République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2020/2021

Analyse d'un système de files d'attente avec distribution générale du temps de service, vacances du serveur et clients impatients

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar

Discipline : MATHEMATIQUES

Spécialité : ASSPA

par **B**erkane **K**helifa [1]

Sous la direction de

Dr. L. Yahiaoui

Soutenue le 13/07/2021 devant le jury composé de

Dr. R. ROUANE	Université Dr Tahar Moulay - Saïda	Président
Dr. L. Yahiaoui	Université Dr Tahar Moulay - Saïda	Encadreur
Dr. M. Kadi	Université Dr Tahar Moulay - Saïda	Examinateur
Dr. N. Ait Ouali	Université Dr Tahar Moulay - Saïda	Examinatrice

1. e-mail : berkanekhelifa82@gmail.com

Table des matières

1 Introduction aux systèmes de files d'attente **10**

1.1 Processus stochastique : quelques définitions 10

1.1.1 Processus de comptage 11

1.1.2 Processus de renouvellement 11

1.1.3 Processus de Poisson 11

1.1.4 Loi exponentielle 12

1.1.5 Processus de Markov 14

1.1.6 Processus de naissance et de mort 18

1.2 Analyse mathématique d'un système de files d'attente 21

1.2.1 Modèle file d'attente simple 21

1.2.2 Structure et discipline de la file 22

1.2.3 Discipline de service 23

1.2.4 Classification des systèmes d'attente 24

1.2.5 Loi de Little 24

1.2.6 Mesures de performance 25

1.3 Types de files d'attente 26

1.3.1 Modèles markoviens : 26

1.3.2 Modèles semi markoviens : 26

2 Systèmes de files d'attente classiques **28**

2.1 Systèmes de files d'attente régis par un modèle markovien de naissance et de mort 28

2.1.1 Système de files d'attente M/M/1 28

2.1.2 Système de files d'attente M/M/c 34
2.1.3 Système de files d'attente M/M/c/K 38
2.2 Système de files d'attente M/G/1 40

3 Système de files d'attente avec distribution générale du temps de service, vacances du serveur et clients impatients 47
3.1 Description et notations du modèle 48
3.2 Caractéristiques du modèle 49
3.2.1 Durée d'une période de vacances τ 49
3.2.2 Nombre de clients en début de période occupé 49
3.2.3 Période d'occupation 51
3.2.4 Probabilité que le système est vide et le serveur en vacences 51
3.2.5 Temps de vacances et d'impatience distribués de façon exponentielle . 52
3.2.6 Nombre de clients à un instant de fin de service 53
3.2.7 Décomposition stochastique 54
3.2.8 Nombre moyen de clients dans le système 55
3.2.9 Proportion de clients servis 56

Table des figures

1.1 Représentation d'une file d'attente. 21
1.2 File d'attente avec plusieurs serveurs. 23
1.3 Capacité d'une file d'attente. 23

2.1 File d'attante M/M/1. 29
2.2 Diagramme de transition d'état M/M/1. 30
2.3 File d'attante M/M/c. 34
2.4 Diagramme de transition d'état M/M/c. 35
2.5 Diagramme de transition d'état M/M/c/K. 39
2.6 Diagramme de transition d'état M/G/1. 42

REMERCIEMENT

En premier lieu, je tiens à remercier mon **Dieu** qui m'a donné le courage et la volonté pour réaliser ce travail.

Je tiens d'exprimer toute ma reconnaissance à mon encadreur de mémoire et mon professeur Monsieur : **LAHCENE Yahiaoui** de m'avoir encadré orienté aidé et conseillé, et ainsi pour la confiance et la liberté qu'il m'a accordé, sans oublier ses efforts fournis pendant ces deux années.

Je remercie **Dr. ROUANE R.** d'avoir accepté à présider le jury de mon mémoire.

Je souhaite également remercier **Dr. N. Ait Ouali** et **Dr. M. Kadi** de m'avoir fait l'honneur d'accepter d'être examinateurs de mon mémoire. Leurs remarques et précieux conseils ont abouti à une amélioration du document de mémoire. Ma gratitude va également aux enseignants du Département de Mathématiques que j'ai eu durant mon cursus.

J'adresse mes sincères remerciements à tous Les membres du laboratoire des Modèles Stochastiques, Statistique et Applications et à la tête le directeur **Pr. A. Kandouci.**

Je remercie mes très chers parents, qui ont toujours été là pour moi. Enfin je remercie mes amis pour les bons mộments passés ensemble et aussi pour leur soutiens et leur encouragements. Et pour finir, merci à toutes les personnes que j'ai oubliées de citer et qui m'ont permis de mener à bien ce mémoire.

Dédicaces

Je dédie ce travail à : Mes parents, qui ont dépensé pour moi sans compter. En reconnaissance de tous les sacrifices consentis par tous et chacun pour me permettre d'atteindre cette étape de ma vie.

A mon frère. A ma sœur. A mes cousins et cousines.

A mes amis "Elias", "Hamlat", "Merzoug", "Kada", "Aziz", et "Hanane", "Fatima".

Introduction générale

La **théorie des files d'attente** est principalement vue comme une branche de la théorie des probabilités appliquées. Les applications sont dans différents domaines, par exemple : les réseaux de transmission, les systèmes informatiques, les réseaux urbains, les banques, la gestion des avions au décollage ou à l'atterrissage, ...

Un système de files d'attente comprend donc un espace de service avec un ou plusieurs dispositifs de service (serveurs) et un espace d'attente dans lequel se forme une éventuelle file d'attente, le processus décrivant le fonctionnement d'un système de files d'attente est processus aléatoire (stochastique).

Pour identifier un système de files d'attente, on a besoin de spécifier le flux d'entrée, le mécanisme de service et la discipline d'attente.

Depuis les travaux d'Erlang [3] Un grand nombre d'applications dans tous les domaines ont été mis en œuvre et publiées. En 1953, David G. Kendall a introduit la notation de Kendall [3] pour décrire les caractéristiques d'un système de file d'attente. en 1957 d'une manière particulièrement élégante et efficace Jackson a traité certains réseaux de files d'attente. En 1961, Thomas L. Saaty [42], auteur de l'un des premiers livres complets sur la théorie des files d'attente. Ensuite c'est les contributions des mathématiciens Khintchine, Palm, Pollaczek et Kolmogorov [43] qui ont vraiment poussés la théorie des files d'attente.

Les systèmes de file d'attente avec des vacances sur serveur ont attiré l'attention de nombreux chercheurs depuis que l'idée a été discutée pour la première fois dans l'article de Levy et Yechiali [8]. Plusieurs enquêtes sur ces modèles de vacances ont été réalisées par Doshi [16], [18] et les livres de Takagi [26], Tian et Zhang [45] sont consacrés à ce sujet.

Zhang et Hou [29] ont analysé un M/G/1 file d'attente avec des vacances de travail et une interruption de vacances. En utilisant la méthode d'une variable supplémentaire et la méthode d'analyse matricielle.

Altman et Yechiali n'ont considéré que l'impatience des clients lorsque les serveurs sont en vacances et indisponibles pour le service. Selvaraju et Goswami analysés impatients clients dans une file d'attente markovienne de serveur unique avec des vacances de travail uniques et multiples.

Mon mémoire est composé de trois chapitres :

Dans **le premier chapitre**, nous présentons les notions de bases de l'étude des systèmes de files d'attente, à savoir les processus stochastiques (Processus de comptage, processus de renouvellement, processus de Poisson, processus de naissance et de mort), et introduisons certaines définitions et notations sur la théorie des files d'attente comme (Notation de Kendall, la loi de Little,...etc.).

Dans **le deuxième chapitre**, nous étudions quelques modèles de files d'attente markovienne et semi markovienne (M/M/1, M/M/c, M/M/c/K, M/G/1) et diverses mesures de performance du système sont dérivées.

Dans **le troisième chapitre**, nous présentons une étude d'un modèle de file d'attente avec distribution générale du temps de service, vacances du serveur et clients impatients. Nous traitons le cas de files d'attente M/G/1 avec vacances multiples, et clients impatients.[16]

Chapitre 1

Introduction aux systèmes de files d'attente

DAns ce chapitre nous avons défini quelques concepts de base utilisée dans les chapitres suivants. Nous avons présenté et défini les processus stochastiques utilisées en théorie des files d'attente, nous avons donné quelques préliminaires sur les files d'attente.

