

Modélisation et analyse stochastiques des réseaux de télécommunications

Laurent DECREUSEFOND et Pascal MOYAL

Table des matières

Chapitre 1. Introduction	13
1.1. Trafic, charge, Erlang, etc.	13
1.2. Notations et nomenclature	20
PREMIÈRE PARTIE. MODÉLISATION À TEMPS DISCRET	25
Chapitre 2. Chaînes de Markov	27
2.1. Définition et exemples	27
2.2. Propriété de Markov forte	31
2.3. Classification des états	34
2.4. Mesures et probabilité invariantes	43
2.5. Calcul pratique de la probabilité invariante	58
2.6. Problèmes	60
2.7. Notes et commentaires	63
DEUXIÈME PARTIE. MODÉLISATION À TEMPS CONTINU	65
Chapitre 3. Processus de Poisson	67
3.1. Définitions	68
3.2. Propriétés	74
3.3. Analogie discret : le processus de Bernoulli	79
3.4. Simulation du processus de Poisson	82
3.5. Processus de Poisson non homogène	83
3.6. Processus de Cox	87
3.7. Problèmes	88
3.8. Notes et commentaires	89
Chapitre 4. Processus de Markov	91

6 Modélisation et analyse des réseaux

4.1. Préliminaires	91
4.2. Construction trajectorielle	92
4.3. Semi-groupe markovien et générateur infinitésimal	97
4.4. Problème de martingale	113
4.5. Réversibilité et applications	117
4.6. MMPP	123
4.7. Problèmes	129
4.8. Notes et commentaires	131
Chapitre 5. Systèmes à attente	133
5.1. Formule de Little	133
5.2. File à un serveur	136
5.3. File à plusieurs serveurs	140
5.4. La file <i>processor sharing</i>	147
5.5. La file $M/M/\infty$	148
5.6. Le processus des départs	149
5.7. Réseaux de files d'attente	151
5.8. Problèmes	159
5.9. Notes et commentaires	162
Chapitre 6. Modèles à pertes	165
6.1. Généralités	165
6.2. Modèle d'Erlang	168
6.3. La file $M/M/1/1 + K$	170
6.4. L'effet <i>trunk</i>	173
6.5. Modèle d'Engset	174
6.6. File IPP/M/S/S	175
6.7. Modèles d'Erlang généralisés	178
6.8. Réseaux hiérarchiques	182
6.9. Un modèle avec renoncement	187
6.10. Un centre d'appels avec mise en attente et clients impatients	193
6.11. Problèmes	195
6.12. Notes et commentaires	196
Annexes	198
A. Cuisine et dépendance	199
A.1. Espace de probabilités, processus	199
A.2. Espérance conditionnelle	209
A.3. Espaces de vecteurs et ordres	214
A.4. Processus à variation finie	218
A.5. Martingales	225
A.6. Transformée de Laplace	241

Table des matières	7
--------------------	---

A.7. Notes et commentaires	242
----------------------------	-----

Bibliographie	243
----------------------	-----

PREMIÈRE PARTIE

Modélisation à temps discret

Chapitre 2

Chaînes de Markov

Lorsque l'on veut rendre compte de l'évolution temporelle d'un système, il faut décrire la façon dont le futur dépend du présent ou du passé. Parmi les principaux exemples, citons celui des équations différentielles (en temps continu) et celui des suites récurrentes (en temps discret). Dans le cas aléatoire à temps discret, l'analogie des suites récurrentes a été traité dans le chapitre ?? dans le cas stationnaire. Un cas particulier, sous des hypothèses supplémentaires d'indépendance, est celui des chaînes de Markov. Derrière une description apparemment simple se cache un outil mathématique redoutable d'efficacité pour les applications et riche de nombreuses propriétés.

On rappelle qu'il est conseillé de lire le petit paragraphe sur les notations de la remarque A.1.1.

2.1. Définition et exemples

Considérons donc une suite de variables aléatoires $X = (X_n, n \geq 0)$ à valeurs dans E fini ou dénombrable et indexée par les entiers, et la filtration $\mathcal{F}_n = \sigma\{X_j, 0 \leq j \leq n\}$ engendrée par cette suite.

Les trajectoires sont des éléments de $E^{\mathbb{N}}$, c'est-à-dire des suites d'éléments de E . Le décalage (voir l'annexe A) est alors défini par :

$$\begin{aligned}\theta : E^{\mathbb{N}} &\longrightarrow E^{\mathbb{N}} \\ (x_0, x_1, \dots) &\longmapsto (x_1, x_2, \dots).\end{aligned}$$

Le décalage est donc la restriction à $E^{\mathbb{N}}$, non bijective, du flot bijectif défini sur $E^{\mathbb{Z}}$ et introduit en section ???. Comme pour le flot, on aura souvent besoin du n^{e} itéré de θ , noté θ^n et défini par :

$$\begin{aligned}\theta^n : E^{\mathbb{N}} &\longrightarrow E^{\mathbb{N}} \\ (x_0, x_1, \dots) &\longmapsto (x_n, x_{n+1}, x_{n+2}, \dots).\end{aligned}$$

Par la suite, on identifiera θ et θ^1 .

DÉFINITION 2.1.— La suite X est une chaîne de Markov lorsque pour tout $n \leq m$, la tribu \mathcal{F}_n est indépendante de la tribu $\sigma(X_m)$ conditionnellement à $\sigma(X_n)$. En d'autres termes, pour toute fonction F et G bornées, on a :

$$\begin{aligned}\mathbf{E}[F(X_0, \dots, X_n) G(X_m) | X_n] \\ = \mathbf{E}[F(X_0, \dots, X_n) | X_n] \mathbf{E}[G(X_m) | X_n].\end{aligned}\quad (2.1)$$

En vertu du théorème A.12, on sait que cette propriété est équivalente à l'indépendance du passé et du futur conditionnellement au présent, ce qui s'exprime par :

$$\mathbf{E}[F(X_0, \dots, X_n) G \circ \theta^n | \mathcal{F}_n] = F(X_0, \dots, X_n) \mathbf{E}[G \circ \theta^n | X_n]. \quad (2.2)$$

En particulier, pour $G = \mathbf{1}_{\{y\}}(X_1)$, pour tout entier n , on obtient :

$$\mathbf{P}(X_{n+1} = y | \mathcal{F}_n) = \mathbf{P}(X_{n+1} = y | X_n).$$

DÉFINITION 2.2.— La chaîne de Markov X est dite homogène lorsque :

$$\mathbf{P}(X_{n+1} = y | X_n = x)$$

ne dépend pas de n mais seulement de x et y . On notera cette quantité $p(x, y)$ et on appelle $P = (\mathbf{P}(X_1 = y | X_0 = x), x, y \in E)$ l'opérateur de transition de X . Si E est de cardinal fini, P s'identifie à une matrice qui a autant de lignes et de colonnes que d'éléments dans E .

EXEMPLE 2.1.— Un rat se déplace dans le labyrinthe à sept cases représenté dans la figure 2.1. Il passe d'une case à l'autre uniformément suivant les possibilités qui lui sont offertes, c'est-à-dire que lorsqu'il y a deux (respectivement trois) sorties dans la case où il se trouve, il va dans chacune des cases possibles avec une probabilité d'un demi (respectivement d'un tiers). Son évolution est sans mémoire : chaque changement ne dépend que de la situation courante, pas du passé. On appelle X_n la position du rat après son n^{e} mouvement, X_0 est sa position initiale.

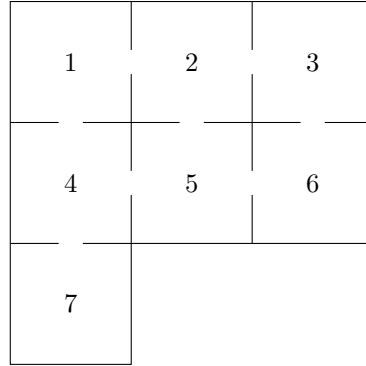


Figure 2.1 – Le labyrinthe

Ici, $E = \{1, 2, 3, 4, 5, 6, 7\}$ et la matrice de transition se déduit aisément de la figure 2.1 :

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

EXEMPLE 2.2.– Partant d'un score vierge, on lance deux dés non pipés. Si la somme est différente de 7, on ajoute cette somme au score courant et on rejoue. Sinon le score s'annule et la partie s'arrête. On appelle X_n le score après le n^e lancer. Il faut distinguer ici deux états 0 si l'on veut que X soit une chaîne homogène. En effet, on peut quitter le 0 du départ mais on ne peut pas quitter le 0 consécutif à un 7. On prend donc espace d'états, $E = \mathbf{N} \cup \{\delta\}$, δ est ce qu'il est convenu d'appeler un point cimetièrè. Les transitions sont alors données pour tout $i \neq \delta$ par :

$$\begin{aligned} \mathbf{P}(X_{n+1} = i + 2 \mid X_n = i) &= \mathbf{P}(X_{n+1} = i + 12 \mid X_n = i) &= 1/36 \\ \mathbf{P}(X_{n+1} = i + 3 \mid X_n = i) &= \mathbf{P}(X_{n+1} = i + 11 \mid X_n = i) &= 2/36 \\ \mathbf{P}(X_{n+1} = i + 4 \mid X_n = i) &= \mathbf{P}(X_{n+1} = i + 10 \mid X_n = i) &= 3/36 \\ \mathbf{P}(X_{n+1} = i + 5 \mid X_n = i) &= \mathbf{P}(X_{n+1} = i + 9 \mid X_n = i) &= 4/36 \\ \mathbf{P}(X_{n+1} = i + 6 \mid X_n = i) &= \mathbf{P}(X_{n+1} = i + 8 \mid X_n = i) &= 5/36 \\ \mathbf{P}(X_{n+1} = \delta \mid X_n = i) &= 1/6 \\ \mathbf{P}(X_{n+1} = \delta \mid X_n = \delta) &= 1. \end{aligned}$$

La définition même d'une chaîne de Markov implique que toute son évolution est déterminée par la loi de la position initiale que l'on notera ν et l'opérateur de transition P .

THÉORÈME 2.1.— Pour tout n , la loi jointe de (X_0, \dots, X_n) est déterminée par la loi de X_0 et P à partir de la formule suivante :

$$\mathbf{P}(X_0 = x_0, \dots, X_m = x_m) = \nu(\{x_0\}) \prod_{l=0}^{m-1} p(x_l, x_{l+1}),$$

pour tout n et tout x_0, \dots, x_n dans E .

REMARQUE.— Par la suite, nous noterons \mathbf{P}_ν la loi d'une chaîne de Markov avec loi initiale ν . Par abus de notation, \mathbf{P}_x représentera la loi de la chaîne si le point de départ est fixe, égal à $x \in E$. Comme E est au plus dénombrable, on peut numéroté les états : utiliser une injection qui existe entre E et \mathbf{N} . On peut alors supposer que $E \subset \mathbf{N}$. On retrouve alors le formalisme des vecteurs et matrices même si l'on est amené à manipuler de tels objets avec un nombre infini de composantes... On considère souvent le « vecteur » π_n défini par $\pi_n(x) = \mathbf{P}(X_n = x)$ pour $x \in E \subset \mathbf{N}$. Il est usuel de le considérer comme un vecteur ligne. Pour tout n , la relation :

$$\begin{aligned} \mathbf{P}(X_{n+1} = j) &= \sum_{i \in E} \mathbf{P}(X_{n+1} = j \mid X_n = i) \mathbf{P}(X_n = i) \\ &= \sum_{i \in E} \mathbf{P}(X_n = i) p(i, j), \quad j \in E, \end{aligned}$$

se lit en notation matricielle :

$$\pi_{n+1} = \pi_n \cdot P \text{ soit } \pi_n = \pi_0 \cdot P^n \quad (2.3)$$

où P^n est la n^{e} puissance de P . En particulier, si π_0 n'est composé que de 0 sauf un 1 en i^{e} position (ce qui revient à travailler sous \mathbf{P}_i) alors pour tout $j \in E$, on a :

$$\mathbf{P}_i(X_n = j) = p^{(n)}(i, j)$$

où $p^{(n)}(i, j)$ est le terme en i^{e} ligne et j^{e} colonne de P^n .

Comme $P^{n+m} = P^n P^m$, on déduit de (2.3) l'équation de Chapman-Kolmogorov :

$$p^{(n+m)}(x, y) = \sum_{z \in E} p^{(n)}(x, z) p^{(m)}(z, y), \quad (2.4)$$

valable pour tout n, m , toute condition initiale et tout état final. On l'a écrite ici sous forme « intrinsèque », c'est-à-dire sans tenir compte de l'injection mentionnée plus haut.

2.1.1. Simulation

Rappelons d'abord comment simuler une loi ν sur un ensemble au plus dénombrable E . Les états sont numérotés grâce à une bijection ϕ entre E et un sous-ensemble de \mathbb{N} . On pose ensuite :

$$r_0 = \nu(\{\phi^{-1}(0)\}) \text{ et } r_n = \sum_{j=0}^n \nu(\{\phi^{-1}(j)\}) = \nu(\phi^{-1}(\{0, \dots, n\})).$$

Algorithme 2.1 . Réalisation d'une variable aléatoire de loi ν

Données : r_0, r_1, \dots
Résultat : un élément de E choisi selon la loi ν
 $x \leftarrow$ réalisation d'une loi uniforme sur $[0, 1]$;
 $n \leftarrow 0$;
tant que $x > r_n$ **faire**
 $n \leftarrow n + 1$
fin
retourner $\phi^{-1}(n)$

Dans une chaîne de Markov, lorsque l'on est à l'état x , on passe à l'état y avec probabilité $p(x, y)$. Pour passer d'une étape à l'autre, il suffit d'appliquer l'algorithme précédent à la loi $\mu = (p(x, y), y \in E)$.

Algorithme 2.2 . Simulation d'une trajectoire d'une chaîne de Markov (ν, P)

Données : ν, P, N
Résultat : une trajectoire à N pas de la chaîne de Markov (ν, P)
Choisir x_0 condition initiale selon ν ;
pour compteur $\leftarrow 1$ **à** N **faire**
 Choisir x_{compteur} selon la loi $(p(x_{\text{compteur}-1}, y), y \in E)$;
fin
retourner x_0, x_1, \dots, x_N

2.2. Propriété de Markov forte

Pour T temps d'arrêt, sur $(T < \infty)$, on définit θ^T par :

$$\theta^T(\omega) = (\omega_{T(\omega)}, \omega_{T(\omega)+1}, \dots).$$

Pour $x \in E$, on définit :

$$\tau_x^1 = \begin{cases} \infty & \text{si } X_n \neq x \text{ pour tout } n > 0; \\ \inf\{n > 0, X_n = x\} & \text{sinon.} \end{cases}$$

$$\tau_x^k = \begin{cases} \infty & \text{si } \tau_x^{k-1} = \infty \text{ ou } X_n \neq x \text{ pour tout } i > \tau_x^{k-1} \\ \inf\{n > \tau_x^{k-1}, X_n = x\} & \text{sinon.} \end{cases}$$

Pour tout k , τ_x^k est l'instant du k^{e} passage de la chaîne X à l'état x .

LEMME 2.2.— Pour x fixé dans E , sur l'événement $\{\tau_x^1 < \infty\}$, on a :

$$\tau_x^k = \tau_x^{k-1} + \tau_x^1 \circ \theta^{\tau_x^{k-1}}. \quad (2.5)$$

Démonstration. Si $\tau_x^{k-1} = \infty$ alors on a ∞ des deux côtés de l'égalité. Sinon, le résultat est immédiat à partir du moment où l'on se persuade que $\theta^{\tau_x^{k-1}}(\omega)$ représente la partie de la trajectoire postérieure à la $(k-1)^{\text{e}}$ visite à l'état x . Par conséquent, la première visite après la $(k-1)^{\text{e}}$ (si elle existe) est la k^{e} visite depuis le début. \square

THÉORÈME 2.3.— Soit T un temps d'arrêt presque sûrement fini et $F : \Omega \rightarrow \mathbf{R}^+$ une variable aléatoire intégrable. On a l'identité suivante :

$$\mathbf{E}[F \circ \theta^T | \mathcal{F}_T] = \mathbf{E}[F | X_0 = X_T]. \quad (2.6)$$

Pour calculer le terme de droite, on calcule $\mathbf{E}[F | X_0 = x] = \phi(x)$ et l'on a :

$$\mathbf{E}[F | X_0 = X_T] = \phi(X_T).$$

Démonstration. Pour $A \in \mathcal{F}_T$, puisque $A \cap \{T = n\} \in \mathcal{F}_n$ et en utilisant (2.2) et les propriétés de l'espérance conditionnelle, il vient :

$$\begin{aligned} \mathbf{E}[F \circ \theta^T \cdot \mathbf{1}_A] &= \sum_{n=0}^{\infty} \mathbf{E}[F \circ \theta^n \cdot \mathbf{1}_{A \cap \{T=n\}}] \\ &= \sum_{n=0}^{\infty} \mathbf{E}[\mathbf{E}[F \circ \theta^n | \mathcal{F}_n] \mathbf{1}_{A \cap \{T=n\}}] \\ &= \sum_{n=0}^{\infty} \mathbf{E}[\mathbf{E}[F | X_0 = X_n] \mathbf{1}_{A \cap \{T=n\}}] \\ &= \mathbf{E}[\mathbf{E}[F | X_0 = X_T] \mathbf{1}_A] \end{aligned}$$

L'égalité est vraie par linéarité pour toutes les fonctions étagées et donc pour toutes les fonctions positives. \square

REMARQUE.— Soit la chaîne de Markov à deux états 0 et 1, et de matrice de transition $p_{0,0} = 0,9, p_{1,1} = 1$. Soit $T = \sup\{n \geq 1, X_n = 0\}$. Sous \mathbf{P}_0 , $T = Y - 1$ où Y suit une loi géométrique de paramètre 0,1. Il s'ensuit T est presque sûrement fini. Or, $\mathbf{P}_0(X_{T+1} = 1 | X_T = 0) = 1$, qui est différent de $\mathbf{P}_0(X_{n+1} = 1 | X_n = 0) = 0,1$.

Cet exemple illustre que l'on ne peut pas supprimer l'hypothèse « T temps d'arrêt » dans la propriété de Markov forte. Il est clair qu'ici T n'est pas un temps d'arrêt car savoir si T est inférieur à n nécessite de connaître la trajectoire après l'instant n pour être sûr que l'on ne repassera pas par 0 après n .

EXEMPLE (Suite de 2.1).— Supposons qu'il y ait un bout de fromage en case 3 et une batterie en case 7. On veut calculer la probabilité que le rat puisse manger avant d'être électrocuté. Introduisons :

$$\tau_3 = \inf\{n \geq 0, X_n = 3\} \text{ et } \tau_7 = \inf\{n \geq 0, X_n = 7\}.$$

On pose $u_i = \mathbf{P}_i(\tau_3 < \tau_7)$. Il est clair que $u_3 = 1$ et que $u_7 = 0$. Pour $i \notin \{3; 7\}$,

$$u_i = \sum_{j=1}^7 \mathbf{P}_i(\tau_3 < \tau_7 | X_1 = j) \mathbf{P}_i(X_1 = j).$$

Puisque i est différent de 3 et 7, l'événement $(\tau_3 < \tau_7)$ est \mathbf{P}_i presque sûrement égal à A_1 où :

$$\begin{aligned} A_l &= \{\omega, \exists i \geq l \text{ tel que } \omega_i = 3 \text{ et } \omega_j \in \{1, 2, 4, 5, 6\} \text{ pour tout } l \leq j < i\} \\ &= \{\text{postérieurement à l'instant } l, \text{ on atteint 3 avant 7}\}. \end{aligned}$$

Comme $\mathbf{1}_{A_1} = \mathbf{1}_{A_0} \circ \theta$, on a :

$$\mathbf{P}_i(\tau_3 < \tau_7 | X_1 = j) = \mathbf{P}_j(\tau_3 < \tau_7).$$

Compte tenu du fait que $\mathbf{P}_i(X_1 = j) = p(i, j)$ on voit que les u_i sont solutions du système linéaire :

$$u_3 = 1, u_7 = 0, u_i = \sum_{j=1}^6 p(i, j) u_j \text{ pour } i \notin \{3; 7\}.$$

La résolution de ce système donne $u_1 = 7/12, u_2 = 3/4, u_4 = 5/12, u_5 = 2/3, u_6 = 5/6$.

Sans fromage et batterie, calculons maintenant le temps moyen d'atteinte de la case 3. Posons $v_i = \mathbf{E}_i[\tau_3]$. Il est clair $v_3 = 0$. Par ailleurs, pour $i \neq 3$, on a :

$$\mathbf{E}_i[\tau_3] = \sum_{j=1}^7 \mathbf{E}_i[\tau_3 | X_1 = j] p(i, j).$$

Si l'on a comme trajectoire $\omega = (1, 2, 5, 2, 5, 6, 3, \dots)$ alors $\tau_3(\omega) = 6$ mais $\tau_3(\theta\omega) = 5$. Plus généralement, on a, conditionnellement à $X_0 \neq 3$, $\tau_3 = \tau_3 \circ \theta + 1$. Par conséquent, on a :

$$v_i = \sum_{j=1}^7 (\mathbf{E}_i [\tau_3 \circ \theta \mid X_1 = j] + 1) p(i, j) = \sum_{j=1}^7 p(i, j) v_j + 1,$$

d'après la relation (2.2). Les v_i sont donc les solutions d'un système linéaire à six équations et six inconnues qu'il ne vous reste plus, cher lecteur, qu'à résoudre.

2.3. Classification des états

On note N_x le nombre de visites à l'état x après le départ :

$$N_x = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_n = x\}}.$$

LEMME 2.4.– Pour tout k , les deux événements $\{N_x \geq k\}$ et $\{\tau_x^k < \infty\}$ coïncident.

Démonstration. $N_x \geq k$ signifie qu'il y a eu plus de k visites à l'état x , ce qui est très exactement équivalent à dire que $\tau_x^k < \infty$. \square

DÉFINITION 2.3.– Un état x est dit récurrent lorsque $\mathbf{P}_x(\tau_x^1 < \infty) = 1$. Sinon, x est dit transient. La chaîne X est dite récurrente (respectivement transiente) si tous ses états sont récurrents (respectivement transients).

LEMME 2.5.– Pour tout couple (x, y) de E , on a :

$$\mathbf{P}_y(\tau_x^k < \infty) = \mathbf{P}_x(\tau_x^1 < \infty)^{k-1} \mathbf{P}_y(\tau_x^1 < \infty). \quad (2.7)$$

En particulier, si $x = y$, $\mathbf{P}_x(\tau_x^k < \infty) = \mathbf{P}_x(\tau_x^1 < \infty)^k$. Par ailleurs, on a :

$$\mathbf{E}_y[N_x] = \frac{\mathbf{P}_y(\tau_x^1 < \infty)}{1 - \mathbf{P}_x(\tau_x^1 < \infty)} = \sum_{n \geq 1} p^{(n)}(y, x). \quad (2.8)$$

☞ Prenant comme instant présent celui de la k^{e} visite à l'état x , d'après la propriété de Markov forte, le passé et le futur conditionnellement à cette visite sont indépendants. Par conséquent, sachant que l'on a déjà visité k fois l'état x , la probabilité que l'on revienne en x une $(k+1)^{\text{e}}$ fois est la même que lors de la première visite en x on revienne au moins une fois. De plus, ces deux événements sont indépendants.

Démonstration. Pour $k > 2$, d'après (2.5) et (2.6), on a :

$$\begin{aligned}
\mathbf{P}_y(\tau_x^k < \infty) &= \mathbf{P}_y(\tau_x^{k-1} < \infty, \tau_x^1 \circ \theta^{\tau_x^{k-1}} < \infty) \\
&= \mathbf{E}_y \left[\mathbf{1}_{\{\tau_x^{k-1} < \infty\}} \mathbf{P}_y(\tau_x^1 \circ \theta^{\tau_x^{k-1}} < \infty \mid \mathcal{F}_{\tau_x^{k-1}}) \right] \\
&= \mathbf{E}_y \left[\mathbf{1}_{\{\tau_x^{k-1} < \infty\}} \mathbf{P}_y(\tau_x^1 < \infty \mid X_0 = X_{\tau_x^{k-1}}) \right] \\
&= \mathbf{E}_y \left[\mathbf{1}_{\{\tau_x^{k-1} < \infty\}} \mathbf{P}_y(\tau_x^1 < \infty \mid X_0 = x) \right] \\
&= \mathbf{E}_y \left[\mathbf{1}_{\{\tau_x^{k-1} < \infty\}} \mathbf{P}_y(\tau_x^1 < \infty) \right] \\
&= \mathbf{P}_y(\tau_x^{k-1} < \infty) \mathbf{P}_y(\tau_x^1 < \infty),
\end{aligned}$$

et on retrouve (2.7) par récurrence.

Maintenant, d'après le théorème de Fubini, on peut écrire :

$$\mathbf{E}_y [N_x] = \sum_{k \geq 1} \mathbf{P}_y(N_x \geq k) = \sum_{k \geq 1} \mathbf{P}_y(\tau_x^k < \infty).$$

La première égalité de (2.8) en découle. Toujours d'après le théorème de Fubini et (2.4), on a :

$$\mathbf{E}_y [N_x] = \mathbf{E}_y \left[\sum_{n=1}^{\infty} \mathbf{1}_{\{X_n=x\}} \right] = \sum_{n=1}^{\infty} \mathbf{E}_y [\mathbf{1}_{\{X_n=x\}}] = \sum_{n=1}^{\infty} p^{(n)}(y, x),$$

d'où le résultat. \square

Le théorème suivant permet de donner les différentes caractérisations de la récurrence et de la transience.

THÉORÈME 2.6.— Soit x un état fixé.

1) Les assertions suivantes sont équivalentes :

- a) x est récurrent,
- b) $\mathbf{P}_x(N_x = \infty) = 1$,
- c) $\mathbf{E}_x [N_x] = \infty$.

2) Les assertions suivantes sont équivalentes :

- a) x est transient,
- b) $\mathbf{P}_x(N_x < \infty) = 1$,
- c) $\mathbf{E}_x [N_x] < \infty$.

Démonstration. Montrons d'abord que $a \Rightarrow b$. D'après (2.4) et le lemme 2.5 :

$$\mathbf{P}_x(N_x > k) = \mathbf{P}_x(\tau_x^k < \infty) = \mathbf{P}_x(\tau_x < \infty)^k, \quad (2.9)$$

et d'après le théorème de convergence monotone :

$$\mathbf{P}_x(N_x = \infty) = \lim_{k \rightarrow \infty} \mathbf{P}_x(N_x > k). \quad (2.10)$$

La récurrence de x signifie $\mathbf{P}_x(\tau_x < \infty) = 1$ et implique donc $N_x = \infty$, \mathbf{P}_x presque sûrement. Par conséquent, x récurrent implique que $\mathbf{P}_x(\tau_x^1 < \infty) = 1$. Par le même raisonnement, x transient implique $\mathbf{P}_x(\tau_x^1 < \infty) < 1$.

$b \Rightarrow c$. Immédiat dans le cas de x récurrent. Pour l'autre cas, utilisons la relation :

$$\mathbf{E}_x[N_x] = \sum_{k \geq 0} \mathbf{P}_x(\tau_x < \infty)^k. \quad (2.11)$$

Comme N_x est fini presque sûrement, d'après (2.10), $\mathbf{P}_x(N_x > k)$ tend vers 0 quand k tend vers l'infini. D'après (2.9) ceci implique que $\mathbf{P}_x(\tau_x < \infty) < 1$ donc que la série converge.

$c \Rightarrow a$. Dans les deux cas, la relation (2.11) permet de conclure. \square

DÉFINITION 2.4.— On dit qu'un état x conduit à un état y et on le note $x \longrightarrow y$, s'il existe un entier strictement positif m tel que $p^{(m)}(x, y) > 0$. Ce qui revient à dire que $\mathbf{P}_x(\tau_y^1 < \infty) = 1$.

THÉORÈME 2.7.— Si x est un état récurrent et $x \longrightarrow y$, alors $y \longrightarrow x$ et y est récurrent.

⚡ Partant de x on sait que l'on va presque sûrement en y , si de y il y a un risque de ne pas revenir en x on va finir par effectivement ne pas y revenir ; on ne fera donc qu'un nombre fini de visites à x ce qui est incompatible avec l'hypothèse de récurrence. De plus, si de y on est presque sûr de revenir en x et que l'on passe une infinité de fois par x on passera vraisemblablement une infinité de fois en y aussi.

Démonstration. Montrons par l'absurde que y conduit à x en écrivant que la probabilité de ne jamais revenir en x en étant parti de x est supérieure à la probabilité de la même chose mais en passant une fois par y :

$$\begin{aligned} \mathbf{P}_x(\tau_x = \infty) &\geq \mathbf{P}_x(\tau_x \circ \theta^{\tau_y} = \infty, \tau_y < \infty) \\ &= \mathbf{P}_x(\tau_y < \infty) \mathbf{P}_y(\tau_x = \infty), \end{aligned}$$

d'après la propriété de Markov forte. Si y ne conduit pas à x , cette quantité est strictement positive ce qui est en contradiction avec la récurrence de x . De même, on a :

$$\begin{aligned}\mathbf{P}_y(\tau_y < \infty) &\geq \mathbf{P}_y(\tau_y \circ \theta^{\tau_x} < \infty, \tau_x < \infty) \\ &= \mathbf{P}_y(\tau_x < \infty)\mathbf{P}_x(\tau_y < \infty) = 1,\end{aligned}$$

donc y est récurrent. \square

THÉORÈME 2.8.— La relation \longrightarrow restreinte aux états récurrents est une relation d'équivalence.

Démonstration. La réflexivité, c'est-à-dire $x \longrightarrow x$, est induite par la définition même d'un état récurrent. La symétrie, c'est-à-dire $x \longrightarrow y \implies y \longrightarrow x$, découle du théorème 2.7. Soit x, y et z trois états de E tels que $x \longrightarrow y$ et $y \longrightarrow z$. Par définition, il existe deux entiers positifs que nous appelons r et s tels que $p^{(r)}(x, y) > 0$ et $p^{(s)}(y, z) > 0$. L'équation de Chapman-Kolmogorov implique que :

$$p^{(r+s)}(x, z) = \sum_{\ell \in E} p^{(r)}(x, \ell)p^{(s)}(\ell, z).$$

Tous les termes de cette somme sont positifs et il existe au moins un terme strictement positif : $p^{(r)}(x, y)p^{(s)}(y, z)$. Nous avons donc trouvé un entier positif, $r + s$, tel que $p^{(r+s)}(x, z) > 0$, d'où le résultat. \square

L'ensemble des points récurrents peut donc être partitionné en classes d'équivalence. Par définition, un état appartenant à une classe conduit à tous les autres états de cette classe et ne conduit à aucun état récurrent appartenant à une autre classe ni à un état transient. En revanche, un état transient peut conduire aussi bien à un état transient qu'à un état récurrent.

DÉFINITION 2.5.— Un sous-ensemble F de E est dit fermé, si pour tout x et y :

$$(x \in F \text{ et } x \longrightarrow y) \implies y \in F.$$

Autrement dit, $\sum_{y \in F} p(x, y) = 1$ pour tout $x \in F$.

THÉORÈME 2.9.— Tout ensemble fermé de cardinal fini contient au moins un point récurrent.

Démonstration. Soit F un ensemble fermé, si tous ses états sont transients, on a :

$$\mathbf{E}_y[N_x] = \mathbf{P}_y(\tau_x^1 < \infty)\mathbf{E}_x[N_x] < \infty$$

pour tout couple (x, y) de F . Comme F est de cardinal fini, $\sum_{x \in F} \mathbf{E}_y [N_x] < \infty$. Or, on a la suite d'identités :

$$\begin{aligned} \sum_{x \in F} \mathbf{E}_y [N_x] &= \sum_{x \in F} \mathbf{E}_y \left[\sum_{n \geq 0} \mathbf{1}_{\{X_n = x\}} \right] = \sum_{n \geq 0} \mathbf{E}_y \left[\sum_{x \in F} \mathbf{1}_{\{X_n = x\}} \right] \\ &= \sum_{n \geq 0} 1 = \infty, \end{aligned}$$

puisque F est fermé. On a abouti à une absurdité, il existe donc au moins un point récurrent. \square

EXEMPLE 2.3.— Il est souvent simple d'avoir une représentation graphique de la matrice de transition d'une chaîne de Markov. Pour ce faire, on construit un graphe orienté dont les sommets correspondent aux états. L'arête (orientée) x, y a comme poids la probabilité de transition de x vers y . Si cette probabilité est nulle, l'arête n'existe pas. Considérons la chaîne de Markov de matrice de transition :

$$P = \begin{pmatrix} 0.8 & 0.1 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{1000} & \frac{999}{1000} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

La représentation graphique est alors celle de la figure 2.2.

Il est clair que les ensembles $\{2, 3\}$ et $\{4, 5, 6\}$ sont des ensembles fermés. Comme ils sont tous deux de cardinal fini, ils contiennent tous deux au moins un point récurrent. A l'intérieur de chacun d'entre eux, les points conduisent tous les uns aux autres, donc ils sont tous récurrents. Si l'on quitte le point 1 pour le point 2 ou 4, on n'est sûr de ne pas y revenir donc la probabilité de ne pas revenir en 1 est $2 * 1/10 > 0$ donc 1 est transient.

Remarquons que même si l'ensemble E est fermé, le théorème 2.9 n'induit pas de contradiction quant au statut de l'état 1. En effet, on sait qu'il existe au moins un point récurrent dans E mais on ne sait pas lequel et on ne peut, *a priori*, rien dire de plus puisque tous les états ne communiquent pas entre eux.

EXEMPLE (Suite de l'exemple 2.1).— Tous les états conduisent les uns aux autres donc le seul sous-ensemble fermé est E lui-même. Comme il est de cardinal fini, il existe au moins un état récurrent donc ils sont tous récurrents.

EXEMPLE (Suite de l'exemple 2.2).— Tant qu'il n'atteint pas δ , le score ne peut que croître donc tous les états de \mathbb{N} sont transients. δ est lui récurrent.

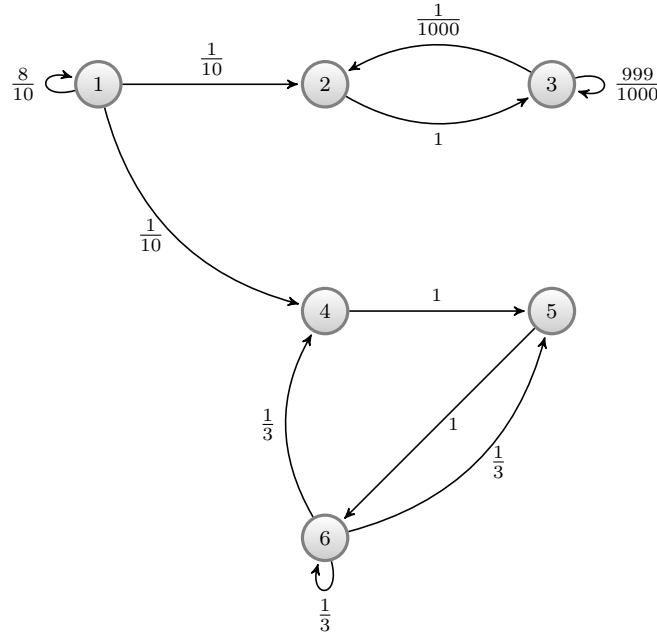


Figure 2.2 – Représentation graphique d'une chaîne de Markov

EXEMPLE.— Un jeu de N cartes est mélangé en le coupant en deux parts qui sont ensuite interverties. Chaque mélange du jeu est représenté par une permutation de $\{1, \dots, N\}$. Si $N = 3$ et que le mélange est représenté par $(3, 2, 1)$, cela signifie que la carte 3 est en position 1, la carte 2 en position 2 et la carte 1 en position 3. On note X_n l'état du paquet de cartes après la n^{e} opération de mélange. L'espace d'états est donc le groupe des permutations de $\{1, \dots, N\}$ dans lui-même, noté \mathfrak{S}_N . Si $X_0 = (3, 2, 1)$ et que la coupe se fait entre la première et la deuxième carte, on a $X_1 = (2, 1, 3)$. En d'autres termes, on a juste fait une permutation circulaire sur les cartes mais on n'a pas changé leur ordre relatif. Pour définir les probabilités de transition, considérons l'ensemble à N éléments :

$$E_1 = \left\{ \sigma \in \mathfrak{S}_N, \exists k \in \{1, \dots, N\}, \sigma = (k+1, k+2, \dots, N, 1, \dots, k) \right\}.$$

Lorsque l'on coupe le paquet au niveau de la k^{e} carte, on applique le cycle $(k+1, k+2, \dots, N, 1, \dots, k)$ à la situation courante. Comme le choix de l'endroit de la coupe est supposé être uniforme sur $\{1, \dots, N\}$, on a :

$$\mathbf{P}(X_1 = \tau \mid X_0 = \sigma) = \frac{1}{N} \text{ si } \tau\sigma^{-1} \in E_1.$$

Les classes d'équivalence de la relation \longrightarrow sont celles de la relation $\sigma \mathfrak{R} \tau \equiv \tau \sigma^{-1} \in E_1$. En d'autres termes, σ conduit à τ si et seulement s'il existe $\rho \in E_1$ tel que $\tau = \rho \sigma$. Il y a donc $(n-1)!$ classes d'équivalence de cardinal n chacune. Toutes ces classes forment des ensembles fermés de cardinal fini qui contiennent donc toutes au moins un point récurrent. Comme à l'intérieur de ces classes les états communiquent tous entre eux, ils sont tous récurrents. La chaîne est donc récurrente.

Lorsque l'espace d'état est infini, on ne peut pas appliquer le théorème 2.9. On introduit alors la notion suivante.

DÉFINITION 2.6.— Une chaîne de Markov est dite irréductible lorsque tous les états conduisent les uns aux autres. En particulier, le plus petit sous-espace fermé est E lui-même et tous les états ont même nature.

REMARQUE.— Si le nombre d'états transients est fini, comme l'on ne passe qu'un nombre fini de fois en chacun d'eux, la chaîne de Markov sera inexorablement contrainte à aller dans une classe de récurrence et à y rester. Remarquons qu'en vertu du lemme 2.7, une classe de récurrence est forcément un sous-ensemble irréductible. Si le nombre d'états transients est infini, le raisonnement précédent ne s'applique plus automatiquement mais les cas dans lesquels on n'atterrit pas automatiquement dans un sous-ensemble fermé irréductible sont hors de notre propos. Pour ce qui nous intéresse (le comportement asymptotique des chaînes de Markov), il n'y a donc pas de perte de généralité à supposer que les chaînes de Markov étudiées sont irréductibles.

Lorsque x est récurrent, on sait que partant de x on reviendra nécessairement en x en un temps fini mais *quid* du temps moyen de retour en x ?

DÉFINITION 2.7.— Un état récurrent x est dit récurrent positif, si $\mathbf{E}_x [\tau_x^1] < \infty$; récurrent nul, si $\mathbf{E}_x [\tau_x^1] = \infty$.

La chaîne X est alors dite récurrente positive (respectivement récurrente nulle) si tous ses états sont récurrents positifs (respectivement récurrents nuls).

La construction suivante est utilisée plusieurs fois par la suite.

DÉFINITION 2.8.— Soit X , une chaîne de Markov irréductible et récurrente sur E et F , un sous-ensemble de E . On note :

$$\tau_F^1 = \inf \left\{ n \geq 1, X_n \in F \right\} \text{ et } \tau_F^{k+1} = \tau_F^k + \tau_F^1 \circ \theta^{\tau_F^k},$$

les instants de visite successifs de la chaîne à l'ensemble F . On considère la suite aléatoire X^F , définie par $X_n^F = X_{\tau_F^n}$, $n \in \mathbb{N}$. On vérifie facilement que X^F est une chaîne de Markov sur F , appelée chaîne de Markov restreinte à F .

THÉORÈME 2.10.— Soit X une chaîne de Markov irréductible et F un sous-ensemble fini de E . Si pour tout $x \in F$, $\mathbf{E}_x [\tau_F^1] < \infty$ alors X est récurrente positive.

Démonstration. Soit pour tout $x \in F$, $\sigma_x = \inf\{n > 0, X_n^F = x\}$ et pour tout $k \in \mathbb{N}^*$, $Y_k = \tau_F^k - \tau_F^{k-1}$. Comme F est fini, X^F est récurrente positive donc $\mathbf{E}_x [\sigma_x] < \infty$ pour tout $x \in F$. Il nous faut prouver que $\mathbf{E}_x [\tau_x] < \infty$. Par construction des variables Y_k , on a :

$$\begin{aligned} \mathbf{E}_x [\tau_x] &= \mathbf{E}_x \left[\sum_{k=1}^{\sigma_x} Y_k \right] \\ &= \sum_{n \geq 1} \mathbf{E}_x \left[\sum_{k=1}^{\sigma_x} Y_k \mathbf{1}_{\{\sigma_x = n\}} \right] \\ &= \sum_{k=1}^{\infty} \mathbf{E}_x \left[Y_k \sum_{n \geq k} \mathbf{1}_{\{\sigma_x = n\}} \right] \\ &= \sum_{k=1}^{\infty} \mathbf{E}_x [Y_k \mathbf{1}_{\{\sigma_x \geq k\}}]. \end{aligned}$$

En utilisant la propriété de Markov forte, on obtient :

$$\begin{aligned} \mathbf{E}_x [Y_k \mathbf{1}_{\{\sigma_x \geq k\}}] &= \mathbf{E}_x \left[\mathbf{E}_x [Y_k | \mathcal{F}_{\tau_{F,k-1}^F}] \mathbf{1}_{\{\sigma_x \geq k\}} \right] \\ &= \mathbf{E}_x \left[\mathbf{E}_{X_{\tau_{F,k-1}^F}} [Y_1] \mathbf{1}_{\{\sigma_x \geq k\}} \right] \\ &\leq \sup_{y \in F} \mathbf{E}_y [Y_1] \mathbf{P}_x(\sigma_x \geq k). \end{aligned}$$

Le *supremum* est fini par hypothèse puisque F est fini. On obtient donc :

$$\mathbf{E}_x [\tau_x] \leq c \sum_{k=1}^{\infty} \mathbf{P}_x(\sigma_x \geq k) = c \mathbf{E}_x [\sigma_x].$$

En vertu de la remarque initiale, cela prouve la positive récurrence de X . □

LEMME 2.11.— Soit X une chaîne de Markov et $h : E \times E \longrightarrow \mathbf{R}$, bornée. Pour tout entier n , on a :

$$\mathbf{E} [h(X_n, X_{n+1}) | \mathcal{F}_n] = P(h(X_n, \cdot))(X_n) = \sum_{y \in E} p(X_n, y) h(X_n, y). \quad (2.12)$$

Démonstration. Comme h est bornée, seul reste le calcul de l'espérance conditionnelle. D'après la propriété de Markov :

$$\mathbf{E}[h(X_n, X_{n+1}) | \mathcal{F}_n] = \mathbf{E}[h(X_n, X_{n+1}) | X_n].$$

Soit maintenant $\phi : E \rightarrow \mathbf{R}$ bornée. On a :

$$\begin{aligned} \mathbf{E}[h(X_n, X_{n+1})\phi(X_n)] &= \int \phi(x) \int h(x, y) d\mathbf{P}_{X_{n+1} | X_n=x}(y) d\mathbf{P}_{X_n}(x) \\ &= \int \phi(x) \sum_{y \in E} h(x, y)p(x, y) d\mathbf{P}_{X_n}(x) \\ &= \mathbf{E}[\phi(X_n)P(h(X_n, \cdot))(X_n)]. \end{aligned}$$

Comme la précédente équation est vraie pour toute fonction ϕ , on en déduit (2.12). \square

THÉORÈME 2.12 (Critère de Foster).— Soit E_0 une partie finie de E . S'il existe une fonction $h : E \rightarrow \mathbf{R}$ telle que l'ensemble $\{x \in E, h(x) < K\}$ soit fini pour tout entier K et que :

$$h(y) \geq \mathbf{E}_y[h(X_1)] \text{ pour tout } y \in E_0^c,$$

alors X est récurrente.

Démonstration. En particulier, h est minorée donc quitte à rajouter une constante, on peut supposer $h \geq 0$. Soit le temps d'arrêt $\tau = \inf\{n, X_n \in E_0\}$ et Y définie par $Y_n = h(X_n) \mathbf{1}_{\{\tau > n\}}$. Montrons que Y est une sur-martingale positive dès que $X_0 \in E_0^c$. Soit $x \in E_0^c$, on a :

$$\begin{aligned} \mathbf{E}_x[h(X_{n+1}) \mathbf{1}_{\{\tau > n+1\}} | \mathcal{F}_n] &= \mathbf{1}_{\{\tau > n+1\}} \mathbf{E}_{X_n}[h(X_{n+1})] \\ &\leq \mathbf{1}_{\{\tau > n\}} h(X_n) = Y_n, \end{aligned}$$

car sur $(\tau > n + 1)$, X_n n'appartient pas à E_0 . Y converge donc presque sûrement vers une variable aléatoire Y_∞ .

Supposons que X soit transiente. Soit $x \notin E_0$, pour tout entier K , l'ensemble $\{x, h(x) < K\}$ est fini donc n'est visité qu'un nombre fini de fois par X donc X n'est pas bornée. Comme Y_∞ est finie, nécessairement τ est fini presque sûrement. Ce qui revient à dire que pour $x \notin E_0$, $\mathbf{P}_x(\tau < \infty) = 1$. Partant de E_0^c , on arrive forcément dans E_0 . Soit on reste dans E_0 pour toujours et comme E_0 est fini, E_0 est récurrent et par irréductibilité la chaîne l'est. Soit la chaîne quitte E_0 mais en vertu de ce que l'on vient de démontrer, elle y reviendra. Le nombre de visites à E_0 est donc infini, ce qui implique encore une fois que E_0 est récurrent donc la chaîne l'est aussi. \square

2.4. Mesures et probabilité invariantes

DÉFINITION 2.9.— Soit E un ensemble dénombrable et P une opérateur de transition sur $E \times E$. Une mesure positive finie ν sur E est dite invariante par rapport à P , si et seulement si :

$$\nu = \nu P \text{ c'est-à-dire } \nu(y) = \sum_{x \in E} \nu(x) p(x, y) \text{ pour tout } y \in E. \quad (2.13)$$

Si de plus $\sum \nu(x) = 1$, ν est une probabilité invariante.

REMARQUE.— Si $\pi_0 = \nu$ alors $\pi_n = \pi_0 P^n = \pi_0$.

THÉORÈME 2.13.— Soit x un état récurrent, alors la mesure ν définie par :

$$\nu(y) = \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1 - 1} \mathbf{1}_{\{X_n = y\}} \right] = \sum_{n=0}^{\infty} \mathbf{P}_x(X_n = y, \tau_x^1 > n)$$

est une mesure invariante.

Démonstration. Montrons d'abord l'égalité des deux expressions de ν . Puisque x est récurrent, τ_x^1 est presque sûrement fini donc $\cup_{n \geq 1} \{\tau_x^1 = n\}$ est une partition de Ω . En déconditionnant sur toutes les valeurs possibles de τ_x^1 et en utilisant le théorème de Fubini, on obtient :

$$\begin{aligned} \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1 - 1} \mathbf{1}_{\{X_n = y\}} \right] &= \sum_{\ell=1}^{\infty} \mathbf{E}_x \left[\sum_{n=0}^{\ell-1} \mathbf{1}_{\{X_n = y\}} \mathbf{1}_{\{\tau_x^1 = \ell\}} \right] \\ &= \sum_{n=0}^{\infty} \mathbf{E}_x \left[\sum_{\ell > n} \mathbf{1}_{\{\tau_x^1 = \ell\}} \mathbf{1}_{\{X_n = y\}} \right] \\ &= \sum_{n=0}^{\infty} \mathbf{E}_x \left[\mathbf{1}_{\{\tau_x^1 > n\}} \mathbf{1}_{\{X_n = y\}} \right]. \end{aligned}$$

Comme sous \mathbf{P}_x , $X_0 = X_{\tau_x^1} = x$, on peut aussi écrire $\nu(y) = \mathbf{E}_x \left[\sum_{n=1}^{\tau_x^1} \mathbf{1}_{\{X_n = y\}} \right]$, ce qui donne par les mêmes calculs avec des bornes différentes :

$$\nu(y) = \sum_{n=1}^{\infty} \mathbf{E}_x \left[\mathbf{1}_{\{\tau_x^1 \geq n\}} \mathbf{1}_{\{X_n = y\}} \right]. \quad (2.14)$$

Remarquons d'ores et déjà que l'événement $\{\tau_x^1 \geq n\}$ appartient à \mathcal{F}_{n-1} puisque c'est le complémentaire de l'événement $\{\tau_x^1 \leq n-1\}$. Pour $y \neq x$, en appliquant les

propriétés de l'espérance conditionnelle et la propriété de Markov forte :

$$\begin{aligned}
& \sum_{n=0}^{\infty} \mathbf{E}_x [\mathbf{1}_{\{\tau_x^1 \geq n\}} \mathbf{1}_{\{X_n=y\}}] \\
&= \sum_{n=1}^{\infty} \mathbf{E}_x [\mathbf{1}_{\{\tau_x^1 \geq n\}} \mathbf{E}_x [\mathbf{1}_{\{X_n=y\}} \mid \mathcal{F}_{n-1}]] \\
&= \sum_{n=1}^{\infty} \mathbf{E}_x [\mathbf{1}_{\{\tau_x^1 \geq n\}} \mathbf{E}_x [\mathbf{1}_{\{X_n=y\}} \mid X_{n-1}]] \\
&= \sum_{n=1}^{\infty} \sum_{z \in E} \mathbf{E}_x [\mathbf{1}_{\{\tau_x^1 \geq n\}} \mathbf{1}_{\{X_{n-1}=z\}} \mathbf{E}_x [\mathbf{1}_{\{X_n=y\}} \mid X_{n-1}=z]] \\
&= \sum_{z \in E} p(z, y) \sum_{n=1}^{\infty} \mathbf{E}_x [\mathbf{1}_{\{\tau_x^1 \geq n\}} \mathbf{1}_{\{X_{n-1}=z\}}] \\
&= \sum_{z \in E} \nu(z) p(z, y).
\end{aligned}$$

Pour $y = x$, il est clair que $\nu(x) = 1$ et d'autre part :

$$\begin{aligned}
\sum_{z \in E} \nu(z) p(z, x) &= \sum_{n=0}^{\infty} \sum_{z \in E} p(z, x) \mathbf{P}_x(X_n = z, \tau_x^1 > n) \\
&= \sum_{n=0}^{\infty} \sum_{z \in E} \mathbf{P}_x(X_n = z, X_{n+1} = x, \tau_x^1 > n) \\
&= \sum_{n=0}^{\infty} \sum_{z \in E} \mathbf{P}_x(X_n = z, \tau_x^1 = n + 1) \\
&= \sum_{n=0}^{\infty} \mathbf{P}_x(\tau_x^1 = n + 1) \\
&= \mathbf{P}_x(\tau_x^1 < \infty) = 1.
\end{aligned}$$

On a donc bien $\nu = \nu P$, et il ne reste qu'à vérifier que $\nu(y) < \infty$ pour tout y . C'est vrai pour $x = y$. Pour $y \neq x$, de deux choses l'une : soit x ne conduit pas à y et alors $\nu(y) = 0$, soit x conduit à y et comme x est récurrent, d'après le théorème 2.7, y conduit à x , c'est-à-dire qu'il existe m tel que $p^{(m)}(y, x) > 0$. Comme ν est invariante, $\nu.P^m = \nu$, ce qui implique que :

$$1 = \nu(x) = \sum_{z \in E} \nu(z) p^{(m)}(z, x) \geq \nu(y) p^{(m)}(y, x),$$

et donc $\nu(y) < \infty$. \square

COROLLAIRE 2.14.— Soit X une chaîne de Markov irréductible et récurrente de mesure invariante ν . Soit F un sous-ensemble de E et X^F la chaîne restreinte à F . Alors, X^F est irréductible et récurrente et admet, comme X , pour mesure invariante celle donnée par le théorème 2.13.

Démonstration. Les deux premiers points sont évidents. Pour $y \in F$, le nombre de visites à y de X^F et de X sont les mêmes donc X et X^F la même mesure invariante donnée par le théorème 2.13. \square

THÉORÈME 2.15.— Si la chaîne de Markov X est irréductible et récurrente alors il existe une unique (à un coefficient près) mesure invariante ν telle que pour tout y , $0 < \nu(y) < \infty$. L'unicité à un coefficient multiplicatif près signifie que si ν et ν' sont deux telles mesures alors il existe $c > 0$ tel que $\nu(x) = c\nu'(x)$ pour tout $x \in E$.

Démonstration. Soit μ une mesure invariante et soit $a \in E$. Soit ν la mesure invariante construite dans le théorème 2.13 avec a comme point de départ. Par construction, $\nu(a) = 1$ donc pour toute mesure invariante μ , $\mu(a) = \nu(a)\mu(a)$. Par définition, pour $z \in E \setminus \{a\}$:

$$\mu(z) = \sum_{y \in E} \mu(y)p(y, z) = \mu(a)p(a, z) + \sum_{y \neq a} \mu(y)p(y, z).$$

En itérant cette relation, on obtient :

$$\mu(z) = \mu(a)p(a, z) + \mu(a) \sum_{y \neq a} p(a, y)p(y, z) + \sum_{i \neq a} \sum_{y \neq a} \mu(x)p(x, y)p(y, z),$$

ce qui peut se récrire de la façon suivante :

$$\begin{aligned} \mu(z) &= \mu(a)\mathbf{P}_a(X_1 = z) \\ &\quad + \mu(a)\mathbf{P}_a(X_1 \neq a, X_2 = z) + \mathbf{P}_\mu(X_0 \neq a, X_1 \neq a, X_2 = z). \end{aligned}$$

Par récurrence sur n , on montre alors que pour tout n :

$$\mu(z) = \mu(a) \sum_{m=1}^n \mathbf{P}_a(\tau_a^1 > m, X_m = z) + \mathbf{P}_\mu \left(\bigcap_{y=0}^n (X_y \neq a) \cap X_n = z \right).$$

La dernière probabilité est un terme positif et l'on reconnaît, quand n tend vers l'infini, dans la première somme la définition de ν donc :

$$\mu(z) \geq \mu(a)\nu(z) \text{ pour tout } z \in E.$$

D'autre part, puisque pour tout n , $\mu.P^n = \mu$, on a aussi :

$$\mu(a) = \sum_x \mu(x)p^{(n)}(x, a) \geq \mu(a) \sum_x \nu(x)p^{(n)}(x, a) = \mu(a)\nu(a) = \mu(a).$$

Par conséquent, l'inégalité intermédiaire est une égalité et comme $\mu(x) \geq \mu(a)\nu(x)$, on doit avoir $\mu(x) = \mu(a)\nu(x)$ chaque fois que n est tel que $p^{(n)}(x, a) > 0$. Étant donné que X est irréductible, un tel entier n existe toujours donc $\mu(x) = \mu(a)\nu(x)$ pour tout $x \in E$. \square

THÉORÈME 2.16.— S'il existe une probabilité invariante ν , alors tous les états vérifiant $\nu(y) > 0$ sont récurrents.

