n, dx_{n+1})f_{n+1}(x_0, \dots, x_{n+1}) \quad (1.2.9)$$

pour  $n = N - 1, N - 2, \dots, 0$ , définissent (par récurrence descendante) pour chaque  $n$  une fonction  $f_n$  qui est mesurable sur  $(F_n, \mathcal{F}_n)$ .

(2) Il existe une probabilité  $\eta_N$  et une seule sur  $(F_N, \mathcal{F}_N)$ , notée aussi (un peu abusivement)  $\mu_0 \otimes Q_1 \otimes \dots \otimes Q_N$ , telle que pour toute fonction  $f$  mesurable sur  $(F_N, \mathcal{F}_N)$ , positive ou bornée, on ait (avec la notation (1.2.9)):

$$\int f d\eta_N = \int f_0 d\mu_0. \quad (1.2.10)$$

(3) Si enfin les  $X_i$  sont des variables aléatoires à valeurs dans  $E_i$ , définies sur un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  arbitraire, et telles que la loi du  $N + 1$ -uplet  $(X_0, \dots, X_N)$

soit  $\eta$ , alors on a (pour  $n = 1, \dots, N$ ):

- la loi de  $X_0$  est  $\mu_0$ ,
- $Q_n(X_{n-1}(\omega), \cdot)$  est une version régulière de  $\mathbb{P}_{X_n/\sigma(X_0, X_1, \dots, X_{n-1})}$
- $(Q_n(X_{n-1}(\omega), \cdot) \otimes Q_{n+1} \otimes \dots \otimes Q_N)(\cdot)$  est une version régulière de  $\mathbb{P}_{(X_n, X_{n+1}, \dots, X_N)/\sigma(X_0, X_1, \dots, X_{n-1})}$ .

(1.2.11)

On combine en général (1.2.9) et (1.2.10), pour écrire:

$$\int f d\eta_N = \int \mu_0(dx_0) \int Q_1(x_0, dx_1) \dots \int Q_N(x_{N-1}, dx_N) f(x_0, \dots, x_N). \quad (1.2.12)$$

Ci-dessus, la convention est – comme d’habitude – que les intégrales sont effectuées de la droite vers la gauche.

**Preuve.** En posant  $Q_n((x_0, \dots, x_{n-1}), B) = Q_n(x_{n-1}, B)$ , on voit qu’on peut considérer  $Q_n$  comme une probabilité de transition de  $(F_{n-1}, \mathcal{F}_{n-1})$  dans  $(E_n, \mathcal{E}_n)$ . Donc la mesurabilité de  $f_n$  découle immédiatement de la mesurabilité de  $f_{n+1}$  et du théorème 1.2.4 appliqué à  $E = F_n$  et  $F = E_{n+1}$ , d’où le (1) par récurrence descendante.

Le (2) se montre par récurrence sur  $N$ : lorsque  $N = 0$  il n’y a rien à montrer. Supposons que ce soit vrai pour  $N - 1$ . En remarquant que les fonctions  $f_n$  associées à  $f$  pour  $n \leq N$  sont aussi les fonctions associées à  $f_{N-1}$  lorsqu’on démarre la récurrence descendante à  $N - 1$ , on voit que  $\int f_0 d\mu_0 = \int f_{N-1} d\eta_{N-1}$ . En d’autres termes, la formule (1.2.10) s’écrit aussi

$$\int f d\eta_N = \int \eta_{N-1}(du) \int Q_N(u, dx_N) f(u, x_N),$$

en notant  $u = (x_0, \dots, x_{N-1})$ . Pour obtenir l’existence et l’unicité de  $\eta_N$  il suffit alors d’appliquer le théorème 1.2.4 à  $E = F_{N-1}$  et  $F = E_N$ .

Enfin pour (3), là encore la première partie de (1.2.11) est évidente. En appliquant (1.2.12) à  $f(x_0, \dots, x_N) = 1_B(x_n, \dots, x_N)g(x_0, \dots, x_{n-1})$ , on voit que

$$\mathbb{E}(g(X_0, \dots, X_{n-1})1_B(X_n, \dots, X_N)) = \mathbb{E}(g(X_0, \dots, X_{n-1})R_n(X_{n-1}, B)),$$

où nous avons posé  $R_n(x, \cdot) = Q_n(x, \cdot) \otimes Q_{n+1} \otimes \dots \otimes Q_N$ . La troisième partie de (1.2.11) en découle, et la seconde partie en est un cas particulier.  $\square$

On peut déjà observer que (1.2.11) nous donne la “propriété de Markov” (1.1.3), et même (1.1.2), pour la suite finie  $(X_n)_{0 \leq n \leq N}$ . Il faut maintenant voir ce qui se passe si on veut construire une suite infinie de variables: c’est un problème plus difficile, qui fait l’objet du paragraphe suivant.

### 1.2.5 Probabilités sur un produit: le cas infini

La situation est encore la suivante: pour  $i \in \mathbb{N}$  on a des espaces mesurables  $(E_i, \mathcal{E}_i)$ , des probabilités de transition  $Q_i$  de  $(E_{i-1}, \mathcal{E}_{i-1})$  dans  $(E_i, \mathcal{E}_i)$  si  $i \geq 1$ , et une probabilité  $\mu_0$



sur  $(E_0, \mathcal{E}_0)$ . Mais, outre les espaces  $(F_N, \mathcal{F}_N)$  définis ci-dessus, nous avons aussi l'espace  $(F_\infty, \mathcal{F}_\infty)$ , construit ainsi: d'une part  $F_\infty = \prod_{i \in \mathbb{N}} E_i$ ; d'autre part  $\mathcal{F}_\infty$  est la plus petite tribu contenant les pavés mesurables  $\prod_{i \in \mathbb{N}} A_i$  (avec  $A_i \in \mathcal{E}_i$ ), ou de manière équivalente contenant les pavés mesurables  $A_0 \times \dots \times A_n \times \prod_{i \geq n+1} E_i$  avec  $n$  fini arbitraire.

Toute fonction  $f$  sur un  $F_n$  sera aussi considérée ci-dessous comme une fonction sur  $F_\infty$ , ne dépendant que des "coordonnées"  $x_0, \dots, x_n$ .

Pour bien comprendre la seconde partie de l'énoncé ci-après, il convient de rappeler que si les  $X_i$  sont des variables aléatoires à valeurs dans les espaces  $(E_i, \mathcal{E}_i)$ , la suite infinie  $(X_i)_{i \in \mathbb{N}}$  peut être considérée comme une variable à valeurs dans  $(F_\infty, \mathcal{F}_\infty)$ : la tribu  $\mathcal{F}_\infty$  est construite précisément pour qu'on ait cette propriété. La "loi" de la suite infinie est donc une probabilité sur  $(F_\infty, \mathcal{F}_\infty)$ . A l'inverse, n'importe quelle probabilité  $\eta$  sur cet espace est la loi d'une suite infinie  $(X_i)_{i \in \mathbb{N}}$  de variables sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ : il suffit de prendre  $\Omega = F_\infty$ ,  $\mathcal{A} = \mathcal{F}_\infty$ ,  $\mathbb{P} = \eta$ , et  $X_i(x_0, x_1, \dots) = x_i$  (les "applications coordonnées").

**Théorème 1.2.6 (Ionescu–Tulcea)** (1) *Il existe une probabilité  $\eta$  et une seule sur  $(F_\infty, \mathcal{F}_\infty)$ , notée aussi  $\mu_0 \otimes Q_1 \otimes \dots \otimes Q_n \otimes \dots$ , telle que pour tout  $N \in \mathbb{N}$  et toute fonction  $f$  mesurable sur  $(F_N, \mathcal{F}_N)$ , positive ou bornée, on ait (avec la mesure  $\eta_N$  sur  $(F_N, \mathcal{F}_N)$  du théorème 1.2.5):*

$$\int f d\eta = \int f d\eta_N. \quad (1.2.13)$$

(2) *Si les  $X_i$  sont des variables aléatoires à valeurs dans  $E_i$ , définies sur un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  arbitraire, et telles que la loi de la suite infinie  $(X_n)_{n \in \mathbb{N}}$  soit  $\eta$ , alors on a pour tout  $n$ :*

$$\left. \begin{aligned} &\bullet \text{ la loi de } X_0 \text{ est } \mu_0, \\ &\bullet Q_n(X_{n-1}(\omega), \cdot) \text{ est une version régulière de } \mathbb{P}_{X_n/\sigma(X_0, X_1, \dots, X_{n-1})} \\ &\bullet \left( Q_n(X_{n-1}(\omega), \cdot) \otimes Q_{n+1} \otimes \dots \right) (\cdot) \text{ est une version régulière de} \end{aligned} \right\} \quad (1.2.14)$$

$$\mathbb{P}_{(X_n, X_{n+1}, \dots)/\sigma(X_0, X_1, \dots, X_{n-1})}.$$

**Preuve.** a) De même qu'une fonction sur  $F_n$  peut être considérée comme une fonction sur  $F_\infty$ , une partie  $A$  de  $F_n$  est identifiée à la partie  $A \times \prod_{i \geq n+1} E_i$  de  $F_\infty$ . Ainsi, la tribu  $\mathcal{F}_n$  sur  $F_n$  peut être considérée comme une tribu sur  $F_\infty$ . Avec cette convention, on pose  $\mathcal{B} = \cup_n \mathcal{F}_n$ : la classe  $\mathcal{B}$  est une algèbre sur  $F_\infty$ , qui engendre la tribu  $\mathcal{F}_\infty$ .

Si  $A \in \mathcal{B}$ , on a  $A \in \mathcal{F}_N$  pour un  $N$  fini, et on pose  $\eta(A) = \eta_N(A)$ , avec  $\eta_N$  définie dans le théorème 1.2.5, qui implique aussi que  $\eta_{N=p}(A) = \eta_N(A)$  pour tout  $p \geq 1$ . Cela définit donc une fonction additive d'ensembles sur  $\mathcal{B}$ , et d'après le théorème de prolongement des mesures on sait qu'elle s'étend en une probabilité sur  $\mathcal{F}_\infty$ , notée encore  $\eta$  et nécessairement unique, si et seulement si on a

$$\eta(A_n) \rightarrow 0 \quad \text{pour toute suite } (A_n) \text{ d'éléments de } \mathcal{B} \text{ décroissant vers } \emptyset. \quad (1.2.15)$$

Si c'est le cas, la probabilité  $\eta$  coïncide avec  $\eta_N$  sur  $\mathcal{F}_N$ , d'où (1.2.13).

b) Pour obtenir (1) il nous reste donc à montrer (1.2.15). Si  $n \geq 1$  et  $A \in \mathcal{F}_N$  pour un  $N \geq n$ , on pose

$$R_n(x_0, \dots, x_{n-1}; A) = \int Q_n(x_{n-1}, dx_n) \dots \int Q_N(x_{N-1}, dx_N) 1_A(x_0, \dots, x_N).$$

D'après le (1) du théorème 1.2.5 cette quantité est  $\mathcal{F}_{n-1}$ -mesurable en  $(x_0, \dots, x_{n-1})$ . On notera que cette quantité ne dépend pas de  $N$ , pourvu que  $N \geq n$  et que  $A \in \mathcal{F}_N$ , de sorte que  $R_n(x_0, \dots, x_{n-1}; A)$  est en fait définie pour tout  $A \in \mathcal{B}$ , et de plus si  $A \in \mathcal{F}_N$  et  $N \geq n$  on a avec  $u = (x_0, \dots, x_{n-1})$ :

$$\eta(A) = \eta_N(A) = \int \eta_n(du) R_n(u, A). \quad (1.2.16)$$

Soit maintenant une suite  $A_j$  dans  $\mathcal{B}$ , décroissant vers  $\emptyset$ . La suite  $\eta(A_j)$  décroît vers une limite  $a \geq 0$ . On va supposer que  $a > 0$  et en déduire une contradiction, ce qui prouvera (1.2.15). D'abord, (1.2.16) appliqué à  $n = 1$  donne

$$\eta(A_j) = \int \mu_0(dx_0) R_1(x_0; A_j) \downarrow a > 0,$$

et comme  $R_1(x_0; A_j)$  décroît et est compris entre 0 et 1, on déduit du théorème de Lebesgue qu'il existe au moins un  $x_0 \in E_0$  tel que  $R_1(x_0; A_j) \downarrow a_1 > 0$ .

Supposons alors qu'on ait trouvé pour un  $n \geq 1$  des points  $x_0, \dots, x_{n-1}$  tels que

$$R_n(x_0, \dots, x_{n-1}; A_j) \downarrow a_n > 0, \quad (1.2.17)$$

lorsque  $j \rightarrow \infty$ . En remarquant que

$$R_n(x_0, \dots, x_{n-1}; A_j) = \int Q_n(x_{n-1}, dx_n) R_{n+1}(x_0, \dots, x_n; A_j) \geq a_n,$$

le même argument que ci-dessus entraîne qu'il existe  $x_n \in E_n$  et  $a_{n+1} > 0$  tels que  $R_{n+1}(x_0, \dots, x_{n-1}, x_n; A_j) \downarrow a_{n+1}$ . On construit ainsi par récurrence une suite  $(x_n)_{n \in \mathbb{N}}$  telle que  $x_n \in E_n$  et qu'on ait (1.2.17) pour tout  $n$ .

Mais chaque  $A_j$  est élément de  $\mathcal{F}_{N_j}$  pour un certain entier  $N_j$ . En revenant à la définition de  $R_n$ , on voit que si  $n > N_j$  le nombre  $R_n(x_0, \dots, x_{n-1}; A_j)$  vaut 1 si  $(x_0, \dots)$  appartient à  $A_j$  et vaut 0 autrement. En comparant ceci à (1.2.17), on voit que nécessairement  $(x_0, \dots) \in A_j$ . Comme ceci est vrai pour tout  $j$ , le point  $(x_0, \dots)$  appartient à  $\bigcap_j A_j = \emptyset$ , qui est vide, d'où une contradiction.

c) Plaçons-nous dans la situation de (2). La première partie de (1.2.14) est encore une fois évidente, et la seconde est un cas particulier de la troisième. Pour celle-ci, posons  $R_n(x, \cdot) = Q_n(x, \cdot) \otimes Q_{n+1} \otimes Q_{n+2} \otimes \dots$  (en vertu de (1), c'est une probabilité sur l'espace  $(H_n, \mathcal{H}_n) = (\prod_{i \geq n} E_i, \otimes_{i \geq n} \mathcal{E}_i)$ ). D'après (1.2.13) et (3) du théorème 1.2.5,  $\int f dR_n(X_{n-1}, \cdot)$  est une version de l'espérance conditionnelle de  $\mathbb{E}(f(X_n, \dots) | \sigma(X_0, \dots, X_{n-1}))$  pour toute fonction  $f$  bornée mesurable sur  $(H_n, \mathcal{H}_{n,m})$  pour un  $m \geq n$ , où  $\mathcal{H}_{n,m} = \otimes_{n \leq i \leq m} \mathcal{E}_i$ . Un argument de classe monotone montre qu'il en est de même si  $f$  est bornée et  $\mathcal{H}_n$ -mesurable, ce qui donne le résultat.  $\square$

**Remarque:** Ce théorème – plutôt difficile – n’est en aucune manière plus simple lorsque les espaces  $E_i$  sont égaux à  $\mathbb{R}$ , ou sont dénombrables, ou même finis ! (sauf bien-sûr s’ils sont réduits à un seul point). De même il n’est pas plus simple lorsque les probabilités de transition sont de la forme  $Q_i(x, dy) = \mu_i(dy)$ , où  $\mu_i$  est une probabilité sur  $(E_i, \mathcal{E}_i)$ . Dans ce dernier cas, en vertu de la partie (2) du théorème précédent la mesure  $\eta$  est la loi d’une suite de variables  $(X_n)$  qui sont *indépendantes*, chaque  $X_n$  étant à valeurs dans  $E_n$  et de loi  $\mu_n$ : comme corollaire, on obtient ainsi l’existence d’une suite de variables indépendantes de lois données.

### 1.3 Chaînes de Markov: définitions

Venons-en maintenant aux chaînes de Markov. On se donne un “espace d’état” mesurable  $(E, \mathcal{E})$ , qui est *a priori* quelconque. Il existe plusieurs définitions possibles d’une chaîne de Markov.

- La notion la plus faible consiste à dire qu’on a un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  et une suite  $(X_n)$  de variables à valeurs dans  $(E, \mathcal{E})$ , telle qu’on ait (1.1.2) (où  $g$  “mesurable sur  $E^\infty$ ” signifie mesurable par rapport à la tribu produit tensoriel  $\mathcal{E}^{\otimes \infty}$ ). Par conditionnements successifs, cette propriété se ramène à (1.1.3). Cette dernière propriété donne moralement la loi conditionnelle de  $X_n$  sachant  $(X_0, \dots, X_{n-1})$ , mais on n’est pas sûr de l’existence de versions régulières de ces lois conditionnelles.

- Une seconde définition possible consiste à dire qu’on a (1.1.3) *avec de plus* des probabilités conditionnelles régulières: cela revient à dire qu’on a des probabilités de transition  $P_n$  de  $(E, \mathcal{E})$  dans lui-même, telles que pour tout  $n$  et toute fonction mesurable bornée ou positive  $f$  sur  $E$ ,

$$\mathbb{E}(f(X_n) | \sigma(X_0, \dots, X_{n-1})) = P_n f(X_{n-1}). \quad (1.3.1)$$

(Comme l’espérance conditionnelle n’est définie qu’à un ensemble de probabilité nulle près, nous omettons systématiquement d’écrire le “p.s.” dans des égalités faisant intervenir les espérances ou probabilités conditionnelles). De ce point de vue, la “loi initiale” (i.e., la loi de  $X_0$ ) est considérée comme donnée.

- Une situation un peu plus générale consiste à supposer que l’espace  $(\Omega, \mathcal{F})$  est muni d’une *filtration*: une suite  $(\mathcal{F}_n)_{n \in \mathbb{N}}$  de sous-tribus de  $\mathcal{F}$ , avec  $\mathcal{F}_n \subset \mathcal{F}_{n+1}$  et  $X_n$  est  $\mathcal{F}_n$ -mesurable pour tout  $n$  (on dit que  $(X_n)$  est *adapté* à la filtration); on remplace alors (1.3.1) par

$$\mathbb{E}(f(X_n) | \mathcal{F}_{n-1}) = P_n f(X_{n-1}). \quad (1.3.2)$$

Lorsque  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ , (1.3.2) se ramène à (1.3.1).

- Une quatrième définition possible consiste à se donner les transitions  $P_n$ , considérées comme décrivant le “mécanisme d’évolution” de la chaîne, mais à laisser la loi initiale arbitraire: pour toute probabilité  $\mu$  sur  $(E, \mathcal{E})$  on a donc une probabilité  $\mathbb{P}_\mu$  sur  $(\Omega, \mathcal{F})$  vérifiant (1.3.2), et telle que  $\mathcal{L}(X_0) = \mu$ .

Les situations 2 et 3 ci-dessus nous conduisent à poser:

**Définition 1.3.1** 1) Soit  $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \in \mathbb{N}}, \mathbb{P})$  un espace probabilisé filtré. Une suite  $(X_n)_{n \in \mathbb{N}}$  de variables aléatoires sur cet espace, à valeurs dans  $(E, \mathcal{E})$ , est dite vérifier la  $(\mathcal{F}_n)$ -propriété de Markov si pour tout  $n$  la variable  $X_n$  est  $\mathcal{F}_n$ -mesurable et si on a (1.3.2) pour une suite de probabilités de transition  $(P_n)$ .

2) Si on a  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$  (donc (1.3.2)=(1.3.1)), on dit simplement qu'on a la propriété de Markov.

3) Si  $P_n = P$  ne dépend pas de  $n$ , la propriété de Markov est dite *homogène*.

La situation 4 ci-dessus n'est réellement intéressante que dans le cas homogène. Dans ce cadre, nous allons donner la "vraie" définition des chaînes de Markov, celle que nous utiliserons dans la suite (on distinguera "chaîne de Markov", définie ci-après, et "propriété de Markov", définie ci-dessus). Cette définition peut sembler compliquée, mais en fait elle permet de rendre la suite beaucoup plus simple.

On part donc d'un ensemble mesurable  $(\Omega, \mathcal{F})$ , muni de variables aléatoires  $(X_n)_{n \in \mathbb{N}}$  à valeurs dans  $(E, \mathcal{E})$ ; on pose  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ , et on suppose que  $\mathcal{F} = \bigvee_n \mathcal{F}_n$ . Cet espace est aussi muni d'une "translation"  $\theta$ , qui est une application de  $\Omega$  dans lui-même telle que

$$X_{n+1} = X_n \circ \theta, \quad \forall n. \quad (1.3.3)$$

On se donne enfin une famille  $(\mathbb{P}_x)_{x \in E}$  de probabilités sur  $(\Omega, \mathcal{F})$ , telle que  $x \mapsto \mathbb{P}_x(A)$  est  $\mathcal{E}$ -mesurable pour tout  $A \in \mathcal{F}$  (donc,  $\mathbb{P}_x(d\omega)$  est une probabilité de transition de  $(E, \mathcal{E})$  dans  $(\Omega, \mathcal{F})$ ). On suppose que

$$\mathbb{P}_x(X_0 = x) = 1 \quad \forall x \in E. \quad (1.3.4)$$

On notera  $\mathbb{E}_x$  l'espérance, relativement à la probabilité  $\mathbb{P}_x$ .

**Définition 1.3.2** Le terme  $\mathbf{X} = (\Omega, \mathcal{F}, (X_n), \theta, (\mathbb{P}_x))$  défini ci-dessus s'appelle une *chaîne de Markov* (sous-entendu: "homogène") s'il existe une probabilité de transition  $P$  de  $(E, \mathcal{E})$  dans lui-même telle que pour tout  $n$ , tout  $x \in E$ , et toute fonction mesurable  $f$  sur  $E$ , bornée ou positive, on a

$$\mathbb{E}_x(f(X_n) | \mathcal{F}_{n-1}) = Pf(X_{n-1}). \quad (1.3.5)$$

Pour toute probabilité  $\mu$  sur  $(E, \mathcal{E})$ , on pose

$$\mathbb{P}_\mu(A) = \int \mu(dx) \mathbb{P}_x(A) \quad \forall A \in \mathcal{F}, \quad (1.3.6)$$

ce qui en vertu de (1.2.4) définit une nouvelle probabilité sur  $(\Omega, \mathcal{F})$ . L'espérance relative à  $\mathbb{P}_\mu$  sera notée  $\mathbb{E}_\mu$ .

**Proposition 1.3.3** Si  $\mathbf{X}$  est une chaîne de Markov, pour toute probabilité  $\mu$  sur  $(E, \mathcal{E})$ , la loi de  $X_0$  sous  $\mathbb{P}_\mu$  (appelée "loi initiale") est  $\mu$ . De plus pour tout  $n$  et toute variable aléatoire réelle  $Y$  sur  $(\Omega, \mathcal{F})$ , positive ou bornée, en notant  $\theta^n$  la  $n$ -ième itérée de  $\theta$ , la variable  $Y \circ \theta^n$  est  $\mathcal{F}$ -mesurable et on a:

$$\mathbb{E}_\mu(Y \circ \theta^n | \mathcal{F}_n) = \mathbb{E}_{X_n}(Y). \quad (1.3.7)$$

**Preuve.** D'abord, si  $A \in \mathcal{E}$ , on a par (1.3.3) et (1.3.4):

$$\mathbb{P}_\mu(X_0 \in A) = \int \mu(dx) \mathbb{P}_x(X_0 \in A) = \int \mu(dx) 1_A(x) = \mu(A),$$

d'où la première assertion.

D'après le théorème des classes monotones 1.2.2, et puisque  $\mathcal{F} = \sigma(X_0, X_1, \dots)$ , pour la seconde assertion il suffit de montrer le résultat quand  $Y$  est de la forme  $Y = \prod_{i=0}^m f_i(X_i)$ , pour des fonctions bornées mesurables  $f_i$ . Dans ce cas, on a  $Y \circ \theta^n = \prod_{i=0}^m f_i(X_{n+i})$ , donc on voit que  $Y \circ \theta^n$  est  $\mathcal{F}$ -mesurable. Si  $Z$  est une variable bornée  $\mathcal{F}_n$ -mesurable, on a

$$\mathbb{E}_\mu(Z Y \circ \theta^n) = \int \mu(dx) \mathbb{E}_x(Z Y \circ \theta^n).$$

Il nous reste donc à montrer que

$$\mathbb{E}_x(Z Y \circ \theta^n) = \mathbb{E}_x(Z \mathbb{E}_{X_n}(Y)). \quad (1.3.8)$$

Cela se montre par récurrence sur  $m$ . Lorsque  $m = 0$  on a  $Y = f_0(X_0)$ , donc  $\mathbb{E}_x(Y) = f_0(x)$ , et (1.3.8) est évident. Supposons donc que (1.3.8) soit vrai pour tout  $n$  et pour toute variable de la forme  $Y = \prod_{i=0}^{m-1} g_i(X_i)$ ; nous allons alors montrer (1.3.8) pour  $Y = \prod_{i=0}^m f_i(X_i)$ . Posons  $U = \prod_{i=0}^{m-1} f_i(X_i)$  et  $V = U P f_m(X_{m-1})$ . En vertu de (1.3.5) on a

$$\begin{aligned} \mathbb{E}_x(Z Y \circ \theta^n) &= \mathbb{E}_x(Z U \circ \theta^n f_m(X_{n+m})) \\ &= \mathbb{E}_x(\mathbb{E}_x(Z U \circ \theta^n f_m(X_{n+m}) | \mathcal{F}_{n+m-1})) \\ &= \mathbb{E}_x(Z U \circ \theta^n P f_m(X_{n+m-1})) \\ &= \mathbb{E}_x(Z V \circ \theta^n) \\ &= \mathbb{E}_x(Z \mathbb{E}_{X_n}(V)), \end{aligned}$$

où la dernière égalité provient de l'hypothèse de récurrence. En appliquant ce qui précède à  $n = 0$  et  $Z = 1$ , on obtient aussi  $\mathbb{E}_y(Y) = \mathbb{E}_y(V)$  pour tout  $y \in E$ , d'où (1.3.8).  $\square$

Ce résultat nous dit que, sous chaque  $\mathbb{P}_\mu$ , le chaîne  $(X_n)$  admet la propriété de Markov homogène; de plus la probabilité de transition est  $P$ : il suffit d'appliquer (1.3.7) à  $Y = f(X_1)$  et de remarquer que  $\mathbb{E}_x(f(X_1)) = P f(x)$  par (1.3.5). Noter qu'on a aussi démontré (1.1.2) pour  $\mathbb{P}_\mu$ .

**Construction d'une chaîne de Markov:** Etant donnée une probabilité de transition  $P$  arbitraire de  $(E, \mathcal{E})$  dans lui-même, nous sommes en mesure de construire une chaîne de Markov  $\mathbf{X}$  de transition  $P$ . Cela peut se faire sur beaucoup d'espaces  $\Omega$  différents, bien-sûr, mais une manière canonique de faire est la suivante:

On prend  $\Omega = E^{\mathbb{N}}$ , donc un point de  $\Omega$  est une suite infinie  $\omega = (x_0, x_1, \dots)$ . On pose

$$X_n(x_0, \dots) = x_n, \quad \theta(x_0, x_1, \dots) = (x_1, \dots).$$

Enfin, on prend pour  $\mathbb{P}_x$  la probabilité construite dans le théorème de Ionescu-Tulcea 1.2.6, correspondant aux  $(E_i, \mathcal{E}_i) = (E, \mathcal{E})$  et  $Q_i = P$  et à  $\mu_0 = \varepsilon_x$  (masse de Dirac en  $x$ ).

La mesurabilité de  $x \mapsto \mathbb{P}_x(A)$  s'obtient par application de (1) du théorème 1.2.5 avec  $n = 0$  et (1.2.13) et encore une fois un argument de classe monotone. Noter que  $\mathbb{P}_\mu$  est obtenue de la même manière, en prenant  $\mu_0 = \mu$ .

La chaîne de Markov  $\mathbf{X}$  ainsi obtenue s'appelle *la chaîne canonique* associée à la transition  $P$ .

## 1.4 Premières propriétés

Ci-dessous on considère une chaîne de Markov  $\mathbf{X} = (\Omega, \mathcal{F}, (X_n), \theta, (\mathbb{P}_x))$  de transition  $P$  sur  $(E, \mathcal{E})$ . On désigne par  $P^n = P \dots P$  la  $n$ -ième itérée de  $P$  au sens de (1.2.5) (par convention  $P^0$  est la transition "identité", c'est-à-dire  $P^0(x, dy) = \varepsilon_x(dy)$ ).

**Proposition 1.4.1** *Pour tous  $n, m \geq 0$  et toute fonction mesurable  $f$  bornée ou positive, on a*

$$\mathbb{E}_\mu(f(X_{n+m})|\mathcal{F}_n) = P^m f(X_n). \quad (1.4.1)$$

De plus, la loi de  $X_m$  sous  $\mathbb{P}_\mu$  est  $\mu P^m$ .

**Preuve.** Pour  $m = 0$  (1.4.1) est évident, et pour  $m = 1$  c'est (1.3.7) (appliqué à  $Y = f(X_1)$ ). Supposons (1.4.1) vraie pour  $m - 1$ , et toute fonction  $f$  et tout  $n$ . On a

$$\mathbb{E}_\mu(f(X_{n+m})|\mathcal{F}_n) = \mathbb{E}_\mu(\mathbb{E}_\mu(f(X_{n+m})|\mathcal{F}_{n+m-1})|\mathcal{F}_n) = \mathbb{E}_\mu(Pf(X_{n+m-1})|\mathcal{F}_n),$$

qui vaut  $P^{m-1}Pf(X_n) = P^m f(X_n)$  d'après l'hypothèse de récurrence. On a donc (1.4.1) pour tout  $n, m \geq 0$ .

Enfin, la loi  $\mu_m$  de  $X_m$  sous  $\mathbb{P}_\mu$  est donnée par  $\mu_m(A) = \mathbb{P}_\mu(X_m \in A)$ . Mais

$$\mathbb{P}_\mu(X_m \in A) = \mathbb{E}_\mu(\mathbb{P}_\mu(X_m \in A|\mathcal{F}_0)) = \mathbb{E}_\mu(P^m(X_0, A)) = \int \mu(dx)P^m(x, A) = \mu P^m(A),$$

où on a utilisé (1.4.1) pour la seconde égalité: on a donc la seconde assertion.  $\square$

La seconde propriété, fondamentale pour la suite, est la "propriété forte de Markov". Rappelons qu'un temps d'arrêt est une application de  $\Omega$  dans  $\overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$  qui vérifie  $\{T \leq n\} \in \mathcal{F}_n$  pour tout  $n \in \mathbb{N}$ , ou de manière équivalente  $\{T = n\} \in \mathcal{F}_n$  pour tout  $n$ . On associe à un temps d'arrêt  $T$  sa "tribu antérieure", qui est

$$\mathcal{F}_T = \{A : A \in \mathcal{F}, A \cap \{T \leq n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}\},$$

et ci-dessus on peut remplacer  $A \cap \{T \leq n\}$  par  $A \cap \{T = n\}$  partout si on veut: on vérifie que  $\mathcal{F}_T$  est une tribu, et si  $T(\omega) = m$  pour tout  $\omega$  alors  $T$  est un temps d'arrêt et  $\mathcal{F}_T = \mathcal{F}_m$ .

De même qu'on a défini les itérées  $\theta^n$  de la translation  $\theta$ , pour tout temps d'arrêt  $T$  on peut définir les applications  $\theta^T$  et  $X_T$  de  $\{T < \infty\}$  dans  $\Omega$  et  $E$  respectivement, par

$$\left. \begin{aligned} \theta^T(\omega) &= \theta^{T(\omega)}(\omega) = \theta^n(\omega) \\ X_T(\omega) &= X_{T(\omega)}(\omega) = X_n(\omega) \end{aligned} \right\} \text{ sur l'ensemble } \{T = n\}. \quad (1.4.2)$$

D'une part on a  $\{X_T \in A\} \cap \{T = n\} = \{X_n \in A\} \cap \{T = n\}$ , qui est  $\mathcal{F}_n$ -mesurable si  $A \in \mathcal{E}$ : donc  $X_T$  est  $\mathcal{F}_T$ -mesurable en restriction à l'ensemble  $\{T < \infty\}$ . D'autre part pour toute variable aléatoire  $Y$  on a  $\{Y \circ \theta^T \in A\} \cap \{T = n\} = \{Y \circ \theta^n \in A\} \cap \{T = n\}$ , qui est  $\mathcal{F}$ -mesurable pour tout borélien  $A$ : donc  $Y \circ \theta^T$  est  $\mathcal{F}$ -mesurable en restriction à l'ensemble  $\{T < \infty\}$ . Ces propriétés de mesurabilité montrent que l'assertion ci-dessous a un sens.

**Proposition 1.4.2 (Propriété forte de Markov)** *Pour tout temps d'arrêt  $T$  et toute variable aléatoire  $Y$  positive ou bornée on a*

$$\mathbb{E}_\mu(Y \circ \theta^T | \mathcal{F}_T) = \mathbb{E}_{X_T}(Y) \quad \text{sur l'ensemble } \{T < \infty\}. \quad (1.4.3)$$

Pour bien comprendre (1.4.3) il convient de remarquer que l'ensemble  $\{T < \infty\}$  est  $\mathcal{F}_T$ -mesurable, donc ci-dessus on peut aussi bien ajouter son indicatrice en facteur dans chaque membre, et à gauche faire passer cette indicatrice à l'intérieur de l'espérance conditionnelle: les deux membres ont donc *a priori* un sens.

En appliquant (1.4.3) à  $Y = f(X_m)$ , on obtient (en utilisant aussi la proposition 1.4.1):

$$\mathbb{E}_\mu(f(X_{T+m}) | \mathcal{F}_T) = P^m f(X_T) \quad \text{sur l'ensemble } \{T < \infty\}. \quad (1.4.4)$$

**Preuve.** Les variables  $U = Y \circ \theta^T$  et  $V = \mathbb{E}_{X_T}(Y)$  sont bien définies sur l'ensemble  $\{T < \infty\}$ , et la seconde est  $\mathcal{F}_T$ -mesurable. Ainsi il suffit de montrer que si  $A \in \mathcal{F}_T$  on a  $\mathbb{E}_\mu(U 1_A 1_{\{T < \infty\}}) = \mathbb{E}_\mu(V 1_A 1_{\{T < \infty\}})$ . Il suffit même de montrer que pour chaque  $n \in \mathbb{N}$  on a

$$\mathbb{E}_\mu(U 1_A 1_{\{T=n\}}) = \mathbb{E}_\mu(V 1_A 1_{\{T=n\}}).$$

Mais le membre de gauche est  $\mathbb{E}_\mu(1_{A \cap \{T=n\}} Y \circ \theta^n)$ , tandis que le membre de droite est  $\mathbb{E}_\mu(1_{A \cap \{T=n\}} \mathbb{E}_{X_n}(Y))$ , et  $A \cap \{T = n\} \in \mathcal{F}_n$ : il suffit alors d'appliquer (1.3.7).  $\square$

Les notions suivantes seront souvent utilisées; elles sont relatives à la transition  $P$ :

**Définition 1.4.3** 1) Une mesure (positive)  $\eta$  sur  $(E, \mathcal{E})$  est dite

- *invariante* si  $\eta P = \eta$ ,
- *sous-invariante* si  $\eta P \leq \eta$ .

2) Une fonction mesurable positive sur  $(E, \mathcal{E})$  est dite

- *harmonique* si  $Pf = f$ ,
- *surharmonique* si  $Pf \leq f$ .

**Proposition 1.4.4** (1) *Si  $f$  est une fonction harmonique on a  $P^n f = f$  pour tout  $n$ , et si de plus  $\mu(f) < \infty$  le processus  $f(X_n)$  est une martingale sous la probabilité  $\mathbb{P}_\mu$ .*

(2) *Si  $f$  est une fonction surharmonique on a  $P^n f \leq f$  pour tout  $n$ , et le processus  $f(X_n)$  est une surmartingale sous toute probabilité  $\mathbb{P}_\mu$ .*

**Preuve.** Comme  $P$ , en tant qu'opérateur sur les fonctions, est positif, les premières assertions de (1) et (2) sont évidentes. De plus  $f(X_n)$  est  $\mathcal{F}_n$ -mesurable, et on a pour toute probabilité initiale  $\mu$ :

$$\mathbb{E}_\mu(f(X_{n+m})|\mathcal{F}_n) = P^m f(X_n).$$

Si  $f$  est surharmonique on a donc  $\mathbb{E}_\mu(f(X_{n+m})|\mathcal{F}_n) \leq f(X_n)$ , ce qui est l'inégalité caractéristique des surmartingales. Quand  $f$  est harmonique, on a  $\mathbb{E}_\mu(f(X_{n+m})|\mathcal{F}_n) = f(X_n)$ , ce qui est l'égalité caractéristique des martingales: pour obtenir que  $f(X_n)$  est une  $\mathbb{P}_\mu$ -martingale il faut de plus que chaque  $f(X_n)$ , ou de manière équivalente  $f(X_0)$ , soit  $\mu$ -intégrable, ce qui revient à dire que  $\mu(f) < \infty$ .  $\square$

**Définition 1.4.5** Si  $f$  est une fonction mesurable positive, son *potentiel* est la fonction

$$Uf = \sum_{n \geq 0} P^n f \quad (\text{avec } P^0 f = f). \quad (1.4.5)$$

Par le théorème de limite monotone, il est facile de voir qu'on définit ainsi une mesure de transition  $U$ , appelée aussi le potentiel de  $P$ , de sorte que  $Uf$  définie ci-dessus est l'action de  $U$  sur  $f$ . Pour toute mesure positive  $\mu$  on a de même la mesure  $\mu U$ . On écrit bien-sûr  $U = \sum_{n \geq 0} P^n$ .

Comme  $Uf = f + \sum_{n \geq 1} P^n f$ , on obtient l'équation (pour  $f \geq 0$ ):

$$U = I + PU = I + UP \quad \text{i.e.,} \quad Uf = f + PUf = f + UPf. \quad (1.4.6)$$

En particulier,  $Uf$  est une fonction surharmonique. De même pour toute mesure (positive)  $\mu$  on a  $\mu U = \mu + \mu PU = \mu + \mu UP$ , donc  $\mu U$  est une mesure sous-invariante.

On notera aussi que, par le théorème de limite monotone, pour toute fonction mesurable positive  $f$  et toute probabilité  $\mu$ ,

$$\mu Uf = \mathbb{E}_\mu \left( \sum_{n \geq 0} f(X_n) \right). \quad (1.4.7)$$

Pour terminer ce paragraphe, examinons la stationnarité de la chaîne  $(X_n)$ . On rappelle qu'un processus  $(X_n)_{n \in \mathbb{N}}$  est *stationnaire* si pour tout  $n$  la loi du  $n+1$ -uplet  $(X_m, X_{m+1}, \dots, X_{m+n})$  ne dépend pas de l'entier  $m$ .

**Proposition 1.4.6** *La chaîne  $(X_n)$  est stationnaire sous la probabilité  $\mathbb{P}_\mu$  si et seulement si la probabilité  $\mu$  est invariante.*

**Preuve.** Si la chaîne est stationnaire sous  $\mathbb{P}_\mu$ , la loi  $\mu P$  de  $X_1$  (cf. proposition 1.4.1) égale la loi  $\mu$  de  $X_0$ , donc  $\mu$  est invariante.

Inversement supposons  $\mu$  invariante. Soit  $n \geq 0$ . La loi  $\zeta_{n,m}$  de  $(X_m, X_{m+1}, \dots, X_{m+n})$  est donnée par  $\zeta_{n,m}(A) = \mathbb{P}_\mu((X_m, X_{m+1}, \dots, X_{m+n}) \in A)$  pour  $A \in \mathcal{E}^{\otimes(n+1)}$ . Si  $Y = 1_A(X_0, \dots, X_n)$  et  $f(x) = \mathbb{E}_x(Y)$ , on a

$$\eta_{n,m}(A) = \mathbb{E}_\mu(Y \circ \theta^m) = \mathbb{E}_\mu(\mathbb{E}_\mu(Y \circ \theta^m)|\mathcal{F}_m) = \mathbb{E}_\mu(f(X_m)) = \mu P^m f$$



par (1.3.7) et la proposition 1.4.1. L'invariance de  $\mu$  entraîne  $\mu P^m f = \mu(f)$ , donc  $\eta_{m,m}$  ne dépend pas de  $m$ .  $\square$

Le fait d'avoir une chaîne stationnaire est crucial pour les applications: par exemple (on le verra en exercice) un grand nombre de “files d'attente” sont des chaînes de Markov, et le fait qu'elles soient stationnaires signifie qu'il n'y a pas engorgement de la file. Une bonne partie de ce chapitre sera en fait consacré à l'étude de l'existence d'une probabilité invariante, rendant la chaîne stationnaire.

## 1.5 Chaînes discrètes: classifications

On dit qu'une chaîne de Markov  $\mathbf{X}$  est *discrète* si son espace d'états  $E$  est *fini ou dénombrable*. Dans ce cas,  $\mathcal{E}$  est toujours la tribu de toutes les parties de  $E$ , et toute fonction sur  $E$  est mesurable. Les points de  $E$  seront – selon l'habitude – notés  $i, j, k$ .

