

# Files d'attente

Dimitri LOZEVE

2014

## Résumé

Initiée dans le domaine des télécommunications, la théorie des files d'attente est un modèle mathématique permettant de représenter une multitude de phénomènes, de la physique à l'économie. Elle est notamment appelée à jouer un grand rôle dans l'étude et la compréhension des échanges rapides entre serveurs dans le cadre des technologies de communication modernes. Dans ce TIPE, nous nous intéresserons particulièrement à l'étude de la convergence des files d'attente, grâce à un cadre mathématique rigoureux, fondé sur l'algèbre et la théorie des chaînes de Markov en temps continu, ainsi que grâce à des simulations numériques pour mieux étudier l'équilibre.

## Table des matières

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Chaînes de Markov en temps continu</b>   | <b>2</b> |
| 1.1      | Définitions . . . . .                       | 2        |
| 1.2      | Semigroupe de transition . . . . .          | 2        |
| 1.3      | Générateur infinitésimal . . . . .          | 2        |
| 1.4      | Distributions stationnaires . . . . .       | 3        |
| 1.5      | Convergence vers l'équilibre . . . . .      | 3        |
| <b>2</b> | <b>Files d'attente finies</b>               | <b>4</b> |
| 2.1      | Définition . . . . .                        | 4        |
| 2.2      | Notation de Kendall . . . . .               | 5        |
| 2.3      | La file $M/M/1/n$ . . . . .                 | 5        |
| 2.4      | La file $M/M/n/n$ (file d'Erlang) . . . . . | 7        |
| <b>3</b> | <b>Files d'attente infinies</b>             | <b>8</b> |
| 3.1      | Problématique . . . . .                     | 8        |
| 3.2      | La file $M/M/1$ . . . . .                   | 9        |
| 3.3      | Simulations du spectre . . . . .            | 9        |

# 1 Chaînes de Markov en temps continu

## 1.1 Définitions

DÉFINITION 1.1.  $\{X(t), t \geq 0\}$  est une *chaîne de Markov en temps continu* à états dans  $\mathcal{S}$  si :

$$\begin{aligned} \forall(t, s, s_1, \dots, s_k, i, j, i_1, \dots, i_k) \in \mathbb{R}_+^{k+2} \times \mathcal{S}^{k+2}, \\ \mathbb{P}(X(t+s) = j \mid X(s) = i, X(s_1) = i_1, \dots, X(s_k) = i_k) \\ = \mathbb{P}(X(t+s) = j \mid X(s) = i) \end{aligned}$$

DÉFINITION 1.2. Une chaîne de Markov en temps continu  $X$  est dite *homogène* si :

$$\begin{aligned} \forall(t, s, i, j) \in \mathbb{R}_+^2 \times \mathcal{S}^2, \\ \mathbb{P}(X(t+s) = j \mid X(s) = i) = \mathbb{P}(X(t) = j \mid X(0) = i). \end{aligned}$$

On considère ici un espace d'états  $\mathcal{S}$  **fini**.

## 1.2 Semigroupe de transition

DÉFINITION 1.3. On définit la matrice  $\mathbf{P}$  par :

$$\forall(i, j, t) \in \mathcal{S}^2 \times (\mathbb{R}_+^*)^2, p_{i,j}(t) = \mathbb{P}(X(t) = j \mid X(0) = i).$$

PROPOSITION 1.1 (Chapman-Kolmogorov).

$$\forall(t, s) \in \mathbb{R}_+^2, \mathbf{P}(t+s) = \mathbf{P}(s)\mathbf{P}(t).$$

En particulier :  $\mathbf{P}(0) = \mathbf{I}$ .

## 1.3 Générateur infinitésimal

**Continuité de  $\mathbf{P}$**  On suppose  $\mathbf{P}$  continue en 0. D'après la propriété de Chapman-Kolmogorov (proposition 1.1),  $\mathbf{P}$  est dérivable à droite sur tout  $\mathbb{R}_+$ .

Comme  $\mathbf{P}$  est continue en 0 et  $\mathbf{P}(0) = \mathbf{I}$ , il existe  $h > 0$  tel que  $\mathbf{P}(h)$  soit inversible.

Soit  $t > 0$ . D'après Chapman-Kolmogorov, on a :

$$\mathbf{P}(t-h)\mathbf{P}(h) = \mathbf{P}(t),$$

d'où :

$$\mathbf{P}(t-h) = \mathbf{P}(t)\mathbf{P}(h)^{-1}.$$

Or  $\lim_{h \rightarrow 0} \mathbf{P}(h) = \mathbf{I}$ , donc  $\mathbf{P}$  est continue à gauche en  $t$ . Donc  $\mathbf{P}$  est continue en tout point de  $\mathbb{R}_+$ .