1.1 Processus stochastique : quelques définitions

Définition 1.1.0.1. *(Processus stochastiques)[36]*
Un processus stochastique $\{X(t), t \in T\}$ est une collection de variables aléatoires définies sur un même espace de probabilité $(\Omega, F, \mathbb{P})$. Le paramètre t est généralement interprété comme le temps et appartient à un ensemble ordonné T .
Généralement $X(t)$ représente l'état du processus stochastique au temps t.
• Si T est dénombrable, i.e $T \subseteq \mathbb{N}$, alors nous disons que $\{X(t), t \in T\}$ est un processus à temps discret. On le dénote par $\{X_n, n \geq 0\}$.
• Si T est un intervalle de $[0; \infty)$, alors le processus stochastique est dit un processus à temps continu. On le dénote par $\{X(t), t \geq 0\}$.
L'ensemble des valeurs de $X(t)$ est appelé l'espace d'état, qui peut également être soit discret (fini ou infini dénombrable) ou continu (un sous-ensemble de $\mathbb{R}$ ou $\mathbb{R}^n$).

1.1.1 Processus de comptage

Définition 1.1.1.1. *(Processus de comptage)[33]*

Un processus stochastiques $\{N(t), t \geq 0\}$ *est dit processus de comptage ou processus de dénombrement si*

$N(t)$ *représente le nombre d'événements se produisant dans l'intervalle* $[0, t]$ *vérifiant :*

- $N(t) \geq 0$, $\forall t \geq 0$.
- $\forall t > s$, $N(t) \geq N(s)$.
- *Pour* $s < t$, $N(t) - N(s)$ *représente le nombre d'événements se produisant dans l'intervalle* $(s, t]$.

1.1.2 Processus de renouvellement

Un processus de renouvellement a pour fonction le dénombrement des occurrences d'un phénomène donné, lorsque les délais entre deux occurrences consécutives sont des variables aléatoires indépendantes et identiquement distribuées.

Définition 1.1.2.1. *[17]*

Un processus de comptage pour lequel les temps entre deux arrivés consécutives sont des variables aléatoires i.i.d, s'appelle processus de renouvellement. Les temps de renouvellement (ou les temps de la n-ième arrivée) sont :

$$A_n = \sum_{i=1}^{n} a_i, \quad n = 1, 2, \ldots$$

avec $a_i, i = 1, 2, \ldots$ *est le temps entre deux arrivées consécutives. Il est facile de voir que le nombre d'arrivées avant le temps t, i.e. le processus*

$$\{N(t), t \geq 0\} = \underset{k}{sup}\{k \in \mathbb{N} : A_k \leq t\}$$

est un processus de comptage.

1.1.3 Processus de Poisson

Le processus de poisson est le plus utilisé dans la théorie des files d'attente. Il modélisera généralement le processus d'arrivée des clients dans un système.

Définition 1.1.3.1. *(Processus de poisson)[11]*

On dit qu'un processus de comptage $\{N(t), t \geq 0\}$ est un processus de Poisson s'il satisfait aux trois conditions suivantes :

C1 ▶ *Le processus est homogène dans le temps : La probabilité d'avoir k évènements dans un intervalle de longueur donné t ne dépend que de t et non pas de la position de l'intervalle par rapport l'axe temporel :*

$$p_k(t) = \mathbb{P}(N(t) = k) = \mathbb{P}(N(t+s) - N(s) = k) \text{ pour tout } s > 0,\ t > 0.$$

C2 ▶ *Le processus $N(t)$ est à accroissement indépendants :*

$$\begin{aligned}\mathbb{P}(N(t+s) - N(s) = k, N(s) = j) &= \mathbb{P}(N(t+s) - N(s) = k)\mathbb{P}(N(s) = j)\\ &= p_k(t)p_j(s)\end{aligned}$$

pour tout $s > 0$, $t > 0$.

C3 ▶ *La probabilité $p_k(\Delta t)$*

$$p_k(\Delta t) = \begin{cases} 1 - \lambda(\Delta t) + o(\Delta t) & \text{si } k = 0 \\ \lambda(\Delta t) + o(\Delta t) & \text{si } k = 1 \\ o(\Delta t) & \text{si } k \geq 2 \end{cases}$$

λ est appelé densité ou intensité du processus. C'est le nombre d'évènements qui apparaissent par unité de temps.

1.1.4 Loi exponentielle

Définition 1.1.4.1. *(loi exponentielle)[2]*

Soit $\mu > 0$. On dit qu'une variable aléatoire réelle continue T suit la loi exponentielle de paramètre μ ($T \sim Exp(\mu)$) si

$$f(t) = \begin{cases} \mu e^{-(\mu t)} & t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Distribution exponentielle :

La fonction de répartition de cette loi est :

$$F(t) = \begin{cases} 1 - e^{-(\mu t)} & t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

L'espérance et la variance d'une loi exponentielle sont :

$$\mathbb{E}(T) = \frac{1}{\mu}.$$

$$Var(T) = \frac{1}{\mu^2}.$$

Théorème 1.1.4.0.1. *[25]*

Soit T une variable aléatoire continue à valeurs dans $\mathbb{R}_*^+$. On a l'équivalence :

(i) il existe $\mu > 0$ tel que $T \sim Exp(\mu)$,

(ii) pour tout $s, t \geq 0$, $\mathbb{P}(T > s + t | T > t) = \mathbb{P}(T > s)$.

C'est très important. (ii) est une propriété qualitative (absence de mémoire). Donc toute v.a. sans mémoire suit nécessairement une loi exponentielle.

Le temps d'attente T, avant la prise d'un premier poisson, d'un pêcheur totalement inexpérimenté est une variable aléatoire sans mémoire (le temps d'attente résiduel ne dépend pas du temps d'attente écoulé). Donc par nature, T suit une loi exponentielle.

La durée de vie D d'un objet qui ne s'use pas est une variable aléatoire sans mémoire. Donc par nature, D suit une loi exponentielle. Etc.

Preuve :

(i) implique (ii), car

$$\mathbb{P}(T > s + t | T > t) = \frac{\mathbb{P}(T > s + t)}{\mathbb{P}(T > t)} = \frac{e^{-\mu(s+t)}}{e^{-\mu t}} = e^{-\mu s} = \mathbb{P}(T > s).$$

Montrons maintenant que (ii) implique (i). Pour cela, introduisons $G(t) = \mathbb{P}(T > t)$.

C'est une fonction décroissante sur $\mathbb{R}^+$, on a $G(0) = 1$ et $G(\infty) = 0$ par hypothèse.

De plus (ii) donne que $G(t + s) = G(t)G(s)$ pour tous $s, t \geq 0$. On en déduit que pour tous $p, q \in \mathbb{N}$, on a

$$G(p) = G(p/q)^q, \quad G(p) = G(1)^p, \quad \text{d'où } G(p/q) = G(1)^{p/q}.$$

En utilisant que $\mathbb{Q}$ est dense dans $\mathbb{R}^+$, que G est décroissante et que $t \mapsto G(1)^t$ est continue, on en déduit que $G(t) = G(1)^t$ pour tout $t \in \mathbb{R}^+$.
De plus, $G(1) > 0$. Sinon, on aurait $G(t) = 0$ pour tout $t > 0$, et donc $T = 0$ p.s.
(or on a supposé T à valeurs dans $\mathbb{R}_*^+$).
Aussi, $G(1) < 1$. Sinon, on aurait $G(t) = 1$ pour tout $t > 0$, et donc $T = \infty$ p.s.
(or on a supposé T à valeurs dans $\mathbb{R}_*^+$).
On pose $\mu = -lnG(1) > 0$ et on a $G(t) = e^{-\mu t}$, i.e. $T \sim Exp(\mu)$. □

1.1.5 Processus de Markov

Définition 1.1.5.1. *Soit $(\Omega, F, \mathbb{P})$, un espace probabilisé, E un ensemble fini ou dénombrable et $T \subset R^+$ un intervalle .*
E : Espace des états. T : Espace de temps.
Soit $\{X(t), t \geq 0\}$ un processus défini sur Ω à valeurs dans E, on dit que $\{X(t), t \geq 0\}$ est un processus de Markov si :
$\forall s, t, u \in T$, avec $(u < s < t)$ et $\forall i, j, x \in E$,
on a :

$$\mathbb{P}(X_t = j/X_s = i, X_u = x) = \mathbb{P}(X_t = j/X_s = i) = P_{ij}(t, s) \qquad (*).$$

(: Propriété d'absence de mémoire ou propriété de Markov).*

Remarque 1.1.5.1. *Si dans la propriété (*), on a en plus $P_{ij}(s, t) = P_{ij}(t - s)$, on dira que le processus de Markov est homogène. Dans ce qui suit, on ne considérera que les processus de Markov homogènes.*
On note $P_{ij}(t) = \mathbb{P}(X_{t+s} = j/X_s = i)$, $s, t \in T; i, j \in E$ et $P(t) = (P_{ij}(t))_{i,j \in E \times E}$ est la matrice de transition du processus de Markov $\{X(t), t \in T\}$.

Proposition 1.1.5.1. *[28]*
Pour $s, t \in T$ et $i \in E$; $\mathbb{P}(X_s = i) > 0$, on a :
1) $\sum_{j \in E} P_{ij}(t) = 1$, $\forall i \in E$.
2) $P_{ik}(t + s) = \sum_{j \in E} P_{ij}(s) P_{jk}(t)$.
(Equations de Chapman-Kolmogorov associées au processus).