### 1.5.1 Notations

Toute mesure  $\mu$  sur  $E$  est caractérisée par les mesures des singletons  $\mu_i = \mu(\{i\})$ , via la formule  $\mu(A) = \sum_{i \in A} \mu_i$ , et toute famille  $(\mu_i)_{i \in E}$  de nombres positifs ou nuls correspond à une mesure. Cette mesure est une probabilité si de plus  $\sum_i \mu_i = 1$ .

Une probabilité de transition  $P$  est aussi caractérisée par les nombres  $p_{ij} = P(i, \{j\})$ , et inversement une double suite  $(p_{ij})$  correspond à une probabilité de transition si et seulement si on a

$$p_{ij} \geq 0, \quad \sum_j p_{ij} = 1.$$

Ainsi on peut identifier  $P = (p_{ij})$  avec une “matrice” infinie, à termes positifs, la somme de chaque ligne valant 1.

Si  $Q = (q_{ij})$  est une autre probabilité de transition, alors le produit  $R = PQ$  est

$$R = (r_{ij}), \quad \text{avec} \quad r_{ij} = \sum_k p_{ik} q_{kj}.$$

On a de même, pour une mesure  $\mu$  et une fonction  $f$ :

$$(\mu P)_i = \sum_j \mu_j p_{ji}, \quad (Pf)(i) = \sum_j p_{ij} f(j),$$

de sorte qu'en identifiant une mesure à une *matrice ligne*, et une fonction à une *matrice colonne*, les différents produits (1.2.3), (1.2.4) et (1.2.5) sont les produits usuels des matrices (infinies, si  $E$  est dénombrable).

Revenons à notre chaîne de Markov de transition  $P$ . On définit les temps d'arrêt suivants:

$$T_i = \inf(n \geq 1 : X_n = i), \quad T_i^1 = T_i, \quad T_i^{n+1} = \inf(m > T_i^n : X_m = i),$$

avec la convention habituelle que l'inf de l'ensemble vide vaut  $+\infty$  (Exercice: montrer que ce sont effectivement des temps d'arrêt). Ensuite, on pose

$$f_{ij}^{(n)} = \mathbb{P}_i(T_j^n < \infty), \quad f_{ij} = f_{ij}^{(1)} = \mathbb{P}_i(T_i < \infty), \quad N_i = \sum_{n=0}^{\infty} 1_{\{i\}}(X_n).$$

Ainsi,  $N_i$  est le nombre (aléatoire) de fois où la chaîne visite l'état  $i$ . Les éléments de la transition itérée  $P^n$  seront notés  $p_{ij}^{(n)}$  (donc  $p_{ij}^{(0)} = \delta_{ij}$ , le symbole de Kronecker). Le potentiel  $U = (u_{ij})$  est donné par

$$u_{ij} = \sum_{n \geq 0} p_{ij}^{(n)} = \mathbb{E}_i(N_j). \quad (1.5.1)$$

### 1.5.2 Première classification

**Définition 1.5.1** Si  $i, j \in E$ , on dit que  $i$  mène à  $j$ , et on écrit  $i \mapsto j$ , si on a soit  $i = j$ , soit  $\mathbb{P}_i(T_j < \infty) > 0$ . On écrit  $i \sim j$  si  $i \mapsto j$  et  $j \mapsto i$ .

**Proposition 1.5.2** On a  $i \mapsto j$  si et seulement si  $u_{ij} > 0$ , et la relation  $i \sim j$  est une relation d'équivalence.

**Preuve.** Si  $j = i$  on sait que  $u_{ij} \geq 1$  (car  $p_{ij}^{(0)} = \delta_{ij}$ ), et que  $i \mapsto j$ .

Supposons  $j \neq i$ . Si  $u_{ij} > 0$  il existe  $n > 0$  avec  $p_{ij}^{(n)} > 0$ , et bien-sûr  $\mathbb{P}_i(T_j < \infty) \geq \mathbb{P}_i(X_n = j) = p_{ij}^{(n)} > 0$ , donc  $i \mapsto j$ . Si inversement  $i \mapsto j$  il existe  $n \geq 1$  avec  $0 < \mathbb{P}_i(T_j = n) \leq \mathbb{P}_i(X_n = j) = p_{ij}^{(n)}$ , donc  $u_{ij} > 0$ .

La relation  $\sim$  est évidemment réflexive et symétrique. Pour montrer la transitivité il suffit de vérifier que si  $i \mapsto j \mapsto k$ , alors  $i \mapsto k$ . Mais, si  $i \mapsto j \mapsto k$ , d'après ce qui précède il existe  $n \geq 0$  avec  $p_{ij}^{(n)} > 0$  et  $m \geq 0$  avec  $p_{jk}^{(m)} > 0$ . Comme  $P^{n+m} = P^n P^m$ , il vient

$$p_{ik}^{(n+m)} = \sum_{l \in E} p_{il}^{(n)} p_{lk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0,$$

d'où le résultat. □

Comme  $\sim$  est une relation d'équivalence, on peut considérer les classes d'équivalence associées, qui forment une partition de  $E$  et qui s'appellent *les classes* de la chaîne de Markov. La signification intuitive d'une classe est donnée dans la:

**Proposition 1.5.3** Soit  $C$  une classe et  $T = \inf\{n \geq 0 : X_n \notin C\}$ . Pour tout  $i \in C$  on a

$$\mathbb{P}_i(T < \infty, \text{ il existe } n > T \text{ avec } X_n \in C) = 0.$$

En d'autres termes, si la chaîne quitte une classe elle ne peut pas y revenir.

**Preuve.** Le résultat de l'énoncé équivaut à dire (puisque  $\mathbb{N}$  et  $C$  sont dénombrables) que

$$\mathbb{P}_i(T < \infty, X_T = k, X_{T+n} = j) = 0 \quad \forall n \in \mathbb{N}, \forall j \in C, \forall k \notin C \quad (1.5.2)$$

(puisque  $X_T \notin C$  si  $T < \infty$  par définition de  $T$ ). Mais  $T$  est un temps d'arrêt, donc le membre de gauche ci-dessus est d'après la propriété de Markov forte:

$$\mathbb{E}_i \left( 1_{\{T < \infty, X_T = k\}} \mathbb{P}_i(X_i \circ \theta^T = j | \mathcal{F}_T) \right) = \mathbb{P}_i(T < \infty, X_T = k) p_{kj}^{(n)} \leq \mathbb{P}_i(T_k < \infty) p_{kj}^{(n)}.$$

Si l'expression ci-dessus est  $> 0$  on a  $p_{kj}^{(n)} > 0$  et  $\mathbb{P}_i(T_k < \infty) > 0$ , donc  $i \mapsto k \mapsto j$  et comme aussi  $j \mapsto i$  (car  $j \in C$ ) on a  $k \in C$ : par suite on a (1.5.2).  $\square$

**Définition 1.5.4** On appelle *période de l'état  $i$*  et on note  $d_i$  le PGCD de l'ensemble  $M_i$  des  $n \geq 1$  tels que  $p_{ii}^{(n)} > 0$ , avec la convention  $d_i = \infty$  si cet ensemble est vide.

**Proposition 1.5.5** Si  $i \sim j$  on a  $d_i = d_j$ .

Etant donnée une classe  $C$  de la chaîne, on appellera donc *période de la classe* la valeur commune des périodes de ses éléments.

**Preuve.** Supposons bien-sûr  $i \neq j$ . On a vu ci-dessus qu'il existe  $n, m \geq 1$  tels que  $a = p_{ij}^{(n)} > 0$  et  $b = p_{ji}^{(m)} > 0$ . En utilisant  $P^{n+l+m} = P^n P^l P^m$ , le même argument que dans la proposition 1.5.2 montre que  $p_{ii}^{(n+l+m)} \geq ab p_{jj}^{(l)}$ . Donc  $l \in M_j$  implique  $n+m+l \in M_i$ , et aussi évidemment  $n+m \in M_i$ : donc  $d_i$  divise  $n+m$  et  $n+m+l$ , donc  $l$ , pour tout  $l \in M_j$ , donc finalement  $d_i$  divise  $d_j$ . Symétriquement,  $d_i$  divise  $d_j$ , d'où le résultat.  $\square$

**Proposition 1.5.6** Soit  $C$  une classe de période  $d$  finie. On peut diviser  $C$  en  $d$  sous-classes non-vides  $C_0, \dots, C_{d-1}$  telles que, si  $i \in C_r$ , alors  $\mathbb{P}_i$ -presque sûrement la chaîne ne visite  $C_l$  qu'en des instants  $n$  tels que  $n+r = l$  modulo  $d$ .

**Preuve.** Soit  $i_0$  fixé arbitrairement dans  $C$ . Si  $j \in C$ , il existe  $m > 0$  avec  $p_{ji_0}^{(m)} > 0$ , et on note  $r_j$  l'unique entier dans  $\{0, 1, \dots, d-1\}$  tel que  $m = r_j$  modulo  $d$ . Pour tout  $n > 0$  avec  $p_{i_0j}^{(n)} > 0$ , on a  $p_{i_0i_0}^{(n+m)} > 0$ , donc  $d$  divise  $n+m$  et donc  $n = r_j$  modulo  $d$ . En d'autres termes, sous  $\mathbb{P}_{i_0}$  la chaîne ne visite  $j$  qu'en des instants de la forme  $r_j + nd$  pour  $n \in \mathbb{N}$ , presque sûrement.

On note alors  $C_r$  l'ensemble des  $j \in C$  tels que  $r_j = r$ . Les  $C_r$  sont deux-à-deux disjoints, de réunion égale  $C$ . De plus, si la chaîne part de  $i_0$ , à l'instant 1 elle est dans  $C_1$ , puis à l'instant 2 dans  $C_2$ , puis dans  $C_3, \dots$ , puis dans  $C_{d-1}$ , puis  $C_0$ , puis  $C_1$ , etc..., jusqu'à ce qu'elle sorte de  $C$ , auquel cas elle n'y revient plus d'après la proposition 1.5.3. Cela montre à l'évidence la dernière partie de l'énoncé, et aussi que chaque sous-classe  $C_r$  est non vide: en effet  $C_0$  contient au moins  $i_0$ , et si  $C_r$  était vide pour un  $r = 1, \dots, d-1$  la chaîne aurait quitté  $C$   $\mathbb{P}_{i_0}$ -p.s. avant l'instant  $r$ , donc ne pourrait pas revenir en  $i_0$ , ce qui contredirait le fait que  $d = d_{i_0}$  est fini.  $\square$

### 1.5.3 Seconde classification

**Définition 1.5.7** Un état  $i$  est dit *récurrent* si  $f_{ii} = \mathbb{P}_i(T_i < \infty) = 1$ , et *transient* sinon.

**Lemme 1.5.8** On a  $f_{ij}^{(n+1)} = f_{ij}f_{jj}^{(n)}$  et  $u_{ij} = \delta_{ij} + \sum_{n \geq 1} f_{ij}^{(n)}$ .

**Preuve.** On a  $T_j^{n+1} = T_j + T_j^n \circ \theta^{T_j}$  sur l'ensemble  $\{T_j < \infty\}$ , et  $T_j^{n+1} = \infty$  sur  $\{T_j = \infty\}$ . On déduit alors, de la propriété de Markov forte et du fait que  $X_{T_j} = j$  si  $T_j < \infty$ , que

$$\begin{aligned} f_{ij}^{(n+1)} &= \mathbb{P}_i(T_j < \infty, T_j^n \circ \theta^{T_j} < \infty) = \mathbb{E}_i(1_{\{T_j < \infty\}} \mathbb{P}_i(T_j^n \circ \theta^{T_j} < \infty | \mathcal{F}_{T_j})) \\ &= \mathbb{E}_i(1_{\{T_j < \infty\}} \mathbb{P}_{X_{T_j}}(T_j^n < \infty)) = \mathbb{E}_i(1_{\{T_j < \infty\}}) f_{jj}^{(n)} = f_{ij}f_{jj}^{(n)}. \end{aligned}$$

Par ailleurs,  $N_j$  étant une variable à valeurs entières, on a  $\mathbb{E}_i(N_j) = \sum_{n \geq 1} \mathbb{P}_i(N_j \geq n)$ . Comme  $u_{ij} = \mathbb{E}_i(N_j)$  et comme  $\{N_j \geq n\} = \{T_j^n < \infty\}$  si  $X_0 \neq j$  et  $\{N_j \geq n\} = \{T_j^{n-1} < \infty\}$  si  $X_0 = j$ , on obtient immédiatement la seconde propriété énoncée.  $\square$

**Théorème 1.5.9** (1) On a les équivalences

$$i \text{ récurrent} \iff u_{ii} = +\infty \iff \mathbb{P}_i(N_i = \infty) = 1. \quad (1.5.3)$$

Dans ce cas, si  $j \sim i$  alors  $j$  est aussi récurrent, et on a  $f_{ij} = 1$  et  $u_{ij} = \infty$  et  $\mathbb{P}_i(N_j = \infty) = 1$ .

(2) On a les équivalences

$$i \text{ transient} \iff u_{ii} < +\infty \iff \mathbb{P}_i(N_i = \infty) = 0. \quad (1.5.4)$$

Dans ce cas, si  $i \sim j$  alors  $j$  est transient et  $u_{ij} < \infty$  et  $\mathbb{P}_i(N_j = \infty) = 0$ .

Ainsi, la récurrence et la transience sont des *propriétés de classe*: on dira que la classe  $C$  est récurrente (resp. transiente) si ses points sont récurrents (resp. transients).

**Preuve.** D'après le lemme,  $f_{ii}^{(n)} = (f_{ii})^n$  et  $u_{ii} = \sum_{n \geq 0} (f_{ii})^n$ , tandis que  $\mathbb{P}_i(N_i = \infty) = \lim_n f_{ii}^{(n)}$ : on a donc les équivalences (1.5.3) et (1.5.4).

Supposons  $i$  récurrent et  $j \sim i$  avec  $j \neq i$ . Comme dans la proposition 1.5.5 il existe  $n$  et  $m$  avec  $a = p_{ij}^{(n)} > 0$  et  $b = p_{ji}^{(m)} > 0$ , donc  $p_{jj}^{(n+m+l)} \geq abp_{ii}^{(l)}$  pour tout  $l \geq 0$ . Par suite  $u_{jj} \geq abu_{ii} = \infty$ , et donc  $j$  est aussi récurrent. On a de même  $u_{ij} \geq au_{jj} = \infty$ . Si on avait  $f_{ij} < 1$ , le temps d'arrêt  $S = T_i + T_j \circ \theta^{T_i}$ , premier temps de passage en  $j$  de la chaîne après être passé par  $i$ , vérifierait

$$\mathbb{P}_j(S < \infty) = \mathbb{E}_j(1_{\{T_i < \infty\}} \mathbb{P}_j(T_j \circ \theta^{T_i} < \infty | \mathcal{F}_{T_i})) = f_{ji}f_{ij} < 1$$

par la propriété de Markov forte; mais évidemment  $\{S = \infty\} \subset \{N_j < \infty\}$ , donc on aurait  $\mathbb{P}_j(N_j = \infty) < 1$ , ce qui contredirait la récurrence de  $j$ : donc  $f_{ij} = 1$ .

Enfin, comme  $\mathbb{E}_i(N_j) = u_{ij} = f_{ij}u_{jj}$  si  $j \neq i$ , la dernière assertion de (2) découle de ce qui précède.  $\square$

**Corollaire 1.5.10** *Soit  $C$  une classe récurrente, et  $\mu$  une probabilité sur  $E$  ne chargeant que  $C$ . On a alors  $\mathbb{P}_\mu(X_n \in C, \forall n) = 1$  et  $\mathbb{P}_\mu(N_i = \infty) = 1$  pour tout  $i \in C$ .*

**Preuve.** Comme  $\mathbb{P}_\mu(A) = \sum_{j \in C} \mu_j \mathbb{P}_j(A)$  pour tout  $A \in \mathcal{F}$ , il suffit de montrer les assertions si  $\mu = \varepsilon_j$  pour un  $j \in C$ : la seconde assertion vient de la proposition précédente, et la première de la combinaison de la seconde assertion et de la proposition 1.5.3.  $\square$

## 1.6 Chaînes discrètes: propriétés ergodiques

Ce qu'on appelle propriétés ergodiques pour une chaîne de Markov concerne le comportement à l'infini, soit de la chaîne elle-même, soit de ses probabilités de transition  $P^n$ .

On part encore d'une chaîne de Markov  $\mathbf{X}$  avec un espace d'état fini ou dénombrable, et on utilise toutes les notations du paragraphe précédent. On commence par étudier les chaînes dites *irréductibles*, i.e. qui ne possèdent qu'une seule classe.

**Proposition 1.6.1** *Si  $\mathbf{X}$  est une chaîne irréductible récurrente (i.e., tous ses états sont récurrents), toute fonction surharmonique bornée est constante.*

**Preuve.** Soit  $f$  surharmonique bornée. D'après la proposition 1.4.4, le processus  $f(X_n)$  est une surmartingale bornée sous  $\mathbb{P}_i$ , donc pour tout temps d'arrêt  $\mathbb{P}_i$ -p.s. fini  $T$  on a  $\mathbb{E}_i(f(X_T)) \leq \mathbb{E}_i(f(X_0)) = f(i)$ . En particulier  $\mathbb{P}_i(T_j < \infty) = f_{ij} = 1$  pour tout  $j$  et  $f(X_{T_j}) = f(j)$  si  $T_j < \infty$ : donc  $f(j) \leq f(i)$ .  $\square$

**Théorème 1.6.2** *Si  $\mathbf{X}$  est une chaîne irréductible récurrente, la mesure  $\mu = (\mu_i)$  définie par*

$$\mu_i = \mathbb{E}_{i_0} \left( \sum_{0 \leq n < T_{i_0}} 1_{\{i\}}(X_n) \right), \quad (1.6.1)$$

*où  $i_0$  est un point arbitraire de  $E$ , est invariante et vérifie  $0 < \mu_i < \infty$ . De plus, une mesure est sous-invariante si et seulement si elle est le produit de  $\mu$  par une constante arbitraire dans  $[0, \infty]$ .*

Cela montre en particulier que si  $\mu'$  est la mesure associée par (1.6.1) à un autre état  $i'_0$ , alors  $\mu'_i = c\mu_i$  pour tout  $i$ , pour une constante  $c \in ]0, \infty[$ . Cela montre aussi que toute mesure sous-invariante est invariante.

**Preuve.** a) Soit  $\eta$  une mesure sous-invariante. Pour tout  $n$  on a  $\eta P^n \leq \eta$ , donc  $\eta_i \geq \sum_j \eta_j p_{ji}^{(n)}$ . Pour tout couple  $(i, j)$  il existe  $n$  avec  $p_{ji}^{(n)} > 0$ , puisque la chaîne est irréductible: donc  $\eta_j > 0 \Rightarrow \eta_i > 0$ , et aussi  $\eta_i < \infty \Rightarrow \eta_j < \infty$ . Par suite, si  $\eta$  n'est ni identiquement nulle ni identiquement infinie (i.e.  $\eta_i = \infty$  pour tout  $i$ ), on a  $0 < \eta_i < \infty$  pour tout  $i$ .

b) Montrons maintenant l'unicité de la mesure sous-invariante (ni identiquement nulle ni identiquement infinie), à une constante multiplicative près. Posons

$$\hat{p}_{ij} = \frac{\eta_j p_{ji}}{\eta_i}, \quad \hat{p}_{ij}^{(n)} = \frac{\eta_j p_{ji}^{(n)}}{\eta_i}.$$

La matrice  $\hat{P} = (\hat{p}_{ij})$  est à éléments positifs et la somme de chaque ligne est  $\leq 1$  puisque  $\eta$  est sous-invariante. Considérons un point  $\Delta$  "extérieur" à  $E$ , et posons  $E_\Delta = E \cup \{\Delta\}$ . Ensuite, introduisons la matrice  $\bar{P} = (\bar{p}_{ij})_{i,j \in E_\Delta}$  suivante:

$$\bar{p}_{ij} = \begin{cases} \hat{p}_{ij} & \text{si } i, j \in E \\ 1 - \sum_{j \in E} \hat{p}_{ij} & \text{si } i \in E, j = \Delta \\ 0 & \text{si } i = \Delta, j \in E \\ 1 & \text{si } i = j = \Delta \end{cases} \quad (1.6.2)$$

Cette matrice est associée à une chaîne de Markov  $\bar{\mathbf{X}}$  à valeurs dans  $E_\Delta$ , et sa puissance  $n$ ème est donnée par (1.6.2) également, à condition de remplacer  $\hat{p}_{ij}$  par  $\hat{p}_{ij}^{(n)}$ . Le potentiel associé  $\bar{u}_{ij}$  vérifie  $\bar{u}_{ij} = \eta_j u_{ji} / \eta_i = \infty$  (car la chaîne initiale  $\mathbf{X}$  est irréductible récurrente) pour tous  $i, j \in E$ , ainsi que  $\bar{u}_{\Delta\Delta} = \infty$  et  $\bar{u}_{\Delta i} = 0$  si  $i \in E$ : en comparant au théorème 1.5.9, on en déduit que  $\bar{\mathbf{X}}$  admet deux classes récurrentes, à savoir  $E$  et  $\{\Delta\}$ : étant donné le corollaire 1.5.10, si on part d'un point de  $E$ , on ne sort jamais de  $E$ , ce qui signifie qu'en fait la somme des lignes de la matrice  $\hat{P}$  vaut 1 (et la construction de la chaîne  $\bar{\mathbf{X}}$  sur un espace étendu  $E_\Delta$  était inutile...)

Supposons maintenant que  $\nu$  soit une autre mesure sous-invariante pour  $P$ , encore non identiquement nulle, ni identiquement infinie. Si  $f(i) = \nu_i / \eta_i$  on a

$$\hat{P}f(i) = \sum_j \frac{\eta_j p_{ji}}{\eta_i} \frac{\nu_j}{\eta_j} = \frac{(\nu P)_i}{\eta_i} \leq f(i)$$

et  $f$  est surharmonique pour la chaîne récurrente irréductible associée à la transition  $\hat{P}$ . Donc  $f$  est constante par la proposition précédente, et  $\nu$  est un multiple de  $\eta$ .

c) Il nous reste à montrer que (1.6.1) définit une mesure invariante, automatiquement ni identiquement nulle, ni identiquement infinie, puisque par construction  $\mu_{i_0} = 1$ . On a

$$\begin{aligned} (\mu P)_i &= \mathbb{E}_{i_0} \left( \sum_{0 \leq n < T_{i_0}} p_{X_n, i} \right) = \sum_{n \geq 0} \mathbb{E}_{i_0} \left( p_{X_n, i} 1_{\{n < T_{i_0}\}} \right) \\ &= \sum_{n \geq 0} \mathbb{E}_{i_0} \left( 1_{\{i\}}(X_{n+1}) 1_{\{n < T_{i_0}\}} \right) \quad (\text{propriété de Markov}) \\ &= \mathbb{E}_{i_0} \left( \sum_{0 \leq n < T_{i_0}} 1_{\{i\}}(X_{n+1}) \right) \\ &= \mathbb{E}_{i_0} \left( \sum_{0 \leq n < T_{i_0}} 1_{\{i\}}(X_n) \right) = \mu_i, \end{aligned}$$

où l'avant-dernière égalité provient de ce que  $\sum_{0 \leq n < T_{i_0}} 1_{\{i\}}(X_{n+1}) = \sum_{0 \leq n < T_{i_0}} 1_{\{i\}}(X_n)$   $\mathbb{P}_{i_0}$ -p.s. (distinguer les cas  $i = i_0$  et  $i \neq i_0$ ), et la dernière est une nouvelle application de (1.6.1).  $\square$

**Corollaire 1.6.3** *Si  $\mathbf{X}$  est une chaîne irréductible récurrente et si  $\mu$  est une mesure invariante ni identiquement nulle ni identiquement infinie, on a deux cas possibles:*

- a)  $\mu(E) < \infty$  et  $\mathbb{E}_i(T_i) < \infty$  pour tout  $i \in E$ ,
- b)  $\mu(E) = \infty$  et  $\mathbb{E}_i(T_i) = \infty$  pour tout  $i \in E$ .

**Preuve.** La mesure  $\mu$  de (1.6.1) vérifie clairement  $\mu(E) = \mathbb{E}_{i_0}(T_{i_0})$ . A cause de "l'unicité", toute les mesures invariantes ni identiquement nulles ni identiquement infinies sont simultanément de masse totale finie (resp. infinie), et le résultat est évident.  $\square$

Ce corollaire montre que le temps moyen de retour dans l'état  $i$ , soit  $m_i = \mathbb{E}_i(T_i)$ , joue un rôle important. Cela nous amène à une définition:

**Définition 1.6.4** On dit que l'état  $i$  est *positif* si  $m_i < \infty$ , et est *nul* si  $m_i = \infty$ .

Cette terminologie curieuse vient du fait que c'est l'inverse  $\frac{1}{m_i}$  qui va jouer un rôle dans la suite: il est positif (resp. nul) si  $i$  est positif (resp. nul).

**Proposition 1.6.5** (1) *Tout état transient est nul.*

(2) *Les états d'une même classe sont, soit tout positifs, soit tous nuls. On dit alors que la classe est positive, resp. nulle.*

**Preuve.** Si  $i$  est transient on a  $\mathbb{P}_i(T_i = \infty) > 0$ , donc *a fortiori*  $m_i = \infty$ : cela donne (1), et pour (2), il suffit donc de considérer le cas d'une classe récurrente  $C$ . On a vu que dans ce cas, si on part d'un point de  $C$ , on ne sort jamais de  $C$ : si on se contente de considérer des mesures initiales portées par  $C$ , on peut donc restreindre l'espace d'état à  $C$  lui-même: on obtient ainsi à l'évidence une chaîne irréductible récurrente, et le résultat découle alors immédiatement du corollaire 1.6.3.  $\square$

Nous pouvons maintenant passer à un premier théorème ergodique.

**Théorème 1.6.6** *Si  $i$  est un état de période  $d_i$  et si  $m_i = \mathbb{E}_i(T_i)$  on a*

$$p_{ii}^{(nd)} \xrightarrow{n \rightarrow \infty} \begin{cases} \frac{d_i}{m_i} & \text{si } m_i < \infty \text{ (donc } d_i < \infty \text{ aussi)} \\ 0 & \text{si } m_i = \infty. \end{cases} \quad (1.6.3)$$

Comme  $p_{ii}^{(n)} = 0$  quand  $n$  n'est pas un multiple de  $d_i$ , on obtient ainsi le comportement asymptotique complet de la suite  $p_{ii}^{(n)}$ : elle tend vers 0 si  $m_i = \infty$ , et elle alterne entre 0 et une suite convergeant vers  $d_i/m_i$  si  $m_i < \infty$ .

**Preuve.** a) Si  $i$  est transient, on a  $u_{ii} < \infty$ , donc la suite  $p_{ii}^{(n)}$  tend vers 0 quand  $n \rightarrow \infty$ , tandis que  $m_i = \infty$  d'après ce qui précède: le résultat est donc démontré. Dans la suite on suppose donc  $i$  récurrent.

b) On pose  $d = d_i$ ,  $m = m_i$ , et

$$h_r = \mathbb{P}_i(T_i = r), \quad A = \{r > 0 : h_r > 0\}, \quad B = \{r > 0 : p_{ii}^{(r)} > 0\}.$$

On sait que  $d$  est le PGCD de  $B$ , qui est non vide (car  $i$  est récurrent) et on va montrer que  $d$  est aussi le PGCD de  $A$  (qui est aussi non vide).

Pour cela, notons  $d_N$  (resp.  $d'_N$ ) le PGCD de  $\{r : 1 \leq r \leq N, p_{ii}^{(r)} > 0\}$  (resp.  $\{r : 1 \leq r \leq N, h_r > 0\}$ ). La suite  $d_N$  décroît vers  $d$ , la suite  $d'_N$  décroît vers le PGCD de  $A$ , donc il suffit de montrer que  $d_N = d'_N$  pour tout  $N$ . Comme  $h_1 = p_{ii}$ , c'est évident pour  $N = 1$ . Supposons alors que  $d_N = d'_N$  pour un  $N \geq 1$  donné. D'après la propriété de Markov forte, on a

$$p_{ii}^{(r)} = \mathbb{P}_i(T_i \leq r, X_r = i) = \mathbb{E}_i \left( 1_{\{T_i \leq r\}} p_{ii}^{(r-T_i)} \right) = \sum_{s=1}^r h_s p_{ii}^{(r-s)}. \quad (1.6.4)$$

En particulier, on a

$$p_{ii}^{(N+1)} = h_{N+1} + \sum_{s=1}^N h_s p_{ii}^{(N+1-s)}.$$

Si alors  $h_{N+1} > 0$ , alors  $d_{N+1}$  et  $d'_{N+1}$  sont tous les deux le PGCD de  $\{d_N, N+1\}$ . Si ensuite  $h_{N+1} = p_{ii}^{(N+1)} = 0$ , on a  $d'_{N+1} = d_{N+1} = d_N$ . Si enfin  $h_{N+1} = 0 < p_{ii}^{(N+1)}$ , alors  $d'_{N+1} = d_N$  tandis que  $d_{N+1}$  est le PGCD de  $\{d_N, N+1\}$ ; mais d'après (1.6.4) il existe  $s \leq N$  avec  $h_s > 0$  et  $p_{ii}^{(N+1-s)} > 0$ , donc  $d_N$  divise  $s$  et aussi  $N+1-s$ , donc *a fortiori*  $N+1$ , et on a donc  $d_{N+1} = d_N$ : on a ainsi montré le résultat voulu.

c) D'après (b), on a  $h_r = p_{ii}^{(r)} = 0$  si  $r$  n'est pas un multiple de  $d$ . Donc si  $k_n = h_{nd}$  et  $p_n = p_{ii}^{(nd)}$ , la relation (1.6.4) s'écrit aussi

$$p_n = \sum_{r=1}^n k_r p_{n-r}. \quad (1.6.5)$$

Posons  $\lambda = \limsup_n p_n$ . Il existe une suite  $n_N$  croissant vers l'infini, telle que  $p_{n_N} \rightarrow \lambda$ . Donc (1.6.5) donne pour tout  $q$  fixé:

$$\begin{aligned} \lambda &= \lim_N \left( k_q p_{n_N-q} + \sum_{1 \leq n \leq n_N, n \neq q} k_n p_{n_N-n} \right) \\ &= k_q \liminf_N p_{n_N-q} + \limsup_N \sum_{1 \leq n \leq n_N, n \neq q} k_n p_{n_N-n} \\ &\leq k_q \liminf_N p_{n_N-q} + \lambda(1 - k_q) \end{aligned}$$

en appliquant le lemme de Fatou aux "fonctions"  $n \mapsto p_{n_N-n} 1_{\{n \neq q\}}$ , positives et bornées par 1, et à la "mesure"  $n \mapsto k_q$ , qui est une probabilité car  $f_{ii} = \sum_{r \geq 1} k_r = 1$  (car  $i$  est récurrent). Cela implique que



$$p_{n_N} \rightarrow \lambda \text{ et } k_q > 0 \implies p_{n_N - q} \rightarrow \lambda.$$

En itérant cette propriété, on voit facilement que

$$q = \alpha_1 q_1 + \dots + \alpha_l q_l, \quad \alpha_i \in \mathbb{N}, \quad k_{q_i} > 0 \implies p_{n_N - q} \rightarrow \lambda. \quad (1.6.6)$$

d) D'après (b),  $d$  est le PGCD de  $A$ , donc le PGCD de  $\{r : k_r > 0\}$  égale 1. Il existe donc un nombre fini  $s_1 < \dots < s_r$  d'entiers tels que  $k_{s_i} > 0$ , et que le PGCD de  $\{s_1, \dots, s_r\}$  soit 1. D'après l'identité de Bezout, cela entraîne l'existence d'entiers relatifs  $\beta_i \in \mathbb{Z}$  tels que  $\beta_1 s_1 + \dots + \beta_r s_r = 1$ . Si  $M = \sup_i |\beta_i|$  et  $s = s_1 + \dots + s_r$  et  $v_0 = Ms^2$ , tout  $v \geq v_0$  s'écrit  $v = sq + v'$  avec  $q, v' \in \mathbb{N}$  et  $v' \leq s - 1$ , et donc  $q \geq Ms$ . En utilisant  $\beta_1 s_1 + \dots + \beta_r s_r = 1$  on obtient alors  $v = \sum_{i=1}^r (q + v' \beta_i) s_i$ , et  $q + v' \beta_i \in \mathbb{N}$ , et (1.6.6) implique

$$v \geq v_0 \implies p_{n_N - v} \rightarrow \lambda. \quad (1.6.7)$$

e) Posons  $l_i = \sum_{n=i+1}^{\infty} k_n$ . Comme  $i$  est récurrent, on a  $l_0 = 1$ , et comme  $\mathbb{P}_i$ -p.s. la variable  $T_i/d$  ne prend que des valeurs entières on a

$$\frac{m}{d} = \mathbb{E}_i \left( \frac{T_i}{d} \right) = \sum_{n \geq 1} \mathbb{P}_i(T_i \geq nd) = \sum_{n=0}^{\infty} l_n. \quad (1.6.8)$$

Par ailleurs (1.6.5) donne  $p_n = \sum_{r=1}^n (l_{r-1} - l_r) p_{n-r}$ , d'où l'on déduit immédiatement

$$\sum_{r=0}^n l_r p_{n-r} = l_0 p_0 = 1 \quad \forall n \geq 0. \quad (1.6.9)$$

Nous distinguons maintenant deux cas:

(A)  $m < \infty$ : Par (1.6.8) on a  $\sum_r l_r < \infty$ , et  $p_n \leq 1$ , donc d'après le théorème de Lebesgue, en utilisant (1.6.7) et en écrivant (1.6.9) pour  $n = n_N - v_0$ , on obtient que  $\sum_{r \geq 0} l_r \lambda = 1$ , donc  $\lambda = \frac{d}{m}$  par (1.6.8).

(B)  $m = \infty$ : D'après le Lemme de Fatou, en utilisant (1.6.7) et en écrivant (1.6.9) pour  $n = n_N - v_0$ , on obtient que  $\sum_{r \geq 0} l_r \lambda \leq 1$ , donc  $\lambda = 0$  par (1.6.8).

Dans le cas (B), en se rappelant que  $\lambda = \limsup_n p_n$ , on voit que  $p_n \rightarrow 0$  et le résultat est démontré. Dans le cas (A), on pose  $\lambda' = \liminf_n p_n$ . On peut répéter la preuve précédente en remplaçant partout  $\lambda$  par  $\lambda'$  (il existe en effet une suite  $n'_N$  telle que  $p_{n'_N} \rightarrow \lambda'$ ) et en changeant les sens des inégalités. On obtient finalement que  $\lambda' = \frac{d}{m}$  également, donc  $\lambda = \lambda'$ , et la suite  $p_n$  converge vers  $\frac{d}{m}$ .  $\square$

Voici maintenant la description complète du comportement asymptotique des éléments de la matrice  $P^n$ :

**Corollaire 1.6.7** *On a les propriétés suivantes, avec  $m_i = \mathbb{E}_i(T_i)$  et  $d_i$  la période de  $i$ :*

(a) *Si  $j$  est un état nul (i.e.  $m_j = \infty$ ), alors  $p_{ij}^{(n)} \rightarrow 0$  pour tout  $i \in E$ .*

(b) *Si  $j$  est un état positif et si  $i \sim j$ , il existe  $r_{ij} \in \{0, \dots, d_j - 1\}$  tel que  $p_{ij}^{(nd_j + r_{ij})} \rightarrow \frac{d_j}{m_j}$ , tandis que  $p_{ij}^{(nd_j + r)} = 0$  pour tout  $r \in \{0, \dots, d_j - 1\}$  avec  $r \neq r_{ij}$ .*

(c) Si  $j$  est un état positif et si  $i$  n'est pas dans la même classe que  $j$ , pour tout  $r \in \{0, \dots, d_j - 1\}$  on a  $p_{ij}^{(nd_j+r)} \rightarrow f_{ij}(r) \frac{d_j}{m_j}$ , où  $f_{ij}(r) = \sum_{n \geq 0} \mathbb{P}_i(T_j = nd_j + r)$ .

**Preuve.** Si on pose  $h_{ij}^{(n)} = \mathbb{P}_i(T_j = n)$ , on a d'après la propriété de Markov forte

$$p_{ij}^{(n)} = \mathbb{P}_i(T_j \leq n, X_n = j) = \mathbb{E}_i \left( \mathbf{1}_{\{T_j \leq n\}} p_{jj}^{(n-T_i)} \right) = \sum_{s=1}^n h_{ij}^{(s)} p_{jj}^{(n-s)}. \quad (1.6.10)$$

Par ailleurs  $\sum_{s \geq 1} h_{ij}^{(s)} = f_{ij} \leq 1$ .

Si  $j$  est nul on a  $p_{jj}^{(n)} \rightarrow 0$ , donc le théorème de Lebesgue appliqué à (1.6.10) donne immédiatement le résultat (a). Dans le cas (b), d'après la proposition 1.5.6 il existe  $r_{ij} \in \{0, \dots, d_j - 1\}$  tel que  $p_{ij}^{(nd_j+r)} = 0$  et  $h_{ij}^{(nd_j+r)} = 0$  pour tout  $r' \in \{0, \dots, d_j - 1\}$  avec  $r' \neq r_{ij}$ . On a aussi  $p_{jj}^{(n)} = 0$  si  $n$  n'est pas un multiple de  $d_j$ , de sorte qu'on peut réécrire (1.6.10) ainsi:

$$p^{(nd_j+r_{ij})} = \sum_{s=1}^n h_{ij}^{(sd_j+r_{ij})} p_{jj}^{((n-s)d_j)},$$

et comme  $p_{jj}^{((n-s)d_j)} \rightarrow \frac{d_j}{m_j}$  un nouvelle application du théorème de Lebesgue donne le résultat.

Enfin dans le cas (c), on peut réécrire (1.6.10) ainsi:

$$p^{(nd_j+r)} = \sum_{s=1}^n h_{ij}^{(sd_j+r)} p_{jj}^{((n-s)d_j)},$$

pour tout  $r \in \{0, \dots, d_j - 1\}$ . Il suffit encore une fois d'appliquer le théorème de Lebesgue pour obtenir le résultat.  $\square$

Le comportement décrit ci-dessus est plutôt compliqué, à cause de la période. On obtient un résultat moins fort, mais bien plus simple, si on considère les moyennes de Césaro:

$$\Pi_N = \frac{1}{N} \sum_{n=1}^N P^n, \quad \text{de composantes} \quad \pi_{ij}^{(N)} = \frac{1}{N} \sum_{n=1}^N p_{ij}^{(n)}. \quad (1.6.11)$$

**Corollaire 1.6.8 (Théorème ergodique en moyenne):** Lorsque  $N \rightarrow \infty$ , la matrice  $\Pi_N$  converge (élément par élément) vers la matrice  $\Pi$  de composantes  $\pi_{ij} = \frac{f_{ij}}{m_j}$  (avec  $\pi_{ij} = 0$  si  $m_j = \infty$ ).

**Preuve.** Rappelons que si une suite  $(a_n)$  converge vers  $a$ , alors  $\frac{1}{N} \sum_{n=1}^N a_n$  converge *a fortiori* vers  $a$ . La convergence  $\pi_{ij}^{(n)} \rightarrow \pi_{ij}$  découle alors immédiatement du corollaire précédent (noter que dans le cas (c) de ce corollaire, on a  $f_{ij} = \sum_{r=0}^{d_j-1} f_{ij}(r)$ ).  $\square$

Pour terminer ce paragraphe, on va déduire des résultats précédents l'existence (ou non...) d'une ou de plusieurs probabilités invariantes pour la chaîne.

**Théorème 1.6.9** *Si  $\mathbf{X}$  est une chaîne irréductible, il existe une probabilité invariante si et seulement si la chaîne est positive. Dans ce cas la probabilité invariante est unique et donnée par  $\mu_i = \frac{1}{m_i}$ .*

**Preuve.** Si  $\mu$  est une probabilité invariante, on a  $\mu P^n = \mu$  pour tout  $n$ , donc aussi  $\mu \Pi_N = \mu$ . Par application du théorème de Lebesgue on obtient alors  $\mu \Pi = \mu$ : comme  $\pi_{ij} = \frac{1}{m_j}$ , on en déduit  $\mu_j = \frac{1}{m_j}$ . Cela implique en particulier qu'on n'a pas  $m_j = \infty$  pour tout  $j$ , donc la chaîne est positive (et  $m_j < \infty$  pour tout  $j$ , cf. proposition 1.6.5).

Inversement si la chaîne est positive, l'existence d'une probabilité invariante découle du corollaire 1.6.3.  $\square$

Le cas non irréductible est évidemment un peu plus compliqué. On note  $C_\alpha$  les *classes positives*, indicées par un ensemble  $A$ . La démonstration du résultat suivant, facile, est laissée au lecteur.

**Théorème 1.6.10** *Il existe au moins une probabilité invariante si et seulement si l'ensemble  $A$  est non vide. Dans ce cas, si on note  $\mu(\alpha) = (\mu(\alpha)_i)$ , avec  $\mu(\alpha)_i = \frac{1}{m_i}$  pour  $i \in C_\alpha$  et  $\mu(\alpha)_i = 0$  sinon, une mesure  $\mu$  est une probabilité invariante si et seulement si elle s'écrit*

$$\mu = \sum_{\alpha \in A} c_\alpha \mu(\alpha),$$

où les  $c_\alpha$  vérifient  $c_\alpha \geq 0$  et  $\sum_{\alpha \in A} c_\alpha = 1$ .