**Dérivabilité de  $\mathbf{P}$**  Soit  $a \geq 0$  et  $h > 0$ .

$$\begin{aligned} \int_a^{a+h} \mathbf{P}(s) ds &= \int_0^h \mathbf{P}(a+u) du \\ &= \mathbf{P}(a) \int_0^h \mathbf{P}(u) du. \end{aligned}$$

$$\text{Or } \frac{1}{h} \int_0^h \mathbf{P}(u) \, du \xrightarrow{h \rightarrow 0} \mathbf{P}(0) = \mathbf{I}.$$

Donc il existe  $h > 0$  tel que  $\int_0^h \mathbf{P}(u) \, du$  soit inversible.

$$\forall a \geq 0, \mathbf{P}(a) = \left( \int_a^{a+h} \mathbf{P}(s) \, ds \right) \left( \int_0^h \mathbf{P}(u) \, du \right)^{-1}$$

et  $\int_a^{a+h} \mathbf{P}(s) \, ds$  est dérivable. Donc pour tout  $a \geq 0$ ,  $\mathbf{P}$  est dérivable en  $a$ . Donc  $\mathbf{P}$  est dérivable sur  $\mathbb{R}_+$ .

### Définition de Q

DÉFINITION 1.4. On définit  $\mathbf{Q} := \mathbf{P}'(0)$ .  $\mathbf{Q}$  est appelé le *générateur infinitésimal* de la chaîne de Markov.

D'après Chapman-Kolmogorov,

$$\forall (t, s) \in \mathbb{R}_+^2, \mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s).$$

En dérivant en  $t = 0$  (car  $\mathbf{P}$  est dérivable) :

$$\forall s \in \mathbb{R}_+, \mathbf{P}'(s) = \mathbf{P}'(0)\mathbf{P}(s) = \mathbf{Q}\mathbf{P}(s).$$

De même,

$$\forall s \in \mathbb{R}_+, \mathbf{P}'(s) = \mathbf{P}(s)\mathbf{Q}.$$

$\mathbf{P}$  est donc solution d'une équation différentielle du premier ordre. On a alors :

$$\begin{aligned} \forall s \in \mathbb{R}_+, \mathbf{P}(s) &= \exp(s\mathbf{Q}) \\ &= \sum_{n=0}^{+\infty} \frac{\mathbf{Q}^n}{n!}. \end{aligned}$$

## 1.4 Distributions stationnaires

DÉFINITION 1.5. Un vecteur  $\pi$  est une *distribution stationnaire* si

$$\forall t \geq 0, \pi\mathbf{P}(t) = \pi.$$

PROPOSITION 1.2.  $\pi$  est une *distribution stationnaire* si, et seulement si,  $\pi\mathbf{Q} = 0$ .

## 1.5 Convergence vers l'équilibre

LEMME 1.3. Pour tout  $(i, j) \in \mathcal{S}^2$ , la fonction  $t \mapsto p_{i,j}(t)$  est uniformément continue.

THÉORÈME 1.4. Soit  $\mathbf{Q}$  un *générateur infinitésimal irréductible et non explosif*, de *semigroupe de transition*  $t \mapsto P(t)$ , et  $\pi$  sa *distribution d'équilibre*. Alors :

$$\forall (i, j) \in \mathcal{S}^2, p_{i,j}(t) \xrightarrow{t \rightarrow +\infty} \pi_j.$$

*Démonstration.* Soit  $(X(t))$  une chaîne de Markov de transitions représentées par le générateur infinitésimal  $\mathbf{Q}$ . On fixe  $h > 0$ . Pour tout  $n \in \mathbb{N}$ , on pose  $Z_n := X(nh)$ . Alors, pour tout  $n \in \mathbb{N}$  et pour toute famille d'états  $(i_0, \dots, i_{n+1}) \in \mathcal{S}^{n+2}$ , on a :

$$\begin{aligned} & \mathbb{P}(Z_{n+1} = i_{n+1} \mid Z_0 = i_0, \dots, Z_n = i_n) \\ &= \mathbb{P}(X((n+1)h) = i_{n+1} \mid X_0 = i_0, \dots, X(nh) = i_n) \\ &= p_{i_n, i_{n+1}}((n+1)h - nh) = p_{i_n, i_{n+1}}(h). \end{aligned}$$

Donc  $(Z_n)$  est une chaîne de Markov à temps discret de matrice de passage  $\mathbf{P}(h)$ . De plus, comme on sait que  $\forall (i, j) \in \mathcal{S}^2, p_{i,j}(h) > 0$ ,  $\mathbf{P}(h)$  est irréductible et apériodique.