Démonstration. Comme $\nu = \nu P^n$, le théorème de Fubini implique que :

$$\sum_{x \in E} \nu(x) \sum_{n \geq 1} p^{(n)}(x, y) = \sum_{n \geq 1} \nu(y) = \infty \text{ si } \nu(y) > 0.$$

D'autre part, d'après le lemme 2.5 :

$$\sum_{n \geq 1} p^{(n)}(x, y) = \frac{\mathbf{P}_x(\tau_y^1 < \infty)}{1 - \mathbf{P}_y(\tau_y^1 < \infty)}.$$

Comme $\mathbf{P}_x(\tau_y^1 < \infty) \leq 1$, on a :

$$\infty \leq \sum_{x \in E} \nu(x) \cdot \frac{1}{1 - \mathbf{P}_y(\tau_y^1 < \infty)},$$

donc $\mathbf{P}_y(\tau_y^1 < \infty) = 1$ puisque ν est finie, ce qui signifie que y est récurrent. \square

THÉORÈME 2.17.— Si X est irréductible et admet ν comme probabilité invariante alors la relation suivante est satisfaite :

$$\nu(x) = \frac{1}{\mathbf{E}_x[\tau_x^1]}.$$

Démonstration. S'il existe x tel que $\nu(x) = 0$ alors comme pour tout n , on a :

$$\nu(x) = \sum_{y \in E} p^{(n)}(y, x)\nu(y),$$

cela signifie que pour tout n et tout y , le produit de $\nu(y)$ et de $p^{(n)}(y, x)$ est nul. Or la chaîne est irréductible donc pour tout y , il existe n_y tel que $p^{(n_y)}(y, x) > 0$, donc $\nu(y) = 0$. Mais alors ν n'est pas une probabilité donc pour tout $x \in E$, $\nu(x) > 0$. D'après le théorème précédent, tous les états sont donc récurrents. On sait donc

que ν est à un coefficient c près donné par le théorème 2.13. Ce coefficient vérifie $c \sum_{y \in E} \nu(y) = 1$, or l'on sait que :

$$\sum_{y \in E} \nu(y) = \sum_{y \in E} \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1-1} \mathbf{1}_{\{X_n=y\}} \right] = \mathbf{E}_x [\tau_x^1],$$

d'après le théorème de Fubini et pour x choisi de façon quelconque dans E . □

Le théorème suivant résume les principaux résultats précédents.

THÉORÈME 2.18.— Si X est irréductible, les trois assertions suivantes sont équivalentes :

- 1) l'un des états est récurrent positif ;
- 2) il existe une probabilité invariante ;
- 3) tous les états sont récurrents positifs.

De plus, la probabilité invariante est donnée par :

$$\nu(y) = \frac{1}{\mathbf{E}_x [\tau_x^1]} \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1-1} \mathbf{1}_{\{X_n=y\}} \right].$$

Démonstration. 1) \Rightarrow 2). En combinant les théorèmes 2.13 et 2.17, on voit que :

$$\nu(y) = \frac{1}{\mathbf{E}_x [\tau_x^1]} \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1-1} \mathbf{1}_{\{X_n=y\}} \right]$$

définit une probabilité invariante. Comme le terme de gauche ne dépend pas de x , on peut choisir $x = y$ et on retrouve bien $\nu(y) = \mathbf{E}_y [\tau_y^1]^{-1}$.

2) \Rightarrow 3). Puisque X est irréductible, on sait que la probabilité invariante est un multiple de celle construite dans le théorème 2.13 et donc que $\pi(y) > 0$ pour tout $y \in E$. D'après le théorème 2.17, cela signifie que tous les états sont récurrents positifs.

3) \Rightarrow 1) est trivial. □

COROLLAIRE 2.19.— Toute chaîne de Markov irréductible sur E de cardinal fini est récurrente positive.

Démonstration. Il existe une mesure invariante μ . Comme l'espace d'états est fini, on peut toujours la normaliser en posant :

$$\nu(x) = \frac{1}{\sum_{y \in E} \mu(y)} \mu(y),$$

et l'on obtient une probabilité invariante. D'après le point 2) du théorème précédent, on en déduit qu'elle est récurrente positive. \square

Quand l'espace d'états est de cardinal infini, on peut utiliser le théorème suivant qui se démontre avec les mêmes outils que le théorème 2.12.

THÉORÈME 2.20 (Critère de Foster).— Supposons qu'il existe $h : E \rightarrow \mathbf{R}$ et $\epsilon > 0$ tels que :

- $\liminf_y h(y) > -\infty$;
- $h(X_1)$ est intégrable ;
- pour tout $y \in E_0^c$, on ait :

$$h(y) - \epsilon \geq \mathbf{E}[h(X_1) \mid X_0 = y].$$

Dans ces conditions, X est récurrente positive.

Soit X une chaîne de Markov irréductible récurrente sur E polonais. Pour tout $x \in E$, on pose $(Y_k^n, k \geq 0) = (X_{k \wedge \tau_x^1})$ et pour tout $n \geq 1$, on définit la n^{e} excursion Y^n de X depuis x par $(Y_k^n, k \geq 0)$, où :

$$Y_k^n = X_{k \wedge \tau_x^1} \circ \theta^{\tau_x^n}.$$

La 0^e excursion coïncide avec X jusqu'à la première visite à x , après Y^0 reste à x . La n^{e} excursion est une chaîne de Markov qui part de x et suit le comportement de la chaîne initiale jusqu'au temps d'atteinte de x suivant. Ensuite, elle reste constamment égale à x . L'évolution de la chaîne est reflétée par Y^{n+1} .

D'après la propriété de Markov forte, les processus $(Y^n, n \geq 0)$ sont indépendants les uns des autres et à partir de $n = 1$, ils ont tous la même loi : pour toute fonction $\psi : E^{\mathbf{N}} \rightarrow \mathbf{R}$:

$$\mathbf{E}[\psi(Y^n)] = \mathbf{E}[\psi(Y^1 \circ \theta^{\tau_x^n})] = \mathbf{E}[\psi(Y^1)].$$

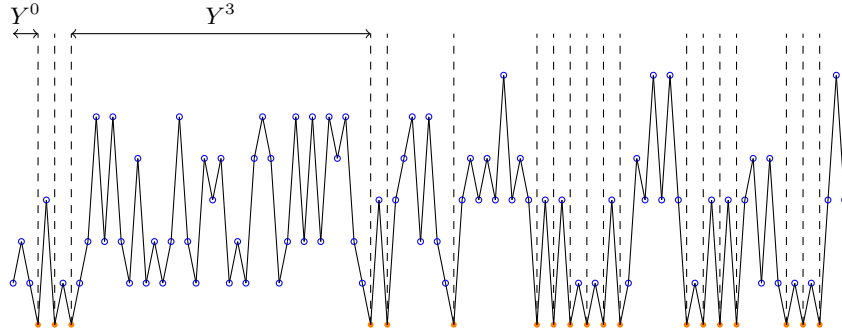


Figure 2.3 – Excursions de la chaîne de Markov de l'exemple 2.1

THÉORÈME 2.21.— Soit X récurrente, irréductible de loi invariante ν . Quelle que soit la condition initiale $x \in E$, pour toute fonction f dans $L^1(\nu)$, pour toute fonction $g \geq 0$ telle que $\sum_y g(y)\nu(y) > 0$, on a :

$$\frac{\sum_{j=0}^n f(X_j)}{\sum_{j=0}^n g(X_j)} \xrightarrow{n \rightarrow \infty} \frac{\sum_{y \in E} f(y)\nu(y)}{\sum_{y \in E} g(y)\nu(y)}, \quad \mathbf{P}_x \text{ presque sûrement.}$$

En particulier, pour $f \in L^1(\nu)$, on a :

$$\frac{1}{n} \sum_{j=0}^n f(X_j) \xrightarrow{n \rightarrow \infty} \sum_{y \in E} f(y)\nu(y), \quad \mathbf{P}_x \text{ presque sûrement.}$$

On peut découper toute fonctionnelle additive en morceaux dépendant de chaque excursion. D'après l'indépendance et l'équidistribution de celles-ci, on peut appliquer la loi forte des grands nombres. Il reste à prouver que les termes de « bord », c'est-à-dire qui dépendent de Y^0 et ceux qui dépendent de l'excursion incomplète disparaissent dans la division par n .

Démonstration. La probabilité invariante est proportionnelle à la mesure invariante donnée dans le théorème 2.13. En particulier, il existe $c > 0$ tel que pour toute fonction $g \geq 0$, on ait :

$$c \sum_{y \in E} g(y)\nu(y) = \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1 - 1} g(X_n) \right].$$

Par homogénéité, on peut supposer que $c = 1$.

Soit $Z = (Z_k, k \geq 1)$ définie par :

$$Z_k = \sum_{n=\tau_x^k}^{\tau_x^{k+1}-1} f(X_n) = \sum_{n=0}^{\tau_x^1-1} f(Y_n^k) = Z_1 \circ \theta^{\tau_x^k}.$$

D'après la propriété de Markov forte, les variables aléatoires $(Z_k, k \geq 1)$ sont indépendantes et identiquement distribuées. De plus, on a :

$$\mathbf{E}_x [|Z_1|] \leq \mathbf{E}_x \left[\sum_{n=0}^{\tau_x^1-1} |f(X_n)| \right] = \sum_{y \in E} |f(y)| \nu(y) < \infty,$$

puisque $f \in L^1(\nu)$. On peut donc appliquer la loi forte des grands nombres, qui stipule que :

$$\frac{1}{n} \sum_{k=0}^{\tau_x^n-1} f(X_k) = \frac{1}{n} \sum_{k=1}^n Z_k \xrightarrow{n \rightarrow \infty} \mathbf{E}_x [Z_1] = \sum_{y \in E} f(y) \nu(y), \quad \mathbf{P}_x \text{ p.s.} \quad (2.15)$$

Si l'on applique ce résultat à $f \equiv 1$, on obtient :

$$\frac{\tau_x^n}{n} \xrightarrow{n \rightarrow \infty} 1, \quad \mathbf{P}_y \text{ presque sûrement.} \quad (2.16)$$

Soit $e(n)$ le nombre de visites à x entre les instants 0 et n . Notons que $e(n)$ est aussi le nombre d'excursions partant de x complètes avant l'instant n . Par définition, $\tau_x^{e(n)} \leq n < \tau_x^{e(n)+1}$, donc :

$$\frac{\tau_x^{e(n)}}{e(n)} \leq \frac{n}{e(n)} < \frac{\tau_x^{e(n)+1}}{e(n)+1} \frac{e(n)+1}{e(n)}.$$

En vertu de (2.16), les termes extrêmes de la ligne précédente tendent p.s. vers 1 donc $n^{-1}e(n)$ aussi. On écrit alors :

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^n f(X_k) = \\ \frac{1}{n} \sum_{k=0}^{\tau_x^1-1} f(X_k) + \frac{e(n)}{n} \frac{1}{e(n)} (Z_1 + \dots + Z_{e(n)}) + \frac{1}{n} \sum_{k=\tau_x^{e(n)}+1}^n f(X_k). \end{aligned} \quad (2.17)$$

Supposons que $f \geq 0$. Le premier terme tend vers 0 \mathbf{P}_y -presque sûrement pour tout $y \in E$. D'après la définition de la convergence presque sûre, cela équivaut à ce que l'on ait pour tout $\epsilon > 0$:

$$\mathbf{P}_y \left(\limsup_n \left(\frac{1}{n} \sum_{k=0}^{\tau_x^1-1} f(X_k) > \epsilon \right) \right) = 0. \quad (2.18)$$

Par conséquent, compte tenu de la propriété de Markov forte, pour tout $\epsilon > 0$, on a :

$$\begin{aligned} \mathbf{P}_y \left(\limsup_n \left(\frac{1}{n} \sum_{k=\tau_x^{e(n)}}^{\tau_x^{e(n+1)}-1} f(X_k) > \epsilon \right) \right) \\ = \mathbf{P}_x \left(\limsup_n \left(\frac{1}{n} \sum_{k=0}^{\tau_x^1-1} f(Y_k^n) > \epsilon \right) \right) \\ = \mathbf{P}_x \left(\limsup_n \left(\frac{1}{n} \sum_{k=0}^{\tau_x^1-1} f(X_k) > \epsilon \right) \right) = 0, \end{aligned}$$

en vertu de (2.18). Comme :

$$\sum_{k=\tau_x^{e(n)}+1}^n f(X_k) \leq \sum_{k=\tau_x^{e(n)}}^{\tau_x^{e(n+1)}-1} f(X_k),$$

on a bien :

$$\frac{1}{n} \sum_{k=\tau_x^{e(n)}}^n f(X_k) \xrightarrow{n \rightarrow \infty} 0, \quad \mathbf{P}_y \text{ presque sûrement.}$$

Pour f de signe quelconque, en appliquant le raisonnement précédent à $|f|$, on montre que les premier et troisième termes de (2.17) tendent presque sûrement vers 0.

D'après la première partie de la preuve (voir (2.15)), le terme médian de (2.17) tend presque sûrement vers $\sum_{y \in E} f(y) \nu(y)$.

Le cas particulier s'obtient en prenant $g = 1$. □

DÉFINITION 2.10.— Un état x est dit périodique s'il existe un entier $\delta \geq 2$ tel que :

$$\sum_{k=1}^{\infty} \mathbf{P}_x(\tau_x^1 = \delta k) = 1. \quad (2.19)$$

Le plus petit δ vérifiant (2.19) est appelé la période de x et nous la noterons $d(x)$. Les états qui ne sont pas périodiques sont appelés apériodiques.

EXEMPLE (Suite de l'exemple 2.1).— Dans ce cas, un peu de réflexion montre que l'on n'atteindra une case de numéro impair que tous les deux pas, de même pour les cases de numéro pair : si le rat part de 1 il ne peut être en case 3, 5 ou 7 au coup d'après. La période est donc 2. On voit que l'on peut classer les états en deux paquets, les cases $\{1, 3, 5, 7\}$ d'une part, les cases $\{2, 4, 6\}$ d'autre part. Le rat sera tous les pas pairs dans le sous-ensemble de départ et tous les pas impairs dans l'autre sous-ensemble.

Plus généralement, on a le théorème suivant.

THÉORÈME 2.22.— Soit X une chaîne de Markov récurrente irréductible de période d . Soit x fixé dans E , il existe une partition de E en d ensembles C_0, C_1, \dots, C_{d-1} tels que :

- x appartient à C_0 ;
- soit $y \in C_r$ et $z \in C_s$, si $p^{(n)}(y, z) > 0$ alors $n = (s - r) \pmod{d}$;
- C_0, \dots, C_{d-1} sont des classes de récurrence irréductibles apériodiques pour la chaîne de Markov de matrice de transition P^d .

La décomposition est unique à une renumérotation près. La chaîne de Markov de matrice de transition P^d est irréductible, récurrente, apériodique. Si sa condition initiale est dans C_r pour $r \in \llbracket 0, d-1 \rrbracket$, alors tous ces éléments sont dans C_r .

La démonstration de ce théorème nécessite deux lemmes techniques.

LEMME 2.23.— Soit a_1, \dots, a_n des nombres entiers premiers dans leur ensemble, tout entier $m \geq \prod_x (1 + a_x)$ peut s'écrire sous la forme :

$$m = \sum_k x_k a_k \text{ avec } x_k \geq 0 \text{ pour tout } k. \quad (2.20)$$

Démonstration. Montrons par récurrence sur n que si a_1, \dots, a_n sont n entiers (non nécessairement premiers entre eux) et que $m \in \mathbf{N}$ s'écrit $m = \sum_k x_k a_k$ alors on peut toujours trouver une autre écriture satisfaisant les conditions de (2.20). Plus précisément, il existe une permutation σ de $\{1, \dots, N\}$ dans lui-même telle que :

$$x_{\sigma(i)} \leq \min_{l \neq \sigma(1), \dots, \sigma(i-1)} (a_l) \text{ pour tout } l \leq n-1.$$

Supposons d'abord que $n = 2$. Comme $m \geq 0$, l'un des deux coefficients x_1 ou x_2 est positif. A une renumérotation près, on peut toujours supposer qu'il s'agit de x_1 . Montrons que l'on peut toujours supposer que $x_1 < a_2$. Si ce n'est pas le cas, on écrit alors $x_1 = ka_2 + r$ avec $0 \leq r < a_2$:

$$\begin{aligned} m &= x_1 a_1 + x_2 a_2 + ka_1 a_2 - ka_1 a_2 \\ &= (x_1 - ka_2) a_1 + (x_2 + ka_1) a_2 = r a_1 + (x_2 + ka_1) a_2. \end{aligned}$$

En conclusion, tout entier m peut s'écrire sous la forme $m = x_1 a_1 + x_2 a_2$ avec $0 \leq x < b$. En particulier, si $m \geq a_1 a_2$, x_2 doit être positif.

Supposons le résultat démontré pour $(n-1)$. A une renumérotation près, on peut toujours supposer que x_1 est positif et appliquer l'hypothèse de récurrence à $m - x_1 a_1$

et aux $(n - 1)$ nombres restants. La renumérotation que l'on a appliquée au cours de cette opération définit la permutation σ .

Maintenant, si a_1, \dots, a_n sont premiers dans leur ensemble, le lemme de Bezout garantit l'existence de la représentation $m = \sum_k x_k a_k$ pour tout entier. D'après la première partie de la démonstration, on peut toujours supposer que $\sum_{k \leq n-1} x_k a_k$ est positive et inférieure à :

$$\sup_x \left(a_1 \dots a_n + \prod_{y \neq x} a_y + \dots \right) \leq a_1 \dots a_n + \sum_x \prod_{y \neq x} a_y + \dots = \prod (1 + a_x) - 1.$$

Par conséquent, pour m supérieur ou égal à $\prod (1 + a_x)$, il existe toujours une écriture de la forme (2.20). \square

LEMME 2.24.— Si x est apériodique alors il existe n_0 tel que si $n \geq n_0$ alors $p^{(n)}(x, x) > 0$.

Démonstration. Définissons l'ensemble :

$$I_x = \left\{ n \in \mathbb{N}, p^{(n)}(x, x) > 0 \right\}.$$

D'après la propriété de Markov, I_x est un semi-groupe : si m et n appartiennent à I_x alors $m + n$ aussi. En effet, on a :

$$p^{(m+n)}(x, x) \geq p^{(m)}(x, x)p^{(n)}(x, x).$$

On ordonne I_x par l'ordre naturel. Soit u_n le nombre de diviseurs communs des n premiers éléments de I_x . $(u_n, n \in \mathbb{N})$ est une suite décroissante positive donc convergente, et puisque x est apériodique, cette limite est 1. Comme u_n est à valeurs entières, il y a nécessairement un rang à partir duquel elle est constante, soit n_0 ce rang et soit a_1, \dots, a_{n_0} les n_0 premiers éléments de I_x . D'après le lemme précédent, pour n assez grand, $n \in I_x$. \square

Démonstration du théorème 2.22. Soit $K_y = \left\{ n, p^{(n)}(x, y) > 0 \right\}$. Pour k et l deux entiers, d'après la propriété de Markov, on a l'identité suivante :

$$\mathbf{P}_x(X_{k+l} = x) \geq \mathbf{P}_x(X_k = y)\mathbf{P}_y(X_l = x).$$

Par conséquent, n ne peut appartenir à K_y que si d divise $n + l$, c'est-à-dire si n s'écrit $\alpha d + r$ où $r \in \{0, \dots, d-1\}$ est le reste de la division de l par d . On définit C_r comme l'ensemble des points de E qui ont le même r . Ces ensembles forment clairement une partition et en prenant $l = 0$, on prouve que $x \in C_0$.

Soit m et n tels que $p^{(m)}(y, z) > 0$ et $p^{(n)}(x, y) > 0$. Comme $p^{(n+m)}(x, z) > 0$, il en découle d'après i) que $n + m \equiv s \pmod{d}$ et comme $n \equiv r \pmod{d}$, le résultat s'ensuit.

L'irréductibilité découle immédiatement du point précédent, l'apériodicité de la définition de la période. \square

Nous pouvons maintenant énoncer le résultat.

THÉORÈME 2.25.— Soit X une chaîne de Markov irréductible, récurrente positive, apériodique, de matrice de transition P et de probabilité invariante ν . Alors, on a :

$$\lim_{n \rightarrow \infty} p^{(n)}(x, y) = \nu(y), \text{ pour tout } x \text{ et tout } y.$$

La démonstration se fait par couplage : on va montrer que deux chaînes de Markov indépendantes de même matrice de transition mais de conditions initiales différentes finissent toujours par se rencontrer. Notons qu'à partir de cet instant de rencontre, elles coïncident en loi.

Démonstration. Sur $E \times E$, on définit la chaîne de Markov $Z_n = (W_n, y_n)$ de matrice de transition :

$$\hat{p}((x_1, x_2), (y_1, y_2)) = p(x_1, y_1)p(x_2, y_2).$$

En d'autres termes, les deux coordonnées W et Y évoluent indépendamment l'une de l'autre selon la loi de la chaîne de Markov originelle.

On va montrer premièrement que Z est une chaîne de Markov irréductible. Comme tous les états sont apériodiques, d'après le lemme 2.24, à partir d'un certain rang M :

$$p^{(l)}(y_1, y_1) > 0 \text{ et } p^{(l)}(x_2, x_2) > 0.$$

Comme X est irréductible et récurrent, il existe $K \geq M$ et $L \geq M$ tels que :

$$p^{(K)}(x_1, x_2) > 0 \text{ et } p^{(L)}(y_1, y_2) > 0.$$

Par conséquent, le chemin :

$$(x_1, y_1) \rightarrow (x_2, y_1) \rightarrow (x_2, y_2)$$

est de probabilité strictement positive pour l'indice $K + L + M$. En effet, d'après la propriété de Markov :

$$\begin{aligned} p^{(K+L)}((x_1, y_1), (x_2, y_2)) \\ \geq p^{(K)}(x_1, x_2)p^{(K)}(y_1, y_1) \cdot p^{(L)}(x_2, x_2)p^{(L)}(y_1, y_2) > 0. \end{aligned}$$

Il est clair que $\hat{\nu}(x, y) = \nu(x)\nu(y)$ définit une probabilité invariante pour la chaîne de Markov Z . Par conséquent d'après le théorème 2.18, tous les états sont récurrents positifs. Soit T le temps d'atteinte de la diagonale de $E \times E$ par Z :

$$\Delta = \{(x, y) \in E \times E, x = y\};$$

$$T = \inf \{n > 0, Z_n \in \Delta\}.$$

Comme Z est irréductible, récurrente, le temps d'atteinte d'un état (x, x) de la diagonale est presque sûrement fini. Comme T est le minimum de tous ces temps, il est aussi presque sûrement fini. Montrons que sur $\{T \leq n\}$, W_n et Y_n ont même loi :

$$\begin{aligned} \mathbf{P}(W_n = y, T \leq n) &= \sum_x \mathbf{E} [\mathbf{1}_{\{W_n=y\}} \mathbf{1}_{\{W_T=x\}} \mathbf{1}_{\{T \leq n\}}] \\ &= \sum_x \mathbf{E} [\mathbf{1}_{\{W_T=x\}} \mathbf{1}_{\{T \leq n\}} \mathbf{E} [\mathbf{1}_{\{W_n=y\}} | \mathcal{F}_T]] \\ &= \sum_x \mathbf{E} [\mathbf{1}_{\{W_T=x\}} \mathbf{1}_{\{T \leq n\}} \mathbf{E}_x [\mathbf{1}_{\{W_{n-T}=y\}}]] \\ &= \sum_x \mathbf{E} [\mathbf{1}_{\{Y_T=x\}} \mathbf{1}_{\{T \leq n\}} \mathbf{E}_x [\mathbf{1}_{\{Y_{n-T}=y\}}]] \\ &= \mathbf{P}(Y_n = y, T \leq n). \end{aligned}$$

D'après ce qui précède :

$$\begin{aligned} \mathbf{P}(W_n = y) &= \mathbf{P}(W_n = y, T \leq n) + \mathbf{P}(W_n = y, T > n) \\ &= \mathbf{P}(Y_n = y, T \leq n) + \mathbf{P}(W_n = y, T > n) \\ &\leq \mathbf{P}(Y_n = y) + \mathbf{P}(W_n = y, T > n). \end{aligned}$$

Symétriquement, on a :

$$\mathbf{P}(Y_n = y) \leq \mathbf{P}(W_n = y) + \mathbf{P}(Y_n = y, T > n),$$

dont on déduit que :

$$|\mathbf{P}(W_n = y) - \mathbf{P}(Y_n = y)| \leq \mathbf{P}(Y_n = y, T > n) + \mathbf{P}(W_n = y, T > n).$$

D'où en sommant sur toutes les valeurs possibles de y , on obtient :

$$\sum_y |\mathbf{P}(Y_n = y) - \mathbf{P}(W_n = y)| \leq 2\mathbf{P}(T > n).$$

Comme T est presque sûrement fini, le terme de droite tend vers 0 quand n croît indéfiniment. Si l'on prend $W_0 = x$ et Y ayant la loi ν , on en déduit :

$$\sum_y |p^{(n)}(x, y) - \nu(y)| \xrightarrow{n \rightarrow \infty} 0,$$

ce qui implique le résultat annoncé. □

REMARQUE.— On remarque que l'hypothèse d'apériodicité ne sert que dans la démonstration de l'irréductibilité de la chaîne de Markov Z . Pour se convaincre que cela est essentiel, regardons encore l'exemple du rat dans son labyrinthe. Formons la chaîne de Markov $Z_n = (X_n, Y_n)$ qui représente les positions de deux rats lâchés dans le même labyrinthe, qui évoluent indépendamment l'un de l'autre selon les mêmes règles que précédemment. Soit C_1 la classe cyclique de 1 et C_2 celle de 2 pour la chaîne de Markov X . Si Z part d'un état de $C_1 \times C_2$ alors Z évolue entre des états de cet ensemble et des états de $C_2 \times C_1$ mais n'atteint jamais d'états de $C_1 \times C_1$, par conséquent Z n'est pas irréductible.

Dans le cas périodique, on a toutefois le résultat suivant :

THÉORÈME 2.26.— Soit X une chaîne de Markov irréductible, récurrente positive, périodique de période d et de probabilité invariante ν . Soit $x \in E$ et C_0, \dots, C_{d-1} les classes cycliques associées à x . Si $y \in C_r$, on a :

$$\lim_{n \rightarrow \infty} p^{(nd+r)}(x, y) = d\nu(y).$$

✎ L'idée est d'appliquer le théorème précédent à la chaîne de matrice de transition P^d . Il faut pour cela déterminer la probabilité invariante de cette chaîne de Markov. On remarque d'après le théorème 2.13 qu'à un coefficient près la probabilité invariante d'un état y est égal à la proportion du nombre de visites à cet état entre deux visites à un état fixe x . Etant donné que dans la chaîne de Markov de matrice P^d on divise par d le nombre de pas, cette proportion est multiplié par d .

Démonstration. Par définition de la période et de C_k , C_k est un sous-ensemble fermé pour la chaîne X^k définie par $X_n^k = X_{nd+k}$ pour $k = 0, \dots, d-1$. Ces chaînes sont irréductibles et récurrentes positives. En vertu du corollaire 2.14, la probabilité invariante ν^k de X^k est proportionnelle à ν , c'est-à-dire qu'il existe α_k tel que $\nu^k(y) = \alpha_k \nu(y)$ pour tout $y \in C_k$. Par ailleurs, puisque ν est la probabilité invariante de X , pour tout k et tout l appartenant à $0, \dots, d-1$, on a :

$$\alpha_k = \mathbf{P}_\nu(X_{nd+k} \in C_k) = \mathbf{P}_\nu(X_{nd+k} \in C_k \cup C_l) = \mathbf{P}_\nu(X_{nd+l} \in C_l) = \alpha_l.$$

Il s'ensuit que $\alpha_k = d^{-1}$. Le dernier point découle du théorème 2.25. \square

Enfin, le dernier résultat utile pour les simulations est le théorème central limite suivant.

THÉORÈME 2.27.— Soit X une chaîne de Markov récurrente positive de probabilité invariante ν . Pour $f : E \times E \rightarrow \mathbf{R}$, on note :

$$Pf : \begin{cases} E & \longrightarrow \mathbf{R} \\ x & \longmapsto Pf(x) = \sum_y f(x, y)p(x, y) = \mathbf{E}_x[f(X_0, x_1)]. \end{cases}$$

Pour toute fonction f telle que $\mathbf{E}_\nu[P(f^2)] < \infty$ on a la convergence en loi suivante :

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n (f(X_{k-1}, X_k) - Pf(X_{k-1})) \xrightarrow[n \rightarrow \infty]{\text{Loi}} \mathcal{N}(0, \sigma^2(f)),$$

où $\mathcal{N}(0, \sigma^2(f))$ représente une loi gaussienne centrée de variance donnée par :

$$\sigma^2(f) = \mathbf{E}_\nu[P(f^2)] - \mathbf{E}_\nu[(Pf)^2].$$

Démonstration. Le lemme 2.12 implique que pour f bornée, la suite :

$$M_n^f = \sum_{j=0}^n f(X_j, X_{j+1}) - \sum_{j=0}^n Pf(X_j, \cdot)(X_j)$$

est une martingale. D'autre part, son processus croissant est défini par :

$$\begin{aligned} \Delta \langle M^f \rangle_n &= \mathbf{E}[(\Delta M_n^f)^2 | \mathcal{F}_n] \\ &= \mathbf{E}\left[\left(f(X_n, X_{n+1}) - Pf(X_n, \cdot)(X_n)\right)^2 | \mathcal{F}_n\right] \\ &= Pf^2(X_n) + Pf(X_n)^2 - 2Pf(X_n)^2 \\ &= \Gamma f(X_n), \end{aligned}$$

où $\Gamma f = Pf^2 - (Pf)^2$ est l'opérateur carré du champ associé à P . Par conséquent, on a :

$$\langle M^f \rangle_n = \sum_{j=0}^n \Gamma f(X_j).$$

Par hypothèse, Γf est intégrable par rapport à ν , le théorème 2.21 implique que :

$$\frac{\langle M^f \rangle_n}{n} \xrightarrow[n \rightarrow \infty]{} \sigma^2(f), \text{ } \mathbf{P}_x \text{ presque sûrement.}$$

Le résultat découle du théorème de la limite centrée pour les incréments de martingales. \square

Si l'on prend comme cas particulier, $f(X_{k-1}, X_k) = \mathbf{1}_{\{X_k = x\}}$, on obtient :

$$\mathbf{P}(\sqrt{n}(N_x^n - \pi(x)) \in [a, b]) \xrightarrow[n \rightarrow \infty]{} \int_a^b \exp(-x^2/2\sigma^2) \frac{dx}{\sigma\sqrt{2\pi}},$$

avec $\sigma^2 = \nu(y) - \sum_x p(x, y)^2 \nu(x)$.

EXEMPLE (Suite de l'exemple 2.1).— C'est le cas le plus simple dans lequel on n'a qu'à résoudre le système $\nu = \nu P$ et $\sum \pi(x) = 1$. Tous calculs faits, on trouve :

$$\nu = \left(\frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{3}{16}, \frac{3}{16}, \frac{1}{8}, \frac{1}{16} \right).$$

EXEMPLE (Suite de l'exemple 2.3).— Il faut restreindre la chaîne de Markov à une quelconque classe d'équivalence de la relation « communique ». Dans ce cas, il est clair que la probabilité invariante est la mesure uniforme sur ces états.

2.5. Calcul pratique de la probabilité invariante

Le principe est simple : la probabilité invariante est l'unique vecteur à composantes positives ou nulles, de poids total 1 qui satisfait l'équation $\nu(P - \text{Id}) = 0$. Si l'on veut résoudre un tel système par ordinateur, il faut prendre garde au fait que ce système est de co-rang 1 : il faut donc supprimer une colonne de P (par exemple, la dernière) et la remplacer par une colonne composée uniquement de 1. Soit \hat{P} la matrice ainsi obtenue. Il nous faut alors résoudre le système :

$$\pi(\hat{P} - \hat{I}) = b, \text{ avec } b = (0, \dots, 0, 1) \text{ et } \hat{I} = \begin{pmatrix} 1 & & & 0 \\ & 1 & (0) & 0 \\ & (0) & \ddots & \\ & & & 0 \end{pmatrix}.$$

Pratiquement, les chaînes que l'on utilise ont un espace d'états fini mais de cardinal très grand (plusieurs milliers d'états). Se pose alors le problème de la détermination de la probabilité invariante. Il s'agit *a priori* « simplement » de résoudre un système linéaire mais la taille de ce système oblige à utiliser des méthodes d'analyse numérique.

2.5.1. Méthode itérative

On a donc à résoudre l'équation $\pi = \pi P$ où P est la matrice de transition. D'après le théorème 2.25, si la chaîne est apériodique alors $\pi_{n+1} = \pi_n P$ tend vers la probabilité invariante. Pratiquement, on prend un π_0 quelconque et on itère. Ce procédé peut s'avérer coûteux si le calcul des coefficients de P est long. Néanmoins, la convergence est extrêmement rapide puisque géométrique de raison le module de la deuxième plus grande valeur propre de P .

Dans le cas où la chaîne est périodique (voir l'exemple du rat) de période d , il faut être plus précautionneux. Le théorème 2.25 nous indique que la suite π_n a d valeurs d'adhérence. Précisément, par définition même des classes cycliques, si π_0 est une masse de Dirac en x , les termes π_{kn} ont des composantes positives seulement pour les états de la classe cyclique de x , les termes π_{kn+j} ont des composantes positives seulement pour les états de la classe cyclique C_j , pour tout $j \in \{1, \dots, d-1\}$.

EXEMPLE (Suite de l'exemple 2.1).– Si la condition initiale est $\pi_0 = (1, 0, \dots)$ alors on a :

$$\begin{aligned} v(2) &= \left(0, \frac{1}{2}, 0, \frac{1}{2}, 0, 0, 0\right) \\ v(3) &= \left(\frac{1}{3}, 0, \frac{1}{6}, 0, \frac{1}{3}, 0, \frac{1}{6}\right) \\ v(4) &= \left(0, \frac{13}{36}, 0, \frac{4}{9}, 0, \frac{7}{36}, 0\right) \\ v(5) &= \left(\frac{29}{108}, 0, \frac{47}{216}, 0, \frac{79}{216}, 0, \frac{4}{27}\right) \\ v(6) &= \left(0, \frac{473}{1296}, 0, \frac{131}{324}, 0, \frac{299}{1296}, 0\right) \\ v(7) &= \left(\frac{997}{3888}, 0, \frac{1843}{7776}, 0, \frac{2891}{7776}, 0, \frac{131}{972}\right) \end{aligned}$$

Un moyen pour éviter ce désagrément est de considérer les sommes de Césaro, $\hat{\pi}_n = d^{-1} \sum_{i=n-d}^n \pi_i$. Cela exige de connaître la période, si c'est impossible on peut alors utiliser la méthode ergodique.

2.5.2. Méthode ergodique

Le théorème 2.21 affirme que pour une chaîne de Markov irréductible et récurrente positive de probabilité invariante ν , on a pour toute condition initiale et tout $x \in E$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{X_k=x\}} = \nu(x).$$

On simule donc une trajectoire de la chaîne de Markov le plus longtemps possible et on compte le pourcentage des fois où l'on passe dans l'état x . Le théorème 2.27 indique que la vitesse de convergence est en $1/\sqrt{n}$ ce qui se compare très désavantageusement avec les deux méthodes précédentes. En échange, on n'a pas à stocker tous les $\nu(x)$ mais seules les fréquences des valeurs qui nous intéressent. Il est en fait très fréquent que seules nous intéressent quelques composantes de ν .

EXEMPLE (File M/GI/1/K).– Il y a un buffer de taille K et l'on note X_n le nombre de clients dans le système juste après le départ du client n . Alors $(X_n, n \in \mathbf{N})$ vérifie la récurrence $X_{n+1} = \min((X_n - 1)^+ + A_{n+1}, K + 1)$. On a donc bien une chaîne de Markov irréductible et nécessairement récurrente puisque l'espace d'états est fini. On ne peut pas facilement calculer les probabilités invariantes par la méthode de la fonction

génératrice à cause des effets de bord. Cependant, pour le dimensionnement du buffer seule nous intéresse la probabilité de perte, c'est-à-dire $\nu(K+1)$ et éventuellement le nombre moyen de clients dans la file. Ces deux quantités s'obtiennent comme suit :

$$\nu(K+1) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{X_k = K+1\}} \text{ et } \hat{N} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k.$$

2.6. Problèmes

EXERCICE 1.— Sur un échiquier de 8 par 8, on place un cavalier dans le coin $A1$, on suppose que le cavalier se déplace au hasard (il choisit une direction au hasard parmi celles possibles à chaque coup) et sans mémoire. On rappelle qu'un cavalier se déplace de deux cases dans une direction (horizontale ou verticale) et d'une case dans l'autre direction. En utilisant la réversibilité et des considérations de symétrie, calculer le temps moyen de retour à la case $A1$.

Même question, si l'on identifie les bords opposés de l'échiquier ; le cavalier se déplace alors sur un tore !

EXERCICE 2.— Construire (dans les cas où c'est possible) une chaîne de Markov à deux états telle que :

- les deux états soient récurrents ;
- les deux états soient transients ;
- l'un des états soit transient, l'autre récurrent ;
- les deux soient transients ;
- les deux soient récurrents nuls.

EXERCICE 3.— On considère la chaîne de Markov à valeurs dans $\{1, 2, 3\}$ dont la matrice de transition est donnée par :

$$\begin{pmatrix} 0 & 1/2 & 1/2 \\ f(p) & 0 & 1-f(p) \\ 1-f(p) & 0 & f(p) \end{pmatrix}$$

où $p \in [0, 1]$ et $f(p)$ est définie par :

$$f(p) = \begin{cases} 0 & \text{si } p \leq 1/4 \\ 2p - 1/2 & \text{si } 1/4 < p \leq 3/4 \\ 1 & \text{si } p \geq 3/4. \end{cases}$$

- 1) Donner la classification des états en fonction des valeurs de p .

2) Pour quelles valeurs de p existe-il une probabilité invariante ? La calculer lorsqu'elle existe.

3) Partant de 2, quelle est, en fonction de p , la valeur du temps moyen de retour en 2 ?

4) Soit h la fonction définie par :

$$h(1) = -1, \quad h(2) = 1, \quad h(3) = 1.$$

Que vaut la limite :

$$\frac{1}{n} \sum_{j=1}^n h(X_j)$$

quand n tend vers $+\infty$ pour $p < 3/4$?

5) Si l'on dispose d'un nombre arbitrairement grand d'exemples de trajectoires, comment sait-on si p est supérieur à $3/4$? Comment sait-on si $p < 1/4$? Comment peut-on estimer p dans le cas où il est compris entre $1/4$ et $3/4$?

EXERCICE 4.— Soit X , une chaîne de Markov irréductible et récurrente sur E , et soit F un sous-ensemble de E . Montrer que la chaîne restreinte à F ($X_n^F, n \in \mathbf{N}$) (voir définition 2.8) est une chaîne de Markov sur E .

EXERCICE 5.— On considère la chaîne de Markov homogène X à deux états A et B de matrice de transition :

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

On cherche le temps de première apparition de la séquence ABA . Pour ce faire, on construit le processus $Y_n = (X_n, X_{n+1}, X_{n+2})$.

1) Montrer que Y est une chaîne de Markov homogène dont on donnera la matrice de transition (sous forme de matrice ou de graphe).

2) Cette chaîne est-elle irréductible ? apériodique ? récurrente positive ?

3) Calculer la probabilité invariante de Y . On pourra numérotiser les états dans l'ordre lexicographique : $AAA = 1, AAB = 2, \dots$

4) En déduire le temps moyen (à l'état invariante) qui s'écoule entre deux occurrences de ABA .

5) On suppose que $X_0 = A, X_1 = B$. Donner les équations qui permettent de calculer $\mathbf{E} [\tau_{ABA}^1]$. Il n'est pas demandé de les résoudre.

EXERCICE 6.— Soit un équipement qui émet, sur une ligne de transmission, des paquets de taille constante. On note T le temps de transmission d'un paquet. Par la suite on considère un modèle à temps discret du système, c'est-à-dire, un modèle pour lequel le temps est divisé en intervalles de longueur constante, que nous supposons égale à T . On appelle chaque intervalle un slot. La ligne de transmission peut introduire des erreurs et

on définit une suite $(Y_n, n \in \mathbf{N})$ tel que $Y_n = 1$ si, à l'instant $n + 1$, la ligne est dans un état pour lequel elle introduit des erreurs et $Y_n = 0$ si, à l'instant $n + 1$, elle est dans un état où elle n'introduit pas des erreurs. On suppose que $(Y_n, n \in \mathbf{N})$ est une chaîne de Markov invariante et que $P(Y_1 = 1 | Y_0 = 1) = 0,9$ et $P(Y_1 = 0 | Y_0 = 0) = 0,1$.

L'émission se fait avec le protocole « arrêt et attente » (*stop and wait*). Selon ce protocole, chaque paquet doit être acquitté. S'il n'y a pas d'erreur, le paquet est acquitté positivement et le paquet suivant peut être transmis. En cas contraire, le paquet doit être retransmis. Pour simplifier le problème nous considérons que l'acquittement arrive instantanément.

1) Calculer la distribution de probabilité invariante de Y_n .

2) On suppose que les paquets arrivent selon un processus géométrique. C'est-à-dire qu'au n^e slot il y a une arrivée avec probabilité q et aucune arrivée avec probabilité $1 - q$. Un paquet peut être transmis au même slot que celui où il arrive. Soit X_n le nombre de paquets dans le système, au slot n , après la transmission (s'il y a un paquet à transmettre). Le couple (X_n, Y_n) est une chaîne de Markov. On ordonne les états dans l'ordre lexicographique, c'est-à-dire :

0 1 2 3 4 5 6 7 8 9 ...
00 01 10 11 20 21 30 31 40 41 ...

3) Trouver Q la matrice de transition de (X_n, Y_n) .

4) Montrer que :

$$\begin{aligned}\nu_0 &= 1, \\ \nu_{2n} &= 9 \left(3\sqrt{q/1-q} \right)^{2n}, \\ \nu_{2n+1} &= 9\nu_{2n}\end{aligned}$$

est une mesure invariante pour la chaîne de Markov (X_n, y_n) .

5) Trouver l'ensemble de valeurs de q pour lesquelles tous les états sont récurrents positifs. Comparer le résultat obtenu avec le résultat de 1. Conclure.

EXERCICE 7.— On considère un paquet de N cartes. Pour le mélanger, on procède de la manière suivante : on choisit une carte au hasard parmi les N et on met cette carte sur le dessus du paquet sans bouger les ordres relatifs des autres.

1) Comment représenter l'état du paquet, noté X_n , après la n^e opération ?

2) En introduisant les permutations particulières :

$$\tau_k = \begin{pmatrix} 1 & 2 & \dots & k-1 & k & k+1 & \dots & N \\ 2 & 3 & \dots & k & 1 & k+1 & \dots & N \end{pmatrix},$$

pour $k \in \{1, \dots, N\}$; écrire les probabilités de transition de X .

L'écriture de τ_k signifie que $\tau_k(1) = 2, \tau_k(2) = 3, \dots$, ce que l'on peut aussi noter :

$$\tau_k = (2, 3, \dots, k, 1, k+1, \dots, n).$$

3) Montrer que cette chaîne est irréductible (on pourra avantageusement raisonner pour de petites valeurs de N comme $N = 4$ par exemple).

4) Montrer qu'au bout d'un nombre suffisamment grand d'opérations on obtient un « bon » mélange, caractérisé par l'équiprobabilité de tous les états possibles du paquet de cartes. (Cette question comporte deux parties : formaliser mathématiquement le problème puis le résoudre.)

EXERCICE 8.— On pose $E = \{1, \dots, 10\}$. On définit sur E , l'addition comme l'addition modulo 10, c'est-à-dire que $10 + 1 = 1$. On considère X , la chaîne de Markov de matrice de transition $P = (p_{i,j})$ donnée par :

$$p_{i,i+1} = p, \quad p_{i,i-1} = 1 - p.$$

On suppose que p n'est égal ni à 0, ni à 1.

- 1) Cette chaîne est-elle irréductible ? récurrente ? apériodique ?
- 2) Quelle est sa probabilité invariante ?

On considère maintenant X_1 et X_2 deux copies indépendantes de cette chaîne. On pose $Y = (X_1, X_2)$.

- 3) Cette chaîne est-elle irréductible ?
- 4) Quelle est sa probabilité invariante ?
- 5) On pose maintenant $Z_n = Y_{2n}$. Cette chaîne est-elle irréductible ? Quels sont ces sous-ensembles fermés ? Est-elle récurrente ? apériodique ?

EXERCICE 9.— Soit $A = \{A_n : n \geq 1\}$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs dans \mathbf{R}^k , soit h une application de $E \times \mathbf{R}^k$ dans E et soit X_0 une variable aléatoire indépendante de la suite A . On définit la suite $X = \{X_n : n \in \mathbf{N}\}$ par X_0 pour $n = 0$ et par $X_n = h(X_{n-1}, A_n)$, pour $n \geq 1$. Montrer que X est une chaîne de Markov.

2.7. Notes et commentaires

Le nombre d'ouvrages traitant des chaînes de Markov est incalculable, on ne saurait les citer tous. Parmi les plus récents et qui s'approchent le plus ou complètent notre approche, on peut noter [BAL 01, GRA 08]. Les chaînes de Markov constituent toujours un champ d'investigation très actif en raison de leur universalité. Les problèmes actuels se focalisent autour du calcul de la vitesse de convergence vers la probabilité stationnaire et son lien avec le « trou spectral », la réduction de l'espace d'états pour calculer plus facilement une approximation de la probabilité invariante, les applications en simulation et statistiques à travers les méthodes MCMC.

Résumé

- Une chaîne de Markov est définie par sa loi initiale ν et son opérateur de transition P .
- Un état récurrent est un état visité une infinité de fois. Un état transient est un état visité un nombre fini de fois.
- Deux états reliés entre eux ont même nature, c'est-à-dire transient ou récurrent.
- Une chaîne est irréductible si tous les états sont liés entre eux.
- Une mesure stationnaire s'identifie à un vecteur ligne solution de l'équation $\pi P = \pi$.
- Si l'on peut trouver π tel que $\sum_{x \in E} \pi(x) = 1$ alors π est une probabilité invariante, la chaîne est récurrente.
- Dans ce cas, quelle que soit la condition initiale, $\mathbf{P}(X_n = x) \xrightarrow{n \rightarrow \infty} \pi(x)$.
- Pour calculer π , on peut au choix résoudre le système $\pi P = \pi, \sum_{x \in E} \pi(x) = 1$ ou considérer la limite de la suite $\pi_{n+1} = \pi_n P, \pi_0$ quelconque.

DEUXIÈME PARTIE

Modélisation à temps continu

Chapitre 3

Processus de Poisson

La modélisation d'un système physique doit obéir à deux contraintes. D'une part, elle doit refléter aussi fidèlement que possible la « réalité » et d'autre part elle doit avoir un rôle prédictif ; autrement dit on doit pouvoir faire des calculs grâce à cette modélisation. Au-delà de la difficulté à déterminer quantitativement et qualitativement les paramètres pertinents d'un système physique, l'expérience prouve que plus l'on voudra un modèle précis moins celui-ci sera utilisable pour les calculs.

Dans le cadre des files d'attente, il nous faut modéliser en premier lieu le processus des arrivées des requêtes. Le processus de Poisson, que nous étudions dans cette section, est le modèle le plus fréquemment utilisé essentiellement parce que c'est l'un des rares avec lequel on peut faire des calculs. Cette modélisation se trouve être très pertinente pour les arrivées des appels téléphoniques à un commutateur, malheureusement il n'en est pas de même pour les autres types de réseaux dans lesquels le trafic est beaucoup plus versatile. Néanmoins, comme nous le verrons à la fin de ce chapitre, le processus de Poisson peut être modifié de façon à refléter dans une certaine mesure cette versatilité.

La définition d'un processus ponctuel et les notations associées sont définies dans A.5.2. Rappelons qu'un processus ponctuel intégrable est une suite strictement croissante de variables aléatoires positives (T_1, T_2, \dots) telle que $T_n \rightarrow \infty$ presque sûrement. Par convention, on adjoint à cette suite la variable aléatoire $T_0 = 0$ presque sûrement. Ces variables aléatoires vont représenter les instants d'arrivées des requêtes au système. On peut tout aussi bien décrire la suite par l'écart de temps qui s'écoule entre deux arrivées successives : $\xi_n = T_{n+1} - T_n$ est la n^{e} interarrivée. La donnée de la suite $(\xi_n, n \in \mathbb{N})$ définit aussi un processus ponctuel par la relation $T_n = \sum_{i \leq n-1} \xi_i$. On notera finalement $N(t)$, le nombre de points (d'arrivées), entre les instants 0 et t .

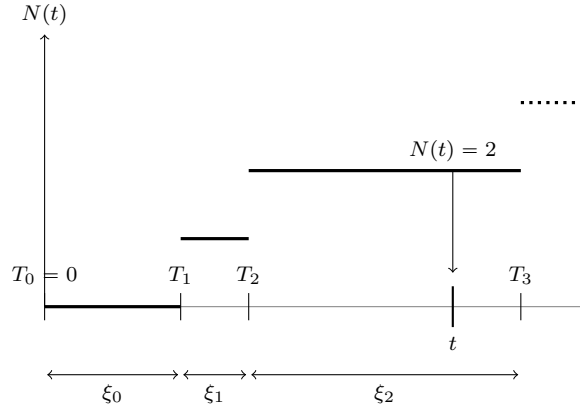


Figure 3.1 – Notations relatives aux processus ponctuels

3.1. Définitions

Le processus de Poisson admet plusieurs caractérisations. Comme n'importe laquelle peut être considérée comme une définition et les autres comme des propriétés, nous donnons à toutes le statut de définition et nous démontrons qu'elles sont équivalentes.

DÉFINITION 3.1.– Le processus ponctuel N est un processus de Poisson d'intensité λ si et seulement si les variables aléatoires $(\xi_n, n \in \mathbf{N})$ sont indépendantes et de même loi exponentielle de paramètre λ .

DÉFINITION 3.2.– Le processus ponctuel N est un processus de Poisson d'intensité λ si et seulement si les deux conditions suivantes sont vérifiées :

- pour tout $t \geq 0$, la variable aléatoire $N(t)$ suit une loi de Poisson de paramètre λt ;
- conditionnellement à $\{N(t) = n\}$, la famille (T_1, \dots, T_n) est uniformément distribuée sur $[0, t]$.

DÉFINITION 3.3.– Le processus ponctuel N est un processus de Poisson d'intensité λ si et seulement si les deux conditions suivantes sont vérifiées :

- pour tout $0 = t_0 < t_1 < \dots < t_n$, les variables aléatoires $(N(t_i + 1) - N(t_i), 1 \leq i \leq n - 1)$ sont indépendantes ;
- pour tout t, s , la variable aléatoire $N(t + s) - N(t)$ suit une loi de Poisson de paramètre λs :

$$\mathbf{P}(N(t + s) - N(t) = k) = \exp(-\lambda s) \frac{(\lambda s)^k}{k!}.$$

DÉFINITION 3.4.— Le processus ponctuel N est un processus de Poisson d'intensité λ si et seulement si pour toute fonction $f : \mathbf{R}^+ \rightarrow \mathbf{R}^+$ (ou pour toute fonction f à support compact dans \mathbf{R}^+), on a l'identité suivante :

$$\mathbf{E} \left[\exp \left(- \sum_{n \geq 1} f(T_n) \right) \right] = \exp \left(- \int_0^\infty (1 - e^{-f(s)}) \lambda \, ds \right).$$

DÉFINITION 3.5.— Le processus ponctuel N est un processus de Poisson d'intensité λ si et seulement si le processus $(N(t) - \lambda t, t \geq 0)$ est une martingale par rapport à la filtration \mathcal{F} engendrée par N , c'est-à-dire $\mathcal{F}_t = \sigma\{N(s), s \leq t\}$.

Afin de montrer l'équivalence entre ces définitions, nous devons introduire trois résultats techniques.