Ainsi, les  $\mu(\alpha)$  sont les probabilités invariantes *extrémales* dans l'ensemble de toutes les probabilités invariantes, qui est un ensemble convexe. La probabilité invariante est unique si et seulement s'il existe une classe positive et une seule. S'il n'y a pas de classe positive, il n'y a pas de probabilité invariante.

Enfin, on peut noter le résultat élémentaire suivant:

**Proposition 1.6.11** *Si l'espace  $E$  est fini, il existe au moins une classe positive (donc récurrente) et donc une probabilité invariante.*

**Preuve.** On a  $\sum_{j \in E} \pi_{ij}^{(N)} = 1$ , et comme  $E$  est fini on peut passer à la limite en  $N$  (cf. corollaire 1.6.8) et on a  $\sum_j \pi_{ij} = 1$ : il existe donc au moins un point  $j$  tel que  $\pi_{ij} > 0$ , donc  $m_j < \infty$ .  $\square$

Si  $E$  est infini, on peut n'avoir que des états transients: par exemple si  $E = \mathbb{N}$  et  $p_{ij} = 1$  si  $j = i + 1$  et  $p_{ij} = 0$  sinon. On peut aussi n'avoir que des états récurrents nuls: par exemple si  $E = \mathbb{Z}$  et  $p_{ij} = \frac{1}{2}$  si  $j = i - 1$  ou  $j = i + 1$ , et  $p_{ij} = 0$  sinon; dans ce cas la chaîne  $\mathbf{X}$  est une "marche aléatoire de Bernoulli", i.e.  $X_n = X_{n-1} + Y_n$  où les  $Y_n$  sont i.i.d. avec  $\mathbb{P}(Y_n = 1) = \mathbb{P}(Y_n = -1) = \frac{1}{2}$ ; il est clair que cette chaîne est irréductible, de période 2, et comme  $p_{00}^{(2n)} = C_{2n}^n \frac{1}{2^n}$  on vérifie que  $u_{00} = \infty$ , de sorte que la chaîne est récurrente; enfin la mesure  $\mu$  définie par  $\mu_i = 1$  pour tout  $i$  est clairement invariante, et comme elle est de masse totale infinie le corollaire 1.6.3 montre que la chaîne est nulle.

## Chapitre 2

# Processus de Markov de saut pur

### 2.1 Processus de Markov généraux

Dans ce chapitre on va aborder les processus à temps continu, indicés par  $\mathbb{R}_+$ . L'essentiel sera consacré aux processus de Markov dits "de saut pur", mais nous commençons par des considérations générales.

#### 2.1.1 Définitions

Soit  $(E, \mathcal{E})$  est espace mesurable quelconque. Exactement comme pour les chaînes de Markov, il y a plusieurs définitions pour les processus de Markov; mais nous allons d'emblée nous situer dans un cadre comparable à celui des définitions 1.3.1 et 1.3.2, dans le cas homogène.

D'abord, on suppose donnée une famille  $(P_t)_{t \geq 0}$  de probabilités de transition de  $(E, \mathcal{E})$  dans lui-même, avec  $P_0 = I$  (l'identité). Cette famille sera appelée un *semi-groupe* si elle vérifie

$$P_{t+s} = P_t P_s \quad \forall s, t \geq 0. \quad (2.1.1)$$

Ensuite, soit  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  un espace probabilisé filtré (les  $\mathcal{F}_t$  sont des sous-tribus de  $\mathcal{F}$  avec  $\mathcal{F}_t \subset \mathcal{F}_s$  si  $t \leq s$ ). Un *processus adapté* est une famille  $(X_t)_{t \geq 0}$  de variables aléatoires à valeurs dans  $(E, \mathcal{E})$ , telle que chaque  $X_t$  soit  $\mathcal{F}_t$ -mesurable. L'analogie (dans le cas homogène seulement) de la définition 1.3.1 est alors:

**Définition 2.1.1** Le processus  $(X_t)$  sur  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  est dit vérifier la  *$(\mathcal{F}_t)$ -propriété de Markov*, s'il existe un semi-groupe de transitions  $(P_t)$  tel que pour tous  $s, t \geq 0$  et toute fonction mesurable  $f$ , positive ou bornée,

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_t) = P_s f(X_t). \quad (2.1.2)$$

Si  $\mathcal{F}_t = \sigma(X_s : s \leq t)$ , on parle simplement de la *propriété de Markov*.

Pour l'analogie de la définition 1.3.2 on part de l'espace  $\Omega$  muni du processus  $(X_t)$ , on pose  $\mathcal{F}_t = \sigma(X_s : s \leq t)$  et  $\mathcal{F} = \bigvee_t \mathcal{F}_t$ , et on suppose donnée une famille  $(\theta_t)_{t \geq 0}$

d'applications de  $\Omega$  dans lui-même, telle que

$$\theta_0 = I, \quad \theta_{s+t} = \theta_s \circ \theta_t, \quad X_{s+t} = X_s \circ \theta_t, \quad \forall s, t \geq 0. \quad (2.1.3)$$

On se donne enfin une famille  $(\mathbb{P}_x)_{x \in E}$  de probabilités sur  $(\Omega, \mathcal{F})$  telle que  $x \mapsto \mathbb{P}_x(A)$  soit  $\mathcal{E}$ -mesurable pour tout  $A \in \mathcal{F}$ , et vérifiant  $\mathbb{P}_x(X_0 = x) = 1$  (cf. (1.3.4)). On définit  $\mathbb{P}_\mu$ , pour toute probabilité  $\mu$  sur  $(E, \mathcal{E})$ , par la formule (1.3.6):  $\mathbb{P}_\mu(A) = \int \mathbb{P}_x(A) \mu(dx)$ .

**Définition 2.1.2** Le terme  $\mathbf{X} = (\Omega, \mathcal{F}, (\mathcal{F}_t), (\theta_t), (X_t), (\mathbb{P}_x))$  est un *processus de Markov* s'il existe un semi-groupe de transitions  $(P_t)$  tel que pour tout  $x \in E$ , tous  $s, t \geq 0$  et toute fonction mesurable  $f$ , positive ou bornée,

$$\mathbb{E}_x(f(X_{t+s})|\mathcal{F}_t) = P_s f(X_t). \quad (2.1.4)$$

On a alors l'analogie de la proposition 1.3.3:

**Proposition 2.1.3** Si  $\mathbf{X}$  est un processus de Markov, pour toute probabilité  $\mu$  sur  $(E, \mathcal{E})$ , la loi de  $X_0$  sous  $\mathbb{P}_\mu$  (loi initiale) est  $\mu$ . De plus pour tout  $t \geq 0$  et toute variable aléatoire réelle  $Y$  sur  $(\Omega, \mathcal{F})$ , positive ou bornée, la variable  $Y \circ \theta_t$  est  $\mathcal{F}$ -mesurable et on a:

$$\mathbb{E}_\mu(Y \circ \theta_t | \mathcal{F}_t) = \mathbb{E}_{X_t}(Y). \quad (2.1.5)$$

**Preuve.** La première propriété se montre comme dans le cas discret. D'après le théorème des classes monotones 1.2.2, pour la seconde assertion il suffit de montrer le résultat quand  $Y$  est de la forme  $Y = \prod_{i=0}^m f_i(X_{t_i})$ , pour des fonctions bornées mesurables  $f_i$  et des temps  $0 = t_0 < \dots < t_m$ . Dans ce cas  $Y \circ \theta_t = \prod_{i=0}^m f_i(X_{t+t_i})$ , donc on voit que  $Y \circ \theta_t$  est  $\mathcal{F}$ -mesurable. Si  $Z$  est une variable bornée  $\mathcal{F}_t$ -mesurable, on voit (comme dans le cas discret encore) qu'il suffit de montrer que pour tout  $x \in E$ ,

$$\mathbb{E}_x(Z Y \circ \theta_t) = \mathbb{E}_x(Z \mathbb{E}_{X_t}(Y)). \quad (2.1.6)$$

Cela se montre par récurrence sur  $m$ . C'est évident pour  $m = 0$ . Supposons (2.1.6) vraie pour tout  $t$  et toute variable de la forme  $Y = \prod_{i=0}^{m-1} g_i(X_{s_i})$ , et montrons qu'elle est vraie pour  $Y = \prod_{i=0}^m f_i(X_{t_i})$ . Posons  $U = \prod_{i=0}^{m-1} f_i(X_{t_i})$  et  $V = U P_{t_m - t_{m-1}} f_m(X_{t_{m-1}})$ . On a

$$\begin{aligned} \mathbb{E}_x(Z Y \circ \theta_t) &= \mathbb{E}_x(Z U \circ \theta_t f_m(X_{t+t_m})) \\ &= \mathbb{E}_x(\mathbb{E}_x(Z U \circ \theta_t f_m(X_{t+t_m}) | \mathcal{F}_{t+t_{m-1}})) \\ &= \mathbb{E}_x(Z U \circ \theta_t P_{t_m - t_{m-1}} f_m(X_{t+t_{m-1}})) \\ &= \mathbb{E}_x(Z V \circ \theta_t) = \mathbb{E}_x(Z \mathbb{E}_{X_t}(V)), \end{aligned}$$

où la dernière égalité provient de l'hypothèse de récurrence. En appliquant ce qui précède à  $t = 0$  et  $Z = 1$ , on obtient aussi  $\mathbb{E}_y(Y) = \mathbb{E}_y(V)$  pour tout  $y$ , d'où (2.1.6).  $\square$

**Construction d'un processus de Markov:** Etant donné un semi-groupe  $(P_t)$  de transitions sur  $E$ , on peut se poser le problème de la construction d'un processus de Markov  $\mathbf{X}$  au sens de la définition 2.1.2 admettant ce semi-groupe pour transitions.

Comme dans le cas discret, il est naturel de considérer la construction canonique suivante: on prend  $\Omega^{\mathbb{R}_+}$ , qui est l'espace de toutes les fonctions  $t \mapsto x(t)$  sur  $\mathbb{R}_+$ , à valeurs dans  $E$ . Puis, pour toute fonction  $x(\cdot)$ , on pose

$$X_t(x(\cdot)) = x(t), \quad \theta_t(x)(\cdot) = x(t + \cdot),$$

de sorte qu'on a (2.1.3). Les tribus  $\mathcal{F}$  et  $\mathcal{F}_t$  sont comme avant la définition 2.1.2. Il reste à définir les probabilités  $\mathbb{P}_x$ : malheureusement le théorème de Ionescu–Tulcea n'est valide que pour une suite *dénombrable strictement ordonnée* de temps (comme  $\mathbb{N}$ ), et nullement pour un ensemble du type  $\mathbb{R}_+$  comme ensemble des temps. Il nous faut nous référer à un théorème d'une autre nature, appelé *théorème de Kolmogorov*, et qui nécessite une certaine structure de l'espace  $E$ . Voici (énoncée seulement) la version du théorème de Kolmogorov relative aux processus de Markov:

**Théorème 2.1.4** *Supposons que  $E$  soit un espace polonais (i.e. métrique, complet, séparable), et  $\mathcal{E}$  sa tribu borélienne. A tout semi-groupe  $(P_t)_{t \geq 0}$  de probabilités de transition de  $(E, \mathcal{E})$  dans lui-même on associe une famille  $\mathbb{P}_x$  de probabilités sur l'espace canonique  $(\Omega, \mathcal{F})$  ci-dessus, telle que  $\mathbf{X} = (\Omega, \mathcal{F}, (\mathcal{F}_t), (\theta_t), (X_t), (\mathbb{P}_x))$  soit un processus de Markov de transitions  $(P_t)$ .*

## 2.1.2 La propriété forte de Markov

Soit  $\mathbf{X}$  un processus de Markov, au sens de la définition 2.1.2. Rappelons qu'un temps d'arrêt, dans le cadre présent, est une application  $T$  de  $\Omega$  dans  $[0, \infty]$  telle que  $\{T < t\} \in \mathcal{F}_t$  pour tout  $t \in \mathbb{R}_+$  (**Attention:** on ne suppose pas ici que la filtration  $(\mathcal{F}_t)$  est "continue à droite", donc la définition des temps d'arrêt n'est pas tout-à-fait standard; la même remarque s'applique à ce qui suit et explique la notation  $\mathcal{F}_{T+}$ ). On lui associe sa "tribu antérieure":

$$\mathcal{F}_{T+} = \{A : A \in \mathcal{F}, A \cap \{T < t\} \in \mathcal{F}_t \ \forall t \in \mathbb{R}_+\}.$$

Ceci définit bien une tribu, et si  $T(\omega) = t$  pour tout  $\omega$  alors  $T$  est un temps d'arrêt et  $\mathcal{F}_{T+} = \mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s$ . Ensuite, sur l'ensemble  $\{T < \infty\}$ , on pose

$$\theta_T(\omega) = \theta_{T(\omega)}(\omega), \quad X_T(\omega) = X_{T(\omega)}(\omega).$$

On dira que le processus vérifie *la propriété forte de Markov* si, pour toute loi initiale  $\mu$ , tout temps d'arrêt  $T$  et toute variable aléatoire  $Y$  positive ou bornée, on a

$$\mathbb{E}_\mu(Y \circ \theta_T | \mathcal{F}_{T+}) = E_{X_T}(Y) \quad \text{sur l'ensemble } \{T < \infty\}. \quad (2.1.7)$$

Comme dans le cas discret, l'ensemble  $\{T < \infty\}$  est  $\mathcal{F}_{T+}$ -mesurable, et donc l'égalité ci-dessus ne fait en fait intervenir  $Y \circ \theta_T$  que sur cet ensemble. Ceci dit, (2.1.7) implique de manière implicite que  $X_T$  et  $Y \circ \theta_T$  sont, en restriction à  $\{T < \infty\}$ , mesurables respectivement par rapport à  $\mathcal{F}_{T+}$  et à  $\mathcal{F}$  (au moins à des ensembles  $\mathbb{P}_x$ -négligeables près pour tout  $x$ ). Ceci peut être faux, mais, même si c'est vrai, la propriété (2.1.7) *n'est pas toujours vraie*.

Il existe un certain nombre de conditions assurant cette propriété. Nous en explicitons une ci-dessous:

**Théorème 2.1.5** Soit  $\mathbf{X}$  un processus de Markov de transitions  $(P_t)$ . Si  $E$  est un espace polonais, la propriété de Markov forte (2.1.7) est valide (et en particulier pour tout  $x$ ,  $X_T$  est  $\mathbb{P}_x$ -p.s. égale à une variable  $\mathcal{F}_{T+}$ -mesurable, et  $Y \circ \theta_T$  est  $\mathbb{P}_x$ -p.s. égale à une variable  $\mathcal{F}$ -mesurable) dès qu'on a les deux propriétés suivantes:

- 1) pour toute fonction continue bornée  $f$  sur  $E$  les fonctions  $P_t f$  sont également continues (on dit que les  $P_t$  sont fellériens),
- 2) pour tout  $x$  et  $\mathbb{P}_x$ -presque tout  $\omega$ , la fonction  $t \mapsto X_t(\omega)$  est continue à droite.

**Preuve.** On peut supposer (par un argument de classe monotone, et puisque  $E$  est polonais) que  $Y = \prod_{i=0}^m f_i(X_{t_i})$  pour des  $f_i$  bornées et continues sur  $E$  et  $0 = t_0 < \dots < t_m$ .

a) Soit  $T$  un temps d'arrêt, et posons

$$T_n = \begin{cases} \frac{k+1}{2^n} & \text{si } \frac{k}{2^n} \leq T < \frac{k+1}{2^n}, \quad k \in \mathbb{N}, \\ \infty & \text{si } T = \infty. \end{cases}$$

On a  $B_{n,k} := \{T_n = k/2^n\} = \{(k-1)/2^n \leq T < k/2^n\}$ , donc  $B_{n,k} \in \mathcal{F}_{k/2^n}$  pour tout  $k \in \mathbb{N}^*$ . On pose aussi  $B = \{T < \infty\}$ , qui égale  $\cup_{k \geq 1} B_{n,k}$  pour tout  $n$ . Enfin, il est clair que  $T_n$  décroît vers  $T$ . Dans la suite "p.s." veut dire  $\mathbb{P}_x$ -p.s. pour tout  $x$ .

b) Montrons d'abord les propriétés de mesurabilité de  $X_T$  et  $Y \circ \theta_T$  sur  $B$ . D'abord, on a  $\{X_{T_n+t_i} \in A\} \cap B = \cup_{k \geq 1} (B_{n,k} \cap \{X_{k/2^n+t_i} \in A\})$ , qui est dans  $\mathcal{F}$  pour tout  $A \in \mathcal{E}$ : donc  $(Y \circ \theta_{T_n})1_B$  est  $\mathcal{F}$ -mesurable et comme  $Y \circ \theta_{T_n}1_B \rightarrow Y \circ \theta_T1_B$  p.s. (à cause de (2) et de la forme spéciale de  $Y$ ), on a la  $\mathcal{F}$ -mesurabilité p.s. de  $Y \circ \theta_T1_B$ .

Par ailleurs fixons  $t > 0$ . Posons  $V_n = X_{T_n}$  sur  $\{T_n < t\}$  et  $V_n = \Delta$  ailleurs (où  $\Delta$  est un point extérieur à  $E$ , isolé topologiquement dans  $E_\Delta = E \cup \{\Delta\}$ ), et de même  $V = X_T$  sur  $\{T < t\}$  et  $V = \Delta$  ailleurs. Comme  $\{V_n \in A\} = \cup_{1 \leq k < t2^n} (B_{n,k} \cap \{X_{k/2^n} \in A\}) \in \mathcal{F}_t$  pour tout  $A \in \mathcal{E}$ , les variables  $V_n$  sont  $\mathcal{F}_t$ -mesurable, donc  $V$ , qui par (2) et  $T_n \downarrow T$  est la limite p.s. des  $V_n$ , est p.s. égale à une variable  $\mathcal{F}_t$ -mesurable. Comme  $\{X_T \in A, T < t\} = \{V \in A\} \in \mathcal{F}_t$  pour  $A \in \mathcal{E}$ , on voit que  $X_T$  en restriction à  $B$  est p.s. égale à une variable  $\mathcal{F}_{T+}$ -mesurable.

c) Soit  $Z$  bornée et  $\mathcal{F}_{T+}$ -mesurable. La variable  $Z1_{B_{n,k}}$  est  $\mathcal{F}_{k/2^n}$ -mesurable, donc

$$\begin{aligned} \mathbb{E}_\mu(Z 1_B Y \circ \theta_{T_n}) &= \sum_{k=1}^{\infty} \mathbb{E}_\mu(Z 1_{B_{n,k}} Y \circ \theta_{k/2^n}) \\ &= \sum_{k=1}^{\infty} \mathbb{E}_\mu(Z 1_{B_{n,k}} E_{X_{k/2^n}}(Y)) \quad (\text{par (2.1.5)}) \\ &= \mathbb{E}_\mu(Z 1_B \mathbb{E}_{X_{T_n}}(Y)), \end{aligned} \tag{2.1.8}$$

Quand  $n \rightarrow \infty$ , le membre de gauche tend vers  $\mathbb{E}_\mu(Z1_B Y \circ \theta_T)$  à cause de (2) et de la forme spéciale de  $Y$ . Par ailleurs le même raisonnement que dans la proposition 2.1.3 montre que  $\mathbb{E}_x(Y) = g_0(x)$ , où on définit les  $g_i$  par récurrence descendante en partant de  $g_m = f_m$  et en posant  $g_i(x) = f_i(x)P_{t_{i+1}-t_i}g_{i+1}(x)$ . Etant donné (1) et la continuité des  $f_i$ , on voit que les  $g_i$  sont également continues et bornées: donc le membre de droite de (2.1.8) converge vers  $\mathbb{E}_\mu(Z1_B E_{X_t}(Y))$ , ce qui achève la preuve.  $\square$

**Corollaire 2.1.6 (Loi 0–1)** *Si on a la propriété de Markov forte, alors pour tout  $x \in E$  et tout  $A \in \mathcal{F}_{0+}$  le nombre  $\mathbb{P}_x(A)$  vaut 0 ou 1.*

**Preuve.** Soit  $A \in \mathcal{F}_{0+}$  et  $f(x) = \mathbb{P}_x(A)$ . Comme  $\theta_0$  est l'identité, en appliquant (2.1.7) à  $T \equiv 0$  on obtient

$$f(x) = \mathbb{E}_x(1_A \circ \theta_0) = \mathbb{E}_x(1_A \mathbb{E}_x(1_A \circ \theta_0 | \mathcal{F}_{0+})) = \mathbb{E}_x(1_A f(X_0)) = f(x)^2,$$

puisque  $\mathbb{P}_x(X_0 = x) = 1$ . □

La condition (1) du théorème 2.1.5 est agréable, puisque qu'elle s'exprime directement en fonction du semi-groupe  $(P_t)$ . Il n'en est pas de même de (2), qui est évidemment fort difficile à vérifier dans le cas général. Nous énonçons ci-dessous (sans la démonstration – difficile) un critère impliquant (2).

**Théorème 2.1.7** *Soit  $(P_t)$  un semi-groupe de probabilités de transition sur l'espace  $(E, \mathcal{E})$ , espace localement compact de type dénombrable muni de ses boréliens. On peut construire un processus de Markov  $\mathbf{X}$  de transitions  $(P_t)$ , dont les trajectoires  $t \mapsto X_t(\omega)$  sont continues à droite, dès que le semi-groupe est fortement fellérien, ce qui signifie qu'il vérifie les deux propriétés suivantes:*

(a) *pour toute fonction continue nulle à l'infini (pour le compactifié d'Alexandrov de  $E$ )  $f$ , les fonctions  $P_t f$  sont également continues nulles à l'infini,*

(b) *pour toute fonction comme ci-dessus,  $P_t f$  converge simplement vers  $f$  quand  $t \rightarrow 0$ .*

On remarquera les conditions sur  $E$ : pour les théorèmes 2.1.4 et 2.1.5 on a besoin de  $E$  polonais, ce qui est le cas si  $E$  est un borélien de  $E = \mathbb{R}^d$  ou de  $\mathbb{R}^N$  (pour la topologie produit). Pour le dernier résultat, la condition est pratiquement que  $E$  est un compact séparable, ou est un ouvert de  $\mathbb{R}^d$  pour  $d$  fini. Noter aussi, à l'inverse, que si  $(X_t)$  est à trajectoires continues à droite, la condition (b) ci-dessus est nécessairement satisfaite.

### 2.1.3 Stationnarité

Soit  $\mathbf{X}$  un processus de Markov à valeurs dans  $(E, \mathcal{E})$ . Comme dans le cas discret, on peut se poser la question de savoir s'il existe une ou plusieurs lois initiales  $\mathbb{P}_\mu$  rendant le processus stationnaire. Rappelons que  $(X_t)$  est *stationnaire* sous la probabilité  $\mathbb{P}_\mu$  si pour tous  $t_1, \dots, t_m$  la loi des variables  $(X_{t+t_1}, \dots, X_{t+t_m})$  sous  $\mathbb{P}_\mu$  ne dépend pas de  $t \geq 0$ .

De même que dans la définition 1.4.3, on dit qu'une mesure  $\eta$  sur  $(E, \mathcal{E})$  est *invariante*, resp. *sous-invariante*, si  $\eta P_t = \eta$ , resp.  $\eta P_t \leq \eta$ , pour tout  $t \geq 0$ .

**Proposition 2.1.8** *Le processus  $(X_t)$  est stationnaire sous la probabilité  $\mathbb{P}_\mu$  si et seulement si la probabilité  $\mu$  est invariante.*

**Preuve.** On reproduit essentiellement la preuve de la proposition 1.4.6. Si le processus est stationnaire sous  $\mathbb{P}_\mu$ , la loi  $\mu P_t$  de  $X_t$  égale la loi  $\mu$  de  $X_0$  pour tout  $t \geq 0$ , donc  $\mu$  est invariante.



Inversement supposons  $\mu$  invariante. Soit  $m \geq 0$  et  $0 = t_0 < \dots < t_m$  et  $A \in \mathcal{E}^{\otimes(m+1)}$ . On pose  $Y = 1_A(X_{t_0}, \dots, X_{t_m})$ . La loi  $\eta_t$  de  $(X_{t+t_0}, \dots, X_{t+t_m})$  est caractérisée par les nombres  $\eta_t(A) = \mathbb{E}_\mu(Y \circ \theta_t)$ . D'après (2.1.5), et si  $f(x) = \mathbb{E}_x(Y)$ , on a alors

$$\eta_t(A) = \mathbb{E}_\mu(f(X_t)) = \mu P_t f,$$

qui vaut  $\mu(f) = \eta_0(A)$  à cause de l'invariance de  $\mu$ .  $\square$

## 2.2 Processus de saut pur: premières propriétés

On va considérer à partir de maintenant la même situation qu'au chapitre 1, c'est-à-dire le cas où l'espace d'état  $E$  est fini ou dénombrable, avec pour  $\mathcal{E}$  la tribu de toutes les parties. Les points de  $E$  seront encore notés  $i, j, k$ . On a vu ci-dessus que les propriétés topologiques de  $E$  jouent un rôle important, et nous munissons ici  $E$  de la *topologie discrète*, pour laquelle tous les points sont isolés.

Ceci dit, au lieu de partir du semi-groupe de transitions, on va inverser la problématique: on part du processus  $(X_t)$ , à valeurs dans  $E$ , donné sur un espace  $\Omega$ . On supposera qu'il est du type suivant:

**Définition 2.2.1** Un processus  $(X_t)$  à valeurs dans un espace fini ou dénombrable  $E$  est dit *de saut pur* si toutes ses trajectoires  $t \mapsto X_t(\omega)$  sont continues à droite.

Plus précisément, pour rester dans le cadre précédent, on a un processus de Markov  $\mathbf{X} = (\Omega, \mathcal{F}, (\mathcal{F}_t), (\theta_t), (X_t), (P_t))$  de saut pur.

Comme dans le chapitre 1, une mesure est un "vecteur ligne"  $\mu = (\mu_i)$ , et la probabilité de transition  $P_t$  est une matrice  $P_t = (p(t)_{ij})_{i,j \in E}$ . Bien entendu  $p(0)_{ij} = \delta_{ij}$ , et comme  $(X_t)$  est à trajectoires continues à droite on a

$$t \mapsto p(t)_{ij} \quad \text{est continue à droite} \tag{2.2.1}$$

(car la fonctions  $1_{\{j\}}$ , comme toutes les fonctions sur  $E$ , sont continues). On verra en fait bien mieux plus tard.

**Proposition 2.2.2** *Tout processus de Markov de saut pur  $\mathbf{X}$  vérifie la propriété de Markov forte.*

**Preuve.** L'espace  $E$  muni de la topologie discrète est polonais, donc il suffit de vérifier les deux conditions du théorème 2.1.5: la première est satisfaite car toute fonction sur  $E$  est continue, et la seconde l'est par hypothèse.  $\square$

On va maintenant étudier de plus près la structure de  $\mathbf{X}$ . D'abord, comme la topologie de  $E$  est la topologie discrète, la continuité à droite entraîne que les trajectoires sont "constantes par morceaux", au sens où pour tout  $t \geq 0$  et tout  $\omega$  on a  $X_s(\omega) = X_t(\omega)$

pour tout  $s$  dans un intervalle  $[t, t'(\omega, t)[$ , où  $t'(\omega, t)$  est un réel strictement plus grand que  $t$ . En particulier, si on définit par récurrence les variables:

$$T_0 = 0, \quad T_{n+1} = \inf(t : t > T_n, X_t \neq X_{T_n}) \quad (2.2.2)$$

(comme d'habitude l'inf d'un ensemble vide vaut  $+\infty$ ), on obtient une suite croissante de temps d'arrêt vérifiant  $T_{n+1} > T_n$  sur  $\{T_n < \infty\}$ . On pose aussi  $T_\infty = \lim_n T_n$ , qu'on appelle le *premier temps d'explosion*. Il se peut bien sûr que  $T_\infty < \infty$ , auquel cas  $X_t$  prend une certaine valeur  $i$  en  $t = T_\infty$  et y reste un temps positif, puis "saute" de nouveau, etc... Si  $\mathbb{P}_i(T_\infty = \infty) = 1$  pour tout  $i \in E$ , on dit que le processus *est sans explosion*.

En utilisant les translations  $(\theta_t)$ , on remarque aussi que

$$T_n - T_{n-1} = T_1 \circ \theta_{T_{n-1}} \quad \text{sur l'ensemble } \{T_{n-1} < \infty\}. \quad (2.2.3)$$

Le fait que  $X_{T_n}$  n'est bien défini que si  $T_n < \infty$  nous incite à considérer un point  $\Delta$  extérieur" à  $E$  et à poser  $E_\Delta = E \cup \{\Delta\}$ , comme on l'a déjà fait plusieurs fois. On peut alors poser pour  $n \geq 0$

$$\left. \begin{array}{lll} S_n = T_{n+1} - T_n, & \xi_n = X_{T_n} & \text{si } T_n < \infty \\ S_n = \infty & \xi_n = \Delta & \text{si } T_n = \infty. \end{array} \right\} \quad (2.2.4)$$

Le théorème de structure suivant est fondamental (rappelons que la loi exponentielle de paramètre  $a \in ]0, \infty[$  est la loi sur  $\mathbb{R}_+$  de densité  $x \mapsto ae^{-ax}$ ; par convention, si  $a = 0$  c'est la masse de Dirac  $\varepsilon_\infty$  en l'infini. Une variable exponentielle de paramètre  $a$  admet l'espérance  $1/a$ , avec la convention  $1/0 = \infty$ ):

**Théorème 2.2.3** *Soit  $\mathbf{X}$  un processus de Markov de saut pur.*

(1) *Sous  $\mathbb{P}_i$ , les variables  $S_0$  et  $\xi_1$  sont indépendantes, et  $S_0$  suit une loi exponentielle dont on note  $a_i \in \mathbb{R}_+$  le paramètre; si on pose  $\pi_{ij} = \mathbb{P}_i(\xi_1 = j)$  pour  $j \in E_\Delta$  on a aussi*

$$\left. \begin{array}{ll} \pi_{ii} = 0, & \begin{array}{l} a_i > 0 \implies \sum_{j \in E} \pi_{ij} = 1, \quad \pi_{i\Delta} = 0 \\ a_i = 0 \implies \sum_{j \in E} \pi_{ij} = 0, \quad \pi_{i\Delta} = 1 \end{array} \end{array} \right\} \quad (2.2.5)$$

pour tout  $i \in E$ . On pose aussi par convention  $a_\Delta = 0$ , et  $\pi_{\Delta\Delta} = 1$  et  $\pi_{\Delta i} = 0$  pour  $i \in E$  (ce qui est cohérent avec (1.2.10)).

(2) *Sous chaque  $\mathbb{P}_\mu$  la chaîne  $(\xi_n)$  à valeurs dans  $E_\Delta$  a la propriété de Markov relativement à la filtration discrète  $(\mathcal{F}_{T_n})_{n \geq 0}$ , avec la transition  $\Pi = (\pi_{ij})$ .*

(3) *Sous chaque  $\mathbb{P}_\mu$ , conditionnellement à  $\mathcal{F}_{T_n}$ , les variables  $S_n$  et  $\xi_{n+1}$  sont indépendantes, la première est de loi exponentielle de paramètre  $a_{\xi_n}$  et la seconde de loi  $(\pi_{\xi_n, j})_{j \in E_\Delta}$ .*

**Preuve.** Noter que (1) est un cas particulier de (3), mais nous allons commencer par prouver (1).

En observant que  $S_0 = t + S_0 \circ \theta_t$  et  $\xi_1 = \xi_1 \circ \theta_t$  et  $X_t = X_0$  sur l'ensemble  $\{S_0 > t\}$ , on obtient pour tous  $s, t \geq 0$ ,  $i \in E$  et  $A \subset E_\Delta$ , par la propriété de Markov en  $t$ :

$$\begin{aligned} \mathbb{P}_i(S_0 > t + s, \xi_1 \in A) &= \mathbb{P}_i(S_0 > t, S_0 \circ \theta_t > s, \xi_1 \circ \theta_t \in A) \\ &= \mathbb{E}_i(1_{\{S_0 > t\}} \mathbb{P}_{X_t}(S_0 > s, \xi_1 \in A)) \\ &= \mathbb{P}_i(S_0 > t) \mathbb{P}_i(S_0 > s, \xi_1 \in A). \end{aligned}$$

En prenant  $A = E_\Delta$ , cela donne  $\mathbb{P}_i(S_0 > t + s) = \mathbb{P}_i(S_0 > t)\mathbb{P}_i(S_0 > s)$ , relation caractéristique des lois exponentielles de paramètre dans  $[0, \infty]$ ; comme  $S_0 > 0$ , le paramètre  $a_i$  de la loi de  $S_0$  sous  $\mathbb{P}_i$  est dans  $\mathbb{R}_+$ . En prenant  $s = 0$  ci-dessus, on obtient aussi  $\mathbb{P}_i(S_0 > t, \xi_1 \in A) = \mathbb{P}_i(S_0 > t)\mathbb{P}_i(\xi_1 \in A)$ : cela entraîne l'indépendance de  $S_0$  et  $\xi_1$  sous  $\mathbb{P}_i$ . Enfin si  $a_i > 0$  on a  $S_0 < \infty$   $\mathbb{P}_i$ -p.s., donc  $\xi_1 = X_{S_0} \in E$   $\mathbb{P}_i$ -p.s., tandis que si  $a_i = 0$  on a  $\mathbb{P}_i(S_0 = \infty) = 1$ , donc  $\mathbb{P}_i(\xi_1 = \Delta) = 1$  et on en déduit (2.2.5).

Sur l'ensemble  $\{T_n < \infty\}$  on a  $S_n = S_0 \circ \theta_{T_n}$  et  $\xi_{n+1} = \xi_1 \circ \theta_{T_n}$ , de sorte que d'après la propriété forte de Markov au tempes  $T_n$  on obtient pour tout borélien  $B$  de  $\mathbb{R}_+ \times E_\Delta$ :

$$\mathbb{P}_\mu((S_n, \xi_{n+1}) \in B | \mathcal{F}_{T_n}) = \mathbb{P}_{X_{T_n}}((S_0, \xi_1) \in B).$$

Donc conditionnellement à  $\mathcal{F}_{T_n}$  et sur l'ensemble  $\mathcal{F}_{T_n}$ -mesurable  $\{T_n < \infty\}$ , on a (3). Sur l'ensemble  $\{T_n = \infty\}$  on a  $S_n = \infty$  et  $\xi_n = \xi_{n+1} = \Delta$ , donc comme  $a_\Delta = 0$  et  $\pi_{\Delta\Delta} = 1$  on a aussi (3) sur cet ensemble. En particulier pour  $A \subset E_\Delta$  il vient  $\mathbb{P}_\mu(\xi_{n+1} \in A | \mathcal{F}_{T_n}) = \sum_{j \in A} \pi_{\xi_n, j}$ , et on a donc (2) également.  $\square$

**Corollaire 2.2.4** *Soit  $\mathbf{X}$  un processus de Markov de saut pur. Conditionnellement à la tribu  $\mathcal{G} = \sigma(\xi_n : n \geq 0)$ , et sous chaque probabilité  $\mathbb{P}_\mu$ , les variables  $(S_n)_{n \geq 0}$  sont indépendantes, chaque  $S_n$  étant de loi exponentielle de paramètre  $a_{\xi_n}$ .*

**Preuve.** Notons  $\eta_i$  la loi exponentielle de paramètre  $a_i$  (rappelons que  $a_\Delta = 0$ ). Il suffit de montrer que pour tout entier  $N$ , tous boréliens  $A_n$  de  $[0, \infty]$  et tous  $i_n \in E_\Delta$ , on a

$$\mathbb{P}_\mu\left(\left(\bigcap_{n=0}^N \{S_n \in A_n\}\right) \cap \left(\bigcap_{n=0}^{N+1} \{\xi_n = i_n\}\right)\right) = \mathbb{E}_\mu\left(\prod_{n=0}^{N+1} 1_{\{i_n\}}(\xi_n) \prod_{n=0}^N \eta_{i_n}(A_n)\right). \quad (2.2.6)$$

Comme  $\mathbb{P}_\mu(S_n \in A_n, \xi_{n+1} = i_{n+1} | \mathcal{F}_{T_n}) = \eta_{\xi_n}(A_n) \pi_{\xi_n, i_{n+1}}$  par le théorème précédent, on vérifie alors immédiatement (2.2.6) par récurrence descendante sur  $N$ .  $\square$

**Le processus “minimal”:** Dans la suite, on va essayer de reconstruire le processus de Markov à partir des suites  $(\xi_n)$  et  $(S_n)$ . En effet, comme  $X_t = X_{T_n}$  sur chaque intervalle  $[T_n, T_{n+1}[$ , on voit d'après (2.2.4) que

$$X_t = \xi_n \quad \text{si } T_n = S_0 + \dots + S_{n-1} \leq t < S_0 + \dots + S_n = T_{n+1}, \quad (2.2.7)$$

Ainsi, les deux suites  $(\xi_n)$  et  $(S_n)$  caractérisent entièrement  $X_t$  sur l'ensemble  $\{t < T_\infty\}$ . On est donc naturellement conduit à introduire un nouveau processus  $(\tilde{X}_t)$  à valeurs dans  $E_\Delta$ , en complétant (2.2.7):

$$\tilde{X}_t = \begin{cases} \xi_n & \text{si } S_0 + \dots + S_{n-1} \leq t < S_0 + \dots + S_n \\ \Delta & \text{si } t \geq S_0 + \dots + S_n + \dots \end{cases} \quad (2.2.8)$$

De manière équivalente, on a aussi:

$$\tilde{X}_t = \begin{cases} X_t & \text{si } t < T_\infty \\ \Delta & \text{si } t \geq T_\infty. \end{cases} \quad (2.2.9)$$

**Proposition 2.2.5** Soit  $\mathbf{X}$  un processus de Markov de saut pur, et  $\tilde{X}_t$  défini ci-dessus. Pour tout  $i \in E$  le processus  $(\tilde{X}_t)$  sur  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P}_i)$  vérifie la  $(\mathcal{F}_t)$ -propriété de Markov avec le semi-groupe de transition  $(\tilde{P}_t = (\tilde{p}(t)_{ij}))$  défini par

$$\tilde{p}(t)_{ij} = \begin{cases} \mathbb{P}_i(X_t = j, t < T_\infty) & \text{si } i, j \in E \\ \mathbb{P}_i(t \geq T_\infty) & \text{si } i \in E, j = \Delta \\ 1 & \text{si } i, j = \Delta \\ 0 & \text{si } i = \Delta, j \in E. \end{cases} \quad (2.2.10)$$

De plus, les variables  $\xi_n$  et  $S_n$  associées à  $\tilde{X}$  sont les mêmes que celles associées à  $X$ .

Le processus  $(\tilde{X}_t)$  s'appelle le *processus minimal* associé à  $X$ .

**Preuve.** La dernière assertion est triviale. Pour la propriété de semi-groupe de  $(\tilde{P}_t)$  et la propriété de Markov de  $\tilde{X}$ , il nous suffit de montrer que si  $k \in E$  et  $s, t \geq 0$  on a :

$$\mathbb{P}_k(\tilde{X}_{s+t} = j | \mathcal{F}_s) = \tilde{p}(t)_{\tilde{X}_s, j}. \quad (2.2.11)$$

Sur l'ensemble  $\mathcal{F}_s$ -mesurable  $\{s \geq T_\infty\}$  on a  $\tilde{X}_s = \tilde{X}_{s+t} = \Delta$ , de sorte que (2.2.11) est évident. Sur le complémentaire  $\{s < T_\infty\}$  on a  $\tilde{X}_{s+t} = \tilde{X}_t \circ \theta_s$ , et aussi  $\tilde{X}_s = X_s$ , donc la proposition 2.1.3 entraîne que le membre de gauche de (2.2.11) égale  $\mathbb{P}_{\tilde{X}_s}(\tilde{X}_t = j)$ , qui vaut bien  $\tilde{p}(t)_{\tilde{X}_s, j}$  lorsque  $j \in E$  et aussi lorsque  $j = \Delta$ .  $\square$

## 2.3 Construction d'un processus de saut pur

Pour construire un processus de Markov de saut pur on peut appliquer le théorème 2.1.4, mais cela nécessite d'une part de connaître le semi-groupe, d'autre part de vérifier que le processus de Markov ainsi construit est effectivement de saut pur, ce qui n'est pas évident du tout. On peut aussi utiliser le théorème 2.2.3 et son corollaire, qui d'une certaine manière fournissent de façon explicite la "dynamique" du processus, ou au moins celle du processus minimal associé. C'est ce que nous allons faire maintenant.

On part d'une famille  $(a_i)_{i \in E_\Delta}$  de réels positifs ou nuls, avec  $a_\Delta = 0$ , et d'une probabilité de transition  $\Pi = (p_{ij})_{i, j \in E_\Delta}$  vérifiant (2.2.5) (et donc  $\pi_{\Delta\Delta} = 1$ ).

Intuitivement, la démarche consiste à construire une chaîne de Markov  $(\xi_n)$  de transition  $\Pi$ , et une suite de variables  $(S_n)$  dont la loi, conditionnellement à  $(\xi_n)_{n \geq 0}$ , est celle décrite au corollaire 2.2.4, puis à définir  $\tilde{X}$  par (2.2.8). Formellement, les choses sont évidemment un peu plus compliquées !