Remarque 1.1.5.2. *La matrice de transition $P(t)$ caractérise le processus de Markov $\{X(t), t \in T\}$, c'est-à-dire :*
A toute matrice stochastique $P(t)$, on peut associer un processus de Markov $\{X(t), t \in T\}$, de loi initiale $(\mathbb{P}(X_0 = i), i \in E) = \Pi_0$ qui va admettre $P(t)$ comme matrice de transition.
En effet, construire ce processus revient à évaluer juste ses lois fini-dimensionnelles en fonction de $P(t)$ et Π_0. C'est-à-dire à évaluer :

$$L = \mathbb{P}[X_{t_n} = a_n, X_{t_{n-1}} = a_{n-1}, ..., X_{t_1} = a_1, X_{t_0} = a_0], \quad \forall t_0, t_1, t_2, ..., t_{n-1}, t_n \in T, \forall a_i \in E.$$

$$L = \mathbb{P}[X_{t_n} = a_n/X_{t_{n-1}} = a_{n-1}, ...X_0 = a_0] \times \mathbb{P}[X_{t_{n-1}} = a_{n-1}/X_{t_{n-2}} = a_{n-2}, ...X_0 = a_0]$$

$$\times ... \times \mathbb{P}[X_{t_1} = a_1/X_0 = a_0] \times \mathbb{P}[X_0 = a_0].$$

$$= \mathbb{P}[X_{t_n} = a_n/X_{t_{n-1}} = a_{n-1}] \times \mathbb{P}[X_{t_{n-1}} = a_{n-1}/X_{t_{n-2}} = a_{n-2}] \times .. \times \mathbb{P}[X_{t_2} = a_2/X_{t_1} = a_1]$$

$$\times \mathbb{P}[X_{t_1} = a_1/X_{t_0} = a_0] \times \mathbb{P}[X_0 = a_0]$$

$$= P_{a_{n-1}a_n}(t_n - t_{n-1}) \times P_{a_{n-2}a_{n-1}}(t_{n-1} - t_{n-2}) \times ... \times P_{a_1 a_2}(t_2 - t_1) \times P_{a_0 a_1}(t_1 - t_0) \times \Pi_0(a_0)$$

avec $\Pi_0 = (\mathbb{P}[X_0 = i], i \in E)$.
On note $\Pi_t = (\mathbb{P}[X_t = i], i \in E)$ la loi t-instantanée du processus $\{X(t), t \in T\}$. On a :

$$\Pi_t = \Pi_0 \times P(t).$$

En effet, pour $i \in E$

$$\begin{aligned}
\Pi_t(i) &= P\left[X_t = i\right] = \textstyle\sum_{j\in E} P\left[X_t = i, X_0 = j\right] \\
&= \textstyle\sum_{j\in E} P\left[X_t = i/X_0 = j\right] \times P\left[X_0 = j\right] \\
&= \textstyle\sum_{j\in E} \Pi_0(j) \times P_{ji}(t). \\
\Rightarrow \quad & \Pi_t(i) = \textstyle\sum_{j\in E} \Pi_0(j) \times P_{ij}(t) \\
\Leftrightarrow \quad & \Pi(t) = \Pi(0) \times P(t).
\end{aligned}$$

Génerateur infinitisimal d'un processus de Markov *:[25]*
On suppose que $\forall (i,j) \in E \times E$, la fonction $P_{ij}(t)$ est continue en 0, c'est-à-dire :

$$\lim_{t \to 0^+} P_{ij}(t) = \begin{cases} 1 & si \quad i = j, \\ 0 & sinon \end{cases}$$

$$= P_{ij}(0).$$

Soit alors $i \geq j$,

- $i \neq j, q_{ij} = \lim\limits_{t \to 0^+} \frac{P_{ij}(t)}{t} = \lim\limits_{t \to 0^+} \left(\frac{P_{ij}(t) - P_{ij}(0)}{t-0}\right) = P'_{ij}(0)$, *(si* $P_{ij}(t)$ *est dérivable en 0).*
- $i = j, q_{ij} = \lim\limits_{t \to 0^+} \left(\frac{P_{ij}(t)-1}{t}\right) = \lim\limits_{t \to 0^+} \left(\frac{P_{ij}(t) - P_{ij}(0)}{t}\right) = P'_{ii}(0)$, *(si* $P_{ii}(t)$ *est dérivable en 0).*

On pose $q_i = -q_{ii} \geq 0$.

On appelle alors la matrice générateur infinitésimal du processus de Markov la matrice suivante :

$$Q = (q_{ij})_{(i,j) \in E \times E} \ .$$

On a

$$P_{ij}(t) = \begin{cases} q_{ij}t + o(t) & si\ i \neq j \\ 1 + q_{ij}t + o(t) & si\ i = j \end{cases} .$$

$$\Rightarrow \quad 1 - P_{ii}(t) = -q_{ii}t + o(t) \quad \Rightarrow \quad 1 - P_{ii}(t) = q_i t + o(t).$$

Remarque 1.1.5.3.

$$\sum_{j \in E} q_{ij} = 0, \quad \forall i \in E.$$

En effet

$\sum_{j \in E} q_{ij} = \sum_{j \in E} (P_{ij}(t))'|_{t=0} = (\sum_{j \in E} P_{ij}(t))'|_{t=0} = (1)' = 0.$

Ainsi

$q_{ii} + \sum_{(j \in E, i \neq j)} q_{ij} = 0 \Rightarrow q_i = -q_{ii} = \sum_{i \neq j} q_{ij}.$

- q_{ij} *est appelé le taux de transition de* i *vers* j.

- q_i *est appelé le taux de transition a partir de* i.

Equations de Chapman-Kolmogorov au processus de Markov :[39]

On a $P_{ik}(s+t) = \sum_{j \in E} P_{ij}(s) \times P_{jk}(t)$, donc

$$\text{i)} \quad \left.\frac{\partial (P_{ik}(s+t))}{\partial t}\right|_{t=0} = \left.\frac{\partial \left(\sum_{j \in E} P_{ij}(s) P_{jk}(t)\right)}{\partial t}\right|_{t=0} = \sum_{j \in E} P_{ij}(s) \times q_{jk}$$

$$\Rightarrow P'_{ik}(s) = \sum_{j \in E} P_{ij}(s) \times q_{jk}, \quad \forall i, k \in E \text{ et } \forall s, t \in T.$$

ii) $\left.\frac{\partial(P_{ik}(s,t))}{\partial s}\right|_{t=0} = \sum_{j \in E} q_{ij} P_{jk}(t) \iff P'_k(t) = \sum_{j \in E} q_{ij} P_{jk}(t), \quad \forall i, k \in E \text{ et } \forall s, t \in T.$

On a alors l'écriture matricielle suivante :

$$P'(t) = Q \times P(t).$$

Proposition 1.1.5.2. *[11]*

L'équation différentielle matricielle $\begin{cases} P'(t) = QP(t) \\ P(0) = I_{E \times E} \end{cases}$ *admet la solution qui s'écrit comme suit :*

$$P(t) = \sum_{n \geq 0} \left(\frac{Q^n t^n}{n!} \right) = I_{E \times E} + \sum_{n \geq 1} \left(\frac{(Qt)^n}{n!} \right) = e^{Qt} (\text{ notation }).$$

Proposition 1.1.5.3. *[30]*

Si E est fini et Q (qui est donc finie) est diagonalisable (c'est-à-dire $\exists B$ inversible et D diagonale telles que $Q = BDB^{-1}$) où

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

avec λ_i, $i = \overline{1, n}$ sont les valeurs propres de Q et B la matrice des vecteurs propres associés aux valeurs propres λ_i.

Alors

$$P(t) = B\Delta(t)B^{-1}, \text{ avec } \Delta(t) = e^{Dt}$$

$$\Delta(t) = e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & e^{\lambda_n t} \end{pmatrix}$$

Lois stationnaires d'un processus de Markov

On dit que $\Pi(t) = \Pi$ est une loi (solution) stationnaire du processus de Markov $\{X(t), t \in T\}$, si elle est solution du système d'équations :

$$(S) \begin{cases} \Pi Q = 0, \\ \sum_{j \in E} \pi_j = 1 \end{cases}$$

avec Q : générateur infinitésimal de $\{X(t), t \in T\}$.

1.1.6 Processus de naissance et de mort

Les processus de naissance et de mort sont des processus stochastiques à temps continu et à espace d'états discrets $n = 0, 1, 2 \ldots$ Ils sont sans mémoire, et à partir d'un état donné n, seules les transitions vers l'un des états voisins $(n+1)$ et $(n-1)$ avec $n \geq 1$ sont possibles. On parle alors de " naissances " et de " morts ". Ces processus sont utilisés pour modéliser les systèmes d'attente et l'évolution de populations.