On admet le théorème de convergence pour les chaînes de Markov à temps discret. On a alors :

$$\forall (i, j) \in \mathcal{S}^2, p_{i,j}(nh) \xrightarrow[n \rightarrow +\infty]{} \pi_j.$$

Soit  $(i, j) \in \mathcal{S}^2$ . On fixe  $\varepsilon > 0$ . Comme  $p_{i,j}$  est uniformément continue, il existe  $h > 0$  tel que :

$$\forall t \in \mathbb{R}_+, |p_{i,j}(t) - p_{i,j}(nh)| \leq \varepsilon/2,$$

où  $n := \mathbb{E}(\frac{t}{h})$ .

De plus, il existe  $N \in \mathbb{N}$  tel que  $\forall n \in \mathbb{N}, n \geq N \implies |p_{i,j}(nh) - \pi_j| \leq \varepsilon/2$ .

Donc, pour tout  $t \in \mathbb{R}_+$ , tel que  $t \geq Nh$ , on a, avec  $n := \mathbb{E}(\frac{t}{h})$  :

$$\begin{aligned} |p_{i,j}(t) - \pi_j| &\leq |p_{i,j}(t) - p_{i,j}(nh)| + |p_{i,j}(nh) - \pi_j| \\ &\leq \varepsilon/2 + \varepsilon/2 \\ &\leq \varepsilon. \end{aligned}$$

Donc  $p_{i,j}(t) \xrightarrow[t \rightarrow +\infty]{} \pi_j$ . □

## 2 Files d'attente finies

### 2.1 Définition

DÉFINITION 2.1. On appelle *file d'attente* un système caractérisé par :

1. un processus d'entrée des clients ;
2. un processus de service ;
3. un nombre de serveurs (ou capacité de service) ;
4. la taille de la file d'attente proprement dite ;
5. une discipline de service.

Les processus aléatoires utilisés peuvent être déterministes, markoviens (c'est-à-dire des processus de Poisson), géométriques, etc. Dans la suite, on étudiera seulement des files d'attentes régies par des processus markoviens.

La discipline de la file d'attente décrit la manière dont les clients sont servis : premier arrivé, premier servi (LIFO, *last in first out*), premier arrivé, dernier servi (FIFO, *first in first out*), ou aléatoire.

Enfin, le nombre de serveurs et la taille de la file sont des éléments de  $\mathbb{N} \cup \{+\infty\}$ .



### 2.3.2 Distribution d'équilibre

En résolvant  $\pi Q = 0$ , on obtient les relations de récurrence suivantes :

$$\begin{cases} \lambda\pi_0 = \mu\pi_1 \\ \forall k \in \llbracket 1, n-1 \rrbracket, (\lambda + \mu)\pi_k = \lambda\pi_{k-1} + \mu\pi_{k+1} \\ \lambda\pi_{n-1} = \mu\pi_n. \end{cases}$$

Donc  $\forall k \in \llbracket 0, n-1 \rrbracket, \pi_k = \rho^k \pi_0$ , où  $\rho = \lambda/\mu$ .

Le vecteur  $\pi$  doit être stochastique, on a donc :

$$\sum_{k=0}^n \pi_k = 1$$

$$\pi_0 = \frac{1 - \rho}{1 - \rho^n}$$

D'où finalement :

$$\forall k \in \llbracket 0, n \rrbracket, \pi_k = \rho^k \frac{1 - \rho}{1 - \rho^n}.$$

### 2.3.3 Convergence vers l'équilibre

Pour contrôler la convergence du semigroupe de transition vers la distribution d'équilibre (théorème 1.4), on peut réaliser des simulations numériques. On a pris ici  $\lambda = 0.4$ ,  $\mu = 0.6$ , et  $n = 10$ .

On constate que chaque ligne de la matrice tend vers le vecteur  $\pi$  calculé précédemment, mais à des vitesses différentes (figure 1).

D'autre part, il est possible de représenter graphiquement la convergence de chaque coefficient de la matrice vers le coefficient correspondant de la distribution d'équilibre (figure 2). Cependant, on constate que la convergence n'est pas identique pour tous les coefficients, notamment du point de vue de la monotonie.