On considère d'abord la chaîne de Markov "canonique"  $\Xi = (\Omega', \mathcal{F}, (\xi_n), \theta', (\mathbb{P}'_i)_{i \in E_\Delta})$  de transition  $\Pi$  (cf. la fin du paragraphe 1.3). Il est évident que  $\mathbb{P}'_i$ -p.s. on a  $\xi_{n+m} = \Delta$  pour tout  $m \geq 0$  si  $\xi_n = \Delta$ , et aussi que  $\mathbb{P}'_\Delta(\xi_0 = \Delta) = 1$ . Ensuite on pose  $\Omega'' = ]0, \infty]^{\mathbb{N}}$ , qu'on munit des variables  $S_n(s_0, s_1, \dots) = s_n$  et de la tribu produit  $\mathcal{F}'' = \sigma(S_n : n \geq 0)$ . Puis, pour toute suite  $\omega' = (x_0, x_1, \dots)$  dans  $E_\Delta$  on note  $\mathbb{P}''(\omega', d\omega'')$  l'unique probabilité sur  $(\Omega'', \mathcal{F}'')$  sous laquelle les  $(S_n)$  sont indépendantes, chaque  $S_n$  étant de loi exponentielle

de paramètre  $a_{x_n}$ . Par un argument de classe monotone on vérifie immédiatement que  $\mathbb{P}''$  est une probabilité de transition de  $(\Omega', \mathcal{F}')$  dans  $(\Omega'', \mathcal{F}'')$ . D'après le théorème 1.2.4, pour tout  $i \in E_\Delta$  on peut munir l'espace produit

$$\Omega = \Omega' \times \Omega'', \quad \mathcal{F}''' = \mathcal{F}' \otimes \mathcal{F}''$$

d'une unique probabilité  $\mathbb{P}_i$  telle que  $\mathbb{P}_i(A \times B) = \int \mathbb{P}'_i(d\omega') \mathbb{P}''(\omega', B) 1_A(\omega')$  pour tous  $A \in \mathcal{F}'$  et  $B \in \mathcal{F}''$ , et sous  $\mathbb{P}_i$  on a les deux propriétés suivantes:

1. la suite  $(\xi_n)$  a la propriété de Markov avec la transition  $\Pi$ ;
2. conditionnellement lorsque la suite  $(\xi_n)$  est fixée, les variables  $(S_n)_{n \geq 0}$  sont indépendantes, et  $S_n$  est de loi exponentielle de paramètre  $a_{\xi_n}$ ;

(ci-dessus,  $\xi_n$  ou  $S_n$ , définis respectivement sur  $\Omega'$  et  $\Omega''$ , sont considérés également comme des fonctions sur  $\Omega$ ).

Il nous reste alors à définir le processus  $(\tilde{X}_t)$  sur  $\Omega$  par la formule (2.2.8), la filtration  $\mathcal{F}_t = \sigma(\tilde{X}_s : s \leq t)$ , la tribu  $\mathcal{F} = \bigvee_t \mathcal{F}_t$ , et le semi-groupe des translations  $(\theta_t)$  comme étant les uniques applications de  $\Omega$  dans lui-même vérifiant identiquement (2.1.3) (avec  $\tilde{X}$  au lieu de  $X$ , bien-sûr; l'existence des  $\theta_t$  est très facile à montrer). On a à l'évidence la troisième propriété:

3. si  $\tilde{X}_t(\omega) = \Delta$  pour un  $t$ , on a aussi  $\tilde{X}_s(\omega) = \Delta$  pour tout  $s \geq t$ .

**Théorème 2.3.1** *Le terme  $\tilde{\mathbf{X}} = (\Omega, \mathcal{F}, (\mathcal{F}_t), (\theta_t), (\tilde{X}_t), (\mathbb{P}_i)_{i \in E_\Delta})$  est un processus de Markov de saut pur, qui est un processus minimal, et qui est associé aux données  $\Pi = (\pi_{ij})$  et  $(a_i)$  par le théorème 2.2.3.*

**Preuve.** La seule chose non évidente est la propriété de Markov (2.1.4) du terme  $\tilde{\mathbf{X}}$ . Pour cela, il suffit de vérifier (2.1.5) lorsque  $\mu = \varepsilon_i$  pour  $i \in E_\Delta$ : en effet, en posant  $P_t(i, A) = \mathbb{P}_i(\tilde{X}_t \in A)$  on aura (2.1.4), et le fait que  $(P_t)$  soit un semi-groupe découle immédiatement d'une itération de la propriété (2.1.5).

On pose  $T_0 = 0$  et  $T_n = S_0 + \dots + S_{n-1}$  pour  $n \geq 1$ , et  $T_\infty = \lim_n T_n$ . Remarquons que toute variable  $\mathcal{F}$ -mesurable  $Y$  vérifie  $Y(\omega) = \alpha$  pour une constante  $\alpha$  sur l'ensemble  $\{\omega : \tilde{X}_0(\omega) = \Delta\}$ , en vertu de la propriété 3 ci-dessus. Donc sur l'ensemble  $\mathcal{F}_t$ -mesurable  $\{T_\infty \leq t\}$  on a  $Y \circ \theta_t = \alpha$ , et aussi  $\tilde{X}_t = \Delta$ , de sorte que (2.1.5) est satisfaite sur cet ensemble.

Il nous reste à montrer la propriété suivante: considérons les ensembles  $\mathcal{F}_t$ -mesurables  $A(n, t) = \{T_n \leq t < T_{n+1}\}$ ; si alors  $Z$  est une variable  $\mathcal{F}_t$ -mesurable bornée et  $Y$  est  $\mathcal{F}$ -mesurable et bornée, alors

$$\mathbb{E}_i(Z Y \circ \theta_t 1_{A(n,t)}) = \mathbb{E}_i \left( Z 1_{A(n,t)} \mathbb{E}_{\tilde{X}_t}(Y) \right), \quad (2.3.1)$$

et, par un argument de classe monotone, il suffit même de montrer ceci lorsque  $Y = f(S_0, \dots, S_m; \xi_0, \dots, \xi_m)$  pour une fonction mesurable bornée  $f$ . Le résultat est intuitivement "évident" par application de la propriété de non-veillessement de la loi exponentielle, mais formellement c'est un peu compliqué...

En restriction à l'ensemble  $A(n, t)$  chaque variable  $\tilde{X}_s$  pour  $s \leq t$  est une fonction de  $(S_0, \dots, S_{n-1}; \xi_0, \dots, \xi_n)$ , donc il est de même de  $Z$ , de sorte qu'on peut écrire  $Z 1_{A(n, t)} = g(S_0, \dots, S_{n-1}; \xi_0, \dots, \xi_n) 1_{A(n, t)}$  pour une fonction mesurable bornée  $g$ .

En utilisant les propriétés 1 et 2 ci-dessus, il est facile de voir que

$$\mathbb{E}_{j_0}(Y) = \sum_{j_1, \dots, j_m \in E_\Delta} \left( \prod_{k=0}^{m-1} \pi_{j_k, j_{k+1}} \right) \int_{]0, \infty]^{m+1}} \left( \prod_{k=0}^m a_{j_k} e^{-a_{j_k} r_k} \right) f(r_0, \dots, r_m; j_0, \dots, j_m) dr_0 \dots dr_m,$$

avec la convention que la mesure  $a_i e^{-a_i r} dr$  est  $\varepsilon_\infty$  si  $a_i = 0$ . Il s'ensuit que le membre de droite de (2.3.1) vaut, avec la convention  $i_0 = i$ :

$$\begin{aligned} & \sum_{i_1, \dots, i_n \in E} \left( \prod_{k=0}^{n-1} \pi_{i_k, i_{k+1}} \right) \int_{]0, \infty]^{n+1}} \left( \prod_{k=0}^n a_{i_k} e^{-a_{i_k} s_k} \right) g(s_0, \dots, s_{n-1}; i_0, \dots, i_n) \\ & \quad 1_{\{s_0 + \dots + s_{n-1} \leq t < s_0 + \dots + s_n\}} ds_0 \dots ds_n \sum_{i_{n+1}, \dots, i_{n+m} \in E_\Delta} \left( \prod_{l=n}^{n+m-1} \pi_{i_l, i_{l+1}} \right) \\ & \quad \int_{]0, \infty]^{m+1}} \left( \prod_{l=n}^{n+m} a_{i_l} e^{-a_{i_l} r_l} \right) f(r_n, \dots, r_{n+m}; i_n, \dots, i_{n+m}) dr_n \dots dr_{n+m}. \end{aligned}$$

Mais l'intégrale de  $a_{i_n} e^{-a_{i_n} s_n}$  par rapport à  $ds_n$ , sur l'ensemble  $\{s_0 + \dots + s_{n-1} \leq t < s_0 + \dots + s_n\}$ , vaut  $\exp(-a_{i_n}(t - (s_0 + \dots + s_{n-1}))) 1_{\{t \geq s_0 + \dots + s_{n-1}\}}$ , de sorte que l'expression précédente égale

$$\begin{aligned} & \sum_{i_1, \dots, i_n \in E; i_{n+1}, \dots, i_{n+m} \in E_\Delta} \left( \prod_{k=0}^{n+m-1} \pi_{i_k, i_{k+1}} \right) \int_{]0, \infty]^{n+1}} \left( \prod_{k=0}^{n-1} a_{i_k} e^{-a_{i_k} s_k} \right) \\ & \quad g(s_0, \dots, s_{n-1}; i_0, \dots, i_n) e^{-a_{i_n}(t - (s_0 + \dots + s_{n-1}))} 1_{\{s_0 + \dots + s_{n-1} \leq t\}} ds_0 \dots ds_{n-1} \\ & \quad \int_{]0, \infty]^{m+1}} \left( \prod_{l=n}^{n+m} a_{i_l} e^{-a_{i_l} r_l} \right) f(r_n, \dots, r_{n+m}; i_n, \dots, i_{n+m}) dr_n \dots dr_{n+m}. \end{aligned}$$

Dans la dernière intégrale on fait les changement de variables  $r_k = s_k$  si  $k \geq n+1$  et  $r_n = s_0 + \dots + s_n - t$  (pour  $s_0, \dots, s_{n-1}$  fixés), de jacobien 1, et l'expression précédente devient

$$\begin{aligned} & \sum_{i_1, \dots, i_n \in E; i_{n+1}, \dots, i_{n+m} \in E_\Delta} \left( \prod_{k=0}^{n+m-1} \pi_{i_k, i_{k+1}} \right) \int_{]0, \infty]^{n+m+1}} \left( \prod_{k=0}^{n+m} a_{i_k} e^{-a_{i_k} s_k} \right) \\ & \quad g(s_0, \dots, s_{n-1}; i_0, \dots, i_n) 1_{\{s_0 + \dots + s_{n-1} \leq t < s_0 + \dots + s_n\}} \\ & \quad f(s_0 + \dots + s_n - t, s_{n+1}, \dots, s_{n+m}; i_n, \dots, i_{n+m}) ds_0 \dots ds_{n+m}. \end{aligned}$$

Comme sur l'ensemble  $A(n, t)$  on a  $(S_k, \xi_k) \circ \theta_t = (S_{k+n}, \xi_{n+k})$  pour  $k \geq 1$  et  $(S_0, \xi_0) \circ \theta_t = (S_0 + \dots + S_n - t, \xi_n)$ , on voit que cette expression égale le premier membre de (2.3.1), et la preuve est terminée.  $\square$

**Condition de “non-explosion”:** Comme il est naturel, la donnée de  $\Pi$  et des  $a_i$  permet de reconstruire le processus jusqu’au temps  $T_\infty$  (temps d’explosion), tandis qu’après cet instant on le pose égal à  $\Delta$  de manière arbitraire. Si on veut un processus à valeurs dans  $E$  lui-même, il est nécessaire (et possible, de diverses manières) de “prolonger” le processus  $\tilde{X}$  au delà de  $T_\infty$ , ce que nous ne ferons pas ici.

Un cas important est celui où il n’y a pas d’explosion, ce qui veut dire que  $\mathbb{P}_i(T_\infty = \infty) = 1$  pour tout  $i \in E$ : en effet dans ce cas le processus de Markov  $\mathbf{X}$  est son propre processus minimal, et le processus  $\tilde{X}$  ci-dessus ne prend  $\mathbb{P}_i$ -p.s. que des valeurs dans  $E$ , lorsque  $i \in E$ . Il est intéressant de noter qu’on a un critère de non-explosion en termes uniquement de  $\Pi$  et des  $a_i$ .

A cet effet, pour tout  $\lambda > 0$  on note  $Q(\lambda) = (q(\lambda)_{ij})_{i,j \in E}$  la matrice donnée par (avec la convention  $e^{-\infty} = 0$ ):

$$q(\lambda)_{ij} = \mathbb{E}_i \left( e^{-\lambda S_0} 1_{\{j\}}(\xi_1) \right) = \begin{cases} \frac{a_i}{a_i + \lambda} \pi_{ij} & \text{si } a_i > 0 \\ 0 & \text{si } a_i = 0 \end{cases} \quad (2.3.2)$$

La seconde égalité ci-dessus provient du théorème 2.2.3-(1) et du fait que la transformée de Laplace d’une variable exponentielle  $S$  de paramètre  $a > 0$  est  $\mathbb{E}(e^{-\lambda S}) = \frac{a}{a + \lambda}$ .

**Lemme 2.3.2** *Les éléments de la puissance nième  $Q(\lambda)^n$  de la matrice  $Q(\lambda)$  sont donnés par  $q(\lambda)_{ij}^{(n)} = \mathbb{E}_i \left( e^{-\lambda T_n} 1_{\{j\}}(\xi_n) \right)$ .*

**Preuve.** Comme  $S_0 = T_1$  c’est évident pour  $n = 1$ . Supposons la propriété vraie pour  $n$ . On a  $T_{n+1} = T_n + S_0 \circ \theta_{T_n}$  et  $\xi_{n+1} = \xi_1 \circ \theta_{T_n}$  et  $X_{T_n} = \xi_n$  si  $T_n < \infty$ , donc d’après la propriété de Markov forte et la convention  $e^{-\infty} = 0$ :

$$\begin{aligned} \mathbb{E}_i \left( e^{-\lambda T_{n+1}} 1_{\{j\}}(\xi_{n+1}) \right) &= \mathbb{E}_i \left( e^{-\lambda T_n} 1_{\{T_n < \infty\}} \left( e^{-\lambda S_0} 1_{\{j\}}(\xi_1) \right) \circ \theta_{T_n} \right) \\ &= \mathbb{E}_i \left( e^{-\lambda T_n} 1_{\{T_n < \infty\}} q(\lambda)_{X_{T_n}, j} \right) \\ &= \sum_{k \in E} \mathbb{E}_i \left( e^{-\lambda T_n} 1_{\{k\}}(X_{T_n}) \right) q(\lambda)_{kj} \\ &= \sum_{k \in E} q(\lambda)_{ik}^{(n)} q(\lambda)_{kj} = q(\lambda)_{ij}^{(n+1)}. \end{aligned}$$

On obtient donc le résultat par récurrence sur  $n$ . □

**Théorème 2.3.3** *Soit  $\mathbf{X}$  un processus de Markov de saut pur, et  $\lambda > 0$  arbitraire.*

(1) *Il y a équivalence entre  $\mathbb{P}_i(T_\infty = \infty) = 1$  et  $\lim_{n \rightarrow \infty} \sum_{j \in E} q(\lambda)_{ij}^{(n)} = 0$ .*

(2) *Il y a équivalence entre la propriété  $\mathbb{P}_i(T_\infty = \infty) = 1$  pour tout  $i \in E$  (i.e.,  $\mathbf{X}$  est sans explosion), et le fait que  $Q(\lambda)f = f$  n’admette que  $f \equiv 0$  comme solution positive bornée.*

**Preuve.** D’après le lemme on a  $Q(\lambda)^n 1(i) = \sum_{j \in E} q(\lambda)_{ij}^{(n)} = \mathbb{E}_i(e^{-\lambda T_n})$ , qui décroît vers  $g(i) = \mathbb{E}_i(e^{-\lambda T_\infty})$ , d’où (1).

Soit  $f$  positive bornée, avec  $Q(\lambda)f = f$ , et  $b = \sup_i f(i)$ . D'une part,  $f = Q(\lambda)^n f \leq bQ(\lambda)^n 1$  pour tout  $n$ , donc  $f \leq bg$ , de sorte que l'implication  $\Rightarrow$  de (2) découle de (1). On a aussi  $0 \leq g \leq 1$ , et  $Q(\lambda)g = \lim_n Q(\lambda)^{n+1} g = g$ , d'où l'implication  $\Leftarrow$  de (2).  $\square$

**Corollaire 2.3.4** *Si  $\sup_{i \in E} a_i < \infty$ , le processus  $\mathbf{X}$  est sans explosion.*

**Preuve.** Soit  $a = \sup_i a_i$ . On a  $\frac{a_i}{a_i + \lambda} \leq \frac{a}{a + \lambda}$ , donc  $Q(\lambda)1 \leq \frac{a}{a + \lambda} 1$  et par suite  $Q(\lambda)^n 1 \leq \left(\frac{a}{a + \lambda}\right)^n \rightarrow 0$ .  $\square$

Terminons par un critère de nature assez différente. On dit qu'un point  $i$  est *absorbant* si  $a_i = 0$  (en effet, dans ce cas on a  $\mathbb{P}_i(X_t = i \ \forall t) = 1$ ).

**Proposition 2.3.5** *Si tous les points de  $E$  sont soit absorbants, soit récurrents pour la chaîne de Markov  $(\xi_n)$  de transition  $\Pi$ , le processus  $\mathbf{X}$  est sans explosion.*

**Preuve.** Soit  $i \in E$ . Si  $i$  est absorbant on a  $\mathbb{P}_i(S_0 = \infty) = 1$ , donc *a fortiori*  $\mathbb{P}_i(T_\infty = \infty) = 1$ . Supposons donc  $i$  non-absorbant, mais récurrent pour la chaîne  $(\xi_n)$ , et notons  $N_k$  le  $k$ ième passage de la chaîne  $(\xi_n)$  en  $i$ , i.e.  $N_1 = \inf(l \geq 1 : \xi_l = i)$  et  $N_{k+1} = \inf(l \geq N_k + 1 : \xi_l = i)$ . L'hypothèse implique  $\mathbb{P}_i(N_k < \infty) = 1$  pour tout  $k$ , et le corollaire 2.2.4 implique que les  $S_{N_k}$  sont des variables indépendantes de loi exponentielle de paramètre  $a_i > 0$ , sous  $\mathbb{P}_i$ . On a bien-sûr  $T_\infty \geq \sum_{k \geq 1} S_{N_k}$ , et cette somme est  $\mathbb{P}_i$ -p.s. infinie d'après la loi des grands nombres (par exemple): d'où le résultat.  $\square$

## 2.4 Les équations de Kolmogorov

On vient de voir qu'on peut reconstruire un processus de Markov de saut pur à partir des quantités  $\Pi$  et  $a_i$ , au moins lorsqu'il n'y a pas d'explosion. On doit donc pouvoir aussi calculer le semi-groupe  $P_t$  à partir des mêmes données, ce qui est l'objet de ce paragraphe. Ci-dessous on part donc d'un processus  $\mathbf{X}$  auquel sont associés  $\Pi = (\pi_{ij})$  et les  $a_i \geq 0$ .

Rappelons que  $p(t)_{ij} = \mathbb{P}_i(X_t = j)$ . Sous  $\mathbb{P}_i$ , il n'y a eu aucun saut entre 0 et  $t$  avec la probabilité  $e^{-a_i t}$ , qui est approximativement égale à  $1 - a_i t$  lorsque  $t$  est "petit"; il y a eu exactement un saut entre 0 et  $t$ , arrivant en  $j \neq i$ , avec la probabilité  $\int_0^t a_i e^{-a_i s} \pi_{ij} e^{-a_j(t-s)} ds$ , approximativement égale à  $a_i \pi_{ij} t$  pour  $t$  petit, et la probabilité pour qu'il y ait eu au moins 2 sauts est de l'ordre de  $t^2$ . En se rappelant que  $p(0)_{ii} = 1$  et  $p(0)_{ij} = 0$  pour  $j \neq i$ , on s'attend donc à ce que  $t \mapsto p(t)_{ij}$  soit dérivable (à droite) en 0, avec la dérivée suivante:

$$p'(0)_{ij} = \begin{cases} -a_i & \text{si } j = i \\ a_i \pi_{ij} & \text{si } j \neq i. \end{cases} \quad (2.4.1)$$

On a en fait bien mieux:



**Théorème 2.4.1** Les fonctions  $t \mapsto p(t)_{ij}$  sont dérivables en tout point de  $\mathbb{R}_+$ , et vérifient les systèmes d'équations différentielles et intégrales suivants:

$$p'(t)_{ij} = -a_i p(t)_{ij} + a_i \sum_k \pi_{ik} p(t)_{kj}, \quad (2.4.2)$$

$$p(t)_{ij} = \delta_{ij} e^{-a_i t} + \sum_k \pi_{ik} \int_0^t a_i e^{-a_i s} p(t-s)_{kj} ds. \quad (2.4.3)$$

Ces équations s'appellent les *premières équations de Kolmogorov*.

**Preuve.** Nous aurons besoin d'une légère extension de la propriété (2.1.7) de Markov forte, qui est satisfaite par  $\mathbf{X}$  (cf. proposition 2.2.2). Soit  $T$  un temps d'arrêt, et  $V$  une fonction bornée sur  $\Omega \times \Omega$ , qui est  $\mathcal{F}_{T+} \otimes \mathcal{F}$ -mesurable. On pose  $V'(\omega) = V(\omega, \theta_T(\omega))$ , qui est évidemment  $\mathcal{F}$ -mesurable. On a alors pour toute loi initiale  $\mu$ :

$$\text{une version de } \mathbb{E}_\mu(V' | \mathcal{F}_{T+}) \text{ sur } \{T < \infty\} \text{ est } \omega \mapsto \int \mathbb{P}_{X_T(\omega)}(d\omega') V(\omega, \omega'). \quad (2.4.4)$$

Cette formule est évidente lorsque  $V$  a la forme produit  $V(\omega, \omega') = U(\omega)U'(\omega')$ , avec  $U$  et  $U'$  respectivement  $\mathcal{F}_{T+}$  et  $\mathcal{F}$ -mesurables, car  $V'(\omega) = U(\omega) U' \circ \theta_T(\omega)$ , et un argument de classe monotone montre qu'elle est vraie pour toute fonction  $\mathcal{F}_{T+} \otimes \mathcal{F}$ -mesurable  $V$ .

On applique alors (2.4.4) à la fonction  $V(\omega, \omega') = 1_{\{j\}}(X_{t-S_0(\omega)}(\omega')) 1_{\{S_0(\omega) \leq t\}}$  et au temps d'arrêt  $T_1 = S_0$ , ce qui donne

$$\begin{aligned} p(t)_{ij} &= \delta_{ij} \mathbb{P}_i(S_0 > t) + \mathbb{E}_i(V') = \delta_{ij} \mathbb{P}_i(S_0 > t) + \mathbb{E}_i \left( 1_{\{S_0 \leq t\}} \mathbb{E}_i(V' | \mathcal{F}_{S_0+}) \right) \\ &= \delta_{ij} \mathbb{P}_i(S_0 > t) + \mathbb{E}_i \left( 1_{\{S_0 \leq t\}} p(t - S_0)_{X_{S_0}, j} \right). \end{aligned}$$

Il suffit alors d'appliquer le théorème 2.2.3-(1) pour obtenir (2.4.3).

Le changement de variable  $t - s = u$  dans (2.4.3) donne

$$p(t)_{ij} = \delta_{ij} e^{-a_i t} + e^{-a_i t} \sum_k \pi_{ik} \int_0^t a_i e^{a_i u} p(u)_{kj} du. \quad (2.4.5)$$

On en déduit (utiliser le théorème de convergence dominée et le fait que  $\sum_k \pi_{ik} \leq 1$ ) que les fonctions  $t \mapsto p(t)_{ij}$  sont continues, puis (par dérivation des intégrales d'intégrands continus par rapport à la borne supérieure) que leurs dérivées satisfont (2.4.2).  $\square$

Noter que (2.4.1) se déduit de (2.4.2), qu'on peut d'ailleurs réécrire sous la forme plus compacte suivante:

$$P'_t = P'_0 P_t. \quad (2.4.6)$$

Une autre manière de se rappeler (2.4.6) consiste à écrire  $\frac{P_{t+s} - P_t}{s} = \frac{P_s - P_0}{s} P_t$  (par la propriété de semi-groupe), et à faire tendre  $s$  vers 0. On peut aussi écrire  $\frac{P_{t+s} - P_t}{s} = P_t \frac{P_s - P_0}{s}$ , de sorte qu'on s'attend aussi à avoir la propriété suivante:

$$P'_t = P_t P'_0. \quad (2.4.7)$$

Ceci revient à écrire

$$p'(t)_{ij} = -p(t)_{ij} a_j + \sum_k p(t)_{ik} a_k \pi_{kj}, \quad (2.4.8)$$

ou, sous forme intégrale:

$$p(t)_{ij} = \delta_{ij} e^{-a_j t} + \sum_k \int_0^t p(s)_{ik} a_k \pi_{kj} e^{-a_j(t-s)} ds \quad (2.4.9)$$

(comparer à (2.4.5)). Bien entendu, (2.4.8) découle de (2.4.9) par dérivation, et ces formules sont connues sous le nom de *secondes équations de Kolmogorov*.

En fait, les secondes équations de Kolmogorov ne sont pas toujours vraies, et on a seulement le résultat suivant:

**Théorème 2.4.2** *Si le processus de Markov de saut pur  $\mathbf{X}$  est sans explosion (i.e.  $\mathbb{P}_i(T_\infty = \infty) = 1$  pour tout  $i$ ), alors on a (2.4.9).*

**Preuve.** Supposons qu'on ait montré l'égalité (2.4.9) pour presque tout  $t$ . La fonction  $\alpha(s) = \sum_k p(s)_{ik} a_k \pi_{kj} e^{a_j s}$  est positive et vérifie  $\int_0^t \alpha(s) ds < \infty$  pour presque tout  $t$ , donc pour tout  $t$ , et par suite  $t \mapsto \int_0^t \alpha(s) ds$  est continue, ce qui entraîne que (2.4.9) a lieu pour tout  $t$ . Il suffit donc de montrer que les transformées de Laplace des deux membres de (2.4.9), en tant que fonctions de  $t$ , sont identiques.

Si  $\lambda > 0$ , en utilisant Fubini, le lemme 2.3.2 et la propriété de Markov forte en chaque  $T_n$ , et comme il n'y a pas d'explosion,

$$\begin{aligned} (\lambda + a_j) \int_0^\infty e^{-\lambda t} p(t)_{ij} dt &= (\lambda + a_j) \mathbb{E}_i \left( \sum_{n \geq 0} 1_{\{j\}}(\xi_n) \int_{T_n}^{T_{n+1}} e^{-\lambda t} dt \right) \\ &= (\lambda + a_j) \sum_{n \geq 0} \mathbb{E}_i \left( 1_{\{j\}}(\xi_n) e^{-\lambda T_n} \int_0^{S_0 \circ T_n} e^{-\lambda s} ds \right) \\ &= (\lambda + a_j) \sum_{n \geq 0} \mathbb{E}_i \left( 1_{\{j\}}(\xi_n) e^{-\lambda T_n} \int_0^\infty a_j e^{-a_j u} du \int_0^\infty e^{-\lambda s} ds \right) \\ &= \sum_{n \geq 0} \mathbb{E}_i \left( 1_{\{j\}}(\xi_n) e^{-\lambda T_n} \right) = \sum_{n \geq 0} q(\lambda)_{ij}^{(n)} \end{aligned} \quad (2.4.10)$$

La transformée de Laplace du membre de droite de (2.4.9), multipliée par  $\lambda + a_j$ , s'écrit

$$\begin{aligned} &(\lambda + a_j) \int_0^\infty e^{-\lambda t} dt \left( \delta_{ij} e^{-a_j t} + \sum_k \int_0^t p(s)_{ik} a_k \pi_{kj} e^{-a_j(t-s)} ds \right) \\ &= \delta_{ij} + (\lambda + a_j) \sum_k a_k \pi_{kj} \int_0^\infty e^{a_j s} p(s)_{ik} ds \int_s^\infty e^{-(\lambda + a_j)t} dt \\ &= \delta_{ij} + \sum_k a_k \pi_{kj} \int_0^\infty e^{-\lambda s} p(s)_{ik} ds = \delta_{ij} + \sum_k \sum_{n \geq 0} q(\lambda)_{ik}^{(n)} \frac{a_k \pi_{kj}}{\lambda + a_k} \end{aligned}$$

en vertu de (2.4.10) appliqué à  $p(s)_{ik}$ . D'après (2.3.2) cette expression vaut

$$\delta_{ij} + \sum_{n \geq 0} q(\lambda)_{ij}^{(n+1)} = \sum_{n \geq 0} q(\lambda)_{ij}^{(n)}.$$

Donc les transformées de Laplace des deux membres de (2.4.9), multipliées par  $\lambda + a_j$ , sont identiques, et par suite (2.4.9) est vraie.

Passons à la preuve de (2.4.8). On pose  $g(t)_{ij} = \sum_k p(t)_{ik} a_k \pi_{kj}$ , de sorte qu'on peut réécrire (2.4.9) ainsi:

$$p(t)_{ij} = e^{-a_j t} \left( \delta_{ij} + \int_0^t g(s)_{ij} e^{a_j s} ds \right). \quad (2.4.11)$$

Par ailleurs, en appliquant (2.4.5) et Fubini, on voit que si  $h(t)_{ij} = \sum_k \pi_{ik} g(t)_{kj}$ ,

$$g(t)_{ij} = a_i e^{-a_i t} \left( \pi_{ij} + \int_0^t e^{a_i u} h(u)_{ij} du \right). \quad (2.4.12)$$

D'abord, (2.4.11) implique que  $s \mapsto g(s)_{ij}$  est intégrable par rapport à la mesure de Lebesgue sur tout compact, donc fini pour presque tout  $s$ ; par suite (2.4.12) montre que  $u \mapsto h(u)_{ij}$  est aussi intégrable sur chaque intervalle  $[0, t]$  tel que  $g(t)_{ij} < \infty$ , donc en fait sur tout compact; par suite (2.4.12) encore montre qu'en fait  $t \mapsto g(t)_{ij}$  est continue (donc finie partout): il suffit alors de dériver (2.4.11) pour obtenir (2.4.8).  $\square$

**Propriétés de martingale.** Rappelons qu'un processus  $M = (M_t)_{t \geq 0}$  à trajectoires continues à droite sur un espace filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  est une *martingale* si pour tous  $s, t \geq 0$  la variable  $M_t$  est  $\mathcal{F}_t$ -mesurable et intégrable et  $\mathbb{E}(M_{t+s} | \mathcal{F}_t) = M_t$ .

On affaiblit cette notion en disant que  $M$  est une *martingale généralisée* si, au lieu d'imposer l'intégrabilité de chaque variable  $M_t$ , on suppose simplement que les espérances conditionnelles  $\mathbb{E}(M_{t+s} | \mathcal{F}_t)$  existent et égalent  $M_t$  pour tout  $s \geq 0$ : par exemple (ce sera la situation étudiée ci-dessous) si  $M_t = M'_t - M''_t$  avec  $M'_t \geq 0$  et  $M''_t \geq 0$  et  $M'_t$  intégrable pour tout  $t$ ; on définit alors  $\mathbb{E}(M_{t+s} | \mathcal{F}_t)$  comme la différence  $\mathbb{E}(M'_{t+s} | \mathcal{F}_t) - \mathbb{E}(M''_{t+s} | \mathcal{F}_t)$ , qui est *a priori* à valeurs dans  $[-\infty, \infty]$ ; dans ce cas,  $M$  sera une martingale généralisée si  $\mathbb{E}(M_{t+s} | \mathcal{F}_t) = M_t$ .

**Proposition 2.4.3** *Si  $\mathbf{X}$  est sans explosion, pour tout  $j$  le processus*

$$M_t = 1_{\{j\}}(X_t) - 1_{\{j\}}(X_0) - \int_0^t (a_{X_u} \pi_{X_u, j}) du \quad (2.4.13)$$

*est une martingale généralisée pour chaque probabilité  $\mathbb{P}_\mu$ .*

**Preuve.** On peut écrire  $M_t = M'_t - M''_t$ , avec  $M'_t = 1_{\{j\}}(X_t)$ : on a  $0 \leq M'_t \leq 1$  et  $M''_t \geq 0$ , et évidemment  $M'_t$  et  $M''_t$  sont  $\mathcal{F}_t$ -mesurables. Il reste à montrer que  $\mathbb{E}_\mu(M_{t+s} | \mathcal{F}_t) = M_t$ . On a par la propriété de Markov et le théorème de Fubini pour les espérances conditionnelles:

$$\begin{aligned} \mathbb{E}_\mu(M_{t+s} | \mathcal{F}_t) &= \mathbb{P}_\mu(X_{t+s} = j | \mathcal{F}_t) - 1_{\{j\}}(X_0) - \int_0^t (a_{X_u} \pi_{X_u, j}) du \\ &\quad - \mathbb{E}_\mu \left( \int_t^{t+s} (a_{X_u} \pi_{X_u, j}) du \right) | \mathcal{F}_t \\ &= M_t + \mathbb{P}_\mu(X_{t+s} = j | \mathcal{F}_t) - 1_{\{j\}}(X_t) - \int_t^{t+s} \mathbb{E}_\mu(a_{X_u} \pi_{X_u, j} | \mathcal{F}_t) du \\ &= M_t + p(s)_{X_t, j} - 1_{\{j\}}(X_t) - \int_0^s \left( \sum_k p(v)_{X_t, k} a_k \pi_{kj} \right) dv. \end{aligned}$$

Il reste alors à appliquer (2.4.9) pour obtenir que ceci vaut  $M_t$ .  $\square$

On peut montrer une “réciproque”, très intéressante mais au delà des objectifs de ce cours: soit un processus  $X$  adapté sur un espace filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ , dont les trajectoires sont à valeurs dans  $E$ , continues à droite et sans explosion, et une probabilité de transition  $\Pi$  et des nombres  $a_i \geq 0$  vérifiant les conditions (2.2.5). Si alors pour tout  $j \in E$  le processus  $M$  associé à  $X$  par (2.4.13) est une martingale généralisée sur l’espace  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ , alors  $X$  satisfait la propriété de Markov avec un semi-groupe vérifiant les (première et seconde) équations de Kolmogorov avec  $\Pi$  et les  $a_i$ .

Plus généralement, à toute fonction bornée  $g$  on peut associer le processus

$$M_t^g = g(X_t) - g(X_0) - \int_0^t (P'_0 g)(X_s) ds$$

(qui égale (2.4.13) quand  $g = 1_{\{j\}}$ ). Alors,  $M^g$  est une martingale sous chaque  $\mathbb{P}_\mu$ .

## 2.5 Classification et Stationnarité

On considère toujours un processus de Markov de saut pur  $\mathbf{X}$ . De plus, et quoique cette hypothèse ne soit pas toujours nécessaire, on le supposera *sans explosion*.

### 2.5.1 Première classification

Exactement comme pour les chaînes de Markov, on peut diviser l’espace d’état  $E$  en “classes”, à condition de bien définir les temps de passage dans les états. On pose

$$R_i = \inf(t : t > S_0, X_t = i), \quad R_i^1 = R_i, \quad R_i^{n+1} = \begin{cases} R_i^n + R_i \circ \theta_{R_i^n} & \text{si } R_i^n < \infty \\ \infty & \text{si } R_i^n = \infty. \end{cases}$$

(Nous utilisons la notation  $R_i$  au lieu de  $T_i$  comme dans le cas des chaînes, puisque ici  $T_n$  désigne le  $n$ ième temps de saut de  $\mathbf{X}$ ).

**Définition 2.5.1** Si  $i, j \in E$  on dit que  $i$  mène à  $j$ , et on écrit  $i \mapsto j$ , si on a soit  $i = j$ , soit  $\mathbb{P}_i(R_j < \infty) > 0$ . On écrit  $i \sim j$  si  $i \mapsto j$  et  $j \mapsto i$ .

On a un critère analogue à la proposition 1.5.2, pour lequel nous devons aussi introduire le potentiel  $U = (u_{ij})$ :

$$u_{ij} = \int_0^\infty p(t)_{ij} dt = \mathbb{E}_i \left( \int_0^\infty 1_{\{j\}}(X_t) dt \right). \quad (2.5.1)$$

**Proposition 2.5.2** *Supposons  $\mathbf{X}$  sans explosion. Il y a équivalence entre les quatre conditions*

(a)  $i \mapsto j$ ,

- (b)  $u_{ij} > 0$ ,
- (c) Il existe  $t > 0$  avec  $p(t)_{ij} > 0$ ,
- (d) pour tout  $t > 0$  on a  $p(t)_{ij} > 0$ ,

et la relation  $i \sim j$  est une relation d'équivalence.

**Preuve.** Si  $j = i$ , on a  $p(t)_{ij} \geq e^{-a_i t} > 0$  par (2.4.3) (par exemple), donc on a clairement (a), (b), (c) et (d).

Supposons maintenant  $j \neq i$ . (d)  $\Rightarrow$  (c)  $\Rightarrow$  (b)  $\Rightarrow$  (a) sont évidents (car  $t \mapsto p(t)_{ij}$  est continue). Comme  $\mathbf{X}$  n'explose pas, on peut écrire  $R_j = T_N$ , où  $N$  est une variable aléatoire à valeurs dans  $\{1, 2, \dots, \infty\}$ , et  $R_j < \infty \Leftrightarrow N < \infty$ . Sous  $\mathbb{P}_i$  et conditionnellement à  $\sigma(\xi_n : n \geq 0)$ , chaque variable  $T_m$  est une somme de  $m$  variables exponentielles ou infinies, donc le support de sa loi est soit  $\{\infty\}$ , soit  $\mathbb{R}_+$ . Donc le support de la loi de  $T_m$  sous  $\mathbb{P}_i$  est soit  $[0, \infty[$ , soit  $[0, \infty]$ , soit  $\{\infty\}$ . Il en est évidemment de même pour la variable  $R_j$ , de sorte que sous (a) le support de la loi de  $R_j$  sous  $\mathbb{P}_i$  est soit  $[0, \infty[$ , soit  $[0, \infty]$ . Comme  $X_t = j$  si  $R_j \leq t < R_j + S_0 \circ \theta_{R_j}$  et comme  $S_0 \circ \theta_{R_j}$  est indépendante de  $R_j$  et exponentielle de paramètre  $a_j$ , il est alors clair que (a)  $\Rightarrow$  (d).

Enfin, la dernière assertion se montre comme dans la proposition 1.5.2, en utilisant  $p(t+s)_{ik} \geq p(t)_{ij}p(s)_{jk}$ .  $\square$

On peut alors considérer les classes d'équivalence associées à la relation  $\sim$  ci-dessus, qu'on appelle simplement les *classes* du processus  $\mathbf{X}$ . Elles se comparent aux classes de la chaîne de transition  $\Pi$  de la manière suivante (*attention*: cette dernière chaîne est à valeurs dans  $E_\Delta$ ; mais  $\{\Delta\}$  en constitue évidemment une classe, et les autres sont toutes entièrement contenues dans  $E$ ):

**Proposition 2.5.3** *Supposons  $\mathbf{X}$  sans explosion. Les classes de  $\mathbf{X}$  sont exactement les classes associées à  $\Pi$  et contenues dans  $E$ .*

**Preuve.** En reprenant la preuve précédente, dans l'égalité  $R_j = T_N$  la variable  $N$  représente le premier instant (entier, supérieur ou égal à 1) où la chaîne  $(\xi_n)$  atteint  $j$ : donc la relation  $i \mapsto j$  signifie la même chose pour  $\mathbf{X}$  et pour la chaîne associée à  $\Pi$ , d'où le résultat.  $\square$

Enfin, vu la proposition 2.5.2, la notion de période n'a ici aucune signification: de ce point de vue, les processus de Markov sont un peu plus simples que les chaînes de Markov.

## 2.5.2 Seconde classification

**Définition 2.5.4** Un état  $i$  est dit *récurrent* si  $\mathbb{P}_i(R_i < \infty) = 1$  ou s'il est absorbant ( $a_i = 0$ ), et *transient* sinon.

**Théorème 2.5.5** *Supposons  $\mathbf{X}$  sans explosion.*

(1) On a les équivalences

$$i \text{ récurrent} \Leftrightarrow u_{ii} = \infty \Leftrightarrow \mathbb{P}_i(\cap_n \{R_i^n < \infty\}) = 1 \Leftrightarrow \mathbb{P}_i(\forall t, \exists s > t \text{ avec } X_s = i) = 1. \quad (2.5.2)$$

Dans ce cas si  $i$  est non-absorbant il est aussi récurrent pour la chaîne  $(\xi_n)$  et si de plus  $i \sim j$  alors  $j$  est aussi récurrent et on a  $\mathbb{P}_i(R_j < \infty) = 1$  et  $u_{ij} = \infty$  et  $\mathbb{P}_i(\forall t, \exists s > t \text{ avec } X_s = j) = 1$ .