Les files d'attente de type Markovien (M/M) sont des cas particuliers très importants de processus de naissance et de mort. Leur étude complète sera effectuée dans le chapitre 2.

Définition 1.1.6.1. *[39]*

Soit un processus stochastique $\{N(t), t \geq 0\}$ à états discrets $n \in \mathbb{N}$, et homogène dans le temps, c'est à dire :

$$\mathbb{P}(N(t+s) = j / N(s) = i) = p_{ij}(t), \text{ ne dépend pas de } s.$$

Le processus $\{N(t), t \geq 0\}$ est un processus de naissance et de mort s'il satisfait les conditions Suivantes :

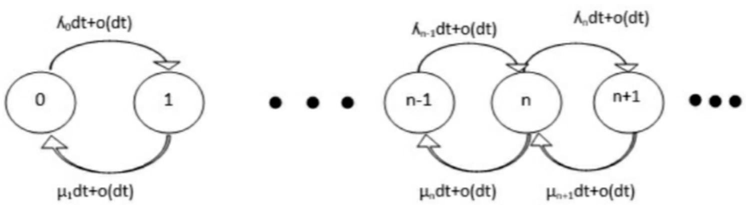
$$\begin{cases} p_{i,i+1}(\Delta t) = \lambda_i \Delta t + o(\Delta t), & i \geq 0, \\ p_{i,i-1}(\Delta t) = \mu_i \Delta t + o(\Delta t), & i \geq 1, \\ p_{i,i}(\Delta t) = 1 - (\lambda_i + \mu_i)\Delta t + o(\Delta t), & i \geq 0, \\ p_{i,j}(\Delta t) = o(\Delta t) & si\ |i-j| \geq 2, \end{cases}$$

$$p_{i,j}(0) = \delta_{i,j} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}.$$

Les coeffcients positifs $\lambda_i > 0$ *et* $\mu_i > 0$, $(\mu_0 = 0)$, *sont appelés taux de transition, plus particulièrement taux de naissance (ou de croissance) pour* λ_i *et taux de mort (ou de décroissance) pour* μ_i.

Régime transitoire :

Soient $p_n = \mathbb{P}(N(t) = n), n \geq 0$, les probabilités d'état



La matrice des transitions correspondante est :

$$Q = \begin{pmatrix} 1-\lambda_0\Delta t & \lambda_0\Delta t & 0 & 0 \\ \mu_1\Delta t & 1-(\lambda_1+\mu_1)\Delta t & \lambda_1\Delta t & 0 \\ 0 & \mu_2\Delta t & 1-(\lambda_2+\mu_2)\Delta t & \lambda_2\Delta t \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & . & . \end{pmatrix}.$$

En appliquant $\mathbb{P}(t+\Delta t) = \mathbb{P}(t) \times Q$, on trouve

$$\begin{aligned} &p_0(t+\Delta t) = (1-\lambda_0\Delta t)p_0(t) + \mu_1\Delta t p_1(t); \\ &p_n(t+\Delta t) = \lambda_{n-1}\Delta t p_{n-1}(t) + (1-(\lambda_n+\mu_n)\Delta t)p_n(t) + \mu_{n+1}\Delta t p_{n+1}(t), \quad n \geq 1. \end{aligned} \tag{1.1}$$

$$\frac{p_n(t+\Delta t) - p_n(t)}{\Delta t} = \lambda_{n-1}p_{n-1}(t) - (\lambda_n+\mu_n)p_n(t) + \mu_{n+1}p_{n+1}(t).$$

On faisant tendre Δt vers 0, on trouve :

$$p'_n(t) = \lambda_{n-1}p_{n-1}(t) - (\lambda_n+\mu_n)p_n(t) + \mu_{n+1}p_{n+1}(t), \quad n \geq 1. \tag{1.2}$$

Pour $n = 0$:

$$p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t). \tag{1.3}$$

Donc :

$$\begin{cases} p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t); \\ p_n'(t) = \lambda_{n-1} p_{n-1}(t) - (\lambda_n + \mu_n) p_n(t) + \mu_{n+1} p_{n+1}(t), \quad n \geq 1. \end{cases} \tag{1.4}$$

Les équations (1.4) sont connues sous le nom "équations différentielles de Kolmogorov " elles permettent de calculer les probabilités d'état $p_n(t)$ si l'on connait les conditions initiales du processus.

Régime stationnaire :

Soit $p_n = \lim\limits_{t \to +\infty} p_n(t)$, qui est la distribution stationnaire du processus étudié. Ces probabilités satisfont le système d'équations de balance suivant :

$$\begin{aligned} \lambda_0 p_0 &= \mu_1 p_1; \\ (\lambda_n + \mu_n) p_n &= \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1}, \quad n \geq 1 \end{aligned} \tag{1.5}$$

avec l'équation de normalisation $\sum\limits_{n=0}^{\infty} p_n = 1$.

De (1.5), on obtient :

$$p_1 = \frac{\lambda_0}{\mu_1} p_0.$$

Pour $n = 1$:

$$(\lambda_1 + \mu_1) p_1 = \lambda_0 p_0 + \mu_2 p_2 \Rightarrow p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0$$

$$\vdots$$

$$p_n = \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} p_0.$$

Pour déduire p_0, on utilise l'équation de normalisation. On obtient le résultat suivant :

$$p_0 = \left[1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} + \dots + \frac{\lambda_0 \lambda_1 \dots \lambda_{n-1}}{\mu_1 \mu_2 \dots \mu_n} + \dots \right]^{-1}.$$

Pour que le régime existe il faut que la somme ci-dessus converge.

1.2 Analyse mathématique d'un système de files d'attente

► L'étude mathématique d'un système de files d'attente se fait le plus souvent par l'introduction d'un processus stochastique, défini de façon appropriée.
En premier lieu, on s'intéresse principalement au nombre de clients $N(t)$, se trouvant dans le système à l'instant t $(t \geq 0)$.
En fonction des quantités qui définissent le système, on cherche à calculer :
⋆ Les probabilités d'état $p_n(t) = \mathbb{P}(N(t) = n)$, qui définissent le **régime transitoire** du processus stochastique $\{N(t), t \geq 0\}$. Il est évident que les fonctions $p_n(t)$ dépendent de l'état initial ou de la distribution initiale du processus.
⋆ Le **régime stationnaire** du processus stochastique est défini par :

$$p_n = \lim_{t \to +\infty} p_n(t) = \lim_{t \to +\infty} \mathbb{P}(N(t) = n), \quad n = 0, 1, 2, \ldots$$

Où, $\{p_n\}_{n \geq 0}$ est appelée distribution stationnaire du processus $\{N(t), t \geq 0\}$.

1.2.1 Modèle file d'attente simple

Le modèle général d'un système de files d'attente peut être résumé comme suit.
Les demandes de service (clients) arrivent à un certain endroit et réclament un certain service. Si un dispositif de service (serveur) est libre, le client qui arrive se dirige vers ce dernier où il est servi. Dans le cas contraire, on a deux possibilités : soit le client quitte le système, soit il prend une place dans une file d'attente. A un moment donné, le client est sélectionné pour le service selon une discipline donnée. Une représentation graphique est donnée par la figure (1.1)

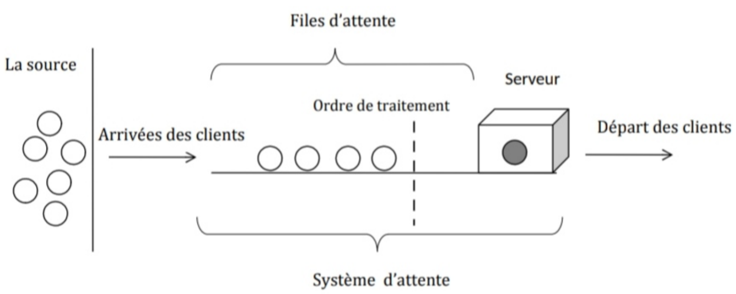


FIGURE 1.1 – Représentation d'une file d'attente.

1.2.2 Structure et discipline de la file

File d'attente : "Lieu" où les clients font la queue avant d'être servis.[4]

Système d'attente : la file d'attente + service en cours.

• **Le processus des arrivées des clients** :

Les arrivées des clients sont caractérisées par l'ensemble des instants d'arrivées de chaque client ou d'un groupe de clients dans le système. La collection de ces instants forment un processus des arrivées. Souvent, on suppose que les temps entre deux arrivées consécutives sont indépendants et identiquement distribués.

• **La source des clients** :

La population source, d'où proviennent les clients, peut être finie ou infinie, unique ou multiples.

• **Nombre de serveurs :**

Une station peut disposer de plusieurs serveurs en parallèle. Soit C (Voire FIG (1.2)) le nombre de serveurs. Dès qu'un client arrive à la station, soit il y a un serveur de libre et le client entre instantanément en service, soit tous les serveurs sont occupés et le client se place dans la file en attente de libération d'un des serveurs. La plupart du temps, les serveurs sont supposés identiques (ils possèdent donc la même distribution) et indépendants les uns des autres.