LEMME 3.1.— La densité de la loi de (T_1, \dots, T_n) est donnée par :

$$d\mathbf{P}_{(T_1, \dots, T_n)}(x_1, \dots, x_n) = \lambda^n \exp(-\lambda x_n) \mathbf{1}_{\mathcal{C}}(x_1, \dots, x_n) \, dx_1 \dots dx_n, \quad (3.1)$$

où :

$$\mathcal{C} = \{(y_1, \dots, y_n) \in (\mathbf{R}^+)^n, 0 \leq y_1 \leq \dots \leq y_n\}.$$

En particulier, T_n suit une loi gamma de paramètres n et λ , définie par :

$$d\mathbf{P}_{T_n}(x) = \lambda^n \exp(-\lambda x) \frac{x^{n-1}}{(n-1)!} \mathbf{1}_{\mathbf{R}^+}(x) \, dx \quad (3.2)$$

Démonstration. On procède par identification :

$$\begin{aligned} & \mathbf{E}[f(T_1, \dots, T_n)] \\ &= \int \dots \int_{(\mathbf{R}^+)^n} f(x_0, x_0 + x_1, \dots, x_0 + \dots + x_{n-1}) \, d\mathbf{P}_{\xi_0}(x_0) \dots d\mathbf{P}_{\xi_{n-1}}(x_{n-1}). \end{aligned}$$

On considère le changement de variables :

$$u_1 = x_0, \, u_2 = x_0 + x_1, \, \dots, \, u_n = x_0 + \dots + x_{n-1},$$

dont le jacobien vaut clairement 1. Les conditions $x_0 \geq 0, \dots, x_{n-1} \geq 0$ reviennent à $0 \leq u_1 \leq u_2 \leq \dots \leq u_n$. On a donc :

$$\mathbf{E}[f(T_1, \dots, T_n)] = \int \dots \int_{(\mathbf{R}^+)^n} f(u_n) \lambda^n e^{-\lambda u_n} \mathbf{1}_{\mathcal{C}}(u_1, \dots, u_n) \, du_1 \dots du_n.$$

La densité de la loi jointe s'en déduit. Si f ne dépend que de T_n , on obtient :

$$\begin{aligned}\mathbf{E}[f(T_n)] &= \int \dots \int_{0 \leq u_1 \dots \leq u_n} f(u_n) \lambda^n \exp(-\lambda u_n) \, du_1 \dots du_n \\ &= \int_0^\infty f(u_n) \lambda^n \exp(-\lambda u_n) \left(\int_0^{u_n} du_{n-1} \int \dots \int_0^{u_2} du_1 \right) du_n \\ &= \int_0^\infty f(u_n) \lambda^n \exp(-\lambda u_n) \frac{u_n^{n-1}}{(n-1)!} du_n.\end{aligned}$$

Le résultat en découle. \square

LEMME 3.2.— Soit X une variable aléatoire de loi de Poisson de paramètre λ . On a :

$$\mathbf{E}[e^{-sX}] = \exp(-\lambda(1 - e^{-s})).$$

Démonstration. Par définition de la loi de Poisson, on a l'identité suivante :

$$\mathbf{E}[e^{-sX}] = \sum_{k=0}^{\infty} \mathbf{E}[e^{-sk}] e^{-\lambda} \frac{\lambda^k}{k!} = \exp(-\lambda + \lambda e^{-s}),$$

d'où le résultat. \square

LEMME 3.3.— Soit (U_1, \dots, U_n) des variables aléatoires indépendantes de loi uniforme sur $[0, t]$. Soit \bar{U} , le réordonnancement de ce n -uplet par ordre croissant, c'est-à-dire :

$$\bar{U}_1(\omega) \leq \bar{U}_2(\omega) \leq \dots \leq \bar{U}_n(\omega), \text{ presque sûrement.}$$

La loi de \bar{U} est donnée par :

$$d\mathbf{P}_{(\bar{U}_1, \dots, \bar{U}_n)}(x_1, \dots, x_n) = \frac{n!}{t^n} \mathbf{1}_{\mathcal{C}}(x_1, \dots, x_n) \, dx_1 \dots dx_n$$

Démonstration. On appelle σ la variable aléatoire à valeurs dans \mathfrak{S}_n , qui représente le réarrangement à faire pour mettre dans l'ordre les différentes valeurs des $U_i(\omega)$, c'est-à-dire que $\sigma(\omega)$ est défini de la manière suivante : si l'on a :

$$U_2(\omega) \leq U_3(\omega) \leq U_1(\omega),$$

alors :

$$\sigma(\omega) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

L'image de i par $\sigma(\omega)$ est l'indice de la variable aléatoire qui est en i^{e} position pour cet ω . On a donc par définition $\bar{U}_i(\omega) = U_{\sigma(\omega)(i)}(\omega)$. Comme les variables aléatoires $U_i, i \in \llbracket 1, n \rrbracket$ sont indépendantes et de même loi, pour tout $\tau \in \mathfrak{S}_n$, on a :

$$d\mathbf{P}_{(U_{\tau(1)}, \dots, U_{\tau(n)})}(u_1, \dots, u_n) = \otimes_{i=1}^n \frac{1}{t} \mathbf{1}_{[0,t]}(u_i) \, du_i.$$

On remarque en particulier que cette loi ne dépend pas de τ . Par conséquent, il vient :

$$\begin{aligned}\mathbf{P}(\sigma = \tau) &= \mathbf{P}(U_{\tau(1)} \leq \dots \leq U_{\tau(n)}) \\ &= \int \dots \int \mathbf{1}_{\mathcal{C}}(u_1, \dots, u_n) \, d\mathbf{P}_{(U_{\tau(1)}, \dots, U_{\tau(n)})}(u_1, \dots, u_n) \\ &= \mathbf{P}(\sigma = \text{Id}).\end{aligned}$$

Ainsi, σ suit une loi uniforme sur \mathfrak{S}_n , c'est-à-dire que $\mathbf{P}(\sigma = \tau) = 1/n!$. Pour calculer la loi du n -uplet \bar{U} on décompose l'espace d'état en la partition $\cup_{\tau \in \mathfrak{S}_n} (\sigma = \tau)$. Pour toute fonction continue bornée, on a l'identité suivante :

$$\begin{aligned}\mathbf{E}[f(\bar{U}_1, \dots, \bar{U}_n)] &= \sum_{\tau \in \mathfrak{S}_n} \mathbf{E}[f(\bar{U}_1, \dots, \bar{U}_n); \sigma = \tau] \\ &= \sum_{\tau \in \mathfrak{S}_n} \mathbf{E}[f(U_{\tau(1)}, \dots, U_{\tau(n)}) \mathbf{1}_{\mathcal{C}}(U_{\tau(1)}, \dots, U_{\tau(n)})] \\ &= \sum_{\tau \in \mathfrak{S}_n} \int \dots \int f(u_1, \dots, u_n) \mathbf{1}_{\mathcal{C}}(u_1, \dots, u_n) \, d\mathbf{P}_{(U_{\tau(1)}, \dots, U_{\tau(n)})}(u_1, \dots, u_n) \\ &= \sum_{\tau \in \mathfrak{S}_n} \int \dots \int f(u_1, \dots, u_n) \mathbf{1}_{\mathcal{C}}(u_1, \dots, u_n) \otimes_{i=1}^n \frac{1}{t} \mathbf{1}_{[0,t]}(u_i) \, du_i \\ &= \frac{n!}{t^n} \int \dots \int f(u_1, \dots, u_n) \mathbf{1}_{\mathcal{C}}(u_1, \dots, u_n) \, du_1 \dots du_n,\end{aligned}$$

d'où le résultat. □

Démonstration des équivalences entre les définitions. Nous allons montrer la chaîne d'implications : $3.1 \implies 3.2 \implies 3.3 \implies 3.4 \implies 3.5 \implies 3.1$.

$3.1 \implies 3.2$.

Montrons d'abord que $N(t)$ suit une loi de Poisson. Comme il est évident que les événements $\{N(t) = k\}$ et $\{T_k \leq t < T_{k+1}\}$ coïncident, on a :

$$\begin{aligned}\mathbf{P}(N(t) = k) &= \mathbf{P}(T_k \leq t < T_{k+1}) \\ &= \iint \mathbf{1}_{\{x \leq t\}} \mathbf{1}_{\{x+y > t\}} \, d\mathbf{P}_{T_k}(x) \, d\mathbf{P}_{\xi_k}(y) \\ &= \int_0^t \left(\int_{t-x}^{\infty} \lambda e^{-\lambda y} \, dy \right) \lambda^k \frac{x^{k-1}}{(k-1)!} \exp(-\lambda x) \, dx \\ &= e^{-\lambda t} \int_0^t \lambda^k \frac{x^{k-1}}{(k-1)!} \, dx = e^{-\lambda t} \frac{(\lambda x)^k}{k!}.\end{aligned}$$

Pour la loi conditionnelle, on procède de façon analogue :

$$\begin{aligned}
& \mathbf{P}((T_1, \dots, T_n) \in A \mid N(t) = n) \mathbf{P}(N(t) = n) \\
&= \mathbf{P}((T_1, \dots, T_n) \in A, T_n \leq t < T_{n+1}) \\
&= \int \dots \int_{0 \leq u_1 \leq \dots \leq u_{n+1}} \mathbf{1}_A(u_1, \dots, u_n) \mathbf{1}_{[u_n, u_{n+1}[}(t) \lambda^{n+1} e^{-\lambda u_{n+1}} \, du_1 \dots du_{n+1} \\
&= \lambda^n \int \dots \int_{0 \leq u_1 \leq \dots \leq u_{n+1}} \mathbf{1}_A(u_1, \dots, u_n) \mathbf{1}_{[u_n, \infty[}(t) \left(\int_t^\infty \lambda e^{-\lambda u_{n+1}} \, du_{n+1} \right) \, du_1 \dots du_n \\
&= \lambda^n e^{-\lambda t} \int \dots \int_{0 \leq u_1 \leq \dots \leq u_{n+1}} \mathbf{1}_A(u_1, \dots, u_n) \mathbf{1}_{[u_n, \infty[}(t) \, du_1 \dots du_n.
\end{aligned}$$

En divisant le terme de droite par $e^{-\lambda t}(\lambda t)^n/n!$, on obtient :

$$\begin{aligned}
& \mathbf{P}((T_1, \dots, T_n) \in A \mid N(t) = n) \\
&= \frac{n!}{t^n} \int \dots \int_{0 \leq u_1 \leq \dots \leq u_{n+1}} \mathbf{1}_A(u_1, \dots, u_n) \mathbf{1}_{[u_n, \infty[}(t) \, du_1 \dots du_n,
\end{aligned}$$

ce qui en vertu du lemme 3.3 signifie que (T_1, \dots, T_n) a conditionnellement à $\{N(t) = n\}$ la même loi que $(\tilde{U}_1, \dots, \tilde{U}_n)$ définie dans ce lemme. On résume ceci en disant que, conditionnellement à $\{N(t) = n\}$, le n -uplet (T_1, \dots, T_n) est uniformément réparti sur $[0, t]$.

3.2 \implies 3.3.

Soit $t_0 = 0 < t_1 < \dots < t_n$ des réels et i_0, \dots, i_{n-1} des entiers. On veut prouver que :

$$\mathbf{P} \left(\bigcap_{l=0}^{n-1} \{N(t_{l+1}) - N(t_l) = i_l\} \right) = \prod_{l=1}^n \mathbf{P}(N(t_{l+1}) - N(t_l) = i_l).$$

On peut toujours écrire :

$$\begin{aligned}
& \mathbf{P} \left(\bigcap_{l=0}^{n-1} \{N(t_{l+1}) - N(t_l) = i_l\} \right) \\
&= \sum_{k \in \mathbf{N}} \mathbf{P} \left(\bigcap_{l=0}^{n-1} \{N(t_{l+1}) - N(t_l) = i_l\} \mid N(t_n) = k \right) \mathbf{P}(N(t_n) = k).
\end{aligned}$$

La seule valeur de k pour laquelle les probabilités conditionnelles de cette somme sont non nulles est $k_0 = \sum_l i_l$. Pour calculer la probabilité conditionnelle correspondante on sait que les impacts entre 0 et t_n sont uniformément répartis. Cette quantité égale donc la probabilité que k points uniformément répartis sur un intervalle se répartissent en i_1 points dans l'intervalle $[0, t_1]$, i_2 points dans l'intervalle $]t_1, t_2]$, etc. Chaque point se trouve dans un intervalle de longueur x avec probabilité x/t_n , par conséquent on a :

$$\begin{aligned} \mathbf{P} \left(\bigcap_{l=0}^{n-1} \{N(t_{l+1}) - N(t_l) = i_l\} \mid N(t_n) = k_0 \right) \\ = \frac{k_0!}{i_1! \dots i_n!} \prod_{l=0}^{n-1} \left(\frac{t_{l+1} - t_l}{t_n} \right)^{i_l}. \end{aligned}$$

Comme $N(t_n)$ suit une loi de Poisson de paramètre λt_n et $k_0 = \sum_{l=0}^{n-1} i_l$, on en déduit que :

$$\begin{aligned} \mathbf{P} \left(\bigcap_{l=0}^{n-1} \{N(t_{l+1}) - N(t_l) = i_l\} \right) \\ = e^{-\lambda t_n} \frac{(\lambda t_n)^{k_0}}{k_0!} \frac{k_0!}{i_1! \dots i_n!} \prod_{l=0}^{n-1} \left(\frac{t_{l+1} - t_l}{t_n} \right)^{i_l} \\ = \prod_{l=0}^{n-1} e^{-\lambda(t_{l+1} - t_l)} \frac{(\lambda(t_{l+1} - t_l))^{i_l}}{i_l!}. \end{aligned}$$

On a pu mettre la probabilité de l'intersection comme un produit de probabilité, par conséquent les variables aléatoires sont indépendantes. En prenant $n = 2$, $t_1 = t$, $t_2 = t + s$, $i_0 = i$ et $i_1 = j$, on obtient :

$$\mathbf{P}(N(t) = i, N(t + s) - N(t) = j) = e^{-\lambda t} \frac{(\lambda t)^i}{i!} e^{-\lambda s} \frac{(\lambda s)^j}{j!}.$$

En sommant sur toutes les valeurs de i , on obtient le résultat voulu.

3.3 \implies 3.4.

Remarquons que si l'on prend $f(s) = \mathbf{1}_{[a,b]}(s)$, on a :

$$\sum_n f(T_n) = N(b) - N(a).$$

Du lemme 3.2, on déduit que le résultat est vrai pour les fonctions indicatrices et par linéarité pour les fonctions étagées (c'est-à-dire les combinaisons linéaires de fonctions

indicatrices). Par convergence monotone, on en déduit que le résultat est vrai pour toute fonction mesurable positive.

3.4 \implies 3.5.

Il suffit d'écrire $N(t+s) = (N(t+s) - N(t)) + N(t)$ et d'utiliser l'indépendance de ces variables aléatoires pour prouver que :

$$\mathbf{E}[N(t+s) | \mathcal{F}_t] = \mathbf{E}[N(t+s) - N(t)] + N(t) = \lambda s + N(t),$$

d'où le résultat.

3.5 \implies 3.1

Comme N n'évolue pas entre ses instants de sauts, pour f bornée, on a :

$$\begin{aligned} f(N(t)) - f(N(s)) &= \sum_{s < r \leq t, \Delta N(r)=1} f(N(r)) - f(N(r_-)) \\ &= \int_s^t (f(N(r_-) + 1) - f(N(r_-))) dN(r) \\ &= \int_s^t (f(N(r_-) + 1) - f(N(r_-))) (dN(r) - \lambda dr) \\ &\quad + \int_s^t (f(N(r_-) + 1) - f(N(r_-))) \lambda dr. \end{aligned}$$

Comme le processus $(r \mapsto f(N(r_-)))$ est prévisible, le théorème A.34 implique que l'intégrale stochastique définit bien une martingale donc on a :

$$t \longmapsto f(N(t)) - \int_0^t (f(N(r_-) + 1) - f(N(r_-))) \lambda dr$$

est une martingale. D'après le théorème 4.15, N est un processus de Markov de générateur infinitésimal :

$$Af(x) = \lambda(f(x+1) - f(x)) \text{ pour } f : \mathbf{N} \rightarrow \mathbf{R} \text{ bornée.}$$

D'après le théorème 4.9, la définition 3.1 est vérifiée. \square

3.2. Propriétés

Il ne faut pas se laisser abuser par la définition 3.2 qui stipule que, conditionnellement au nombre de points sur un intervalle, les impacts sont uniformément répartis sur cet intervalle. Certes, lorsque l'on observe une trajectoire, on connaît t et le nombre

d'impacts dans cet intervalle, on devrait donc observer des points uniformément répartis. Pourtant, on observe une répartition similaire à celle de la figure 3.2. C'est le phénomène de « clusterisation » : sur une trajectoire, les arrivées donnent l'impression d'être groupées. C'est d'ailleurs ce que l'on constate dans les magasins où après une période d'inactivité, plusieurs clients arrivent presque en même temps.



Figure 3.2 – Quatre trajectoires d'un processus de Poisson

✎ En fait, la notion même de répartition uniforme est floue, il ne faut pas la confondre avec la constance de l'écart entre les arrivées. Comme le montre la figure 3.2, l'uniformité dans la répartition se « voit » sur plusieurs trajectoires : il n'y a ici presque plus de parties de $[0, 1]$ qui ne comporte pas d'impact dans l'une ou l'autre des trajectoires. Il y a une différence essentielle entre le phénomène de « clusterisation » apparente du processus de Poisson et le régime de « rafales » (ou *burst*) que l'on observe dans les grands réseaux haut débit : dans le premier cas, le débit moyen instantané (calculé en moyennant sur un grand nombre de trajectoires) ne dépend pas du temps (puisque'il est égal à λ) tandis que dans l'autre cas, il va grandement varier avec le temps (penser à de la vidéo à débit variable). On ne peut donc pas représenter de tels trafics par un processus de Poisson.

THÉORÈME 3.4. – Soit N un processus de Poisson d'intensité λ . Pour toute fonction f à support compact, on a :

$$\mathbf{E} \left[\exp \left(- \sum_{n \geq 1} f(T_n) \right) \right] = \exp \left(- \int_{\mathbf{R}^+} (1 - e^{-f(s)}) \lambda \, ds \right) \quad (3.3)$$

Démonstration. Si l'on prend $f(s) = \alpha \mathbf{1}_{[a,b]}(s)$, on a :

$$\sum_n f(T_n) = \alpha(N(b) - N(a)).$$

On sait que $N(b) - N(a)$ suit une loi de Poisson de paramètre $\lambda(b-a)$. Par conséquent, on a :

$$\begin{aligned} \mathbf{E} \left[\exp \left(- \sum_{n \geq 1} f(T_n) \right) \right] &= \sum_{n=0}^{\infty} e^{-\alpha n} e^{-\lambda(b-a)} \frac{(\lambda(b-a))^n}{n!} \\ &= \exp(-\lambda(b-a) + \lambda(b-a)e^{-\alpha}) \\ &= \exp \left(- \int \left(1 - e^{-f(s)} \right) \lambda \, ds \right). \end{aligned}$$

Par indépendance des incréments, l'équation (3.3) est donc vraie pour les fonctions étagées. Par convergence dominée, elle l'est aussi pour les fonctions à support compact. \square

3.2.1. Superposition, amincissement

Lorsque l'on dispose de deux processus ponctuels N^1 et N^2 , la superposition de ces processus est le processus ponctuel, noté $N = N^1 + N^2$, dont les impacts sont ceux de N^1 et ceux de N^2 .

En vertu du théorème 3.4, le résultat suivant est immédiat.

THÉORÈME 3.5.— La superposition de deux processus de Poisson indépendants est un processus de Poisson d'intensité somme.

REMARQUE.— En particulier, lors de la superposition de deux processus de Poisson indépendants, deux clients ne peuvent arriver simultanément. Ce résultat reste vrai pour n'importe quels processus ponctuels indépendants.

Supposons maintenant qu'un processus de Poisson N d'intensité λ soit éclaté en deux processus N^1 et N^2 selon un tirage de Bernoulli de paramètre p , c'est-à-dire qu'à chaque point de N , on décide qu'il appartient à N^1 avec probabilité p et à N^2 avec probabilité $1 - p$, ce tirage au sort étant indépendant de tout le reste et en particulier des précédents tirages au sort (voir la figure 3.3 dans laquelle le chiffre au-dessus ou au-dessous de chaque point représente le numéro du flux auquel ce point est attribué). On dit que le processus de Poisson est aminci.



Figure 3.3 – Amincissement d'un processus de Poisson

Le théorème 3.4 assure qu'un processus de Poisson sur \mathbf{R}^+ est un cas particulier de processus de Poisson spatial (voir section ??). Par conséquent, le théorème ?? assure le résultat qui suit.

THÉORÈME 3.6.— Les processus N^1 et N^2 résultants de l'amincissement de N , sont deux processus de Poisson indépendants d'intensités respectives λp et $\lambda(1 - p)$.

3.2.2. Paradoxe de l'autobus

Il est en revanche un résultat spécifique de la dimension 1, connu sous le nom de paradoxe de l'autobus (ou paradoxe de l'inspection). Interprétons les points d'un processus de Poisson comme les instants d'arrivée des bus à un arrêt donné. Soit :

$$\begin{cases} W(t) &= T_{N(t)+1} - t, \\ Z(t) &= t - T_{N(t)}, \\ \xi(t) &= W(t) + Z(t) = T_{N(t)+1} - T_{N(t)}, \end{cases}$$

respectivement le temps d'attente du bus si l'on arrive en t à l'arrêt, le temps écoulé depuis le passage du dernier bus et l'intervalle de temps entre le bus que l'on prend et celui que l'on a raté.

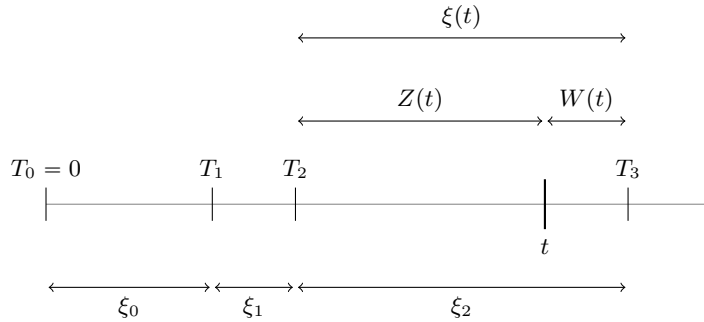
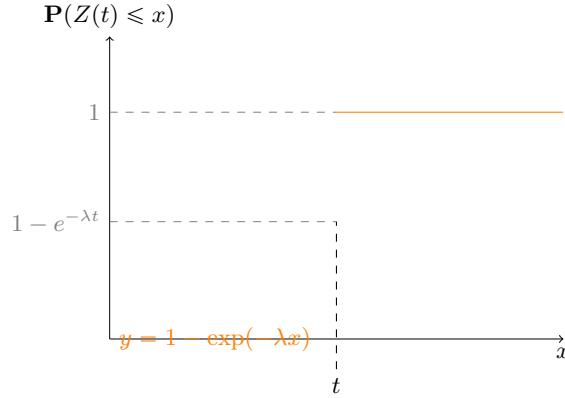


Figure 3.4 – Notations

THÉORÈME 3.7 (Paradoxe de l'autobus).— $W(t)$ suit une loi exponentielle de paramètre λ et est indépendante de $Z(t)$, dont la loi est donnée par :

$$\mathbf{P}(Z(t) \leq x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } 0 \leq x < t; \\ 1 & \text{si } x \geq t. \end{cases}$$

Figure 3.5 – Fonction de répartition de $Z(t)$

⚠ Ceci peut sembler paradoxal puisque la valeur moyenne de W_t , c'est-à-dire le temps moyen d'attente, est donc $1/\lambda$ alors que le temps moyen qui s'écoule entre les passages de deux bus vaut aussi $1/\lambda$. Cette propriété est une manifestation de ce qu'il est convenu d'appeler l'absence de mémoire de la loi exponentielle, sur laquelle nous reviendrons dans le chapitre suivant. En fait, tout se passe comme si, au moment où l'on arrive à l'arrêt de bus, le compteur qui égrène le temps qui s'écoule entre deux bus était remis à zéro et que l'on recomptait un temps de loi exponentielle jusqu'à la prochaine arrivée.

A proprement parler, cette approche est mathématiquement fautive car si ξ_n suit bien une loi exponentielle de paramètre λ , $\xi(t)$ n'a pas la loi de ξ_n puisqu'en fonction de ω le numéro du bus que l'on vient de rater change. Conditionnellement à $\{N(t) = n\}$, $\xi(t)$ a effectivement une loi exponentielle de paramètre λ mais contrairement à ce que l'on pourrait croire, lorsque l'on déconditionne, la loi de $\xi(t)$ n'est plus une loi exponentielle.

Démonstration du théorème 3.7. Notons que $Z(t)$, par définition même, ne peut pas être plus grand que t . Pour $0 \leq x < t$ et $y \geq 0$, en utilisant au passage l'indépendance de T_n et ξ_n pour tout $n \in \mathbf{N}$ on a :

$$\begin{aligned} \mathbf{P}(Z(t) \leq x, W(t) > y) &= \sum_{n \in \mathbf{N}} \mathbf{P}(Z(t) \leq x, W(t) > y, N(t) = n) \\ &= \sum_{n \in \mathbf{N}} \mathbf{P}(t - T_n \leq x, T_n + \xi_n - t > y, T_n \leq t < T_n + \xi_n). \end{aligned}$$

Pour $n = 0$, cette dernière probabilité vaut :

$$\mathbf{P}(t \leq x, \xi_0 > t + y) = 0$$

puisque l'on a supposé que $x < t$. La somme précédente s'écrit donc :

$$\begin{aligned}
& \sum_{n \geq 1} \iint \mathbf{1}_{\{t-u \leq x\}} \mathbf{1}_{\{u+v-t > y\}} \mathbf{1}_{\{u \leq t\}} d\mathbf{P}_{T_n}(u) d\mathbf{P}_{\xi_{n+1}}(v) \\
&= \sum_{n \geq 1} \int_{t-x}^t \lambda^n e^{-\lambda u} \frac{u^{n-1}}{(n-1)!} \left(\int_{t+y-u}^{\infty} \lambda e^{\lambda v} dv \right) du \\
&= \sum_{n \geq 1} \lambda^n e^{-\lambda(t+y)} \int_{t-x}^t \frac{u^{n-1}}{(n-1)!} du \\
&= \sum_{n \geq 1} \lambda^n e^{-\lambda(t+y)} \left[\frac{t^n}{n!} - \frac{(t-x)^n}{n!} \right] \\
&= e^{-\lambda(t+y)} \left(\sum_{n \geq 1} \frac{(\lambda t)^n}{n!} - \sum_{n \geq 1} \frac{(\lambda(t-x))^n}{n!} \right) \\
&= e^{-\lambda(t+y)} \left(e^{\lambda t} - 1 - (e^{\lambda(t-x)} - 1) \right) \\
&= e^{-\lambda y} (1 - e^{-\lambda x}).
\end{aligned}$$

Comme $\lim_{x \nearrow t} (1 - e^{-\lambda x}) = 1 - e^{-\lambda t} < 1$, on en déduit qu'il y a un saut dans la fonction de répartition de $Z(t)$ et par conséquent que $\mathbf{P}(Z(t) = t) = e^{-\lambda t} > 0$. \square

REMARQUE.— Le calcul de l'espérance de $\xi(t)$ se fait de la manière suivante :

$$\begin{aligned}
\mathbf{E}[\xi(t)] &= \mathbf{E}[W(t)] + \mathbf{E}[Z(t)] \\
&= \frac{1}{\lambda} + t \cdot e^{-\lambda t} + \int_0^t \lambda s e^{-\lambda s} ds \\
&= \frac{1}{\lambda} (2 - e^{-\lambda t}).
\end{aligned}$$

Par conséquent, lorsque t grandit, la moyenne de $\xi(t)$ tend vers $2/\lambda$ et le temps moyen d'attente, qui vaut $1/\lambda$, en représente la moitié, conformément à l'intuition.

3.3. Analogie discret : le processus de Bernoulli

L'analogie du processus de Poisson à temps discret est défini de la manière suivante.

DÉFINITION 3.6.— Soit $p \in]0, 1[$ et $(\tilde{\xi}_n, n \in \mathbf{N})$, une suite de variables aléatoires indépendantes et identiquement distribuées, de loi géométrique de paramètre p , c'est-à-dire :

$$\mathbf{P}(\tilde{\xi}_0 = k) = p(1-p)^{k-1}; \quad k \geq 1.$$

On pose alors $\tilde{T}_0 = 0$ et $\tilde{T}_{n+1} - \tilde{T}_n = \tilde{\xi}_n$ pour tout $n \in \mathbb{N}$. Le processus ponctuel $(T_0, T_1, \dots, T_n, \dots)$ ainsi défini est appelé processus de Bernouilli.

REMARQUE.— Comme pour le processus de Poisson, on peut définir le processus aléatoire \tilde{N} , à trajectoires cadlag, défini pour tout $t \geq 0$ par :

$$\tilde{N}(t) = \sum_{n \in \mathbb{N}} \mathbf{1}_{\{\tilde{T}_n \leq t\}},$$

qui compte le nombre de points du processus jusqu'à t . Il est clair que $\tilde{N}(t)$ suit pour tout $t \geq 0$, la loi binomiale $\mathcal{B}([t], p)$, où $[\cdot]$ désigne la partie entière.

Nous pouvons vérifier que le processus de Bernouilli satisfait à l'analogie du paradoxe de l'autobus. Pour tout $k \in \mathbb{N}$, notons par analogie avec la section 3.2.2 :

$$\begin{cases} \tilde{W}_k &= k - \tilde{T}_{\tilde{N}(k)}, \text{ le temps écoulé depuis le dernier point avant } k; \\ \tilde{Z}_k &= \tilde{T}_{\tilde{N}(k)+1}, \text{ le temps à attendre avant le premier point après } k. \end{cases}$$

Par définition, \tilde{W}_k est nul s'il y a précisément un point en k , alors que \tilde{Z}_k est nécessairement strictement positif. On a le résultat suivant.

THÉORÈME 3.8.— Pour tout $k \geq 0$, \tilde{W}_k suit la loi géométrique de paramètre p et est indépendante de \tilde{Z}_k , dont la loi est donnée par :

$$\mathbf{P}(\tilde{Z}_k = i) = \begin{cases} p(1-p)^i & \text{si } i < k; \\ (1-p)^i & \text{si } i = k; \\ 0 & \text{si } i > k. \end{cases}$$

Démonstration. Le schéma de la preuve est le même que celui de la preuve du théorème 3.7. Fixons $j > 1$. Il est clair par définition que \tilde{Z}_k ne peut pas être supérieur à k . Par ailleurs, pour tout $i < k$, $\tilde{Z}_k = i$ si $\tilde{N}(k)$ est strictement positif (sinon $k = i$) et strictement inférieur à $k - i$, car il n'y a aucun point entre l'instant $k - i$ et l'instant k . Par conséquent :

$$\begin{aligned} \mathbf{P}(\tilde{Z}_k = i, \tilde{W}_k = j) &= \sum_{1 \leq n \leq k-i} \mathbf{P}(\tilde{Z}_k = i, \tilde{W}_k = j, \tilde{N}(k) = n) \\ &= \sum_{1 \leq n \leq k-i} \mathbf{P}(\tilde{T}_n = k - i, \tilde{\xi}_n = i + j) \\ &= \sum_{1 \leq n \leq k-i} \mathbf{P}(\tilde{T}_n = k - i) p(1-p)^{i+j-1}, \end{aligned}$$

par indépendance. Or, $\tilde{T}_n = k - i$ revient à dire qu'il y a un point à $k - i$ (et donc un succès à une expérience de Bernouilli de paramètre p) et si $k - i - 1 > 0$, qu'il

y a eu $n - 1$ succès lors de $k - i - 1$ expériences de Bernouilli précédentes, toutes indépendantes. Autrement dit :

$$\mathbf{P}(\tilde{T}_n = k - i) = p\mathbf{P}(B = n - 1) = pC_{k-i-1}^{n-1}p^{n-1}(1-p)^{k-i-1-(n-1)},$$

où B est une variable aléatoire de loi binomiale $\mathcal{B}(k - i - 1, p)$ et en comprenant éventuellement $C_0^0 = 1$. Pour tout $i < k$, on a donc :

$$\begin{aligned} \mathbf{P}(\tilde{Z}_k = i, \tilde{W}_k = j) &= \sum_{1 \leq n \leq k-i} pC_{k-i-1}^{n-1}p^{n-1}(1-p)^{k-i-1-(n-1)}p(1-p)^{i+j-1} \\ &= p^2(1-p)^{i+j-1} \sum_{0 \leq n \leq k-i-1} C_{k-i-1}^n p^n (1-p)^{k-i-1-n} \\ &= p(1-p)^{j-1}p(1-p)^i, \end{aligned} \quad (3.4)$$

d'après la formule du binôme. Par ailleurs, $\tilde{Z}_k = k$ revient à dire que $\tilde{N}(k) = 0$ et donc :

$$\begin{aligned} \mathbf{P}(\tilde{Z}_k = k, \tilde{W}_k = j) &= \mathbf{P}(\tilde{\xi}_0 = i + j) \\ &= p(1-p)^{i+j-1} = p(1-p)^{j-1}(1-p)^i. \end{aligned} \quad (3.5)$$

De (3.4) et (3.5), on déduit que \tilde{Z}_k et \tilde{W}_k sont indépendantes et suivent les lois annoncées. \square

⌘ Cette propriété d'oubli se comprend aisément : une variable aléatoire $\tilde{\xi}_i$ compte le nombre d'essais de Bernouilli indépendants, nécessaires pour obtenir un succès. La variable aléatoire \tilde{W}_k compte le nombre d'essais encore nécessaires à partir de k pour obtenir le premier succès après k . Les essais étant indépendants, on voit bien là encore que l'attente « capitalisée » entre le dernier succès jusqu'à k n'augmente pas la chance de succès à chaque tentative après k et donc \tilde{W}_k a la même loi que n'importe lequel des $\tilde{\xi}_i$.

Le dernier résultat est donc intuitivement clair et il peut nous aider à comprendre le paradoxe de l'autobus en temps continu. En effet, le processus de Poisson n'est autre qu'une version macroscopique de processus de Bernouilli. Plus précisément, on fixe $\lambda > 0$ et pour tout entier $n > 1$ tel que $\lambda/n < 1$, on note \tilde{N}^n , un processus de Bernouilli de paramètre λ/n , de variables associées $\tilde{\xi}_0^n, \tilde{\xi}_1^n, \dots$. Finalement, on définit pour tout n le processus ponctuel \tilde{N}^n comme suit :

$$\left\{ \begin{array}{l} \bar{\xi}_i^n = \tilde{\xi}_i^n/n, \quad i \in \mathbf{N}; \\ \bar{T}_0^n = 0, \quad \bar{T}_{i+1}^n - \bar{T}_i^n = \bar{\xi}_i^n, \quad i \in \mathbf{N}; \\ \bar{N}^n(t) = \sum_{i \in \mathbf{N}} \mathbf{1}_{\{\bar{T}_i^n \leq t\}}. \end{array} \right.$$

A partir du processus de Bernouilli de paramètre λ , pour obtenir \bar{N}^n , on divise donc par n la probabilité d'occurrence d'un point à chaque instant et l'on compense par une « accélération du temps » de facteur n , en divisant les temps interpoints par n . On a alors le résultat suivant.

THÉORÈME 3.9.— La suite de processus $\{\bar{N}^n, n > \lfloor \lambda \rfloor\}$ converge en loi vers le processus de Poisson de paramètre λ .

Démonstration. La notion de convergence en loi pour les processus est assez lourde à définir. Nous n'entrerons pas ici dans ces détails. En fait, il suffit dans notre cas de vérifier que les interpoints de \bar{N}^n tendent en loi vers ceux du processus de Poisson de paramètre λ . C'est clairement le cas puisque pour tout $i \in \mathbf{N}$ et tout t , on a :

$$\mathbf{P}(\bar{\xi}_i^n > t) = \mathbf{P}(\tilde{\xi}_i^n > nt) = \left(1 - \frac{\lambda}{n}\right)^{\lfloor nt \rfloor},$$

et la dernière quantité tend vers $e^{-\lambda t}$ quand n tend vers l'infini. \square

3.4. Simulation du processus de Poisson

La définition 3.1 permet de simuler des trajectoires d'un processus de Poisson d'intensité λ simplement en effectuant des tirages successifs de variables aléatoires de loi exponentielle de paramètre λ .

Algorithme 3.1 . Réalisation d'une trajectoire d'un processus de Poisson (méthode 1)

Données : λT

Résultat : une trajectoire $(t_n, n \geq 1)$ sur $[0, T]$ d'un processus de Poisson d'intensité λ

$t \leftarrow 0;$

$n \leftarrow 0;$

tant que $t \leq T$ **faire**

$x \leftarrow$ réalisation d'une $\varepsilon(\lambda);$

$t \leftarrow t + x;$

$t_n \leftarrow t;$

$n \leftarrow n + 1$

fin

retourner t_1, t_2, \dots, t_n

La définition 3.2 permet de simuler une trajectoire sur $[0, t]$ en effectuant un tirage d'une loi de Poisson de paramètre λt , dont on note le résultat par k , puis d'effectuer

k tirages uniformes sur $[0, t]$. Une fois ordonné, le plus petit de ces tirages peut être assimilé au premier impact, le deuxième plus petit à T_2 , etc.

Algorithme 3.2 . Réalisation d'une trajectoire d'un processus de Poisson (méthode 2)

Données : λT

Résultat : une trajectoire $(t_n, n \geq 1)$ sur $[0, T]$ d'un processus de Poisson d'intensité λ

$n \leftarrow$ réalisation d'une variable aléatoire de loi de Poisson de paramètre λT ;

pour $i = 1, \dots, n$ **faire**

$u_i \leftarrow$ réalisation d'une $U([0, 1])$;

fin

$(t_1, \dots, t_n) \leftarrow$ tri en croissant de (u_1, \dots, u_n) ;

retourner t_1, t_2, \dots, t_n

Dans les deux méthodes, si l'on veut ensuite prolonger la trajectoire sur $[t, t + s]$, on recommence la même chose sur $[t, t + s]$ seulement, puisque la définition 3.3 garantit que la trajectoire sur $[t, t + s]$ est indépendante de celle sur $[0, t]$.

La méthode 1 est la plus avantageuse pour les « grandes » valeurs de λ , parce que simuler une loi de Poisson pour λ grand se révèle impossible dès lors que $\exp(-\lambda)$ devient plus petit que la précision numérique de l'ordinateur. En revanche, comme nous le voyons dans le chapitre ??, la méthode 2 est la seule qui s'adapte à la simulation d'un processus de Poisson spatial.

3.5. Processus de Poisson non homogène

DÉFINITION 3.7.— N est un processus de Poisson non homogène d'intensité $\lambda(s)$ où s varie dans \mathbf{R}^+ si et seulement si pour toute fonction $f : \mathbf{R}^+ \rightarrow \mathbf{R}^+$, on a l'identité suivante :

$$\exp\left(-\sum_{n \geq 1} f(T_n)\right) = \exp\left(-\int \left(1 - e^{-f(s)}\right) \lambda(s) \, ds\right). \quad (3.6)$$

Cette définition doit être rapprochée de la définition 3.4. Si λ est une fonction constante alors on retrouve la définition d'un processus de Poisson d'intensité λ . Cette classe de processus modélise des trafics dans lesquels le débit varie au cours du temps mais de façon connue (λ est une fonction déterministe). On suppose toujours que λ est

continue à droite pourvue de limites à gauche en tout point. Posons :

$$a(t) = \int_0^t \lambda(s) \, ds;$$

$$\tau(t) = \inf\{s \geq 0, a(s) \geq t\}.$$

Le graphe de τ s'obtient en prenant le symétrique de celui de a par rapport à la première bissectrice. Par définition, $a(\tau(t)) = t$. La relation inverse, $\tau(a(t)) = t$ n'est vraie que si λ ne s'annule pas sur un intervalle. En effet, dans ce cas, a présente un intervalle de constance et τ n'est plus continue.

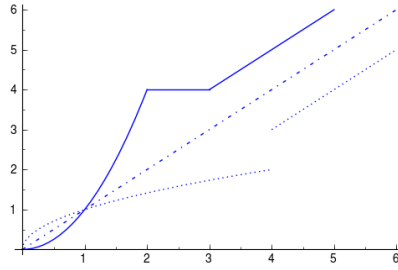


Figure 3.6 – Graphe de a et de son inverse τ (en pointillés)

THÉORÈME 3.10.— Soit N un processus de Poisson non homogène d'intensité $\lambda(s) \, ds$ alors le processus \tilde{N} défini par :

$$\tilde{N}(t) = \sum_{s: \Delta N(s)=1} \mathbf{1}_{\{a(s) \leq t\}}, \quad (3.7)$$

est un processus de Poisson homogène d'intensité ds .

✎ En d'autres termes, quand N saute à l'instant t , \tilde{N} a un saut à l'instant $a(t)$.

Démonstration. Il suffit de montrer que $M(t) = N(\tau_t)$ vérifie les propriétés de la définition 3.3. Or, pour tout $0 = t_0 < t_1 < \dots < t_n$, on l'identité suivante :

$$M(t_{i+1}) - M(t_i) = \sum_n \mathbf{1}_{] \tau_{t_i}, \tau_{t_{i+1}}]}(T_n).$$

Comme τ est déterministe, par définition du processus de Poisson non homogène, on a :

$$\begin{aligned}
\mathbf{E} \left[\exp \left(- \sum_j \alpha_j (M(t_{j+1}) - M(t_j)) \right) \right] &= \\
&= \mathbf{E} \left[\exp \left(- \sum_n \sum_j \alpha_j \mathbf{1}_{\{\tau_{t_j}, \tau_{t_{j+1}}\}}(T_n) \right) \right] \\
&= \exp \left(- \sum_j \int_{\tau_{t_j}}^{\tau_{t_{j+1}}} (1 - e^{-\alpha_j}) \lambda(s) \, ds \right) \\
&= \exp \left(- \sum_j (1 - e^{-\alpha_j}) [a(\tau_{t_{j+1}}) - a(\tau_{t_j})] \right) \\
&= \exp \left(- \sum_j \alpha_j (t_{j+1} - t_j) \right),
\end{aligned}$$

d'où le résultat. \square

Avec les notations précédentes, on en déduit l'algorithme de simulation d'un processus de Poisson non homogène.

Algorithme 3.3 . Réalisation d'une trajectoire d'un processus de Poisson non homogène

Données : a, T

Résultat : une trajectoire $(t_n, n \geq 1)$ sur $[0, T]$ d'un processus de Poisson d'intensité $\lambda(s) \, ds$

$s_1, \dots, s_n \leftarrow$ réalisation d'un processus de Poisson d'intensité 1 sur $a(T)$;

retourner $t_i = a(s_i), i = 1, \dots, n$

THÉORÈME 3.11.— Le processus ponctuel N est un processus de Poisson non homogène d'intensité $\lambda(s) \, ds$ si et seulement si le processus \tilde{N} est une martingale où :

$$\tilde{N} : t \longmapsto N(t) - \int_0^t \lambda(s) \, ds.$$

Démonstration. Soit f à support compact. D'après la formule d'Itô (A.14), on a :

$$\begin{aligned}
\exp \left(\int_0^t f(r) \, d\tilde{N}(r) \right) &= 1 + \int_0^t \exp \left(\int_0^s f(r) \, d\tilde{N}(r) \right) f(s) \, d\tilde{N}(s) \\
&\quad + \int_0^t \exp \left(\int_0^s f(r) \, d\tilde{N}(r) \right) \left(e^{f(s)} - 1 - f(s) \right) \, dN(s).
\end{aligned}$$

En vertu du théorème A.34, l'intégrale stochastique du terme de droite est une martingale, donc en prenant les espérances, il reste :

$$\begin{aligned} \mathbf{E} \left[\exp \left(\int_0^t f(r) d\tilde{N}(r) \right) \right] \\ = 1 + \mathbf{E} \left[\int_0^t \exp \left(\int_0^s f(r) d\tilde{N}(r) \right) (e^{f(s)} - 1 - f(s)) dN(s) \right] \\ = 1 + \mathbf{E} \left[\int_0^t \exp \left(\int_0^s f(r) d\tilde{N}(r) \right) (e^{f(s)} - 1 - f(s)) \lambda(s) ds \right]. \end{aligned}$$

En posant $\phi(t) = \exp(\int_0^t f(r) d\tilde{N}(r))$, il vient alors :

$$\phi(t) = 1 + \int_0^t \phi(s) u(s) ds,$$

où $u(s) = (e^{f(s)} - 1 - f(s))\lambda(s)$. En résolvant l'équation différentielle, on obtient :

$$\phi(t) = \exp \left(\int_0^t (e^{f(s)} - 1 - f(s)) \lambda(s) ds \right),$$

soit en simplifiant les deux côtés par $\exp \left(- \int_0^t f(s) \lambda(s) ds \right)$:

$$\mathbf{E} \left[\exp \left(\int_0^t f(r) dN(r) \right) \right] = \exp \left(\int_0^t (e^{f(s)} - 1) \lambda(s) ds \right).$$

Réciproquement, en appliquant (3.6) à :

$$f = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{]t_i, t_{i+1}]},$$

on obtient :

$$\begin{aligned} \mathbf{E} \left[\exp \left(- \sum_i \alpha_i (N(t_{i+1}) - N(t_i)) \right) \right] \\ = \exp \left(- \lambda \sum_i \int_0^\infty (1 - e^{-\alpha_i \mathbf{1}_{]t_i, t_{i+1}]}(s)}) ds \right). \end{aligned}$$

Remarquons alors que la fonction $s \rightarrow (1 - e^{-\alpha_i \mathbf{1}_{]t_i, t_{i+1}]}(s)})$ s'annule en dehors de l'intervalle $]t_i, t_{i+1}]$ et vaut $1 - e^{-\alpha_i}$ sur cet intervalle. On a donc :

$$\mathbf{E} \left[\exp \left(- \sum_i \alpha_i (N(t_{i+1}) - N(t_i)) \right) \right] = \exp \left(- \sum_i (1 - e^{-\alpha_i}) \int_{t_i}^{t_{i+1}} \lambda(s) ds \right).$$

On obtient donc que la transformée de Laplace du vecteur aléatoire $(N(t_i + 1) - N(t_i), 1 \leq i \leq n-1)$ est le produit des transformées de Laplace de chaque composante (puisque qu'elle s'écrit comme produit de fonctions dépendant chacune d'un seul des α_i), donc les variables aléatoires sont indépendantes. Par classe monotone, on en déduit que $N(t + s) - N(t)$ est indépendante de $\mathcal{F}_t = \sigma(N(r), r \leq t)$. Dans le cas où $n = 2, t_1 = a, t_2 = b$ on obtient de la formule précédente que :

$$\mathbf{E} [\exp(-\alpha (N(b) - N(a)))] = \exp(-(1 - e^{-\alpha})) \int_a^b \lambda(s) \, ds,$$

donc $N(b) - N(a)$ suit une loi de Poisson de paramètre $\int_a^b \lambda(s) \, ds$. Par conséquent :

$$\mathbf{E} [N(t + s) - N(t) | \mathcal{F}_t] = \mathbf{E} [N(t + s) - N(t)] = \int_t^{t+s} \lambda(s) \, ds,$$

donc \tilde{N} est une martingale. □

3.6. Processus de Cox

Les processus de Cox sont des processus de Poisson pour lesquels on choisit aléatoirement l'intensité. De ce qui précède, on voit que cela revient à mettre une mesure de probabilité sur les fonctions positives (de préférence cadlag).

DÉFINITION 3.8.— Soit M une variable aléatoire à valeurs dans \mathbb{D} . Le processus ponctuel N est un processus de Cox d'intensité M si et seulement si pour toute fonction f à support compact, on a :

$$\mathbf{E} \left[\exp \left(- \sum_n f(T_n) \right) \mid M \right] = \exp \left(- \int (1 - e^{-f(s)}) M(s) \, ds \right).$$

Par conséquent, pour calculer la transformée de Laplace, il faut déconditionner par rapport à M , d'où :

$$\mathbf{E} \left[\exp \left(- \sum_n f(T_n) \right) \right] = \mathbf{E} \left[\exp \left(- \int (1 - e^{-f(s)}) M(s) \, ds \right) \right].$$

Les résultats précédents de martingales s'étendent sans changement autre que celui de la filtration, qui doit ici être prise égale à $\mathcal{F}_0 = \sigma(M)$, $\mathcal{F}_t = \sigma(M) \vee \sigma(N(r), r \leq t)$.

L'exemple de processus de Cox qui nous sera le plus utile est celui des MMPP, voir la section 4.6.

3.7. Problèmes

EXERCICE 10.— Soit N un processus de Poisson d'intensité λ , on note T_n le n^{e} instant de saut. Par convention, $T_0 = 0$. Soit $(Z_n, n \geq 1)$, une suite de variables aléatoires de même loi telles que pour tout n , T_n et Z_n sont indépendantes. Soit g la densité de la loi commune aux Z_n .

1) Montrer que pour toute fonction f :

$$E[f(T_n, Z_n)] = \int_0^{+\infty} \int f(t, z) g(z) \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} dz dt.$$

2) En déduire que :

$$E\left[\sum_{n \geq 1} f(T_n, Z_n)\right] = \lambda \int_0^{+\infty} \int f(t, z) g(z) dz dt.$$

3) On suppose que les communications téléphoniques d'un abonné durent un temps aléatoire de loi exponentielle de moyenne trois minutes. Ces durées sont indépendantes entre elles. Au siècle dernier, le coût d'une communication était fonction de sa durée t selon la formule suivante :

$$c(t) = \alpha \text{ si } t \leq t_0 \text{ et } c(t) = \alpha + \beta(t - t_0) \text{ si } t \geq t_0.$$

Déduire de ce qui précède que le coût moyen d'une heure totale de communication est donné par :

$$\lambda \int_0^{t_0} c(t) \lambda e^{-\lambda t} dt$$

avec $\lambda = 20$. (Indication : considérer $Z_n = T_{n+1} - T_n$ et expliquer pourquoi on peut appliquer le résultat précédent.)

Application numérique : pour les communications locales, en 1999, on avait les paramètres suivants : $t_0 = 3$ minutes, $\alpha = 0,11$ euro et $\beta = 0,043$ euro par minute. Pour les communications nationales, $t_0 = 39$ secondes et $\beta = 0,17$ euro par minute. α était le même. En tarif réduit, diviser β par 2. En mettant $t_0 = 1$ minute et $\alpha = 0,15$ euro, à combien s'élevait le prix de la seconde supplémentaire en téléphonie mobile dans un forfait dont le montant pour 1 heure de communication était de 23,62 euros ?

EXERCICE 11.— Un distributeur automatique de billets enregistre les instants de début et de fin de requêtes de ces clients mais évidemment par leur heure d'arrivée dans la file. Un nouveau cycle d'activité ayant commencé à 7h30, on a relevé les données du tableau 11.

Supposons que les arrivées aient lieu selon processus de Poisson, que peut-on dire de l'instant d'arrivée T_1 du client numéro 1 ? En particulier, calculer son espérance.

Client numéro	Début de service	Fin de service
0	7h30	7h34
1	7h34	7h40
2	7h40	7h42
3	7h45	7h50

Tableau 3.1 –

EXERCICE 12.– Une compagnie d'assurance doit payer pour des sinistres au rythme de cinq par jour. On suppose que les instants d'occurrence des sinistres suivent un processus de Poisson et que les montants des dégâts sont indépendants les uns des autres, de loi exponentielle, de moyenne cinq cents euros. On introduit $(X_i, i \geq 1)$ une suite de variables aléatoires indépendantes équidistribuées, de loi exponentielle de moyenne $1/\mu = 3\,000$ euros et N un processus de Poisson d'intensité $\lambda = 5 \text{ jours}^{-1}$, indépendant des X_i .

- 1) Que représente $Z = \sum_{i=1}^{N(365)} X_i$?
- 2) Calculer le montant total moyen dépensé par an par la compagnie d'assurance.
- 3) Calculer $E[e^{-sZ}]$.
- 4) En déduire la variance de Z .

3.8. Notes et commentaires

Pour des résultats plus approfondis sur les processus ponctuels en dimension quelconque, on pourra consulter [BRE 81, DAL 03, LAS 95]. Pour les convergences en loi de processus ponctuels, on pourra consulter l'annexe D de [ROB 03].

Résumé

- Un processus de Poisson est un processus qui représente un flux aléatoire à débit moyen constant dans le temps.
- Son intensité λ représente le nombre moyen d'arrivées par unité de temps.
- La superposition de deux processus de Poisson est un processus de Poisson. L'amincissement d'un processus de Poisson est un processus de Poisson.
- Un processus de Poisson représente bien les événements issus de l'activité humaine (téléphone, instants d'arrivée dans un magasin, ouverture de session mail, ouverture de page *web*, etc.) pas ceux de l'activité d'une machine (envoi de paquets, messages de signalisation, etc.).

Chapitre 4

Processus de Markov

La modélisation markovienne d'un système dynamique se fait naturellement en utilisant une chaîne de Markov dont les durées de séjour dans chaque état deviennent aléatoires. Cette description est insuffisante pour établir les propriétés mathématiques intéressantes. On a, pour cela, recours au formalisme des processus markoviens avec leur semi-groupe et leur générateur infinitésimal. Pour aller plus loin et prouver notamment des identités comme PASTA ou ses avatars, on a besoin de caractériser les processus de Markov comme solutions de problèmes de martingales. Dans ce chapitre, nous passons en revue ces différentes caractérisations et montrons qu'elles sont équivalentes.

Dans tout ce chapitre, E désigne un espace au plus dénombrable muni de la topologie discrète. On rappelle qu'il est conseillé de lire le petit paragraphe sur les notations de la remarque A.1.1.

4.1. Préliminaires

Nous commençons par énoncer deux lemmes techniques sur la loi exponentielle, qui nous seront utiles dans les calculs qui suivent.

LEMME 4.1.– Soit U et V , deux variables aléatoires indépendantes, de lois respectives $\varepsilon(\lambda)$ et $\varepsilon(\mu)$, où $\lambda, \mu > 0$. Les deux propriétés suivantes sont satisfaites :

- (i) $\mathbf{P}(U \leq V) = \lambda/(\lambda + \mu)$;
- (ii) $U \wedge V \sim \varepsilon(\lambda + \mu)$.

Démonstration. (i) La densité du couple de variables aléatoires (U, V) est donnée pour tout (u, v) par :

$$f_{(U,V)}(u, v) = \lambda e^{-\lambda u} \mu e^{-\mu v} \mathbf{1}_{\mathbf{R}^+}(u) \mathbf{1}_{\mathbf{R}^+}(v).$$

On notant le sous-ensemble $A = \{(u, v) \in \mathbf{R}^2; u \leq v\}$, on peut donc écrire :

$$\begin{aligned} \mathbf{P}(U \leq V) &= \int \int_A f_{(U,V)}(u, v) du dv \\ &= \int_0^{+\infty} \lambda e^{-\lambda u} \left\{ \int_u^{+\infty} \mu e^{-\mu v} dv \right\} du \\ &= \int_0^{+\infty} \lambda e^{-(\lambda+\mu)u} du = \frac{\lambda}{\lambda + \mu}. \end{aligned}$$

(ii) Il suffit de voir que, pour tout $x \geq 0$, on a :

$$\mathbf{P}(U \wedge V \geq x) = \mathbf{P}(\{U \geq x\} \cap \{V \geq x\}) = e^{-\lambda x} e^{-\mu x} = e^{-(\lambda+\mu)x}.$$

D'où le résultat. \square

LEMME 4.2 (Propriété d'oubli de la loi exponentielle).— Soit U , une variable aléatoire de loi $\varepsilon(\lambda)$, où $\lambda > 0$. Soit $t \geq 0$. Conditionnellement à $\{U \geq t\}$, la variable aléatoire $U - t$ suit aussi la loi $\varepsilon(\lambda)$.