(2) On a les équivalences

$$i \text{ transient} \Leftrightarrow u_{ii} < \infty \Leftrightarrow \mathbb{P}_i(\cup_n \{R_i^n = \infty\}) = 1 \Leftrightarrow \mathbb{P}_i(\exists t, \forall s > t \text{ on a } X_s \neq i) = 1. \quad (2.5.3)$$

Dans ce cas  $i$  est non-absorbant et si  $i \sim j$  alors  $j$  est transient et  $u_{ij} < \infty$  et  $\mathbb{P}_i(\exists t, \forall s > t \text{ on a } X_s \neq j) = 1$ .

**Preuve.** Si  $i$  est absorbant, il est évident qu'on a les quatre assertions de (2.5.2), et tout  $j \sim i$  vérifie  $j = i$ .

Supposons donc  $i \sim j$  non-absorbants. Comme  $\mathbf{X}$  est non-explosif, on a  $\lim_n R_i^n = \infty$ , donc clairement

$$\cap_n \{R_i^n < \infty\} = \{\forall t, \exists s > t \text{ avec } X_s = i\}, \quad \cup_n \{R_i^n = \infty\} = \{\exists t, \forall s > t \text{ on a } X_s \neq i\}, \quad (2.5.4)$$

Noter aussi que

$$\int_0^\infty 1_{\{i\}}(X_s) ds = S_0 1_{\{X_0=i\}} + \sum_{n \geq 1} S_0 \circ \theta_{R_i^n} 1_{\{R_i^n < \infty\}}. \quad (2.5.5)$$

Posons  $N_i^0 = 0$  et  $N_i^{k+1} = \inf(n > N_i^k : \xi_n = i)$  (temps de passage successifs de  $(\xi_n)$  en  $i$ ). Il est clair que  $R_i^n = T_{R_i^n}$  et  $R_i^n$  est fini si et seulement si  $N_i^n$  est fini (les  $N_i^n$  ici correspondent, pour la chaîne  $(\xi_n)$ , aux  $T_i^n$  du paragraphe 1.5.1). Ainsi on a

$$f_{ij} := \mathbb{P}_i(R_j < \infty) = \mathbb{P}_i(N_i^1 < \infty),$$

et d'après le lemme 1.5.8 il vient

$$f_{ij}^{(n+1)} := \mathbb{P}_i(R_j^{n+1} < \infty) = \mathbb{P}_i(N_j^{n+1} < \infty) = f_{ij}(f_{jj})^n. \quad (2.5.6)$$

Comme  $S_0 \circ \theta_{R_i^n}$  est exponentielle de paramètre  $a_i$ , donc d'espérance  $1/a_j$ , conditionnellement à  $R_i^n < \infty$ , il découle de (2.5.1), (2.5.5) et (2.5.6) que

$$\begin{aligned} u_{ij} &= \delta_{ij} \mathbb{E}_i(S_0) + \sum_{n \geq 1} \mathbb{E}_i(S_0 \circ \theta_{R_j^n} 1_{\{R_j^n < \infty\}}) \\ &= \frac{1}{a_j} \left( \delta_{ij} + \sum_{n \geq 1} \mathbb{P}_i(R_j^n < \infty) \right) = \frac{1}{a_j} \left( \delta_{ij} + f_{ij} \sum_{n \geq 1} (f_{jj})^{n-1} \right). \end{aligned} \quad (2.5.7)$$

Compte tenu de (2.5.4), (2.5.6) et (2.5.7), on obtient les équivalences dans (1) et (2), ainsi que le fait que  $i$  est récurrent pour  $\mathbf{X}$  si et seulement s'il est récurrent pour  $(\xi_n)$  (rappelons que  $i$  est non-absorbant, donc  $a_i > 0$ ). Au vu de (2.5.4) et (2.5.7) et du théorème 1.5.9, les dernières assertions de (1) et (2) sont aussi immédiates.  $\square$

### 2.5.3 Propriétés ergodiques

On va maintenant étudier le comportement ergodique du processus  $\mathbf{X}$ , c'est-à-dire l'existence de limites pour les  $p(t)_{ij}$  lorsque  $t \rightarrow \infty$ , au moins dans le cas non-explosif. Les résultats sont plus simples que dans le cas discret, puisqu'il n'y a pas de notion de période. Nous allons nous contenter d'un résultat partiel, qui n'explique pas la limite obtenue. Pour simplifier on supposera aussi  $\mathbf{X}$  sans explosion.

**Théorème 2.5.6** *Supposons  $\mathbf{X}$  sans explosion. Pour tous  $i, j \in E$  les quantités  $p(t)_{ij}$  convergent vers une limite  $\alpha_{ij}$  lorsque  $t \rightarrow \infty$ . De plus:*

- (1) *Si  $j$  est transient on a  $\alpha_{ij} = 0$  pour tout  $i$ .*
- (2) *Si  $j$  est récurrent on a  $\alpha_{ij} = \mathbb{P}_i(R_j < \infty)\alpha_{jj}$ ; si de plus  $i \sim j$ , on a  $\alpha_{ij} = \alpha_{jj}$ .*
- (3) *On a  $\sum_j \alpha_{ij} \leq 1$ .*

**Preuve.** On fera la démonstration seulement dans le cas où on a l'hypothèse supplémentaire suivante, dont on peut en fait se dispenser: à savoir que  $A = \sup_i a_i < \infty$ .

Soit  $s > 0$ . La matrice  $P_s$  est la probabilité de transition de la chaîne de Markov (à temps discret)  $(X_{ns})_{n \in \mathbb{N}}$ , et évidemment  $(P_s)^n = P_{ns}$ . De plus  $p(s)_{jj} \geq e^{-a_j s} > 0$ , donc la période de  $j$  pour cette chaîne est 1. En vertu du théorème 1.6.6 et de son corollaire 1.6.7, la suite  $p(ns)_{ij}$  converge vers une limite  $\gamma(s)_{ij}$  quand  $n \rightarrow \infty$ .

Soit maintenant  $t > 0$ ; il existe un entier  $n$  tel que  $ns \leq t < (n+1)s$ , et  $P_t = P_{ns}P_{t-ns}$ . Avec la notation  $g(u)_{ij}$  de la preuve du théorème 2.4.2, il vient alors par (2.4.9):

$$p(t)_{ij} = p(ns)_{ij}e^{-a_j(t-ns)} + \sum_k p(ns)_{ik} \int_0^{t-ns} g(u)_{kj} e^{-a_j(t-ns-u)} du.$$

D'une part on en déduit que  $p(t)_{ij} \geq p(ns)_{ij}e^{-a_j s}$ ; donc pour tout  $s > 0$  fixé, et comme  $n \rightarrow \infty$  lorsque  $t \rightarrow \infty$ , on obtient

$$b_{ij} := \liminf_t p(t)_{ij} \geq \gamma(s)_{ij}e^{-a_j s}.$$

D'autre part notre hypothèse additionnelle implique que  $g(u)_{kj} \leq A$  pour tout  $k$ , donc on a aussi  $p(t)_{ij} \leq p(ns)_{ij} + sA$ , et par suite

$$c_{ij} := \limsup_t p(t)_{ij} \leq \gamma(s)_{ij} + sA.$$

En rassemblant ces deux résultats, on voit que  $c_{ij} - b_{ij} \leq sA + 1 - e^{-a_j s}$  pour tout  $s$ , et comme cette dernière quantité tend vers 0 quand  $s \rightarrow 0$  on en déduit que  $c_{ij} = b_{ij}$ : cela entraîne que  $p(t)_{ij}$  converge vers une limite  $\alpha_{ij}$  ( $= b_{ij}$ ) quand  $t \rightarrow \infty$ , et évidemment on a en fait  $\gamma(s)_{ij} = \alpha_{ij}$  pour tout  $s$ .

Comme  $\sum_j p(t)_{ij} = 1$ , (3) découle de ce qui précède et du lemme de Fatou. (2) provient immédiatement du fait que si  $j$  est transient on a  $u_{ij} = \int_0^\infty p(t)_{ij} dt < \infty$  (cf. théorème 2.5.5). Supposons enfin  $j$  récurrent. En utilisant la propriété (2.4.4) avec le temps d'arrêt  $R_j$  et la variable  $V(\omega, \omega') = 1_{\{j\}}(X_{t-R_j(\omega)}(\omega'))$ , on voit que

$$p(t)_{ij} = \mathbb{P}_i(R_j \leq t, X_t = j) = \mathbb{E}_i \left( 1_{\{R_j \leq t\}} p(t - R_j)_{jj} \right).$$

Lorsque  $t \rightarrow \infty$ , la variable  $1_{\{R_j \leq t\}} p(t - R_j)_{jj}$  converge simplement vers  $1_{\{R_j < \infty\}} \alpha_{jj}$ , donc la première partie de (2) provient du théorème de Lebesgue, et la seconde de la première et du fait que  $\mathbb{P}_i(R_j < \infty) = 1$  si  $i \sim j$  (cf. théorème 2.5.5-(1)).  $\square$

## 2.5.4 Probabilités invariantes

Nous nous contenterons d'examiner le problème de l'existence et de la détermination des probabilités invariantes (les lois initiales qui rendent le processus  $\mathbf{X}$  stationnaire, d'après la proposition 2.1.8). On pourrait utiliser le théorème ergodique 2.5.6, mais outre le fait qu'on n'a pas explicité les limites dans ce théorème (et qu'il n'a été montré que lorsque  $\sup_i a_i < \infty$ ), il est intéressant de relier les probabilités invariantes aux termes  $\Pi$  et  $a_i$ .

**Théorème 2.5.7** *Supposons  $\mathbf{X}$  sans explosion. Si  $\mu = (\mu_i)$  est une probabilité sur  $E$ , il y a équivalence entre les assertions suivantes:*

- (i)  $\mu$  est invariante.
- (ii) On a l'équation matricielle  $\mu P'_0 = 0$ .
- (iii) On a  $\mu_i a_i = \sum_j \mu_j a_j \pi_{ji}$  pour tout  $i$ .

**Preuve.** La matrice  $P'_0$  étant donnée par (2.4.1), il est clair que (ii) = (iii).

On va utiliser les matrices  $Q(\lambda) = (q(\lambda)_{ij})$  de (2.3.2) pour  $\lambda > 0$ , et le lemme 2.3.2. D'après Fubini, la propriété de Markov forte et le fait que  $\mathbb{E}_j(1 - e^{-\lambda S_0}) = \frac{\lambda}{a_j + \lambda}$ , on obtient (rappelons que  $e^{-\infty} = 0$ ):

$$\begin{aligned} \lambda \int_0^\infty e^{-\lambda t} p(t)_{ij} dt &= \mathbb{E}_i \left( \lambda \int_0^\infty e^{-\lambda t} 1_{\{j\}}(X_t) dt \right) \\ &= \mathbb{E}_i \left( \sum_{n \geq 0} 1_{\{j\}}(X_{T_n}) 1_{\{T_n < \infty\}} \lambda \int_{T_n}^{T_{n+1}} e^{-\lambda t} dt \right) \\ &= \sum_{n \geq 0} \mathbb{E}_i \left( 1_{\{j\}}(X_{T_n}) e^{-\lambda T_n} (1 - e^{-\lambda S_0 \circ \theta_{T_n}}) \right) \\ &= \sum_{n \geq 0} \mathbb{E}_i \left( 1_{\{j\}}(X_{T_n}) e^{-\lambda T_n} \frac{\lambda}{a_j + \lambda} \right) = \sum_{n \geq 0} q(\lambda)_{ij}^{(n)} \frac{\lambda}{a_j + \lambda}. \end{aligned}$$

Si on a (i) il est clair que

$$\int_0^\infty e^{-\lambda t} \sum_j \mu_j p(t)_{ji} dt = \frac{1}{\lambda} \mu_i \quad (2.5.8)$$

pour tout  $\lambda > 0$ . Supposons à l'inverse (2.5.8); le membre de gauche (resp. de droite) est la transformée de Laplace en  $\alpha$  de la fonction positive continue  $t \mapsto \sum_j \mu_j p(t)_{ji}$  (resp.  $t \mapsto \mu_i$ ), donc par un résultat classique ces deux fonctions sont égales: ainsi, (i) équivaut à (2.5.8) pour tout  $\lambda > 0$  et tout  $i$ .

D'après le calcul précédent et Fubini, (2.5.8) et donc (i) équivalent à

$$\mu_i(a_i + \lambda) = \lambda \sum_j \mu_j \sum_{n \geq 0} q(\lambda)_{ji}^{(n)} \quad (2.5.9)$$



pour tout  $i$ . Mais  $\sum_{n \geq 0} Q(\lambda)^n = I + \left( \sum_{n \geq 0} Q(\lambda)^n \right) Q(\lambda)$ , donc (i) équivaut aussi à

$$\mu_i(a_i + \lambda) = \mu_i \lambda + \sum_k \left( \lambda \sum_j \mu_j \sum_{n \geq 0} q(\lambda)_{jk}^{(n)} \right) q(\lambda)_{ki} \quad (2.5.10)$$

pour tout  $i$ . En injectant (2.5.9) dans (2.5.10), cela revient à écrire

$$\mu_i(a_i + \lambda) = \mu_i \lambda + \sum_k \mu_k(a_k + \lambda) q(\lambda)_{ki}$$

pour tout  $i$ . Vu (2.3.2), ceci n'est autre que (iii), et on a donc le résultat.  $\square$

En d'autres termes, si on associe à toute mesure  $\mu$  la mesure  $\mu'$  définie par  $\mu'_i = \mu_i a_i$ , la mesure  $\mu$  est invariante pour le semi-groupe de  $\mathbf{X}$  si et seulement si la mesure  $\mu'$  est invariante pour la transition  $\Pi$  (en restriction à  $E$ ). Mais il faut faire attention: si  $\mu$  est une probabilité il se peut que  $\mu'$  n'en soit pas une, et même soit infinie; inversement si  $\Pi$  admet une probabilité invariante  $\mu'$ , la mesure associée  $\mu$  n'est pas nécessairement une probabilité, et peut être infinie.

Une manière de trouver les probabilités invariantes est alors la suivante: d'abord, si  $i$  est absorbant, la masse de Dirac  $\varepsilon_i$  est évidemment invariante. Ensuite, on détermine les mesures invariantes  $\mu'$  pour  $\Pi$  (un problème *a priori* plus simple...) portées par une classe d'équivalence de  $\Pi$  (donc vérifiant  $\mu'_i = 0$  si  $i$  est absorbant); on pose alors  $\mu''_i = \mu'_i / a_i$  (avec  $0/0 = 0$ ); si  $C = \sum_i \mu''_i < \infty$ , la mesure  $\mu_i = \mu''_i / C$  est une probabilité invariante pour  $\mathbf{X}$ . Enfin, toute combinaison linéaire convexe de probabilités invariantes est encore une probabilité invariante.

**Corollaire 2.5.8** *Supposons  $\mathbf{X}$  sans explosion. Si la chaîne  $(\xi_n)$  associée à  $\Pi$  est irréductible récurrente, le processus  $\mathbf{X}$  est aussi irréductible récurrent, et admet au plus une probabilité invariante.*

**Preuve.** Dans ce cas on sait que les états sont tous non-absorbants. De plus le théorème 2.5.5 entraîne que  $\mathbf{X}$  possède une seule classe, dont tous les états sont récurrents. Soit  $\mu$  et  $\nu$  deux probabilités invariantes pour  $(P_t)$ . Alors  $\mu'$  et  $\nu'$  sont invariantes pour  $\Pi$ , et d'après le théorème 1.6.2 on a  $\mu'_i = \alpha \nu'_i$  pour une constante  $\alpha \geq 0$ , et comme  $a_i > 0$  et  $\sum_i \mu_i = \sum_i \nu_i = 1$  on a nécessairement  $\mu = \nu$ .  $\square$

**Remarque:** Rappelons encore que si la chaîne  $(\xi_n)$  est irréductible récurrente positive on n'est pas sûr de l'existence d'une probabilité invariante pour  $\mathbf{X}$ . Si à l'inverse  $\mathbf{X}$  est irréductible et possède une probabilité invariante, tous les états sont récurrents pour  $\mathbf{X}$  (car en passant à la limite dans  $\mu P_t = \mu$  on obtient  $\sum_j \mu_i \alpha_{ij} = \mu_j$ , donc  $\alpha_{ij} > 0$  pour au moins un couple  $(i, j)$  et on applique le théorème 2.5.6), donc tous les états sont aussi récurrents pour  $(\xi_n)$ ; mais ils ne sont pas nécessairement positifs pour cette chaîne.

## Chapitre 3

# Processus de Poisson

### 3.1 Processus de Lévy

Les processus de Poisson sont des cas particuliers d'une classe de processus appelés processus de Lévy (ou, processus à accroissements indépendants et stationnaires; ce sont les accroissements qui sont stationnaires), dont on va donner quelques propriétés élémentaires.

**Définition 3.1.1** On appelle  $(\mathcal{F}_t)$ -processus de Lévy un processus  $X = (\mathbf{X}_t)_{t \geq 0}$  défini sur un espace probabilisé filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ , à valeurs dans  $\mathbb{R}^d$ , tel que chaque  $X_t$  soit  $\mathcal{F}_t$ -mesurable et qu'on ait:

- (i)  $X_0 = 0$  et les trajectoires  $t \mapsto X_t(\omega)$  sont continues à droites et avec des limites à gauche,
- (ii) pour tous  $s, t \geq 0$ , la variable  $X_{t+s} - X_t$  est indépendante de la tribu  $\mathcal{F}_t$ , et de même loi que  $X_s$ .

Si  $\mathcal{F}_t = \sigma(X_s : s \leq t)$ , on dit simplement "processus de Lévy".  $\square$

**Théorème 3.1.2** Soit  $X$  un  $(\mathcal{F}_t)$ -processus de Lévy, et  $\mu_t$  la loi de  $X_t$ .

(1) La famille  $(\mu_t)_{t \geq 0}$  est un semi-groupe de convolution, ce qui signifie que pour tout  $s, t \geq 0$  on a

$$\mu_{t+s} = \mu_t \star \mu_s. \quad (3.1.1)$$

(2) Le processus  $X$  satisfait la propriété de Markov relativement à la filtration  $(\mathcal{F}_t)$ , et avec le semi-groupe de transition  $(P_t)$  défini par

$$P_t(x, A) = \int 1_A(x + y) \mu_t(dy). \quad (3.1.2)$$

**Preuve.** Comme  $X_{t+s} = X_t + (X_{t+s} - X_t)$  et comme les variables  $X_t$  et  $X_{t+s} - X_t$  sont indépendantes et de lois respectives  $\mu_t$  et  $\mu_s$ , (1) est évident.

On définit  $P_t$  par (3.1.2). Il est évident que chaque  $P_t$  est une probabilité de transition de  $\mathbb{R}^d$  dans lui-même. De plus

$$\begin{aligned} P_{t+s}(x, A) &= \int 1_A(x+y) \mu_{t+s}(dy) = \int 1_A(x+u+v) \mu_t(du) \mu_s(dv) \\ &= \int \mu_t(du) \left( \int 1_A(x+u+v) \mu_s(dv) \right) \\ &= \int \mu_t(du) P_s(x+u, A) = (P_t P_s)(x, A), \end{aligned}$$

donc  $(P_t)$  est une semi-groupe. Enfin on a

$$\mathbb{E}(f(X_{t+s}) | \mathcal{F}_t) = \mathbb{E}(f(X_t + (X_{t+s} - X_t)) | \mathcal{F}_t) = \int \mu_s(dy) f(X_t + y) = P_s f(X_t),$$

où la seconde égalité provient de (ii) dans la définition ci-dessus.  $\square$

Habituellement on suppose en plus que  $\Omega$  est muni d'un semi-groupe de translations, c'est-à-dire d'une famille  $(\theta_t)_{t \geq 0}$  d'applications mesurables de  $(\Omega, \mathcal{F})$  dans lui-même, vérifiant (comparer avec (2.1.3)):

$$\theta_0 = I, \quad \theta_{s+t} = \theta_s \circ \theta_t, \quad X_{s+t} - X_s = X_t \circ \theta_s, \quad \forall s, t \geq 0. \quad (3.1.3)$$

On dit alors que le terme  $\mathbf{X} = (\Omega, \mathcal{F}, (\mathcal{F}_t), (\theta_t), (X_t), \mathbb{P})$  est un *processus de Lévy*.

Par exemple, on peut prendre pour  $\Omega$  l'espace canonique, qui ici est l'espace de toutes les fonctions  $t \mapsto x(t)$  de  $\mathbb{R}_+$  dans  $\mathbb{R}^d$  vérifiant  $x(0) = 0$  et continues à droite avec des limites à gauche. Puis, pour toute fonction  $x(\cdot)$  on pose

$$X_t(x) = x(t), \quad \theta_t(x)(\cdot) = x(t + \cdot) - x(t),$$

de sorte qu'on a (3.1.3). Enfin  $\mathcal{F}_t = \sigma(X_s : s \leq t)$  et  $\mathcal{F} = \bigvee_t \mathcal{F}_t$ . Dans ce cas, on dit que  $\mathbf{X}$  est un processus de Lévy *canonique*.

Comme par hypothèse  $X_t \rightarrow X_0 = 0$  lorsque  $t \rightarrow 0$ , on a clairement

$$\mu_t \rightarrow \varepsilon_0 \quad \text{étroitement quand } t \rightarrow 0. \quad (3.1.4)$$

Soit  $\mathbf{X}$  un processus de Lévy (au sens ci-dessus). D'après le théorème précédent il vérifie la propriété de Markov, mais ce n'est pas un processus de Markov au sens de la définition 2.1.2 puisqu'il n'y a qu'une seule probabilité. Il est facile d'associer à  $\mathbf{X}$  un processus de Markov  $\mathbf{X}'$  au sens de cette définition 2.1.2. On peut par exemple poser

$$\Omega' = \Omega \times \mathbb{R}^d, \quad X'_t(\omega, x) = x + X_t(\omega), \quad \theta'_t(\omega, x) = (\theta_t(\omega), x + X_t(\omega)),$$

puis

$$\mathcal{F}' = \mathcal{F} \otimes \mathcal{R}^d, \quad \mathcal{F}'_t = \mathcal{F}_t \otimes \mathcal{R}^d$$

(où  $\mathcal{R}^d$  est la tribu borélienne de  $\mathbb{R}^d$ ), et enfin

$$\mathbb{P}'_y(d\omega, dx) = \mathbb{P}(d\omega) \otimes \varepsilon_y(dx) \quad \forall y \in \mathbb{R}^d.$$

On vérifie immédiatement la

**Proposition 3.1.3** *Si  $\mathbf{X}$  est un processus de Lévy, et avec la construction ci-dessus, le terme  $\mathbf{X}' = (\Omega', \mathcal{F}', (\mathcal{F}'_t), (\theta'_t), (X'_t), (\mathbb{P}'_y)_{y \in \mathbb{R}^d})$  est un processus de Markov, de semi-groupe de transition  $(P_t)$  donné par (3.1.2).*

Remarquer que si  $P_t$  est donné par (3.1.2) il est clair que si  $f$  est continue et tend vers 0 à l'infini, il en est de même de  $P_t f$ , qui de plus converge simplement vers  $f$  quand  $t \rightarrow 0$  si on a (3.1.4). Par conséquent la partie (1) du théorème ci-dessous est un corollaire immédiat du théorème 2.1.7 (sauf pour le fait que les trajectoires ont des limites à gauche), compte tenu de l'identification de  $X$  et  $X'$  faite auparavant. Quant à la partie (2), elle est plus difficile et énoncée à titre de référence.

**Théorème 3.1.4** (1) *Si  $(\mu_t)$  est un semi-groupe de convolution de probabilités sur  $\mathbb{R}^d$  vérifiant (3.1.4), il existe un processus de Lévy  $\mathbf{X}$  tel que chaque  $\mu_t$  soit la loi de  $X_t$ .*

(2) *Dans ce cas, les probabilités  $\mu_t$  appartiennent à la classe de probabilités appelées lois indéfiniment divisibles, et pour toute loi indéfiniment divisible  $\mu$  il existe un semi-groupe de convolution  $(\mu_t)$  et un seul, vérifiant (3.1.4) et tel que  $\mu_1 = \mu$ .*

Terminons par la propriété forte de Markov:

**Théorème 3.1.5** *Soit  $\mathbf{X}$  un processus de Lévy. Le processus de Markov associé  $\mathbf{X}'$  vérifie la propriété forte de Markov (cf. (2.1.7)), et  $\mathbf{X}$  vérifie la version suivante de la propriété de Markov forte: pour tout temps d'arrêt  $T$  pour la filtration  $(\mathcal{F}_t)$  et toute variable aléatoire  $Y$  positive ou bornée on a*

$$\mathbb{E}(Y \circ \theta_T | \mathcal{F}_{T+}) = \mathbb{E}(Y) \quad \text{sur l'ensemble } \{T < \infty\}. \quad (3.1.5)$$

Une énoncé équivalent à (3.1.5) pour toute variable  $Y$  consiste à dire ceci: pour toute variable aléatoire  $Y'$ , la translatée  $Y' \circ \theta_T$  est indépendante de  $\mathcal{F}_{T+}$  et de même loi que  $Y'$ , conditionnellement à  $T < \infty$ . Cet énoncé entraîne bien-sûr (3.1.5), et réciproquement (3.1.5) appliqué à  $Y = f(Y')$  pour  $f$  mesurable arbitraire donne l'énoncé ci-dessus.

**Preuve.** Comme on a (3.1.4) il est clair que  $P_t f$  est continue pour toute fonction continue bornée  $f$ . Par suite la première assertion découle du théorème 2.1.5.

Pour (3.1.5) on considère une variable positive (resp. bornée)  $\mathcal{F}$ -mesurable et un  $(\mathcal{F}_t)$ -temps d'arrêt  $T$ . Il s'agit de montrer que pour toute variable  $\mathcal{F}_{T+}$ -mesurable  $Z$  positive (resp. bornée), on a

$$\mathbb{E}(Z Y \circ \theta_T 1_{\{T < \infty\}}) = \mathbb{E}(Y) \mathbb{E}(Z 1_{\{T < \infty\}}). \quad (3.1.6)$$

On peut considérer aussi  $Y$ ,  $Z$  et  $T$  comme étant définis sur  $\Omega'$  (par exemple,  $T(\omega, x) = T(\omega)$ ): dans ce cas  $T$  est un  $(\mathcal{F}'_t)$ -temps d'arrêt,  $Y$  est  $\mathcal{F}'$ -mesurable et  $Z$  est  $\mathcal{F}'_{T+}$ -mesurable. La propriété de Markov forte pour  $\mathbf{X}'$  donne

$$\mathbb{E}'_0(Z Y \circ \theta'_T 1_{\{T < \infty\}}) = \mathbb{E}'_0(Z 1_{\{T < \infty\}}) \mathbb{E}'_{X'_T}(Y). \quad (3.1.7)$$

Mais  $\mathbb{E}'_y(U) = \mathbb{E}(U)$  pour toute variable  $U$  vérifiant  $U(\omega, x) = U(\omega)$  et tout  $y$ . Donc le second membre de (3.1.7) coïncide avec celui de (3.1.6). Par ailleurs  $Y \circ \theta'_T(\omega, x) = Y \circ \theta_T(\omega)$ , donc les premiers membres de ces deux relations coïncident également.  $\square$

## 3.2 Le processus de Poisson

### 3.2.1 Les processus de comptage

**Définition 3.2.1** Un *processus de comptage* est un processus  $N = (N_t)_{t \geq 0}$  dont les trajectoires  $t \mapsto N_t(\omega)$  sont à valeurs dans  $\overline{\mathbb{N}} = \{0, 1, \dots, \infty\}$ , croissantes, continues à droite, nulles en 0, et de sauts de taille 1.

On associe au processus de comptage  $N$  ses temps de saut successifs  $T_n$  et les intervalles inter-sauts  $S_n$ , par les formules suivantes (avec la convention  $\infty - \infty = \infty$ ):

$$T_n = \inf(t : N_t = n), \quad S_n = T_{n+1} - T_n, \quad \forall n \in \mathbb{N}, \quad (3.2.1)$$

et on pose aussi  $T_\infty = \lim_n T_n$ , le “temps d’explosion”. Les temps  $T_0$  et  $T_\infty$  ne sont *pas* des temps de saut. On a  $N_t = \infty$  si et seulement si  $t \geq T_\infty$ , et le processus  $N$  est dit *non-explosif* si  $T_\infty = \infty$  identiquement (ou au moins p.s., si on a une probabilité).

Remarquer que  $T_n < T_{n+1}$  si  $T_n < \infty$ , et en particulier  $T_1 > 0$ . On peut imaginer que les  $T_n$  pour  $n \geq 1$  sont les temps d’occurrence d’un certain événement, et  $N_t$  est le nombre d’occurrences sur l’intervalle  $]0, t]$ : cela explique la terminologie “processus de comptage”.

Si (3.2.1) donne les  $T_n$  et les  $S_n$  en fonction de  $N$ , on peut à l’inverse reconstruire le processus  $N$  à partir des  $T_n$  (une suite de variables avec  $T_0 = 0$ ,  $T_{n+1} > T_n$  si  $T_n < \infty$  et  $T_{n+1} = \infty$  si  $T_n = \infty$ ), ou des variables  $S_n$  strictement positives, grâce aux formules

$$N_t = \sum_{n \geq 1} 1_{\{T_n \leq t\}}, \quad T_n = S_0 + \dots + S_{n-1} \quad \text{si } n \geq 1. \quad (3.2.2)$$

(l’indexation des  $S_n$ , en partant de  $S_0$ , est la même que pour les processus de Markov de saut pur).

Si enfin on a une filtration  $(\mathcal{F}_t)$ , en remarquant que  $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$  et que  $\{T_n \leq t\} = \{N_t \geq n\}$ , on voit qu’on a:

$$N \text{ est adapté} \iff \text{les } T_n \text{ sont des temps d'arrêt.} \quad (3.2.3)$$

### 3.2.2 Trois définitions équivalentes

Il existe plusieurs définitions possibles – équivalentes – du (ou des) processus de Poisson. La première n’est pas la plus “élémentaire”, mais elle suit naturellement le premier paragraphe.

**Définition 3.2.2** On appelle  $(\mathcal{F}_t)$ -*processus de Poisson* un processus de comptage  $N$  sur un espace probabilisé filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  qui est aussi un  $(\mathcal{F}_t)$ -processus de Lévy. Lorsque  $\mathcal{F}_t = \sigma(N_s : s \leq t)$  on dit simplement *processus de Poisson*.

Ainsi, un processus de Poisson prend ses valeurs dans  $\overline{\mathbb{N}}$  (comme processus de comptage) et dans  $\mathbb{R}^d$  (comme processus de Lévy), donc  $d = 1$  et en fait il prend ses valeurs

dans  $\mathbb{N}$ . Il est donc non-explosif. Noter aussi qu'un  $(\mathcal{F}_t)$ -processus de Poisson est *a fortiori* un processus de Poisson.

Donnons maintenant, sous forme de propriétés équivalentes à la définition 3.2.1, deux autres définitions des processus de Poisson.

**Théorème 3.2.3** *Soit  $N$  est un processus de comptage adapté à une filtration  $(\mathcal{F}_t)$ . Il y a équivalence entre*

(1)  *$N$  est une  $(\mathcal{F}_t)$ -processus de Poisson,*

(2) *Il existe  $\lambda \geq 0$  tel que le processus  $N_t - \lambda t$  soit une martingale relativement à  $(\mathcal{F}_t)$ .*

*Si de plus  $\mathcal{F}_t = \sigma(N_s : s \leq t)$ , ces deux propriétés sont aussi équivalentes à*

(3) *les  $S_n$  sont des variables indépendantes de loi exponentielle de même paramètre  $\lambda \geq 0$  (rappelons qu'une variable exponentielle de paramètre 0 est identiquement infinie).*

*Enfin dans ce cas, le nombre  $\lambda$  est le même dans (2) et dans (3), et est appelé le paramètre du processus de Poisson.*

Une conséquence importante de ce théorème est que pour tout  $\lambda \geq 0$  il existe un processus de Poisson de paramètre  $\lambda$ , et un seul en loi (cela découle aussi du théorème – non démontré – 3.1.4): il suffit de construire une suite  $(S_n)$  comme dans (3), et de poser (3.2.2).

La preuve sera faite en plusieurs étapes, en même temps que nous démontrerons quelques autres propriétés intéressantes.

**Lemme 3.2.4** *On a (1)  $\Rightarrow$  (3).*

**Preuve.** Si on a (1),  $N$  est de saut pur et vérifie la propriété de Markov, donc d'après le théorème 2.2.3 la variable  $S_0$  est exponentielle avec un certain paramètre  $\lambda \geq 0$ . Si  $\lambda = 0$  on a  $S_0 = \infty$  p.s., donc aussi  $S_n = \infty$  p.s. d'après (3.2.1), et on a (3).

Supposons alors  $\lambda > 0$ . En utilisant les translation  $\theta_t$  associées à  $N$  comme dans (3.1.3), on voit que  $S_n = S_0 \circ \theta_{T_n}$ . La propriété (3.1.5) entraîne alors que  $S_n$  est indépendante de  $\mathcal{F}_{T_n+}$  et de même loi que  $S_0$ , conditionnellement à  $\{T_n < \infty\}$ : on en déduit d'abord par récurrence sur  $n$  que  $T_n < \infty$  p.s., puis donc que  $S_n$  est indépendante de  $\mathcal{F}_{T_n+}$  et *a fortiori* de  $S_0, \dots, S_{n-1}$ : on a donc (3).  $\square$

D'après cette preuve, on peut considérer le processus de Poisson comme un processus de Markov de saut pur à valeurs dans  $\mathbb{N}$ ; avec ce point de vue, les caractéristiques  $\Pi$  et  $a_i$  qui lui sont associées dans le théorème 2.2.3 sont

$$a_i = \lambda, \quad \pi_{ij} = \begin{cases} 1 & \text{si } j = i + 1 \\ 0 & \text{sinon.} \end{cases} \quad (3.2.4)$$

Pour le résultat suivant, rappelons que la loi gamma de paramètre  $\lambda > 0$  et d'indice  $n$  est la loi de densité

$$f_{n,\lambda}(s) = \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s} \mathbf{1}_{\mathbb{R}_+}(s). \quad (3.2.5)$$

Donc si  $n = 1$  c'est la loi exponentielle de paramètre  $\lambda$ . La transformée de Laplace en est

$$\phi_{n,\lambda}(v) := \int e^{-vs} f_{n,\lambda}(s) ds = \left( \frac{\lambda}{\lambda + v} \right)^n.$$

**Lemme 3.2.5** *Si  $N$  est un processus de Poisson de paramètre  $\lambda > 0$ , la loi de  $N_t$  est la loi de Poisson de paramètre  $\lambda t$  (i.e.,  $\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$  pour  $n \in \mathbb{N}$ ), et la loi de  $T_n$  est la loi gamma de paramètre  $\lambda$  et d'indice  $n$ .*

**Preuve.** Comme  $T_n$  est la somme de  $n$  variables exponentielles de paramètre  $\lambda$  et indépendantes, sa transformée de Laplace est  $v \mapsto (\phi_{1,\lambda}(v))^n = \phi_{n,\lambda}(v)$ , et on a la seconde assertion. On a aussi

$$\mathbb{P}(N_t = n) = \mathbb{P}(T_n \leq t < T_{n+1}) = \mathbb{P}(T_n \leq t) - \mathbb{P}(T_{n+1} \leq t).$$

Une intégration par parties utilisant les fonctions  $u(s) = e^{-\lambda s}$  et  $v(s) = \frac{\lambda^n s^n}{n!}$  donne

$$\int_0^t f_{n,\lambda}(s) ds = u(t)v(t) - u(0)v(0) + \int_0^t f_{n+1,\lambda}(s) ds,$$

et  $u(t)v(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$  et  $u(0)v(0) = 0$ , d'où la première assertion.  $\square$

**Lemme 3.2.6** *On a (1)  $\Rightarrow$  (2) et le paramètre  $\lambda$  dans (2) est le même que dans (3).*

**Preuve.** Supposons (1). D'après les lemmes précédents on a (3) avec un paramètre  $\lambda \geq 0$  et si  $\lambda > 0$  la variable  $N_t$  est intégrable et d'espérance  $\lambda t$ ; si  $\lambda = 0$  on a  $S_0 = \infty$  p.s., donc  $N_t = 0$  p.s., donc en fait  $\mathbb{E}(N_t) = \lambda t$  dans tous les cas.

Posons  $M_t = N_t - \lambda t$ , qui est  $\mathcal{F}_t$ -mesurable et d'intégrale nulle. En utilisant le fait que  $N_{t+s} - N_t$  est indépendante de  $\mathcal{F}_t$  et de même loi que  $N_s$ , on voit que

$$\mathbb{E}(M_{t+s} - M_t | \mathcal{F}_t) = \mathbb{E}(N_{t+s} - N_t | \mathcal{F}_t) - \lambda s = \mathbb{E}(N_s) - \lambda s = 0,$$

d'où (3) avec le paramètre  $\lambda$ .  $\square$

**Lemme 3.2.7** *Si  $\mathcal{F}_t = \sigma(N_s : s \leq t)$ , on a (3)  $\Rightarrow$  (1).*

**Preuve.** Il s'agit de montrer la propriété suivante: étant donnée une suite  $(S_n)_{n \in \mathbb{N}}$  de variables indépendantes et exponentielles de paramètre  $\lambda$ , le processus  $N$  défini par (3.2.2) est un processus de Poisson, relativement à la filtration  $\mathcal{F}_t = \sigma(N_s : s \leq t)$ . Si  $\lambda = 0$  on a en fait  $N \equiv 0$  et il n'y a rien à montrer, donc nous supposons  $\lambda > 0$ .

D'abord, d'après la loi des grands nombres  $T_n/n \rightarrow 1/\lambda > 0$  p.s., donc  $T_n \rightarrow \infty$  p.s. et le processus  $N$  est non-explosif.

On peut voir le résultat comme un cas particulier du théorème 2.3.1, avec les données (3.2.4), mais il est aussi simple (ou compliqué...!) de reprendre la même démonstration plutôt que d'adapter ce théorème. Comme toute variable  $\mathcal{F}_t$ -mesurable est égale, en restriction à chaque ensemble  $\{N_t = n\}$ , à une fonction de  $(S_0, \dots, S_{n-1})$ , il suffit (à cause

du caractère non-explosif) de montrer que pour tous  $m, n \in \mathbb{N}$ , tous  $s, t \geq 0$  et toute fonction borélienne bornée  $g$  sur  $\mathbb{R}_+^n$ ,

$$\mathbb{E}(Z \mathbf{1}_{\{N_t=n\}} \mathbf{1}_{\{N_{t+s}-N_t=m\}}) = \mathbb{E}(Z \mathbf{1}_{\{N_t=n\}}) \mathbb{P}(N_s = m). \quad (3.2.6)$$

On a calculé déjà  $\mathbb{P}(N_s = m)$  dans le lemme 3.2.5), mais comme cette quantité est  $\mathbb{P}(T_m \leq s < T_{m+1})$  on peut aussi l'écrire ainsi:

$$\mathbb{P}(N_s = m) = \int_{]0, \infty[^{m+1}} \left( \prod_{k=0}^m \lambda e^{-\lambda r_k} \right) f(r_0, \dots, r_m) dr_0 \dots dr_m,$$

avec  $f(r_0, \dots, r_m) = \mathbf{1}_{\{r_0 + \dots + r_{m-1} \leq s < r_0 + \dots + r_m\}}$ . Donc le membre de droite de (3.2.6) vaut

$$\int_{]0, \infty[^{n+1}} \left( \prod_{k=0}^n \lambda e^{-\lambda s_k} \right) g(s_0, \dots, s_{n-1}) \mathbf{1}_{\{s_0 + \dots + s_{n-1} \leq t < s_0 + \dots + s_n\}} ds_0 \dots ds_n \\ \int_{]0, \infty[^{m+1}} \left( \prod_{l=0}^m \lambda e^{-\lambda r_l} \right) f(r_0, \dots, r_m) dr_0 \dots dr_m.$$

L'intégrale de  $s_n \mapsto \lambda e^{-\lambda s_n}$  sur l'ensemble  $\{s_0 + \dots + s_{n-1} \leq t < s_0 + \dots + s_n\}$  égale  $\exp(-\lambda(t - (s_0 + \dots + s_{n-1}))) \mathbf{1}_{\{t \geq s_0 + \dots + s_{n-1}\}}$ , donc l'expression précédente vaut

$$\int_{]0, \infty[^{n+m+1}} \lambda^n e^{-\lambda t} \left( \prod_{l=0}^m \lambda e^{-\lambda r_l} \right) g(s_0, \dots, s_{n-1}) \mathbf{1}_{\{s_0 + \dots + s_{n-1} \leq t\}} f(r_0, \dots, r_m) \\ ds_0 \dots ds_{n-1} dr_0 \dots dr_m.$$

On fait les changement de variables  $s_{n+k} = r_k$  si  $k \geq 1$  et  $s_n = r_0 + t - (s_0 + \dots + s_{n-1})$ , de jacobien 1, et l'expression précédente devient

$$\int_{]0, \infty[^{n+m+1}} \left( \prod_{k=0}^{n+m} \lambda e^{-\lambda s_k} \right) g(s_0, \dots, s_{n-1}) \mathbf{1}_{\{s_0 + \dots + s_{n-1} \leq t < s_0 + \dots + s_n\}} \\ f(s_0 + \dots + s_n - t, s_{n+1}, \dots, s_{n+m}) ds_0 \dots ds_{n+m},$$

qui égale clairement le premier membre de (3.2.6), et la preuve est terminée.  $\square$

La preuve de l'implication (2)  $\Rightarrow$  (1) nécessite une version élémentaire du "calcul stochastique" discontinu. On se place sur un espace probabilisé filtré. Un processus  $X$  est dit à variation finie s'il est la différence de deux processus croissant positifs, continus à droite, nuls en 0, soit  $X_t = X'_t - X''_t$ . La "variation" de  $X$  est le processus  $X'_t + X''_t$ . On peut intégrer par rapport à un tel processus  $X$  tout processus  $H = (H_t)$  borné et à trajectoires continues à gauche, par exemple:

$$\int_0^t H_s dX_s = \int_0^t H_s dX'_s - \int_0^t H_s dX''_s,$$

où les deux intégrales de droite sont des "intégrales de Stieltjes" (la première, par exemple, est l'intégrale de  $s \mapsto H_s$  par rapport à la mesure dont la fonction de répartition est  $t \mapsto X'_t$ ); tout ceci se fait " $\omega$  par  $\omega$ ".