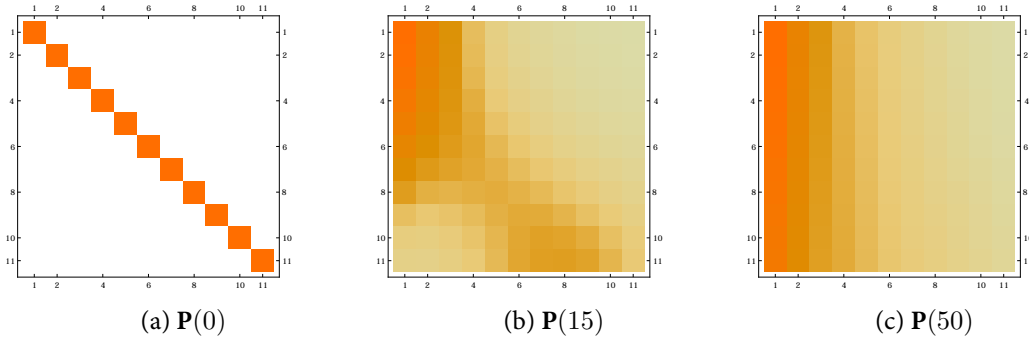


FIG. 1 : Convergence de chaque ligne de  $M/M/1/n$

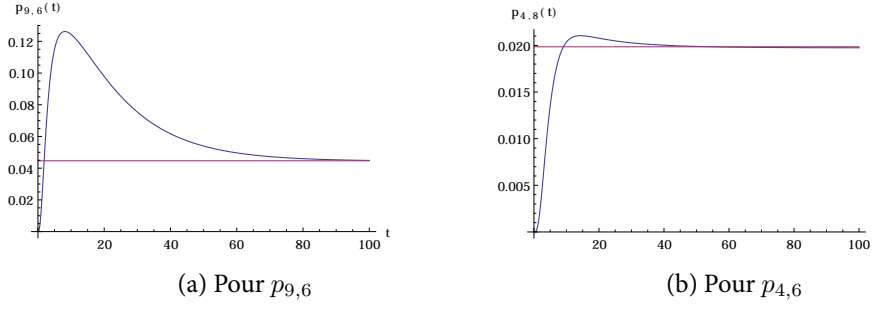


FIG. 2 : Exemples de convergence de M/M/1/n

## 2.4 La file M/M/n/n (file d'Erlang)

### 2.4.1 Générateur infinitésimal

Le générateur infinitésimal de la file M/M/n/n, ou *file d'Erlang*, est le suivant :

$$\mathbf{Q} = \begin{pmatrix}
 -\lambda & \lambda & & & & (0) \\
 \mu & -\lambda - \mu & \lambda & & & \\
 & 2\mu & -\lambda - 2\mu & \lambda & & \\
 & & 3\mu & \ddots & \ddots & \\
 (0) & & & \ddots & -\lambda - (k-1)\mu & \lambda \\
 & & & & k\mu & -k\mu
 \end{pmatrix}$$

Cette fois-ci, le taux de service est proportionnel au nombre de clients dans la file. Cela montre l'existence de plusieurs serveurs, capables de traiter l'ensemble des clients présents dans la file simultanément. La file ne pouvant accueillir plus de  $n$  clients, il y a donc  $n$  serveurs.

### 2.4.2 Distribution d'équilibre

En résolvant la relation de récurrence fournie par  $\pi \mathbf{Q} = 0$ , on trouve :

$$\begin{cases}
 \lambda \pi_0 = \mu \pi_1 \\
 \forall k \in \llbracket 1, n-1 \rrbracket, (\lambda + k\mu) \pi_k = \lambda \pi_{k-1} + (k+1) \mu \pi_{k+1} \\
 \lambda \pi_{n-1} = n \mu \pi_n.
 \end{cases}$$

La résolution de la relation de récurrence donne :

$$\forall k \in \llbracket 0, n \rrbracket, \pi_k = \frac{\rho^k}{k!} \pi_0.$$

Et comme  $\pi$  est un vecteur stochastique :

$$\sum_{k=0}^n \frac{\rho^k}{k!} \pi_0 = 1,$$

$$\forall k \in \llbracket 0, n \rrbracket, \pi_k = \frac{\rho^k}{k!} \left( \sum_{k=0}^n \frac{\rho^k}{k!} \right)^{-1}.$$

### 2.4.3 Convergence vers l'équilibre

De manière similaire au cas de la file  $M/M/1/n$ , on peut simuler le comportement asymptotique de la file pour  $\lambda = 0.4$ ,  $\mu = 0.6$ , et  $n = 10$ . Les figures 3 et 4 représentent cette convergence pour la matrice entière et pour certains coefficients.