Démonstration. Il suffit de calculer, pour tout $x \geq 0$, la probabilité conditionnelle :

$$\mathbf{P}(U \geq t + x \mid U \geq t) = \frac{\mathbf{P}(U \geq t + x)}{\mathbf{P}(U \geq t)} = \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x}.$$

D'où le résultat. \square

✎ Si la durée de vie de chaque humain suivait une loi exponentielle, le résultat précédent signifierait que la probabilité d'atteindre 90 ans quand on a déjà plus de 80 ans est la même que celle d'atteindre l'âge de 10 ans au moment de sa naissance. Comme il y a sensiblement moins d'octogénaires qui deviennent nonagénaires que de bébés qui atteignent le collège, on ne peut raisonnablement pas représenter la durée de vie d'un humain par une telle loi.

4.2. Construction trajectorielle

DÉFINITION 4.1.— Soit ν une mesure de probabilité sur E et $Q = (q(x, y), x, y \in E)$ une famille de réels tels que pour tout $(x, y) \in E$, les propriétés suivantes soient satisfaites :

$$q(x, y) \geq 0 \text{ et } \sum_{y \neq x} q(x, y) = 1. \quad (4.1)$$

Le processus de Markov X de paramètres (ν, Q) est ainsi construit :

- $X(0)$ est une variable aléatoire de loi ν ;
- si $X(0) = x$, on considère $\xi_1 = T_1 \sim \varepsilon(q(x, x))$ indépendante de $X(0)$;
 $X(t) = x$ pour $t < T_1$; ensuite, on considère \hat{X}_1 une variable aléatoire indépendante de $(X(0), T_1)$ telle que :

$$\mathbf{P}(\hat{X}_1 = y) = q(x, y);$$

- soit x_1 la valeur de \hat{X}_1 ; on considère $\xi_2 \sim \varepsilon(q(x_1, x_1))$ indépendante de $(X(0), \xi_1, \hat{X}_1)$. $X(t) = \hat{X}_1$ pour $T_1 \leq t < T_2$;
- on continue cette construction jusqu'à l'infini.

REMARQUE.— Une variable aléatoire de loi exponentielle de paramètre 0 doit être comprise comme égale presque sûrement à l'infini. Pour tout point $x \in E$ tel que $q(x, x) = 0$, (4.1) impose que $q(x, y) = 0$ pour tout y . Le point x est donc un point « cimetière », puisque l'on n'en sort jamais.

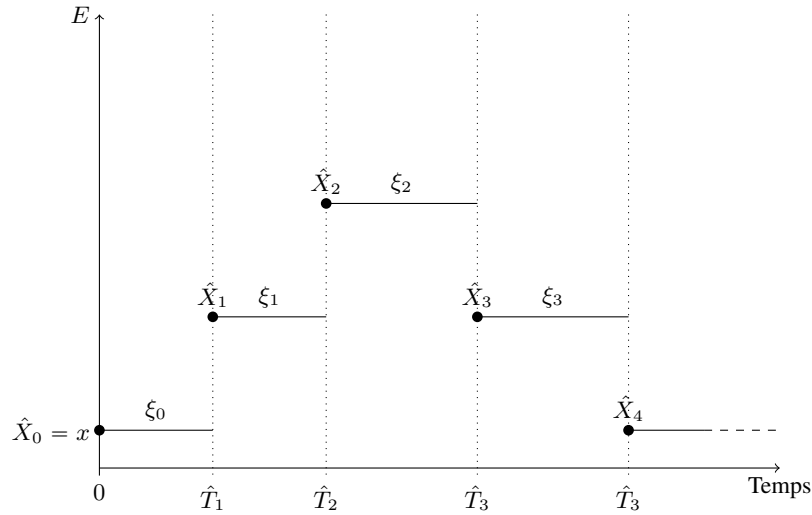


Figure 4.1 – Réalisation d'une trajectoire d'un processus de Markov de paramètres (ν, Q)

REMARQUE.— Compte tenu des hypothèses d'indépendance, il est évident que la suite $(\hat{X}_n, n \geq 0)$ est une chaîne de Markov de matrice de transition \hat{Q} définie par :

$$\hat{Q}(x, y) = \begin{cases} q(x, y) & \text{si } x \neq y \\ 0 & \text{si } x = y. \end{cases}$$

Cette chaîne de Markov s'appelle la chaîne incluse. Il ne peut pas y avoir de transition d'un état vers lui-même puisque l'on observe les positions de X seulement lors de ses changements d'état.

EXEMPLE 4.1 (File M/M/1).— Dans ce modèle, les arrivées ont lieu selon un processus de Poisson d'intensité λ et les temps de séjour sont indépendants et de loi exponentielle de paramètre μ . Le processus étudié $X = (X(t), t \geq 0)$ est le processus qui compte le nombre de clients dans le système à chaque instant. L'espace d'états E est égal à l'espace des entiers naturels. S'il y a $i \neq 0$ clients dans le système, le prochain événement est un départ ou une arrivée. En vertu du paradoxe de l'autobus (théorème 3.7), la prochaine arrivée aura lieu après une durée de loi exponentielle de paramètre λ . Par ailleurs, le théorème 4.3 ci-après nous apprend que le prochain départ aura lieu après un temps exponentiel de paramètre μ . En vertu du lemme 4.1, le temps de séjour dans l'état i suit donc une loi exponentielle de paramètre $\lambda + \mu$ donc $q(i, i) = \lambda + \mu$ pour $i \neq 0$. Si $i = 0$, il ne peut pas y avoir de départ donc $q(0, 0) = \lambda$. Toujours en vertu du lemme 4.1, la probabilité de passer de l'état i à l'état $(i + 1)$ correspond à la probabilité qu'une variable aléatoire de loi exponentielle de paramètre λ se termine avant une variable aléatoire de loi exponentielle de paramètre μ donc :

$$q(i, i + 1) = \frac{\lambda}{\lambda + \mu} \text{ et de même } q(i, i - 1) = \frac{\mu}{\lambda + \mu}.$$

Si $i = 0$, on obtient $q(0, 1) = 1$.

Il nous reste donc à démontrer notre assertion sur la loi du temps de service résiduel. Le processus des départs n'est pas de façon évidente un processus de Poisson puisqu'il n'y a pas de départs quand la file est vide. On ne peut donc pas *a priori* appliquer le paradoxe de l'autobus (voir le théorème 3.7). Néanmoins, si l'on conditionne par le fait d'être dans une période d'activité alors les départs forment un processus de Poisson. Nous formalisons cette idée dans le théorème suivant.

THÉORÈME 4.3.— Soit $R(t)$ le temps à venir entre t et le prochain départ. Pour tout $t \geq 0$, $R(t)$ suit, conditionnellement à $\{X(t) > 0\}$, la loi $\varepsilon(\mu)$.

Démonstration. On note $T'_0 < T'_1 < \dots$ les instants successifs de sortie du système et pour tout $t \geq 0$, $B_t \in \mathbb{N}^*$, l'indice du dernier client entré dans un système vide avant t . Autrement dit, T_{B_t} représente l'instant de début de la dernière période d'activité débutée avant t . A partir de là, on marque les instants de sortie du serveur par un processus ponctuel de la façon suivante :

$$\tilde{T}_0^t = T_{B_t}; \quad \tilde{T}_k^t = T'_{B_t+k-1}; \quad k > 1,$$

le k^{e} instant de sortie du serveur à partir du début de la dernière période d'activité débutée avant t . Pour tout $s \geq T_{B_t}$, on note également :

$$M_s^t = \sum_{k \in \mathbb{N}^*} \mathbf{1}_{\{\tilde{T}_k^t \leq s\}}$$

et pour tout $u \geq 0$:

$$\tilde{M}_u^t = M_{T_{B_t} + u}^t.$$

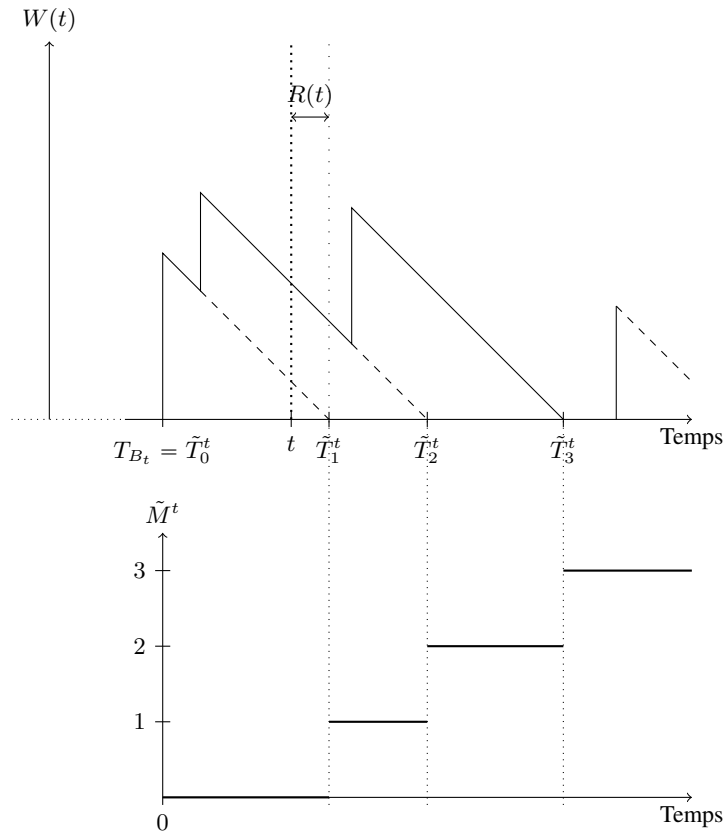


Figure 4.2 – Notations

Les notations sont un peu compliquées, mais l'idée est simple : le processus ponctuel $(\tilde{M}_u^t, u \geq 0)$ marque les instants de sortie du serveur à partir du début de la dernière période d'activité débutée avant t , instant auquel on initialise l'échelle de temps. Le

nombre de sorties du serveur entre le début de la période d'activité et t est donné par $\tilde{M}_{t-T_{B_t}}^t$. Sur $\{X(t) > 0\}$, le système n'est jamais vide entre T_{B_t} et t et pour tout $1 \leq k \leq \tilde{M}_{t-T_{B_t}}^t + 1$, on a donc :

$$\tilde{T}_k^t - \tilde{T}_{k-1}^t = \sigma_{B_t+k-1},$$

le temps de service du $k - 1^{\text{e}}$ client depuis le début de la période d'activité. Ces temps de service sont indépendants les uns des autres, indépendants du passé avant T_{B_t} et tous de loi $\varepsilon(\mu)$ pour $k \leq \tilde{M}_{t-T_{B_t}}^t$. Donc, conditionnellement à $\{X(t) > 0\}$, le processus $(\tilde{M}_u^t, u \geq 0)$ est égal en loi sur l'intervalle $[0, \tilde{T}_{\tilde{M}_{t-T_{B_t}}^t}^t]$ à un processus de Poisson de paramètre μ , dont $\tilde{T}_0 < \tilde{T}_1 < \dots$ représentent les points et $\tilde{\xi}_0, \tilde{\xi}_1, \dots$ les tailles des intervalles entre les points. Pour tout x , on peut donc écrire que :

$$\begin{aligned} \mathbf{P}(R(t) \geq x \mid X(t) > 0) \\ = \sum_{k \in \mathbf{N}} \mathbf{P}\left(\left\{\tilde{T}_k + \tilde{\xi}_k - (t - T_{B_t}) \geq x\right\} \cap \{\tilde{T}_k \leq t - T_{B_t} < \tilde{T}_{k+1}\}\right), \end{aligned}$$

et l'on peut poursuivre le calcul comme dans la preuve du théorème 3.7 pour conclure la preuve. \square

DÉFINITION 4.2.— Un processus de Markov X est dit irréductible (respectivement récurrent, transient) si et seulement si la chaîne incluse \hat{X} est irréductible (respectivement récurrente, transiente).

Une autre construction équivalente est possible. Elle est plus artificielle mais permet de démontrer plus facilement certaines propriétés mathématiques.

DÉFINITION 4.3.— Un processus X de paramètres (ν, Q) est régulier lorsque :

$$\|q\|_\infty = \sup_{x \in E} q(x, x) < \infty.$$

Soit X un processus de Markov régulier de paramètres (ν, Q) . On pose :

$$\tilde{q}(x, y) = \begin{cases} \frac{q(x, x)q(x, y)}{\|q\|_\infty} & \text{si } x \neq y, \\ 1 - \frac{q(x, x)}{\|q\|_\infty} & \text{si } x = y. \end{cases}$$

On considère \tilde{X} la chaîne de Markov de loi initiale ν et de matrice de transition \tilde{Q} . On considère aussi N un processus de Poisson d'intensité $\|q\|_\infty$ indépendant de \tilde{X} .

Le processus $Y(t) = \tilde{X}_{N(t)}$ a même loi que X . En effet, dans les trajectoires de ce processus, on peut rester pendant plusieurs transitions dans le même état. Soit $x \in E$ et :

$$\tilde{\tau}_{x^c} = \inf\{n > 0, \tilde{X}_n \neq x\}.$$

Conditionnellement à $\tilde{X}(0) = x$, $\tilde{\tau}_{x^c}$ est indépendante de N et suit une loi géométrique de paramètre $q(x, x)/\|q\|_\infty$. Soit $\tau_{x^c} = \inf\{t > 0, Y(t) \neq x\}$. On a alors :

$$\mathbf{P}(\tau_{x^c} \geq t \mid Y(0) = x) = \mathbf{P}\left(\sum_{j=1}^{\tilde{\tau}_{x^c}} \xi_j \geq t\right),$$

où les ξ_j sont les interpoints de N , donc des variables aléatoires indépendantes entre elles et indépendantes de $\tilde{\tau}_{x^c}$, de loi exponentielle de paramètre $\|q\|_\infty$. Du lemme 4.1, on déduit que le temps de séjour dans l'état x suit bien une loi exponentielle de paramètre $q(x, x)$. De plus, au moment d'un saut de \tilde{X} , on remarque que :

$$\mathbf{P}(\tilde{X}_1 = y \mid \tilde{X}_1 \neq \tilde{X}_0 = x) = \frac{\tilde{q}(x, y)}{1 - \tilde{q}(x, x)} = q(x, y) = \mathbf{P}(\hat{X}_1 = y \mid \hat{X}_0 = x).$$

REMARQUE.— On déduit aussi de cette construction que pour s fixé, il n'y a presque sûrement pas de saut : $\mathbf{P}(\Delta X(s) > 0) = 0$. En effet, on a :

$$\begin{aligned} \mathbf{P}(\Delta X(s) > 0 = 0) &\leq \mathbf{P}(\Delta N(s) = 1) \\ &= \mathbf{E}\left[\int \mathbf{1}_s(x) dN(x)\right] = \int \mathbf{1}_s(x) \|q\|_\infty dx = 0. \end{aligned}$$

Par conséquent, $\mathbf{P} \otimes d$ s -presque sûrement, $X(s) = X(s^-)$. En effet, la mesure de Lebesgue ne voit pas les sauts puisqu'ils sont en nombre au plus dénombrable.

4.3. Semi-groupe markovien et générateur infinitésimal

DÉFINITION 4.4.— Soit X , un processus à valeurs dans E et à trajectoires cadlag et soit $\mathcal{F}_t = \sigma\{X_u, u \leq t\}$. Le processus X vérifie la propriété de Markov simple lorsque pour tout $t, s \geq 0$, on a :

$$\mathbf{E}[f(X(t+s)) \mid \mathcal{F}_t] = \mathbf{E}[f(X(t+s)) \mid X(t)]. \quad (4.2)$$

Le processus X est dit homogène lorsque pour tout $t \geq 0$, pour tout $x \in E$, on a :

$$\mathbf{E}[f(X(t+s)) \mid X(t) = x] = \mathbf{E}[f(X(s)) \mid X(0) = x]. \quad (4.3)$$

Soit X un processus à trajectoires cadlag satisfaisant (4.2) et (4.3). Pour $f \in l^\infty(E)$, pour $x, y \in E$, on pose :

$$P_t(x, y) = \mathbf{P}(X(t) = y \mid X(0) = x);$$

$$P_t f(x) = \sum_{y \in E} f(y) P_t(x, y) = \mathbf{E}[f(X(t)) \mid X(0) = x].$$

THÉORÈME 4.4 (Equation de Kolmogorov).— Pour tout $t \geq 0$, P_t est continu de $l^\infty(E)$ dans lui-même. Par ailleurs, pour tout $t, s \geq 0$, $P_{t+s} = P_t P_s = P_s P_t$.

REMARQUE.— Cette identité s'entend soit comme une identité entre opérateurs, c'est-à-dire que pour toute $f \in l^\infty(E)$, pour tout $x \in E$, on a :

$$P_{t+s}f(x) = P_t(P_sf)(x)$$

où comme une identité entre matrices (même si la notion de matrice de taille infinie reste floue) :

$$P_{t+s}(x, y) = \sum_{z \in E} P_t(x, z) P_s(z, y).$$

La famille $(P_t, t \geq 0)$ un semi-groupe d'opérateurs : tous ces opérateurs satisfont une propriété de stabilité analogue à celle d'un groupe mais chaque élément de la famille n'a pas nécessairement d'inverse pour le produit de composition.

Preuve du théorème 4.4. On remarque d'abord que d'après la définition de P_t , on a :

$$P_t \mathbf{1} = \mathbf{1} \text{ et } |P_t f| \leq P_t |f|.$$

De plus, d'après les propriétés de l'espérance conditionnelle, si $f \geq 0$ alors $P_t f \geq 0$ donc en particulier, si $f \leq g$, on a $P_t f \leq P_t g$. Par conséquent si f est bornée, $P_t f$ l'est aussi :

$$|P_t f(x)| \leq P_t |f|(x) \leq P_t (\|f\|_\infty \mathbf{1})(x) = \|f\|_\infty P_t \mathbf{1}(x) = \|f\|_\infty.$$

Pour tout $t \geq 0$, P_t est donc continu de $l^\infty(E)$ dans lui-même. Comme il est vrai que :

$$\sigma(X(0)) \vee \mathcal{F}_t = \mathcal{F}_t.$$

D'après la propriété d'emboîtement des espérances conditionnelles, on a :

$$\begin{aligned} P_{t+s}f(x) &= \mathbf{E}[f(X(t+s)) \mid X(0) = x] \\ &= \mathbf{E}[\mathbf{E}[f(X(t+s)) \mid \mathcal{F}_t] \mid X(0) = x] \\ &= \mathbf{E}[(P_sf)(X(t)) \mid X(0) = x] \\ &= P_t(P_sf)(x). \end{aligned}$$

On obtient la deuxième identité en conditionnant par \mathcal{F}_s . □

Comme le processus X a des trajectoires cadlag, le théorème de convergence dominée assure que pour tout $x \in E$, $P_t f(x)$ tend vers $f(x)$.

DÉFINITION 4.5.— On dit qu'un semi-groupe P est fortement continu lorsque pour tout $f \in l^\infty(E)$, on a :

$$\lim_{t \rightarrow 0} \sup_{x \in E} |P_t f(x) - f(x)| = 0.$$

DÉFINITION 4.6.— Soit P un semi-groupe fortement continu. Soit $\text{Dom } A$ l'ensemble des fonctions f pour lesquelles $t^{-1}(P_t f(x) - f(x))$ a une limite quand t tend vers 0. Le générateur infinitésimal A du semi-groupe P est défini sur $\text{Dom } A$ par :

$$Af(x) = \lim_{t \rightarrow 0} \frac{1}{t} (P_t f(x) - f(x)). \quad (4.4)$$

THÉORÈME 4.5.— Soit P un semi-groupe fortement continu de générateur infinitésimal A .

(i) Pour $f \in \text{Dom } A$ et $t \geq 0$, la fonction :

$$x \mapsto \int_0^t P_s f(x) \, ds$$

appartient à $\text{Dom } A$ et :

$$P_t f - f = A \int_0^t P_s f \, ds. \quad (4.5)$$

(ii) Pour $f \in \text{Dom } A$ et $t \geq 0$, la fonction $P_t f$ appartient à $\text{Dom } A$ et :

$$\frac{d}{dt} P_t f = A P_t f = P_t A f. \quad (4.6)$$

(iii) Pour $f \in \text{Dom } A$ et $t \geq 0$, l'identité suivante est satisfaite :

$$P_t f - f = \int_0^t A P_s f \, ds = \int_0^t P_s A f \, ds. \quad (4.7)$$

Démonstration. Par définition de A , il faut montrer que :

$$\frac{1}{h} (P_h - \text{Id}) \left(\int_0^t P_s f(x) \, ds \right) \text{ converge quand } h \rightarrow 0.$$

Or :

$$\begin{aligned}
 (P_h - \text{Id}) \left(\int_0^t P_s f(x) \, ds \right) &= \int_0^t P_h P_s f(x) \, ds - \int_0^t P_s f(x) \, ds \\
 &= \int_0^t P_{s+h} f(x) \, ds - \int_0^t P_s f(x) \, ds \\
 &= \int_h^{t+h} P_s f(x) \, ds - \int_0^t P_s f(x) \, ds \\
 &= \int_t^{t+h} P_s f(x) \, ds - \int_0^h P_s f(x) \, ds.
 \end{aligned}$$

Comme P est un semi-groupe fortement continu, la continuité en 0 implique la continuité en s , c'est-à-dire que pour $f \in l^\infty(E)$, l'application $(s \mapsto P_s f)$ est continue de \mathbf{R}^+ dans $l^\infty(E)$. Par conséquent, on a la limite suivante :

$$\lim_{h \rightarrow 0} \frac{1}{h} \left(\int_t^{t+h} P_s f(x) \, ds - \int_0^h P_s f(x) \, ds \right) = P_t f(x) - f(x).$$

Pour $h > 0$, on pose $A_h f = h^{-1}(P_h f - f)$. On a facilement :

$$A_h P_t f = h^{-1}(P_{t+h} f - P_t f) = P_t(h^{-1}(P_h f - f)) = P_t A_h f.$$

Comme $f \in \text{Dom } A$, $A_h f$ tend vers Af quand h tend vers 0 donc $A_h P_t f$ converge aussi, ce qui revient à dire que $P_t f$ appartient à $\text{Dom } A$ et que (4.6) est vraie.

L'identité (4.7) est une conséquence immédiate de (4.5) et (4.6). □

✎ En fait, tout cela s'interprète très facilement. Formellement, (4.6) implique que $P_t f = \exp(tA)f$. Cette écriture a un sens bien connu si A est une matrice. Le fait que A puisse être un opérateur (une matrice de « taille infinie ») oblige à quelques circonvolutions mathématiques, mais l'essentiel est là. Il est, par exemple, évident sur cette écriture que A et P_t commutent (voir (4.6)) et que $A^{-1}f = \int_0^\infty P_t f \, dt$ (penser à ce que vaut $\int_0^\infty \exp(at) \, dt$ quand a est réel négatif).

THÉORÈME 4.6.— Soit X un processus à valeurs dans E au plus dénombrable, à trajectoires cadlag et satisfaisant les propriétés (4.2) et (4.3). Le couple (ν, A) détermine complètement la loi de X . En particulier, pour tout entier $n \geq 1$, pour toutes

$f_j, j = 1, \dots, n$, fonctions bornées définies sur E et pour tout $t_1 < \dots < t_n$, on a :

$$\begin{aligned} & \mathbf{E} \left[\prod_{j=1}^n f_j(X(t_j)) \right] \\ &= \int_E P_{t_1} (f_1 P_{t_2-t_1} (f_2 \dots P_{t_{n-1}-t_{n-2}} (f_{n-1} P_{t_n-t_{n-1}} f_n)) \dots) (x) d\nu(x). \end{aligned} \quad (4.8)$$

Démonstration. En vertu du théorème 4.5, connaître P est équivalent à connaître A . Tenons pour acquise l'identité (4.8), en l'appliquant à $f_j = \mathbf{1}_{A_j}$ où A_j est un sous-ensemble quelconque de E , on voit que le terme de gauche est :

$$\mathbf{P}(X(t_1) \in A_1, \dots, X(t_n) \in A_n).$$

Les lois fini-dimensionnelles de X sont donc complètement caractérisées par (ν, A) . D'après le théorème de prolongement, on en déduit que la loi de X est bien complètement caractérisée par (ν, A) . Reste à démontrer (4.8). Pour $n = 1$, c'est la définition de P . Supposons que le résultat soit vrai pour n . Par conditionnement et d'après la définition de P , on a :

$$\begin{aligned} & \mathbf{E} \left[\prod_{j=1}^{n+1} f_j(X(t_j)) \right] \\ &= \mathbf{E} \left[\prod_{j=1}^n f_j(X(t_j)) \mathbf{E}[f_{n+1}(X(t_{n+1})) | \mathcal{F}_{t_n}] \right] \\ &= \mathbf{E} [f_1(X(t_1)) \dots f_{n-1}(X(t_{n-1})) (f_n P_{t_{n+1}-t_n} f_{n+1})(X(t_n))] . \end{aligned}$$

Comme $f_n P_{t_{n+1}-t_n} f_{n+1}$ est bornée, le résultat vient par récurrence. \square

THÉORÈME 4.7.— Soit X un processus à valeurs dans E au plus dénombrable, à trajectoires cadlag et satisfaisant les propriétés (4.2) et (4.3). Le processus X satisfait la propriété de Markov forte : pour toute fonction $F : D(\mathbf{R}^+, E) \rightarrow \mathbf{R}$ bornée, pour tout temps d'arrêt τ , on a :

$$\mathbf{E}[F(\theta_\tau X) | \mathcal{F}_\tau] = \mathbf{E}[F(X) | X(0) = X(\tau)]. \quad (4.9)$$

Démonstration. Supposons dans un premier temps que τ prenne ses valeurs dans $\mathfrak{T} = \{t_j, j \geq 1\}$ dénombrable. Les ensembles $\{\tau \leq t\}$ et $\{\tau > t\}$ appartiennent, par définition d'un temps d'arrêt, à \mathcal{F}_t , donc :

$$\{\tau = t\} = \left(\{\tau \leq t\} \cap \bigcap_{s < t, s \in \mathfrak{T}} \{\tau > s\} \right) \in \mathcal{F}_t.$$

On peut alors suivre la preuve de Markov fort (voir (2.6)) établie pour les chaînes de Markov (en remplaçant n par t_n), pour obtenir que pour toute fonction $f \in l^\infty(E)$:

$$\mathbf{E}[f(X(\tau + s)) | \mathcal{F}_\tau] = \mathbf{E}[f(X(\tau + s) | X(\tau))] = P_s f(X(\tau)). \quad (4.10)$$

Pour τ temps d'arrêt quelconque, on considère la suite de temps d'arrêt $\{\tau_n, n \geq 1\}$ définie par :

$$\tau_n = \sum_{k=0}^{\infty} \frac{k+1}{2^n} \mathbf{1}_{]k2^{-n}, (k+1)2^{-n}]}(\tau).$$

Cette suite converge en décroissant vers τ . Comme X a des trajectoires continues à droite $X(\tau_n + s)$ tend presque sûrement vers $X(\tau + s)$. Par conséquent, pour tout $A \in \mathcal{F}_\tau \subset \mathcal{F}_{\tau_n}$, par convergence dominée, on a :

$$\begin{aligned} \mathbf{E}[f(X(\tau + s)) \mathbf{1}_A] &= \lim_{n \rightarrow \infty} \mathbf{E}[f(X(\tau_n + s)) \mathbf{1}_A] \\ &= \lim_{n \rightarrow \infty} \mathbf{E}[P_s f(X(\tau_n)) \mathbf{1}_A] = \mathbf{E}[P_s f(X(\tau)) \mathbf{1}_A], \end{aligned}$$

donc (4.10) reste vraie pour τ quelconque.

Par conditionnements successifs, pour $0 < s_1 < \dots < s_k$, on a donc :

$$\begin{aligned} \mathbf{E} \left[\prod_{j=1}^k f_j(X(\tau + s_j)) | \mathcal{F}_\tau \right] &= (P_{s_k - s_{k-1}} f_k \dots P_{s_1} f_1)(X(\tau)) \\ &= \mathbf{E} \left[\prod_{j=1}^k f_j(X(\tau + s_j)) | X(\tau) \right]. \end{aligned}$$

Or, on a aussi (voir (4.8)) :

$$\mathbf{E} \left[\prod_{j=1}^k f_j(X(s_j)) | X(0) = x \right] = (P_{s_k - s_{k-1}} f_k \dots P_{s_1} f_1)(x).$$

Par conséquent :

$$\mathbf{E}[F \circ \theta_\tau | X(\tau) = x] = \mathbf{E}[F | X(0) = x],$$

pour les fonctions F de la forme $\prod_j f_j$. Par classe monotone, ce résultat reste vrai pour toutes les fonctions bornées de $D(\mathbf{R}^+, E)$ dans \mathbf{R} , d'où (4.9). \square

DÉFINITION 4.7.— Soit X un processus de Markov régulier de paramètres (ν, Q) . Pour $f \in l^\infty(E)$, on pose :

$$A_Q f(x) = q(x, x) \sum_{y \neq x} (f(y) - f(x)) q(x, y).$$

Si l'on identifie f au vecteur colonne $(f(x), x \in E)$ (après avoir ordonné les éléments de E ce qui est possible puisque E s'injecte dans \mathbb{N}), on peut réécrire l'identité précédente sous la forme d'un produit matriciel en introduisant la matrice, notée également A_Q et définie par :

$$A_Q(x, y) = \begin{cases} -q(x, x) & \text{si } x = y, \\ q(x, y)q(x, x) & \text{si } x \neq y. \end{cases} \quad (4.11)$$

On notera que $A_Q(x, y) = (A_Q \mathbf{1}_y)(x)$ pour tous x, y .

THÉORÈME 4.8.— Soit X un processus de Markov régulier de paramètres (ν, Q) . Le processus X satisfait la propriété de Markov simple (4.2) et d'homogénéité (4.3). Le semi-groupe associé est fortement continu. Son générateur infinitésimal est A_Q et son domaine est $l^\infty(E)$.

Démonstration. Partons de la deuxième construction des trajectoires de X . Dans ce cas, connaître \mathcal{F}_t revient à connaître en particulier le nombre de sauts de N avant t et la valeur de X après le dernier saut de N avant t . De par la construction même des trajectoires, ces deux quantités sont les deux seules utiles pour déterminer la suite de la trajectoire. On a donc :

$$\begin{aligned} \mathbf{E}[f(X(t+s)) | \mathcal{F}_t] &= \mathbf{E}\left[f(X(t+s)) | N(t), \tilde{X}_{N(t)}\right] \\ &= \sum_{n \geq 0} \mathbf{E}\left[f(X(t+s)) \mathbf{1}_{\{N(t+s)-N(t)=n\}} | N(t), \tilde{X}_{N(t)}\right] \\ &= \sum_{n \geq 0} \mathbf{E}\left[f(\tilde{X}_{N(t)+n}) \mathbf{1}_{\{N(t+s)-N(t)=n\}} | N(t), \tilde{X}_{N(t)}\right]. \end{aligned}$$

Comme \tilde{X} et N sont indépendants et N a des accroissements indépendants, on a :

$$\mathbf{E}[f(X(t+s)) | \mathcal{F}_t] = \sum_{n \geq 0} \tilde{Q}^{(n)} f(\tilde{X}_{N(t)}) e^{-\|q\|_\infty s} \frac{(\|q\|_\infty s)^n}{n!}.$$

Par ailleurs :

$$\begin{aligned}
\mathbf{E}[f(X(t+s)) | X(t)] &= \sum_{n \geq 0} \mathbf{E}[\mathbf{E}[f(X(t+s)) \mathbf{1}_{\{N(t+s)-N(t)=n\}} | N(t), X(t)] | X(t)] \\
&= \sum_{n \geq 0} \mathbf{E}[\mathbf{E}[f(\tilde{X}_{N(t)+n}) \mathbf{1}_{\{N(t+s)-N(t)=n\}} | N(t), \tilde{X}_{N(t)}] | X(t)] \\
&= \mathbf{E}\left[\sum_{n \geq 0} \tilde{Q}^{(n)} f(\tilde{X}_{N(t)}) e^{-\|q\|_\infty s} \frac{(\|q\|_\infty s)^n}{n!} | X(t)\right] \\
&= \sum_{n \geq 0} \tilde{Q}^{(n)} f(\tilde{X}_{N(t)}) e^{-\|q\|_\infty s} \frac{(\|q\|_\infty s)^n}{n!},
\end{aligned}$$

d'après la première partie de la preuve et puisque $X(t) = \tilde{X}_{N(t)}$. La propriété de Markov simple est donc bien vérifiée. On déduit aussi de la dernière équation que :

$$\mathbf{E}[f(X(t+s)) | X(t) = x] = \sum_{n \geq 0} \tilde{Q}^{(n)} f(x) e^{-\|q\|_\infty s} \frac{(\|q\|_\infty s)^n}{n!}.$$

Le terme de droite ne dépend pas de t donc la propriété d'homogénéité est aussi vérifiée. \square

THÉORÈME 4.9.— Soit X un processus à valeurs dans E au plus dénombrable, satisfaisant (4.2) et (4.3). Soit ν la loi de $X(0)$ et A son générateur infinitésimal. Le processus X est un processus de Markov de paramètres (ν, Q_A) où :

$$Q_A(x, y) = \begin{cases} |A(x, x)| & \text{si } y = x, \\ \frac{A(x, y)}{|A(x, x)|} & \text{si } y \neq x. \end{cases} \quad (4.12)$$

Démonstration. Posons $T_0 = 0$ et pour tout entier n :

$$\begin{cases} T_{n+1} &= \inf\{t > T_n, X(t) \neq X(T_n)\} \\ \xi_n &= T_{n+1} - T_n, \\ \hat{X}_n &= X(T_n), \end{cases}$$

avec la convention habituelle : $\inf \emptyset = \infty$.

Soit pour tout $u \geq 0$, $g(u) = \mathbf{P}(T_1 > u | X_0 = x)$. Montrons que g est solution de l'équation fonctionnelle caractéristique de la fonction exponentielle. D'après les

propriétés (4.2) et (4.3), on a :

$$\begin{aligned}
 g(u+v) &= \mathbf{P}(T_1 > u+v \mid X_0 = x) \\
 &= \mathbf{E} \left[\mathbf{1}_{\{X(s)=x, s \in [0, u]\}} \mathbf{1}_{\{X(t)=x, t \in [u, u+v]\}} \right] \\
 &= \mathbf{E} \left[\mathbf{1}_{\{X(s)=x, s \in [0, u]\}} \mathbf{E} \left[\mathbf{1}_{\{X(t)=x, t \in [u, u+v]\}} \mid \mathcal{F}_u \right] \right] \\
 &= \mathbf{E} \left[\mathbf{1}_{\{X(s)=x, s \in [0, u]\}} \mathbf{E} \left[\mathbf{1}_{\{X(t)=x, t \in [0, v]\}} \mid X_0 = x \right] \right] \\
 &= \mathbf{E} \left[\mathbf{1}_{\{X(s)=x, s \in [0, u]\}} \right] g(v) \\
 &= g(u)g(v).
 \end{aligned}$$

Comme g est borné, on en déduit l'existence de $q(x) \geq 0$ tel que :

$$g(u) = \exp(-q(x)u).$$

Le temps de séjour dans l'état initial suit donc une loi exponentielle.

Par définition d'un temps d'arrêt, l'événement $\{T_1 > u\}$ appartient à \mathcal{F}_u . De plus, sur $\{T_1 > u\}$, $X(u) = X(0)$ donc :

$$\begin{aligned}
 \mathbf{P}(\hat{X}_1 = y, T_1 > u \mid X(0) = x) &= \mathbf{E}_x \left[\mathbf{1}_{[u, \infty[}(T_1) \mathbf{E}_x \left[\mathbf{1}_{\{y\}}(\hat{X}_1) \mid \mathcal{F}_u \right] \right] \\
 &= \mathbf{E}_x \left[\mathbf{1}_{[u, \infty[}(T_1) \mathbf{E}_x \left[\mathbf{1}_{\{y\}}(\hat{X}_1) \mid X_0 = x \right] \right].
 \end{aligned}$$

Comme la quantité $\mathbf{P}(\hat{X}_1 = y \mid X_0 = x)$ est déterministe, elle sort de l'espérance. Par conséquent, si l'on pose :

$$\mathbf{P}(\hat{X}_1 = y \mid X(0) = x) = q(x, y),$$

on peut écrire :

$$\mathbf{P}(\hat{X}_1 = y, T_1 > u \mid X(0) = x) = q(x, y) \exp(-(q(x)u)). \quad (4.13)$$

On déduit de (4.13) que conditionnellement à $X(0) = \hat{X}_0$, \hat{X}_1 et T_1 sont indépendantes. On a donc bien prouvé ce que l'on voulait jusqu'au premier saut : temps de séjour de loi exponentielle dépendant de l'état initial, puis choix du nouvel état de façon indépendante de ce temps de séjour.

Supposons prouvé que pour $j \leq n-1$, pour tout $y \in E$, tout $u \geq 0$, on a l'identité :

$$\begin{aligned}
 \mathbf{P}(\hat{X}_{j+1} = y, \xi_j > u \mid \hat{X}_0, \dots, \hat{X}_j, T_1, \dots, T_j) \\
 = q(\hat{X}_j, y) \exp(-q(\hat{X}_j, \hat{X}_j)u). \quad (4.14)
 \end{aligned}$$

Etant donné que la connaissance de $\hat{X}_0, \dots, \hat{X}_n, T_1, \dots, T_n$ est équivalente à celle de tout le passé du processus jusqu'à l'instant T_n , nous avons :

$$\begin{aligned} \mathbf{P}(\hat{X}_{n+1} = y, \xi_n > u \mid \hat{X}_0, \dots, \hat{X}_n, T_1, \dots, T_n) \\ = \mathbf{P}(\hat{X}_{n+1} = y, \xi_n > u \mid \mathcal{F}_{T_n}). \end{aligned}$$

D'après la propriété forte de Markov (voir le théorème 4.7), on peut transformer cette quantité de la manière suivante :

$$\begin{aligned} \mathbf{P}(\hat{X}_{n+1} = y, \xi_n > u \mid \mathcal{F}_{T_n}) &= \mathbf{P}(\hat{X}_{n+1} = y, \xi_n > u \mid \hat{X}_n) \\ &= \mathbf{P}_{\hat{X}_n}(\hat{X}_1 = y, \xi_0 > u) \\ &= q(\hat{X}_n, y) \exp(-q(\hat{X}_n, \hat{X}_n)u), \end{aligned}$$

d'après (4.13). La relation (4.14) est donc vérifiée au rang n . Cette relation signifie que l'on peut construire les trajectoires de X comme dans la définition 4.1.

Par identification, (4.12) découle de (4.11) et du théorème 4.8. \square

EXEMPLE (File M/M/1).— En vertu de (4.11) et des considérations de l'exemple 4.1, le générateur infinitésimal du nombre de clients dans la file M/M/1 est donné pour tout $i \in \mathbf{N}$ par :

$$\begin{cases} A(i, i+1) &= \lambda, \\ A(i, i) &= -(\lambda + \mu \mathbf{1}_{[1, +\infty[}(i)), \\ A(i, i-1) &= \mu \text{ si } i > 0. \end{cases}$$

A cette écriture en ligne, on préfère souvent la représentation matricielle :

$$A = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & (0) \\ & & \ddots & & \\ (0) & \mu & -(\lambda + \mu) & \lambda & \\ & & & \ddots & \end{pmatrix},$$

ou la représentation graphique de la figure 4.3.

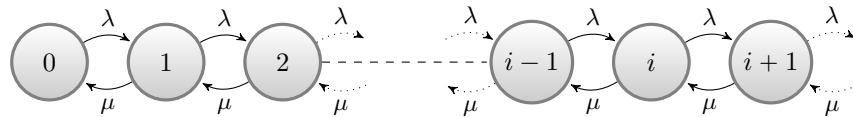


Figure 4.3 – Représentation graphique des taux de transition de la file M/M/1

THÉORÈME 4.10.— Soit X un processus de Markov régulier de paramètres (ν, Q) . On note \tilde{X} la chaîne incluse. Les propriétés suivantes sont équivalentes :

- \tilde{X} est irréductible;
- pour tout $x, y \in E$, il existe $t > 0$ tel que $p_t(x, y) > 0$;
- pour tout $x, y \in E$, pour tout $t > 0$, $p_t(x, y) > 0$.

Démonstration. En partant de la deuxième construction des trajectoires de X , la chaîne \tilde{X} est indépendante de N donc on a :

$$\begin{aligned}
 p_t(x, y) &= \mathbf{P}(X(t) = y \mid X(0) = x) \\
 &= \sum_{n=0}^{\infty} \mathbf{P}(X(t) = y, N(t) = n \mid X(0) = x) \\
 &= \sum_{n=0}^{\infty} \mathbf{P}(N(t) = n) \mathbf{P}(\tilde{X}_n = y \mid \tilde{X}_0 = x) \\
 &= \sum_{n=0}^{\infty} e^{-\|q\|_{\infty} t} \frac{(\|q\|_{\infty} t)^n}{n!} \tilde{Q}^{(n)}(x, y).
 \end{aligned}$$

D'après la définition de l'irréductibilité, l'équivalence est évidente. \square

DÉFINITION 4.8.— Une mesure μ sur E est dite invariante pour le processus de Markov X lorsque $X(0) \sim \mu$ implique que $X(t) \sim \mu$ pour tout $t \geq 0$.

THÉORÈME 4.11.— Soit X un processus de Markov régulier de paramètre (ν, Q) . Une mesure μ est invariante si et seulement si elle satisfait les équations :

$$\int Af \, d\mu = 0, \text{ pour tout } f \in l^{\infty}(E). \quad (4.15)$$

En notation matricielle, cela revient à : $\mu A = 0$, où μ est le vecteur ligne $(\mu(x), x \in E)$.

Démonstration. Dire que $X(t)$ et $X(0)$ ont même loi signifie que pour tout $f \in l^{\infty}(E)$, $\mathbf{E}[f(X(t))] = \mathbf{E}[f(X(0))]$. Or :

$$\mathbf{E}[f(X(t))] = \mathbf{E}[\mathbf{E}[f(X(t)) \mid X(0)]] = \mathbf{E}[P_t f(X(0))] = \int_E P_t f(x) \, d\mu(x).$$

Donc, μ est une mesure invariante si et seulement si :

$$\int_E P_t f(x) \, d\mu(x) = \int_E f(x) \, d\mu(x).$$

Par dérivation, on en déduit (4.15). Réciproquement, si (4.15) est vraie, d'après (4.7) et le théorème de Fubini :

$$\begin{aligned} \int_E (P_t f(x) - f(x)) \, d\mu(x) &= \int_E \int_0^t P_s A f(x) \, ds \, d\mu(x) \\ &= \int_0^t \left(\int_E P_s A f(x) \, d\mu(x) \right) \, ds \\ &= 0. \end{aligned}$$

Comme E est un espace discret, l'équation (4.15) s'écrit :

$$\sum_{x \in E} \left(\sum_{z \in E} A(x, z) f(z) \right) \mu(x) = 0.$$

Pour tout $y \in E$, en prenant $f = \mathbf{1}_y$, il vient donc :

$$\sum_{x \in E} \mu(x) A(x, y) = 0.$$

Dans le langage matriciel, cela signifie exactement que le produit du vecteur ligne μ par la matrice A est nul. \square

De manière analogue au cas discret, posons :

$$\tau_x^1 = \inf \{t > 0, X(t) = x\},$$

avec la convention $\tau_x^1 = \infty$ si $X(t) \neq x$ pour tout $t > 0$.

THÉORÈME 4.12.— Soit X un processus de Markov régulier irréductible récurrent. Il existe une unique mesure invariante à un facteur multiplicatif près. Cette mesure est proportionnelle à l'une des trois mesures suivantes :

(i) pour tout $y \in E$:

$$\mu(y) = \mathbf{E}_x \left[\int_0^{\tau_x^1} \mathbf{1}_{\{X(s)=y\}} \, ds \right], \quad (4.16)$$

où $x \in E$ est quelconque mais fixé ;

(ii) pour tout $y \in E$, $\mu(y) = \hat{\mu}(y)/q(y, y)$ où $\hat{\mu}$ est une mesure invariante de la chaîne incluse \tilde{X} ;

(iii) μ , solution de l'équation matricielle $\mu A = 0$.

Démonstration. Pour prouver le point (i), il faut vérifier que pour tout $t > 0$, pour tout $f \in l^\infty(E)$, l'identité suivante est satisfaite :

$$\int_E P_t f(x) \, d\mu(x) = \int_E f(x) \, d\mu(x). \quad (4.17)$$

Or, avec l'expression (4.16), on a :

$$\begin{aligned} \int_E P_t f(y) \, d\mu(y) &= \sum_{y \in E} \mathbf{E}_x \left[\int_0^{\tau_x^1} P_t f(y) \mathbf{1}_{\{X(s)=y\}} \, ds \right] \\ &= \sum_{y \in E} \mathbf{E}_x \left[\int_0^\infty P_t f(X(s)) \mathbf{1}_{\{X(s)=y\}} \mathbf{1}_{\{s < \tau_x^1\}} \, ds \right] \\ &= \mathbf{E}_x \left[\int_0^\infty \mathbf{E}[f(X(s+t)) \mid \mathcal{F}_s] \mathbf{1}_{\{s < \tau_x^1\}} \, ds \right] \\ &= \int_0^\infty \mathbf{E}_x [\mathbf{E}[f(X(s+t)) \mathbf{1}_{\{s < \tau_x^1\}} \mid \mathcal{F}_s]] \, ds \\ &= \int_0^\infty \mathbf{E}_x [f(X(s+t)) \mathbf{1}_{\{s < \tau_x^1\}}] \, ds \\ &= \mathbf{E}_x \left[\int_t^{\tau_x^1+t} f(X(s)) \, ds \right] \\ &= \mathbf{E}_x \left[\int_t^{\tau_x^1} f(X(s)) \, ds \right] + \mathbf{E}_x \left[\int_{\tau_x^1}^{\tau_x^1+t} f(X(s)) \, ds \right]. \end{aligned}$$

D'après la propriété de Markov et en remarquant que, par définition de τ_x^1 et par continuité à droite de X , on a $X(\tau_x^1) = x$, on obtient :

$$\begin{aligned} \int_E P_t f(y) \, d\mu(y) &= \mathbf{E}_x \left[\int_t^{\tau_x^1} f(X(s)) \, ds \right] + \mathbf{E}_x \left[\int_0^t f(X(s)) \, ds \right] \\ &= \mathbf{E}_x \left[\int_0^{\tau_x^1} f(X(s)) \, ds \right] \\ &= \sum_{y \in E} \mathbf{E}_x \left[\int_0^{\tau_x^1} f(X(s)) \mathbf{1}_{\{X(s)=y\}} \, ds \right] \\ &= \sum_{y \in E} \mathbf{E}_x \left[\int_0^{\tau_x^1} f(y) \mathbf{1}_{\{X(s)=y\}} \, ds \right] = \int_E f(y) \, d\mu(y). \end{aligned}$$

Pour le point (ii), rappelons qu'avec les notations du chapitre 2, on a :

$$\hat{\tau}_x^1 = \inf\{n > 0, \hat{X}_n = x\}.$$

En utilisant la première construction des trajectoires de X , on obtient :

$$\mathbf{E}_x \left[\int_0^{\tau_x^1} \mathbf{1}_{\{X(s)=y\}} \, ds \right] = \sum_{n=1}^{\infty} \mathbf{E}_x \left[\xi_n \mathbf{1}_y(\hat{X}_{n-1}) \mathbf{1}_{\{n \leq \hat{\tau}_x^1\}} \right].$$

L'événement $\{n \leq \hat{\tau}_x^1\}$ est le complémentaire de $\{\tau_x^1 < n\} = \{\tau_x^1 \leq n-1\}$, donc il est $\hat{\mathcal{F}}_{n-1}$ mesurable (avec des notations évidentes). Par construction :

$$\mathbf{E} \left[\xi_n \mid \hat{\mathcal{F}}_{n-1} \right] = \mathbf{E} \left[\xi_n \mid \hat{X}_{n-1} \right] = \frac{1}{q(\hat{X}_{n-1}, \hat{X}_{n-1})}.$$

On en déduit que :

$$\begin{aligned} \mathbf{E}_x \left[\xi_n \mathbf{1}_y(\hat{X}_{n-1}) \mathbf{1}_{\{n \leq \hat{\tau}_x^1\}} \right] &= \mathbf{E}_x \left[\mathbf{E} \left[\xi_n \mid \hat{\mathcal{F}}_{n-1} \right] \mathbf{1}_y(\hat{X}_{n-1}) \mathbf{1}_{\{n \leq \hat{\tau}_x^1\}} \right] \\ &= \frac{1}{q(y, y)} \mathbf{P}_x(\hat{X}_{n-1} = y, \hat{\tau}_x^1 \geq n). \end{aligned}$$

Pour $\hat{\mu}$ mesure stationnaire de \hat{X} , on obtient alors :

$$\mathbf{E}_x \left[\int_0^{\tau_x^1} \mathbf{1}_{\{X(s)=y\}} \, ds \right] = \frac{1}{q(y, y)} \sum_{n=1}^{\infty} \mathbf{P}_x(\hat{X}_{n-1} = y, \hat{\tau}_x^1 \geq n) = \frac{1}{q(y, y)} \hat{\mu}(y).$$

Le point (iii) a déjà été démontré dans le théorème 4.11. □

EXEMPLE (File M/M/1).— Pour cet exemple, la résolution du système :

$$(\pi(0), \pi(1), \dots, \pi(i), \dots) \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & (0) \\ & & \ddots & & \\ & (0) & \mu & -(\lambda + \mu) & \lambda \\ & & & \ddots & \end{pmatrix} = 0$$

équivalent aux équations :

$$\begin{aligned}
 -\lambda\pi(0) + \mu\pi(1) &= 0 \\
 \lambda\pi(0) - (\lambda + \mu)\pi(1) + \mu\pi(2) &= 0 \\
 &\vdots \\
 \lambda\pi(i-1) - (\lambda + \mu)\pi(i) + \mu\pi(i+1) &= 0 \\
 &\vdots
 \end{aligned}$$

En additionnant ces équations deux à deux successivement, on obtient :

$$-\lambda\pi(0) + \mu\pi(1) = 0, -\lambda\pi(1) + \mu\pi(2) = 0, \dots, -\lambda\pi(i) + \mu\pi(i+1) = 0,$$

soit, pour tout entier i :

$$\pi(i+1) = \rho\pi(i) \text{ avec } \rho = \lambda/\mu.$$

Par conséquent, $\pi(i) = \rho^i\pi(0)$. On sait de l'étude de la file M/GI/1 que la file M/M/1 n'est stable (le processus comptant le nombre de clients est récurrent) que si $\rho < 1$. On retrouve bien ici sur la forme particulière des mesures invariantes, que l'on ne peut construire une probabilité invariante que si $\rho < 1$. Dans ce cas, la probabilité invariante est donnée par $\pi(i) = \rho^i(1 - \rho)$. Cela correspond à une loi géométrique de paramètre ρ , décalée de 1.

Compte tenu de la propriété de Markov forte du théorème 4.12, la preuve du résultat suivant est rigoureusement similaire à celle de son analogue dans le cas discret, le théorème 2.21.

THÉORÈME 4.13.— Soit X un processus de Markov régulier irréductible récurrent. On note π l'unique probabilité invariante. Pour tout $f \in L^1(\pi)$, on a :

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X(s)) \, ds = \sum_{x \in E} f(x) \pi(x).$$

REMARQUE.— Notons que ce théorème permet d'approcher π . En effet, il suffit de mener la simulation d'une trajectoire, en suivant par exemple la construction trajectorielle du début de ce chapitre et de calculer la proportion du temps passé par le processus dans chaque état. Cela donnera pour un temps long une approximation de π qui permettra ensuite de calculer les espérances de fonctions plus générales à l'état stationnaire. Il reste à déterminer ce que « long » signifie, donc d'identifier la vitesse de cette convergence. Cela dépasse le cadre de cet ouvrage, mais n'en est pas moins un sujet particulièrement actuel de recherche.

THÉORÈME 4.14.— Soit X un processus de Markov régulier irréductible. Si X est transient alors :

$$p_t(x, x) \xrightarrow[t \rightarrow \infty]{} 0, \text{ pour tout } x \in E.$$

Si X est récurrent de probabilité invariante π :

$$p_t(x, y) \xrightarrow[t \rightarrow \infty]{} \pi(y) \text{ et } \mathbf{E}_x [\tau_x^1] = \frac{1}{\pi(x)},$$

pour tout $x \in E$.

Démonstration. Nous ne développons la preuve que dans le cas récurrent, le cas transient se traite de manière analogue. Soit $h > 0$ et $X_n^h = X(nh)$. D'après le théorème A.12 :

$$\begin{aligned} \mathbf{E} [f(X_n^h) | X_j^h, j = 0, \dots, n-1] &= \mathbf{E} [f(X_n^h) | X_{n-1}^h] \\ &= \mathbf{E} [f(X(nh)) | X((n-1)h)] = P_h f(X((n-1)h)). \end{aligned}$$

La suite X^h est donc une chaîne de Markov d'opérateur de transition P_h . En vertu du théorème 4.10, X^h est irréductible et d'après (4.17), π en est une probabilité invariante. D'après le théorème 2.25 :

$$p_{nh}(x, y) \xrightarrow[n \rightarrow \infty]{} \pi(y).$$

Comme l'on a :

$$\begin{aligned} |p_t(x, y) - p_s(x, y)| &= \left| \int_s^t A p_u(x, y) \, du \right| \\ &\leq \int_s^t |A p_u(x, y)| \, du \\ &\leq \|q\|_\infty (t - s), \end{aligned}$$

on en déduit que :

$$|p_t(x, y) - \pi(y)| \leq \|q\|_\infty |t - nh| + |p_{nh}(x, y) - \pi(y)|.$$

Soit $\epsilon > 0$, on choisit h tel que $h\|q\|_\infty < \epsilon$ et n_0 tel que :

$$n \geq n_0 \implies |p_{nh}(x, y) - \pi(y)| \leq \epsilon.$$

Si $t \geq n_0 h$, il existe n tel que $|t - nh| \leq h$ et alors $|p_t(x, y) - \pi(y)| \leq 2\epsilon$. Donc, $\lim_{t \rightarrow \infty} p_t(x, y) = \pi(y)$.