Lorsque de plus  $H$  s'écrit  $H_t = \sum_{n \geq 0} \alpha_n 1_{]T_n, T_{n+1}]}(t)$ , pour une suite de variables  $\alpha_n$  uniformément bornées, et une suite strictement croissante  $(T_n)$  tendant vers l'infini, on a évidemment

$$\int_0^t H_s dX_s = \sum_{n \geq 0} \alpha_n (X_{T_{n+1}} \wedge t - X_{T_n} \wedge t), \quad (3.2.7)$$

où la somme ci-dessus ne comporte qu'un nombre fini (aléatoire) de termes non nuls.

**Lemme 3.2.8** *En plus des notations et hypothèses ci-dessus, supposons que les  $T_n$  soient des temps d'arrêt, que les  $\alpha_n$  soient  $\mathcal{F}_{T_n}$ -mesurables, que  $X$  soit une martingale et que sa variation  $X'_t + X''_t$  soit intégrable pour tout  $t$ . Alors le processus  $H \bullet X_t = \int_0^t H_s dX_s$  donné par (3.2.7) est également une martingale continue à droite.*

**Preuve.** Posons  $Y = H \bullet X$ . Si  $C$  est une borne pour  $H$ , on a  $|Y_t| \leq C(X'_t + X''_t)$ , qui est intégrable, et (3.2.7) montre que  $Y_t$  est  $\mathcal{F}_t$ -mesurable (cf. la partie (b) de la preuve du théorème 2.1.5), et la continuité à droite de  $Y$  est évidente.

Il reste à montrer que si  $s < t$  on a  $\mathbb{E}(Y_t - Y_s | \mathcal{F}_s) = 0$ . Vu (3.2.7), il suffit de montrer que  $\mathbb{E}(U_n | \mathcal{F}_s) = 0$  pour chaque  $n$ , où

$$U_n = \alpha_n (X_{T_{n+1}} \wedge t - X_{T_n} \wedge t - X_{T_{n+1}} \wedge s + X_{T_n} \wedge s).$$

Il suffit même de le montrer séparément sur les trois éléments  $A_n = \{T_{n+1} < s\}$ ,  $B_n = \{T_n < s \leq T_{n+1}\}$ ,  $C_n = \{s \leq T_n\}$ . Sur  $A_n$  on a  $U_n = 0$ , d'où le résultat. Sur  $B_n$  on a  $U_n = \alpha_n (X_{T_{n+1}} \wedge t - X_s)$  et  $\alpha_n 1_{B_n}$  est  $\mathcal{F}_s$ -mesurable, donc  $\mathbb{E}(U_n 1_{B_n} | \mathcal{F}_s) = 0$  par le théorème d'arrêt pour les martingales à temps continu, continues à droite. Sur  $C_n$  on a  $U_n = \alpha_n (X_{T_{n+1}} \wedge t - X_{T_n} \wedge t) 1_{\{T_n < t\}}$ , et  $\alpha_n 1_{\{T_n < t\}}$  est  $\mathcal{F}_{(T_n \wedge t)_+}$ -mesurable, donc une nouvelle application du théorème d'arrêt montre que  $\mathbb{E}(U_n 1_{C_n} | \mathcal{F}_s) = 0$ .  $\square$

**Lemme 3.2.9** *On a (2)  $\Rightarrow$  (1).*

**Preuve.** On part d'un processus de comptage  $N$ , tel que le processus  $M_t = N_t - \lambda t$  soit une martingale relativement à la filtration  $(\mathcal{F}_t)$ , pour un certain nombre  $\lambda \geq 0$ . En particulier  $\mathbb{E}(M_t) = \mathbb{E}(M_0) = 0$ , donc  $M_t$  et aussi  $N_t$  sont à valeurs finies p.s.: par suite  $N$  est non-explosif.

Si  $\lambda = 0$ , on a aussi  $\mathbb{E}(N_t) = \mathbb{E}(M_t) = 0$ . Donc en fait  $N \equiv 0$  et le résultat est évident. On va donc supposer  $\lambda > 0$  dans la suite. L'idée consiste à appliquer le lemme précédent avec la martingale  $M$  (qui vérifie toutes les hypothèses, et en particulier sa variation  $M'_t + M''_t = N_t + \lambda t$  est intégrable), et avec le processus  $H_t = (u - 1)u^{N_{t-}}$ , où  $u \in [0, 1]$  et  $N_{t-}$  désigne la limite à gauche de  $N$  au temps  $t$ . Il est clair que  $|H_t| \leq 1$ , et on a aussi  $H_t = \sum_{n \geq 0} \alpha_n 1_{]T_n, T_{n+1}]}(t)$  avec les  $T_n$  qui sont les temps de saut de  $N$  et  $\alpha_n = (u - 1)u^n$ . Par suite le processus

$$Y_t = \int_0^t H_s dM_s = \int_0^t H_s dN_s - \lambda \int_0^t H_s ds$$

est une martingale.

D'autre part le processus  $t \mapsto \int_0^t H_s dN_s$  est constant sur chaque intervalle  $[T_n, T_{n+1}[$ , et au temps  $T_n$  (avec  $n \geq 1$ ) il saute de la quantité  $\alpha_{n-1} = (u-1)u^{n-1} = u^n - u^{n-1}$ . Le processus  $u^{N_t}$  a exactement les mêmes propriétés, donc ces deux processus ne diffèrent que par leurs valeurs à l'instant 0, et donc

$$u^{N_t} = 1 + \int_0^t H_s dN_s.$$

En rassemblant ces deux résultats, on voit que le processus  $Z_t = u^{N_t} - 1 - \lambda \int_0^t H_s ds$  est une martingale nulle en 0. Donc en utilisant le théorème de Fubini pour les espérances conditionnelles, pour tous  $s, t \geq 0$  on arrive à

$$\begin{aligned} \mathbb{E} \left( u^{N_{t+s} - N_t} | \mathcal{F}_t \right) &= \mathbb{E} \left( 1 + u^{-N_t} (u^{N_{t+s}} - u^{N_t}) | \mathcal{F}_t \right) \\ &= 1 + u^{-N_t} \mathbb{E} \left( Z_{t+s} - Z_t + \lambda \int_t^{t+s} H_r dr | \mathcal{F}_t \right) \\ &= 1 + \lambda(u-1) e^{-N_t} \int_0^s \mathbb{E} \left( u^{N_{t+r}} | \mathcal{F}_t \right) dr \\ &= 1 + \lambda(u-1) \int_0^s \mathbb{E} \left( u^{N_{t+r} - N_t} | \mathcal{F}_t \right) dr. \end{aligned} \quad (3.2.8)$$

Cela prouve que  $\phi_u(s) = \mathbb{E}(u^{N_{t+s} - N_t} | \mathcal{F}_t)$  (pour  $t, u$  fixés) vérifie l'équation intégrale

$$\phi_u(s) = 1 + \lambda(u-1) \int_0^s \phi_u(r) dr. \quad (3.2.9)$$

Par ailleurs  $\phi_u$  est une fonction continue à droite de  $s$ , donc nécessairement  $\phi_u(s) = e^{\lambda(u-1)s}$ , c'est-à-dire

$$\mathbb{E}(u^{N_{t+s} - N_t} | \mathcal{F}_t) = e^{\lambda(u-1)s}. \quad (3.2.10)$$

Ci-dessus à droite, et en tant que fonction de  $u$ , on obtient la fonction génératrice de la loi de Poisson de paramètre  $\lambda s$ , et comme la fonction génératrice caractérise la loi on en déduit que la loi conditionnelle de  $N_{t+s} - N_t$  sachant  $\mathcal{F}_t$  est la loi de Poisson de paramètre  $\lambda s$ : cela prouve (1), et aussi la seconde partie du lemme 3.2.5.

Nous avons été trop vite, car (3.2.8) et la fonction  $\phi_u$  elle-même (qui est aléatoire) ne sont vraies ou définies qu'à des ensembles négligeables près. Mais cela peut se justifier: on peut considérer la loi conditionnelle si  $\mathcal{F}_t$  du processus  $(N'_s = N_{t+s} - N_t)_{s \geq 0}$ , soit  $\mathbb{P}' = \mathbb{P}'(\omega; d\omega')$ . Ensuite on prend la version de  $\phi_u$  donnée par  $\phi_u(s)(\omega) = \int \mathbb{P}'(\omega, d\omega') u^{N'_s(\omega')}$ , qui d'une part vérifie (3.2.9) presque sûrement, pour chaque  $s$  fixé, et d'autre part est continue à droite en  $s$ : par suite elle vérifie (3.2.9) pour tout  $s$ , en dehors d'un ensemble négligeable, et on en déduit (3.2.10), donc  $\phi_u(s) = e^{\lambda(u-1)s}$ , en dehors d'un ensemble négligeable. Enfin les deux membres de cette égalité sont (pour  $s$  fixé) continus en  $u$ : donc on a  $\phi_u(s) = e^{\lambda(u-1)s}$  pour tout  $u \in [0, 1]$ , en dehors d'un ensemble négligeable, ce qui prouve que pour presque tout  $\omega$ , sous  $\mathbb{P}'(\omega, \cdot)$  la loi de  $N'_s$  est la loi de Poisson de paramètre  $\lambda s$ : à ce point, la démonstration est correcte et achevée.  $\square$

En rassemblant tous ces lemmes, on obtient le théorème 3.2.3.

### 3.2.3 Trois autres caractérisations

Il existe une autre manière de considérer un processus de comptage, comme étant la “fonction de répartition” d’une mesure aléatoire sur  $\mathbb{R}_+$ . En effet au processus  $N$  défini sur l’espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$  on peut associer la mesure  $\mu = \mu(\omega, dt)$  sur  $\mathbb{R}_+$  par

$$\mu(\omega, dt) = \sum_{n \geq 1} 1_{\{T_n(\omega) < \infty\}} \varepsilon_{T_n(\omega)}(dt), \quad (3.2.11)$$

et on a  $N_t(\omega) = \mu(\omega, [0, t])$ . Une telle mesure (aléatoire), somme de masses de Dirac en des points  $T_n$  deux-à-deux distincts, s’appelle une “répartition ponctuelle” su  $\mathbb{R}_+$  (ou plutôt, sur  $]0, \infty[$ , puisqu’on a exclu  $T_0 = 0$  de la somme). Clairement,  $\omega \mapsto \mu(\omega, A)$  est mesurable pour tout borélien  $A$ .

Les trois définitions précédentes pour les processus de Poisson sont adaptées aux processus indicés par  $\mathbb{R}_+$  (ou éventuellement  $\mathbb{R}$ ) mais n’ont aucun sens lorsque l’espace des “temps” est  $\mathbb{R}^q$ , ou un espace “abstrait”. Nous proposons ci-dessous trois autres définitions équivalentes aux précédentes, et qui s’étendront aux répartitions ponctuelles sur un espace plus général que  $\mathbb{R}_+$ . Là encore, ces définitions supplémentaires seront formulées comme des propriétés caractéristiques.

Pour formuler ce théorème nous avons encore besoin de notations. On note  $m$  la mesure de Lebesgue sur  $\mathbb{R}_+$ . Pour tout borélien  $A$  on note  $\mu_A$  la restriction de  $\mu$  à  $A$ , i.e.  $\mu_A(B) = \mu(A \cap B)$ . On considère alors la tribu

$$\mathcal{F}(A) = \sigma(\mu_A(B) : B \in \mathcal{R}_+) \quad (3.2.12)$$

qui est “engendrée” par la restriction  $\mu_A$  ( $\mathcal{R}_+$  est la tribu borélienne de  $\mathbb{R}_+$ ). Il est clair que  $\mathcal{F}([0, t]) = \sigma(N_s, s \leq t)$ .

Pour toute fonction borélienne positive sur  $\mathbb{R}_+$  on note  $\mu(f) = \mu(f)(\omega)$  l’intégrale de  $f$  par rapport à la mesure  $\mu(\omega, \cdot)$ , c’est-à-dire

$$\mu(f) = \sum_{n \geq 1} f(T_n) 1_{\{T_n < \infty\}}.$$

Il est donc évident que  $\omega \mapsto \mu(f)(\omega)$  est mesurable.

Enfin, si  $E \in \mathcal{R}_+$  vérifie  $0 < m(E) < \infty$ , on dit qu’une variable  $S$  est “uniformément répartie sur  $E$ ” si sa loi est la restriction de  $m$  à  $E$ , normalisée par  $m(E)$  (i.e.,  $\mathbb{P}(S \in A) = \frac{m(A \cap E)}{m(E)}$ ). Si  $S_1, \dots, S_m$  sont  $m$  variables indépendantes uniformément réparties sur  $E$ , on appelle “réarrangement croissant” des  $S_i$  la suite  $R_1, \dots, R_m$  de variables telles que  $R_1 \leq \dots \leq R_m$  et que les deux ensembles  $\{R_i : 1 \leq i \leq m\}$  et  $\{S_i : 1 \leq i \leq m\}$  soient les mêmes. Il est facile de voir qu’il existe une “version mesurable” de ces variables; en fait, en dehors d’un ensemble de mesure nulle les  $S_i$  sont toutes différentes, donc  $R_1 < \dots < R_m$  p.s.

**Théorème 3.2.10** *Soit  $N$  un processus de comptage sur un espace  $(\Omega, \mathcal{F}, \mathbb{P})$ , et  $\mu$  la mesure aléatoire (répartition ponctuelle) associée par (3.2.11). Fixons également une partition borélienne  $(E_n)_{n \geq 1}$  de  $\mathbb{R}_+$  avec  $m(E_n) < \infty$  pour tout  $n$ . Il y a équivalence entre*

- (1)  $N$  est un processus de Poisson de paramètre  $\lambda \geq 0$ .

(4) Pour toute famille  $(A_i)_{1 \leq i \leq n}$  de boréliens de  $\mathbb{R}_+$ , deux-à-deux disjoints, les variables  $\mu(A_i)$  ( $= \mu(\omega, A_i)$ ) sont indépendantes, et  $\mathbb{E}(\mu(A)) = \lambda m(A)$  pour tout borélien  $A$ .

(5) Les tribus  $\mathcal{F}(E_n)$  sont indépendantes; de plus, pour chaque  $n$  la loi de  $\mu(E_n)$  est de Poisson de paramètre  $\lambda m(E_n)$ ; enfin, conditionnellement au fait que  $\mu(E_n) = m$  pour un  $m \geq 1$ , si on note  $T_{n_1} < \dots < T_{n_m}$  les temps successifs  $T_n$  appartenant à  $E_m$ , la loi de  $(T_{n_1}, \dots, T_{n_m})$  est la même que la loi du réarrangement croissant de  $m$  variables indépendantes uniformément réparties sur  $E_n$ .

(6) Pour toute fonction borélienne positive  $f$  l'application  $\omega \mapsto \mu(f)(\omega)$  est mesurable, et on a (avec  $e^{-\infty} = 0$ ):

$$\mathbb{E}(e^{-\mu(f)}) = \exp\left(-\lambda \int_0^\infty (1 - e^{-f(t)}) dt\right). \quad (3.2.13)$$

Noter que dans (5), la partition  $(E_n)$  est totalement arbitraire. Cette propriété fournit une nouvelle méthode de construction d'un processus de Poisson: on prend par exemple  $E_n = ]n-1, n]$ ; puis on construit des suites  $(M_n)_{n \geq 1}$  et  $(R_{n,m})_{n \geq 1, m \geq 1}$  de variables globalement indépendantes, telles que chaque  $M_n$  suive la loi de Poisson de paramètre  $\lambda$  et chaque  $R_{n,m}$  suive la loi uniforme sur  $E_n$ . Il reste alors à poser

$$\mu = \sum_{n \geq 1} \sum_{m=1}^{M_n} \varepsilon_{R_{n,m}}, \quad \text{où une somme "vide" vaut 0,} \quad (3.2.14)$$

et  $N_t = \mu(]0, t])$ .

Dans (6), l'application  $f \mapsto \mathbb{E}(e^{-\mu(f)})$  s'appelle la *fonctionnelle de Laplace* de la mesure aléatoire  $\mu$ .

Là encore la preuve sera faite en une série de lemmes. Noter que si  $\lambda = 0$ , les quatre conditions (1), (4), (5) et (6) impliquent chacune que  $N \equiv 0$ , et il n'y a donc rien à montrer. Donc dans la suite on suppose toujours  $\lambda > 0$ .

**Lemme 3.2.11** *On a (1)  $\Rightarrow$  (6).*

**Preuve.** Soit  $f_A(t) = f(t)1_{\{t \leq A\}}$ . D'une part  $\mu(f_A)$  croît vers  $\mu(f)$  quand  $A \rightarrow \infty$ , donc  $e^{-\mu(f_A)}$  décroît vers  $e^{-\mu(f)}$  en restant borné par 1, donc  $\mathbb{E}(e^{-\mu(f_A)}) \rightarrow \mathbb{E}(e^{-\mu(f)})$ . D'autre part  $\int_0^\infty (1 - e^{-f_A(t)}) dt \rightarrow \int_0^\infty (1 - e^{-f(t)}) dt$ , toujours d'après le théorème de Lebesgue. Par suite il suffit de montrer (3.2.13) pour chaque  $f_A$ , donc en fait pour une fonction borélienne positive  $f$  vérifiant  $f(t) = 0$  pour tout  $t > A$ , où  $A$  est un réel arbitraire. Dans ce cas,

$$\mathbb{E}(e^{-\mu(f)}) = \sum_{n=0}^{\infty} a_n, \quad \text{où } a_n = \mathbb{E}\left(e^{-\sum_{j=1}^n f(T_j)} 1_{\{N_A=n\}}\right).$$

D'après la propriété (3) du théorème 3.2.3, il vient

$$a_n = \int_{\mathbb{R}_+^{n+1}} \left(\prod_{i=0}^n \lambda e^{-\lambda s_i}\right) e^{-\sum_{j=1}^n f(s_0 + \dots + s_{j-1})} 1_{\{s_0 + \dots + s_{n-1} \leq A < s_0 + \dots + s_n\}} ds_0 \dots ds_n$$

$$\begin{aligned}
&= \lambda^n e^{-\lambda A} \int_{\mathbb{R}_+^n} e^{-\sum_{j=1}^n f(s_0+\dots+s_{j-1})} \mathbf{1}_{\{s_0+\dots+s_{n-1} \leq A\}} ds_0 \dots ds_{n-1} \\
&= \lambda^n e^{-\lambda A} \alpha, \quad \text{où } \alpha = \int_{\mathbb{R}_+^n} e^{-\sum_{j=1}^n f(t_j)} \mathbf{1}_{\{t_1 < t_2 < \dots < t_n \leq A\}} dt_1 \dots dt_n,
\end{aligned}$$

où la seconde égalité est obtenue en intégrant  $s_n \mapsto \lambda e^{-\lambda s_n}$  sur l'ensemble  $\{t - (s_0 + \dots + s_{n-1}), \infty[$ , et la seconde en faisant les changements de variables  $t_i = s_0 + \dots + s_{i-1}$ .

Soit  $\mathcal{S}_n$  l'ensemble des permutations de  $\{1, \dots, n\}$ . Si  $\sigma \in \mathcal{S}_n$ , on a  $\sum_{i=1}^n f(t_{\sigma(i)}) = \sum_{i=1}^n f(t_i)$ . Par suite en faisant les changements de variables  $t'_i = t_{\sigma(i)}$  dans l'intégrale définissant  $\alpha$ , on voit qu'on a

$$\alpha = \int_{\mathbb{R}_+^n} e^{-\sum_{j=1}^n f(t_j)} \mathbf{1}_{\{t_{\sigma(1)} < t_{\sigma(2)} < \dots < t_{\sigma(n)} \leq A\}} dt_1 \dots dt_n.$$

Comme  $\cup_{\sigma \in \mathcal{S}_n} \{(t_1, \dots, t_n) : t_{\sigma(1)} < t_{\sigma(2)} < \dots < t_{\sigma(n)} \leq A\} = ]0, A]^n$  à un ensemble de mesure nulle près pour  $dt_1 \dots dt_n$ , on en déduit que

$$\alpha = \frac{1}{n!} \int_{]0, A]^n} e^{-\sum_{i=1}^n f(t_i)} dt_1 \dots dt_n = \frac{1}{n!} \left( \int_0^A e^{-f(t)} dt \right)^n.$$

Il vient alors

$$\mathbb{E}(e^{-\mu(f)}) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda A} \left( \int_0^A e^{-f(t)} dt \right)^n = \exp \left( -\lambda A + \lambda \int_0^A e^{-f(t)} dt \right),$$

et comme  $f(t) = 0$ , donc  $e^{-f(t)} = 1$ , si  $t > A$ , on obtient (3.2.13).  $\square$

**Lemme 3.2.12** (6) implique (4) et aussi que pour tout borélien  $A$  la variable  $\mu(A)$  suit une loi de Poisson de paramètre  $\lambda m(A)$  si  $m(A) < \infty$  et est p.s. infinie si  $m(A) = \infty$ .

**Preuve.** Soit  $A_i$  des boréliens disjoint, et  $u_i \geq 0$ . Une application de (3.2.13) à la fonction  $f(t) = \sum_{i=1}^n u_i \mathbf{1}_{A_i}(t)$  donne

$$\mathbb{E} \left( \prod_{i=1}^n e^{-u_i \mu(A_i)} \right) = \prod_{i=1}^n \exp(-\lambda \mu(A_i) (1 - e^{-u_i})).$$

Cela prouve d'abord que la transformée de Laplace multi-dimensionnelle de la variable vectorielle  $(\mu(A_i))$  est le produit des transformées de Laplace uni-dimensionnelles, donc les variables  $\mu(A_i)$  sont indépendantes. Cela prouve aussi que  $\mathbb{E}(e^{-u\mu(A)}) = e^{-\lambda\mu(A)(1-e^{-u})}$  pour tout  $u \geq 0$ , d'où la dernière assertion: lorsque  $m(A) < \infty$ , car la transformée de Laplace de la loi de Poisson de paramètre  $\theta$  est  $u \mapsto \exp(-\theta(1-e^{-u}))$ ; lorsque  $m(A) = 0$ , car on a alors  $\mathbb{E}(e^{-u\mu(A)}) = 0$ , donc  $e^{-u\mu(A)} = 0$  p.s., donc  $\mu(A) = \infty$  p.s.

En particulier, on a  $\mathbb{E}(\mu(A)) = \lambda m(A)$ , ce qui achève de prouver (4).  $\square$

**Lemme 3.2.13** On a (4)  $\Rightarrow$  (1).

**Preuve.** La propriété (4) implique que pour tous  $s, t \geq 0$  la variable  $N_{t+s} - N_t = \mu([t, t+s])$  est indépendante des variables  $N_r = \mu([0, r])$  pour  $r \leq t$ , donc de la tribu  $\mathcal{F}_t = \sigma(N_r : r \leq t)$ , et par ailleurs elle est d'espérance  $\lambda s$ . Donc  $M_r = N_r - \lambda r$  vérifie

$$\mathbb{E}(M_{t+s} - M_t | \mathcal{F}_t) = \mathbb{E}(N_{t+s} - N_t | \mathcal{F}_t) - \lambda s = \mathbb{E}(N_{t+s} - N_t) - \lambda s = 0.$$

Donc  $M$  est une martingale, et il suffit d'appliquer l'implication (2)  $\Rightarrow$  (1) du théorème 3.2.3.  $\square$

**Lemme 3.2.14** *On a (1)  $\Leftrightarrow$  (5).*

**Preuve.** On a vu que (1)  $\Leftrightarrow$  (6). Par ailleurs la propriété (5) est une description de la construction de la mesure  $\mu$ , comme nous l'avons vu en (3.2.14). Il suffit donc de montrer que si on construit  $\mu$  par (3.2.14), alors cette mesure vérifie (6). Soit  $f$  une fonction borélienne positive. A cause de l'indépendance des variables  $(M_n, R_{n,m})$  et au vu de leurs lois, il est clair que

$$\begin{aligned} \mathbb{E}(e^{-\mu(f)}) &= \mathbb{E}\left(\prod_{n \geq 1} \prod_{m=1}^{M_n} e^{-f(R_{n,m})}\right) \\ &= \prod_{n \geq 1} \left( \mathbb{P}(M_n = 0) + \sum_{j \geq 1} \mathbb{P}(M_n = j) \mathbb{E}\left(\prod_{m=1}^j e^{-f(R_{n,m})}\right) \right) \\ &= \prod_{n \geq 1} \left( \mathbb{P}(M_n = 0) + \sum_{j \geq 1} \mathbb{P}(M_n = j) \prod_{m=1}^j \mathbb{E}\left(e^{-f(R_{n,m})}\right) \right) \\ &= \prod_{n \geq 1} \sum_{j \geq 0} \frac{\lambda^j m(E_n)^j}{j!} e^{-\lambda m(E_n)} \left( \frac{1}{m(E_n)} \int_{E_n} e^{-f(t)} dt \right)^j \\ &= \prod_{n \geq 1} e^{-\lambda \int_{E_n} (1 - e^{-f(t)}) dt} = e^{-\lambda \int_0^\infty (1 - e^{-f(t)}) dt}, \end{aligned}$$

d'où le résultat.  $\square$

Ce lemme achève la preuve du théorème 3.2.10.

### 3.3 Répartitions ponctuelles de Poisson

Nous allons maintenant considérer la généralisation des processus de Poisson au cas où l'espace des "temps" est remplacé par un espace abstrait. Les bonnes définitions seront alors des généralisations des propriétés (4), (5) ou (6) du théorème 3.2.10.

On considère donc un espace mesurable  $(E, \mathcal{E})$  et une mesure aléatoire  $\mu = \mu(\omega, dx)$ , et qui est la somme d'un nombre fini ou infini dénombrable (éventuellement aléatoire) de masses de Dirac. Une telle mesure s'appelle une *répartition ponctuelle* et peut évidemment s'écrire comme

$$\mu(\omega, dx) = \sum_{n \geq 1}^{M(\omega)} \varepsilon_{T_n}(\omega), \quad (3.3.1)$$

où  $M$  est à valeurs dans  $\overline{\mathbb{N}}$ , et les  $T_n$  sont à valeurs dans  $E$ . Si de plus les points sont tous distincts, i.e.  $T_n(\omega) \neq T_m(\omega)$  pour tous  $1 \leq n < m \leq M(\omega)$ , on dit que la répartition ponctuelle est *simple*.

Il faut préciser maintenant les propriétés de mesurabilité: on suppose que  $M$  et les  $T_n$  sont  $\mathcal{F}$ -mesurables. Dans ce cas,

$$\mu(f) = \sum_{n=1}^M f(T_n) \quad (3.3.2)$$

est aussi  $\mathcal{F}$ -mesurable pour toute fonction positive mesurable  $f$  sur  $(E, \mathcal{E})$ , de sorte que  $\mu$  est une mesure de transition de  $(\Omega, \mathcal{F})$  dans  $(E, \mathcal{E})$ . On peut alors définir sa *fonctionnelle de Laplace* par

$$f \mapsto \mathbb{E}(e^{-\mu(f)}) \quad \text{pour } f \text{ mesurable positive sur } (E, \mathcal{E}). \quad (3.3.3)$$

et sa *mesure intensité* (qui est une mesure sur  $(E, \mathcal{E})$ ) par

$$A \mapsto \mathbb{E}(\mu(A)) \quad \forall A \in \mathcal{E}. \quad (3.3.4)$$

**Remarque 1:** On pourrait à l'inverse partir de  $\mu$ , supposée être d'une part une répartition ponctuelle, d'autre part une mesure de transition de  $(\Omega, \mathcal{F})$  dans  $(E, \mathcal{E})$ . Peut-on écrire  $\mu$  comme (3.3.1) avec  $M$  et les  $T_n$  mesurables ? C'est évident pour  $M$ , qui vaut  $\mu(E)$ , mais pour les  $T_n$  une première difficulté tient à ce que chaque  $T_n$  n'est défini que sur l'ensemble où  $M \geq n$ ; et même si  $M(\omega) = \infty$  pour tout  $\omega$  il reste difficile de prouver qu'on peut trouver des  $T_n$  mesurables, et c'est même faux si on ne fait pas des hypothèses supplémentaires sur  $\mathcal{E}$  (à savoir, que  $\mathcal{E}$  est séparable et contient les singletons  $\{x\}$ ). Ici, nous supposons *a priori* la mesurabilité des  $T_n$ .  $\square$

**Définition 3.3.1** Une répartition ponctuelle  $\mu$  est dite

(i) *complètement aléatoire* si pour tous  $A_i \in \mathcal{E}$  deux-à-deux disjoints les variables  $(\mu(A_i))$  sont indépendantes,

(ii) *de Poisson* si sa fonctionnelle de Laplace s'écrit

$$\mathbb{E}(e^{-\mu(f)}) = \exp\left(-\int_E (1 - e^{-f(x)})\nu(dx)\right), \quad (3.3.5)$$

où  $\nu$  est une mesure sur  $(E, \mathcal{E})$ .

**Théorème 3.3.2** Si  $\mu$  est une répartition ponctuelle de Poisson associée à la mesure  $\nu$  dans (3.3.5), alors elle est complètement aléatoire. De plus pour tout  $A \in \mathcal{E}$  la variable  $\mu(A)$  est p.s. infinie si  $\nu(A) = \infty$ , et suit une loi de Poisson de paramètre  $\nu(A)$  si  $\nu(A) < \infty$ , et  $\nu$  est aussi la mesure intensité de  $\mu$ .

**Preuve.** La preuve est exactement la même que celle du lemme 3.2.12 (où  $\lambda m$  joue le rôle de  $\nu$ ).  $\square$

Ainsi, un processus de Poisson de paramètre  $\lambda$  est une répartition ponctuelle de Poisson sur  $\mathbb{R}_+$ , de mesure intensité  $\nu = \lambda m$ . Au vu du théorème 3.2.10 on peut penser que la réciproque est vraie aussi, et que tout cela est équivalent à la construction décrite dans la propriété (5) et adaptée à la situation présente. Voici une formalisation de cette construction:

*Propriété (5')*: La mesure  $\mu$  satisfait cette propriété si elle s'écrit comme dans (3.2.14), c'est-à-dire  $\mu = \sum_{n \geq 1} \sum_{m=1}^{M_n} \varepsilon_{R_{n,m}}$ , où

- (i) les variables  $M_n$  et  $R_{m,n}$  sont toutes indépendantes;
  - (ii) les variables  $M_n$  suivent des lois de Poisson de paramètres  $\nu(E_n)$ , pour une certaine mesure  $\nu$  sur  $(E, \mathcal{E})$  et une partition  $\mathcal{E}$ -mesurable  $(E_n)_{n \geq 1}$  de  $E$  avec  $\nu(E_n) < \infty$ ;
  - (iii) les variables  $R_{n,m}$  sont à valeurs dans  $E_n$  et de loi  $\nu_n$ , où on a posé  $\nu_n(A) = \frac{\nu(E_n \cap A)}{\nu(E_n)}$ .
- 

Cette propriété dépend *a priori* de la partition choisie. Elle implique évidemment que  $\nu$  est  $\sigma$ -finie. Le résultat suivant montre que dans ce cas, si la propriété est vraie pour une partition  $(E_n)$ , elle est vraie pour toute autre partition  $(E'_n)$  satisfaisant bien-sûr  $\nu(E'_n) < \infty$ .

**Théorème 3.3.3** *Si la répartition ponctuelle  $\mu$  vérifie (5') avec une certaine mesure  $\sigma$ -finie  $\nu$  sur  $(E, \mathcal{E})$ , elle est de Poisson et  $\nu$  est sa mesure intensité.*

*Réciproquement si  $\mu$  est une répartition ponctuelle de Poisson de mesure intensité  $\nu$   $\sigma$ -finie, et si  $(E, \mathcal{E})$  est un espace polonais, on a (5') pour toute partition  $(E_n)$  vérifiant  $\nu(E_n) < \infty$ .*

**Preuve.** La preuve de la partie directe est ici encore exactement la même que la preuve de (5)  $\Rightarrow$  (6) dans le lemme 3.2.14, en remplaçant partout la mesure  $\lambda m$  par la mesure  $\nu$ .

La réciproque est plus délicate, et en rapport avec la remarque 1. On fixe une partition arbitraire vérifiant les conditions requises, et on pose  $M_n = \mu(E_n)$ . En fait, si on définit les tribus  $\mathcal{F}(A)$  comme dans (3.2.12) (pour  $A \in \mathcal{E}$ ), (5') revient à dire que

a) les  $M_n$  sont de lois de Poisson de paramètres  $\nu(E_n)$  et les tribus  $\mathcal{F}(E_n)$  sont indépendantes,

b) et, conditionnellement à  $\{M_n = p\}$ , on peut trouver des variables  $R_{n,1}, \dots, R_{n,p}$  indépendantes et de loi  $\nu_n$ , telles que la restriction de  $\mu$  à  $E_n$  s'écrive  $\mu_n = \sum_{m=1}^p \varepsilon_{R_{n,m}}$ .

Il est clair que le théorème 3.3.2 entraîne (a). Mais (b) n'est pas évident. On résout la difficulté en notant que la donnée des  $R_{n,m}$  (ordonnés de manière arbitraire) satisfaisant  $\mu_n = \sum_{m=1}^p \varepsilon_{R_{n,m}}$  est équivalente à la donnée des  $\mu(A)$  pour tout  $A \subset E_n$ , donc si  $\mu$  et  $\mu'$  sont deux répartitions ponctuelles telles que les lois de  $(\mu(A_1), \dots, \mu(A_q))$  et  $(\mu'(A_1), \dots, \mu'(A_q))$  coïncident pour tout choix de  $q$  et des  $A_i \subset E_n$ , alors si  $\mu'$  vérifie la propriété (b) ci-dessus il en est de même de  $\mu$ .

Mais la loi de  $(\mu(A_1), \dots, \mu(A_q))$  est entièrement caractérisée par sa transformée de



Laplace  $q$ -dimensionnelle

$$(u_1, \dots, u_q) \mapsto \mathbb{E}(\exp(-\sum_{i=1}^q u_i \mu(A_i)))$$

(pour  $u_i \geq 0$ ), qui est en fait la fonctionnelle de Laplace pour la fonction  $\sum_{i=1}^q u_i 1_{A_i}$ . En d'autres termes la propriété (5') ne dépend que de la fonctionnelle de Laplace. Par suite la réciproque découle de la partie directe.

(*Remarque:* le fait que  $(E, \mathcal{E})$  est polonais est utilisé – de manière implicite – quand on écrit que la donnée des  $R_{n,m}$  équivaut à celle des  $\mu(A)$ ; le lecteur vérifiera que, quand  $E = \mathbb{R}_+$ , ce résultat est essentiellement trivial, dû aux faits que les  $T_n$  sont bien définis partout et qu'on peut utiliser la “fonction de répartition”  $N_t = \mu([0, t])$ : cf. lemme 3.1.3.)  
□

En ce qui concerne la réciproque du théorème 3.3.2, elle n'est pas toujours vraie non plus: il faut des hypothèses supplémentaires.

**Théorème 3.3.4** *Supposons que  $(E, \mathcal{E})$  soit un espace polonais muni de ses boréliens. Soit  $\mu$  une répartition ponctuelle simple et complètement aléatoire, dont la mesure intensité  $\nu$  est  $\sigma$ -finie et n'a pas d'atome (i.e.  $\nu(\{x\}) = 0$  pour tout  $x \in E$ ). Alors,  $\mu$  est une répartition de Poisson.*

L'hypothèse sur  $E$  est de nature technique. En revanche le fait que  $\nu$  n'ait pas d'atome est primordial (si  $a \in E$ , la mesure  $\mu(\omega, dx) = \varepsilon_a(dx)$  – déterministe – est complètement aléatoire mais n'est pas de Poisson car  $\mu(E)$  n'est ni infinie ni de loi de Poisson). La “simplicité” de  $\mu$  est aussi essentielle (si  $\mu$  est une répartition de Poisson, la mesure  $2\mu$  est aussi une répartition ponctuelle complètement aléatoire, mais pas de Poisson).

**Preuve.** i) On va d'abord montrer qu'il suffit de prouver le résultat quand  $\nu$  est une mesure finie.

Supposons le résultat vrai pour toute répartition simple complètement aléatoire de mesure intensité finie et sans atome. Soit une partition mesurable  $(E_n)$  de  $E$  telle que  $\nu(E_n) < \infty$ , et  $\mu_n$  la répartition ponctuelle définie par  $\mu_n(A) = \mu(A \cap E_n)$ , dont la mesure intensité  $\nu_n$  est clairement  $\nu_n(A) = \nu(A \cap E_n)$ . La mesure  $\nu_n$  est finie et sans atome, et la répartition  $\mu_n$  est simple et évidemment complètement aléatoire, donc on a

$$\mathbb{E}(e^{-\mu_n(f)}) = \exp\left(-\int (1 - e^{-f(x)})\nu_n(dx)\right)$$

pour toute  $f$  borélienne positive. Mais l'hypothèse sur  $\mu$  entraîne que les  $\mu_n(f)$  sont indépendantes si  $f$  est étagée, donc pour  $f$  positive mesurable quelconque par le théorème de limite monotone. Donc

$$\mathbb{E}(e^{-\mu(f)}) = \prod_n \exp\left(-\int (1 - e^{-f(x)})\nu_n(dx)\right) = \exp\left(-\int (1 - e^{-f(x)})\nu(dx)\right)$$

et donc  $\mu$  est de Poisson.

ii) On suppose donc dans la suite que  $\nu$  est une mesure finie. Soit  $A \in \mathcal{E}$ . Nous allons montrer que  $\mu(A)$  suit une loi de Poisson de paramètre  $\nu(A)$ .

L'espace  $E$  est muni d'une distance  $d$  et contient une suite  $(x_k)_{k \geq 1}$  qui est dense. Si  $B(x, \varepsilon) = \{y : d(x, y) < \varepsilon\}$  on a  $\nu(B(x, \varepsilon)) \rightarrow 0$  lorsque  $\varepsilon > 0$  puisque  $\nu(\{x\}) = 0$ , donc  $\varepsilon_{n,k} = (1/n) \wedge \sup(\varepsilon : \nu(B(x_k, \varepsilon)) \leq 1/n)$  vérifie  $\varepsilon_{n,k} > 0$  et  $\nu(B(x_k, \varepsilon_{n,k})) \leq 1/n$ . Pour tout  $x$  il existe une suite  $x_{n_m}$  tendant vers  $x$ , et on a  $\nu(B(x, \varepsilon)) \leq 1/n$  pour un  $\varepsilon \in ]0, 1/2n[$ . Si  $d(x, x_{n_m}) \leq \varepsilon/2$ , on a donc  $B(x_{n_m}, \varepsilon/2) \subset B(x, \varepsilon)$ , de sorte que  $\varepsilon_{n, n_m} > \varepsilon/2$ , et donc  $x \in B(x_{n_m}, \varepsilon_{n, n_m})$ . Par suite on voit que  $\cup_{k \geq 1} B(x_k, \varepsilon_{n, k}) = E$ .

On pose alors  $A_{n,1} = A \cap B(x_0, \varepsilon_{n,0})$  et, par récurrence sur  $k$ ,  $A_{n,k} = A \cap B(x_k, \varepsilon_{n,k}) \cap (\cup_{1 \leq i \leq k-1} A_{n,i})^c$  pour  $k \geq 2$ . On obtient ainsi une partition  $(A_{n,i})_{i \geq 1}$  de  $A$ , chaque  $A_{n,i}$  étant de diamètre  $\leq 2/n$ , mesurable, et avec  $\nu(A_{n,i}) \leq 1/n$ .