Une station particulière est la station IS (infinité servers) dans laquelle le nombre de serveurs est infini. Cette station ne comporte donc pas de file d'attente.

Dès qu'un client s'y présente, il trouve en effet instantanément un serveur disponible et entre donc directement en service. Elle permet de représenter des systèmes pour lesquels le nombre de serveurs est toujours supérieur au nombre de clients qui peuvent s'y trouver.

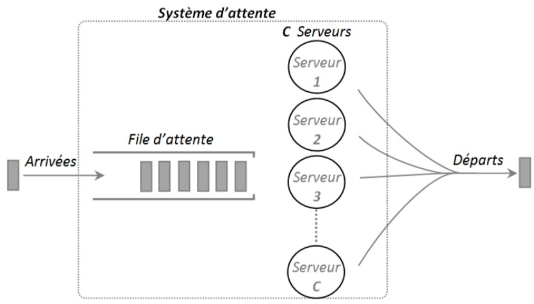


FIGURE 1.2 – File d'attente avec plusieurs serveurs.

- **Capacité de la file** :

Elle représente le nombre maximal de clients dans le système. Un client arrivant et trouvant ce nombre de clients présents dans le système sera perdu.

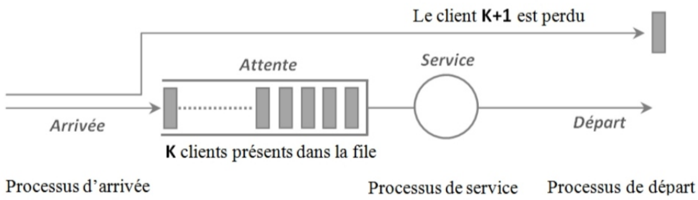


FIGURE 1.3 – Capacité d'une file d'attente.

1.2.3 Discipline de service

Elle spécifie la manière avec laquelle le serveur sélectionne le prochain client à servir. Cependant, plusieurs possibilités existent quant à l'ordre selon lequel les clients seront servis. Les principales disciplines de service sont :

FIFO (first in, first out) : Cette discipline est la plus usuelle. Les clients quittent le système dans l'ordre suivant lequel ils sont entrés.

LIFO (last in, first out) : Le dernier client dans la file est le premier à être servi.

RANDOM (aléatoire) : Le prochain client qui sera servi est choisi aléatoirement dans la file d'attente :

Prioritaire : Les clients sont servis suivant un attribut qui leur est associé.

PS (Processor Sharing) : les clients sont servis de manière égale. La capacité du système est partagée entre les clients.

1.2.4 Classification des systèmes d'attente

Pour la classification des systèmes de files d'attente, on a recours à une notation symbolique (NOTATION DE KENDALL) comprenant six symboles rangés dans l'ordre **A/B/c/m/n/Z** où **A** et **B** décrivent respectivement la distribution des temps entre deux arrivées successives et la distribution des temps de service, **c** est le nombre de serveurs (montés en parallèle), **m** est la capacité du système. Le dernier symbole peut être supprimé si **m** $=\infty$.
n : population des usagers.
Z : discipline de service c'est la façon dont les clients sont ordonnés pour être servi.
• Pour spécifier les distributions **A** et **B**, on introduit les symboles suivants :
M : inter-arrivées des clients sont indépendamment, identiquement distribuées selon une loi exponentielle. Il correspond à un processus de Poisson ponctuel (propriété sans mémoire).
E_k : Ce symbole désigne un processus où les intervalles de temps entre deux arrivées successives sont des variables aléatoires indépendantes et identiquement distribuées suivant une loi d'Erlang d'ordre k.
H_k : distribution hyperexponentielle de degré k .
D : les temps inter-arrivées des clients ou les temps de service sont constants et toujours les mêmes.
G : Inter-arrivées de clients ont une distribution générale et peuvent être dépendantes.

1.2.5 Loi de Little

La loi de Little est une relation très générale qui s'applique à une grande classe de systèmes.
Elle ne concerne que le régime permanent du système. Aucune hypothèse sur les variables aléatoires qui caractérisent le système (temps d'inter - arrivées, temps de service,. . .) n'est nécessaire. La seule condition d'application de la loi de Little est que le système soit stable.
Le débit du système est alors indifféremment, soit le débit d'entrée, soit le débit de sortie : $\lambda_s = \lambda_e = \lambda$. La loi de Little s'exprime telle que dans la propriété suivante :

Théorème 1.2.5.0.2. *[4](Formule de Little)*
Le nombre moyen de clients, le temps moyen passé dans le système et le débit moyen d'un système stable en régime permanent se relient de la façon suivante :

$$\overline{n} = \lambda_e \overline{W}_s \ .$$

Où λ_e est le taux d'entrée dans le système ($\lambda_e = \lambda$ pour une file(M/M/1)).

1.2.6 Mesures de performance

On note λ le taux d'arrivée des clients. Cela signifie que l'espérance mathématique de la durée séparant deux arrivées successives est $\mathbb{E}(A) = \frac{1}{\lambda}$.
On note μ le taux de service des clients. Cela signifie que l'espérance de la durée de service est $\mathbb{E}(S) = \frac{1}{\mu}$.
L'intensité du trafic s'exprime de la manière suivante :
$\rho = \frac{\lambda}{\mu} = \frac{\mathbb{E}(S)}{\mathbb{E}(A)} =$ temps moyen de service / temps moyen entre deux arrivées successives.
La distribution stationnaire du processus stochastique introduit permet d'obtenir les caractéristiques d'exploitation du système, telles que :
Le temps d'attente d'un client W, le temps de séjour d'un client dans le système W_s, le taux d'occupation des dispositifs de service, la durée de la période d'activité, le nombre de clients dans le système N, nombre de clients dans la files d'attente N_f.
Les mesures de performance sont :
- Le nombre moyen de clients dans le système $\overline{n}$;
- Le nombre moyen de clients dans la file d'attente $\overline{n}_f$;
- Le temps moyen d'attente d'un client $\overline{W}$;
- Le temps moyen de séjour d'un client dans le système $\overline{W}_s$;

Soient encore des relations (formules de Little) :
$\overline{n} = \lambda \overline{W}_s$; $\overline{n}_f = \lambda \overline{W}$; $\overline{W}_s = \overline{W} + 1/\mu$; $\overline{W} = \frac{\overline{n}_f}{\lambda}$; $\overline{n} = \overline{n}_f + \frac{\lambda}{\mu}$.

1.3 Types de files d'attente

1.3.1 Modèles markoviens :

Les modèles markoviens caractérisent les systèmes dans lesquels les deux quantités stochastiques principales, qui sont le temps inter-arrivées et la durée de service, sont des variables aléatoires indépendantes et exponentiellement distribuées. La propriété d'absence de mémoire de la loi exponentielle facilite l'étude de ces modèles. L'étude mathématique de tels systèmes se fait par l'introduction d'un processus stochastique approprié. Ce processus est souvent le processus de naissance et de mort $\{N(t), t \geq 0\}$ qui défini comme étant le nombre de clients dans le système à l'instant t. L'évolution temporelle du processus markovien est complètement définie grâce à la propriété d'absence de mémoire.

1.3.2 Modèles semi markoviens :

En l'absence de l'exponentialité c'est à dire lorsque l'on s'écarte de l'hypothèse d'exponentialité de l'une des deux quantités stochastiques :
Le temps des inter-arrivées et la durée de service, ou en prenant en compte certaines spécificités des problèmes par l'introduction de paramètres supplémentaires, on aboutit à un modèle semi markovien. La combinaison de tous ces facteurs rend l'étude mathématique du modèle très délicate, voire impossible. On essaye alors de se ramener à un processus de Markov judicieusement choisi à l'aide de l'une des méthodes d'analyse suivantes :

1. Méthode des étapes d'Erlang :

Son principe est d'approximer toute loi de probabilité ayant une transformée de Laplace rationnelle par une loi de Cox (mélange de lois exponentielles). Cette dernière possède la propriété d'absence de mémoire par étapes.

2. Méthode des variable supplémentaires :

Elle consiste à compléter l'information sur le processus $\{N(t), t \geq 0\}$ en lui donnant un caractère Markovien, ce qui nous ramène à l'étude du processus $\{N(t), \xi(t), t \geq 0\}$, où $\xi(t)$ sont dites alors variables supplémentaires.

3. Méthode de la chaîne de Markov induite :

Elle consiste à choisir une séquence d'instants $1, 2, 3, \ldots, n$ (déterministes ou aléatoires) telle que la chaîne induite $\{q_n, n \geq 0\}$, où q_n est le nombre de clients dans le système à l'instant $n \geq 0$, soit markovienne et homogène.