On sait que la probabilité invariante est proportionnelle à la mesure définie dans (4.16). Or :

$$\sum_{y \in E} \mathbf{E}_x \left[\int_0^{\tau_x^1} \mathbf{1}_{\{y\}}(X(s)) \, ds \right] = \mathbf{E}_x [\tau_x^1],$$

donc en appliquant (4.16) avec $y = x$, on obtient :

$$\pi(x) = \frac{1}{\mathbf{E}_x [\tau_x^1]},$$

d'où le résultat. □

4.4. Problème de martingale

DÉFINITION 4.9.— Soit ν une mesure de probabilité sur E dénombrable et A un opérateur continu de $l^\infty(E)$ dans lui-même, c'est-à-dire qu'il existe $c > 0$ tel que pour tout $f \in l^\infty(E)$, l'identité suivante soit satisfaite :

$$\|Af\|_\infty = \sup_{x \in E} |Af(x)| \leq c \|f\|_\infty.$$

On note alors :

$$\|A\|_\infty = \sup_{\|f\|_\infty=1} \|Af\|_\infty.$$

THÉORÈME 4.15.— Le processus X est un processus de Markov régulier de loi initiale ν et de générateur infinitésimal A , si et seulement si la loi de $X(0)$ est ν et pour toute fonction bornée f , le processus :

$$M^f : t \mapsto f(X(t)) - f(X(0)) - \int_0^t Af(X(s)) \, ds$$

est une martingale locale. Pour tout t , son crochet s'écrit alors :

$$\langle M^f, M^f \rangle_t = \int_0^t (Af^2(X(s)) - 2f(X(s))Af(X(s))) \, ds. \quad (4.18)$$

Démonstration. Soit $f \in l^\infty(E)$, supposons dans un premier temps que les processus :

$$t \mapsto \int_0^t Af(X(s)) \, ds \text{ et } t \mapsto \int_0^t Af^2(X(s)) \, ds$$

soient bornés. On remarque que le processus $(s \mapsto Af(X(s)))$ est adapté. Pour tout $s \geq 0$, on a :

$$\begin{aligned}
& \mathbf{E}[M^f(t+s) | \mathcal{F}_t] \\
&= \mathbf{E}[f(X(t+s)) | \mathcal{F}_t] - \int_0^{t+s} \mathbf{E}[Af(X(u)) | \mathcal{F}_t] \, du \\
&= \mathbf{E}[f(X(t+s)) | X(t)] - \int_0^t Af(X(u)) \, du - \int_t^{t+s} \mathbf{E}[Af(X(u)) | \mathcal{F}_t] \, du \\
&= P_s f(X(t)) - \int_0^t Af(X(u)) \, du - \int_0^s P_u Af(X(t)) \, du.
\end{aligned}$$

D'après (4.5) :

$$P_s f(X(t)) - \int_0^s P_u Af(X(t)) \, du = f(X(t)),$$

donc :

$$\mathbf{E}[M^f(t+s) | \mathcal{F}_t] = f(X(t)) - \int_0^t Af(X(u)) \, du = M^f(t),$$

ce qui signifie que M^f est une martingale.

Pour $f \in l^\infty(E)$, le processus $(t \mapsto f^2(X(t)))$ est intégrable. D'une part, d'après la première partie de la preuve on peut écrire :

$$f^2(X(t)) - f^2(X(0)) = \int_0^t Af^2(X(s)) \, ds + M^{f^2}(t). \quad (4.19)$$

D'autre part, la formule d'Itô (voir théorème A.23) pour les martingales cadlag à variation finie, donne :

$$\begin{aligned}
f^2(X(t)) - f^2(X(0)) &= 2 \int_0^t f(X(s^-)) \, dM^f(s) \\
&\quad + 2 \int_0^t f(X(s^-)) Af(X(s)) \, ds + [f \circ X, f \circ X]_t. \quad (4.20)
\end{aligned}$$

En comparant (4.19) et (4.20), par définition du crochet, on en déduit que :

$$[f \circ X, f \circ X]_t = \int_0^t Af^2(X(s)) \, ds - 2 \int_0^t f(X(s^-)) Af(X(s)) \, ds.$$

Comme ce processus est continu et adapté, il est prévisible, d'où (4.18). Noter que l'on peut remplacer $f(X(s^-))$ par $f(X(s))$ dans la deuxième intégrale car il n'y a pas de saut fixe en s donc ds presque sûrement $X(s^-) = X(s)$, voir la remarque 4.2.

Traisons maintenant le cas général. Pour tout entier n , posons :

$$\tau_n = \inf \left\{ t \geq 0, \int_0^t |Af(X(s))| \, ds > n \text{ ou } \int_0^t |Af^2(X(s))| \, ds > n \right\},$$

et $M_{\tau_n}^f(t) = M^f(\tau_n \wedge t)$. Notons par construction que $|M_{\tau_n}^f(t)| \leq 2\|f\|_\infty + n$ donc $M_{\tau_n}^f$ est uniformément intégrable. Par ailleurs, comme A est continu de $l^\infty(E)$ dans lui-même, τ_n tend vers l'infini donc $\{\tau_n, n \geq 1\}$ réduit M^f . On remarque enfin que pour tout temps d'arrêt τ :

$$\begin{aligned} M_\tau^f(t) &= f(X(t \wedge \tau)) - f(X(0)) - \int_0^{t \wedge \tau} Af(X(s)) \, ds \\ &= f(X^\tau(t)) - f(X^\tau(0)) - \int_0^{t \wedge \tau} Af(X^\tau(s)) \, ds, \end{aligned}$$

où $X^\tau(s) = X(\tau \wedge s)$. Par conséquent, on peut appliquer ce qui précède à $M_{\tau_n}^f$ et X^{τ_n} . Il en résulte que $M_{\tau_n}^f$ est une martingale de carré intégrable dont le crochet est donné par (4.18).

Pour la réciproque, admettons le résultat du lemme 4.16 ci-après. On va montrer que pour toute $f \in l^\infty(E)$, pour tout $t, s \geq 0$, on a :

$$\mathbf{E}[f(X(t+s)) | \mathcal{F}_t] = \lim_{n \rightarrow \infty} (\text{Id} - n^{-1}A)^{-[ns]} f(X(t)). \quad (4.21)$$

L'espérance conditionnelle sachant \mathcal{F}_t apparaît donc comme dépendant uniquement de $X(t)$ et de s d'où la propriété de Markov simple et l'homogénéité. Soit A' le générateur infinitésimal de X . D'après la première partie de la preuve :

$$t \mapsto \int_0^t (A - A')f(X(s)) \, ds$$

est une martingale nulle en 0. Qui plus est, ce processus est continu et à variation finie. Cela implique que ce processus est nul $\mathbf{P} \otimes dt$ -presque sûrement. Par ailleurs :

$$(A - A')f(x) = \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t (A - A')f(X(s)) \, ds = 0,$$

donc $A = A'$. Finalement, comme A est continu de $l^\infty(E)$ dans lui-même et :

$$\sup_{x \in E} |a(x, x)| = \sup_{x \in E} |A \mathbf{1}_{\{x\}}(x)| \leq c,$$

on en déduit que X est un processus de Markov régulier.

Montrons maintenant (4.21). Quitte à localiser X , on peut toujours supposer que $E^{\lambda, f}$ est une martingale uniformément intégrable. On utilise (4.22) pour $\lambda = n$, on en déduit :

$$f(X(t)) = \mathbf{E} \left[\int_0^\infty e^{-ns} ((n \text{Id} - A)f)(X(t+s)) \, ds \mid \mathcal{F}_t \right].$$

On peut appliquer ce résultat à $(\text{Id} - \frac{1}{n}A)^{-1}f$:

$$\begin{aligned} \left(\text{Id} - \frac{1}{n}A\right)^{-1} f(X(t)) &= n\mathbf{E} \left[\int_0^\infty e^{-ns} f(X(t+s)) \, ds \mid \mathcal{F}_t \right] \\ &= \mathbf{E} \left[\int_0^\infty e^{-s} f(X(t+s/n)) \, ds \mid \mathcal{F}_t \right]. \end{aligned}$$

Par récurrence, pour tout entier k , il vient donc :

$$\begin{aligned} &\left(\text{Id} - \frac{1}{n}A\right)^{-k} f(X(t)) \\ &= \mathbf{E} \left[\int_0^\infty \dots \int_0^\infty e^{-(s_1+\dots+s_k)} f(X(t+n^{-1}(s_1+\dots+s_k))) \, ds_1 \dots ds_k \mid \mathcal{F}_t \right]. \end{aligned}$$

Soit $\{Z_k, k \geq 1\}$ une suite de variables aléatoires indépendantes, indépendantes de X , de loi exponentielle de paramètre 1. Pour tout réel u non nul, la loi forte des grands nombres indique que :

$$u \frac{1}{[nu]} \sum_{j=1}^{[nu]} Z_j \xrightarrow{n \rightarrow \infty} u\mathbf{E}[Z_1] = u.$$

On a alors :

$$\begin{aligned} &\int_0^\infty \dots \int_0^\infty e^{-(s_1+\dots+s_{[nu]})} f(X(t+n^{-1}(s_1+\dots+s_k))) \, ds_1 \dots ds_k \\ &= \mathbf{E} \left[f \left(X \left(t + u(nu)^{-1} \sum_{j=1}^{[nu]} Z_j \right) \right) \right] \xrightarrow{n \rightarrow \infty} f(X(t+u)). \end{aligned}$$

Par convergence dominée, on obtient (4.21). \square

REMARQUE.— Le calcul du crochet est intéressant en lui-même car il est à la base de nombreuses méthodes d'approximations de processus, en particulier les limites fluides (ou approximations champ-moyen) et les approximations diffusions.

LEMME 4.16.— Soit X un processus cadlag. Pour toute $f \in l^\infty(E)$, le processus :

$$M^f : t \mapsto f(X(t)) - f(X(0)) - \int_0^t Af(X(s)) \, ds$$

est une martingale locale si et seulement si pour tout $\lambda \in \mathbf{R}$, le processus :

$$E^{\lambda, f} : t \mapsto e^{-\lambda t} f(X(t)) + \int_0^t e^{-\lambda s} (\lambda f(X(s)) - Af(X(s))) \, ds$$

est une martingale locale. En particulier, si $\{\tau_n, n \geq 1\}$ réduit $E^{\lambda, f}$, on obtient l'identité suivante :

$$f(X^{\tau_n}(t)) = \mathbf{E} \left[\int_0^\infty e^{-\lambda s} (\lambda f(X^{\tau_n}(t+s)) - Af(X^{\tau_n}(t+s))) \, ds \mid \mathcal{F}_t \right]. \quad (4.22)$$

Démonstration. Quitte à localiser, on peut toujours supposer que les processus $E^{\lambda, f}$ et M^f sont uniformément intégrables. On pose pour tout $t \geq 0$, $U(t) = \exp(-\lambda t)$. Ce processus est continu, à variation bornée, d'après la formule d'intégration par parties, pour tout t , on a :

$$U(t)M^f(t) = \int_0^t U(s) \, dM^f(s) - \int_0^t M^f(s) \lambda e^{-\lambda s} \, ds. \quad (4.23)$$

Par hypothèse, M^f est une martingale donc par construction de l'intégrale stochastique, $\int U \, dM^f$ aussi. Notons maintenant que :

$$\begin{aligned} \int_0^t \int_0^s Af(X(u)) \, du \lambda e^{-\lambda s} \, ds &= \int_0^t Af(X(u)) \int_u^t \lambda e^{-\lambda s} \, ds \, du \\ &= \int_0^t Af(X(u)) (e^{-\lambda u} - e^{-\lambda t}) \, du. \end{aligned} \quad (4.24)$$

En reportant (4.24) dans (4.23), on voit que le processus :

$$\begin{aligned} t \mapsto U(t)M^f(t) + \int_0^t M^f(s) \lambda e^{-\lambda s} \, ds \\ = e^{-\lambda t} f(X(t)) + \int_0^t f(X(s)) \lambda e^{-\lambda s} \, ds - \int_0^t Af(X(s)) e^{-\lambda s} \, ds \end{aligned}$$

est une martingale, d'où le résultat. \square

4.5. Réversibilité et applications

Dans cette section, $(X(t), t \geq 0)$ désigne un processus de Markov homogène, régulier, irréductible, récurrent à trajectoires cadlag sur l'espace dénombrable E , d'opérateur de transition Q et de générateur infinitésimal A . Dans la suite, on dira que $(X(t), t \geq 0)$ est stationnaire s'il admet une probabilité invariante π et si la loi de $X(0)$ est π . On peut toujours supposer que $\pi(x) > 0$ pour tout $x \in E$.

DÉFINITION 4.10.— On suppose que $(X(t), t \geq 0)$ est stationnaire. Pour tout $T > 0$, le processus retourné de $(X(t), t \geq 0)$ à partir de T est le processus $(\bar{X}^T(t), t \in [0, T])$ défini pour tout $t \in [0, T]$ par :

$$\bar{X}^T(t) = X((T-t)^-) = \lim_{s \searrow t} X((T-s)).$$

Rappelons (voir la définition A.5 pour plus de détails) que $l^2(\mathbf{N}, \pi)$ est l'espace de Hilbert des suites de carré intégrable pour la mesure π .

LEMME 4.17.— Si A est le générateur infinitésimal d'un processus de Markov régulier X alors A est continu de $l^2(\mathbf{N}, \pi)$ dans lui-même. Il admet donc un adjoint \bar{A} dans $l^2(\mathbf{N}, \pi)$ défini par :

$$\bar{A}(x, y) = A(y, x) \frac{\pi(y)}{\pi(x)}.$$

Démonstration. Rappelons que $A(x, x) = -\sum_{y \neq x} A(x, y) < 0$ donc on a :

$$\sum_{y \in E} |A(x, y)| = 2|A(x, x)|.$$

Par conséquent, la mesure ν définie par :

$$\nu(y) = \frac{1}{2|A(x, x)|} |A(x, y)|, \text{ pour tout } y \in E,$$

est une mesure de probabilité sur E . Par conséquent :

$$\begin{aligned} \left(\sum_{y \in E} A(x, y) f(y) \right)^2 &\leq 4A(x, x)^2 \left(\sum_{y \in E} \frac{|A(x, y)|}{2|A(x, x)|} f(y) \right)^2 \\ &\leq 2|A(x, x)| \sum_{y \in E} |A(x, y)| |f(y)|^2 \\ &\leq 2\|A\|_\infty \sum_{y \in E} |A(x, y)| |f(y)|^2, \end{aligned}$$

d'après l'inégalité de Jensen et la régularité de A . Par conséquent :

$$\begin{aligned} \sum_{x \in E} (AF)(x)^2 \pi(x) &\leq 2\|A\|_\infty \sum_{x \in E} \sum_{y \in E} |A(x, y)| |f(y)|^2 \pi(x) \\ &= 2\|A\|_\infty \sum_{y \in E} (1 + |A(y, y)|) \pi(y) f(y)^2 \\ &\leq 2\|A\|_\infty (1 + \|A\|_\infty) \|f\|_{l^2(\pi)}^2. \end{aligned}$$

Donc, A est continu de $l^2(\mathbf{N}, \pi)$ dans lui-même. Par conséquent, l'adjoint \bar{A} de A , défini pour tous $u, v \in l^2(\mathbf{N}, \pi)$ par :

$$\langle Au, v \rangle_{l^2(\mathbf{N}, \pi)} = \langle u, Av \rangle_{l^2(\mathbf{N}, \pi)} \quad (4.25)$$

est aussi continu de $l^2(\mathbf{N}, \pi)$ dans lui-même. De plus :

$$\bar{A}(x, y) = A(y, x) \frac{\pi(y)}{\pi(x)}.$$

En effet : $\bar{A}(x, y) = (\bar{A} \mathbf{1}_{\{x\}})(y)$ donc pour $u = \mathbf{1}_{\{y\}}$, (4.25) donne :

$$(A \mathbf{1}_{\{y\}})(x)\pi(x) = (\bar{A} \mathbf{1}_{\{x\}})(y)\pi(y), \text{ soit } A(x, y)\pi(x) = \bar{A}(y, x)\pi(y),$$

d'où le résultat. \square

THÉORÈME 4.18.— Sous les hypothèses de cette section, pour tout $T > 0$ le processus $(\bar{X}^T(t), t \in [0, T])$ est markovien et à trajectoires cadlag, de générateur infinitésimal \bar{A} .

Démonstration. Les trajectoires du processus $(\bar{X}^T(t), t \geq 0)$ sont presque sûrement continues à droite, car par existence presque sûre d'une limite à gauche en tout point pour X , on a presque sûrement :

$$\begin{aligned} \lim_{h \searrow 0} \bar{X}^T(t+h) &= \lim_{h \searrow 0} \lim_{h' \searrow 0} X(T - (t+h) - h') \\ &= \lim_{\epsilon \searrow 0} X(T - (t+\epsilon)) = X(T-t)^- = \bar{X}^T(t), \end{aligned}$$

et l'existence d'une limite à gauche se démontre de manière similaire. On prend donc la limite à gauche en $T-t$ pour rendre cadlag le processus retourné.

Compte tenu du théorème 4.15, pour tout $f \in l^\infty(E)$, nous devons prouver que :

$$t \mapsto f(\bar{X}^T(t)) - f(\bar{X}^T(0)) - \int_0^t \bar{A}f(\bar{X}^T(r)) \, dr$$

est une martingale locale. La filtration est bien évidemment celle engendrée par les trajectoires de \bar{X}^T et non celles de X , c'est-à-dire que l'on a $\bar{\mathcal{F}}_t = \sigma\{\bar{X}^T(s), s \leq t\}$. En utilisant un raisonnement de classe monotone, il faut et il suffit que l'on prouve que pour $0 \leq s_1 < \dots < s_n \leq s < t$, pour toute fonction $\varphi : E^n \rightarrow \mathbf{R}$ bornée, on a :

$$\begin{aligned} \mathbf{E} \left[\left(f(\bar{X}^T(t)) - f(\bar{X}^T(s)) - \int_s^t \bar{A}f(\bar{X}^T(r)) \, dr \right) \varphi(\bar{X}^T(s_1), \dots, \bar{X}^T(s_n)) \right] \\ = 0. \end{aligned}$$

Comme X est markovien, on peut réécrire cette identité sous la forme :

$$\begin{aligned} 0 = \mathbf{E} \left[\left(f(X(T-t)) - f(X(T-s)) - \int_s^t (\bar{A}f)(X(T-r)) \, dr \right) \right. \\ \left. \times \mathbf{E} [\varphi(\bar{X}^T(s_1), \dots, \bar{X}^T(s_n)) | X(T-s)] \right]. \end{aligned}$$

D'après le théorème A.7, il suffit que pour ψ bornée, l'on démontre :

$$\mathbf{E} \left[\left(f(X(T-t)) - f(X(T-s)) - \int_{T-t}^{T-s} (\bar{A}f)(X(r)) \, dr \right) \psi(X(T-s)) \right] = 0. \quad (4.26)$$

Or, en introduisant le semi-groupe associé à X et en se rappelant que par stationnarité, la loi de $X(T-s)$ est celle de $X(0)$, c'est-à-dire π , on obtient d'une part :

$$\begin{aligned} \mathbf{E} [(f(X(T-t)) - f(X(T-s))) \psi(X(T-s))] \\ = \int_E \psi P_{t-s} f \, d\pi - \int_E f \psi \, d\pi \end{aligned} \quad (4.27)$$

et d'autre part, d'après (4.25) et (4.5), on a :

$$\begin{aligned} \mathbf{E} \left[\int_s^t \bar{A}f(X(T-r)) \, dr \, \psi(X(T-s)) \right] &= \int_E \int_s^t P_{r-s} \psi \bar{A}f \, dr \, d\pi \\ &= \int_0^{t-s} \int_E P_r \psi \bar{A}f \, d\pi \, dr \\ &= \int_0^{t-s} \int_E A P_r \psi f \, d\pi \, dr \\ &= \int_E \int_0^{t-s} A P_r \psi \, dr \, f \, d\pi \\ &= \int_E (P_{t-s} \psi - \psi) f \, d\pi. \end{aligned} \quad (4.28)$$

En faisant la différence de (4.27) et (4.28), on obtient (4.26). \square

Remarquons en particulier que le générateur de $(\bar{X}^T(t), t \geq 0)$ ne dépend pas de T , ce qui va s'avérer très précieux dans la construction suivante.

DÉFINITION 4.11.— On dit que le processus $(X(t), t \geq 0)$ est réversible, s'il est stationnaire et pour tout $T > 0$, de même loi que son processus retourné sur $[0, T]$.

Pour tout $T > 0$, $X(0)$ et $\bar{X}^T(0)$ ont même loi π . Comme la loi d'un processus de Markov est entièrement caractérisée par son générateur et sa loi initiale, on déduit du théorème 4.18 que $(X(t), t \geq 0)$ est réversible si et seulement si pour tout $x, y \in E$, on a :

$$\pi(x)A(x, y) = \pi(y)A(y, x). \quad (4.29)$$

Cette équation, appelée de balance locale, est une propriété « miroir » : le flux de passage de x à y est le même que celui de y à x .

THÉORÈME 4.19.— Réciproquement, pour tout processus markovien X , s'il existe une probabilité π vérifiant (4.29) alors X admet π pour probabilité stationnaire.

Démonstration. Il suffit pour cela de vérifier que $\pi A = 0$. Or, pour tout $x \in E$, on a :

$$\pi A(x) = \sum_{y \in E} \pi(y) A(y, x) = \sum_{y \in E} \pi(x) A(x, y) = \pi(x) \sum_{y \in E} A(x, y) = 0.$$

Qui plus est, X est un processus de Markov réversible. \square

Une classe importante de processus markoviens, dans laquelle on pourra ranger pratiquement tous les processus rencontrés dans la suite, a automatiquement la propriété de réversibilité dès qu'il existe une probabilité invariante : ce sont les processus de naissance et de mort.

DÉFINITION 4.12.— Un processus markovien $(X(t), t \geq 0)$ homogène à valeurs dans E , où $E = \mathbb{N}$ ou $\llbracket 0, n \rrbracket$, est dit de naissance et de mort si son générateur infinitésimal est tridiagonal sur E , c'est-à-dire que pour tout $i \in E$, on a :

$$A(i, j) = 0 \text{ pour tout } j \text{ tel que } |j - i| \geq 2.$$

Cette terminologie est héritée de la dynamique des populations. Si $(X(t), t \geq 0)$ compte la taille d'une population donnée à chaque instant, les sauts de $(X(t), t \geq 0)$ ne se produisent qu'au moment d'une naissance (de i à $i + 1$) ou d'une mort (de i à $i - 1$).

THÉORÈME 4.20.— Tout processus de naissance et mort stationnaire de loi π est réversible.

Démonstration. Comme A est tridiagonale, il suffit de vérifier que :

$$\pi(i - 1) A(i - 1, i) = \pi(i) A(i, i - 1)$$

pour tout $i \in E$, ce que nous faisons par récurrence. Tout d'abord, on a :

$$\pi(0) A(0, 1) = -\pi(0) A(0, 0) = \pi(1) A(1, 0),$$

puisque $\pi A(0) = 0$. Ensuite, si $\pi(i-1)A(i-1, i) = \pi(i)A(i, i-1)$ pour un indice $i \geq 1$ tel que $i+1 \in E$, on a :

$$\begin{aligned}\pi(i)A(i, i+1) &= \pi(i)(-A(i, i) - A(i, i-1)) \\ &= -\pi(i)A(i, i) - \pi(i-1)A(i-1, i) \\ &= \pi(i+1)A(i+1, i),\end{aligned}$$

puisque $\pi A(i) = 0$. Le résultat est démontré. \square

Le théorème suivant permet de calculer très facilement de nombreuses lois stationnaires dans des cas pratiques.

THÉORÈME 4.21 (Théorème de Kelly).— Soit $(X(t), t \geq 0)$ un processus de Markov réversible sur E , de générateur infinitésimal A et de probabilité invariante π . Soit $F \subset E$. On définit, pour un certain $\alpha \geq 0$, la matrice \tilde{A} suivante sur $E \times E$:

$$\begin{aligned}\tilde{A}(x, y) &= \begin{cases} \alpha A(x, y) & \text{si } x \in F, y \in E \setminus F; \\ A(x, y) & \text{sinon pour } x \neq y; \end{cases} \\ \tilde{A}(x, x) &= - \sum_{y \neq x} \tilde{A}(x, y) \text{ pour tout } x \in E.\end{aligned}$$

Alors, le processus de Markov $(\tilde{X}(t), t \geq 0)$ de générateur \tilde{A} est réversible, de probabilité invariante $\tilde{\pi}$ donnée pour tout $i \in E$ par :

$$\pi(x) = \begin{cases} C\pi(x) & \text{si } x \in F; \\ C\alpha\pi(x) & \text{si } x \in E \setminus F, \end{cases}$$

où $C = \left(\sum_{k \in F} \pi(k) + \alpha \sum_{k \in E \setminus F} \pi(k) \right)^{-1}$ est la constante de normalisation.

Démonstration. On montre très facilement que $\tilde{\pi}$ définit bien une probabilité sur E . Il suffit de vérifier qu'elle est réversible, ce qui est immédiat en remarquant que :

$$\tilde{\pi}(x)\tilde{A}(x, y) = \alpha C\pi(x)A(x, y) = \tilde{\pi}(y)\tilde{A}(y, x)$$

pour tout $x \in F$ et $y \in E \setminus F$. \square

Si l'on pose $\alpha = 0$, on interdit donc au processus de quitter le sous-ensemble F , tout en gardant la propriété de réversibilité. Nous verrons une application de ce résultat au chapitre 6.

4.6. MMPP

Les processus de Poisson modulé, *Markov Modulated Poisson Process* (MMPP), correspondent à une classe de processus markoviens qui généralisent les processus de Poisson tout en gardant des caractéristiques calculables. Ils apparaissent naturellement dans la modélisation des processus de débordement, voir chapitre 6. On a pendant un temps pensé qu'ils pouvaient servir de modèles pour les flux de données. Même si cette approche semble tombée en désuétude, elle n'en demeure pas moins intéressante, voir le chapitre 1.

DÉFINITION 4.13.— Soit $(J(t), t \geq 0)$ un processus de Markov stationnaire à valeurs dans $E = \llbracket 1, m \rrbracket$. Soit Q son générateur infinitésimal et ν sa probabilité stationnaire. Soit λ une fonction de E dans \mathbf{R}^+ . Le processus ponctuel N est un MMPP de paramètres Q et λ si et seulement si pour toute fonction f à support compact, on a :

$$\mathbf{E} \left[\exp \left(- \int_0^t f(s) \, dN(s) \right) \right] = \mathbf{E} \left[\exp \left(- \int_0^t (1 - e^{-f(s)}) \lambda(J(s)) \, ds \right) \right].$$

▮ Un MMPP est donc un processus de Cox dont l'intensité, variable dans le temps, dépend de l'évolution d'un processus de Markov à espace d'états fini.

Cela revient à dire que lorsque le processus des phases J est dans la phase j , les arrivées du processus ponctuel N se font selon un processus de Poisson d'intensité $\lambda(j)$. Quand le processus des phases change d'état, selon la dynamique induite par son générateur infinitésimal, l'intensité du processus de Poisson des arrivées change. La figure 4.4 représente une réalisation d'un MMPP à deux phases.

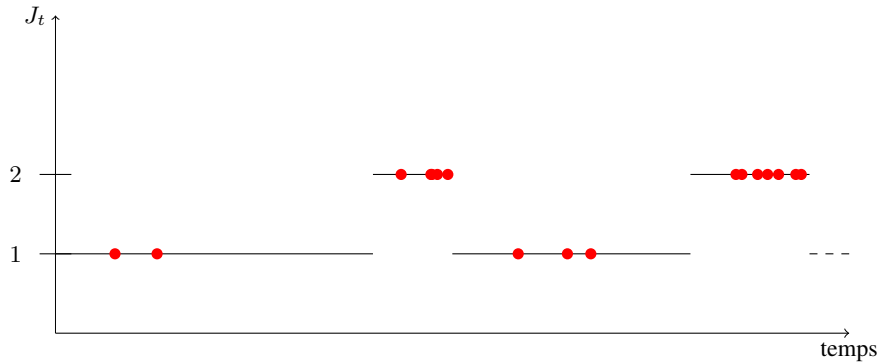


Figure 4.4 – Une réalisation d'un MMPP à deux phases

Le processus N n'est pas markovien seul car à l'instant t , le temps avant la prochaine arrivée suit, certes, une loi exponentielle, mais dont le paramètre dépend de la phase. Le processus (N, J) est markovien et son générateur s'écrit sous forme de matrices par blocs. On note Λ , la matrice diagonale dont le coefficient en (i, i) est $\lambda(i)$. Avec ces notations, le générateur infinitésimal de (N, J) est déterminé par :

$$\begin{pmatrix} Q - \Lambda & \Lambda & & & \\ 0 & Q - \Lambda & \Lambda & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

En effet, les événements sont de deux types : arrivée ou changement de phase. Une arrivée engendre une transition d'un état (i, ϕ) vers $(i + 1, \phi)$. Un changement de phase laisse la première composante inchangée et modifie la deuxième. On en déduit un algorithme (voir 4.1) de simulation de trajectoires de MMPP. On suppose toujours que l'espace d'états est $E = \llbracket 1, m \rrbracket$.

EXEMPLE 4.2.— Un autre exemple de MMPP est donné par le processus de débordement d'une file $M_\lambda/M_\mu/S/S$. Dans un tel système, tant que les serveurs ne sont pas tous occupés, rien ne débord. Lorsque tous les serveurs travaillent, les nouvelles arrivées sont refusées et forment le processus de débordement. Pour représenter cela sous forme d'un MMPP, il suffit de considérer comme processus des phases le processus J qui compte le nombre de serveurs occupés et la fonction λ définie par :

$$\lambda(i) = \begin{cases} 0 & \text{si } i < S, \\ \lambda & \text{si } i = S. \end{cases}$$

Pour qu'elle soit utilisable en pratique, une classe de processus doit être stable par superposition : si l'on superpose deux processus de la même classe (par exemple Poisson, Cox ou MMPP), il est souhaitable que le processus « somme » soit de la même classe. Nous savons déjà que les processus de Poisson satisfont cette propriété. Nous allons l'étendre au cas des MMPP. Cela nécessite d'introduire de nouvelles notations.

DÉFINITION 4.14.— Le produit de Kronecker de deux matrices A et B , noté $A \otimes B$, est la matrice :

$$\begin{pmatrix} A(1, 1)B & A(1, 2)B & \dots & A(1, n)B \\ \vdots & \vdots & & \vdots \\ A(n, 1)B & \dots & \dots & A(n, n)B \end{pmatrix}.$$

La somme de Kronecker est alors définie par $A \oplus B = (A \otimes \text{Id}_B) + (\text{Id}_A \otimes B)$ où Id_A est la matrice identité de même taille que A .

Algorithme 4.1 . Réalisation d'une trajectoire d'un MMPP**Données :** $m, Q, \lambda, \{i_0\}, T$ **Résultat :** une trajectoire $(t_n, n \geq 1)$ sur $[0, T]$ d'un MMPP (m, Q, λ) d'état initial $\{i_0\}$.**pour** $i = 1, \dots, m$ **faire** $r_{i0} = \lambda_i$; **pour** $j = 1, \dots, m$ **faire** $r_{i,j} = r_{i,j-1} + q(i, j)$; **fin****fin**phase $\leftarrow i_0$; $t \leftarrow 0$; $n \leftarrow 0$;**tant que** $t \leq T$ **faire** $x \leftarrow$ réalisation d'une $\varepsilon \left(\lambda(\text{phase}) + \sum_j q(\text{phase}, j) \right)$; $t \leftarrow t + x$; $u \leftarrow$ réalisation d'une variable aléatoire de loi uniforme sur $[0, 1]$; **si** $u \leq r_{i0}$ **alors** $t_n \leftarrow t$ **sinon** $j \leftarrow 1$; **tant que** $u > r_{i,j}$ **faire** $j \leftarrow j + 1$; **fin** phase $\leftarrow j$; **fin** $n \leftarrow n + 1$ **fin****retourner** t_1, t_2, \dots, t_n Si A est diagonale de taille n , on peut lui associer la fonction :

$$a : \begin{cases} \llbracket 1, n \rrbracket & \longrightarrow \mathbf{R} \\ i & \longmapsto A(i, i). \end{cases}$$

Pour deux matrices diagonales A_1 et A_2 donnant lieu à deux fonctions a_1 et a_2 définies respectivement sur $\llbracket 1, n_1 \rrbracket$ et $\llbracket 1, n_2 \rrbracket$, la matrice $A_1 \oplus A_2$ est la matrice qui correspond à la fonction :

$$a_1 \oplus a_2 : \begin{cases} \llbracket 1, n_1 \rrbracket \times \llbracket 1, n_2 \rrbracket & \longrightarrow \mathbf{R} \\ (i, j) & \longmapsto A_1(i, i) + A_2(j, j) = a_1(i) + a_2(j). \end{cases}$$

THÉORÈME 4.22.— Soit $J_1 = (E_1, A_1, \nu_1)$ et $J_2 = (E_2, A_2, \nu_2)$ deux processus de Markov indépendants. Le processus « produit » $J = (J_1, J_2)$ est un processus de Markov de caractéristiques $(E_1 \times E_2, A_1 \oplus A_2, \nu_1 \otimes \nu_2)$.

Démonstration. Il est évident que l'espace d'états est $E = E_1 \times E_2$. Compte tenu du théorème 4.15, il suffit de démontrer que pour toute fonction $f \in l^\infty(E)$, le processus :

$$t \rightarrow \Theta^f(t) = f(J_1(t), J_2(t)) - \int_0^t (A_1 \oplus A_2)f(J_1(s), J_2(s)) \, ds$$

est une martingale. Supposons d'abord que $f = f_1 \otimes f_2$ avec f_1 et f_2 bornées. Puisque J_1 et J_2 sont des processus de Markov, pour $i = 1, 2$, pour tout $t \geq 0$, on peut écrire :

$$f_i(J_i(t)) = f_i(J_i(0)) + \int_0^t A_i f_i(J_i(s)) \, ds + M_i(t),$$

où M_1 et M_2 sont des martingales. La formule d'Itô donne alors :

$$\begin{aligned} f_1(J_1(t))f_2(J_2(t)) - f_1(J_1(0))f_2(J_2(0)) \\ = \int_0^t f_1(J_1(s))A_2f_2(J_2(s)) \, ds + \int_0^t f_2(J_2(s))A_1f_1(J_1(s)) \, ds \\ + \sum_{s \leq t} \Delta f_1(J_1(s))\Delta f_2(J_2(s)) + M(t), \end{aligned}$$

où M est une martingale. Comme J_1 et J_2 sont indépendants, presque sûrement ils n'ont aucun instant de sauts commun, donc l'avant-dernier terme est nul. On peut donc réécrire :

$$\begin{aligned} f_1(J_1(t))f_2(J_2(t)) - f_1(J_1(0))f_2(J_2(0)) \\ = \int_0^t (A_1 \oplus A_2)(f_1 \otimes f_2)(J_1(s), J_2(s)) \, ds + M(t). \end{aligned}$$

Le résultat est donc prouvé pour $f = f_1 \otimes f_2$. Par linéarité, pour $f = \sum_{k=1}^n f_{1k} \otimes f_{2k}$ et $\psi \in \mathcal{F}_t$ bornée, on a donc :

$$\mathbf{E} [\Theta^f(t+s)\psi] = \mathbf{E} [\Theta^f(t)\psi], \text{ pour tout } s \geq 0. \quad (4.30)$$

Pour toute fonction f bornée, il existe une suite de fonctions $(f^l, l \geq 1)$ de la forme $f^l = \sum_{k=1}^{n_l} f_{1k}^l \otimes f_{2k}^l$, qui converge uniformément vers f . Comme A_1 et A_2 sont des noyaux markoviens, $\|A_i f\|_\infty \leq \|f\|_\infty$. Par conséquent, d'après le théorème de convergence dominée, Θ^{f^l} converge dans $l^\infty(E)$ vers Θ^f . Comme ψ est bornée, (4.30)

est vraie aussi pour toute fonction f bornée, par conséquent Θ^f est une martingale. Finalement, il vient :

$$\begin{aligned}\nu_1 \otimes \nu_2(A_1 \oplus A_2) &= \nu_1 \otimes \nu_2(A_1 \otimes \text{Id}_2 + \text{Id}_1 \otimes A_2) \\ &= \nu_1 A_1 \otimes \nu_2 + \nu_1 \otimes \nu_2 A_2 \\ &= 0\end{aligned}$$

et donc $\nu_1 \otimes \nu_2$ est une mesure invariante du processus (J_1, J_2) . \square

Il est intéressant d'écrire « à la main » le générateur du processus produit et de retrouver la forme $A_1 \oplus A_2$ (voir l'exercice 15).

THÉORÈME 4.23.— Soit N_i , $i = 1, \dots, K$ des MMPP indépendants tels que pour tout i , J_i est à valeurs dans $E_i = \llbracket 1, m_i \rrbracket$, de générateur infinitésimal A_i et de probabilité invariante ν_i . On note Λ_i la matrice diagonale des taux pour le MMPP N_i .

Le processus superposition N est un MMPP de processus des phases J , dont le générateur infinitésimal est donné par :

$$A = A_1 \oplus A_2 \oplus \dots \oplus A_K$$

et de fonction de taux $\lambda = \lambda_1 \otimes \lambda_2 \otimes \dots \otimes \lambda_K$. La probabilité invariante du processus des phases est $\nu_1 \otimes \dots \otimes \nu_K$.

✎ Donnons une idée intuitive de ce qui se passe dans le cas de deux MMPP. Tant qu'aucun processus de phases ne change d'état, les clients arrivent selon la superposition de deux processus de Poisson indépendants, donc selon un processus de Poisson d'intensité somme. On a donc bien sur des intervalles aléatoires, un processus de Poisson dont l'intensité est modulée par les couples de phases sous-jacents. Toutes les combinaisons de phases sont *a priori* possibles donc il y a $m_1.m_2$ phases et intensités possibles, ce que traduit le fait que λ est définie sur l'espace produit $E_1 \times E_2$. Il faudrait ensuite s'assurer que le processus des phases est bien markovien. La démonstration par les transformées de Laplace est plus abstraite mais largement plus facile que ce début de raisonnement.

Démonstration. Nous pouvons nous contenter de faire la preuve pour $K = 2$, le cas général s'en déduit par récurrence. D'après la définition d'un MMPP, pour $i = 1, 2$ et f_1, f_2 des fonctions bornées, pour tout $t \geq 0$, on a :

$$f_i(N_i(t)) = f_i(N_i(0)) + \int_0^t f_i(N_i(s)) \lambda_i(J_i(s)) \, ds + M^i(t),$$

où M^i est une martingale. Comme les processus sont indépendants, les martingales M^i sont indépendantes et donc leur crochet est nul. Par conséquent, d'après la formule d'intégration par parties (A.12), on a :

$$\begin{aligned} f_1(N^1(t))f_2(N^2(t)) - f_1(N^1(0))f_2(N^2(0)) \\ = \int_0^t f_2(N_2(s)) \, dM^1(s) + \int_0^t f_1(N_1(s)) \, dM^2(s) \\ + \int_0^t f_1(N_1(s))f_2(N_2(s))(\lambda_1(J_1(s)) + \lambda_2(J_2(s))) \, ds. \end{aligned}$$

En vertu du théorème A.32, on a prouvé que :

$$\begin{aligned} t \longmapsto (f_1 \otimes f_2)(N_1(t), N_2(t)) - (f_1 \otimes f_2)(N_1(0), N_2(0)) \\ - \int_0^t (f_1 \otimes f_2)(N_1(s), N_2(s))(\lambda_1(J_1(s)) + \lambda_2(J_2(s))) \, ds \end{aligned}$$

est une martingale. Par densité, ce résultat reste vrai pour toutes les fonctions bornées sur $\mathbf{N} \times \mathbf{N}$, en particulier pour les fonctions de la forme $(n_1, n_2) \mapsto f(n_1 + n_2)$ avec f bornée de \mathbf{N} dans \mathbf{R} . Il s'ensuit que le processus $N_1 + N_2$ est un processus de Cox de compensateur $\lambda_1(J_1) + \lambda_2(J_2)$. En vertu du théorème 4.22, cela correspond bien au processus intensité associé au processus produit (J_1, J_2) . \square

THÉORÈME 4.24 (PASTA modifiée).— Soit $N = (E, J, \nu)$ un MMPP et $(\psi(t), t \geq 0)$ un processus \mathcal{F}^N prévisible borné. On a l'identité :

$$\lim_{t \rightarrow \infty} \frac{1}{N_t} \int_0^t \psi(s) \, dN_s = \frac{1}{\sum_{j \in E} \lambda(j)\nu(j)} \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \psi(s)\lambda(J(s)) \, ds, \quad (4.31)$$

dès lors que chacune des limites existe.

Démonstration. D'après le théorème A.37, comme ψ est bornée, on a :

$$\lim_{t \rightarrow \infty} \frac{1}{N(t)} \int_0^t \psi(s) \, dN(s) = \lim_{t \rightarrow \infty} \frac{t}{\int_0^t \lambda(J(s)) \, ds} \frac{1}{t} \int_0^t \psi(s)\lambda(J(s)) \, ds.$$

Comme J est un processus de Markov ergodique, on a aussi :

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \lambda(J(s)) \, ds = \sum_{j \in E} \lambda(j)\nu(j),$$

d'où le résultat. \square

4.7. Problèmes

EXERCICE 13.– Soit $(J(t), t \geq 0)$ le processus de Markov à valeurs dans $\{A, B\}$ dont le générateur infinitésimal est donné par :

$$\begin{pmatrix} -\sigma_A & \sigma_A \\ \sigma_B & -\sigma_B \end{pmatrix}$$

- 1) Déterminer la probabilité stationnaire π de J .
- 2) Quelle est la loi du temps de séjour à l'état A ? On construit maintenant un processus d'arrivées de la manière suivante : pendant les périodes où $J(t) = A$ il n'y a pas d'arrivées, pendant les périodes où $J(t) = B$, les arrivées se font un processus de Poisson d'intensité λ . Notons comme d'habitude par N_t le nombre de clients qui arrivent jusqu'à l'instant t . On admet qu'il ne peut y avoir simultanément changement de phase (c'est-à-dire changement d'état du processus J et une arrivée).
- 3) Ecrire les coefficients non nuls du générateur infinitésimal du processus de Markov (J, N) .
- 4) Ce processus admet-il une probabilité stationnaire? On appelle ce processus, un processus IPP (*Interrupted Poisson Process*), c'est un modèle possible de la voix humaine (la phase A représentant les périodes de silence et la phase B les périodes de parole). On veut maintenant étudier la file $IPP/M/1/1$. On appelle $1/\mu$ le temps moyen de service. Pour ce faire, on étudie le processus (X, J) où X représente le nombre de clients dans la file. On énumère les états dans l'ordre lexicographique :

$$(0, A), (0, B), (1, A), (1, B).$$

- 5) Ecrire le générateur infinitésimal de (X, J) .
- 6) Ce processus admet-il une probabilité stationnaire? Expliciter les valeurs de trois de ses composantes en fonction de la quatrième.
- 7) Exprimer la probabilité stationnaire de X en fonction de celle de (X, J) . On suppose maintenant que :

$$\lambda = 3, \mu = 1, \sigma_A = 1, \sigma_B = 2.$$

- 8) Calculer complètement les probabilités stationnaires de (X, J) puis de X .
- 9) Quelle est la probabilité de blocage à l'état stationnaire?
- 10) Quelle est la probabilité de perte à l'état stationnaire?

EXERCICE 14.– Pour tout entier n et tout réel positif β , on rappelle que :

$$\int_0^{+\infty} x^n \beta e^{-\beta x} dx = \frac{n!}{\beta^n}.$$

On considère une file d'attente dans laquelle les temps de service sont indépendants et de même loi. Une requête a une probabilité p que son temps de service suive une loi exponentielle de paramètre μ_α et une probabilité $q = 1 - p$ que son temps de service suive une loi exponentielle de paramètre μ_β . Formellement, si X_1, X_2 sont deux variables aléatoires indépendantes de loi exponentielle de paramètres respectifs μ_α et μ_β , si Y est une variable aléatoire indépendante de X_1 et X_2 telle que $P(Y = 1) = p = 1 - P(Y = 0)$, alors la variable aléatoire $X = X_Y = X_1 \mathbf{1}_{\{Y=1\}} + X_2 \mathbf{1}_{\{Y=2\}}$ a la loi d'une durée de service.

Les arrivées forment un processus de Poisson d'intensité λ . Il y a un seul serveur et un buffer de taille K .

- 1) Montrer que le temps moyen de service est donné par :

$$p \frac{1}{\mu_\alpha} + q \frac{1}{\mu_\beta}.$$

On notera cette quantité par $1/\mu$.

- 2) Montrer que la variance du temps moyen de service est donnée par :

$$p \frac{2}{\mu_\alpha^2} + q \frac{2}{\mu_\beta^2} - \frac{1}{\mu^2}.$$

Afin d'étudier les performances de cette file d'attente, on considère le processus de Markov X qui représente le nombre de clients dans le système et la phase du client en service. L'espace d'état est donc :

$$E = \{0\} \cup \{(i, \gamma), 1 \leq i \leq K + 1, \gamma \in \{\alpha, \beta\}\}.$$

Par exemple, lorsque X est dans l'état $(5, \alpha)$, cela signifie qu'il y a quatre clients dans le buffer et que le client en service est en phase α donc que son temps de service est distribué comme une exponentielle de paramètre μ_α .

- 3) Pour $K = 0$, écrire le générateur infinitésimal de X .
- 4) Pour $K = 0$, calculer la probabilité de perte à l'état stationnaire.
- 5) Comparer au résultat obtenu pour la file M/M/1/1 pour la même charge.

On suppose dorénavant que le buffer est de taille infinie.

- 1) Quelle est la condition d'existence d'une probabilité stationnaire ? (Aucun calcul à faire !)
- 2) Quel est le nombre moyen de clients dans le système à l'état stationnaire ?
- 3) Si $\mu_\alpha = 1$ et μ est fixé, quelle est la relation entre p et μ_β ?
- 4) Représenter les variations du nombre moyen de clients quand p varie de 0 à $1/\mu$.
- 5) Ecrire les coefficients non nuls du générateur infinitésimal de X .

6) Soit π le vecteur représentant la probabilité stationnaire, on pose $x_0 = \pi(0)$, $x_i = (\pi(i, \alpha), \pi(i, \beta))$. Ecrire les équations satisfaites par les x_i en utilisant des produits matriciels par blocs. Il n'est pas demandé de résoudre le système obtenu.

EXERCICE 15.— Soit $J_1 = (E_1, A_1, \nu_1)$ et $J_2 = (E_2, A_2, \nu_2)$ deux processus de Markov indépendants. Ecrire « à la main » le générateur du processus produit $J = (J_1, J_2)$, pour retrouver la conclusion du théorème 4.22.

4.8. Notes et commentaires

Il existe de nombreux livres sur les processus de Markov à espaces d'états discrets et leurs applications notamment aux files d'attente, par ordre plus ou moins chronologique, citons de façon non exhaustive [ASM 03, CIN 75, KEL 79, KLE 76]. Ces livres n'abordent pas les problèmes de martingales à cause de leur complexité. Cet outil est néanmoins à la base de nombreux travaux notamment sur les limites fluides comme l'on en trouvera de magnifiques exemples dans [ROB 03]. Les livres de référence plus mathématiques sur ce sujet sont [BRE 81, DEL 76, ETH 86, JAC 79]. Nous avons délibérément choisi cette approche un peu plus formelle que celle utilisée dans le chapitre sur les chaînes de Markov à la fois pour son élégance mais aussi pour préparer le terrain pour les lecteurs qui voudraient se lancer dans la théorie, autrement plus ardue, des processus de Markov à espaces d'états continus.

Pour en savoir plus sur les MMPP, on pourra se référer à [FIS 93, NEU 94] et à leurs références. Noter toutefois que l'approche par les martingales développée ici est originale.

Résumé

– Un processus de Markov à valeurs dans un espace d'états discret est un processus dont les trajectoires sont faites de sauts régulés par une chaîne de Markov et dont les durées de séjour dans chaque état sont de loi exponentielle. Le paramètre de la loi exponentielle dépend de l'état. Entre deux sauts, le processus n'évolue pas.

– Son générateur infinitésimal A est une matrice éventuellement infinie dont les coefficients s'interprètent par :

- $A(x, x)$ est l'opposé du paramètre de la loi exponentielle qui régit le temps de séjour dans l'état x . Attention au signe moins !

- $A(x, y)/|A(x, x)|$ est la probabilité que le processus aille en y lorsqu'il quitte x .

– La nature des états (récurrents, transients) du processus de Markov est la même que la nature des états de la chaîne de Markov incluse.

– La probabilité invariante π s'obtient en résolvant les équations $\pi A = 0$ et $\sum_{x \in E} \pi(x) = 1$ où π est écrite sous forme de vecteur ligne.

– La probabilité stationnaire par une multiplication de tous les coefficients du générateur infinitésimal par un même réel positif. La dynamique change : les temps dans chaque état sont plus courts, les transitions restent les mêmes. En modélisation de files d'attente, on tire profit de cela en choisissant la durée moyenne de service comme l'unité de temps, c'est-à-dire $\mu = 1$.

– Si le processus est irréductible, récurrent, de probabilité invariante π , on a :

$$\frac{1}{t} \int_0^t f(X(s)) \, ds \xrightarrow{t \rightarrow \infty} \sum_{x \in E} f(x) \pi(x)$$

$$\mathbf{P}(X(t) = y \mid X(0) = x) \xrightarrow{t \rightarrow \infty} \pi(y), \quad \forall x, y \in E.$$

– Le temps de retour à un point se déduit de la probabilité invariante :

$$\mathbf{E}_x [\tau_x^1] = \frac{1}{\pi(x)}.$$

– Les trajectoires d'un processus de Markov se simulent en utilisant la construction de la définition 4.1.

Chapitre 5

Systèmes à attente

Tous les systèmes de télécommunications réels sont des processus à pertes puisque toutes les mémoires ont des capacités finies, donc limitées. On entend par système à attente, un système dans lequel le dimensionnement est tel que les pertes par débordement sont négligeables et pour lesquels le critère pertinent d'évaluation des performances est le temps d'attente.

Avant de nous lancer dans l'étude détaillée de ces systèmes, il est important de mentionner une formule bien connue sous le nom de formule de Little.

5.1. Formule de Little

On considère un système à attente dans lequel les clients arrivent à des instants $(T_n, n \geq 1)$, restent un temps $(W_n, n \geq 1)$ et en sortent aux dates $(D_n = T_n + W_n, n \geq 1)$. On note N , le processus ponctuel des entrées et X le processus qui compte le nombre de clients présents dans le système à t . A l'instant 0, le système est supposé vide, donc $X(0) = 0$. L'important est que le système soit conservatif : autant de travail en sortie qu'en entrée.

THÉORÈME 5.1 (Formule de Little).— On suppose que N est asymptotiquement linéaire, c'est-à-dire qu'il existe $\lambda > 0$ tel que :

$$\frac{N(t)}{t} \xrightarrow[t \rightarrow \infty]{} \lambda, \text{ p.s.}$$

On suppose que la suite W est ergodique, c'est-à-dire que :

$$\frac{1}{n} \sum_{j=1}^n W_j \xrightarrow[n \rightarrow \infty]{} W, \text{ p.s.}$$

Sous ces hypothèses, on a :

$$X = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X(s) \, ds = \lambda W.$$

L'existence de cette dernière limite est montrée dans la preuve.

Démonstration. Fixons $t \geq 0$ et appliquons la formule d'intégration par parties (A.12) aux processus X et $t \mapsto t$. Comme le deuxième est continu, il n'y a pas de parties « crochet » et l'on a :

$$tX(t) = \int_0^t X(s) \, ds + \int_0^t s \, dX(s).$$

Or les sauts de X sont ceux de N et D , le processus des départs. Par conséquent :

$$\begin{aligned} \int_0^t X(s) \, ds &= \int_0^t (t-s) \, dN(s) - \int_0^t (t-s) \, dD(s) \\ &= \sum_{T_n \leq t} (t - T_n) - \sum_{T_n + W_n \leq t} (t - T_n - W_n). \end{aligned}$$

Il faut maintenant distinguer les clients qui sont entrés puis sortis avant t , de ceux qui sont entrés avant t mais pas encore sortis à t . D'où :

$$\begin{aligned} \int_0^t X(s) \, ds &= \sum_{T_n + W_n \leq t} (t - T_n - (t - T_n - W_n)) + \sum_{T_n + W_n > t, T_n \leq t} (t - T_n) \\ &= \sum_{T_n + W_n \leq t} W_n + \sum_{T_n + W_n > t, T_n \leq t} (t - T_n). \end{aligned}$$

Pour les instants T_n tels que $(T_n + W_n > t, T_n \leq t)$, on a évidemment $0 \leq t - T_n \leq W_n$ donc :

$$\sum_{T_n + W_n \leq t} W_n \leq \int_0^t X(s) \, ds \leq \sum_{T_n + W_n \leq t} W_n + \sum_{T_n + W_n > t, T_n \leq t} W_n = \sum_{T_n \leq t} W_n.$$

Par définition de $N(t)$, on a :

$$\sum_{T_n \leq t} W_n = \sum_{n=1}^{N(t)} W_n,$$

donc :

$$\frac{1}{t} \sum_{T_n \leq t} W_n = \frac{N(t)}{t} \frac{1}{N(t)} \sum_{n=1}^{N(t)} W_n \xrightarrow{t \rightarrow \infty} \lambda W,$$

d'après les deux hypothèses. Il reste à prouver que l'on a la même limite pour la borne inférieure. Pour cela, il faut contrôler le nombre de clients entrés avant t et pas encore sortis à cet instant. Remarquons que :

$$\begin{aligned}\frac{W_n}{n} &= \frac{1}{n} \left(\sum_{j=1}^n W_j - \sum_{j=1}^{n-1} W_j \right) \\ &= \frac{1}{n} \sum_{j=1}^n W_j - \frac{n-1}{n} \frac{1}{n-1} \sum_{j=1}^{n-1} W_j \\ &\xrightarrow{n \rightarrow \infty} 0, \text{ p.s.}\end{aligned}$$

Par conséquent :

$$\frac{W_n}{T_n} = \frac{W_n}{n} \frac{N(T_n)}{T_n} \xrightarrow{n \rightarrow \infty} 0. \lambda = 0, \text{ p.s.}$$

Raisonnons à trajectoire fixée. Pour tout $\varepsilon > 0$, il existe m (dépendant de la trajectoire) tel que $W_n \leq \varepsilon T_n$ pour $n \geq m$. Pour $n \geq m$, $T_n + W_n \leq (1 + \varepsilon)T_n$. Si $n \geq m$ et $T_n \leq t(1 + \varepsilon)^{-1}$ alors le client n est sorti à l'instant t donc :

$$\sum_{T_n + W_n \leq t} W_n \geq \sum_{j=m}^{N(t(1+\varepsilon)^{-1})} W_j = \sum_{j=1}^{N(t(1+\varepsilon)^{-1})} W_j - \sum_{j=1}^{m-1} W_j.$$

Le même raisonnement que pour la borne supérieure donne :

$$\liminf_{t \rightarrow \infty} \sum_{T_n + W_n \leq t} W_n \geq \lambda W (1 + \varepsilon)^{-1}.$$

Comme ce résultat est vrai pour tout $\varepsilon > 0$, on en déduit qu'il est encore vrai pour $\varepsilon = 0$. D'après le célèbre théorème des « gendarmes », on en déduit le résultat. \square

EXEMPLE.— Prenons comme système le serveur d'une file à un seul serveur. Le temps de séjour y est égal au temps de service. Comme il y a 0 ou 1 client dans le serveur, le nombre moyen de clients correspond au taux d'occupation du serveur noté τ . De la formule de Little, on déduit $\tau = \lambda \times 1/\mu$. En d'autres termes, la charge est bien le pourcentage du temps où le serveur est occupé.