Avec la notation (3.3.1), on pose

$$\Delta = \inf(d(T_n, T_m) : 1 \leq n < m \leq M).$$

Comme  $\mathbb{E}(M) = \mathbb{E}(\mu(E)) = \nu(E) < \infty$ , on a  $M < \infty$  p.s., et la répartition étant simple on a aussi  $T_n \neq T_m$  si  $1 \leq n < m \leq M$ . Par suite  $\Delta > 0$  p.s. et, pour tout  $u \geq 0$ , on a

$$\mathbb{E} \left( e^{-u\mu(A)} 1_{\{\Delta > 2/n\}} \right) \rightarrow \mathbb{E}(e^{-u\mu(A)}) \quad (3.3.6)$$

si  $n \rightarrow \infty$ . Si  $\Delta > 2/n$ , chaque ensemble  $A_{n,i}$  contient au plus un seul point  $T_k$ , puisqu'il est de diamètre  $\leq 2/n$ , et on a  $\mu(A_{n,i}) = 1 \wedge \mu(A_{n,i})$ , tandis que  $\mu(A) = \sum_{i \geq 1} \mu(A_{n,i})$ . Par suite

$$\alpha_n := \mathbb{E} \left( e^{-u\mu(A)} 1_{\{\Delta > 2/n\}} \right) = \mathbb{E} \left( e^{-u \sum_{i \geq 1} 1 \wedge \mu(A_{n,i})} 1_{\{\Delta > 2/n\}} \right),$$

et comme  $\Delta > 0$  p.s., on obtient aussi lorsque  $n \rightarrow \infty$ :

$$\left| \alpha_n - \mathbb{E} \left( e^{-u \sum_{i \geq 1} 1 \wedge \mu(A_{n,i})} \right) \right| \leq \mathbb{P}(\Delta > \frac{2}{n}) \rightarrow 0. \quad (3.3.7)$$

Comme  $\mu$  est complètement aléatoire, les variables  $1 \wedge \mu(A_{n,i})$  sont indépendantes pour  $i = 1, \dots$ , et prennent la valeur 1 avec la probabilité  $\delta_i^n = \mathbb{P}(\mu(A_{n,i}) \geq 1) \leq \mathbb{E}(\mu(A_{n,i})) = \nu(A_{n,i}) \leq \frac{1}{n}$ , et 0 avec la probabilité  $1 - \delta_i^n$ . Par suite

$$\beta_n := \mathbb{E} \left( e^{-u \sum_{i \geq 1} 1 \wedge \mu(A_{n,i})} \right) = \prod_{i \geq 1} (1 - \delta_i^n (1 - e^{-u})).$$

Si  $0 \leq x \leq 1/n$  on a  $-x(1+1/n) \leq \log(1-x) \leq -x$  si  $n \geq 2$ , donc comme  $\delta_i^n (1 - e^{-u}) \leq 1/n$  il vient

$$\exp \left( -\left(1 + \frac{2}{n}\right)(1 - e^{-u}) \sum_i \delta_i^n \right) \leq \beta_n \leq \exp \left( -(1 - e^{-u}) \sum_i \delta_i^n \right).$$

En combinant ceci avec (3.3.6) et (3.3.7), on obtient que  $\sum_i \delta_i^n$  converge vers une limite finie  $\delta$  et que  $\mathbb{E}(e^{-\mu(A)}) = \exp(-(1 - e^{-u})\delta)$ . Ceci étant vrai pour tout  $u \geq 0$ , on en déduit que  $\mu(A)$  suit une loi de Poisson de paramètre  $\delta$ , et comme de plus  $\mathbb{E}(\mu(A)) = \nu(A)$  on a nécessairement  $\delta = \nu(A)$ .

iii) Si  $f$  est mesurable positive étagée, de la forme  $f = \sum_{i=1}^m u_i 1_{A_i}$  pour des  $A_i \in \mathcal{E}$  deux-à-deux disjoints et  $u_i \geq 0$ , les variables  $\mu(A_i)$  sont indépendantes (puisque  $\mu$  est complètement aléatoire) et de lois de Poisson de paramètres  $\nu(A_i)$  d'après ii), donc

$$\begin{aligned} \mathbb{E}(e^{-\mu(f)}) &= \prod_{i=1}^m \mathbb{E}(e^{-u_i \mu(A_i)}) = \prod_{i=1}^m e^{-(1-e^{-u_i})\nu(A_i)} \\ &= e^{-\sum_{i=1}^m \int (1-e^{-f(x)}) 1_{A_i}(x) \nu(dx)} = e^{-\int (1-e^{-f(x)}) \nu(dx)}. \end{aligned}$$

Toute fonction mesurable positive  $f$  étant limite croissante de fonctions étagées, Par limite monotone les deux membres extrêmes ci-dessus sont égaux pour  $f$ , et on a le résultat.  $\square$

### 3.4 Processus de Poisson composé

**Définition 3.4.1** On appelle  $(\mathcal{F}_t)$ -processus de Poisson composé un processus défini sur un espace filtré  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ , à valeurs dans  $\mathbb{R}^d$ , qui est un  $(\mathcal{F}_t)$ -processus de Lévy et dont les trajectoires  $t \mapsto X_t(\omega)$  sont constantes par morceaux. Si  $\mathcal{F}_t = \sigma(X_s : s \leq t)$  on dit simplement que  $X$  est un processus de Poisson composé.

On associe à  $\mathbf{X}$  ses temps de saut successifs  $T_n$

$$T_0 = 0, \quad T_{n+1} = \begin{cases} \inf(t : t > T_n, X_t \neq X_{T_n}) & \text{si } T_n < \infty \\ \infty & \text{si } T_n = \infty, \end{cases} \quad (3.4.1)$$

( $T_0$  n'est pas un temps de saut; comme d'habitude  $T_\infty = \lim_n T_n$ ) et les tailles  $Y_n$  des sauts correspondants, c'est-à-dire

$$Y_n = \begin{cases} X_{T_n} - X_{T_n-} = X_{T_n} - X_{T_{n-1}} & \text{si } T_n < \infty \\ 0 & \text{si } T_n = \infty, \end{cases} \quad (3.4.2)$$

Par ailleurs, on dit qu'un processus  $(X_t)$  vérifie la propriété  $A$  si on peut l'écrire

$$X_t = \sum_{n \geq 1} Z_n 1_{\{U_n \leq t\}}, \quad (3.4.3)$$

où les  $(U_n)_{n \geq 1}$  sont les points d'un processus de Poisson de paramètre  $\lambda \geq 0$  et où les variables  $Z_n$  sont indépendantes, de même loi  $G$  et indépendantes de la suite  $(T_n)$

**Théorème 3.4.2** (i) Si le processus  $(X_t)$  vérifie la propriété  $A$  avec  $\lambda$  et  $G$ , c'est un processus de Poisson composé. De plus la fonction caractéristique de la variable  $X_t$  est

$$\mathbb{E}(e^{i \langle u, X_t \rangle}) = \exp -\lambda t \int (1 - e^{i \langle u, x \rangle}) G(dx), \quad (3.4.4)$$

et la répartition ponctuelle  $\mu = \sum_{n \geq 1} 1_{\{U_n < \infty\}} \varepsilon_{(U_n, Z_n)}$  est de Poisson (sur  $\mathbb{R}_+ \times \mathbb{R}^d$ ), simple, et de mesure intensité  $\nu(dt, dx) = \lambda dt \otimes G(dx)$ .

(ii) Si le processus  $(X_t)$  est de Poisson composé, il vérifie la propriété  $A$  avec  $U_n = T_n$  et  $Z_n = Y_n$ . Dans ce cas on a  $G(\{0\}) = 0$  si  $\lambda > 0$ , et  $G = \varepsilon_0$  si  $\lambda = 0$ .

Le lecteur remarquera la différence entre (i) et (ii): si on a (i), les  $U_n$  ne sont pas forcément des temps de saut du processus  $X$ , puisqu'on peut avoir  $Z_1 = 0$  avec une probabilité positive. On peut donc "réaliser" le même processus de Poisson composé sous la forme (3.4.3) de différentes manières. Lorsque  $T_n = U_n$  et  $Z_n = Y_n$ , on dit qu'on a la *réalisation minimale*. Le premier temps de saut, par exemple, du processus défini par (3.4.3) est  $T_1 = \inf(U_n : n \geq 1, Z_n \neq 0)$ .

Lorsque le processus de Poisson composé est à valeurs réelles et a des sauts d'amplitude 1, i.e. si  $Y_n = 1$  sur  $\{T_n < \infty\}$ , c'est un processus de Poisson: le théorème précédent découle alors des théorèmes 3.2.3 et 3.2.10.

Noter que pour un processus de Poisson composé avec sa réalisation minimale, soit on a  $\lambda = 0$  et alors  $X \equiv 0$ , soit on a  $\lambda > 0$  et  $T_n < \infty$  pour tout  $n$ , et  $T_\infty = \infty$ : les variables  $Y_n$  sont alors bien définies partout (ou au moins presque partout).

La démonstration du théorème suit encore d'une série de lemmes simples.

**Lemme 3.4.3** *Si  $X$  est un processus de Poisson composé, les variables  $T_1$  et  $Y_1$  sont indépendantes, et  $T_1$  est de loi exponentielle de paramètre  $\lambda \geq 0$  (rappelons que si  $\lambda = 0$ , cela signifie que  $T_1 = \infty$  p.s.). De plus si  $\lambda > 0$  on a  $G(\{0\}) = 0$ , et si  $\lambda = 0$  on a  $G = \varepsilon_0$ , où  $G$  désigne la loi de  $Y_1$ .*

**Preuve.** La preuve est essentiellement la même que celle du théorème 2.2.3-(1). On peut toujours supposer que l'espace  $\Omega$  est muni de translations  $(\theta_t)$  vérifiant (3.1.3), donc  $T_1 = t + T_1 \circ \theta_t$  et  $Y_1 = Y_1 \circ \theta_t$  sur l'ensemble  $\{T_1 > t\}$ . Par suite d'après (3.1.5) on a pour  $A \in \mathcal{R}^d$  et  $s, t \geq 0$ :

$$\mathbb{P}(T_1 > t+s, Y_1 \in A) = \mathbb{P}(T_1 > t, T_1 \circ \theta_t > s, Y_1 \circ \theta_t \in A) = \mathbb{P}(T_1 > t) \mathbb{P}(T_1 > s, Y_1 \in A).$$

Si  $s = 0$  on obtient  $\mathbb{P}(T_1 > t, Y_1 \in A) = \mathbb{P}(T_1 > t) \mathbb{P}(Y_1 \in A)$  (puisque  $T_1 > 0$  identiquement), donc  $T_1$  et  $Y_1$  sont indépendants. Si  $A = \mathbb{R}^d$  on obtient  $\mathbb{P}(T_1 > t+s) = \mathbb{P}(T_1 > t) \mathbb{P}(T_1 > s)$ , donc  $T_1$  suit une loi exponentielle dont on note  $\lambda$  le paramètre. Soit aussi  $G$  la loi de  $Y_1$ .

Enfin si  $\lambda = 0$  on a  $T_1 = \infty$  p.s., donc  $Y_1 = 0$  p.s. (cf. (3.4.2)). Si  $\lambda > 0$  on a  $T_1 < \infty$  p.s., donc  $Y_1 \neq 0$  p.s. et  $G(\{0\}) = 0$ .  $\square$

**Lemme 3.4.4** *On a la partie (ii) du théorème 3.4.2.*

**Preuve.** Reprenons les notations du lemme précédent. Si  $\lambda = 0$ , il n'y a rien à montrer puisque  $T_n = \infty$  et  $Y_n = 0$  pour tout  $n \geq 1$ . On suppose donc  $\lambda > 0$ , donc  $T_1 < \infty$ .

On a  $T_{n+1} - T_n = T_1 \circ \theta_{T_n}$  et  $Y_{n+1} = Y_1 \circ \theta_{T_n}$  sur  $\{T_n < \infty\}$ . D'après la propriété (3.1.5) on a donc que  $(T_{n+1} - T_n, Y_{n+1})$  a même loi que  $(T_1, Y_1)$ , conditionnellement par rapport à  $\sigma(T_i, Y_i : i \leq n)$  et en restriction à  $\{T_n < \infty\}$ . On en déduit d'abord par récurrence sur  $n$  que  $T_n < \infty$  p.s. pour tout  $n$ , puis que la suite  $(T_n)_{n \geq 1}$  et les variables  $Y_i$  sont toutes indépendantes, que les  $Y_i$  ont toutes même loi  $G$ , et que les  $(T_n)$  sont les temps de saut d'un processus de Poisson de paramètre  $\lambda$  (appliquer (3)  $\Rightarrow$  (1) dans le théorème 3.2.3).  $\square$

**Lemme 3.4.5** *On a la partie (i) du théorème 3.4.2.*

**Preuve.** Soit  $X$  vérifiant la propriété A avec  $\lambda$  et  $G$ , et soit la répartition ponctuelle  $\mu = \sum_{n \geq 1} 1_{\{U_n < \infty\}} \varepsilon_{(U_n, Z_n)}$ . Cette répartition est évidemment simple. On va voir qu'elle satisfait la propriété (5') avec  $\nu(dt, dx) = \lambda dt \otimes G(dx)$ . On prend la partition de  $\mathbb{R}_+ \times \mathbb{R}^d$  constituée des  $E_n = [n-1, n[ \times \mathbb{R}^d$ , et  $\nu_n(A) = \frac{\nu(A \cap E_n)}{\nu(E_n)}$  (on a  $\nu(E_n) = \lambda$  ici).

D'après le théorème 3.2.10 (ou le théorème 3.3.3 appliqué à la répartition ponctuelle de Poisson  $\sum_{n \geq 1} 1_{\{U_n < \infty\}} \varepsilon_{U_n}$  sur  $\mathbb{R}_+$ ), les  $U_n$  peuvent être construites ainsi: on se donne des variables indépendantes  $M_n$  de loi de Poisson de paramètre  $\lambda$ , et des variables  $(V_{n,m})_{n,m \geq 1}$  indépendantes entre elles et des  $M_n$ , telles que chaque  $V(n, m)$  soit uniformément répartie sur  $]n-1, n]$ . On a alors

$$\sum_{n \geq 1} 1_{\{U_n < \infty\}} \varepsilon_{U_n} = \sum_{n \geq 1} \sum_{m=1}^{M_n} \varepsilon_{V_{n,m}}.$$

Ensuite on considère une double suite  $(Z'_{n,m})$  de variables de loi  $G$ , indépendantes entre elles et indépendantes des  $M_n$  et des  $V_{n,m}$ . Une manière de construire la mesure  $\mu$  consiste alors à poser

$$\mu = \sum_{n \geq 1} \sum_{m=1}^{M_n} \varepsilon_{(V_{n,m}, Z'_{n,m})}.$$

Il est clair que les variables  $((V_{n,m}, Z'_{n,m}))_{n,m \geq 1}$  sont indépendantes entre elles et indépendantes des  $M_n$ , et chaque variable  $(V_{n,m}, Z'_{n,m})$  est de loi  $\nu_n$ . Cela démontre que  $\mu$  satisfait la propriété (5') avec la mesure  $\nu$ , et par suite  $\mu$  est de Poisson d'intensité  $\nu$ . Enfin, si  $u \in \mathbb{R}^s$  on a

$$\begin{aligned} \mathbb{E}(e^{i \langle u, X_t \rangle}) &= \sum_{n \geq 0} \mathbb{E} \left( e^{i \sum_{m=1}^n \langle u, Y_m \rangle} 1_{\{U_n \leq t < U_{n+1}\}} \right) \\ &= \sum_{n \geq 0} \left( \int e^{i \langle u, x \rangle G(dx)} \right)^n \mathbb{P}(U_n \leq t < U_{n+1}) \\ &= \sum_{n \geq 0} \left( \int e^{i \langle u, x \rangle G(dx)} \right)^n \frac{\lambda^n t^n}{n!} e^{-\lambda t} \\ &= e^{-\lambda t} \int (1 - e^{i \langle u, x \rangle}) G(dx), \end{aligned}$$

d'où (3.4.4). □

Ces trois lemmes nous donnent le théorème 3.4.1.

Il existe aussi une caractérisation des processus de Poisson composé analogue à la propriété (2) du théorème 3.2.3 pour les processus de Poisson. Soit  $(X_t)$  un processus sur  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  à valeurs dans  $\mathbb{R}^d$ , continu à droite et dont les trajectoires sont constantes par morceaux. On lui associe les  $T_n$  et les  $Y_n$  par (3.4.1) et (3.4.2). Pour tout  $A \in \mathcal{R}^d$  on pose

$$N_t^A = \sum_{n \geq 1} 1_A(Y_n) 1_{\{T_n \leq t\}}. \quad (3.4.5)$$

**Proposition 3.4.6** *Si  $(X_t)$  est un  $(\mathcal{F}_t)$ -processus de Poisson composé, auquel on associe  $\lambda$  et  $G$  comme dans le théorème 3.4.2-(ii), alors pour tout  $A \in \mathcal{R}^d$  le processus  $M_t^A = N_t^A - \lambda G(A)t$  est une  $(\mathcal{F}_t)$ -martingale.*

**Preuve.** Encore une fois, on peut supposer qu'il existe des translations  $(\theta_t)$  vérifiant (3.1.3). Le processus  $N^A$  est un processus de comptage qui vérifie  $N_{t+s}^A - N_t^A = N_s^A \circ \theta_t$  et qui est clairement adapté à la filtration  $(\mathcal{F}_t)$ . Donc le théorème 3.2.3 implique que c'est un  $(\mathcal{F}_t)$ -processus de Poisson dont on note  $\lambda_A$  le paramètre, et  $N_t^A - \lambda_A t$  est une martingale. Il suffit donc de vérifier que  $\lambda_A = \lambda G(A)$ . Mais  $\lambda_A = \mathbb{E}(N_1^A)$ , et si  $\mu$  est la répartition ponctuelle  $\mu = \sum_{n \geq 1} 1_{\{T_n < \infty\}} \varepsilon_{(T_n, Y_n)}$  on a  $N_1^A = \mu([0, 1] \times A)$ , qui est d'espérance  $\nu([0, 1] \times A) = \lambda G(A)$ .  $\square$

En fait la propriété précédente caractérise les processus de Poisson composé, et on a le résultat suivant, qui se montre comme le lemme 3.2.9 (en un peu plus compliqué):

**Théorème 3.4.7** *Soit  $(X_t)$  un processus adapté sur  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  dont les trajectoires sont nulles en 0, continues à droite et constantes par morceaux, et  $N^A$  les processus associés par (3.4.5). S'il existe  $\lambda \geq 0$  et une probabilité  $G$  sur  $\mathbb{R}^d$  tels que pour tout  $A \in \mathcal{R}^d$  le processus  $N_t^A - \lambda G(A)t$  soit une  $(\mathcal{F}_t)$ -martingale, alors le processus  $(X_t)$  est un  $(\mathcal{F}_t)$ -processus de Poisson composé.*

## Chapitre 4

# Les processus de renouvellement

### 4.1 Introduction

Ce qu'on désigne sous le terme de "processus de renouvellement" est une généralisation des processus de Poisson, obtenue en remplaçant la loi exponentielle des intervalles "inter-sauts" par une loi  $F$  arbitraire. De manière précise, on pose:

**Définition 4.1.1** Un *processus de renouvellement* est un processus de comptage  $N = (N_t)$ , donc de la forme

$$N_t = \sum_{n \geq 1} 1_{\{T_n \leq t\}}, \quad T_n = S_0 + \dots + S_{n-1} \quad \text{si } n \geq 1 \quad (4.1.1)$$

(comme en (3.2.2)), où les  $S_n$  sont *indépendantes* et  $S_0$  est de loi  $F_0$  et les  $S_n$  pour  $n \geq 1$  sont de loi commune  $F$ , avec  $F_0$  et  $F$  portées par  $]0, \infty[$ .

Un peu à la manière des processus de Markov, on parle d'un processus de renouvellement *associé* à  $F$ , la loi  $F_0$  de  $S_0$  jouant le rôle de "condition initiale" et pouvant varier. On peut réaliser un processus de renouvellement de loi  $F$  sur l'espace canonique  $\Omega = \mathbb{R}_+^N$  muni de la tribu produit  $\mathcal{F}$ , des variables canoniques  $S_n((x_0, x_1, \dots) = x_n$ , et de la probabilité produit  $\mathbb{P}_{F_0} = F_0 \otimes F \otimes \dots \otimes F \otimes \dots$ . Dans la suite on écrira toujours  $\mathbb{P}_{F_0}$  (même si on n'est pas sur l'espace canonique) de façon à bien préciser la loi initiale, qui peut changer, alors que la loi  $F$  est supposée *fixée* une fois pour toutes. Lorsque  $F_0 = \varepsilon_x$  est la masse de Dirac en  $x \in ]0, \infty[$ , on écrit aussi  $\mathbb{P}_x$ .

Un processus de Poisson de paramètre  $\lambda$  est un processus de renouvellement si  $\lambda > 0$ , auquel cas  $F_0 = F$  est la loi exponentielle de paramètre  $\lambda$ ; inversement un processus de renouvellement ayant ces lois est un processus de Poisson.

La terminologie "renouvellement" provient de l'application suivante: un composant d'une machine a une durée de vie aléatoire de loi  $F$ ; chaque fois qu'il tombe en panne on le remplace; les durée de vie successives sont notées  $S_1, S_2, \dots$ ; la durée de vie "résiduelle"  $S_0$  du composant en place à l'instant initial 0 est de loi  $F_0$  (qui dépend du temps pendant lequel ce composant a fonctionné avant l'instant pris comme instant initial): alors,  $N_t$

représente le nombre de fois où il a fallu changer le composant avant l'instant  $t$ . Les temps  $T_n$  auxquels on change le composant sont appelés les *temps de renouvellement*.

On remarque que bien-sûr  $T_n < \infty$  pour tout  $n$ . A l'inverse, si

$$m = \mathbb{E}(S_1) = \int xF(dx) \quad (4.1.2)$$

(qui peut être fini ou infini), la loi des grands nombres pour les variables positives entraîne que  $(S_1 + \dots + S_n)/n$  converge p.s. vers  $m$ , donc

$$T_n \rightarrow \infty \quad \text{p.s.} \quad (4.1.3)$$

On a donc presque sûrement  $N_t < \infty$  pour tout  $t$ , et  $N_t \uparrow \infty$  quand  $t \rightarrow \infty$ .

## 4.2 Quelques propriétés élémentaires

### 4.2.1 Le potentiel

D'un certain point de vue, le processus de renouvellement est entièrement déterminé par la suite  $(T_n)$ . Vu la définition 4.1.1, il est évident que la suite  $(T_n)_{n \geq 1}$  constitue sous la probabilité  $\mathbb{P}_{F_0}$  ce qu'on appelle une *marche aléatoire*, c'est-à-dire une *chaîne de Markov* de probabilité de transition

$$P(x, A) = \int 1_A(x + y)F(dy), \quad \text{i.e.} \quad P(x, \cdot) = \varepsilon_x \star F, \quad (4.2.1)$$

et aussi de loi initiale  $F_0$  (= la loi de  $T_1$ , puisque nous partons du temps 1 pour la chaîne  $(T_n)$ ). Comme d'habitude pour une chaîne de Markov, le potentiel est le noyau  $U = \sum_{n \geq 0} P^n$ , et (comme dans la situation des processus de Lévy: une marche aléatoire n'est rien d'autre que la version "temps discret" d'un processus de Lévy) la puissance  $P^n$  est  $P^n(x, \cdot) = \varepsilon_x \star F^{n\star}$ . En particulier la mesure  $U(0, dx)$ , notée simplement  $U(dx)$ , est

$$U = \sum_{n \geq 0} F^{n\star}, \quad \text{i.e.} \quad U(A) = \mathbb{E}_{F_0} \left( \sum_{n \geq 1} 1_A(T_n - T_1) \right). \quad (4.2.2)$$

Cette mesure s'appelle le *potentiel recentré* du processus de renouvellement, et ne dépend pas de la mesure initiale.

**Lemme 4.2.1** *Pour tout  $t > 0$  il existe une constante  $\alpha_t < \infty$  telle que*

$$U([s, s + t]) \leq \alpha_t \quad \forall s \geq 0. \quad (4.2.3)$$

*En particulier  $\mathbb{E}_{F_0}(N_t) \leq U([0, t]) < \infty$  pour tout  $t < \infty$ .*

**Preuve.** Si on montre (4.2.3) pour un  $t > 0$  et un  $\alpha_t$ , on l'aura évidemment pour tout  $t'$ , avec la constante  $\alpha_{t'} = n\alpha_t$  si  $(n - 1)t \leq t' < nt$ . On choisit alors  $t > 0$  tel que  $\beta = F(]t, \infty]) > 0$ . On a



$$\begin{aligned}
U([s, s+t]) &= \mathbb{E}_{F_0} \left( \sum_{n \geq 1} 1_{\{T_1+s \leq T_n \leq T_1+s+t\}} \right) \\
&= \sum_{n \geq 1} \mathbb{P}_{F_0}(T_1+s \leq T_n \leq T_1+s+t). \tag{4.2.4}
\end{aligned}$$

On a aussi  $\{T_1+s \leq T_n \leq T_1+s+t\} \cap \{S_n > t\} \subset \{T_1+s \leq T_n \leq T_1+s+t < T_{n+1}\}$ . Comme  $S_n$  est indépendante de  $T_1, \dots, T_n$ , il vient alors

$$\mathbb{P}_{F_0}(T_1+s \leq T_n \leq T_1+s+t) \beta \leq \mathbb{P}_{F_0}(T_1+s \leq T_n \leq T_1+s+t < T_{n+1}).$$

En remarquant que les  $\{T_1+s \leq T_n \leq T_1+s+t < T_{n+1}\}$  pour  $n \geq 1$  constituent une partition de l'ensemble  $\Omega$ , en sommant ces inégalités sur  $n \geq 1$  on déduit alors de (4.2.4) que  $U([s, s+t]) \beta \leq 1$ , et comme  $\beta > 0$  on a (4.2.3) pour le  $t$  choisi ci-dessus avec  $\alpha_t = 1/\beta < \infty$ .

Enfin, (4.2.4) appliqué avec  $s = 0$  donne

$$U([0, t]) = \mathbb{E}_{F_0} \left( \sum_{n \geq 1} 1_{\{T_1 \leq T_n \leq T_1+t\}} \right) \geq \mathbb{E}_{F_0} \left( \sum_{n \geq 1} 1_{\{T_n \leq t\}} \right) = \mathbb{E}_{F_0}(N_t),$$

d'où le dernier résultat. □

En particulier, la mesure  $U$  est une mesure de Radon, donc  $\sigma$ -finie, et on peut en faire le produit de convolution avec n'importe quelle probabilité sur  $\mathbb{R}_+$ . Au vu de la première partie de (4.2.2), il est alors évident que  $U$  vérifie l'équation du renouvellement associée à  $F$  pour les mesures, c'est-à-dire

$$U = \varepsilon_0 + F \star U. \tag{4.2.5}$$

Si maintenant on tient compte de la loi initiale, on obtient le *potentiel* du processus de renouvellement, qui est la mesure

$$U_{F_0} = F_0 \star U, \quad \text{i.e.} \quad U_{F_0}(A) = \mathbb{E}_{F_0} \left( \sum_{n \geq 1} 1_A(T_n) \right) \tag{4.2.6}$$

En particulier  $U_{\varepsilon_x}$  n'est autre que la mesure  $U(x, \cdot)$ .

L'objectif de ce chapitre est essentiellement d'étudier le comportement des mesures  $U$  ou  $U_{F_0}$  "à l'infini", c'est-à-dire par exemple les limites éventuelles de  $U(t+A)$  où  $t+A = \{t+x : x \in A\}$  lorsque  $t \rightarrow \infty$ . On étudiera aussi le comportement asymptotique des processus qu'on va maintenant introduire.

## 4.2.2 Le processus des âges

On part d'un processus de renouvellement  $(N_t)$  de loi inter-sauts  $F$  et de loi initiale  $F_0$ . On pose aussi  $T_0 = 0$ , et

$$\left. \begin{aligned}
V_t &= \inf(T_n - t : n \geq 1, T_n > t) = \inf(s : N_s > N_t) - t \\
W_t &= \sup(t - T_n : n \geq 0, T_n \leq t) = \begin{cases} t & \text{si } N_t = 0 \\ t - \sup(s : N_s < N_t) & \text{si } N_t \geq 1. \end{cases}
\end{aligned} \right\} \tag{4.2.7}$$

Le processus  $(V_t)$  s'appelle le processus des "temps avant renouvellement", tandis que  $(W_t)$  s'appelle le "processus des âges".

Nous allons d'abord montrer que les deux processus  $(V_t)$  et  $(W_t)$  vérifient la propriété de Markov, relativement aux deux filtrations qu'ils engendrent:

$$\mathcal{G}_t = \sigma(V_s : s \leq t), \quad \mathcal{F}_t = \sigma(W_s : s \leq t). \quad (4.2.8)$$

**Lemme 4.2.2** (i) *Les processus  $(V_t)$  et  $(W_t)$  sont continus à droite et "linéaires par morceaux", de pentes respectives  $-1$  et  $+1$ , et à valeurs dans  $]0, \infty[$  et  $[0, \infty[$  respectivement.*

(ii) *On a les inclusions et égalités  $\mathcal{F}_t = \sigma(N_s : s \leq t) \subset \mathcal{G}_t = \mathcal{F}_t \vee \sigma(T_{N_t+1})$ .*

**Preuve.** (i) est évident. Posons  $\mathcal{F}'_t = \sigma(N_s : s \leq t)$ . Comme  $\{W_t > x\}$  est vide si  $x > t$  et égale  $\{N_t = N_{t-x}\}$  si  $x \leq t$ , il est clair que  $\mathcal{F}_t \subset \mathcal{F}'_t$ .

Montrons ensuite  $\mathcal{F}'_t \subset \mathcal{F}_t$ . Etant donné (4.1.1), il suffit de prouver que chaque  $T_n$  est un temps d'arrêt pour la filtration  $(\mathcal{F}_t)$ . C'est évident pour  $n = 0$ , et supposons le vrai pour un certain  $n$ . On a  $\{T_{n+1} \leq t\} = \{T_n < t\} \cap \{W_t \leq (t - T_n)^+\}$  (où  $x^+$  désigne la partie positive de  $x \in \mathbb{R}$ ), et  $\{T_n < s\} \in \mathcal{F}_s$  pour tout  $s$  par hypothèse, donc la variable  $(t - T_n)^+$  est  $\mathcal{F}_t$ -mesurable: donc  $\{T_{n+1} \leq t\} \in \mathcal{F}_t$ , et  $T_{n+1}$  est aussi un  $(\mathcal{F}_t)$ -temps d'arrêt.

Puis on prouve que  $\mathcal{F}'_t \subset \mathcal{G}_t$ , et comme ci-dessus il suffit de montrer que chaque  $T_n$  est un  $(\mathcal{G}_t)$ -temps d'arrêt. C'est évident pour  $n = 0$ , et supposons le vrai pour un certain  $n$ . Comme  $(V_t)$  est continu à droite, on a que  $V_{T_n}$  est  $\mathcal{G}_{T_n+}$ -mesurable (cf. la preuve du théorème 2.1.5). On a aussi  $T_{n+1} = T_n + V_{T_n}$  d'après (4.2.7). Par suite  $\{T_{n+1} < t\} = \cup_{r \in [0, t] \cap \mathbb{Q}} (\{T_n < r\} \cap \{V_{T_n} < t - r\})$ , qui appartient clairement à  $\mathcal{G}_t$ , et  $T_{n+1}$  est un  $(\mathcal{G}_t)$ -temps d'arrêt.

Ensuite, on a  $T_{N_t+t} = t + V_t$ , donc  $\sigma(T_{N_t+1}) \subset \mathcal{G}_t$ . Il reste donc à montrer que  $\mathcal{G}_t \subset \mathcal{F}_t \vee \sigma(T_{N_t+1})$ . Mais si  $s \leq t$  on a  $\{V_s \leq x\} = \cup_{n \geq 0} \{T_n \leq s < T_{n+1} \leq x + s\}$ , qui appartient à  $\mathcal{F}_{s+x}$ , donc à  $\mathcal{F}_t$  pour tout  $x \in ]0, t - s]$ ; enfin si  $x > t - s$  il vient  $\{t - s < V_s \leq x\} = \{V_s \leq t - s\}^c \cap \{T_{N_t+1} \leq x\}$ , qui est dans  $\mathcal{F}_t \vee \sigma(T_{N_t+1})$ . Ainsi  $V_s$  est  $\mathcal{F}_t \vee \sigma(T_{N_t+1})$ -mesurable pour tout  $s \leq t$ .  $\square$

Remarquons que l'inclusion  $\mathcal{F}_t \subset \mathcal{G}_t$  est stricte. En effet, la tribu  $\mathcal{F}_0 = \sigma(N_0)$  est la tribu triviale, tandis que la variable  $T_1 = V_0$  est  $\mathcal{G}_0$ -mesurable.

**Proposition 4.2.3** *Sous chaque probabilité  $\mathbb{P}_{F_0}$  le processus  $(V_t)$  est de loi initiale  $F_0$  et est markovien relativement à la filtration  $(\mathcal{G}_t)$  et au semi-groupe de transition  $(P_t^V)$  sur  $]0, \infty[$  donné par*

$$P_t^V(x, A) = \begin{cases} 1_A(x - t) & \text{si } x > t \\ \mathbb{E}_F(1_A(V_{t-x})) & \text{si } 0 < x \leq t. \end{cases} \quad (4.2.9)$$

**Preuve.** On a  $V_0 = T_1$ , d'où la première assertion. Comme  $(V_t)$  est continu à droite, la fonction  $t \mapsto \mathbb{E}_F(1_A(V_t))$  est mesurable. Il en découle que (4.2.9) définit pour chaque  $t$  une probabilité de transition sur  $]0, \infty[$ .

On va montrer que pour tous  $n \in \mathbb{N}$ ,  $s, t \geq 0$ , et  $f_n$  et  $g$  boréliennes bornées,

$$\mathbb{E}_{F_0} \left( f_n(T_1, \dots, T_{n+1}) g(V_{t+s}) 1_{\{N_s=n\}} \right) = \mathbb{E}_{F_0} \left( f_n(T_1, \dots, T_{n+1}) 1_{\{N_s=n\}} P_t^V g(V_s) \right) \quad (4.2.10)$$

En sommant sur  $n \in \mathbb{N}$  et en se rappelant que, d'après le lemme précédent, toute variable  $\mathcal{G}_s$ -mesurable est en restriction à l'ensemble  $\{N_s = n\}$  une fonction de  $(T_1, \dots, T_{n+1})$ , on en déduira la propriété de Markov cherchée, soit  $\mathbb{E}_{F_0}(g(V_{s+t})|\mathcal{G}_s) = P_t^V g(V_s)$ ; enfin, en prenant l'espérance dans cette égalité, lorsque  $F_0 = \varepsilon_x$ , cela donne aussi la propriété de semi-groupe pour  $(P_t^V)$ .

Il suffit bien-sûr de montrer (4.2.10) séparément lorsque la fonction  $f_n$  est nulle sur l'ensemble  $\{(t_1, \dots, t_{n+1}) : t_{n+1} \leq s+t\}$ , ou sur son complémentaire. Dans le premier cas l'égalité (4.2.10) est évidente, puisque sur  $\{N_s = n, T_{n+1} > s+t\}$  on a  $V_{s+t} = V_s - t > 0$ , donc  $P_t^V g(V_s) = g(V_s - t) = g(V_{s+t})$ . On se place maintenant dans le second cas, ce qui revient à ajouter dans les espérances des deux membres de (4.2.10) l'indicatrice  $1_{\{T_{n+1} \leq s+t\}}$ .

On notera  $\widehat{S}_n$  la variable  $\widehat{S}_n = (S_n, S_{n+1}, \dots)$ , qui prend ses valeurs dans l'espace  $(E, \mathcal{E}) = (\mathbb{R}^{\mathbb{N}}, \mathcal{R}^{\mathbb{N} \otimes})$  et dont la loi, sous chaque  $\mathbb{P}_{F_0}$ , est  $Q = F^{\mathbb{N} \otimes}$ . Comme toute variable aléatoire,  $V_u$  peut s'écrire  $V_u = h_u(T_1, \widehat{S}_1)$  pour une fonction mesurable  $h_u$  sur l'espace produit  $(\mathbb{R} \times E, \mathcal{R} \otimes \mathcal{E})$ . On remarque aussi que sur l'ensemble  $\{N_s = n, T_{n+1} \leq s+t\}$  on a  $V_{s+t} = h_{s+t-T_{n+1}}(S_{n+1}, \widehat{S}_{n+2})$ . Comme le couple  $(S_{n+1}, \widehat{S}_{n+2})$  est indépendant de  $(T_1, \dots, T_{n+1})$  et de loi le produit  $F \otimes Q$ , le membre de gauche de (4.2.10) s'écrit alors

$$\begin{aligned} & \mathbb{E}_{F_0} \left( f_n(T_1, \dots, T_{n+1}) 1_{\{T_n \leq s < T_{n+1} \leq s+t\}} g \circ h_{s+t-T_{n+1}}(S_{n+1}, \widehat{S}_{n+2}) \right) \\ &= \mathbb{E}_{F_0} \left( f_n(T_1, \dots, T_{n+1}) 1_{\{T_n \leq s < T_{n+1} \leq s+t\}} \int g \circ h_{s+t-T_{n+1}}(s_0, \widehat{s}_1) F(ds_0) Q(d\widehat{s}_1) \right) \end{aligned}$$

L'intégrale ci-dessus vaut en fait  $P_t^V g(V_s)$ , puisque  $\int g \circ h_{t-x}(s_0, \widehat{s}_1) F(ds_0) Q(d\widehat{s}_1) = P_t^V g(x)$  si  $x \leq t$  d'après (4.2.9), et  $V_s = T_{n+1} - s$  sur  $\{N_s = n\}$ . On en déduit alors l'égalité (4.2.10), et la preuve est terminée.  $\square$

Pour le processus des âges  $(W_t)$  la situation est un peu plus compliquée. C'est toujours un processus de Markov, mais en général non-homogène. L'homogénéité n'est assurée que lorsque la loi initiale est  $F_0 = F$ . Notons aussi  $G_x$  la loi conditionnelle de  $S_1 - x$  sachant que  $S_1 > x$ , pour  $x \geq 0$ , c'est-à-dire

$$G_x(A) = \begin{cases} \frac{1}{F(]x, \infty[)} \int 1_A(s-x) 1_{\{s>x\}} F(ds) & \text{si } F(]x, \infty[) > 0 \\ F'(A) & \text{sinon,} \end{cases} \quad (4.2.11)$$

où  $F'$  est une probabilité arbitraire (par exemple  $F' = F$ ). Remarquer qu'on a aussi  $G_0 = F$ .

**Proposition 4.2.4** *Sous la probabilité  $\mathbb{P}_F$  le processus  $(W_t)$  est de loi initiale  $\varepsilon_0$  et est markovien relativement à la filtration  $(\mathcal{F}_t)$  et au semi-groupe de transition  $(P_t^W)$  sur  $\mathbb{R}_+$  donné par*

$$P_t^W(x, A) = \mathbb{P}_{G_x}(W_t \in A). \quad (4.2.12)$$

**Preuve.** La première assertion est évidente. Pour le reste, on utilise la notation  $\widehat{S}_n$  de la preuve précédente. Chaque variable  $W_u$  s'écrit  $W_u = h_u(T_1, \widehat{S}_1)$ , où  $h_u$  est une fonction mesurable sur l'espace produit  $(\mathbb{R} \times E, \mathcal{R} \otimes \mathcal{E})$ . Par suite  $P_t^W$ , défini par (4.2.12), vaut

$$P_t^W(x, A) = \int 1_A(h_t(t, \widehat{s})) G_x(dt) Q(d\widehat{s}).$$

Comme  $G_x(dy)$  est une probabilité de transition de  $\mathbb{R}_+$  dans lui-même, il est clair que chaque  $P_t^W$  est aussi une probabilité de transition.

Exactement comme dans la preuve précédente, et puisque toute variable  $\mathcal{F}_s$ -mesurable est, en restriction à l'ensemble  $\{N_s = n\}$ , une fonction de  $(T_1, \dots, T_n)$ , il nous reste à montrer que pour tous  $n \in \mathbb{N}$ ,  $s, t \geq 0$ , et  $f_n$  et  $g$  boréliennes bornées,

$$\mathbb{E}_F \left( f_n(T_1, \dots, T_n) g(W_{t+s}) 1_{\{N_s = n\}} \right) = \mathbb{E}_F \left( f_n(T_1, \dots, T_n) 1_{\{N_s = n\}} P_t^W g(W_s) \right) \quad (4.2.13)$$

Sur l'ensemble  $\{N_s = n\} = \{T_n \leq s < T_{n+1}\}$  on a  $W_{s+t} = h_t(T_n + S_n - s, \widehat{S}_{n+1})$ . Comme  $\widehat{S}_{n+1}$  est indépendant de  $(T_1, \dots, T_n, S_n)$  et de loi  $Q$ , la partie gauche de (4.2.13) vaut

$$\begin{aligned} & \mathbb{E}_F \left( f_n(T_1, \dots, T_n) 1_{\{T_n \leq s < T_n + S_n\}} \int g \circ h_t(T_n + S_n - s, \widehat{s}) Q(d\widehat{s}) \right) \\ &= \mathbb{E}_F \left( f_n(T_1, \dots, T_n) 1_{\{T_n \leq s\}} F(\cdot - T_n, \infty) \int g \circ h_t(u, \widehat{s}) G_{s-T_n}(du) Q(d\widehat{s}) \right). \end{aligned}$$

(Si  $n \geq 1$  cela est vrai quelle que soit la loi initiale  $F_0$ , en fait; mais pour  $n = 0$  on a besoin que  $F_0 = F$ ). L'intégrale ci-dessus vaut en fait  $P_t^W g(t - T_n)$ . Comme  $W_s = t - T_n$  sur l'ensemble  $\{T_n \leq s < T_{n+1}\}$ , l'expression précédente vaut aussi

$$\mathbb{E}_F \left( f_n(T_1, \dots, T_n) 1_{\{T_n \leq s\}} F(\cdot | t - T_n, \infty) P_t^W g(W_s) \right),$$

qui égale le membre de droite de (4.2.13).  $\square$

En particulier si  $(N_t)$  est un processus de Poisson de paramètre  $\lambda > 0$ , alors  $F_0 = F$  est la loi exponentielle de paramètre  $\lambda$ , de sorte que  $G_x = F$  pour tout  $x$ . La probabilité  $P_t^W(x, \cdot)$  prend alors la forme  $P_t^W(x, A) = \mathbb{P}(W_t \in A) = F(A)$ , qui ne dépend pas de  $x$ : le temps d'attente du prochain saut après  $s$  est donc exponentiel, et indépendant du passé avant  $s$ . Cette propriété, évidente, vient de ce que le processus  $(N_{s+t} - N_s)_{t \geq 0}$  est indépendant de  $\mathcal{F}_s$  et de même loi que  $(N_t)$ .