Chapitre 2

Systèmes de files d'attente classiques

Dans ce chapitre nous avons étudié la file d'attentes Markoviennes (M/M/1, M/M/c, M/M/c/K) et les mesures de performance de chaque file, ensuite, nous avons étudié le système de files d'attentes non Markoviennes m/G/1 par la méthode de la chaîne induite.

2.1 Systèmes de files d'attente régis par un modèle markovien de naissance et de mort

Les modèles Markoviens sont des systèmes où les temps entre deux arrivées successives et les durées de service sont des variables aléatoires indépendantes et exponentiellement distribuées.
On s'intéresse au nombre $N(t)$ de clients se trouvant dans le système à l'instant t. On introduit donc le processus stochastique

$$\{N(t), t \geq 0\}. \tag{2.1}$$

2.1.1 Système de files d'attente M/M/1

Description du modèle :
Le système d'attente M/M/1 est un système formé d'une file de capacité infinie, d'un unique serveur et la discipline d'attente est **FIFO**. Les clients arrivent vers le système selon un processus de Poisson de taux $\lambda > 0$ (nombre moyen de clients arrivant pendant une unité de temps), le taux de service est μ (nombre moyen de clients servis pendant une unité de

temps)

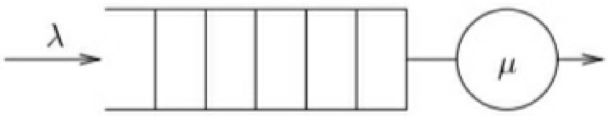


FIGURE 2.1 – File d'attente M/M/1.

La file peut être considérée comme un processus de naissance et de mort, pour lequel :

Les taux des arrivés λ_n et de service μ_n sont :

$$\lambda_n = \lambda, \quad \forall n \geq 0,$$

$$\mu_n = \begin{cases} \mu & n \neq 0 \\ 0 & n = 0 \end{cases}.$$

Le système est stable si :

$$\rho = \frac{\lambda}{\mu} < 1.$$

- Si $\rho > 1$ le nombre de client tend vers l'infini donc le système n'est pas stable.

Analyse du modèle :

L'état du système à la date t peut être décrit par le processus stochastique (2.1). Grâce aux propriétés fondamentales du processus de Poisson et de la loi exponentielle, on a pour un petit intervalle du temps Δt les probabilités suivantes :

$\mathbb{P}$(exactement une arrivée pendant Δt)= $\lambda\Delta t + o(\Delta t)$;

$\mathbb{P}$(aucune arrivée pendant Δt)=$1 - \lambda\Delta t + o(\Delta t)$;

$\mathbb{P}$(deux arrivées ou plus pendant Δt)= $o(\Delta t)$;

$\mathbb{P}$(exactement un départ pendant Δt / N(t) > 0)= $\mu\Delta t + o(\Delta t)$;

$\mathbb{P}$(aucun départ pendant Δt / N(t) > 0)= $1 - \mu\Delta t + o(\Delta t)$;

$\mathbb{P}$(deux départs ou plus pendant Δt)=$o(\Delta t)$.

• Ces probabilités ne dépendent ni de temps t ni de l'état $N(t)$ dans lequel le système se trouve.

Soient $p_{ij}(\Delta t) = \mathbb{P}(N(t + \Delta t) = j/N(t) = i)$;i,j =0,1,2,... Ces *probabilités de transition* ne dépendent pas de l'instant t. On suppose que les arrivées et les départs sont mutuellement indépendants.

Régime transitoire :

Soit $p_n(t) = \mathbb{P}(N(t) = n)$. Le graphe des transitions se présente de la manière suivante

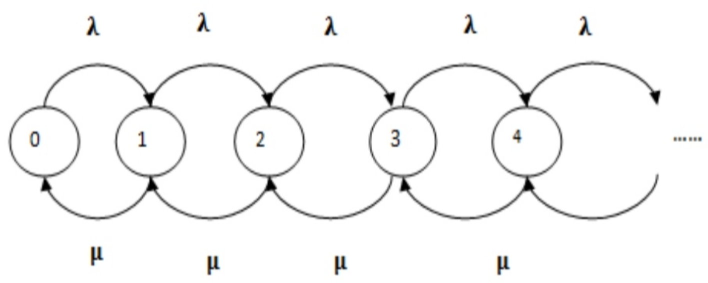


FIGURE 2.2 – Diagramme de transition d'état M/M/1.

A partir du graphe des transitions, on obtient :

$$\begin{cases} p_0(t+\Delta t) = \mu(\Delta t)p_1(t) + (1-\lambda\Delta t)p_0(t); \\ p_n(t+\Delta t) = \mu(\Delta t)p_{n+1}(t) + \lambda(\Delta t)p_{n-1}(t) + (1-(\lambda+\mu)\Delta t)p_n(t), \quad n \geq 1. \end{cases}$$

Puis, les équations de Kolmogorov :

$$\begin{cases} p_0'(t) = -\lambda p_0(t) + \mu p_1(t); \\ p_n'(t) = -(\lambda+\mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t), \quad n \geq 1. \end{cases} \tag{2.2}$$

Ces équations permettent, en principe, de calculer les probabilités d'état $p_n(t)$, si l'on connaît en plus les conditions initiales du processus, c'est-à-dire la distribution de $N(0)$.

Régime stationnaire :

Il est démontré que $\lim\limits_{t\to\infty} p_n(t) = p_n$, $n \geq 0$, existent et sont indépendantes de l'état initial du processus (2.1), et $\lim\limits_{t\to\infty} p_n'(t) = 0$, $n \geq 0$. De (2.2), on obtient le système d'équations de balance suivant :

$$\begin{cases} \mu p_1 = \lambda p_0, \\ \lambda p_{n-1} + \mu p_{n+1} = (\lambda+\mu)p_n, \quad n \geq 1, \end{cases} \tag{2.3}$$

avec $\sum_{n=0}^{\infty} p_n = 1$.

La résolution du système (2.3) (la résolution du modèle) s'effectue de la manière suivante :

$$p_1 = \frac{\lambda}{\mu} p_0.$$

Pour $n = 1$:

$$\lambda p_0 + \mu p_2 = (\lambda + \mu) p_1 \Longrightarrow p_2 = \left(\frac{\lambda}{\mu}\right)^2 p_0.$$

Pour $n > 1$:

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0.$$

Pour trouver la probabilité p_0, on utilise l'équation de normalisation.

En effet :

$\sum_{n=0}^{\infty} p_n = 1 \Longrightarrow p_0 + \frac{\lambda}{\mu} p_0 + \left(\frac{\lambda}{\mu}\right)^2 p_0 + \ldots = 1;$

$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \ldots},$$

où $1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \ldots$ est une progression géométrique de raison $\frac{\lambda}{\mu}$. Elle convergente si $\frac{\lambda}{\mu} < 1$, et est égale à $\frac{1}{1-\frac{\lambda}{\mu}}$. Alors

$$p_0 = 1 - \frac{\lambda}{\mu}.$$

D'où

$$p_n = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^n.$$

$\rho = \frac{\lambda}{\mu}$ est l'intensité du trafic. $\rho < 1$ est la condition d'existence du régime stationnaire.

Encore, $p_n = (1 - \rho)\rho^n$, $n \geq 0$, est la distribution stationnaire du nombre de clients dans le système M/M/1.

Caractéristiques du système M/M/1 :

▶ **Le nombre moyen de clients dans le système** $\overline{n}$:

Soit $N = \lim_{t\to\infty} N(t)$.

$$\begin{aligned}
\overline{n} = \mathbb{E}(N) &= \sum_{n=0}^{\infty} n p_n \\
&= (1-\rho)\sum_{n=0}^{\infty} n\rho^n \\
&= (1-\rho)\rho\sum_{n=1}^{\infty} \frac{d}{d\rho}(\rho)^n \\
&= (1-\rho)\rho\frac{d}{d\rho}\sum_{n=1}^{\infty}(\rho)^n \\
&= (1-\rho)\rho\frac{d}{d\rho}\left(\frac{1}{1-\rho}\right) \\
&= (1-\rho)\rho\left(\frac{1}{(1-\rho)^2}\right) \\
&= \frac{\rho}{1-\rho}.
\end{aligned}$$

Donc

$$\overline{n} = \frac{\rho}{1-\rho} = \frac{\lambda}{\mu-\lambda}.$$

► **Le nombre moyen de clients dans la file d'attente** $\overline{n}_f$:

Soit $N_f = \lim_{t\to\infty} N_f(t)$, où $N_f(t)$ est le nombre de clients dans la file d'attente à la date t. La variable N_f est définie de la manière suivante : $N_f = \begin{cases} 0 & N=0 \\ N-1 & N\geq 1 \end{cases}$.

$$\begin{aligned}
\overline{n}_f = \mathbb{E}(N_f) &= \sum_{n=1}^{\infty}(n-1)p_n \\
&= \frac{\rho^2}{1-\rho} \\
&= \frac{\lambda^2}{\mu(\mu-\lambda)}.
\end{aligned}$$