EXEMPLE.— Revenons sur l'équation (1.2) associée à la figure 1.3. On peut considérer que l'on a un système où les clients rentrent aux instants T_n et sortent aux instants $T_n + Y_n$. Le temps de traversée du client n est donc Y_n . D'après la loi forte des grands nombres, on a :

$$\frac{T_n}{n} = \frac{1}{n} \sum_{j=1}^n (X_j + Y_j) \xrightarrow{n \rightarrow \infty} \frac{1}{\mu} + \frac{1}{\tau}, \text{ presque sûrement.} \quad (5.1)$$

Si l'on note N le nombre d'arrivées entre 0 et t , on a évidemment :

$$T_n \leq t < T_{n+1} \iff N(t) = n,$$

donc :

$$\frac{N(t)}{T_{N(t)+1}} \leq \frac{N(t)}{t} \leq \frac{N(t)}{T_{N(t)}}.$$

Comme $N(t)$ tend vers l'infini presque sûrement, le théorème « dit des gendarmes » et (5.1) impliquent que $N(t)/t$ tend vers $1/(1/\mu + 1/\tau) = \lambda$. Par construction, Q représente le pourcentage du temps où le serveur est actif, d'après la formule de Little on obtient : $Q = \lambda/\mu$. On a donc bien montré que dans ce cas, la charge est bien égale au produit du nombre moyen d'arrivées par le temps moyen de traitement.

REMARQUE.— On remarque que dans la démonstration de la formule de Little, il n'est nullement fait mention de la discipline de service. Il s'ensuit que le temps moyen d'attente est le même dans la discipline FIFO que dans la discipline LIFO. Ce phénomène pour le moins paradoxal signifie seulement que le temps moyen d'attente n'est qu'une mesure bien pauvre en renseignements sur le comportement du système. En revanche, elle est souvent très facile à calculer. En fait, la différence entre les disciplines de service se verra sur la variance et plus généralement sur la loi du temps d'attente.

5.2. File à un serveur

THÉORÈME 5.2.— Pour la file d'attente $M_\lambda/M_\mu/1/\infty$ -FIFO, on pose $\rho = \lambda/\mu$. La condition de stabilité au sens du théorème ?? est donnée par $\rho < 1$. Dans ce cas, la probabilité invariante π est définie par :

$$\pi(0) = 1 - \rho, \pi(n) = \rho^n(1 - \rho), n \geq 1.$$

Le nombre moyen de clients dans le système à l'état stationnaire est $\rho(1 - \rho)^{-1}$.

Démonstration. On a vu dans l'exemple 4.1 que le processus $X = (X(t), t \geq 0)$ désignant le nombre de clients dans le système (aussi appelé processus de congestion du système) est markovien de générateur infinitésimal :

$$A = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & & \ddots & & \\ (0) & \mu & -(\lambda + \mu) & \lambda & \\ & & & \ddots & \end{pmatrix}.$$

L'unique probabilité stationnaire π du processus X sur \mathbf{N} (si elle existe) vérifie le système matriciel :

$$\pi A = 0 \text{ et } \pi \cdot \mathbf{e} = 1, \quad (\text{S})$$

où $\mathbf{e} = (1, 1, \dots, 1)$ et $x.y$ le produit scalaire euclidien de deux vecteurs lignes x et y . La première égalité du système (S) équivaut aux équations :

$$\begin{cases} -\lambda\pi(0) + \mu\pi(1) = 0 \\ \lambda\pi(0) - (\lambda + \mu)\pi(1) + \mu\pi(2) = 0 \\ \vdots \\ \lambda\pi(i-1) - (\lambda + \mu)\pi(i) + \mu\pi(i+1) = 0 \\ \vdots \end{cases}$$

En additionnant ces équations deux à deux successivement, on obtient :

$$-\lambda\pi(0) + \mu\pi(1) = 0, -\lambda\pi(1) + \mu\pi(2) = 0, \dots, -\lambda\pi(i) + \mu\pi(i+1) = 0,$$

soit, pour tout entier i ,

$$\pi(i+1) = \rho\pi(i) \text{ avec } \rho = \lambda/\mu.$$

Par conséquent, $\pi(i) = \rho^i \pi(0)$. Il s'ensuit alors de la deuxième égalité de (S) que sous la condition :

$$\rho < 1, \quad (5.2)$$

on a :

$$\pi(0) = \frac{1}{\sum_{i \in \mathbf{N}} \rho^i} = 1 - \rho \text{ et donc } \pi(i) = \rho^i (1 - \rho), \text{ pour tout } i \geq 1.$$

Cela signifie, en d'autres termes, qu'à la condition (5.2) qui est, d'après le théorème ??, la condition de stabilité du système et donc de récurrence de 0 pour X , l'unique probabilité stationnaire de X est la loi d'une variable aléatoire $X_\infty = Z - 1$, où Z suit la loi géométrique de paramètre $1 - \rho$.

En particulier, le nombre de client moyen dans le système à l'équilibre est donné par :

$$\mathbf{E}[X_\infty] = \mathbf{E}[Z - 1] = \frac{1}{1 - \rho} - 1 = \frac{\rho}{1 - \rho}. \quad (5.3)$$

□

On suppose dorénavant que la condition de stabilité (5.2) est satisfaite.

Pour ce système, une métrique de performance intéressante est la proportion de client entrant dans un système vide ou, ce qui revient au même comme nous allons le

voir, la proportion du temps pendant laquelle le serveur est au repos. En rappelant que l'on note $T_1 < T_2 < \dots$, les instants d'entrée des clients dans la file, la limite en temps long de la proportion de clients entrant dans un système vide est donnée par :

$$\begin{aligned}
 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{\{0\}}(X(T_n^-)) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbf{1}_{\{0\}}(X(t^-)) dt \\
 &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbf{1}_{\{0\}}(X(t)) dt \\
 &= \mathbf{E}[\mathbf{1}_{\{0\}}(X_\infty)] \\
 &= \pi(0) = 1 - \rho,
 \end{aligned} \tag{5.4}$$

où la première égalité presque sûre découle de la propriété PASTA pour la fonction mesurable $\mathbf{1}_{\{0\}}$ (voir le théorème A.38), la deuxième du fait que X est presque sûrement à trajectoires cadlag et la troisième de l'ergodicité du processus X (théorème 2.21). Par exemple, un système de charge 1/2 fonctionne donc la moitié du temps et sur un long intervalle de temps, la moitié des clients trouvent à leur entrée un système vide.

5.2.1. Temps d'attente, temps de séjour

Intéressons-nous au temps d'attente (temps passé dans la salle d'attente) et au temps de séjour (temps total passé dans le système) des clients dans le système. Nous reprenons les notations du chapitre ?? et notons pour tout $n \geq 1$, TA_n le temps d'attente du client C_n (qui coïncide, en FIFO, avec la quantité de travail du serveur à l'entrée de C_n étudiée dans le chapitre ??). D'après le théorème ??, à la condition (5.2) il existe, pour toute discipline de service admissible, un temps d'attente stationnaire TA et par voie de conséquence, un temps de séjour stationnaire TS dans le système (attente + service), donné par $\text{TS} = \text{TA} + \sigma$. La formule de Little nous apprend alors que :

$$\mathbf{E}[\text{TS}] = \frac{\mathbf{E}[X_\infty]}{\lambda} = \frac{1}{\mu - \lambda}, \tag{5.5}$$

avec (5.3). On en déduit alors que :

$$\mathbf{E}[\text{TA}] = \mathbf{E}[\text{TS}] - \mathbf{E}[\sigma] = \frac{1}{\mu - \lambda} = \frac{\rho}{\mu - \lambda}. \tag{5.6}$$

Alors même que la congestion moyenne (5.3) ne dépend que de la charge, comparons les temps d'attente et de séjour des deux systèmes de même charge $M_\lambda/M_\mu/1$ et $M_{\alpha\lambda}/M_{\alpha\mu}/1$, où α est un paramètre positif. En ajoutant l'exposant α aux paramètres du deuxième système, on déduit de (5.5) et (5.6) que :

$$\mathbf{E}[\text{TS}^\alpha] = \frac{1}{\alpha} \mathbf{E}[\text{TS}] \text{ et } \mathbf{E}[\text{TA}^\alpha] = \frac{1}{\alpha} \mathbf{E}[\text{TA}].$$

Ainsi, les temps d'attente et de séjour moyens sont par exemple divisés par deux dès que les temps de service et d'interarrivée sont divisés par deux. Les paramètres de sortie conservent ainsi le facteur d'échelle des paramètres d'entrée.

Rappelons que l'on peut en fait obtenir explicitement la loi du temps d'attente et du temps de séjour stationnaires à travers leur transformée de Laplace. Il suffit pour cela de spécialiser les résultats du théorème ?? au cas où les temps de service suivent la loi $\varepsilon(\mu)$. On a donc le théorème suivant.

THÉORÈME 5.3.— Le temps d'attente stationnaire T_A et le temps de séjour stationnaire T_S dans une file M/M/1 admettent respectivement pour transformée de Laplace les fonctions définies pour tout s par :

$$\mathbf{E}[e^{-sT_A}] = (1 - \rho) \left(\frac{\mu + s}{\mu(1 - \rho) + s} \right), \quad (5.7)$$

$$\mathbf{E}[e^{-sT_S}] = (1 - \rho) \left(\frac{\mu}{\mu(1 - \rho) + s} \right). \quad (5.8)$$

La propriété PASTA permet aussi de montrer que la loi stationnaire de la chaîne incluse coïncide avec celle du processus.

LEMME 5.4.— Sous la condition $\rho < 1$, la chaîne de Markov $\hat{X} = (X(T_n^-), n \geq 0)$ est irréductible, récurrente positive de loi stationnaire π .

Démonstration. Il est évident que la chaîne est irréductible et récurrente, soit $\hat{\pi}$ sa loi invariante. D'après le théorème 4.13, pour $f \in l^\infty(\mathbf{N})$, on a :

$$\frac{1}{N} \sum_{n=1}^N f(X(T_n^-)) \xrightarrow{N \rightarrow \infty} \int_{\mathbf{N}} f \, d\hat{\pi}.$$

Comme $N(t)$ est à valeurs dans \mathbf{N} et tend vers l'infini quand t tend vers l'infini ; par extraction, on a :

$$\frac{1}{N(t)} \sum_{n=1}^{N(t)} f(X(T_n^-)) \xrightarrow{t \rightarrow \infty} \int_{\mathbf{N}} f \, d\hat{\pi}.$$

Or d'après le théorème A.38, il vient :

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{N(t)} \sum_{n=1}^{N(t)} f(X(T_n^-)) &= \lim_{t \rightarrow \infty} \frac{1}{N(t)} \int_0^t f(X(s^-)) \, dN(s) \\ &= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X(s^-)) \, ds = \int_{\mathbf{N}} f \, d\pi. \end{aligned}$$

On en déduit que pour toute fonction bornée f , on a :

$$\int_{\mathbf{N}} f \, d\pi = \int_{\mathbf{N}} f \, d\hat{\pi}.$$

Par identification, il s'ensuit que $\hat{\pi} = \pi$. □

REMARQUE.— L'installation d'un lien est un investissement qu'il faut amortir. Par conséquent, le lien doit être utilisé le plus possible, c'est-à-dire supporter une charge la plus proche possible de 1. Malheureusement, d'après (5.6), une telle charge induit des délais importants et donc un service de qualité limitée. Heureusement, dans la vraie vie, les liens sont chargés à 10 % de leur capacité...

REMARQUE.— Un même volume d'informations peut être transmis par des paquets de taille moyenne m émis à un taux λ ou par des paquets de taille moyenne cm émis à un taux λ/c . La charge, ρ , sera la même et donc le nombre moyen de clients en attente dans la file sera le même. Par conséquent, la taille des mémoires nécessaires (les mémoires physiques n'étant pas infinies) sera d'autant plus grande que la taille moyenne des paquets est grande. Cela n'est pas une évidence *a priori*, en effet le volume total d'octets transmis est le même ! A charge constante, la taille moyenne des paquets a aussi une influence sur le délai de séjour. En effet, $E[TS] = 1/\mu(1 - \rho)$. Transmettre des informations dans des grands paquets impose donc des délais de transmission grands. Il semblerait alors que la solution soit de choisir de petits paquets. Malheureusement, les petits paquets ont aussi leur côté négatif. Un paquet est constitué d'informations utiles et d'informations de contrôle. Comme les informations de contrôle sont, en général, de taille constante pour un protocole donné, limiter la taille des paquets revient alors à limiter le champ d'informations utiles. Par conséquent, plus les paquets sont petits, plus l'efficacité (définie comme le rapport entre les informations utiles transmises et la taille totale du paquet) diminue.

5.3. File à plusieurs serveurs

Nous abordons maintenant la file markovienne à S serveurs ($S \geq 1$) et à salle d'attente de capacité illimitée, noté $M_\lambda/M_\mu/S/\infty$ -FIFO avec la nomenclature habituelle. Ce modèle est représenté par le processus $(X^{(S)}(t), t \geq 0)$ comptant le nombre total de clients dans le système, à valeurs dans \mathbf{N} . Ce processus est naturellement markovien, nous déterminons ci-après sa matrice de transition :

– pour tout $i \in \llbracket 1, S - 1 \rrbracket$, si le processus est dans l'état i il y a i clients dans le système qui sont tous en service. Donc le temps de séjour en i (que nous noterons Y_i pour tout i) est le minimum d'une variable aléatoire $U \sim \varepsilon(\lambda)$ comptant le temps d'interarrivée en cours et de i variables aléatoires indépendantes et identiquement distribuées V_1, V_2, \dots, V_i de loi $\varepsilon(\mu)$, comptant les temps de service résiduels des i clients en service. Ensuite, le processus saute en $i - 1$ si $Y_i = V_k$ pour un certain k et en $i + 1$ si $Y_i = U$;

– pour tout $i \geq S$, dans l'état i les S serveurs sont occupés et $i - S$ clients sont en attente. Ainsi, Y_i égale le minimum de variables aléatoires V_1, V_2, \dots, V_S de loi $\varepsilon(\mu)$ et de U de loi $\varepsilon(\lambda)$, en conservant les notations précédentes. On saute alors en $i - 1$ ou en $i + 1$ comme dans le cas précédent ;

– Y_0 est de loi $\varepsilon(\lambda)$ et l'on quitte forcément l'état 0 pour aller en 1.

D'après la définition 4.1, la matrice de transition $Q^{(S)}$ du processus s'écrit donc :

– $q^{(S)}(0, 0) = \lambda$ et $q(0, 1) = 1$;

– pour tout $i \in \llbracket 1, S - 1 \rrbracket$, $q^{(S)}(i, i) = i\mu + \lambda$ et les seules transitions non nulles sont :

$$q^{(S)}(i, i + 1) = \frac{\lambda}{i\mu + \lambda} \text{ et } q^{(S)}(i, i - 1) = \frac{i\mu}{i\mu + \lambda} ;$$

– pour tout $i \geq S$, $q^{(S)}(i, i) = S\mu + \lambda$ et les seules transitions non nulles sont :

$$q^{(S)}(i, i + 1) = \frac{\lambda}{S\mu + \lambda} \text{ et } q^{(S)}(i, i - 1) = \frac{S\mu}{S\mu + \lambda} .$$

Ainsi, d'après 4.11 le processus $(X^{(S)}(t), t \geq 0)$ est markovien de générateur infinitésimal :

$$A^{(S)} = \begin{pmatrix} -\lambda & \lambda & & & & \\ \mu & -(\lambda + \mu) & \lambda & & & \\ 0 & 2\mu & -(\lambda + 2\mu) & \lambda & & \\ & & \ddots & & & \\ & (0) & S\mu & -(\lambda + S\mu) & \lambda & \\ & & S\mu & -(\lambda + S\mu) & \lambda & \\ & & & & \ddots & \end{pmatrix} \quad (0)$$

La démonstration du théorème suivant reproduit celle du théorème 5.2, elle est donc laissée au lecteur.

THÉORÈME 5.5.– Sous la condition de stabilité $\rho < S$, l'unique probabilité stationnaire $\pi^{(S)}$ du processus $(X^{(S)}(t), t \geq 0)$ s'écrit :

$$\begin{aligned} \pi^{(S)}(0) &= \left(\sum_{k=0}^{S-1} \frac{\rho^k}{k!} + \frac{S\rho^S}{S!(S-\rho)} \right)^{-1}, \\ \pi^{(S)}(i) &= \frac{\rho^i}{i!}, \pi(0) \text{ pour tout } i \in \llbracket 1, S - 1 \rrbracket, \\ \pi^{(S)}(i) &= \frac{\rho^i}{S^{i-S} S!}, \pi(0) \text{ pour tout } i \geq S. \end{aligned} \quad (5.9)$$

Sous la condition de stabilité $\rho < S$, on peut déduire du théorème 5.5, comme en (5.4), la limite du taux de clients qui doivent attendre pour être servis. Celle-ci est donnée par la formule d'Erlang-C :

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{\{[S, +\infty)\}} \left(X^{(S)}(T_n -) \right) &= \mathbf{E} \left[\mathbf{1}_{\{[S, +\infty)\}} \left(X_{\infty}^{(S)} \right) \right] \\ &= \sum_{i=S}^{+\infty} \pi^{(S)}(i) = \frac{\pi^{(S)}(0)}{S!S^{-S}} \sum_{i=1}^{+\infty} \left(\frac{\rho}{S} \right)^i \\ &= \left(\sum_{k=0}^{S-1} \frac{\rho^k}{k!} + \frac{S\rho^S}{S!(S-\rho)} \right)^{-1} \frac{S\rho^S}{S!(S-\rho)} \end{aligned}$$

Cette dernière formule est appelée Formule d'Erlang-C. On note :

$$C(S, \rho) = \left(\sum_{k=0}^{S-1} \frac{\rho^k}{k!} + \frac{S\rho^S}{S!(S-\rho)} \right)^{-1} \frac{S\rho^S}{S!(S-\rho)}.$$

On peut donc implémenter un algorithme simple de dimensionnement d'une file à plusieurs serveurs étant fixée une contrainte de qualité de service en termes de proportion de clients mis en attente.

Algorithme 5.1 . Dimensionnement d'un système multiserveur avec taux de mise en attente garanti

Données : ρ, p

Résultat : le nombre de serveurs optimal S , étant donné une charge ρ et la garantie d'une proportion p d'utilisateurs mis en attente.

$S \leftarrow 1$;

tant que $p < C(S, \rho)$ **faire**

$S \leftarrow S + 1$;

fin

retourner S

Le nombre moyen de clients dans le système à l'équilibre est donné par :

$$\begin{aligned}
 \mathbf{E} \left[X_{\infty}^{(S)} \right] &= \sum_{i=0}^{\infty} i \pi^{(S)}(i) \\
 &= \pi^{(S)}(0) \left(\sum_{i=0}^S \frac{i \rho^i}{i!} + \frac{S^{S-1} \rho}{S!} \sum_{i=S+1}^{\infty} i \left(\frac{\rho}{S} \right)^{i-1} \right) \\
 &= \pi^{(S)}(0) \rho \left(\sum_{i=0}^{S-1} \frac{\rho^i}{i!} + \frac{S^{S-1}}{S!} \frac{d}{dx} \left(\sum_{i=S+1}^{\infty} x^i \right) \left(\frac{\rho}{S} \right) \right) \\
 &= \pi^{(S)}(0) \rho \left(\sum_{i=0}^{S-1} \frac{\rho^i}{i!} + \frac{\rho^S}{S!} \frac{S^2 - S\rho + S}{(S - \rho)^2} \right).
 \end{aligned}$$

D'après la formule de Little, à l'état stationnaire le temps de séjour $Ts^{(S)}$ et le temps d'attente $TA^{(S)}$ ont donc pour moyennes respectives :

$$\mathbf{E} \left[Ts^{(S)} \right] = \frac{\pi^{(S)}(0)}{\mu} \left(\sum_{i=0}^{S-1} \frac{\rho^i}{i!} + \frac{\rho^S}{S!} \frac{S^2 - S\rho + S}{(S - \rho)^2} \right), \quad (5.10)$$

$$\mathbf{E} \left[TA^{(S)} \right] = \frac{1}{\mu} \left[\pi^{(S)}(0) \left(\sum_{i=0}^{S-1} \frac{\rho^i}{i!} + \frac{\rho^S}{S!} \frac{S^2 - S\rho + S}{(S - \rho)^2} \right) - 1 \right]. \quad (5.11)$$

La formule (5.11) permet de dimensionner un système, étant garanti un temps d'attente moyen.

Algorithme 5.2 . Dimensionnement d'un système multiserveur avec temps d'attente garanti

Données : ρ, μ, Ts

Résultat : le nombre de serveurs optimal S , étant donné une charge ρ et la garantie d'un temps d'attente moyen Ts .

$S \leftarrow 1$;

tant que

$$Ts < \frac{1}{\mu} \left[\left(\sum_{i=0}^{S-1} \frac{\rho^i}{i!} + \frac{S \rho^S}{S! (S - \rho)} \right)^{-1} \left(\sum_{i=0}^{S-1} \frac{\rho^i}{i!} + \frac{\rho^S}{S!} \frac{S^2 - S\rho + S}{(S - \rho)^2} \right) - 1 \right]$$

faire

| $S \leftarrow S + 1$;

fin

retourner S

Par ailleurs, comme dans le cas de la file à un seul serveur, on peut calculer la loi du temps d'attente à travers sa transformée de Laplace.

THÉORÈME 5.6. – Si $\rho < S$, le temps d'attente W_n converge en loi vers $TA^{(S)}$ dont la transformée de Laplace est donnée par :

$$\mathbf{E} \left[e^{-sTA^{(S)}} \right] = 1 - C(S, \rho) + \pi^{(S)}(0) \frac{\rho^S}{S!} \frac{S\mu}{s + (S - \rho)\mu}. \quad (5.12)$$

Ce qui équivaut à dire que :

$$d\mathbf{P}_{TA^{(S)}}(x) = (1 - C(S, \rho))\delta_0(x) + \pi_0 \frac{\rho^S}{S!} S\mu \exp(-(S - \rho)\mu x) \mathbf{1}_{\mathbf{R}^+}(x) dx.$$

▮ Cette dernière équation signifie que le temps d'attente est nul avec probabilité $1 - C(S, \rho)$ (ce qui est évident) et que conditionnellement à ce qu'il soit positif, $TA^{(S)}$ est exponentiellement distribué de paramètre $(S - \rho)\mu$.

Démonstration. Comme les durées de service suivent des lois exponentielles, conditionnellement à ce que tous les serveurs soient pris, les interdéparts sont non seulement tous exponentiellement distribués de paramètres $S\mu$ mais de plus les interdéparts sont indépendants les uns des autres. On a donc :

$$W_n \stackrel{\text{loi}}{=} \sum_{j=1}^{\hat{X}_n - S + 1} \eta_j \mathbf{1}_{\{\hat{X}_n \geq S\}},$$

où $(\eta_k, k \geq 0)$ est une suite de variables aléatoires indépendantes et identiquement distribuées de loi exponentielle de paramètre $S\mu$. Par conséquent :

$$\mathbf{E} \left[e^{-sW_n} \mid \hat{X}_n \right] = \left(\frac{S\mu}{s + S\mu} \right)^{\hat{X}_n - S + 1} \mathbf{1}_{\{\hat{X}_n \geq S\}} + \mathbf{1}_{\{\hat{X}_n < S\}}.$$

D'après le théorème 5.4, \hat{X}_n converge en loi vers X . Comme cette espérance conditionnelle est une fonction bornée de \hat{X}_n , si $\rho < S$, $\mathbf{E} \left[e^{-sW_n} \right]$ converge simplement pour tout s vers :

$$\mathbf{E} \left[\left(\frac{S\mu}{s + S\mu} \right)^{\hat{X}_n - S + 1} \mathbf{1}_{\{\hat{X}_n \geq S\}} + \mathbf{1}_{\{\hat{X}_n < S\}} \right] = \mathbf{E} \left[e^{-sTA^{(S)}} \right].$$

On obtient donc :

$$\begin{aligned}
 \mathbf{E} \left[e^{-s\text{TA}^{(S)}} \right] &= 1 - C(S, \rho) + \sum_{j=S}^{\infty} \left(\frac{\rho}{S} \frac{S\mu}{s + S\mu} \right)^{j-S+1} \\
 &= 1 - C(S, \rho) + \pi^{(S)}(0) \frac{\rho^S}{S!} \frac{S\mu}{s + S\mu} \sum_{j=0}^{\infty} \left(\frac{\rho\mu}{s + S\mu} \right)^j \\
 &= 1 - C(S, \rho) + \pi_0 \frac{\rho^S}{S!} \frac{S\mu}{s + (S - \rho)\mu}.
 \end{aligned}$$

□

5.3.1. Comparaison de systèmes

A partir d'une file d'attente simple, nous comparons ici qualitativement trois types d'opérations visant à en améliorer les performances : le multiplexage des ressources, le parallélisme et le débit d'exécution. On considère les quatre systèmes suivants, tous soumis à des arrivées poissonniennes d'intensité λ , de clients demandant des temps de service de loi $\varepsilon(\mu)$. La charge égale donc toujours $\rho = \frac{\lambda}{\mu}$ Erlang et l'on supposera que $\rho < 2$. On suppose de plus que tous les serveurs considérés travaillent en FCFS et que les salles d'attente sont toujours de capacité illimitée :

- le système A a un serveur travaillant à vitesse unité et correspond donc à une file d'attente simple $M_\lambda/M_\mu/1$;
- dans le système B, on dirige les clients avec probabilité 1/2 (de manière indépendantes entre les clients et indépendantes de toutes les autres variables aléatoires) vers l'un ou l'autre de deux systèmes indépendants fonctionnant en parallèle, chacun à un serveur travaillant à vitesse unité. D'après le théorème 3.6, chacune des deux files en parallèle est donc une file $M_{\frac{\lambda}{2}}/M_\mu/1$;
- le système C est une file d'attente unique à deux serveurs, chacun travaillant à vitesse unité. C est donc exactement une file $M_\lambda/M_\mu/2$;
- le système D a un serveur travaillant à vitesse double : le client C_n qui demande un temps de service de durée σ_n met effectivement un temps $\frac{\sigma_n}{2}$ pour être servi. Il est alors facile de voir que les temps de service des clients sont indépendantes et identiquement distribués de loi $\varepsilon(2\mu)$ et D correspond donc à une file d'attente $M_\lambda/M_{2\mu}/1$.

Dans la suite, nous ajoutons respectivement les exposants A, B, C et D aux caractéristiques des différents systèmes.

D'après (5.5) le temps de séjour moyen d'un client à l'équilibre dans le système D est donné par :

$$\mathbf{E} [\text{Ts}^D] = \frac{1}{2\mu - \lambda}, \quad (5.13)$$

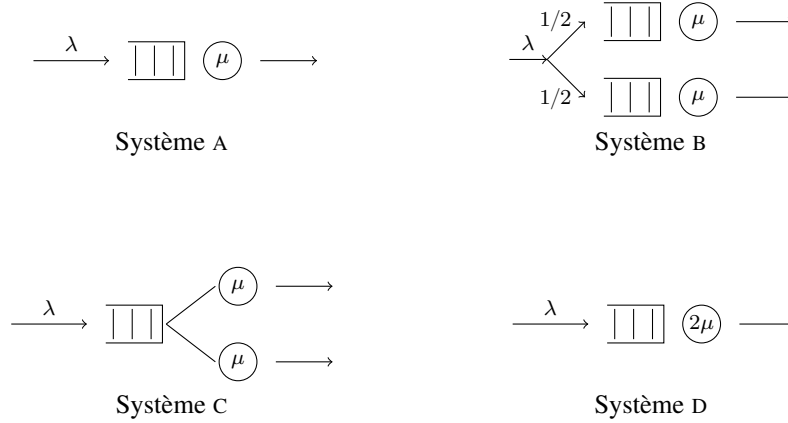


Figure 5.1 – Les quatre systèmes comparés ; la charge est ρ dans A et $\rho/2$ dans les trois autres

et dans le système A, à condition que $\rho < 1$, on a :

$$\mathbf{E}[\text{Ts}^A] = \frac{1}{\mu - \lambda}. \quad (5.14)$$

Considérons maintenant le système B. Soit X^1 et X^2 , les processus comptant respectivement le nombre de clients dans les deux files en parallèle et pour tout t , $X_t^B = X_t^1 + X_t^2$ le nombre de clients dans l'ensemble du système à t . Comme $\rho < 2$, X^1 et X^2 sont ergodiques et en notant X_∞^1 et X_∞^2 leurs limites en loi, d'après (5.3), on a :

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_s^B ds &= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_s^1 ds + \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_s^2 ds \\ &= \mathbf{E}[X_\infty^1] + \mathbf{E}[X_\infty^2] \\ &= \frac{\lambda/2}{\mu - \lambda/2} + \frac{\lambda/2}{\mu - \lambda/2} = \frac{2\lambda}{2\mu - \lambda}. \end{aligned} \quad (5.15)$$

D'après le théorème ??, il existe à la condition $\rho < 2$ un temps d'attente et donc un temps de séjour stationnaire Ts^B pour ce système, vérifiant la convergence presque sûre :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \text{Ts}_n^B = \mathbf{E}[\text{Ts}^B],$$

où pour tout n , Ts_n^B désigne le temps de séjour proposé au n^{e} client. La formule de Little s'applique donc pour ce système. Avec (5.15), il vient :

$$\mathbf{E}[\text{Ts}^B] = \frac{2}{2\mu - \lambda}. \quad (5.16)$$

Finalement, d'après (5.10) et (5.9), on a :

$$\begin{aligned}
 \mathbf{E}[W^c] &= \frac{1}{\mu} \pi(0) \left(1 + \rho + \frac{\rho^2}{2} \frac{4 - 2\rho + 2}{(2 - \rho)^2} \right) \\
 &= \frac{1}{\mu} \frac{2 - \rho}{2 + \rho} \left(\frac{4}{(2 + \rho)(2 - \rho)} \right) \\
 &= \frac{4\mu}{(2\mu - \lambda)(2\mu + \lambda)}.
 \end{aligned} \tag{5.17}$$

Regroupons (5.13), (5.14), (5.16) et (5.17). On retrouve que, dès que $\rho < 2$:

$$\begin{aligned}
 \mathbf{E}[Ts^B] - \mathbf{E}[Ts^C] &= \frac{2\lambda}{(2\mu - \lambda)(2\mu + \lambda)}; \\
 \mathbf{E}[Ts^C] - \mathbf{E}[Ts^D] &= \frac{1}{2\mu + \lambda}.
 \end{aligned}$$

Ainsi :

$$\mathbf{E}[Ts^D] < \mathbf{E}[Ts^C] < \mathbf{E}[Ts^B] < \mathbf{E}[Ts^A]; \tag{5.18}$$

$$\mathbf{E}[Ts^D] < \frac{1}{2} \mathbf{E}[Ts^A], \tag{5.19}$$

où (5.19) et la dernière inégalité de (5.18) n'ont de sens que si de plus $\rho < 1$. Du point de vue du temps passé dans le système, il est donc préférable de doubler la vitesse d'exécution que de doubler le nombre de ressources, ce qui est à son tour plus performant qu'une mise en parallèle de deux systèmes. En outre, d'après (5.19) on fait mieux que doubler les performances d'un système à un serveur en doublant sa vitesse d'exécution.

5.4. La file *processor sharing*

Nous abordons succinctement le cas de la file markovienne $M_\lambda/M_\mu/PS$, étudiée dans le cas général en section ?? . Les clients entrent suivant un processus de Poisson d'intensité λ , demandent des temps de service de loi $\varepsilon(\mu)$ et sont immédiatement servis par un serveur selon la discipline *processor sharing* : tous sont servis simultanément, à une vitesse inversement proportionnelle au nombre de clients présents.

Notons X^{PS} , le processus comptant le nombre de clients dans le système. Il est facile de se convaincre que ce processus est Markovien et de calculer son générateur. Pour tout $t \geq 0$, notons comme précédemment :

- $W(t)$ le temps résiduel avant la prochaine arrivée après t ;
- $R(t)$, le temps résiduel avant le prochain départ après t .

Soit $i \geq 1$. Sur $\{X^{\text{PS}}(t) = i\}$, notons $R_1(t), R_2(t), \dots, R_i(t)$, les temps de service résiduels des i clients à t , comptés en unités de temps. D'après le paradoxe de l'autobus, ces temps de services sont indépendants et de loi $\varepsilon(\mu)$. Tant qu'il n'y a pas de nouvelle arrivée après t (c'est-à-dire jusqu'à $t + W(t)$), le serveur travaille à vitesse $1/i$ (ce qui multiplie l'échelle de temps par ce facteur) et donc du point de vue du serveur les temps de service résiduels des clients suivent la loi $\varepsilon(\mu/i)$ et alors :

– si $\min_{j=1, \dots, i} R_j(t) \leq W(t)$, le serveur travaille à la même vitesse jusqu'au prochain départ, qui a lieu avant la prochaine arrivée et donc $R(t)$ est le minimum de i variables aléatoires indépendantes de loi $\varepsilon(\mu/i)$;

– si $W(t) < \min_{j=1, \dots, i} R_j(t)$, à l'instant $t + W(t)$ il y a une nouvelle arrivée et donc $i + 1$ clients dans le système. Toujours d'après le paradoxe de l'autobus, la loi des caractéristiques du système ne change pas si l'on tire au sort de nouveau les temps de services des $i + 1$ clients, suivant la loi $\varepsilon(\mu)$. Alors, $R(t)$ a la même loi que le minimum de $i + 1$ variables aléatoires indépendantes de loi $\varepsilon(\mu/(i + 1))$, à moins qu'il n'y ait une autre arrivée avant le premier départ. Le cas échéant, on tire de nouveau au sort les $i + 2$ temps de service suivant la même loi $\varepsilon(\mu)$, etc.

En conclusion, dans tous les cas la loi de $R(t)$ suit conditionnellement à $\{X^{\text{PS}}(t) = i\}$, la loi $\varepsilon(\mu)$. Comme $R(t)$ suit pour tout t la loi $\varepsilon(\lambda)$, comme pour les systèmes précédent, on obtient le théorème suivant.

THÉORÈME 5.7.— Le processus X^{PS} a la même loi que le processus de congestion X de la file $M_\lambda/M_\mu/1$. Il admet donc la même loi stationnaire, à condition que $\rho < 1$. En particulier, d'après la formule de Little le temps de séjour moyen à l'équilibre est le même que celui de la file $M/M/1$.

⚠ Attention, cela ne signifie pas pour autant que les processus X et X^{PS} ont presque sûrement la même trajectoire ! Cela n'est pas le cas, l'identité précédente n'est vraie qu'en loi.

5.5. La file $M/M/\infty$

La file $M/M/\infty$ est évidemment une chimère théorique puisqu'aucun système n'aura un nombre infini de serveurs. En revanche, il existe de nombreuses situations où cet objet apparaît, par exemple dans le théorème ??.

Soit $(X^\infty(t), t \geq 0)$ le processus comptant le nombre de clients dans le système (et donc en service). Comme précédemment, on vérifie aisément pour tout $i \in \mathbb{N}$ que :

- le processus $(X^\infty(t), t \geq 0)$ séjourne en i durant un temps de loi $\varepsilon(i\mu + \lambda)$;
- il quitte l'état i pour l'état $i + 1$ avec probabilité $\lambda/(\lambda + i\mu)$ et si $i \geq 1$, pour l'état $i - 1$ avec probabilité $\lambda/(\lambda + i\mu)$.

Le processus $(X^\infty(t), t \geq 0)$ est donc Markovien, de générateur infinitésimal A^∞ donné par :

$$\begin{pmatrix} -\lambda & \lambda & & & & \\ \mu & -(\lambda + \mu) & \lambda & & & \\ 0 & 2\mu & -(\lambda + 2\mu) & \lambda & & \\ & & \ddots & \ddots & \ddots & \\ (0) & & & i\mu & -(\lambda + i\mu) & \lambda \\ & & & (i+1)\mu & -(\lambda + (i+1)\mu) & \lambda \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}. \quad (0)$$

En revanche, la théorie développée dans le chapitre 4 ne s'applique pas telle quelle car les coefficients du générateur infinitésimal ne sont pas bornés. En utilisant une représentation totalement différente de cette file (voir l'exemple ??), on peut toutefois en déduire que l'unique probabilité stationnaire π^∞ de $(X^\infty(t), t \geq 0)$ vérifie :

$$\pi^\infty(i) = \frac{\rho^i e^{-\rho}}{i!},$$

où comme précédemment, $\rho = \lambda/\mu$, ce qui signifie en d'autres termes que la congestion limite X_∞^∞ suit une loi de Poisson de paramètre ρ . En particulier, la congestion moyenne du système à l'équilibre vaut :

$$\mathbf{E}[X_\infty^\infty] = \rho. \quad (5.20)$$

5.6. Le processus des départs

Nous considérons ici l'un des trois modèles considérés dans les sections précédentes (file à un, plusieurs ou une infinité de serveur). Nous voyons ici qu'à la condition de stabilité, la loi du processus des départs est identifiable. Rappelons que l'on note pour tout $n \in \mathbf{N}$, T_n l'instant de la n^e arrivée et T'_n l'instant du n^e départ. On note encore $(A(t), t \geq 0)$ le processus des arrivées et $(D(t), t \geq 0)$, le processus de comptage des départs :

$$D(t) = \sum_{n \in \mathbf{N}} \mathbf{1}_{T'_n \leq t}.$$

A première vue, l'on pourrait penser que le processus des départs est tel que le temps entre deux départs est de loi $\varepsilon(\mu)$. Nous allons voir que cela n'est pas le cas. Commentons par une heuristique sur la file M/M/1 pour réfuter cette idée. Plaçons-nous à un instant de sortie de la file, disons T'_n l'instant du départ de C_n . Alors, T'_{n+1} égale :

- $T'_n + \sigma_{n+1}$ si C_{n+1} est dans la file au départ de C_n ;
- $T_{n+1} + \sigma_{n+1}$, si C_n laisse un système vide derrière lui.

Donc, si l'on admet que $T'_{n+1} - T'_n$ est indépendant du passé de T'_n (une propriété qui sera démontrée par la suite), on a :

$$\begin{aligned}\mathbf{E}[T'_{n+1} - T'_n] &= \mathbf{E}[\sigma_{n+1}] + \mathbf{E}[(T_{n+1} - T'_n) \mathbf{1}_{X(T'_n)=0}] \\ &= \frac{1}{\mu} + \frac{1}{\lambda} \mathbf{P}(X(T'_n) = 0),\end{aligned}$$

d'après le paradoxe de l'autobus. Ainsi, à l'équilibre, le temps moyen entre deux départs vaut :

$$\frac{1}{\mu} + \frac{1}{\lambda} \pi(0) = \frac{1}{\lambda},$$

autant que le temps moyen entre deux arrivées !

Le résultat suivant précise ce résultat et montre qu'à l'équilibre le processus des départs est en fait, comme celui des arrivées, un processus de Poisson d'intensité λ .

THÉORÈME 5.8 (Théorème de Burke).— A l'état stationnaire, le processus des départs est un processus de Poisson d'intensité λ . De plus, pour tout $s \geq 0$, les départs après s sont indépendants des arrivées avant s , autrement dit, $\sigma(D_u; u \geq s)$ est indépendant de $\sigma(A(u); u \leq s)$.

Démonstration. Dans tous les cas, le processus de congestion $(X(t), t \geq 0)$ est de naissance et de mort. Il est par conséquent réversible, de probabilité stationnaire π . En particulier, en supposant que X est issu de la loi π , il a pour tout $T \geq 0$ la même loi que son processus retourné $(\bar{X}^T(t), t \geq 0)$.

Il est clair que pour tout $n \in \mathbf{N}$, T'_n (respectivement T_n) est le n^{e} instant de saut vers le bas (respectivement vers le haut) du processus X . Par ailleurs, pour tout $T \geq 0$ les instants de sauts vers le bas de X jusqu'à T correspondent aux instants de sauts vers le haut du processus retourné : le processus original descend de 1 quand le processus retourné remonte d'autant. Plus précisément, si l'on note $\tilde{T}'_n = T'_n \wedge T$, $n \in \mathbf{N}$ les points du processus $(A(t), t \geq 0)$ restreint à $[0, T]$, on a pour commencer :

$$\begin{aligned}\tilde{T}'_1 &= \inf \{t \leq T; X(t) = X(t^-) - 1\} \\ &= T - \sup \{s \leq T; X(T-s) = X((T-s)^-) - 1\} \\ &= T - \sup \{s \leq T; \bar{X}^T(s) = \bar{X}^T(s^-) + 1\}.\end{aligned}$$

Le dernier *supremum* est le dernier instant de saut vers le haut du processus retourné avant $T \geq 0$. Comme X est réversible, ce *supremum* égale en loi le temps écoulé depuis le dernier instant de saut vers le haut de X (c'est-à-dire la dernière arrivée) avant $T \geq 0$. D'après le théorème 4.3, on a donc l'égalité en loi suivante :

$$\tilde{T}'_1 \stackrel{\mathcal{L}}{=} T - T_{A(T)} \stackrel{\mathcal{L}}{=} T \wedge Y_1,$$

où Y_1 est une variable aléatoire de loi $\varepsilon(\lambda)$. De la même façon, on montre que :

$$\tilde{T}'_2 - \tilde{T}'_1 \stackrel{\mathcal{L}}{=} (T - \tilde{T}'_1) \wedge Y_2,$$

où Y_2 est indépendante de Y_1 et de loi $\varepsilon(\lambda)$ et ainsi de suite.

Ceci montre que $(A(t), t \geq 0)$ égale en loi un processus de Poisson sur $[0, T]$. Comme ceci est vrai pour tout $T \geq 0$, $(A(t), t \geq 0)$ est bien un processus de Poisson.

Finalement, la propriété d'indépendance découle, naturellement, de celle du processus des arrivées. \square

5.7. Réseaux de files d'attente

Dans cette section, nous présentons succinctement les principaux résultats sur le modèle le plus simple de réseau de files d'attente : un ensemble de files d'attente dans lequel les clients qui sortent d'une file sont susceptibles d'en rejoindre une autre pour recevoir un nouveau service.

5.7.1. Réseaux de Jackson ouverts

On considère le système suivant :

- un processus de Poisson de paramètre $\lambda > 0$, qui est interprété comme le processus des arrivées de l'extérieur dans le système ;
- un ensemble de N files d'attente, où la i^{e} file est de type $M_{\mu_i}/S_i/\infty$ -FIFO : les clients qui entrent dans la i^{e} file demandent des temps de service indépendants et identiquement distribués de loi $\varepsilon(\mu_i)$ à un groupe de S_i serveurs. La loi des interarrivées dans la file n'est pas connue, car elle dépend *a priori* des autres files, comme on va le voir ;
- une matrice markovienne P de taille $N + 1$: pour tout $i \in \llbracket 0, N \rrbracket$, $P_{ij} \in [0, 1]$ et $\sum_{j=0}^N P_{ij} = 1$. La matrice P est appelée matrice de routage du système : dès qu'un client a fini son service dans la file i , il fait un tirage au sort indépendant de tous les autres paramètres pour décider de la prochaine file dans laquelle il va demander un service. Pour tout $j \in \llbracket 0, N \rrbracket$, il rejoint alors la file j avec probabilité P_{ij} . La « file 0 » représente ici l'extérieur du système : les clients faisant le chemin de 0 vers j arrivent directement de l'extérieur dans la file j et ceux qui vont de i à 0 sortent du système après avoir visité la file i . Notons que si $P_{ii} > 0$ pour un certain i , un client est susceptible de se remettre en queue dans la file i juste après avoir reçu un service dans cette même file. On suppose que P vérifie les deux conditions :

$$P_{00} = 0, \tag{5.21}$$

et pour tout $i \in \llbracket 1, N \rrbracket$, il existe $n \in \mathbb{N}$ et un n -uplet $\{i_1, i_2, \dots, i_n\}$ d'éléments de $\llbracket 1, N \rrbracket$ contenant i et tel que :

$$P_{0i_1} P_{i_1 i_2} \dots P_{i_{n-1} i_n} P_{i_n 0} > 0. \quad (5.22)$$

La condition (5.22) garantit donc que toute file i fait partie d'un chemin possible, allant de l'extérieur à l'extérieur.

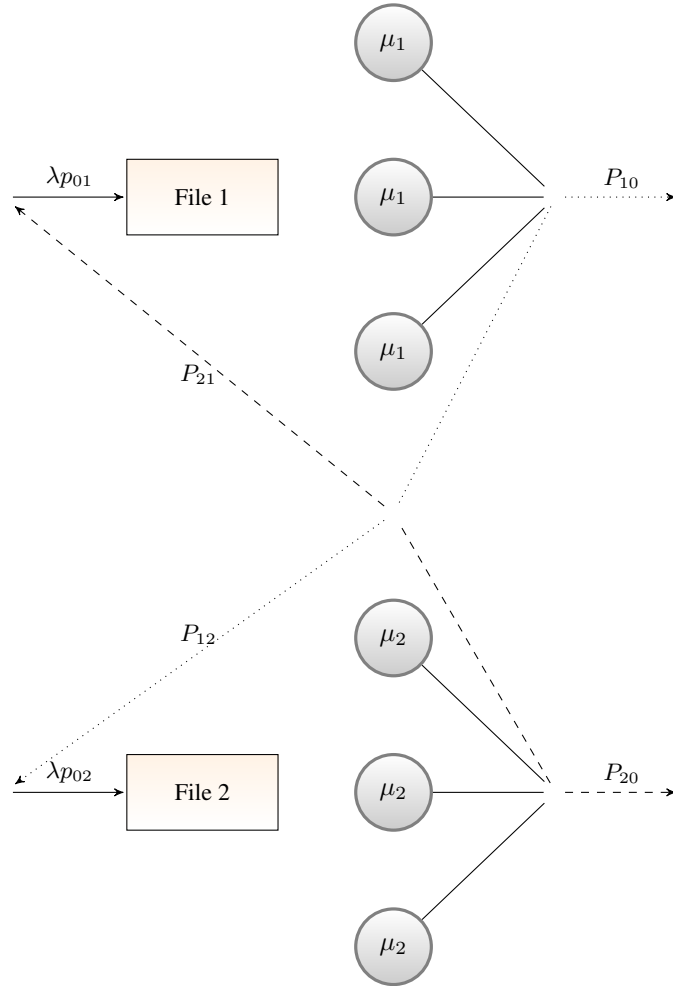


Figure 5.2 – Un réseau de Jackson à deux files

Le système est décrit par le processus :

$$(X(t), t \geq 0) = ((X_1(t), X_2(t), \dots, X_N(t)), t \geq 0),$$

où pour tout i , $X_i(t)$ compte le nombre de clients dans la i^{e} file à l'instant t . Comme le trafic dans chaque file dépend de l'état des autres files, il est facile de voir que chaque processus $(X_i(t), t \geq 0)$ n'est pas markovien. C'est le cas en revanche du processus $(X(t), t \geq 0)$, comme nous le montrons dans le lemme suivant. On appelle (voir les notations dans l'annexe A) pour tout $i \in \mathbf{N}$, $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ et pour tout $k \in \mathbf{N}$, $\mu_i(k) = \mu_i(k \wedge S_i)$.

LEMME 5.9.— Le processus $(X(t), t \geq 0)$ décrivant le réseau de Jackson ouvert est markovien de générateur infinitésimal A^{j0} défini pour tout $x = (x(1), \dots, x(N))$ par :

$$\begin{cases} A^{j0}(x, x + \mathbf{e}_j - \mathbf{e}_i) &= \mu_i(x(i))P_{ij}; \\ A^{j0}(x, x + \mathbf{e}_j) &= \lambda P_{0j}; \\ A^{j0}(x, x - \mathbf{e}_i) &= \mu_i(x(i))P_{i0}, \end{cases},$$

tous les autres coefficients $A^{j0}(x, y)$ sont nuls et les coefficients diagonaux $A^{j0}(x, x)$ égalent l'opposée de la somme des $A^{j0}(x, y)$ pour $y \neq x$.

Démonstration. Supposons que le processus $(X(t), t \geq 0)$ se trouve dans l'état x à t . Alors, le processus peut quitter directement x seulement pour les états suivants :

- $x + \mathbf{e}_j - \mathbf{e}_i$ si $x(i) > 0$ et si l'un des clients en service dans la file i termine son service, puis se rend dans la file j ;
- $x - \mathbf{e}_i$ si $x(i) > 0$ et si l'un des clients en service dans la file i termine son service, puis sort du système ;
- $x + \mathbf{e}_j$ si un client entre de l'extérieur vers la file j .

D'après le théorème 3.6, le processus d'entrée de l'extérieur vers la file j est Poissonien d'intensité λP_{0j} . Le temps résiduel avant le prochain point de ce processus suit donc la loi $\varepsilon(\lambda P_{0j})$. Ensuite, le temps résiduel avant la prochaine fin de service de chacun des $x(i) \wedge S_i$ services en cours dans la file i suit, comme vu précédemment, la loi $\varepsilon((x(i) \wedge S_i)\mu_i) = \varepsilon(\mu_i(x(i)))$. En notant pour tout i l'événement :

$$B_i = \{\text{Le premier service qui se termine est un service de la file } i\},$$

la probabilité que l'on quitte effectivement x pour $x - \mathbf{e}_i + \mathbf{e}_j$ (respectivement $x - \mathbf{e}_i$) est donc donnée par $\mathbf{P}(B_i \cap \{\text{le client quitte } i \text{ pour } j\}) = \mathbf{P}(B_i)P_{ij}$ (respectivement $\mathbf{P}(B_i)P_{i0}$). Ceci conclut la preuve d'après les propriétés des lois exponentielles vues précédemment. \square

Nous aurons besoin du résultat technique suivant pour caractériser l'état stationnaire de $(X(t), t \geq 0)$.

LEMME 5.10.– Le système suivant :

$$\lambda_j = \lambda P_{0j} + \sum_{i=1}^N \lambda_i P_{ij}, \quad (5.23)$$

d'inconnues $(\lambda_1, \lambda_2, \dots, \lambda_N)$ et appelé équation du trafic, admet une unique solution dans $(\mathbf{R}^+)^N$.

Démonstration. Comme la matrice P est markovienne, il existe naturellement une unique chaîne de Markov $(M_n, n \in \mathbf{N})$ à valeurs dans $\llbracket 0, N \rrbracket$ et de matrice de transition P . Pour tout couple (i, j) d'éléments de $\llbracket 1, N \rrbracket$, il existe d'après (5.22) deux familles finies $\{i_1, \dots, i_n\}$ et $\{j_1, \dots, j_p\}$ d'éléments de $\llbracket 1, N \rrbracket$, comprenant respectivement i et j , telles que :

$$P_{0i_1} \dots P_{i_{n-1}i_n} P_{i_n 0} > 0 \text{ et } P_{0j_1} \dots P_{j_{p-1}j_p} P_{j_p 0} > 0.$$

Donc, avec les notations du chapitre 2, il existe un entier $q < n + p$ tel que la probabilité que $(M_n, n \in \mathbf{N})$ passe de i à j en q pas vérifie :

$$p^{(q)}(i, j) \geq P_{0i_1} \dots P_{i_{n-1}i_n} P_{i_n 0} P_{0j_1} \dots P_{j_{p-1}j_p} P_{j_p 0} > 0.$$

La chaîne $(M_n, n \in \mathbf{N})$ est donc irréductible sur l'espace d'état fini $\llbracket 0, N \rrbracket$. Elle est donc récurrente positive et d'après le théorème 2.15 il existe à un coefficient multiplicatif près une unique mesure stationnaire ν sur $\llbracket 0, N \rrbracket$ qui représentée en ligne, vérifie l'équation matricielle :

$$\nu = \nu P. \quad (5.24)$$

Par ailleurs, si $\lambda \{\lambda_1, \dots, \lambda_N\}$ est solution de (5.23), comme P est markovienne, d'après (5.21), on a :

$$\begin{aligned} \sum_{i=1}^N \lambda_i P_{i0} &= \sum_{i=1}^N \lambda_i \left(1 - \sum_{j=1}^N P_{ij} \right) \\ &= \sum_{i=1}^N \lambda_i - \sum_{j=1}^N \sum_{i=1}^N \lambda_i P_{ij} \\ &= \sum_{i=1}^N \lambda_i - \sum_{j=1}^N (\lambda_j - \lambda P_{0j}) \\ &= \lambda \sum_{i=1}^N P_{0i} \\ &= \lambda. \end{aligned} \quad (5.25)$$

Ceci montre bien que $\{\lambda_1, \dots, \lambda_N\}$ est solution de (5.23) si et seulement si $\Lambda = (\lambda, \lambda_1, \dots, \lambda_N)$ est solution de (5.24). L'unique solution de (5.23) est donc l'unique mesure invariante de première composante λ . \square

Pour tout $i \in \llbracket 1, N \rrbracket$, à la condition que $\lambda_i < \mu_i S_i$ le processus de congestion de la file $M_{\lambda_i}/M_{\mu_i}/S_i$ admet la probabilité invariante π^i suivante :

$$\pi^i(0) = \left(\sum_{k=0}^{\infty} \frac{(\lambda_i)^k}{\prod_{j=1}^k \mu_i(j)} \right)^{-1}; \quad (5.26)$$

$$\pi^i(k) = \frac{(\lambda_i)^k}{\prod_{j=1}^k \mu_i(j)} \pi^i(0), \quad i \geq 1. \quad (5.27)$$

Nous pouvons donc énoncer le résultat suivant.