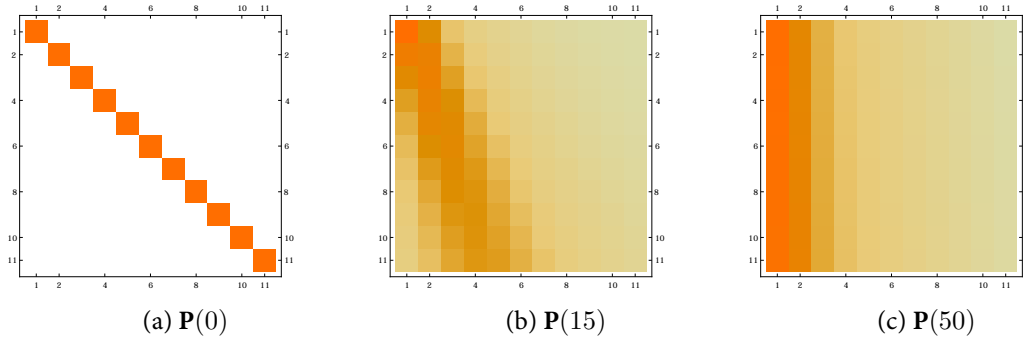


FIG. 3 : Convergence de chaque ligne de  $M/M/n/n$

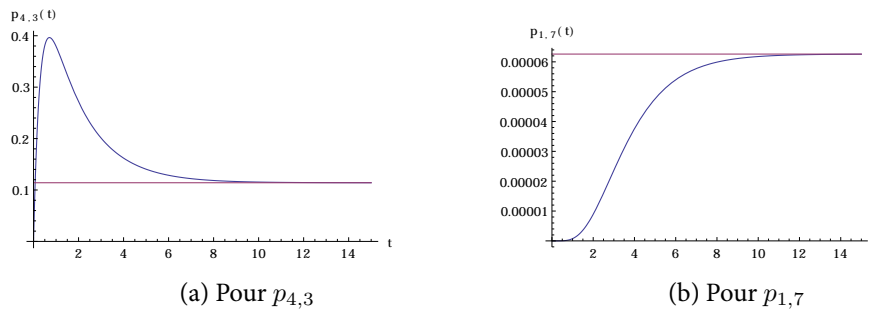


FIG. 4 : Exemples de convergence de  $M/M/n/n$

## 3 Files d'attente infinies

### 3.1 Problématique

Il semble naturel dès le départ d'étendre cette définition des files d'attente à des cas où la file serait infinie : elle ne refuse plus les clients passé un certain nombre. On serait tenté pour cela de reprendre la même définition des semigroupes de transitions et des générateurs infinitésimaux ; mais il est alors nécessaire d'interpréter les matrices infinies qui surgissent.

Il s'agit alors d'interpréter le générateur infinitésimal  $\mathbf{Q}$  comme un opérateur borné d'un espace de Hilbert de dimension infinie. Cette définition, alliée à certaines propriétés, permettra ainsi de justifier les opérations réalisées sur ces matrices, et notamment la définition de la quantité  $\exp t\mathbf{Q}$ , pour  $t \geq 0$ .

Enfin, les principales propriétés des générateurs infinitésimaux (et donc des semigroupes de transition), notamment la notion de *convergence*, est liée à leur spectre. Il serait donc intéressant d'étudier les relations qu'entretient la matrice infinie de  $M/M/1$  avec les matrices finies des files  $M/M/1/n$ .



### 3.2 La file M/M/1

La file M/M/1 est une « prolongation » de la file M/M/1/n. Son générateur infinitésimal s'écrit donc naturellement :

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

On pose :  $\forall n \in \mathbb{N}, \pi_n := \rho^n$ , où  $\rho = \lambda/\mu$ . La propriété suivante du vecteur  $\pi$  sera utile dans la suite :

$$\forall n \in \mathbb{N}, \mu\pi_n = \lambda\pi_{n+1}.$$

On définit ensuite l'espace  $\mathcal{H}$  dans lequel on travaille :

$$\mathcal{H} := l^2(\pi) = \left\{ (u_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \sum_{n \geq 0} u_n^2 \pi_n \text{ converge} \right\}.$$

On munit cet espace du produit scalaire :

$$(u, v) \mapsto \langle u, v \rangle := \sum_{n=0}^{+\infty} u_n v_n \pi_n.$$

Cette application est bien un produit scalaire car tous les coefficients du vecteur  $\pi$  sont strictement positifs.  $\mathcal{H}$  est donc un espace préhilbertien réel de dimension infinie.