Ou bien

$$\overline{n}_f = \overline{n} - \rho.$$

► **Le temps moyen de séjour d'un client dans le système** :

Le temps moyen d'attente $\overline{W}$ et le temps moyen de séjour $\overline{W}_s$ peuvent être calculé soit à l'aide de formules de Little, soit à partir de la distribution stationnaire du système. Soit $\overline{W}_s$ la durée de séjour d'un client dans le système.

$$\overline{W}_s = \mathbb{E}(W_s) = \sum_{n=0}^{\infty} \mathbb{E}(W_s/A_n)\mathbb{P}(A_n).$$

Où A_n est l'événement tel qu'il y a $n \geq 0$ clients dans le système à l'instant d'arrivée d'un nouveau client. On a que $\mathbb{E}(W_s/A_n) = \frac{n+1}{\mu}$ et $\mathbb{P}(A_n) = \mathbb{P}(N = n) = (1-\rho)\rho^n$. Alors

$$\begin{aligned}
\overline{W}_s &= \sum_{n=0}^{\infty}(n+1)(1-\rho)\frac{\rho^n}{\mu} = \frac{1-\rho}{\mu}\sum_{n=0}^{\infty}(n+1)\rho^n \\
&= \frac{1-\rho}{\mu}\sum_{n=0}^{\infty} n\rho^n + \frac{1-\rho}{\mu}\sum_{n=0}^{\infty}\rho^n \\
&= \frac{1-\rho}{\mu}\frac{\rho}{(1-\rho)^2} + \frac{1-\rho}{\mu}\frac{1}{(1-\rho)} \\
&= \frac{\rho}{\mu(1-\rho)} + \frac{1}{\mu} \\
&= \frac{1}{\mu-\lambda}.
\end{aligned}$$

► **Le temps moyen d'attente d'un client :**$\overline{W}$

$$\overline{W} = \mathbb{E}(W) = \sum_{n=0}^{\infty} \mathbb{E}(W/A_n)\mathbb{P}(A_n)$$

et

$$\mathbb{E}(W/A_n) = \frac{n}{\mu};$$

$$\begin{aligned}
\overline{W} &= \sum_{n=0}^{\infty}\frac{n}{\mu}(1-\rho)\rho^n \\
&= \frac{1-\rho}{\mu}\sum_{n=0}^{\infty} n\rho^n \\
&= \frac{1-\rho}{\mu}\frac{\rho}{(1-\rho)^2} \\
&= \frac{\lambda}{\mu(\mu-\lambda)}.
\end{aligned}$$

2.1.2 Système de files d'attente M/M/c

Description du modèle :

les clients arrivent vers le système selon un processus de Poisson de taux $\lambda > 0$. Le service est assuré par $c \geq 1$ serveurs montés en parallèle. A l'arrivée d'un client, si l'un des serveurs est libre, le client commence immédiatement son service. Dans le cas contraire (tous les serveurs sont occupés par le service), le client prend place dans la file d'attente, commune pour tous les serveurs. La capacité d'attente est illimitée (le nombre de positions d'attente est infini). Lorsqu'un serveur se libère, le client en tête de la file d'attente occupe le serveur libéré. Par conséquent, la discipline d'attente est FIFO. Les temps de service sont exponentiellement distribués de moyenne finie $1/\mu$. Les durées entre deux arrivées consécutives et les durées de service sont mutuellement indépendantes.

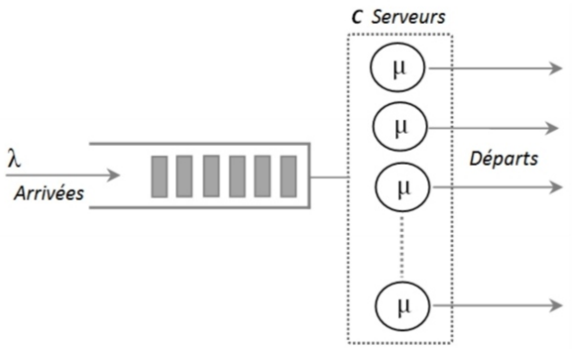


FIGURE 2.3 – File d'attante M/M/c.

Nous avons donc un modèle de file d'attente ou les arrivées et les départs sont modélisés par un processus de naissance et de mort où :

$$\lambda_n = \lambda, \quad \forall n \geq 0,$$

$$\mu_n = \begin{cases} 0 & \text{si } n = 0\,; \\ n\mu & \text{si } n < c\,; \\ c\mu & \text{si } n \geq c. \end{cases}$$

Stabilité du système :

La condition de stabilité de cette file est $\lambda < c\mu$, et exprime le fait que le nombre moyen de clients qui arrivent à la file par unité de temps doit être inferieur au nombre moyen de clients que les serveurs de la file sont capables de traiter par unité de temps.

$$\rho = \frac{\lambda}{c\mu} < 1.$$

Graphe de transition :

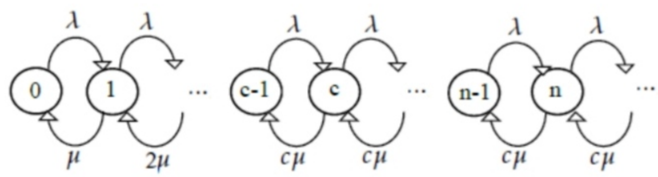


FIGURE 2.4 – Diagramme de transition d'état M/M/c.

Régime transitoire :

Le système d'équations de Kolmogorov pour les probabilités d'état $p_n(t) = \mathbb{P}(N(t) = n)$, $n \geq 0$, se présente de la manière suivante :

$$\begin{cases} p_0'(t) = -\lambda p_0(t) + \mu p_1(t); \\ p_n'(t) = \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + (n+1)\mu p_{n+1}(t), \quad 1 \leq n < c; \\ p_n'(t) = \lambda p_{n-1}(t) - (\lambda + c\mu)p_n(t) + c\mu p_{n+1}(t), \quad n \geq c. \end{cases}$$

Régime stationnaire :

Soit $\lim_{t\to\infty} p_n(t) = p_n$, $n \geq 0$. Cette distribution stationnaire satisfait les équations de balance

Si : $n < c$

$$\begin{cases} n = 0 : & \lambda p_0 = \mu p_1; \\ n = 1 : & \lambda p_1 + \mu p_1 = \lambda p_0 + 2\mu p_2; \\ & \vdots \\ n = (c-1) : & \lambda p_{c-1} + (c-1)\mu p_{c-1} = \lambda p_{c-2} + c\mu p_c. \end{cases}$$

Donc :

$$\begin{cases} p_1 = (\frac{\lambda}{\mu})p_0 \\ p_2 = (\frac{\lambda}{2\mu})p_1 = \frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2 p_0 \\ p_3 = (\frac{\lambda}{3\mu})p_2 = \frac{1}{3\times 2}\left(\frac{\lambda}{\mu}\right)^3 p_0 \\ \quad\vdots \\ p_c = \frac{1}{c!}(\frac{\lambda}{\mu})^c p_0. \end{cases}$$

Si : $n \geq c$

$$\begin{cases} \lambda p_c + c\mu p_c = \lambda p_{c-1} + c\mu p_{c+1} \\ \lambda p_{c+1} + c\mu p_{c+1} = \lambda p_c + c\mu p_{c+2} \\ \vdots \\ \vdots \end{cases}$$

$$\begin{cases} p_c = \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c p_0 \\ p_{c+1} = \frac{1}{c!c}\left(\frac{\lambda}{\mu}\right)^{c+1} p_0 \\ p_{c+2} = \frac{1}{c!c^2}\left(\frac{\lambda}{\mu}\right)^{c+2} p_0 \\ \quad\vdots \\ p_n = \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n p_0. \end{cases}$$

D'où

$$p_n = \begin{cases} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_0, & \text{si } 1 \leq n \leq c; \\ \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n p_0, & \text{si } n \geq c. \end{cases}$$

On remarque que pour $n = c$, les deux formules donnent la même valeur.

Pour calculer la probabilité pour que le système est vide p_0, on applique l'équation de normalisation $\sum_{n=0}^{\infty} p_n = 1$.