THÉORÈME 5.11.— Soit $\{\lambda_1, \dots, \lambda_N\}$ l'unique solution de l'équation du trafic (5.23). On suppose que la condition de stabilité :

$$\lambda_i < \mu_i C_i, \text{ pour tout } i = 1, \dots, N, \quad (5.28)$$

est satisfaite. Soit π^i , la mesure de probabilité sur \mathbf{N} définie par (5.26) et (5.27). Alors, le processus $(X(t), t \geq 0)$ décrivant le réseau de Jackson ouvert admet sur \mathbf{N}^N l'unique probabilité stationnaire π^{JO} définie pour tout $x = (x(1), \dots, x(N)) \in \mathbf{N}^N$ par :

$$\pi^{\text{JO}}(x(1), \dots, x(N)) = \prod_{i=1}^N \pi^i(x(i)). \quad (5.29)$$

Démonstration. Dans ce qui suit, on fait la convention que $\bar{A}_{xy} = A_{xy}^{\text{JO}} = 0$ si l'une des composantes de y est négative. Pour tout x et $y \in \mathbf{N}^N$, définissons :

$$\bar{A}(x, x) = A^{\text{JO}}(x, x) \text{ et } \bar{A}(x, y) = \frac{\pi^{\text{JO}}(x) A^{\text{JO}}(x, y)}{\pi^{\text{JO}}(y)} \text{ pour } x \neq y,$$

où π^{JO} est la probabilité définie par (5.29) et A^{JO} est le générateur infinitésimal défini dans le lemme 5.9. Pour tout i et j , on a :

$$\bar{A}(x, x + \mathbf{e}_i) = \lambda_i P_{i0}$$

$$\bar{A}(x, x - \mathbf{e}_j) = \frac{\mu_j(x(j))}{\lambda_j} \lambda P_{0j} \text{ pour tout } x \text{ tel que } x(j) \geq 1,$$

ainsi que :

$$\begin{aligned}
 \bar{A}(x, x - \mathbf{e}_j + \mathbf{e}_i) &= \frac{\pi^{\text{IO}}(x - \mathbf{e}_j + \mathbf{e}_i) A^{\text{IO}}(x - \mathbf{e}_j + \mathbf{e}_i, x)}{\pi^{\text{IO}}(x)} \\
 &= \frac{\pi^i(x(j) - 1) \pi^i(x(i) + 1)}{\pi^j(x(j)) \pi^i(x(i))} \mu_i(x(i) + 1) P_{ij} \\
 &= \frac{\mu_j(x(j))}{\lambda_j} \frac{\lambda_i}{\mu_i(x(i) + 1)} \mu_i(x(i) + 1) P_{ij} \\
 &= \frac{\lambda_i}{\lambda_j} \mu_j(x(j)) P_{ij},
 \end{aligned}$$

pour tout i et j tels que $i \neq j$ et tout x tel que $x(j) \geq 1$. Pour tout $x \in \mathbf{N}^N$, formons les sommes suivantes :

$$\begin{aligned}
 \sum_{y \neq x} \bar{A}(x, y) &= \sum_{j=1}^N \sum_{i=1; i \neq j}^N \bar{A}(x, x - \mathbf{e}_j + \mathbf{e}_i) + \sum_{i=1}^N \bar{A}(x, x + \mathbf{e}_i) + \sum_{j=1}^N \bar{A}(x, x - \mathbf{e}_j) \\
 &= \sum_{j=1}^N \left(\frac{\mu_j(x(j))}{\lambda_j} \lambda P_{0j} + \lambda_i \sum_{i=1; i \neq j}^N \frac{\mu_j(x(j))}{\lambda_j} P_{ij} \right) + \sum_{i=1}^N \lambda_i P_{i0} \\
 &= \sum_{j=1}^N \frac{\mu_j(x(j))}{\lambda_j} (\lambda_j - \lambda_j P_{jj}) + \sum_{i=1}^N \lambda_i P_{i0} \\
 &= \sum_{j=1}^N \mu_j(x(j)) (1 - P_{jj}) + \lambda,
 \end{aligned}$$

où l'avant-dernière égalité est une conséquence de l'équation du trafic et la dernière découle de (5.25). D'autre part :

$$\begin{aligned}
 \sum_{y \neq x} A^{\text{RJ}}(x, y) &= \sum_{i=1}^N \sum_{j=1; j \neq i}^N A^{\text{RJ}}(x, x - \mathbf{e}_j + \mathbf{e}_i) + \sum_{i=1}^N A^{\text{RJ}}(x, x - \mathbf{e}_i) + \sum_{j=1}^N A^{\text{RJ}}(x, x + \mathbf{e}_j) \\
 &= \sum_{i=1}^N \mu_i(x(i)) (1 - P_{ii}) + \sum_{j=1}^N \lambda P_{0j} = \sum_{i=1}^N \mu_i(x(i)) (1 - P_{ii}) + \lambda = \sum_{y \neq x} \bar{A}(x, y).
 \end{aligned}$$

On a donc $\bar{A}(x, x) = A^{\text{RJ}}(x, x)$ et l'on conclut par réversibilité (voir théorème 4.19).

□

Le théorème précédent, grand classique de la théorie des files d'attente, est appelé théorème de la forme produit : si le réseau de Jackson ouvert est stable, il se comporte à l'équilibre comme un système de N files d'attente indépendantes, où la i^{e} file est une file $M_{\lambda_i}/M_{\mu_i}/S_i$. Ce résultat, assez contre-intuitif, a un intérêt évident pour la simulation, car il indique que l'étude à l'équilibre d'un réseau de Jackson se ramène à celle de files d'attente simples à plusieurs serveurs. A l'équilibre, tout se passe donc comme si les N files fonctionnaient indépendamment et comme des files markoviennes classiques (alors que, rappelons-le, chaque file considérée seule n'est pas une file M/M). Le théorème de Burke appliqué à chacune de N files $M_{\lambda_i}/M_{\mu_i}/S_i$; $i = 1, \dots, N$ donnerait que le débit de sortie de chaque file égale λ_i et donc le débit d'entrée dans la file j soit donné par la quantité :

$$\lambda P_{0j} + \sum_{i=1}^N \lambda_i,$$

soit λ_j d'après l'équation du trafic : le résultat est cohérent.

A tout instant t , conditionnellement à ce que $X(t) = (x_1, x_2, \dots, x_N)$, on peut montrer par les techniques de calcul habituelles sur la loi exponentielle, que le temps résiduel avant la première sortie de la file i vers l'extérieur après t suit la loi $\varepsilon(\mu_i(x(i))P_{i0})$. Par conséquent, le temps résiduel avant la prochaine sortie du réseau pris dans sa globalité vers l'extérieur suit la loi $\varepsilon\left(\sum_{i=1}^N \mu_i(x(i))P_{i0}\right)$. Il est donc naturel de définir le débit de sortie instantané à t par la variable aléatoire :

$$D(t) = \sum_{i=1}^N \mu_i(X_i(t)) P_{i0}.$$

Le débit moyen de sortie à l'équilibre est donc donné par :

$$\begin{aligned} \mathbf{E}[D(\infty)] &= \mathbf{E}\left[\sum_{i=1}^N \mu_i(X_i(\infty)) P_{i0}\right] \\ &= \sum_{(x(1), \dots, x(N)) \in \mathbf{N}^N} \sum_{i=1}^N \mu_i(x(i)) P_{i0} \pi^{\text{RJ}}(x(1), \dots, x(N)), \end{aligned}$$

où $X(\infty) = (X_1(\infty), \dots, X_N(\infty))$ est la variable aléatoire distribuée sur \mathbf{N}^N suivant la loi π^{RJ} . On a alors l'analogie suivant du théorème de Burke.

THÉORÈME 5.12.— Dans un réseau de Jackson ouvert à l'équilibre, le débit moyen de sortie égale l'intensité du processus d'entrée :

$$\mathbf{E}[D(\infty)] = \lambda.$$

5.7.2. Réseaux de Jackson fermés

Le réseau de Jackson fermé est l'analogue du réseau ouvert, en supposant maintenant que le réseau n'est pas « nourri » par un processus de Poisson exogène, et qu'aucune file n'est reliée à l'extérieur : K clients (où K est fixé) circulent donc indéfiniment dans un réseau de N files M_{μ_i}/S_i , suivant les mêmes modalités que dans le réseau ouvert. Le système est donc totalement décrit par :

- la taille K de la population du réseau ;
- les N files M_{μ_i}/S_i ;
- une matrice de routage P , markovienne et de taille N satisfaisant une propriété d'irréductibilité analogue à (5.22) : pour tout i et $j \in \llbracket 1, N \rrbracket$, il existe $n \in \mathbf{N}$ et un n -uplet $\{i_1, i_2, \dots, i_n\}$ d'élément de $\llbracket 1, N \rrbracket$ contenant i et j et tel que :

$$P_{i_1} P_{i_1 i_2} \dots P_{i_{n-1} i_n} > 0. \quad (5.30)$$

Soit l'ensemble :

$$\mathcal{A} = \left\{ x = (x(1), \dots, x(N)) \in \mathbf{N}^N; \sum_{i=1}^N x(i) = K \right\}.$$

Le processus $(X(t), t \geq 0)$, défini comme dans la section précédente, est markovien sur \mathcal{A} , de générateur défini pour tout x par :

$$A^{\text{JF}}(x, x + \mathbf{e}_j - \mathbf{e}_i) = \mu_i(x(i))P_{ij}; \text{ si } x(i) \geq 1,$$

où tous les autres termes sont nuls sauf le terme diagonal, qui est l'opposé de la somme des autres termes de la même ligne.

Comme précédemment, l'introduction d'une chaîne de Markov sur $\{1, \dots, N\}$ de matrice de transition P permet de conclure qu'il existe, à un coefficient multiplicatif près, une unique solution $\lambda \in (\mathbf{R}^+)^N$ à l'équation matricielle :

$$\lambda = \lambda P. \quad (5.31)$$

On a alors le résultat suivant.

THÉORÈME 5.13.— Le processus $(X(t), t \geq 0)$ décrivant le réseau de Jackson fermé admet une unique probabilité stationnaire π^{JF} , définie pour tout $x \in \mathcal{A}$ par :

$$\pi^{\text{JF}}(x(1), \dots, x(N)) = C \prod_{i=1}^N \prod_{\ell=1}^{x(i)} \frac{\lambda_i}{\mu_i(\ell)},$$

où $\lambda = (\lambda_1, \dots, \lambda_N)$ est une solution quelconque de (5.31) et C est la constante de normalisation :

$$C = \left(\sum_{(x(1), \dots, x(N)) \in \mathcal{A}} \prod_{i=1}^N \prod_{\ell=1}^{x(i)} \frac{\lambda_i}{\mu_i(\ell)} \right)^{-1}.$$

Un réseau de Jackson fermé est donc beaucoup moins confortable à étudier que dans le cas ouvert. On peut en effet remarquer que, au contraire du cas ouvert, la forme précédente n'est pas une forme produit à cause de l'expression de la constante C . De plus, le calcul numérique de cette constante pour un grand réseau est très coûteux algorithmiquement.

5.8. Problèmes

EXERCICE 16.— On considère la file M/M/1 avec la discipline de service suivante : une proportion p des clients (les « polis ») se place normalement en queue, tandis qu'une proportion $q = 1 - p$ « d'impolis » double tout le monde et prend la première place de la file d'attente. La catégorie (« polis » ou « impolis ») d'un client est indépendante de son instant d'arrivée. Cette discipline est non préemptive, c'est-à-dire que l'on n'interrompt pas le service en cours. L'intensité du processus d'arrivée global est λ , le temps moyen de service est $1/\mu$.

On note $(X(t), t \geq 0)$ le nombre de clients dans le système (file + serveur) à l'instant t .

- 1) Donner le générateur infinitésimal de X . Est-il différent du générateur infinitésimal de la file M/M/1 avec discipline FIFO (*First In First Out*) ?
- 2) En déduire la condition de stabilité du système et la distribution à l'état stationnaire de X .
- 3) Quel est le temps moyen d'attente à l'état stationnaire ?
- 4) Quelles sont la nature et l'intensité du processus d'arrivée des clients « impolis » ?
- 5) On note W^p le temps d'attente d'un client poli donné, on note X le nombre de clients dans le système à son arrivée et $N(W^p)$ le nombre de clients « impolis » qui arrivent pendant son temps d'attente. Montrer qu'en loi, on a :

$$W^p \stackrel{\text{loi}}{=} \sum_{j=1}^X \eta_j + \sum_{l=1}^{N(W^p)} \sigma_l, \quad (5.32)$$

où $(\eta_j, j \geq 0)$ et $(\sigma_l, l \geq 0)$ sont deux suites indépendantes entre elles et indépendantes de X , de variables aléatoires indépendantes de loi exponentielle de paramètre μ .

- 6) Expliquer pourquoi X et $N(W^p)$ sont indépendantes conditionnellement à W^p .
- 7) On suppose dorénavant que X a la loi stationnaire identifiée dans la deuxième question, montrer que :

$$\mathbf{E} \left[\sum_{j=1}^X \eta_j \right] = \frac{\rho}{(1 - \rho)\mu}.$$

8) Montrer que :

$$\mathbf{E} \left[e^{-s \sum_{l=1}^{N(W^p)} \sigma_l} \mid W^p = v \right] = e^{-\lambda q v \frac{s}{\mu+s}}.$$

9) En déduire que si X a la loi stationnaire identifiée dans la deuxième question, on a :

$$\mathbf{E} \left[e^{-\lambda q W^p \frac{s}{\mu+s}} \mathbf{E} \left[e^{-s \sum_{j=1}^X \eta_j} \mid W^p \right] \right].$$

(On ne demande pas de calculer explicitement $\mathbf{E} [e^{-s W^p}]$.)

10) En déduire, par dérivation, le temps moyen d'attente d'un client « poli ».

11) On note W^i le temps d'attente d'un client « impoli » à l'état stationnaire et W le temps d'attente d'un client quelconque à l'état stationnaire – on considère prouvé que le temps d'attente du n^{e} client impoli converge en loi vers W^i , de même pour W . Expliquer pourquoi l'on a :

$$\mathbf{E} [W] = p \mathbf{E} [W^p] + q \mathbf{E} [W^i].$$

12) En déduire le temps moyen d'attente d'un client « impoli ».

EXERCICE 17.– On considère que l'on a une station d'essence à deux pompes et un emplacement d'attente :

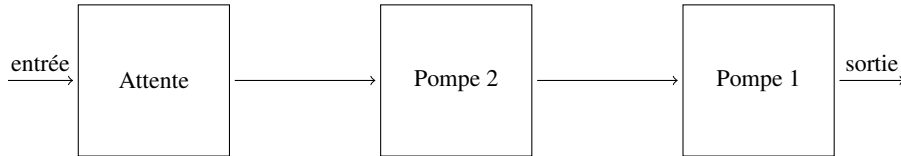


Figure 5.3 – Station d'essence à trois pompes en série

Les clients arrivent à la pompe selon un processus de Poisson d'intensité λ . Si les deux pompes sont libres, le client va en pompe 1. Si la pompe 1 est prise, le client va en pompe 2. Si les deux pompes sont occupées, le client peut se mettre dans la place d'attente. Si la place d'attente est occupée, le client passe son chemin. Un client en pompe 2 doit attendre que la pompe 1 soit libre pour sortir de la station d'essence. Dans le cas où le client à la pompe 2 finit avant celui de la pompe 1, les deux sortent en même temps de la station au moment où le client de la pompe 1 a fini. S'il y a, à ce moment-là, un client en attente, il va instantanément en pompe 1 et commence à remplir son réservoir. Dans le cas où la pompe 1 est libre mais la pompe 2 occupée, personne ne peut entrer en pompe 1. Les temps de remplissage des réservoirs suivent une loi exponentielle de paramètre μ . On modélise ce système par un processus de Markov $X = (X_1, X_2, X_3)$ où X_1 vaut 0 ou 1 et représente le nombre de clients à la

pompe 1, X_2 représente le nombre de clients à la pompe 2 et X_3 représente le nombre de clients dans la place d'attente (X_3 vaut 0 ou 1).

- 1) Quel est l'espace d'état ?
- 2) Ecrire les composantes du générateur infinitésimal ?
- 3) Existe-il une probabilité stationnaire ?
- 4) Si oui, quelles sont les équations qui la caractérise ?
- 5) Quel est le pourcentage de clients qui ne peuvent pas entrer dans la station ?
- 6) Quel est le pourcentage de clients qui ne peuvent pas entrer dans la station parce qu'elle est mal faite : le pourcentage des clients qui ne peuvent pas être servis bien que la pompe 1 soit libre.

EXERCICE 18.— Dans un magasin, les clients, en moyenne vingt par heure, arrivent à la caisse selon un processus de Poisson. Tant qu'il y a moins de deux clients en attente, il y a une seule caisse d'ouverte. Le temps de service à cette caisse est exponentiellement distribué, le temps moyen de service est de cinq minutes. A partir de trois clients en attente, une deuxième caisse ouvre. Les deux caisses partagent la même file d'attente. La deuxième caisse referme dès qu'il n'y a plus que deux clients en attente. On note X_t le nombre de clients présents à l'instant t .

- 1) Ecrire le générateur infinitésimal de X .
- 2) Trouver la probabilité stationnaire si elle existe.
- 3) Quel est le nombre moyen de clients dans le système ?
- 4) Quel est le nombre moyen de clients en attente ?
- 5) Quel est le temps moyen d'attente ?
- 6) Quel est le pourcentage du temps où la deuxième caisse est ouverte ?

EXERCICE 19.— Une agence bancaire comporte $S = 5$ employés. Le nombre d'appels est en moyenne de vingt par heure, la durée moyenne d'un appel est de six minutes. Les arrivées forment un processus de Poisson d'intensité λ et les temps de service sont indépendants et de loi exponentielle de paramètre μ . Dans les questions 1 à 5, on ne vous demande pas de démontrer les formules, juste de trouver celles qu'il faut utiliser et de faire les calculs numériques. Ecrivez bien formellement tous les calculs que vous avez à faire, il y a pas mal de résultats partiels qui servent plusieurs fois.

- 1) Dans le cas où $N = 5$ et où il n'y a pas d'attente, quel est le pourcentage de client dont l'appel échoue faute de conseiller libre ?
- 2) L'agence s'équipe d'un dispositif de mise en attente, c'est-à-dire un disque qui vous fait patienter en attendant qu'un conseiller se libère. On suppose que tous les clients attendent aussi longtemps que nécessaire. Quelle est la probabilité que la probabilité d'avoir à attendre ?
- 3) Est-elle égale à la probabilité de blocage dans le système sans attente ? Pourquoi ?

- 4) Quel est le temps moyen d'attente ?
- 5) Quelle est la probabilité que le temps d'attente soit supérieur à deux minutes ?

On supprime maintenant le système de mise en attente et on considère qu'un pourcentage $100.p$ d'appels nécessitent un reroutage vers un centre de traitement plus spécialisé. Au niveau du PABX (et des N liaisons louées), cela se traduit par une occupation de deux liaisons : une pour l'appel entrant, l'autre pour l'appel rerouté. On note X^1 le nombre d'appels directs, utilisant une seule liaison ; X^2 représente le nombre d'appels qui ont besoin de deux liaisons. N est maintenant à déterminer. X^1 est limité par S le nombre de conseillers, tandis que X^2 est seulement limité par la condition $X^1 + 2X^2 \leq N$. On considère, en effet, que le dimensionnement du centre de traitement spécialisé est tel que les appels qui lui sont destinés aboutissent et on néglige le temps nécessaire au conseiller local pour décider du reroutage. D'autre part, lorsque tous les conseillers sont occupés, les appels sont tous dirigés vers le centre de traitement spécialisé.

- 6) Ecrire le générateur infinitésimal de $X = (X^1, X^2)$.
- 7) Le processus X admet-il une probabilité stationnaire ?
- 8) Ecrire les équations permettant de la déterminer.
- 9) A N fixé, quelle est la taille de la matrice à inverser ?

EXERCICE 20.— Retrouver (5.20) avec la formule de Little.

EXERCICE 21.— Démontrer le théorème 5.13 en utilisant le théorème 4.19.

5.9. Notes et commentaires

Nous n'avons qu'effleuré l'étude des files d'attente markoviennes à buffer infini. Depuis le temps que ce sujet a intéressé nombre de chercheurs, il existe nombre d'articles dévolus à ce problème. Nous n'avons en particulier pas parlé de trafic élastique dans les files *processor sharing*. Ce cadre pourtant riche et très bien étudié dans [BON 11] permet d'obtenir une formule pour les délais, insensible à la loi du temps de service.

L'étude des réseaux à forme produit repose essentiellement sur le lemme de Kelly que l'on trouvera dans [KEL 79]. L'équation de balance locale que cela induit est souvent trop restrictive pour être satisfaite. De nombreuses versions plus faibles mais toutefois moins contraignantes que les équations globales induites par la recherche des solutions de $\pi A = 0$, existent. On en trouvera de nombreux exemples dans [CHA 99].

Résumé

– La probabilité stationnaire d’une file M/M/1 d’intensité λ en entrée et de temps de service moyen $1/\mu$ est donnée par :

$$\pi(0) = 1 - \rho, \pi(n) = \rho^n(1 - \rho) \text{ où } \rho = \lambda/\mu.$$

– Le temps d’attente à l’état stationnaire vaut 0 avec probabilité $1 - \rho$. Conditionnellement à ce qu’il soit strictement positif, le temps d’attente suit une loi exponentielle de paramètre $\mu - \lambda$. Sa valeur moyenne est $1/(\mu - \lambda)$.

– Dans la file M/M/S, la condition de stabilité est $\rho < S$. La probabilité stationnaire est donnée par (5.9).

– Les comparaisons de files d’attente ne peuvent se faire qu’à charge égale.

– Dans un réseau de Jackson ouvert, la probabilité stationnaire du système est le produit des probabilités stationnaires de chacune des files vues comme une file M/M/1 avec comme paramètres d’entrée, les composantes de la solution de l’équation d’équilibre des flux, voir lemme 5.10.

Chapitre 6

Modèles à pertes

6.1. Généralités

Un système à pertes est un système qui comporte moins de ressources que d'utilisateurs potentiels ; tout client arrivant à un moment où les ressources du système sont toutes occupées est perdu. Il s'agit donc de trouver le nombre de ressources adéquat de façon à ne pas trop pénaliser le client ou le fournisseur. Dans toute la suite, on désignera toujours par :

- $N(t)$, le nombre de clients qui se sont présentés à l'entrée du système jusqu'à l'instant t ;
- $Y(t)$, le nombre de clients qui sont effectivement rentrés dans le système entre 0 et t . Par conséquent, $X^d(t) = N(t) - Y(t)$ est le nombre de clients perdus à l'instant t ;
- $X(t)$, le nombre de clients dans le système (serveurs + salle d'attente) à l'instant t ;
- S le nombre de serveurs et K le nombre de places dans la salle d'attente.

EXEMPLE 6.1.— Le réseau téléphonique mobile de deuxième génération, plus connu sous le nom de GSM, repose techniquement sur le TDMA (*Time Division Multiple Access*).

Pour une fréquence, on divise le temps en périodes de durée égales et constantes, appelées *slots*. Les *slots* sont regroupés par paquet de huit pour former une trame. La voix d'un appel est numérisée de sorte qu'acheminer une communication revient à transporter des bits d'information. Pour acheminer une communication, les octets d'un appel sont regroupés par paquets et on attribue à un appel un *slot* fixe pendant toute la durée de la communication. Sur la figure 6.1, c'est le quatrième *slot* qui est attribué à

un appel. Un équipement émission/réception ne peut donc acheminer que huit appels simultanément. Il importe donc à l'opérateur d'en déterminer le nombre pour chaque station de base, c'est-à-dire les antennes qui décorent les toits des immeubles. En fait, pour tenir compte de la signalisation, un équipement permettra d'acheminer sept appels simultanés, deux équipements quatorze, trois équipements vingt-deux, etc.

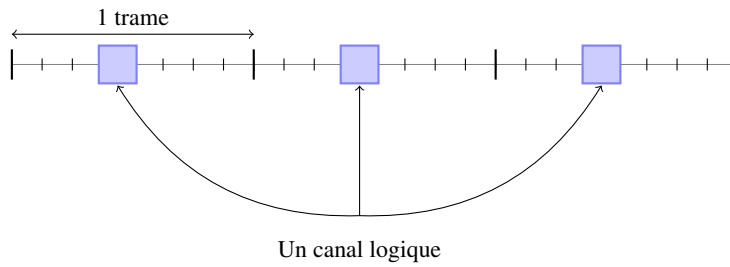


Figure 6.1 – Principe du TDMA

EXEMPLE 6.2.– Un multiplexeur est l'un des éléments-clés des réseaux de données. Il permet d'acheminer les communications qui arrivent de N ports vers N ports de sortie. Comme un même port de sortie peut être sollicité simultanément par plusieurs flux, il importe d'installer des mémoires pour stocker temporairement les données en attendant que le canal soit libre. Le problème est ici de trouver la taille des mémoires de sorte que la probabilité de perte de bits soit inférieure à un certain seuil. Ce seuil dépend de la nature du flux de données. *Grosso modo*, les flux voix et vidéo tolèrent des pertes de l'ordre de un pour mille tandis que les flux de données ne devraient tolérer aucune perte. En pratique, on considère que quelque chose qui arrive avec une probabilité inférieure à un pour un milliard ne se produira de fait pas.

Ces deux exemples sont de même nature mais le deuxième soulève deux problèmes de taille : il n'est pas évident de choisir un modèle de trafic d'entrée qui puisse refléter les différences (de débit, par exemple, pour ne citer qu'elles) qui existent entre des flux de voix, de vidéo et paroles ; par ailleurs la petitesse de la probabilité de perte interdit toute simulation et exige un modèle analytique précis. Ils nous amènent aussi à préciser et distinguer trois probabilités.

DÉFINITION 6.1.– La probabilité de congestion d'un système à $S + K$ ressources est la proportion asymptotique du temps où toutes les ressources soient occupées :

$$P_B = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbf{1}_{\{X(t)=S+K\}} dt.$$

La probabilité de perte est la probabilité de blocage du point de vue des clients :

$$P_p = \lim_{t \rightarrow \infty} \frac{1}{N(t)} \sum_{j=1}^{N(t)} \mathbf{1}_{\{X(T_j^-) = S+K\}} \quad (6.1)$$

La probabilité de dépassement de seuil n'existe en fait que lorsque le nombre de ressources est infini, elle est alors définie par :

$$P_S = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbf{1}_{\{X(t) > S\}} dt.$$

Ces trois probabilités sont en général distinctes. En effet, considérons une situation avec un seul serveur, des arrivées et des temps de service déterministes, de durée respective $\rho < 1$ et 1.

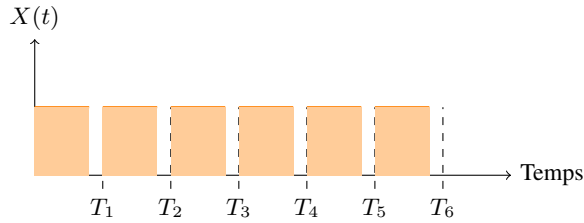


Figure 6.2 – L'évolution temporelle de la file $D/D/1/1$

D'après les définitions, la probabilité de blocage est égale au pourcentage d'aire grisée, soit ρ . La probabilité de perte est nulle puisqu'aucun client n'arrive avec le serveur occupé. En revanche, si les arrivées sont poissonniennes, les probabilités de perte et de blocage coïncident en vertu de la propriété PASTA.

THÉORÈME 6.1.— Si les arrivées suivent un processus de Poisson alors $P_B = P_p$.

Démonstration. Il suffit d'appliquer le théorème A.38 à $\psi(s) = \mathbf{1}_{\{X(s) = S+K\}}$. \square

La probabilité de perte est de loin la plus cruciale mais aussi la plus difficile à évaluer. Les probabilités de blocage et de dépassement sont beaucoup plus facilement calculables et c'est pourquoi elles sont souvent utilisées comme succédané à la probabilité de perte.

6.2. Modèle d'Erlang

On suppose que les arrivées se font selon un processus de Poisson d'intensité λ , que les temps de service sont indépendantes et identiquement distribuées de loi exponentielle de paramètre $\varepsilon(\mu)$, et que la capacité de la salle d'attente est nulle, soit $c = 0$, de telle sorte que tout client est servi si et seulement si l'un des serveurs est libre à son arrivée, et perdu sinon. On étudie en d'autres termes la file $M_\lambda/M_\mu/S/S$ et l'on note comme d'habitude sa charge $\rho = \frac{\lambda}{\mu}$.

Comme précédemment, on décrit le système par le processus $(X^E(t), t \geq 0)$ comptant le nombre de clients dans le système. Sur son espace d'état $\llbracket 0, S \rrbracket$, $(X^E(t), t \geq 0)$ a les mêmes transitions que le processus $(X^S(t), t \geq 0)$ décrivant la file $M_\lambda/M_\mu/S/\infty$ (voir section 5.3), sauf en l'état S , que le processus ne quitte que pour sauter en $S - 1$, puisqu'alors aucun client n'est accepté. Le générateur infinitésimal de $(X^E(t), t \geq 0)$ est donc donné par :

$$A^E = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\mu + \lambda) & \lambda & & \\ & & \ddots & & \\ & & & k\mu & -(\lambda + k\mu) & \lambda \\ & (0) & & & \ddots & \\ & & & & & (S-1)\mu & -((S-1)\mu + \lambda) & \lambda \\ & & & & & S\mu & -S\mu \end{pmatrix}. \quad (0)$$

Il est alors aisé de calculer comme d'habitude l'unique probabilité stationnaire π^E du processus, qui satisfait :

$$\pi^E(i) = \frac{\lambda}{i\mu} \pi^E(i-1), \quad i \in \llbracket 1, S \rrbracket;$$

$$\sum_{i=0}^S \pi^E(i) = 1,$$

ce qui équivaut à :

$$\pi^E(i) = \frac{\rho^i / i!}{\sum_{k=0}^S \rho^k / k!} \quad \text{pour } i \in \llbracket 0, S \rrbracket. \quad (6.2)$$

Par ailleurs, comme le processus $(X^E(t), t \geq 0)$ est de naissance et de mort, il est réversible.

REMARQUE.— Nous pouvons retrouver cette probabilité invariante en remarquant que la file $M_\lambda/M_\mu/S/S$ est une file $M_\lambda/M_\mu/\infty$ contrainte à ne pas dépasser S clients. En d'autres termes, $(X^E(t), t \geq 0)$ est la version tronquée, contrainte à rester dans $\llbracket 0, S \rrbracket$, du processus réversible $(X^\infty(t), t \geq 0)$ représentant le nombre de clients dans la $M_\lambda/M_\mu/\infty$. Pour tout $i \in \llbracket 0, S \rrbracket$, le lemme de Kelly permet de calculer π^E :

$$\pi^E(i) = \frac{1}{\sum_{k=0}^S \pi^\infty(k)} \pi^\infty(i) = \frac{\rho^i/i!}{\sum_{k=0}^S \rho^k/k!}.$$

D'après (6.1), la probabilité de perte pour ce système vaut :

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbf{1}_{\{S\}}(X(T_j^-)) &= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbf{1}_{\{S\}}(X(s^-)) ds \\ &= \pi^E(S). \end{aligned}$$

Nous avons donc établi le résultat suivant.

THÉORÈME 6.2.— La probabilité de perte de la file $M_\lambda/M_\mu/S/S$ de charge ρ est donnée par :

$$\text{Er}[\rho, S] = \frac{\rho^S/S!}{\sum_{i=0}^S \rho^i/i!}.$$

REMARQUE.— On montre que cette probabilité de perte ne dépend pas de la loi de service : cette expression est encore vraie pour la file $M/GI/S/S$.

Pour l'évaluation numérique de cette probabilité, on pourra utiliser l'approximation suivante :

$$\text{Er}[\rho, S] \approx \exp S \log \frac{\rho u}{S} + S - \rho u \sqrt{\frac{u + \rho u(1-u)^2}{S}}$$

avec u défini par :

$$u = \frac{(S + \rho + 1) - \sqrt{(S + \rho + 1)^2 - 4\rho S}}{2\rho};$$

ou alors la relation de récurrence :

$$\frac{1}{\text{Er}[\rho, S]} = 1 + \frac{S}{\rho \text{Er}[\rho, S-1]}.$$

En pratique, ayant mesuré ρ , il s'agit de trouver S tel que $\text{Er}[\rho, S]$ soit inférieur au seuil désiré pour garantir la qualité de service. On considère généralement que ce seuil

est de 0,001 pour le réseau téléphonique fixe et 0,02 pour le réseau GSM. L'algorithme est alors le suivant.

Algorithme 6.1 . Calcul du nombre de serveurs pour garantir une probabilité de perte

Données : ρ, ϵ
Résultat : S tel que $\text{Er}[\rho, S] \leq \epsilon$
 $S \leftarrow 0;$
 $x \leftarrow 1;$
tant que $x < \epsilon^{-1}$ **faire**
 $S \leftarrow S + 1;$
 $x \leftarrow 1 + \frac{S}{\rho};$
fin
retourner S

L'expérience montre que pour une perte de un pour mille, le nombre de serveurs nécessaire est (très) approximativement égal à $\rho + 3\sqrt{\rho}$. Si les arrivées et les temps de service étaient déterministes, ρ représenterait le nombre d'appels simultanés donc aussi le nombre de serveurs nécessaires. Le terme $3\sqrt{\rho}$ s'interprète donc comme une garantie contre les fluctuations aléatoires du trafic.

6.3. La file M/M/1/1 + K

Nous considérons maintenant une file d'attente à un serveur travaillant en FCFS, dont la salle d'attente a une capacité limitée K . Comme précédemment, le processus d'arrivées est poissonien d'intensité λ , les clients demandent des temps de service indépendants et identiquement distribués de loi $\varepsilon(\mu)$ et nous notons $\rho = \lambda/\mu$, la charge du système.

Le processus $(X^k(t), t \geq 0)$ comptant le nombre de clients dans le système est à valeurs dans $\llbracket 0, 1 + K \rrbracket$. Il satisfait la dynamique suivante :

- pour tout $i \in \llbracket 0, K \rrbracket$, le temps de séjour en i et les probabilités de saut partant de i sont clairement les mêmes que pour la file $M_\lambda/M_\mu/1/\infty$;
- une fois en $1 + K$, le système est plein et ne peut plus accepter de clients. Le processus ne peut quitter cet état que pour l'état K , après un temps de séjour égal au temps de service résiduel du client alors en service, de loi $\varepsilon(\mu)$ tant que le processus vaut $1 + K$.

Il est donc facile de voir que $(X^K(t), t \geq 0)$ est markovien et admet sur $\llbracket 0, 1 + K \rrbracket$ le générateur infinitésimal :

$$A^K = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & (0) \\ & & \ddots & & \\ (0) & \mu & -(\lambda + \mu) & \lambda & \\ & & \mu & -\mu & \end{pmatrix}.$$

Le processus $(X^K(t), t \geq 0)$ étant à valeurs dans un ensemble fini, il est nécessairement ergodique et admet donc une unique probabilité stationnaire π^K . Celle-ci est réversible puisque $(X^K(t), t \geq 0)$ est de naissance et de mort. On la retrouve, soit directement en résolvant le système :

$$\begin{cases} \pi^K A^K &= \mathbf{0} \\ \pi^K \mathbf{e} &= 1, \end{cases}$$

soit encore en appliquant le lemme de Kelly : $(X^K(t), t \geq 0)$ est la troncature à l'ensemble $\llbracket 0, 1 + K \rrbracket$ du processus réversible $(X(t), t \geq 0)$ comptant le nombre de clients dans la $M_\lambda/M_\mu/1/\infty$. On retrouve la probabilité donnée pour tout $i \in \llbracket 0, 1 + K \rrbracket$ par :

$$\begin{aligned} \pi^K(i) &= \frac{\pi(i)}{\sum_{k \in \llbracket 0, 1 + K \rrbracket} \pi(k)} \\ &= \frac{(1 - \rho)\rho^i}{\sum_{k \in \llbracket 0, 1 + K \rrbracket} (1 - \rho)\rho^k} \\ &= \begin{cases} \frac{\rho^i - \rho^{i+1}}{1 - \rho^{K+2}} & \text{si } \rho \neq 1, \\ \frac{1}{K + 2} & \text{si } \rho = 1, \end{cases} \end{aligned} \quad (6.3)$$

et l'on remarque au passage que l'unique probabilité invariante et la probabilité uniforme sur $\llbracket 0, 1 + K \rrbracket$ dans le cas critique $\rho = 1$.

THÉORÈME 6.3.— Dans une file $M_\lambda/M_\mu/1/1+K$, la probabilité de perte est donnée par la formule :

$$\frac{\rho^{K+1} - \rho^{K+2}}{1 - \rho^{K+2}} \text{ si } \rho \neq 1.$$

Si $\rho = 1$, la perte est donnée par $1/(K + 2)$.

Démonstration. La probabilité de perte pour ce système est donnée par la limite du taux de clients perdus, c'est-à-dire les clients trouvant un système plein à leur arrivée.

D'après la propriété PASTA et l'ergodicité de $(X^K(t), t \geq 0)$, celle-ci s'écrit :

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{\{1+K\}}(X^K(T_n-)) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbf{1}_{\{1+K\}}(X^K(t)) dt \\ &= \mathbf{E}[\mathbf{1}_{\{1+K\}}(X_\infty^K)] \\ &= \pi^K(1+K). \end{aligned}$$

Le résultat découle alors de (6.3). \square

Si la charge ρ est strictement inférieure à 1, la probabilité de perte peut être comparée à la probabilité de dépassement de seuil de la file $M_\lambda/M_\mu/1/\infty$ correspondante. Avec les notations de la section 5.2, la probabilité que ce dernier système ait une congestion supérieure ou égale à $1+K$ à l'équilibre s'écrit :

$$\mathbf{P}(X_\infty \geq 1+K) = \sum_{i=1+K}^{\infty} \pi(i) = (1-\rho) \sum_{i=1+K}^{\infty} \rho^i = \rho^{1+K}, \quad (6.4)$$

alors que :

$$P_{\text{PERTE}} = \pi^K(0) \rho^{1+K} < \rho^{1+K}. \quad (6.5)$$

Ainsi, il est ainsi possible d'estimer la perte d'un système à buffer fini par la probabilité de dépassement de seuil de la file à buffer infini correspondante. Ce faisant, on surestime la perte d'après (6.4) et (6.5), mais cela peut s'avérer être le seul calcul que l'on sache faire en pratique car le système à buffer infini est globalement plus simple à décrire par des lois de probabilité connues. Qui plus est, un dimensionnement pessimiste tant qu'il n'est pas exagéré, il permet de se protéger contre les fluctuations de trafic ou contre les incertitudes sur l'estimation de la charge.

Dimensionnement du buffer

Il est intéressant de se demander quelle est la taille de buffer optimale que l'on doit déployer pour garantir à l'utilisateur une qualité de service en termes de probabilité de perte de paquets. On peut déjà vérifier que la perte du système est une fonction décroissante du nombre de serveurs. En effet, la fonction définie pour tout $\rho \in \mathbf{R}^+ \setminus \{1\}$ par :

$$f_\rho : \begin{cases} \mathbf{R}^+ & \longrightarrow \mathbf{R}^+ \\ x & \longmapsto \frac{\rho^{x+1} - \rho^{x+2}}{1 - \rho^{x+2}}. \end{cases}$$

admet pour dérivée la fonction définie pour tout x par :

$$f'_\rho(x) = \frac{\rho^{x+1} \ln \rho}{(1 - \rho^{x+2})^2} (1 - \rho) > 0,$$

et les choses se passent bien sûr de la même façon pour $\rho = 1$.

Le dimensionnement optimal du buffer est donc donné par l'algorithme suivant (dans le cas $\rho \neq 1$).

Algorithme 6.2 . Calcul du nombre minimal de places dans le buffer pour garantir une probabilité de perte $\leq \varepsilon$.

Données : ρ, ε

Résultat : K tel que $P_{\text{PERTE}} \leq \varepsilon$

$K \leftarrow 1$;

tant que

$$\varepsilon < \frac{\rho^{K+1} - \rho^{K+2}}{1 - \rho^{K+2}}$$

faire

fin

retourner K

6.4. L'effet *trunk*

Appliquons nos résultats au GSM. Dans le langage du protocole, un *slot* s'appelle un TCH (*traffic channel*). On considère usuellement que la charge par mobile est de 0,025 erlang et que la probabilité de perte acceptable est 0,02. Compte tenu de ce qui précède, on obtient les résultats suivants.

Nb émetteur-récepteurs	1	2	3	4	5	6	7
Nb TCH	7	14	22	29	37	45	52
Capacité	2,9	8,2	15	21	28	35,5	42
Trafic écoulé par TCH	0,41	0,59	0,68	0,72	0,76	0,79	0,81
Nb de mobiles	116	328	596	840	1128	1424	1680

Tableau 6.1 – Mise en évidence de l'effet *trunk*

Ce tableau appelle deux commentaires. Il importe de prendre conscience de la grande variabilité des résultats. Une augmentation raisonnable du nombre de ressources induit une grande augmentation de la capacité. Par exemple, si l'on passe de un à deux émetteurs-récepteurs, la capacité est multipliée par presque trois. Pour écouler plus de trafic, on n'a pas besoin d'augmenter proportionnellement le nombre de TCH : pour écouler 42 erlangs au lieu de 21, on a besoin de 52 au lieu de $2 * 29 = 58$.

L'autre remarque est connue sous le nom d'effet *trunk* : plus on augmente le nombre de serveurs, plus la charge écoulee par chaque serveur est importante. Journalistiquement, cela s'appelle aussi « les économies d'échelle ». Par conséquent, lorsque le seuil

de perte est fixé, on cherchera à avoir un petit nombre de systèmes à grands nombres de serveurs plutôt qu'un grand nombre de systèmes à faible nombre de serveurs.

6.5. Modèle d'Engset

Pour pouvoir utiliser le modèle d'Erlang, il faut implicitement supposer que le nombre de sources potentielles est sinon infini du moins très grand comme c'est le cas, par exemple, en téléphonie. En revanche, lorsque le nombre de sources est faible devant le nombre de serveurs et que l'on suppose qu'une source en service ne peut plus émettre jusqu'à la fin de son service, on ne peut plus considérer que l'intensité d'arrivée des clients est indépendante de l'état du système.

Dans le modèle dit d'Engset, on suppose que l'on a M sources indépendantes qui génèrent chacune des requêtes selon un processus de Poisson d'intensité λ . On a toujours S serveurs ($S \leq M$) et les lois de service sont toujours supposées être exponentielles de paramètre μ , la salle d'attente est encore de taille 0. A la différence du modèle d'Erlang, lorsque k sources sont « en service », seules $M - k$ sources sont susceptibles d'émettre une requête. Par conséquent le taux d'arrivées instantané est $(M - k)\lambda$. De fait, X est un processus de Markov de sauts pur de générateur infinitésimal A donné par :

$$\begin{pmatrix} -\lambda_0 & \lambda_0 & & & & \\ \mu_1 & -(\mu_1 + \lambda_1) & \lambda_1 & & & \\ & & \ddots & & & \\ & & & \mu_k & -(\lambda_k + \mu_k) & \lambda_k \\ & & & & \ddots & \\ & & & & & \mu_{S-1} & -(\mu_{S-1} + \lambda_{S-1}) & \lambda_{S-1} \\ & & & & & & \mu_S & -\mu_S \end{pmatrix} \quad (0)$$

où l'on a posé $\lambda_i = (M - i)\lambda$, $\mu_i = i\mu$ et $\rho = \lambda/\mu$. La distribution invariante de ce processus est définie par :

$$\nu A = 0 \iff \begin{cases} -\lambda_0 \nu_0 + \mu \nu_1 & = 0 \\ \lambda_{i-1} \nu_{i-1} - (\lambda_i + \mu_i) \nu_i + \mu_{i+1} \nu_{i+1} & = 0 \end{cases}$$

On obtient alors :

$$\nu_i = \frac{\rho^i C_M^i}{\sum_{j=0}^S C_M^j \rho^j}.$$

THÉORÈME 6.4.— Dans le modèle d'Engset, la probabilité de perte est donnée par :

$$\text{Eng}[\rho, S, M] = \frac{\rho^S C_{M-1}^S}{\sum_{j=0}^S C_{M-1}^j \rho^j}.$$

Démonstration. D'après la construction du modèle, le processus d'arrivées est un processus ponctuel de Poisson dont l'intensité varie dans le temps comme une fonction de $X(s)$: son intensité est donnée par :

$$\Lambda(t) = \int_0^t (M - X(s)) \lambda \, ds.$$

Comme les sauts de $N(t)$ sont de hauteur 1, le crochet de la martingale $N - \Lambda$ est aussi Λ . Soit le processus adapté borné, $\psi(s) = \mathbf{1}_{\{X(s)=S\}}$. D'après le théorème A.37 :

$$\begin{aligned} P_p &= \lim_{t \rightarrow \infty} \frac{1}{N(t)} \int_0^t \psi(s_-) \, dN(s) = \lim_{t \rightarrow \infty} \frac{1}{\Lambda(t)} \int_0^t \psi(s) \, d\Lambda(s) \\ &= \lim_{t \rightarrow \infty} \frac{t}{\Lambda(t)} \frac{1}{t} \int_0^t \psi(s) (M - X(s)) \lambda \, ds. \end{aligned}$$

Comme X est markovien ergodique, on a les limites presque sûres suivantes :

$$\begin{aligned} \frac{1}{t} \Lambda(t) &\xrightarrow{t \rightarrow \infty} (M - \sum_{j=0}^S j \nu(j)) \lambda \\ \text{et } \frac{1}{t} \int_0^t \psi(s) (M - X(s)) \lambda \, ds &\xrightarrow{t \rightarrow \infty} \lambda (M - S) \nu(S). \end{aligned}$$

Par conséquent :

$$\text{Eng}[\rho, S, M] = \frac{\lambda (M - S) \nu(S)}{\lambda \sum_{j=0}^S (M - j) \nu(j)} = \frac{\rho^S C_{M-1}^S}{\sum_{j=0}^S C_{M-1}^j \rho^j}.$$

En d'autres termes, la probabilité de pertes d'un système à M machines est la probabilité de blocage d'un système à $M - 1$ machines, le reste étant inchangé par ailleurs. \square

6.6. File IPP/M/S/S

Les résultats précédents peuvent laisser accroire que la perte dépend uniquement de la charge. Il n'en est rien comme l'illustre l'exemple suivant.

Un processus IPP (*Interrupted Poisson Process*) est un cas particulier de processus MMPP (voir 4.6) où le processus de phases J est à deux états : A et B . Le générateur infinitésimal de J est donc de la forme :

$$Q_J = \begin{pmatrix} -\sigma_A & \sigma_A \\ \sigma_B & -\sigma_B \end{pmatrix},$$

où $1/\sigma_A$ et $1/\sigma_B$ sont les temps moyens de séjour dans les phases A et B respectivement. Sa probabilité invariante, notée ν se calcule facilement et l'on obtient :

$$\nu(A) = \frac{\sigma_B}{\sigma_A + \sigma_B} \text{ et } \nu(B) = \frac{\sigma_A}{\sigma_A + \sigma_B}.$$

La file IPP/M/S/S est donc la file à S serveurs, sans buffer, avec des temps de service de loi exponentielle et un processus d'arrivées IPP. Le processus X comptant le nombre de serveurs occupés n'est pas markovien mais le processus (X, J) l'est. On numérote les états par ordre lexicographique et on note Λ la matrice des intensités d'arrivées :

$$\Lambda = \begin{pmatrix} \lambda_A & 0 \\ 0 & \lambda_B \end{pmatrix}.$$

L'intensité moyenne du flux d'entrée est donc :

$$\lambda = \lambda_A \frac{\sigma_B}{\sigma_A + \sigma_B} + \lambda_B \frac{\sigma_A}{\sigma_A + \sigma_B}. \quad (6.6)$$

Le générateur infinitésimal de (X, J) s'écrit donc :

$$A = \begin{pmatrix} Q_J - \Lambda & \Lambda & & & \\ \mu \text{Id} & (Q_J - \Lambda - \mu \text{Id}) & \Lambda & & \\ & 2\mu \text{Id} & (Q_J - \Lambda - 2\mu \text{Id}) & \Lambda & \\ & & \ddots & \ddots & \ddots \\ & & & S\mu \text{Id} & (Q_J - S\mu \text{Id}) \end{pmatrix}$$

Le processus de Markov est à espace d'états fini, évidemment irréductible, il admet une probabilité invariante π solution de l'équation $\pi A = 0$ et $\pi \cdot \mathbf{e} = 1$. Pour simplifier les calculs, on introduit les vecteurs lignes à deux composantes $x_n = (\pi(n, A), \pi(n, B))$ pour $n = 0, \dots, S$. Les équations correspondant à $\pi A = 0$ deviennent ainsi S équations vectorielles 2 par 2 :

$$\begin{aligned} x_0(Q_J - \Lambda) + \mu x_1 &= 0 \\ x_0\Lambda + x_1(Q_J - \Lambda - \mu \text{Id}) + 2\mu x_2 &= 0 \\ &\vdots \\ x_{S-2}\Lambda + x_{S-1}(Q_J - \lambda - (S-1)\mu \text{Id}) + S\mu x_S &= 0 \\ x_{S-1}\Lambda + x_S(Q_J - S\mu \text{Id}) &= 0. \end{aligned}$$

Des $(S - 1)$ premières équations, on tire :

$$\begin{aligned} x_1 &= \frac{1}{\mu} x_0 (Q_J - \Lambda) \\ x_2 &= -\frac{1}{2\mu} (x_0 \Lambda + x_1 (Q_J - \Lambda - \mu \text{Id})) \\ &\vdots \\ x_S &= -\frac{1}{S\mu} (x_{S-2} \Lambda + x_{S-1} (Q_J - \Lambda - (S-1)\mu \text{Id})) \\ x_S &= x_{S-1} \Lambda (Q_J - S\mu \text{Id})^{-1}. \end{aligned}$$

Posons :

$$\begin{aligned} R_1 &= \frac{1}{\mu} (Q_J - \Lambda), \\ R_n &= -\frac{1}{n\mu} (R_{n-2} \Lambda + R_{n-1} (Q_J - \Lambda - (n-1)\mu \text{Id})) \text{ pour } n = 2, \dots, S. \end{aligned} \quad (6.7)$$

On peut alors écrire :

$$x_n = x_{n-1} R_n \text{ et } x_0 (R_S - R_{S-1} \Lambda (Q_J - S\mu \text{Id})^{-1}) = 0.$$

Cela définit deux équations sur les deux composantes de x_0 . En fait, une seule est utile puisque si le système est rang 2, la seule solution est 0 ce qui est exclu. On en déduit l'algorithme de résolution suivant.

Algorithme 6.3 . Calcul de la probabilité invariante de la file IPP/M/S/S

Données : Λ, μ, S, Q_J

Résultat : π telle que $\pi A = 0$ et $\pi \cdot \mathbf{e} = 1$

Calculer R_1, \dots, R_S selon l'équation (6.7);

$x_0(0, A) \leftarrow 1$;

Trouver $x_0(0, B)$ tel que x_0 satisfasse $x_0 (R_S - R_{S-1} \Lambda (Q_J - S\mu \text{Id})^{-1}) = 0$;

Calculer $x_n = x_{n-1} R_n$ pour $n = 1, \dots, S$;

Calculer $m = \sum_{n=0}^S x_n \cdot \mathbf{e}$;

$\pi \leftarrow x/m$;

retourner π

THÉORÈME 6.5.— Dans une file IPP/M/S/S de probabilité invariante, la probabilité de perte est donnée par :

$$\frac{\lambda_A}{\lambda} \pi(S, A) + \frac{\lambda_B}{\lambda} \pi(S, B), \quad (6.8)$$

où λ est l'intensité moyenne donnée par (6.6).

Démonstration. On applique le théorème 4.24 à $\psi(s) = \mathbf{1}_{\{S\}}(X(s))$, il vient que la probabilité de perte est donnée par :

$$\frac{1}{\lambda} \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbf{1}_{\{S\}}(X(s)) \lambda(J(s)) \, ds = \frac{1}{\lambda} (\lambda_A \pi(S, A) + \lambda_B \pi(S, B)),$$

d'où le résultat. \square

Fixer la charge revient à fixer λ mais suivant les valeurs relatives de λ_A , λ_B , σ_A et σ_B on obtient des probabilités de perte très différentes comme l'illustre le tableau 6.2. Nous avons choisi pour les calculs $S = 10$ serveurs et une charge de 5 Erlang. La formule d'Erlang donnerait une perte égale à 0,018.

Paramètres	Perte	Blocage
$\begin{pmatrix} \sigma_A = 1/2 & \sigma_B = 1/2 \\ \lambda_A = 10 & \lambda_B = 0 \end{pmatrix}$	0,1	0,05
$\begin{pmatrix} \sigma_B = 1/10 & \sigma_A = 9/10 \\ \lambda_A = 50 & \lambda_B = 0 \end{pmatrix}$	0,66	0,07
$\begin{pmatrix} \sigma_B = 1/100 & \sigma_A = 99/100 \\ \lambda_A = 500 & \lambda_B = 0 \end{pmatrix}$	0,96	0,01
$\begin{pmatrix} \sigma_B = 1/10 & \sigma_A = 9/10 \\ \lambda_A = 1 & \lambda_B = 5,444 \end{pmatrix}$	0,02	0,05

Tableau 6.2 – Probabilités de pertes à charge constante dans une file IPP/M/S/S

Il apparaît à la lecture de ce tableau que l'on peut augmenter la probabilité de perte en gardant une charge constante. On pourrait croire que le facteur déterminant devient alors la variance du processus d'entrée puisque plus l'on augmente la variabilité de celui-ci, plus la perte augmente. C'est certainement un critère mais ce n'est hélas pas le seul. Fixer la variance et la charge revient à imposer deux équations que doivent satisfaire les quatre paramètres. Il reste donc deux degrés de liberté que l'on peut utiliser pour faire varier la probabilité de perte comme bon nous semble.

On remarque aussi que la probabilité de blocage est désespérément loin de la probabilité de perte, il y a donc une erreur monumentale à assimiler l'une à l'autre.

6.7. Modèles d'Erlang généralisés

6.7.1. Canaux de garde

Les canaux de garde sont un système bien connu des téléphonistes qui permet de définir une priorité entre plusieurs flux sans altérer trop la probabilité de perte du flux le moins prioritaire.