On montre alors :

$$\forall (x, y) \in \mathcal{H}^2, \langle Qx, y \rangle = \langle x, Qy \rangle.$$

$$\langle Qx, y \rangle \leq 2(\lambda + \mu) \|x\| \|y\|.$$

$\mathbf{Q}$  est symétrique et bornée. Par le théorème spectral :

$$\text{Sp}(\mathbf{Q}) \subset [-2(\lambda + \mu), 2(\lambda + \mu)].$$

### 3.3 Simulations du spectre

On peut alors étudier par des simulations numériques la convergence du spectre de M/M/1/n vers le spectre de M/M/1 lorsque  $n$  tend vers  $+\infty$ . Pour  $\lambda = 0.5$  et  $\mu = 0.6$ , on obtient le graphe de la figure 5. On constate bien que le spectre de M/M/1 est inclus dans l'intervalle  $[-2(\lambda + \mu), 0]$ .

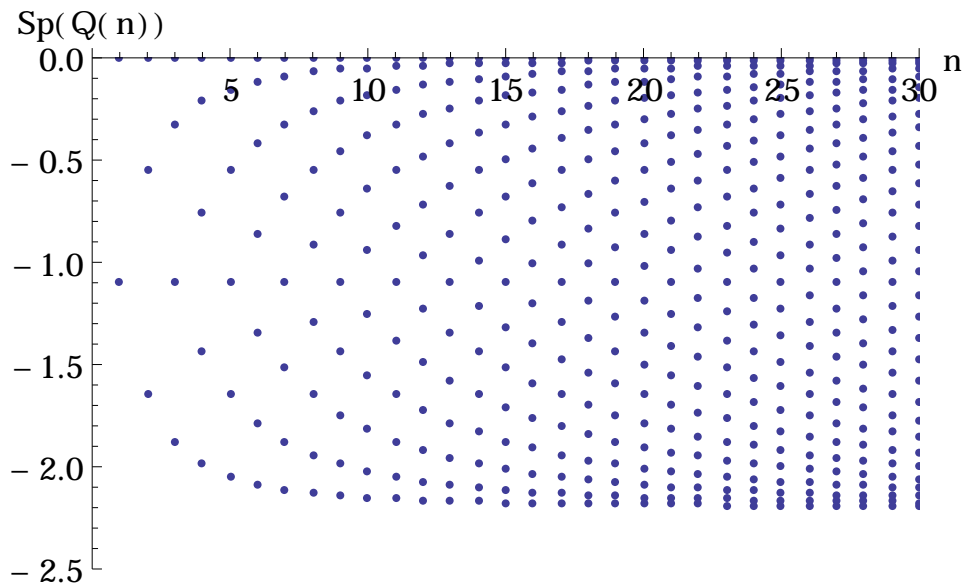


FIG. 5 : Convergence du spectre de  $M/M/1/n$  vers le spectre de  $M/M/1$

## Références

- [1] Ivo Adan and Jacques Resing, *Queueing theory*, 2002.
- [2] Pierre Brémaud, *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*, Springer, 1999.
- [3] Daniel Flipo, *Chaînes de Markov*, <http://daniel.flipo.free.fr/cours/markov.pdf>, Consulté en novembre 2013.
- [4] John G. Kemeny and J. Laurie Snell, *Finite Markov chains*, Springer, 1960.
- [5] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp, *Denumerable Markov chains*, Springer, 1966.
- [6] Carl D. Meyer, *Matrix analysis and applied linear algebra*, SIAM, 2000.
- [7] J. R. Norris, *Markov chains*, Cambridge University Press, 1997.
- [8] Daniel W. Stroock, *An introduction to Markov processes*, Springer, 2005.
- [9] Ward Whitt, *Continuous-time Markov chains*, [www.columbia.edu/~ww2040](http://www.columbia.edu/~ww2040), 2006, Consulté en février 2014.
- [10] Moshe Zukerman, *Introduction to queueing theory and stochastic teletraffic models*, <http://www.ee.cityu.edu.hk/~zukerman/classnotes.pdf>, 2011, Consulté en décembre 2013.