### 4.3 Le théorème de renouvellement

L'objectif de ce paragraphe est de montrer le théorème suivant, appelé *théorème de renouvellement*. Ce théorème concerne le comportement asymptotique des mesures translatées  $\varepsilon_{-t} \star U$  lorsque  $t \rightarrow \infty$ , où  $U$  désigne le potentiel recentré d'un processus de renouvellement de loi inter-sauts  $F$ . Remarquer en effet que pour tout borélien  $A$  de  $\mathbb{R}$  on a

$$U(t + A) = \int 1_A(x - t) U(dx) = (\varepsilon_{-t} \star U)(A). \quad (4.3.1)$$

Plus généralement, pour toute fonction  $f$  positive, ou bornée et à support compact, le produit de convolution  $U \star f$  est la fonction

$$U \star f(t) = \int f(t-x)U(dx)$$

(ceci a un sens, en vertu du lemme 4.2.1, qui implique notamment que  $U$  est une mesure de Radon). Pour toute fonction  $f$  sur  $\mathbb{R}$  on définit sa “symétrique”  $\widehat{f}$  par  $\widehat{f}(x) = f(-x)$ . D’après (4.3.1), on a donc  $U(t+A) = (U \star \widehat{1_A})(t)$ . Plus généralement, pour toute fonction  $f$  positive, ou bornée et à support compact, on a donc

$$U \star f(t) = (\varepsilon_{-t} \star U)(\widehat{f}). \quad (4.3.2)$$

Supposons d’abord que  $F$  soit portée par un ensemble de la forme  $E_a = \{na : n \in \mathbb{N}\}$ , pour un  $a > 0$ . Cela veut dire que presque sûrement  $S_1$  est un multiple entier de  $a$ , et il en est évidemment de même de chaque variable  $T_n - T_1$ . Par suite la mesure  $U$  est aussi portée par  $E_a$ , et la mesure translatée  $\varepsilon_t \star U$  sera portée par l’ensemble  $E_a - t = \{na - t : n \in \mathbb{N}\}$ . On ne peut donc pas espérer en général que les mesures  $\varepsilon_t \star U$  convergent en un sens quelconque vers une limite lorsque  $t \rightarrow \infty$ . Pour obtenir un résultat, il nous faudra donc faire une hypothèse sur le support  $Sup(F)$  de la loi  $F$ : on rappelle que  $Sup(F)$  est le plus petit fermé de  $\mathbb{R}$  dont la  $F$ -mesure du complémentaire soit nulle.

**Définition 4.3.1** Une probabilité  $F$  sur  $\mathbb{R}$  est dite *non-arithmétique* si le plus petit sous-groupe fermé (additif) de  $\mathbb{R}$  contenant  $Sup(F)$  est  $\mathbb{R}$  lui-même (rappelons qu’un sous-groupe additif de  $\mathbb{R}$  est une partie non vide  $A$  de  $\mathbb{R}$  telle que  $a - b \in A$  pour tout  $a, b \in A$ ).

**Lemme 4.3.2** *Si  $F$  est une probabilité sur  $\mathbb{R}$ , trois cas sont possibles:*

- (a)  $F = \varepsilon_0$ , et alors le plus petit sous-groupe fermé contenant  $Sup(F)$  est  $\{0\}$ .
- (b)  $F$  est portée par l’ensemble  $E'_a = \{na : n \in \mathbb{Z}\}$ , pour un  $a > 0$ , et alors  $E'_a$  est le plus petit sous-groupe fermé contenant  $Sup(F)$ .
- (c)  $F$  est non-arithmétique.

**Preuve.** Ce résultat est évident, une fois remarqué que les seuls sous-groupes fermés de  $\mathbb{R}$  sont  $\{0\}$ ,  $E'_a$  pour un  $a > 0$ , et  $\mathbb{R}$  lui-même.  $\square$

Il est aussi évident qu’on est dans le cas (c) si et seulement si  $Sup(F)$  contient au moins deux points  $a, b$  qui sont “rationnellement indépendants”, ce qui signifie que  $a \neq 0$  et  $b/a$  n’est pas rationnel. En particulier si  $F$  admet une densité par rapport à la mesure de Lebesgue (ou, bien plus généralement, admet une partie diffuse non nulle), alors elle est non-arithmétique.

Dans le cas des processus de renouvellement, le cas (a) ci-dessus est exclu puisque  $F$  est portée par  $]0, \infty[$ . Donc, soit  $F$  est portée par un ensemble de la forme  $E_a$ , soit elle est non-arithmétique.

**Théorème 4.3.3** *Soit un processus de renouvellement de loi inter-sauts  $F$  de moyenne  $m$  (cf. (4.1.2)) et de potentiel recentré  $U$ . Si la probabilité  $F$  est non-arithmétique, les mesures  $\varepsilon_{-t} \star U$  convergent vaguement lorsque  $t \rightarrow \infty$  vers la mesure nulle si  $m = \infty$ , et vers la mesure  $\frac{1}{m} \mu$  si  $m < \infty$ , où  $\mu$  désigne la mesure de Lebesgue sur  $\mathbb{R}$ .*

On conviendra de dire que  $\frac{1}{m} \mu = 0$  si  $m = \infty$ . Rappelons qu'une famille de mesures de Radon  $\nu_t$  sur  $\mathbb{R}$  converge vaguement vers une limite  $\nu$  (encore une mesure de Radon), si on a  $\nu_t(f) \rightarrow \nu(f)$  lorsque  $t \rightarrow \infty$ , pour toute fonction  $f$  continue à support compact. Vu (4.3.2), ce théorème signifie donc que, lorsque  $t \rightarrow \infty$ ,

$$U \star f(t) \rightarrow \frac{1}{m} \int f(x) dx \quad (= 0 \text{ si } m = \infty) \quad (4.3.3)$$

pour toute fonction continue à support compact  $f$ , puisque les intégrales de  $f$  et de  $\hat{f}$  par rapport à la mesure de Lebesgue sont égales.

Commençons par des lemmes. Le premier est appelé lemme de Choquet–Deny.

**Lemme 4.3.4** *Si  $F$  est une probabilité sur  $\mathbb{R}$ , non-arithmétique, toute fonction continue bornée  $f$  qui vérifie  $\int f(x+y)F(dy) = f(x)$  pour tout  $x$  est constante.*

**Preuve.** On pose

$$H_0(x) = \int (f(x+y) - f(x))^2 F(dy), \quad H_n(x) = \int H_0(x+y) F^{n\star}(dy)$$

et  $S_n = \sum_{k=0}^n H_k$ . L'hypothèse sur  $f$  implique  $H_0(x) = \int f(x+y)^2 F(dy) - f(x)^2$ , donc

$$\begin{aligned} S_n(x) &= \sum_{k=0}^n \left( \int f(x+y)^2 F^{(k+1)\star}(dy) - \int f(x+y)^2 F^{k\star}(dy) \right) \\ &= \int f(x+y)^2 F^{(n+1)\star}(dy) - f(x)^2 \leq C^2, \end{aligned}$$

si  $C = \sup |f|$ . D'autre part on a par Fubini et Cauchy–Schwarz:

$$\begin{aligned} H_1(x) &= \int F(dy) \int (f(x+y+z) - f(x+y))^2 F(dz) \\ &= \int F(dz) \int (f(x+y+z) - f(x+y))^2 F(dy) \\ &\geq \int \left( \int (f(x+y+z) - f(x+y)) F(dy) \right)^2 F(dz) \\ &= \int (f(x+z) - f(x))^2 F(dz) = H_0(x), \end{aligned}$$

où l'avant-dernière égalité provient encore de l'hypothèse faite sur  $f$ . Donc  $H_1 \geq H_0$  et on en déduit facilement que la suite  $H_n(x)$  est croissante en  $n$  pour tout  $x$ . En particulier  $H_n(x) \geq H_0(x)$ , donc  $S_n(x) \geq nH_0(x)$ . Comme  $S_n(x) \leq 2C$  il faut donc que  $H_0(x) = 0$  pour tout  $x$ .

Mais  $H_0(x) = 0$  implique que  $f(x + y) = f(x)$  pour  $F$ -presque tout  $y$ . Comme  $f$  est continue, cela implique que  $f(x + y) = f(x)$  pour tout  $y \in \text{Sup}(F)$  et tout  $x$ : donc tout point  $y \in \text{Sup}(F)$  est une période pour la fonction  $f$ . L'ensemble des périodes d'une fonction continue est à l'évidence un sous-groupe fermé de  $\mathbb{R}$ , donc il contient le plus petit sous-groupe fermé contenant  $\text{Sup}(F)$ , qui égale  $\mathbb{R}$  puisque  $F$  est non-arithmétique: donc tout  $y \in \mathbb{R}$  est une période pour  $f$ , ce qui signifie que  $f$  est constante.  $\square$

**Lemme 4.3.5** *Si  $\phi(x) = F(]x, \infty[) 1_{\{x \geq 0\}}$ , le produit de convolution  $U \star \phi$  est l'indicatrice de  $\mathbb{R}_+$ .*

**Preuve.** Rappelons que  $U$  ne charge que  $\mathbb{R}_+$ , et vérifie  $U([0, T]) \leq C(T + 1)$  par le lemme 4.2.1. Comme de plus  $\phi$  est continue à droite et vérifie  $0 \leq \phi \leq 1_{\mathbb{R}_+}$ , la fonction  $h = U \star \phi$  est  $h(t) = \int \phi(t - x) 1_{[0, t]}(x) U(dx)$ , et est aussi continue à droite (théorème de Lebesgue) et vérifie  $0 \leq h(t) \leq C(t + 1)$  pour  $t \geq 0$ . Elle est donc entièrement caractérisée par sa transformée de Laplace  $\bar{h}(u) = \int_0^\infty h(x) e^{-ux} dx$  pour  $u > 0$ . Il suffit donc de montrer que  $\bar{h}(u) = 1/u$ , qui est la transformée de Laplace de  $1_{\mathbb{R}_+}$ .

La transformée de Laplace d'un produit de convolution étant le produit des transformées de Laplace, on a  $\bar{h}(u) = \bar{\phi}(u) \sum_{n \geq 0} \bar{F}(u)^n$ , où  $\bar{F}(u) = \int e^{-ut} F(dt)$  et

$$\bar{\phi}(u) = \int_0^\infty e^{-ut} h(t) dt = \int F(dx) \int_0^x e^{-ut} dt = \frac{1}{u} (1 - \bar{F}(u)).$$

On a donc le résultat.  $\square$

**Preuve du théorème 4.3.3.** On pose  $\nu_t = \varepsilon_{-t} \star U$ . D'après le lemme 4.2.1, pour tout compact  $K$  de  $\mathbb{R}$  on a  $\sup_t \nu_t(K) < \infty$ . D'après un résultat classique sur la convergence vague, cela implique que la famille  $\{\nu_t : t \in \mathbb{R}\}$  est relativement compacte pour la topologie vague des mesures. En d'autres termes, il existe des suites  $(t_n)$  tendant vers l'infini, telles que  $\nu_{t_n}$  converge vaguement, et de plus si les limites ainsi obtenues sont toutes les mêmes, disons  $\nu$ , alors  $\nu_t$  converge vaguement vers  $\nu$  quand  $t \rightarrow \infty$ .

Par suite il nous reste à montrer que si  $t_n \rightarrow \infty$  et si  $\nu_{t_n}$  converge vaguement vers  $\nu$ , alors  $\nu = \frac{\mu}{m}$ .

Soit  $f$  continue à support contenu dans le compact  $[-T, T]$ , et  $g(x) = \int f(x + y) \nu(dy)$ . D'abord il est évident (par le théorème de Lebesgue, et puisque  $\nu$  est de Radon) que  $g$  est continue. Ensuite,  $y \mapsto f(x + y)$  est aussi continue à support compact, donc  $g(x)$  est la limite de

$$g_n(x) = \int f(x + y) \nu_{t_n}(dy) = \int f(x - t_n + y) U(dy). \quad (4.3.4)$$

Si  $C = \sup |f|$ , on a donc  $|g_n(x)| \leq CU([-T + t_n + x, T + t_n + x]) \leq C\alpha_{2T}$  par le lemme 4.2.1. On a donc aussi  $|g(x)| \leq C\alpha_{2T}$ . Enfin l'équation (4.2.5) implique

$$g_n(x) = f(x - t_n) + \int F(dz) \int f(x + z - t_n + y) U(dy) = f(x - t_n) + \int g_n(x + z) F(dz),$$

et comme  $g_n$  converge vers  $g$  en restant uniformément borné et que  $f(x - t_n)$  est nul pour  $n$  assez grand, on obtient en passant à la limite que  $g(x) = \int g(x + z) F(dz)$ . Le lemme

de Choquet–Deny entraîne alors que la fonction  $g$  est constante. En d’autres termes, on a montré que

$$\int f(x+y)\nu(dy) = \int f(y)\nu(dy) \quad \forall x,$$

et pour toute fonction continue à support compact: cela entraîne que la mesure de Radon  $\nu$  est invariante par translation, donc en fait est de la forme  $\nu = \alpha\mu$  pour une certaine constante  $\alpha \geq 0$ .

La seule chose restant à montrer est donc que  $\alpha = 1/m$ . Pour cela on va utiliser le lemme 4.3.5, après avoir rappelé que la fonction  $\phi$  qui y figure vérifie  $\mu(\phi) = m$ . On va distinguer deux cas:

- On a  $m = \infty$ : Pour chaque  $p \in \mathbb{N}$  soit  $f_p$  une fonction continue et vérifiant  $1_{[-p,p]} \leq f_p \leq 1_{[-p-1,p+1]}$ . La fonction  $\widehat{\phi}f_p$  est positive bornée à support compact; elle n’est pas forcément continue, mais elle est continue à gauche avec des limites à droite et donc n’est discontinue qu’en au plus un nombre dénombrable de points; comme la mesure limite  $\nu = \alpha\mu$  ne charge pas les points, la suite  $\nu_{t_n}(\widehat{\phi}f_p)$  converge vers  $\alpha\mu(\widehat{\phi}f_p)$ . Par ailleurs  $\nu_{t_n}(\widehat{\phi}) = 1$  d’après (4.3.2) et le lemme 4.2.12, donc  $\alpha\mu(\widehat{\phi}f_p) \leq 1$  pour tout  $p$ . En passant à la limite en  $p$  (théorème de limite monotone), on voit que  $\alpha\mu(\widehat{\phi}) = \alpha\mu(\phi) = \alpha m \leq 1$ . Comme  $m = \infty$  ceci n’est possible que si  $\alpha = 0$ , et le résultat voulu est prouvé.

- On a  $m < \infty$ : La fonction  $x \mapsto \phi(x)/m$  est la densité d’une probabilité  $\eta$ , donc la mesure  $\eta \star U$  admet la densité  $x \mapsto (U \star \phi)(x)/m$ , égale à  $\frac{1}{m} 1_{\mathbb{R}_+}$  par le lemme 4.2.12. Cela revient à dire que  $m\eta \star U$  est la restriction de  $\mu$  à  $\mathbb{R}_+$ , et par suite  $m\eta \star \nu_{t_n} = m\eta \star U \star \varepsilon_{-t_n}$  est la restriction de  $\mu$  à  $[-t_n, \infty[$ , qui converge vaguement vers  $\mu$  puisque  $t_n \rightarrow \infty$ . Par ailleurs si  $f$  est continue à support compact, on a vu que les  $g_n$  de (4.3.4) convergent vers la constante  $\nu(f)$  en restant uniformément bornées, de sorte que  $(m\eta \star \nu_{t_n})(f) = m\eta(g_n)$  converge vers  $m\nu(f)$ . Il en découle que  $m\nu = \mu$ , soit  $\alpha = 1/m$ .  $\square$

Le théorème 4.3.3 est notre résultat fondamental, mais il n’est pas tout-à-fait suffisant pour les applications. Une conséquence immédiate – mais fort utile – en est le corollaire suivant.

**Corollaire 4.3.6** *Sous les hypothèses du théorème 4.3.3, on a la convergence (4.3.3) pour toute fonction  $f$  vérifiant l’une des hypothèses suivantes:*

(i)  *$f$  est bornée à support compact, et l’ensemble des points où elle est discontinue est de mesure de Lebesgue nulle.*

(ii)  *$f$  est positive et Lebesgue-intégrable, et les deux sommes*

$$S_n^+ = \frac{1}{n} \sum_{i \in \mathbb{Z}} \sup_{x \in [i/n, (i+1)/n[} f(x), \quad S_n^- = \frac{1}{n} \sum_{i \in \mathbb{Z}} \inf_{x \in [i/n, (i+1)/n[} f(x) \quad (4.3.5)$$

*convergent vers la même limite (nécessairement  $\mu(f)$ ) quand  $n \rightarrow \infty$ . On dit alors que  $f$  est directement Riemann intégrable.*

Bien entendu, on a aussi (4.3.3) pour toute combinaison linéaire finie de fonctions des



types (i) ou (ii) ci-dessus. Noter que toute fonction positive décroissante et Lebesgue-intégrable est de type (ii).

**Preuve.** Pour (i), c'est un résultat classique relatifs aux convergence (vague ou étroite) des mesures.

Si  $f$  est de type (ii), elle est coincée entre les fonctions  $f_n^+$  et  $f_n^-$ , égales sur chaque intervalle  $I_{n,i} = ]i/n, (i+1)/n]$  respectivement au sup et à l'inf de  $f$  sur  $I_{n,i}$ . Notons aussi  $f_{N,n}^+ = f_n^+ 1_{[-N,N]}$  et  $f_{N,n}^- = f_n^- 1_{[-N,N]}$ . Chacune de ces dernières fonctions est constante par morceaux et à support compact, donc de type (i), et par suite  $U \star f_{N,n}^+(t) \rightarrow \frac{1}{m} \mu(f_{N,n}^+)$  par exemple, quand  $t \rightarrow \infty$ . De plus, (ii) entraîne clairement que les intégrales de Lebesgue de  $f_{N,n}^+$  et  $f_{N,n}^-$  convergent vers  $\mu(f)$  lorsque  $n, N \rightarrow \infty$ .

On a  $f \geq f_{N,n}^-$ , de sorte que  $\liminf_t U \star f(t) \geq \frac{1}{m} \mu(f_{N,n}^-)$  pour tous  $n, N$ , donc aussi  $\liminf_t U \star f(t) \geq \frac{1}{m} \mu(f)$ .

L'inégalité inverse est un peu plus difficile. Fixons  $n$  tel que  $S_n^+ < \infty$ . Si  $\varepsilon > 0$  il existe  $N$  tel que  $S_n'^+ = \frac{1}{n} \sum_{i \in \mathbf{Z}, |i| \geq nN} \sup_{x \in ]i/n, (i+1)/n]} f(x)$  soit inférieur à  $\varepsilon$ . En utilisant le lemme 4.2.1, on voit alors que

$$U \star f(t) \leq U \star f_{N,n}^+(t) + \varepsilon \alpha_{1/n}.$$

Donc  $\limsup_t U \star f(t) \leq \frac{1}{m} \mu(f_{N,n}^+) + \varepsilon \alpha_{1/n}$ . En faisant  $\varepsilon \rightarrow 0$  (donc  $N \rightarrow \infty$ ), on arrive à  $\limsup_t U \star f(t) \leq \frac{1}{m} \mu(f_n^+)$ . Mais (ii) implique aussi clairement que  $\mu(f_n^+) \rightarrow \mu(f)$  quand  $n \rightarrow \infty$ , donc en définitive  $\limsup_t U \star f(t) \leq \frac{1}{m} \mu(f)$ , et on a le résultat.  $\square$

Comme application de ce corollaire, on voit donc que, si  $F$  est non-arithmétique, pour tout borélien  $A$  de  $\mathbb{R}$  dont la frontière est de mesure de Lebesgue nulle,

$$U(t + A) \rightarrow \frac{\mu(A)}{m}. \quad (4.3.6)$$

## 4.4 Quelques applications

### 4.4.1 Comportement asymptotique du potentiel

On peut s'intéresser au comportement asymptotique de  $U_{F_0}$ : cf. (4.2.6). Le résultat est le même que pour le potentiel recentré:

**Théorème 4.4.1** *Soit  $U_{F_0}$  le potentiel d'un processus de renouvellement de loi inter-sauts  $F$  et de loi initiale  $F_0$ . Si la probabilité  $F$  est non-arithmétique et de moyenne  $m$ , on a*

$$U_{F_0} \star f(t) \rightarrow \frac{1}{m} \int f(x) dx \quad (= 0 \text{ si } m = \infty) \quad (4.4.1)$$

pour toute fonction  $f$  vérifiant les conditions (i) ou (ii) du corollaire 4.3.6. En particulier  $\varepsilon_{-t} \star U_{F_0}$  converge vaguement vers  $\mu/m$  si  $t \rightarrow \infty$ .

**Preuve.** Pour  $f$  comme ci-dessus, on sait que  $g(t) = U \star f(t)$  converge vers  $\mu(f)/m$ . Mais  $U_{F_0} \star f(t) = \int g(t-x) F_0(dx)$ , et la fonction  $g$  est bornée, de sorte que le résultat découle d'une simple application du théorème de Lebesgue.  $\square$

#### 4.4.2 L'équation de renouvellement

Nous avons introduit en (4.2.5) une première équation de renouvellement. Plus généralement, on appelle *équation de renouvellement avec second membre*  $\eta$ , pour les mesures, l'équation

$$\rho = F \star \rho + \eta. \quad (4.4.2)$$

Dans cette équation  $\eta$  est une mesure de Radon portée par  $\mathbb{R}_+$ . L'inconnue est  $\rho$ , également une mesure de Radon portée par  $\mathbb{R}_+$ .

**Théorème 4.4.2** *Si  $\eta$  est une mesure de Radon portée par  $\mathbb{R}_+$ , dont la transformée de Laplace  $\bar{\eta}(u) = \int e^{-ux} \eta(dx)$  est finie pour tout  $u > u_0$  pour un  $u_0 \geq 0$  donné, l'équation de renouvellement (4.4.2) admet une solution et une seule  $\rho$  dans l'ensemble des mesures de Radon admettant une transformée de Laplace  $\bar{\rho}(u)$  finie pour tout  $u > u_0$ , et cette solution est donnée par*

$$\rho = U \star \eta. \quad (4.4.3)$$

**Preuve.** La transformée de Laplace de la mesure  $U$  est  $\bar{U}(u) = \sum_{n \geq 0} \bar{F}(u)^n = \frac{1}{1 - \bar{F}(u)}$ , qui est finie pour tout  $u > 0$  puisque  $\bar{F}(u) = \int e^{-ux} F(dx) < 1$  lorsque  $u > 0$  (rappelons que  $F$  est portée par  $]0, \infty[$ ). Donc la mesure  $\rho$  définie par (4.4.3), qui est évidemment portée par  $\mathbb{R}_+$ , vérifie  $\bar{\rho}(u) = \bar{U}(u) \bar{\eta}(u) < \infty$  pour  $u > u_0$ ; en particulier,  $\rho$  est de Radon. Enfin, le fait que  $\rho$  vérifie (4.4.2) découle immédiatement de (4.2.5).

Pour l'unicité, soit  $\rho$  une solution satisfaisant les propriétés requises. On a  $\bar{\rho}(u) = \bar{F}(u) \bar{\rho}(u) + \bar{\eta}(u)$ , donc  $\bar{\rho}(u) = \frac{\bar{\eta}(u)}{1 - \bar{F}(u)}$  pour  $u > u_0$ . Comme la transformée de Laplace en restriction à un intervalle  $]u_0, \infty[$  sur laquelle elle est finie caractérise entièrement la mesure portée par  $\mathbb{R}_+$ , cela prouve l'unicité.  $\square$

On a aussi une équation de renouvellement du même type, mais avec des fonctions positives à support dans  $\mathbb{R}_+$ , qui s'écrit

$$g = F \star g + h. \quad (4.4.4)$$

La fonction  $h$  est encore appelée le *second membre* de l'équation. Comme  $g$  et  $h$  peuvent être considérées comme les densités de deux mesures  $\rho$  et  $\eta$  portées par  $\mathbb{R}_+$ , cette équation se ramène en fait à (4.4.2), et on a le

**Corollaire 4.4.3** *Si  $h$  est une fonction borélienne positive à support dans  $\mathbb{R}_+$ , dont la transformée de Laplace  $\bar{h}(u) = \int e^{-ux} h(x) dx$  est finie pour tout  $u > u_0$  pour un  $u_0 \geq 0$  donné, l'équation de renouvellement (4.4.4) admet une solution  $g$  dans l'ensemble des fonctions boréliennes positives à support dans  $\mathbb{R}_+$  et admettant une transformée de Laplace  $\bar{g}(u)$  finie pour tout  $u > u_0$ , donnée par*

$$g = U \star h. \quad (4.4.5)$$

*Cette solution est unique à un ensemble de mesure de Lebesgue nulle près. Si de plus  $h$  est continue et bornée, la fonction  $g$  ci-dessus est aussi continue.*

Le comportement asymptotique des solutions des équations de renouvellement se déduit aisément du théorème 4.3.3 et de son corollaire 4.3.6. Par exemple si  $h$  est une fonction positive vérifiant (i) ou (ii) du corollaire 4.3.6, la solution (4.4.5) de l'équation de renouvellement vérifie  $g(t) \rightarrow \mu(h)/m$  quand  $t \rightarrow \infty$  dès que  $F$  est non-arithmétique.

### 4.4.3 Comportement asymptotique du processus des âges

Nous allons maintenant revenir aux processus  $(V_t)$  et  $(W_t)$  de (4.2.7).

**Proposition 4.4.4** *Pour toute fonction borélienne bornée  $f$  sur  $\mathbb{R}$ , la fonction  $g(t) = 1_{\{t \geq 0\}} \mathbb{E}_F(f(V_t))$  vérifie l'équation de renouvellement (4.4.4) avec le second membre  $h$  donné par*

$$h(t) = 1_{\{t \geq 0\}} \int f(x-t) 1_{\{x > t\}} F(dx). \quad (4.4.6)$$

**Preuve.** On a aussi  $h(t) = \mathbb{E}_F(f(V_t) 1_{\{T_1 > t\}})$  si  $t \geq 0$ , puisque  $V_t = T_1 - t$  si  $T_1 > t$ . En reprenant les notations de la preuve de la proposition 4.2.3, on a  $V_u = h_u(T_1, \widehat{S}_1)$  donc aussi  $V_t = h_{t-T_1}(S_1, \widehat{S}_2)$  sur l'ensemble  $\{T_1 \leq t\}$ . Donc comme  $(S_1, \widehat{S}_2)$  est indépendant de  $T_1$  et a même loi que le couple  $(S_0, \widehat{S}_1)$  sous  $\mathbb{P}_F$ , il vient pour  $t \geq 0$ :

$$\begin{aligned} g(t) &= h(t) + \mathbb{E}_F \left( f \circ h_{t-T_1}(S_1, \widehat{S}_2) 1_{\{T_1 \leq t\}} \right) \\ &= h(t) + \int E_F(f(V_{t-x})) 1_{\{x \leq t\}} F(dx), \end{aligned}$$

qui vaut  $h(t) + F \star g(t)$ . Si enfin  $t < 0$  on a  $g(t) = 0 = h(t) + F \star g(t)$ .  $\square$

Pour le résultat suivant, on utilise la même notation qu'au lemme 4.2.12:

$$\phi(s) = F(]s, \infty[) 1_{\{s \geq 0\}}. \quad (4.4.7)$$

Rappelons que  $\mu(\phi) = m$ , de sorte que si  $m < \infty$  la fonction  $\frac{1}{m} \phi$  est une densité de probabilité portée par  $\mathbb{R}_+$ .

**Proposition 4.4.5** *Supposons  $F$  non-arithmétique, de moyenne  $m$ . Pour tout  $x > 0$  et toute fonction continue bornée  $f$  sur  $\mathbb{R}$ , on a quand  $t \rightarrow \infty$ :*

$$P_t^V f(x) \rightarrow \begin{cases} \frac{1}{m} \int f(s) \phi(s) ds & \text{si } m < \infty \\ 0 & \text{si } m = \infty. \end{cases} \quad (4.4.8)$$

**Preuve.** Comme  $P_t^V(x, \cdot)$  ne charge que  $]0, \infty[$  et comme  $\phi(s) = 0$  si  $s < 0$ , on peut supposer que  $f(x) = 0$  si  $x \leq 0$ . D'après les propriétés de la convergence étroite, il suffit même de considérer le cas où  $0 \leq f \leq 1$  et le support de  $f$  est un compact de  $]0, \infty[$ .

Avec les notations  $g$  et  $h$  de la proposition précédente, on a  $P_t^V f(x) = g(t-x)$  pour  $t > x$  d'après (4.2.9), donc il suffit de montrer la convergence de  $g(t)$  vers la droite de (4.4.8). Comme  $h$  et  $g$  sont bornées, le corollaire 4.4.3 entraîne que  $g = U \star h$  à un ensemble

de mesure de Lebesgue nulle près. De plus  $h$  et  $g$  sont continues à droite puisque  $f$  est continue, donc en fait  $g = U \star h$  partout. Il existe  $T > 0$  tels que  $0 \leq f \leq 1_{]0, T]}$ . D'après Fubini et comme  $f(t) = 0$  si  $t \leq 0$ , il vient aussi

$$\mu(h) = \int_0^\infty dt \int f(x-t)F(dx) = \int F(dx) \int_0^x f(x-t)dt = \int F(dx) \int_0^x f(s)ds = \int f(s)\phi(s)ds.$$

Distinguons les deux cas suivants:

- On a  $m = \infty$ : Il est clair que  $h(t) \leq F(]t, t+T])$ , donc  $g = U \star h$  vérifie

$$g(t) \leq \int F(]t-x, t+T-x])U(dx) = (U \star F)(]t, t+T]) \leq U(]t, t+T])$$

en vertu de (4.2.5). Le résultat découle alors de (4.3.6).

- On a  $m < \infty$ : Vu le calcul de  $\mu(h)$  fait ci-dessus, il suffit d'après le corollaire 4.3.6 de montrer que la fonction  $h$  est directement Riemann-intégrable, sachant qu'elle est Lebesgue-intégrable (car  $\mu(h) \leq \int_0^\infty \phi(s)ds = m$ ). Comme  $f$  est continue à support dans  $]0, T]$ , on a que  $\varepsilon_n = \sup_{s,t: |s-t| \leq 1/n} (|f(t) - f(s)|)$  tend vers 0, et si  $\frac{i}{n} < s \leq t \leq \frac{i+1}{n}$  on a  $|f(x-t) - f(x-s)| \leq \varepsilon_n 1_{]i/n, \infty[}(x)$ . Comme dans (4.4.6) on peut supprimer l'indicatrice  $1_{\{x>t\}}$  puisque  $f(s) = 0$  pour  $s \leq 0$ , on en déduit que pour  $\frac{i}{n} < s \leq t \leq \frac{i+1}{n}$  on a  $|h(t) - h(s)| \leq \varepsilon_n \phi(i/n)$ . Il en découle que pour tout  $i \in \mathbb{N}$ ,

$$\sup_{x \in [\frac{i}{n}, \frac{i+1}{n}[} h(x) - \inf_{x \in [\frac{i}{n}, \frac{i+1}{n}[} h(x) \leq \varepsilon_n \phi\left(\frac{i}{n}\right).$$

Par suite, en utilisant aussi  $h(s) = 0$  si  $s < 0$ , les deux sommes  $S_n^+$  et  $S_n^-$  associées à la fonction  $h$  par (4.3.5) vérifient

$$S_n^+ - S_n^- \leq \varepsilon_n \frac{1}{n} \sum_{i \geq 0} \phi\left(\frac{i}{n}\right) \leq \varepsilon_n \left(\frac{1}{n} + \int_0^\infty \phi(s)ds\right) = \varepsilon_n \left(\frac{1}{n} + m\right),$$

puisque  $\phi$  est décroissante. Par suite  $S_n^+ - S_n^- \rightarrow 0$ , et la preuve est achevée.  $\square$

Ce résultat nous permet d'étudier la stationnarité éventuelle du processus  $(V_t)$  des temps avant renouvellement:

**Théorème 4.4.6** *Supposons  $F$  non-arithmétique et de moyenne  $m$ . Il existe une loi initiale  $F_0$  sous laquelle le processus  $(V_t)$  est stationnaire si et seulement si  $m < \infty$ . Dans ce cas la stationnarité a lieu si et seulement si  $F_0$  est la loi de densité  $\phi(\cdot)/m$  (cf. 4.4.7).*

**Preuve.** En vertu de la proposition 2.1.8, il suffit de montrer que le semi-groupe  $(P_t^V)$  des transitions du processus markovien  $(V_t)$  admet une probabilité invariante  $\eta$  si et seulement si  $m < \infty$ , et que dans ce cas l'unique probabilité invariante est  $F_0(ds) = \frac{1}{m} \phi(s) ds$ .

Supposons d'abord que  $\eta$  soit une probabilité invariante. Pour toute fonction positive  $f$  on a  $\eta P_t^V f = \eta(f)$ . Si en plus  $f$  est continue, à support compact dans  $\mathbb{R}_+$ , les  $P_t^V f(x)$  sont évidemment majorés par une constante indépendante de  $x$  et  $t$ , et convergent vers 0

(resp.  $F_0(f)$ ) si  $m = \infty$  (resp.  $m < \infty$ , avec  $F_0(ds) = \frac{1}{m} \phi(s) ds$ ). D'après le théorème de Lebesgue on en déduit que  $\eta(f)$  vaut 0 (resp.  $F_0(f)$ ). Lorsque  $m = \infty$  cela entraîne que  $\eta$  est la mesure nulle, d'où une contradiction, et il n'existe pas de probabilité invariante. Si  $m < \infty$ , cela entraîne que  $\eta = F_0$ , donc si  $F_0$  est une probabilité invariante, c'est la seule.

Il reste donc à montrer que si  $m < \infty$  et  $\eta(dx) = \phi(x)dx$ , alors  $\eta P_t^V f = \eta(f)$  pour tout  $t$  et toute fonction  $f$  comme dans la proposition 4.4.5. Avec les notations de la proposition 4.4.4, on a alors  $P_t^V f(x) = f(x-t)1_{\{x>t\}} + g(t-x)$  et  $g = U \star h$ . Par suite

$$\begin{aligned} \eta P_t^V f &= \int_t^\infty \phi(s) f(s-t) ds + \int_0^t \phi(s) g(t-s) ds \\ &= \int_0^\infty \phi(t+s) f(s) ds + U \star \phi \star h(t) \\ &= \int_0^\infty \phi(t+s) f(s) ds + \int_0^t h(t-s) ds. \end{aligned}$$

où dernière égalité vient du lemme 4.3.5. Mais (4.4.6) implique

$$\int_0^t h(t-s) ds = \int_0^t ds \int f(x-s) F(dx) = \int F(dx) \int_{x-t}^x f(s) ds = \int_0^\infty f(s) (\phi(s) - \phi(s+t)) ds,$$

et par suite  $\eta P_t^V f = \int_0^\infty \phi(s) f(s) ds = \eta(f)$ .  $\square$

Pour terminer nous traitons rapidement le cas du processus  $(W_t)$ .

**Proposition 4.4.7** *Pour toute fonction borélienne bornée  $f$  sur  $\mathbb{R}$ , la fonction  $g(t) = 1_{\{t \geq 0\}} \mathbb{E}_F(f(W_t))$  vérifie l'équation de renouvellement (4.4.4) avec le second membre  $h = f\phi$ .*

**Preuve.** En reprenant les notations de la preuve de la proposition 4.2.4, on a  $W_u = h_u(T_1, \hat{S}_1)$  donc aussi  $W_t = h_{t-T_1}(S_1, \hat{S}_2)$  sur l'ensemble  $\{T_1 \leq t\}$ . Donc comme  $(S_1, \hat{S}_2)$  est indépendant de  $T_1$  et a même loi que le couple  $(S_0, \hat{S}_1)$  sous  $\mathbb{P}_F$ , et comme  $W_t = t$  si  $T_1 > t$ , il vient pour  $t \geq 0$ , avec  $h = f\phi$ :

$$\begin{aligned} g(t) &= \mathbb{P}_F(T_1 > t) f(t) + \mathbb{E}_F \left( f \circ h_{t-T_1}(S_1, \hat{S}_2) 1_{\{T_1 \leq t\}} \right) \\ &= h(t) + \int E_F(f(W_{t-x})) 1_{\{x \leq t\}} F(dx), \end{aligned}$$

qui vaut  $h(t) + F \star g(t)$ . Si enfin  $t < 0$  on a  $g(t) = 0 = h(t) + F \star g(t)$ .  $\square$

**Proposition 4.4.8** *Supposons  $F$  non-arithmétique, de moyenne  $m$ . Pour tout  $x > 0$  et toute fonction continue bornée  $f$  sur  $\mathbb{R}$ , on a quand  $t \rightarrow \infty$ :*

$$P_t^W f(x) \rightarrow \begin{cases} \frac{1}{m} \int f(s) \phi(s) ds & \text{si } m < \infty \\ 0 & \text{si } m = \infty. \end{cases} \quad (4.4.9)$$

**Preuve.** Là encore, on peut se limiter au cas où  $f$  est à support compact dans  $[0, \infty[$  (puisque  $P_t^W(x, \cdot)$  ne charge que  $\mathbb{R}_+$ ), et vérifie  $0 \leq f \leq 1$ . Avec les notations  $g$  et  $h$  de la proposition précédente, la fonction  $h$  est bornée à support compact dans  $\mathbb{R}_+$  et est continue à droite, donc vérifie la condition (ii) du corollaire 4.3.6. Par suite  $g(t)$  converge vers  $\mu(h)/m$  quand  $t \rightarrow \infty$ , c'est-à-dire vers le membre de droite de (4.4.9). Etant donné (4.2.12), on a (par le même argument que dans la proposition précédente):

$$P_t^W f(x) = f(t)G_x(]t, \infty]) + \int g(t-s)1_{\{s \leq t\}}G_x(ds), \quad (4.4.10)$$

et le résultat en découle par le théorème de Lebesgue.  $\square$

En ce qui concerne l'existence de probabilités invariantes pour le semi-groupe  $(P_t^W)$ , la situation est radicalement différente de celle de  $(P_t^V)$ . En effet si  $m = \infty$  on montre comme dans le théorème 4.4.6 qu'il n'y a pas de probabilité invariante. Si maintenant  $m < \infty$ , la proposition précédente montre que pour tout  $x$  et tout  $\varepsilon > 0$  on a pour  $t$  assez grand:  $P_t^W(x, [0, \varepsilon]) > 0$ , de sorte que si  $\eta$  est une probabilité invariante on a aussi  $\eta([0, \varepsilon]) > 0$  pour tout  $\varepsilon > 0$ . Mais  $W_t = t$  si et seulement si  $T_1 > t$ , de sorte que d'après (4.2.12),

$$P_t^W(x, \{t\}) = G_x(]t, \infty]) = \frac{\phi(t+x)}{\phi(x)}$$

lorsque  $\phi(x) > 0$ . Si alors  $\alpha = \sup\{s : \phi(s) > 0\}$ , la probabilité invariante  $\eta$  vérifie

$$\eta(\{t\}) \geq \int \frac{\phi(t+x)}{\phi(x)} 1_{\{t+x < \alpha\}} \eta(dx),$$

qui est strictement positif pour tout  $t$  assez petit dès que  $\eta([0, \varepsilon]) > 0$  pour tout  $\varepsilon > 0$ , ce qui est toujours vrai comme on vient de le voir. Mais il n'existe pas de probabilité telle que  $\eta(\{t\}) > 0$  pour tout  $t$  proche de 0. Par suite on a:

**Théorème 4.4.9** *Le semi-groupe  $(P_t^W)$  n'admet jamais de probabilité invariante.*

“Moralement”, toutefois, il est possible en modifiant la définition de  $W_t$  d'en faire un processus de Markov stationnaire, avec la même loi invariante que  $V_t$  (à cause de la proposition 4.4.8): il faut remplacer  $W_t$  par  $W'_t$ , qui égale  $W_t$  si  $t \geq T_1$  et qui vaut  $W'_t = t + Z$  si  $T_1 > t$ , où  $Z$  est une variable indépendante de tout le reste et qui admet également pour loi la loi invariante de  $V_t$ . Le calcul, facile à faire, est laissé au lecteur.