Nous illustrons cette notion par la gestion des *hand-over* dans le réseau GSM. Lorsqu'un appel issu d'un téléphone mobile bouge, il arrive un moment où la BTS qui gère cet appel perd le contact radio avec le mobile. Il est donc nécessaire de passer le relais à une autre BTS. Cette opération, appelée *hand-over*, nécessite de lourdes opérations du plan contrôle et l'on doit s'assurer qu'il y a suffisamment de ressources dans la nouvelle cellule pour acheminer ce nouvel appel.

Lorsque l'on prend en compte la mobilité des usagers, il est nécessaire d'introduire la durée de séjour dans la cellule de chaque usager. Pour pouvoir faire des calculs, on suppose que le temps de traversée d'une cellule suit une loi exponentielle de paramètre α . Le temps de communication d'un mobile vu de la BTS suit donc le minimum de la durée de communication et du temps de séjour. Compte tenu des propriétés de la loi exponentielle, cette durée de communication suit donc une loi exponentielle de paramètre $\mu + \alpha$. La probabilité qu'un usager « quitte » la cellule sur un *hand-over* est la probabilité qu'une variable aléatoire de loi exponentielle de paramètre μ se termine avant une variable aléatoire de loi exponentielle de paramètre α . On sait que cela vaut $\theta = \mu(\mu + \alpha)^{-1}$.

Considérons un ensemble de cellules de caractéristiques identiques et faisons le bilan des entrées sorties dans une cellule donnée. Dans la figure 6.3, λ_f est le taux d'appels nouveaux ou « frais », λ_{HO} le nombre moyen d'appels en *hand-over* qui arrive dans la cellule étudiée, p la probabilité de perte dans la cellule.

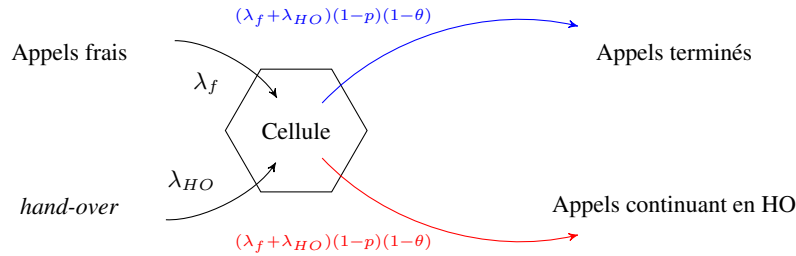


Figure 6.3 – Entrées-sorties d'une cellule

Comme le système est conçu de sorte que p soit négligeable devant 1, on a à l'équilibre :

$$\lambda_{HO} = (\lambda_f + \lambda_{HO})\theta \text{ soit } \lambda_{HO} = \frac{\theta}{1-\theta}\lambda_f.$$

La charge totale que doit traiter la BTS est donc :

$$\begin{aligned}
 \rho &= (\lambda_f + \lambda_{HO}) \times \frac{1}{\mu + \alpha} \\
 &= \lambda_f \left(1 + \frac{\theta}{1 - \theta}\right) \frac{1}{\mu + \alpha} \\
 &= \lambda_f \frac{1}{1 - \theta} \frac{1}{\mu + \alpha} \\
 &= \lambda_f \frac{\mu + \alpha}{\mu} \frac{1}{\mu + \alpha} \\
 &= \frac{\lambda_f}{\mu}.
 \end{aligned}$$

On ne peut toutefois pas appliquer la formule d'Erlang pour dimensionner le système avec *hand-over* car ceux-ci ont des besoins de qualité de service différents. Il est en effet bien plus désagréable d'avoir sa communication interrompue que de ne pouvoir en initier une. Pour donner une priorité aux appels en *hand-over*, on modifie la fonction de contrôle d'accès : on choisit $g < S$; tant que le nombre de serveurs libres est supérieur à g on accepte les appels frais et les appels en *hand-over*. Dès qu'il reste moins de g canaux disponibles, on n'accepte plus que les appels en *hand-over*. Pour calculer les probabilités de perte, on représente le système par un processus de Markov X comptant les serveurs occupés à chaque instant. C'est un processus de Markov dont les transitions peuvent se représenter par le diagramme 6.4 où $\lambda^+ = \lambda_f + \lambda_{HO}$ et $\mu^+ = \mu + \alpha$.

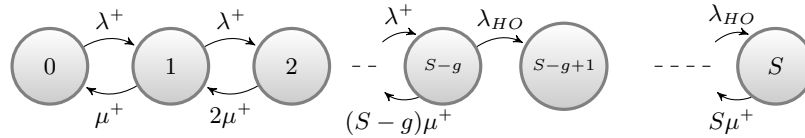


Figure 6.4 – Les transitions du processus « nombre de serveurs occupés » avec canaux de garde

A partir de là, on en déduit la probabilité invariante ν puis les probabilités de perte par la propriété PASTA. Puisque le trafic des appels frais est toujours un processus de Poisson d'intensité λ_f , le théorème A.38 implique que la probabilité de perte des appels frais est donnée par $\sum_{j=S-g}^S \nu(j)$ et celle des appels en *hand-over* par $\nu(S)$. On obtient les résultats numériques suivants.

S	30	32	34	34
g	0	2	4	3
Perte appels frais	0,8457%	1,0275%	1,0332%	0,6584%
Pertes HO	0,8457%	0,0278%	0,0008%	0,0028%

Tableau 6.3 – Pertes des appels frais et des appels en hand-over en fonction du nombre de canaux de garde

Il est tout à fait remarquable que pour de petites valeurs de g , on donne aux appels en *hand-over* une probabilité de perte bien inférieure à celle des appels frais sans pour autant trop pénaliser ces derniers.

6.7.2. Systèmes multiclassés

Dans de nombreuses situations, les clients ne réclament pas tous la même quantité de ressources. Dans ce cas là, on utilise la formule d'Erlang multiclassée. Soit un système comportant K classes de clients et S ressources. Les clients de classe i arrivent selon un processus de Poisson d'intensité λ_i et ont une durée de communication de loi exponentielle de moyenne $1/\mu_i$. On pose $\rho_i = \lambda_i/\mu_i$. Un client de classe i consomme s_i ressources. Les nombres de clients n_i de chaque classe sont donc soumis à la contrainte : $\sum_{i=1}^K n_i s_i \leq S$. On étudie le processus X qui comptabilise le nombre de ressources occupées. Son espace d'états est :

$$E = \{(n_1, \dots, n_K) \in \mathbf{N}, \sum_{j=1}^k n_j s_j \leq S\}.$$

THÉORÈME 6.6. – La probabilité invariante de ce système est donnée par :

$$\nu(n_1, \dots, n_K) = \frac{1}{G} \prod_{j=1}^k \rho^{n_j} / n_j!,$$

où G est la constante de normalisation :

$$G = \sum_{(n_1, \dots, n_K) \in E} \prod_{j=1}^k \rho^{n_j} / n_j!.$$

Démonstration. Si $S = \infty$, les composantes de X sont des processus de Markov indépendants qui évoluent respectivement comme le nombre de serveurs occupés dans une file $M_{\lambda_i}/M_{\mu_i}/\infty$. Ce sont donc des processus réversibles (voir définition 4.11) de loi invariante ν_i , loi de Poisson de paramètre ρ_i (voir l'exemple ??). Le processus X est donc réversible, sa probabilité invariante est le produit tensoriel de ces probabilités ν_i .

Pour S fini, X est juste la restriction à E de la dynamique précédente. En vertu du lemme de Kelly, on sait alors que la loi invariante de X est la restriction, renormalisation de la probabilité invariante dans le cas infini. \square

Soit ϵ_i le vecteur dont toutes les composantes sont nulles sauf la i^{e} qui vaut 1. Il y a perte d'un appel de classe i lorsque $X \in E$ mais $X + \epsilon_i$ n'appartient pas à E . La probabilité de perte p_i d'une classe donnée i est donc selon la propriété PASTA égale :

$$p_i = \sum_{n \in E, n + \epsilon_i \notin E} \nu(n_1, \dots, n_K).$$

EXEMPLE 6.3 (Interface A-bis).— Les progrès des codeurs de voix impliquent que là où un TCH par trame était nécessaire pour acheminer la communication, il n'y en a dorénavant besoin que qu'un demi. Cependant, certains appels ont toujours besoin d'un *slot* complet. Comme la théorie précédente nécessite des nombres entiers de ressources, il faut compter ici le nombre de demi-slots occupés. On a donc deux classes de clients, l'une avec $s_1 = 1$ et l'autre $s_2 = 2$. Prenons comme exemple une cellule avec trente TCH soit soixante demi-slots. Le tableau suivant donne les taux de perte de chaque classe en fonction des charges.

		Charge des appels de classe 1			
		10	20	30	40
Charge des appels de classe 2	10	0,004 %	0,24	2,47 %	8,34 %
		0,1 %	0,55 %	5,29 %	16,71 %
	20	2,62 %	7,48 %	13,9 %	20,62 %
		5,53 %	14,99 %	26,54 %	37,66 %

Tableau 6.4 – Taux de perte des différentes classes en fonction de la charge

6.8. Réseaux hiérarchiques

On a déjà vu que plus l'on augmente le nombre de serveurs moindre est la perte. Une bonne mesure de cette rentabilité peut être le coût de l'Erlang transporté : considérons un système à S serveurs, sans salle d'attente pour lequel on fixe une borne supérieure α à la probabilité de perte. Le trafic maximal $\rho_{S,\alpha}$ qui peut passer dans ce système est défini par l'équation :

$$\text{Er}[\rho_{S,\alpha}, S] = \alpha.$$

Le rendement du système est alors défini comme le rapport de $\rho_{S,\alpha}$ par S :

$$R = \frac{\rho_{S,\alpha}}{S}.$$

Comme le montre la figure 6.5, le rendement augmente quand le nombre de serveurs augmente.

Mais si l'on prend en compte le coût de l'installation d'un serveur, il vient un moment où l'augmentation du coût compense l'augmentation du gain de rendement. La solution mise en œuvre dans le réseau téléphonique classique consiste à hiérarchiser le réseau. Les abonnés sont tous reliés à un commutateur que l'on qualifiera ici de commutateur de niveau 0, les commutateurs de niveau 0 sont reliés à des commutateurs reliés à des commutateurs de niveau 1 en moindre nombre. On continue ainsi jusqu'au niveau 3 dans la hiérarchie. Pour mémoire, du temps de sa splendeur, le réseau téléphonique commuté français comportait environ 1 500 commutateurs de niveau 0, quelques centaines de commutateurs de niveau 1 et 2 et sept commutateurs de niveau 3.

L'établissement d'une communication se faisait en tentant toujours d'utiliser des jonctions de plus bas niveau possible. S'il existait une jonction entre les deux commutateurs de niveau 0, on la prenait. Sinon, le commutateur de niveau 0 envoyait la gestion de l'appel à son commutateur de niveau 1, qui lui-même essayait d'acheminer l'appel en restant à son niveau, etc.

Si le RTC en lui-même est aujourd'hui bien obsolète, il n'empêche que nombre de ses inventions ont été reprises dans des systèmes modernes, notamment les canaux de garde et les réseaux hiérarchiques.

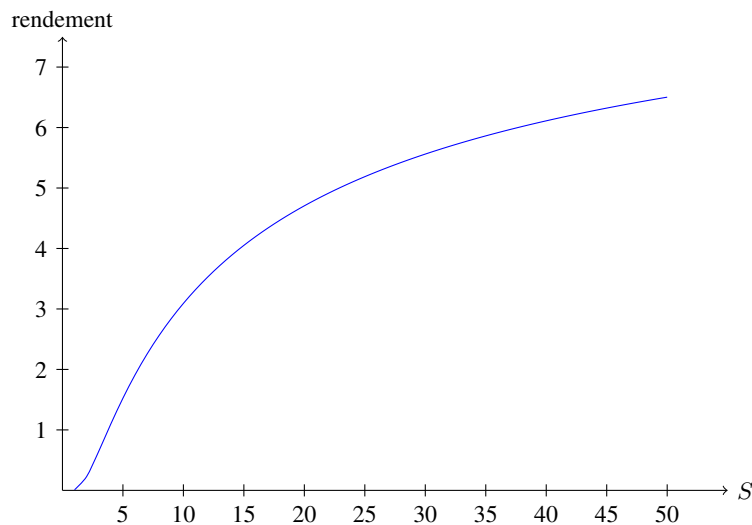


Figure 6.5 – Le rendement en fonction de S

Prenons le cas du GSM. En milieu urbain, les cellules, c'est-à-dire la zone gérée par une antenne ou station de base, sont petites et les *hand-over* fréquents pour les usagers « rapides » tels les automobilistes. On ajoute donc des antennes plus puissantes que les autres, qui couvrent une zone plus large et qui vont offrir deux services. Premièrement, prendre en charge les appels des mobiles rapides pour ainsi diminuer la fréquence des *hand-over*. Deuxièmement, acheminer les appels en débordement des microcellules qu'elles couvrent.

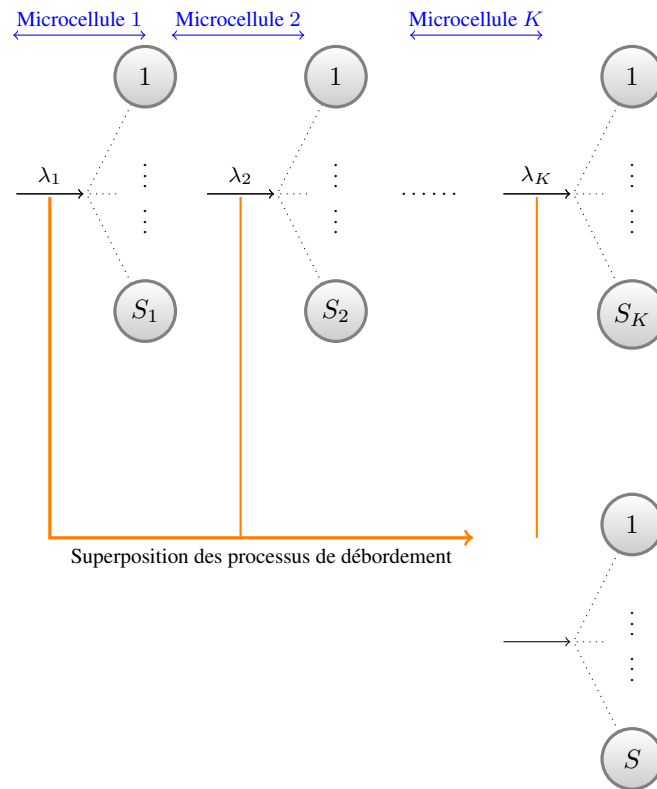


Figure 6.6 – Le trafic écoulé par la macrocellule est la somme des trafics de débordement des microcellules

L'une des questions qui se pose est celle du dimensionnement du nombre d'émetteurs-récepteurs dans la cellule de plus haut niveau. En effet, le processus des arrivées en débordement d'une microcellule n'est pas un processus de Poisson mais un MMPP (voir 4.6, exemple 4.2). Pour la microcellule numéro i , le processus des phases du processus de débordement a pour générateur infinitésimal Q_i celui du nombre de clients

dans une file $M_{\lambda_i}/M_{\mu_i}/S_i/S_i$ et pour fonction de taux λ_i donnée par :

$$\lambda_i(j) = 0 \text{ pour } j < S_i \text{ et } j = S_i.$$

Compte tenu du théorème 4.23, le processus des phases du MMPP global noté J a pour générateur $Q = Q_1 \oplus \dots \oplus Q_K$, la fonction de taux $\lambda = \lambda_1 \oplus \dots \oplus \lambda_K$.

On a vu que la superposition de MMPP est encore un MMPP et un processus de Poisson est un cas particulier de MMPP. Par conséquent, le dimensionnement de la macrocellule revient à étudier la file MMPP/M/S/S. Notons X le nombre de serveurs occupés et J le processus des phases. Ce processus n'est pas markovien parce que sans savoir la phase, on ne peut pas savoir dans combien de temps aura lieu la prochaine arrivée. Dans ces conditions, l'astuce consiste à augmenter l'espace d'états du processus. En l'occurrence, le processus (J, X) est markovien. Son générateur infinitésimal A s'écrit par blocs sous la forme suivante :

$$\begin{pmatrix} Q - \Lambda & \Lambda & & & & \\ \mu I_m & Q - \Lambda - \mu I_m & & & & \\ & 2\mu I_m & \Lambda & & & \\ & & Q - \Lambda - 2\mu I_m & \Lambda & & \\ & & \vdots & \vdots & \ddots & \\ & (0) & S\mu I_m & Q - \Lambda - S\mu I_m & \Lambda & \\ & & & S\mu I_m & Q - \Lambda - S\mu I_m & \Lambda \\ & & & & \ddots & \\ & & & & & S\mu I_m & Q - S\mu I_m \end{pmatrix} \quad (0)$$

La probabilité invariante s'écrit comme d'habitude :

$$\pi A = 0, \quad \pi \mathbf{e} = 1,$$

où \mathbf{e} est le vecteur à $(S+1).m$ composantes toutes égales à 1. On note π la probabilité stationnaire de ce processus. On pose $x_i = (\pi(i, 1), \dots, \pi(i, m))$. Les x_i sont solutions du système :

$$\begin{cases} x_0(Q_{mc} - \Lambda_{mc}) + \theta x_1 = 0 \\ x_0\Lambda_{mc} + x_1(Q_{mc} - \Lambda_{mc} - \theta I_m) + 2\theta x_2 = 0 \\ x_{n-1}\Lambda_{mc} + x_n(Q_{mc} - \Lambda_{mc} - n\theta I_m) + (n+1)\theta x_{n+1} = 0 \end{cases}$$

La dernière équation est :

$$x_{L-1}\Lambda_{mc} + x_L(Q_{mc} - L\theta n I_m) = 0. \quad (6.9)$$

De ces équations, on tire :

$$\begin{cases} x_n = x_0 R_n \text{ avec } R_{-1} = 0, R_0 = I_m, \\ R_{n+1} = -\frac{1}{\theta(n+1)}(R_{n-1}\Lambda_{mc} + R_n(Q_{mc} - \Lambda_{mc} - n\theta I_m)). \end{cases} \quad (6.10)$$

Cela détermine les x_i en fonction de x_0 . Pour déterminer x_0 , nous avons deux possibilités.

– *Méthode 1.* La probabilité stationnaire du nombre de serveurs occupés dans chaque cellule est donnée par la formule d'Erlang :

$$\nu(i) = \frac{\rho^i / i!}{\sum_{j=0}^S \rho^j / j!}.$$

Comme les cellules sont indépendantes les unes des autres, la probabilité stationnaire du processus des phases est le produit tensoriel de ces trois vecteurs :

$$\nu_{mc} = \nu \otimes \nu \otimes \nu.$$

On remarque que :

$$\sum_{n=0}^L \pi(n, j) = \nu_{mc}(j).$$

On introduit alors f_n le vecteur ligne à m composantes dont seule la n^e est non nulle et vaut 1. On introduit aussi e_n le transposé de f_n . La dernière relation s'écrit alors :

$$\sum_{n=0}^L x_0 R_n \cdot e_j = \nu_{mc}(j).$$

Comme $e_j \cdot f_j$ est la matrice dont le seul coefficient non nul est le coefficient (j, j) , on a $\sum_{j=1}^m e_j \cdot f_j = Id_m$. D'autre part :

$$\nu_{mc} = \sum_{j=1}^m \nu_{mc}(j) f_j,$$

on a donc :

$$x_0 \sum_{l=0}^L R_l = \nu_{mc}. \quad (6.11)$$

– *Méthode 2.* En utilisant (6.10) dans (6.9), on obtient :

$$x_0 (R_{L-1} \Lambda_{mc} + R_L (Q_{mc} - L\theta)) = 0.$$

Cette équation détermine toutes les composantes de x_0 sauf une. Ensuite, il faut utiliser la condition de normalisation pour calculer la valeur de cette composante.

Enfin, en vertu du théorème 4.24, la probabilité de perte est donnée par :

$$\left(\nu_{mc} \Lambda_{mc} e \right)^{-1} x_L \Lambda_{mc} e, \quad (6.12)$$

où $e = \sum_{n=1}^m e_n$.

6.9. Un modèle avec renoncement

On modélise ici une limitation forcée de la charge du système, non pas par la limitation de sa capacité, mais par l'introduction d'un filtre aléatoire à l'entrée : plus le système est engorgé, moins le client a de chance d'y entrer.

Avec les notations habituelles, on considère un système $M_\lambda/M_\mu/S$, dans lequel les clients qui arrivent font un tirage au sort (indépendant d'un client à l'autre et tous les autres paramètres) pour décider de leur entrée dans le système. On note $(X^R(t), t \geq 0)$, le processus comptant le nombre de clients dans le système (et tous les paramètres seront affectés de l'exposant R pour « renoncement »). Pour le n -ème client entrant, le tirage au sort est une expérience de Bernoulli de probabilité $p(X^R(T_n^-))$, dépendant donc de la congestion à l'arrivée du client. Si la variable de Bernoulli vaut 1, le client entre et attendra dans le système jusqu'à la fin de son service. Sinon, le client ne rentre même pas dans le système et est définitivement perdu. Il est donc cohérent de supposer que la fonction $p(\cdot)$ soit décroissante et nous considérerons par exemple ici que :

$$p(n) = \frac{1}{n+1}, \quad (6.13)$$

qui est la probabilité d'entrée principalement considérée dans la littérature. Dans la suite, on appellera « client entrant », un client arrivant et entrant effectivement dans le système.

Nous allons décrire le comportement stationnaire du système par l'étude du processus $(X^R(t), t \geq 0)$. Commençons par remarquer que l'on a dans ce cas, l'analogie du paradoxe de l'autobus.

LEMME 6.7.— Soit pour tout instant t , $\tilde{W}(t)$ le temps résiduel à t avant la prochaine arrivée d'un client entrant. Pour tout $i \in \mathbf{N}$, $\tilde{W}(t)$ suit la loi $\varepsilon(\lambda p_i)$ conditionnellement à $X^R(t) = i$, c'est-à-dire que pour $x \geq 0$, on a :

$$\mathbf{P}(\tilde{W}(t) \leq x \mid X^R(t) = i) = 1 - e^{-\lambda p_i x}.$$

Démonstration. Notons à nouveau $T_0 = 0$ et T_1, T_2, \dots les instants d'arrivées des clients et pour tout $t \geq 0$, $N(t)$ le nombre de clients arrivés jusqu'à t et $Z(t)$, le nombre d'arrivées nécessaires après t pour voir entrer effectivement le premier client après t . Alors, il est facile de voir que :

$$\tilde{W}(t) = T_{N(t)+Z(t)} - t.$$

D'autre part, $Z(t)$ suit conditionnellement à $\{X^R(t) = i\}$ une loi géométrique de paramètre p_i , puisqu'alors, chaque client entre à partir de t avec probabilité p_i indépendamment des autres et ce jusqu'à la première entrée effective après t .

Pour tout x , on peut donc écrire :

$$\begin{aligned}
& \mathbf{P} \left(\tilde{W}(t) \geq x \mid X^{\mathbf{R}}(t) = i \right) \\
&= \sum_{k \geq 1} \mathbf{P} \left(\tilde{W}(t) \geq x \mid Z(t) = k; X^{\mathbf{R}}(t) = i \right) \mathbf{P} (Z(t) = k \mid X^{\mathbf{R}}(t) = i) \\
&= \sum_{k \geq 1} \mathbf{P} \left(\tilde{W}(t) \geq x \mid Z(t) = k; X^{\mathbf{R}}(t) = i \right) (1 - p_i)^{k-1} p_i \\
&= p_i \sum_{k \geq 1} (1 - p_i)^{k-1} \mathbf{P} (T_{N(t)+k} - t \geq x) \\
&= p_i \sum_{k \geq 1} (1 - p_i)^{k-1} \sum_{j=1}^{\infty} \mathbf{P} (T_{j+k} - t \geq x; N(t) = j).
\end{aligned} \tag{6.14}$$

D'après le paradoxe de l'autobus :

$$p_i \sum_{j=1}^{\infty} \mathbf{P} (T_{j+1} - t \geq x; N(t) = j) = p_i e^{-\lambda x}. \tag{6.15}$$

D'autre part :

$$\begin{aligned}
& p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \mathbf{P} (T_k - t \geq x; N(t) = 0) \\
&= p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \mathbf{P} (T_k \geq t + x; \xi_1 > t) \\
&= p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \{ \mathbf{P} (\xi_1 > t + x) + \mathbf{P} (T_k \geq t + x; t < \xi_1 < t + x) \} \\
&= p_i \sum_{k \geq 2} (1 - p_i)^{k-1} e^{-\lambda(t+x)} \\
&\quad + p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \int_t^{t+x} \lambda e^{-\lambda u} \int_{t+x-u}^{\infty} \lambda^{k-1} \frac{v^{k-2}}{(k-2)!} e^{-\lambda v} \mathrm{d} v \mathrm{d} u \\
&= e^{-\lambda(t+x)} (1 - p_i) + \int_t^{t+x} \lambda e^{-\lambda u} (1 - p_i) e^{-\lambda p_i(t+x-u)} \mathrm{d} u \\
&= -p_i e^{-\lambda(t+x)} + e^{-\lambda t} e^{-\lambda p_i x}. \tag{6.16}
\end{aligned}$$

Ensuite :

$$\begin{aligned}
& p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \sum_{j=1}^{\infty} \mathbf{P}(T_{j+k} - t \geq x; N(t) = j) \\
&= p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \sum_{j=1}^{\infty} \mathbf{P}(T_j \leq t; T_j + \xi_{j+1} > t + x) \\
&\quad + p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \sum_{j=1}^{\infty} \mathbf{P}(T_{j+k} \geq t + x; T_j \leq t; t < T_j + \xi_{j+1} \leq t + x).
\end{aligned} \tag{6.17}$$

Or, comme on l'a vu dans la preuve du paradoxe de l'autobus :

$$\begin{aligned}
& p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \sum_{j=1}^{\infty} \mathbf{P}(T_j \leq t; T_j + \xi_{j+1} > t + x) \\
&= p_i \sum_{k \geq 2} (1 - p_i)^{k-1} e^{-\lambda(t+x)} (e^{\lambda t} - 1) \\
&= (e^{-\lambda x} - e^{-\lambda(t+x)}) (1 - p_i),
\end{aligned} \tag{6.18}$$

alors que :

$$\begin{aligned}
& p_i \sum_{k \geq 2} (1 - p_i)^{k-1} \sum_{j=1}^{\infty} \mathbf{P}(T_{j+k} \geq t + x; T_j \leq t; t < T_j + \xi_{j+1} \leq t + x) \\
&= p_i \int_0^t \sum_{j=1}^{\infty} \left(\lambda^j \frac{u^{j-1}}{(j-1)!} \right) e^{-\lambda u} \int_{t-u}^{t+x-u} \lambda e^{-\lambda v} \\
&\quad \left(\int_{t+x-u-v}^{\infty} \sum_{k=2}^{\infty} \left((1 - p_i)^{k-1} \lambda^{k-1} \frac{w^{k-2}}{(k-2)!} \right) e^{-\lambda w} dw \right) dv du. \\
&= \int_0^t \lambda e^{-\lambda p_i(t+x-u)} \int_{t-u}^{t+x-u} \lambda (1 - p_i) e^{-\lambda(1-p_i)v} dv du \\
&= (e^{-\lambda t} e^{-\lambda p_i x} - e^{-\lambda(t+x)}) \int_0^t \lambda e^{\lambda u} du \\
&= e^{-\lambda p_i x} - e^{-\lambda x} - e^{-\lambda t} e^{-\lambda p_i x} + e^{-\lambda(t+x)}.
\end{aligned} \tag{6.19}$$

En rassemblant (6.15), (6.16), (6.17), (6.18) et (6.19) dans (6.14), on obtient bien que :

$$\begin{aligned}
 \mathbf{P} \left(\tilde{W}(t) \geq x \mid X^R(t) = i \right) \\
 &= p_i e^{-\lambda x} - p_i e^{-\lambda(t+x)} + e^{-\lambda t} e^{-\lambda p_i x} + \left(e^{-\lambda x} - e^{-\lambda(t+x)} \right) (1 - p_i) \\
 &\quad + e^{-\lambda p_i x} - e^{-\lambda x} - e^{-\lambda t} e^{-\lambda p_i x} + e^{-\lambda(t+x)} \\
 &= e^{-\lambda p_i x}.
 \end{aligned}$$

□

Comme précédemment, on déduit du lemme 6.7 que le processus $(X^R(t), t \geq 0)$ est markovien de générateur A^R donné par :

$$\begin{pmatrix}
 -\lambda p_0 & \lambda p_0 & & & & & \\
 \mu & -(\lambda p_1 + \mu) & \lambda p_1 & & & & \\
 0 & \mu & -(\lambda p_2 + \mu) & \lambda p_2 & & & \\
 & & \ddots & \ddots & \ddots & & \\
 & (0) & & \mu & -(\lambda p_i + \mu) & \lambda p_i & \\
 & & & & \mu & -(\lambda p_{i+1} + \mu) & \lambda p_{i+1} \\
 & & & & \ddots & \ddots & \ddots
 \end{pmatrix} \quad (0)$$

On peut alors vérifier qu'à la condition de stabilité :

$$\sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} p_j < +\infty, \quad (6.20)$$

l'unique probabilité stationnaire π^R du processus de congestion est définie par :

$$\begin{aligned}
 \pi^R(0) &= \left(1 + \sum_{i=1}^{\infty} \rho^i \prod_{j=0}^{i-1} p_j \right)^{-1} ; \\
 \pi^R(i) &= \rho^i \left(\prod_{j=0}^{i-1} p_j \right) \pi(0), \quad i \geq 1.
 \end{aligned} \quad (6.21)$$

REMARQUE.— Dans le cas classique, souvent privilégié dans la littérature, où :

$$p_i = \frac{1}{i+1}, \quad i \geq 0,$$

on a $\pi(i) = e^{-\rho} \rho^i / i!$ pour tout $i \geq 0$. Cela signifie que la congestion stationnaire X_{∞} suit une loi $\mathcal{P}(\rho)$, comme pour la file à une infinité de serveurs.

Probabilité de perte

Définissons le processus $(Y(t), t \geq 0)$ par récurrence sur les instants d'arrivée de la manière suivante. On fixe :

$$Y(t) = 0; t \in [0, T_1[,$$

puis pour tout $i \geq 1$:

$$Y(T_i) = \begin{cases} 1 & \text{si le client } C_i \text{ entre dans le système;} \\ 0 & \text{sinon,} \end{cases}$$

et :

$$Y(t) = Y(T_i), t \in [T_i, T_{i+1}[.$$

Il est alors facile de vérifier que le processus couple $((X^R(t), Y(t)), t \geq 0)$ est markovien sur $\mathbf{N} \times \{0, 1\}$; en effet, c'est un processus à trajectoires cadlag (on remarquera que $(Y(t), t \geq 0)$ est constant par morceaux) dont on peut écrire le générateur \tilde{A}^R comme suit :

– pour tout $i \geq 0$, le processus peut quitter l'état $(i, 0)$ pour l'état $(i + 1, 1)$ si une arrivée d'un client effectivement entrant se produit et si $i \geq 1$, pour l'état $(i - 1, 0)$ si un service se termine donc :

$$\tilde{A}^R((i, 0), (i + 1, 1)) = \lambda p_i;$$

$$\tilde{A}^R((i, 0), (i - 1, 0)) = \mu \text{ pour } i \geq 1;$$

– le processus a les mêmes transitions de $(i, 1)$ vers $(i + 1, 1)$ et $(i - 1, 1)$;
– un autre saut possible se fait de $(i, 1)$ vers $(i, 0)$, lorsqu'un client arrive et ne rentre pas donc :

$$\tilde{A}^R((i, 1), (i + 1, 1)) = \lambda p_i;$$

$$\tilde{A}^R((i, 1), (i, 0)) = \lambda(1 - p_i);$$

$$\tilde{A}^R((i, 1), (i - 1, 1)) = \mu \text{ pour } i \geq 1,$$

où l'on applique un résultat analogue au lemme 6.7 pour la deuxième transition.

Nous résolvons le système :

$$\begin{cases} \tilde{\pi}^R \tilde{A}^R &= \mathbf{0}, \\ \tilde{\pi}^R \mathbf{e} &= 1, \end{cases}$$

où \tilde{A}^R est le générateur du processus sur $\mathbf{N} \times \{0, 1\}$ et $\tilde{\pi}^R$ est une probabilité sur $\mathbf{N} \times \{0, 1\}$.

Appelons X_∞^R et Y_∞^R , les variables aléatoires limites des deux processus, si elles existent. Le cas échéant, $\tilde{\pi}^R$ doit vérifier pour tout $i \in \mathbf{N}$ que :

$$\begin{aligned}\pi^R(i) &= \mathbf{P}(X_\infty = i) = \mathbf{P}(X_\infty = i; Y_\infty = 0) + \mathbf{P}(X_\infty = i; Y_\infty = 1) \\ &= \tilde{\pi}^R(i, 0) + \tilde{\pi}^R(i, 1).\end{aligned}\quad (6.22)$$

En particulier, une unique solution au système précédent existe nécessairement et donc le processus $((X^R(t), Y(t)), t \geq 0)$ est ergodique, à la condition de stabilité (6.20). En effet, les séries mises en jeu dans le calcul de $\tilde{\pi}^R$ sont alors nécessairement sommables d'après (6.22). Remarquons ensuite que pour tout $i \in \mathbf{N}$:

$$\mu\pi^R(i+1) - \lambda p_i \pi^R(i) = 0,$$

ce qui donne avec (6.22) que :

$$\mu\tilde{\pi}^R(i+1, 0) + \mu\tilde{\pi}^R(i+1, 1) - \lambda p_i \tilde{\pi}^R(i, 0) - \lambda p_i \tilde{\pi}^R(i, 1) = 0. \quad (6.23)$$

On lit par ailleurs sur le générateur \tilde{A}^R que :

$$-\lambda p_i \tilde{\pi}^R(i, 0) - \mu \tilde{\pi}^R(i, 0) + \lambda \tilde{\pi}^R(i, 1) - \lambda p_i \tilde{\pi}^R(i, 1) + \mu \tilde{\pi}^R(i+1, 0) = 0,$$

ce qui en combinant avec (6.23) implique que pour tout $i \geq 0$, on a :

$$\begin{aligned}\tilde{\pi}^R(i+1, 0) &= (\rho p_i + 1) \tilde{\pi}^R(i, 0) + \rho(p_i - 1) \tilde{\pi}^R(i, 1); \\ \tilde{\pi}^R(i+1, 1) &= -\tilde{\pi}^R(i, 0) + \rho \tilde{\pi}^R(i, 1).\end{aligned}$$

On a donc la relation de récurrence matricielle :

$$\begin{pmatrix} \tilde{\pi}^R(i+1, 0) \\ \tilde{\pi}^R(i+1, 1) \end{pmatrix} = \begin{pmatrix} \rho p_i + 1 & \rho(p_i - 1) \\ -1 & \rho \end{pmatrix} \begin{pmatrix} \tilde{\pi}^R(i, 0) \\ \tilde{\pi}^R(i, 1) \end{pmatrix} \quad (6.24)$$

$$= M_i \begin{pmatrix} \tilde{\pi}^R(i, 0) \\ \tilde{\pi}^R(i, 1) \end{pmatrix}. \quad (6.25)$$

Pour tout $i \geq 1$, notons :

$$A_i = \prod_{j=0}^{i-1} M_j$$

et A_i^1 (respectivement A_i^2) la première (respectivement deuxième) ligne de A_i et rappelons les relations (6.22) et $\sum_{i \in \mathbf{N}} (\tilde{\pi}^R(i, 0) + \tilde{\pi}^R(i, 1)) = 1$. La probabilité $\tilde{\pi}^R$ est

donc complètement définie par :

$$\begin{aligned}
 \begin{pmatrix} \tilde{\pi}^R(i, 0) \\ \tilde{\pi}^R(i, 1) \end{pmatrix} &= A_i \begin{pmatrix} \tilde{\pi}^R(0, 0) \\ \tilde{\pi}^R(0, 1) \end{pmatrix} \\
 &= A_i \left\{ \tilde{\pi}^R((0, 0)) \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ \pi^R(0) \end{pmatrix} \right\}, i \geq 0; \\
 \tilde{\pi}^R(0, 0) &= \left[1 - \sum_{i \in \mathbf{N}} \left\{ A_i^1 \begin{pmatrix} 0 \\ \pi^R(0) \end{pmatrix} + A_i^2 \begin{pmatrix} 0 \\ \pi^R(0) \end{pmatrix} \right\} \right] \\
 &\quad \cdot \left[\sum_{i \in \mathbf{N}} \left\{ A_i^1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + A_i^2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \right]^{-1}.
 \end{aligned}$$

On peut donc calculer numériquement $\tilde{\pi}^R$ en estimant les valeurs des séries de la formule précédente. La probabilité de perte est alors donnée, d'après la propriété PASTA et l'ergodicité du processus $((X(t), Y(t)), t \geq 0)$, par la formule :

$$\begin{aligned}
 P_{\text{PERTE}}^R &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{\{0\}}(Y(T_n^-)) \\
 &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \sum_{i=0}^{\infty} \mathbf{1}_{\{(i, 0)\}}((X^R(T_n^-), Y(T_n^-))) \\
 &= \sum_{i=0}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \mathbf{1}_{\{(i, 0)\}}((X^R(t), Y(t))) dt \\
 &= \sum_{i=0}^{\infty} \tilde{\pi}^R((i, 0)).
 \end{aligned}$$

6.10. Un centre d'appels avec mise en attente et clients impatientes

On a vu que le modèle d'Erlang modélise par exemple un centre d'appels à s postes, où les appels des clients n'aboutissent que si l'une des lignes est libre au moment de leur prise d'appel. Comme l'on va le voir ici, on peut enrichir le modèle précédent dans le cas, où les clients peuvent être mis en attente. Il est alors cohérent de supposer que les clients sont susceptibles de s'impatienter et de raccrocher avant d'avoir vu leur appel aboutir.

Le modèle choisi est donc celui d'une file d'attente avec capacité infinie et clients impatientes : on considère plus précisément une file $M_\lambda/M_\mu/S/S+M_\alpha$, c'est-à-dire que les temps d'interarrivées et les durées d'appels sont exponentiels et les temps de patience des clients $(D_n, n \in \mathbf{Z})$ avant d'accéder à un opérateur sont indépendants

et identiquement distribués de loi $\varepsilon(\alpha)$, où $\alpha > 0$. Soulignons bien ici, comme dans les modèles avec impatience étudiés au chapitre ??, que l’impatience des clients dure jusqu’à l’accès à un opérateur et tout client qui n’a pas raccroché avant cet instant, poursuit alors son appel jusqu’à son terme : avec la terminologie de la section ??, la patience court jusqu’au début du service. Dans toute cette section, les clients sont servis en FCFS.

Notons $(X^1(t), t \geq 0)$, le processus comptant le nombre de clients dans le système à chaque instant (et ajoutons l’exposant ¹ pour « impatience » à tous les paramètres du système). A t , $X^1(t)$ compte les clients en service et les clients en attente, aussi bien ceux qui atteindront un serveur un jour que ceux qui s’impatienteront avant d’y parvenir. Ce processus peut quitter l’état $i \in \mathbb{N}$ pour visiter les états suivants :

- l’état $i + 1$, si une arrivée se produit ;
- l’état $i - 1$, si $i > S$ et l’un des i clients en attente quitte la salle d’attente, soit parce qu’un serveur s’est libéré, soit parce que le temps de patience du client a expiré.

Comme les temps résiduels de patience des $i - S$ clients en attente sont indépendants et suivent à tout instant la loi $\varepsilon(\alpha)$, il est donc aisé de vérifier que $(X^1(t), t \geq 0)$ est markovien, de générateur infinitésimal donné par :

- $A^1(i, i + 1) = \lambda$ pour tout $i \geq 1$;
- $A^1(i, i - 1) = i\mu$ pour tout $i \in \llbracket 1, S \rrbracket$;
- $A^1(i, i - 1) = S\mu + (i - S)\alpha$ pour tout $i \geq S + 1$.

On calcule alors, comme habituellement, la probabilité stationnaire π^1 de $(X^1, t \geq 0)$:

$$\begin{aligned} \pi^1(i) &= \frac{\rho^i}{i!} \pi^1(0) \text{ pour tout } i \in \llbracket 0, S \rrbracket ; \\ \pi^1(i) &= \frac{\lambda^{i-S} \rho^S}{\prod_{j=1}^{i-S} (S\mu + j\alpha) S!} \pi^1(0) \text{ pour tout } i \geq S + 1 ; \\ \pi^1(0) &= \left(\sum_{i=0}^S \frac{\rho^i}{i!} + \sum_{i=1}^{\infty} \frac{\lambda^i \rho^S}{\prod_{j=1}^i (S\mu + j\alpha) S!} \right)^{-1} . \end{aligned}$$

Estimation de la probabilité de perte

Comme dans le modèle d’Erlang, on cherche à dimensionner le système en déterminant la valeur optimale de S pour garantir une probabilité de perte cible. Le calcul exact de cette probabilité est fastidieux et repose sur des arguments techniques qui dépassent le cadre de cet exposé. Nous pouvons néanmoins en donner une estimation heuristique.

Dans ce système, le client C_n est perdu si et seulement s'il trouve dans le système un temps d'attente W_n supérieur à son temps de patience D_n . La probabilité de perte est donc donnée par :

$$P_{\text{PERTE}}^I = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{1}_{\{[0, D_n]\}}(W_n)$$

D'après les résultats de la section ??, il existe un temps d'attente W_∞ jointement stationnaire avec les temps de patience. A supposer que les variables aléatoires limites W_∞ et D soient effectivement indépendantes (comme le sont W_n et D_n pour tout n), le théorème ?? nous donne donc que :

$$\begin{aligned} P_{\text{PERTE}}^I &= \mathbf{P}((W_\infty, D) \in \{(x, y); x \geq y\}) \\ &= \int_{\mathbf{R}_+} \int_0^x \alpha e^{-\alpha y} dy P_{W_\infty}(dx) \\ &= 1 - \mathcal{L}_{W_\infty}(\alpha), \end{aligned}$$

où \mathcal{L}_{W_∞} est la transformée de Laplace de W_∞ .

6.11. Problèmes

EXERCICE 22.— On considère le concentrateur suivant.

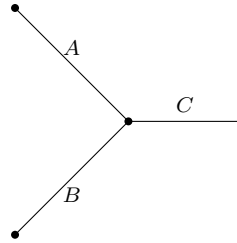


Figure 6.7 – Concentrateur simple

Les clients de classe 1 sont ceux qui empruntent les liaisons A et C. Les clients de classe 2 sont ceux qui empruntent les liaisons B et C. On note S_A , S_B , S_C les capacités de chacun des liens. On pose $X_t = (X_t^1, X_t^2)$, avec X_t^i le nombre d'appels de classe i en cours. Les arrivées des appels de classe i forment un processus de Poisson d'intensité λ_i , $i = 1, 2$. Les durées des appels de classe i sont exponentiellement distribuées de moyenne $1/\mu_i$. On pose $\rho_i = \lambda_i/\mu_i$.

- 1) On suppose dans un premier temps que $S_A = S_B = S_C = \infty$. Ecrire le générateur infinitésimal de X .
- 2) Quelle est sa probabilité stationnaire ?
- 3) Montrer qu'il est réversible.
- 4) Décrire l'espace d'états S lorsque les capacités sont finies.
- 5) En déduire la probabilité stationnaire π de X lorsque les capacités sont toutes finies.
- 6) Montrer que la probabilité p_i de blocage (et donc de pertes) des appels de classe i est de la forme :

$$p_i = 1 - \frac{\sum_{(n_1, n_2) \in S_i} \pi(n_1, n_2)}{\sum_{(n_1, n_2) \in S} \pi(n_1, n_2)},$$

où S_i est un sous-ensemble de S que l'on précisera.

- 7) Application numérique : $S_A = S_B = 2$, $S_C = 3$, $\rho_1 = \rho_2 = 2$. Calculer p_1 .
- 8) Comparer cette probabilité de perte à la quantité :

$$1 - (1 - E[\rho_1, S_A])(1 - E[\rho_1, S_C]).$$

Pourquoi considère-t-on cette quantité ?

6.12. Notes et commentaires

Les calculs explicites pour le dimensionnement avec *hand-over* sont issus d'un cours de X. Lagrange, ceux pour l'interface A-bis ont été faits par N. Dailly. Le dimensionnement des réseaux hiérarchiques est un ancien problème. Les anciennes méthodes reposaient sur le faisceau équivalent de Wilkinson et la méthode Kuczura (voir [IVE 01]). L'approche par les processus MMPP date de [MEI 89]. Son application aux réseaux radio mobiles est inspirée de [LAG 96].

Résumé

– La probabilité de perte n'est égale à la probabilité de blocage que si le processus d'arrivée est un processus de Poisson. Dans les autres cas, on peut se référer au théorème A.34.

– Dans la file M/M/S/S, la probabilité de perte est donnée par la formule d'Erlang :

$$\text{Er}[\rho, S] = \frac{\rho^S / S!}{\sum_{i=0}^S \rho^i / i!}.$$

– Dans le modèle d'Engset, la probabilité de perte est donnée par :

$$\text{Eng}[\rho, S, M] = \frac{\rho^S C_{M-1}^S}{\sum_{j=0}^S C_{M-1}^j \rho^j}.$$

où S est le nombre de serveurs et M le nombre de clients.

– La probabilité de perte ne dépend pas uniquement de la charge, voir le cas de la file IPP/M/S/S.

Bibliographie

- [ARC 10] ARCEP, La qualité des services de voix et de données des réseaux mobiles (2G et 3G) en France métropolitaine, ARCEP, 2010.
- [ASM 03] ASMUSSEN S., *Applied probability and queues*, vol. 51, Springer-Verlag, New York, 2003.
- [BAC 84] BACCELLI F., BOYER P., HÉBUTERNE G., « Single-server queues with impatient customers », *Advances in Applied Probability*, vol. 16, n° 4, p. 887-905, 1984.
- [BAC 89] BACCELLI F., MAKOWSKI A., « Multidimensional stochastic ordering and associated random variables », *Operations Research*, vol. 37, n° 3, p. 478-487, 1989.
- [BAC 02] BACCELLI F., BRÉMAUD P., *Elements of Queueing Theory*, Springer-Verlag, Berlin, 2002.
- [BAC 09a] BACCELLI F., BLASZCZYSZYN B., *Stochastic Geometry and Wireless Networks*, vol. I, Now publishers, Hanover, 2009.
- [BAC 09b] BACCELLI F., BLASZCZYSZYN B., *Stochastic Geometry and Wireless Networks*, vol. II, Now publishers, Hanover, 2009.
- [BAL 01] BALDI P., MAZLIAK L., PRIOURET P., *Martingales et chaînes de Markov*, Hermann, Paris, 2001.
- [BON 11] BONALD T., FEUILLET T., *Performances des réseaux et des systèmes informatiques*, Hermès Sciences, Paris, 2011.
- [BOR 84] BOROVKOV A.A., *Asymptotic methods in queueing theory*, John Wiley & Sons, Chichester, 1984.
- [BOR 92] BOROVKOV A.A., FOSS S.G., « Stochastically recursive sequences and their generalizations », *Siberian Advances in Mathematics*, vol. 2, n° 1, p. 16-81, 1992.
- [BOR 94] BOROVKOV A.A., FOSS S.G., « Two ergodicity criteria for stochastically recursive sequences », *Acta Applicandae Mathematicae*, vol. 34, n° 1-2, p. 125-134, 1994.
- [BOR 98] BOROVKOV A.A., *Ergodicity and stability of stochastic processes*, John Wiley & Sons, Chichester, 1998.

- [BRA 90] BRANDT A., FRANKEN P., LISEK B., *Stationary stochastic models*, vol. 78, Akademie-Verlag, Berlin, 1990.
- [BRE 81] BRÉMAUD P., *Point processes and queues, martingale dynamics*, Springer-Verlag, New York, 1981.
- [CHA 99] CHAO X., MIYAZAWA M., PINEDO M., *Queueing networks : customers, signals, and product form solutions*, John Wiley & Sons, Chichester, 1999.
- [CHU 01] CHUNG K., *A course in probability theory*, Academic Press, San Diego, CA, 2001.
- [CIN 75] ÇINLAR E., *Introduction to stochastic processes*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
- [CRO 96] CROVELLA M., BESTAVROS A., « Self-Similarity in World Wide Web Traffic : Evidence and Causes », *ACM Sigmetrics Conference on Measurement and Modeling of Computer Systems*, 1996.
- [DAL 03] DALEY D.J., VERE-JONES D., *An introduction to the theory of point processes.*, vol. I, Springer-Verlag, New York, 2003.
- [DEL 76] DELLACHERIE C., MEYER P., *Probabilités et Potentiel*, vol. 2, Masson, Paris, 1976.
- [DOU 02] DOUKHAN P., OPPENHEIM G., TAQQU M., DIR., *Long range dependence : theory and applications*, Birkhäuser, Boston, 2002.
- [ETH 86] ETHIER S., KURTZ T., *Markov Processes : Characterizations and Convergence*, John Wiley & Sons, New York, 1986.
- [FIS 93] FISCHER W., MEIER-HELLSTERN K., « The Markov-modulated Poisson process (MMPP) cookbook », *Performance Evaluation*, vol. 18, n° 2, p. 149-171, 1993.
- [FLI 81] FLIPO D., « Comparaison des disciplines de service des files d'attente $G/G/1$ », *Annales de l'Institut Henri Poincaré. Section B.*, vol. 17, n° 2, p. 191-212, 1981.
- [FOS 81] FOSS S.G., « Comparison of service disciplines in multichannel systems with waiting », *Sibirskii Matematicheskii Zhurnal*, vol. 22, n° 1, p. 190-197, 1981.
- [GRA 08] GRAHAM C., *Chaînes de Markov*, Dunod, Paris, 2008.
- [HAE 08] HAENGGI M., GANTI R., « Interference in Large Wireless Networks », *Foundations and Trends in Networking*, vol. 3, n° 2, p. 127-248, 2008.
- [HOU 02] HOUDRÉ C., PRIVAULT N., « Concentration and deviation inequalities in infinite dimensions via covariance representations », *Bernoulli*, vol. 8, n° 6, p. 697-720, 2002.
- [IVE 01] IVERSEN V., *Teletraffic engineering and network planning*, Technical University of Denmark, 2001.
- [JAC 79] JACOD J., *Calcul stochastique et problèmes de martingales*, Springer-Verlag, Berlin, 1979.
- [KAL 98] KALLENBERG O., *Foundations of Modern Probability*, Springer-Verlag, New York, 1998.
- [KEL 79] KELLY F., *Reversibility and stochastic networks*, John Wiley & Sons, Chichester, 1979.

- [KLE 76] KLEINROCK L., *Queueing systems*, John Wiley & Sons, Chichester, 1976.
- [LAG 96] LAGRANGE X., GODLEWSKI P., « Performance of a hierarchical cellular network with mobility dependent hand-over strategies », *Vehicular Technology Conference*, p. 1868-1872, 1996.
- [LAS 95] LAST G., BRANDT A., *Marked point processes on the real line*, Springer-Verlag, New York, 1995.
- [LEL 94] LELAND W., TAQQU M., WILLINGER W., WILSON D., « On the Self-similar Nature of Ethernet Traffic », *IEEE/ACM transactions on networking*, vol. 2, n° 1, p. 1-15, 1994.
- [LOY 62] LOYNES R.M., « Stationary waiting-time distributions for single-server queues », *Annals of Mathematical Statistics*, vol. 33, p. 1323-1339, 1962.
- [MEI 89] MEIER-HELLSTERN K., « The analysis of a queue arising in overflow models », *IEEE Transactions on Communications*, vol. 37, n° 4, p. 367-372, 1989.
- [MOY 08a] MOYAL P., « Convex comparison of service disciplines in real time queues », *Operations Research Letters*, vol. 36, n° 4, p. 496-499, 2008.
- [MOY 08b] MOYAL P., « Stability of a processor-sharing queue with varying throughput », *Journal of Applied Probability*, vol. 45, n° 4, p. 953-962, 2008.
- [MOY 10] MOYAL P., « The queue with impatience : construction of the stationary workload under FIFO », *Journal of Applied Probability*, vol. 47, n° 2, p. 498-512, 2010.
- [NEU 94] NEUTS M., *Matrix-geometric solutions in stochastic models*, Dover Publications, New York, 1994.
- [NEV 84] NEVEU J., « Construction de files d'attente stationnaires », *Modelling and performance evaluation methodology (Paris, 1983)*, vol. 60 de *Lecture Notes in Control and Information Science*, p. 31-41, Springer, Berlin, 1984.
- [NOR 94] NORROS I., « A Storage Model with Self-Similar Inputs », *Queueing Systems*, vol. 16, p. 387-396, 1994.
- [PAN 88] PANWAR S., TOWSLEY D., WOLF J., « Optimal scheduling policies for a class of queues with customer deadlines to the beginning of service », *Journal of the Association for Computing Machinery*, vol. 35, n° 4, p. 832-844, 1988.
- [PRI 09] PRIVAULT N., *Stochastic analysis in discrete and continuous settings with normal martingales*, vol. 1982, Springer-Verlag, Berlin, 2009.
- [RIG 98] RIGAULT C., *Principes de commutation numérique*, Hermès, Paris, 1998.
- [ROB 03] ROBERT P., *Stochastic networks and queues*, vol. 52, Springer-Verlag, Berlin, 2003.
- [RUD 80] RUDIN W., *Analyse réelle et complexe*, Masson, Paris, 1980.
- [SAM 94] SAMORODNITSKY G., TAQQU M., *Stable non-Gaussian random processes*, Chapman & Hall, New York, 1994.
- [SHE 97] SHERMAN R., TAQQU M., WILLINGER W., « Proof of a fundamental result in self-similar traffic modeling », *Computer Communication Review*, vol. 27, n° 2, 1997.

- [STO 83] STOYAN D., *Comparison methods for queues and other stochastic models*, John Wiley & sons, Chichester, 1983.
- [THO 00] THORISSON H., *Coupling, stationarity, and regeneration*, Springer-Verlag, New York, 2000.
- [WU 00] WU L., « A new modified logarithmic Sobolev inequality for Poisson point processes and several applications », *Probability Theory and Related Fields*, vol. 118, n° 3, p. 427-438, 2000.
- [YOS 95] YOSIDA K., *Functional analysis*, Springer-Verlag, Berlin, 1995.