En effet :

$$p_0 = \left(\sum_{n=0}^{c-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \underbrace{\sum_{k=0}^{\infty} \frac{1}{c!c^k}\left(\frac{\lambda}{\mu}\right)^{c+k}}_{\star} \right)^{-1}.$$

La deuxième somme ($\star$) peut être réécrite de la manière suivante

$$\frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c\left(\underbrace{1+\frac{\lambda}{\mu c}+\left(\frac{\lambda}{\mu c}\right)^2+\dots}_{\star\star}\right).$$

La somme ($\star\star$) possède une limite égale à $\frac{1}{1-\frac{\lambda}{\mu c}}$ si $\frac{\lambda}{\mu c}<1$. Par conséquent, le système considéré est en régime stationnaire si $\rho=\frac{\lambda}{\mu c}<1$,

ρ est l'intensité globale du trafic. On obtient ainsi

$$p_0=\left(\sum_{n=0}^{c-1}\frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n+\frac{(\lambda/\mu)^c}{c!(1-\frac{\lambda}{\mu c})}\right)^{-1}.$$

Encors,

$$p_0=\left(\sum_{n=0}^{c-1}\frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n+\frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c\sum_{n=c}^{\infty}\rho^{n-c}\right)^{-1},$$

et

$$p_n=\frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c\left(\frac{\lambda}{\mu c}\right)^{n-c}p_0=\rho^{n-c}p_c.$$

Remarque 2.1.2.1. *la distribution stationnaire peut s'obtenir rapidement en appliquant la relation établie pour les processus de naissance et de mort.*

En effet

$p_n=\frac{\lambda_0\lambda_1\dots\lambda_{n-1}}{\mu_1\mu_2\dots\mu_n}p_0$, *pour $n\leq c$ il vient*

$$p_n=\frac{\lambda\times\lambda\times\dots\lambda\times\lambda\times\lambda\times\dots\times\lambda}{\mu\times 2\mu\times\dots\times(c-1)\mu\times c\mu\times c\mu\times c\mu\dots\times c\mu}p_0=\frac{\lambda^c}{c!\mu^c}\left(\frac{\lambda}{c\mu}\right)^{n-c}p_0=\frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n p_0.$$

Caractéristiques du système M/M/c :

▶ **Le nombre moyen de clients dans le système $\overline{n}$**

$$\overline{n}=\sum_{n=0}^{\infty}np_n=\sum_{n=1}^{c-1}\frac{n(\lambda/\mu)^n}{n!}p_0+\sum_{n=c}^{\infty}\frac{n(\lambda/\mu)^n}{c!c^{n-c}}p_0$$

$$\overline{n}=\frac{\lambda}{\mu}+\frac{(\lambda/\mu)^{c+1}}{c!c\left(1-\frac{\lambda}{\mu c}\right)^2}p_0.$$

▶ **Le nombre moyen de clients dans la file d'attente** $\overline{n}_f$

$$\overline{n}_f = \sum_{k=0}^{\infty} k p_{c+k} = \frac{(\lambda/\mu)^c}{c!} \sum_{k=0}^{\infty} k \left(\frac{\lambda}{\mu c}\right)^k p_0$$

$$\overline{n}_f = \frac{(\lambda/\mu)^{c+1}}{c!c\left(1 - \frac{\lambda}{\mu c}\right)^2} p_0.$$

▶ **Le temps moyen de séjour d'un client dans le système** $\overline{W}_s$

$$\overline{W}_s = \frac{\overline{n}}{\lambda} = \frac{1}{\mu} + \frac{(\lambda/\mu)^c}{c!c\mu\left(1 - \frac{\lambda}{\mu c}\right)^2} p_0.$$

▶ **Le temps moyen d'attente d'un client** $\overline{W}$

$$\overline{W} = \frac{\overline{n}_f}{\lambda} = \frac{c\mu(\lambda/\mu)^c}{c!(c\mu - \lambda)^2} p_0.$$

2.1.3 Système de files d'attente M/M/c/K

Description du modèle :

A présent, supposons que dans le système M/M/c, le nombre de positions d'attente est limité (égal à K). A l'arrivée d'un client, si tous les serveurs et toutes les positions d'attente sont occupées, le client quitte le système définitivement sans recevoir le service.

$$\lambda_n = \lambda, \quad 0 \leq n \leq K,$$

$$\mu_n = \mu \times min\{n, c\}, \quad 1 \leq n \leq K.$$

Graphe de transition :

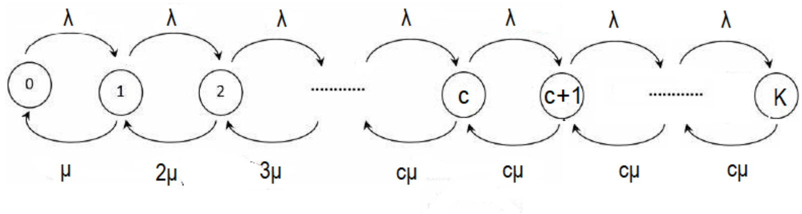


FIGURE 2.5 – Diagramme de transition d'état M/M/c/K.

Régime transitoire :

Soient $p_n(t) = \mathbb{P}(N(t) = n)$, $0 \leq n \leq K$. Le système d'équations de Kolmogorov pour les probabilités s'états s'obtient à partir du graphe des transitions ci-dessus. En effet,

$$\begin{cases} p_0'(t) = -\lambda p_0(t) + \mu p_1(t); \\ p_n'(t) = \lambda p_{n-1}(t) - (\lambda + n\mu)p_n(t) + (n+1)\mu p_{n+1}(t), \quad 1 \leq n < c; \\ p_n'(t) = \lambda p_{n-1}(t) - (\lambda + c\mu)p_n(t) + c\mu p_{n+1}(t), \quad c \leq n < K; \\ p_K'(t) = \lambda p_{K-1}(t) - c\mu p_K(t). \end{cases}$$

Régime stationnaire :

Soient $p_n = \lim_{t\to\infty} p_n(t)$, $0 \leq n \leq K$. La distribution stationnaire p_n satisfait le système d'équations de balance suivant :

$$\begin{cases} 0 = -\lambda p_0 + \mu p_1; \\ 0 = \lambda p_{n-1} - (\lambda + n\mu)p_n + (n+1)\mu p_{n+1}, \quad 1 \leq n < c; \\ 0 = \lambda p_{n-1} - (\lambda + c\mu)p_n + c\mu p_{n+1}, \quad c \leq n < K; \\ 0 = \lambda p_{K-1} - c\mu p_K. \end{cases}$$

La résolution de ce système, nous donne :

$$p_n = \begin{cases} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_0, & 1 \leq n < c; \\ \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n p_0 = \frac{1}{c!}\left(\frac{\lambda}{\mu c}\right)^{n-c}\left(\frac{\lambda}{\mu}\right)^c p_0, & c \leq n \leq K. \end{cases}$$

La mesure importante de ce système est la probabilité de perte, qui est la probabilité pour que le système se trouve dans l'état K :

$$p_K = \frac{1}{c!}\left(\frac{\lambda}{\mu c}\right)^{K-c}\left(\frac{\lambda}{\mu}\right)^c p_0.$$

La probabilité p_0 s'obtient à partir de l'équation de normalisation $\sum_{n=0}^{K} p_n = 1$:

$$p_0 = \left(\sum_{n=1}^{c-1} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n + \sum_{n=c}^{K} \frac{1}{c!} \left(\frac{\lambda}{\mu}\right)^c \left(\frac{\lambda}{\mu c}\right)^{n-c} \right)^{-1}.$$

Dans le cas particulier où $K = c$ (système à demandes refusées), la distribution stationnaire du processus $\{N(t), t \geq 0\}$ correspondant (formule d'Erlang) est

$p_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_0$, $0 \leq n \leq c$, où $p_0 = \left(\sum_{n=0}^{c} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n \right)^{-1}$.

On a également

$$\mathbb{P}(perte) = \frac{1}{c!} \left(\frac{\lambda}{\mu}\right)^c p_0.$$

Caractéristiques du système :

► **Le nombre moyen de clients dans la file d'attente $\overline{n}_f$**

On démontre que

$$\overline{n}_f = \sum_{n=1}^{K-c} n p_{c+n} = \frac{\lambda}{\mu} \frac{(\lambda/\mu)^c}{c!c} p_0 \left(1 + 2\left(\frac{\lambda}{\mu c}\right) + 3\left(\frac{\lambda}{\mu c}\right)^2 + \ldots + (K-c)\left(\frac{\lambda}{\mu c}\right)^{K-c-1} \right)$$

$$= \frac{(\lambda/\mu)^{c+1}}{(c-1)!} \frac{1 - \left(\frac{\lambda}{\mu c}\right)^{K-c} \left(1 + (K-c)(1 - \frac{\lambda}{\mu c})\right)}{\left(c - \frac{\lambda}{\mu}\right)^2} p_0.$$

L'application des relations de Little fournit d'autres mesures de performance
$\overline{W} = \frac{\overline{n}_f}{\lambda}$; $\overline{W}_s = \frac{\overline{n}}{\lambda}$; $\overline{n} = \overline{n}_f + \frac{\lambda}{\mu}$.

2.2 Système de files d'attente M/G/1

Pour décrire l'état d'un système de type M/G/1 à la date t, il faut connaître non seulement le nombre de clients qui se trouvent dans le système à la date t, mais également le temps de service, déjà écoulé $R(t)$ du client qui est en train d'être servi. On peut alors montrer que le processus bidimensionnel $\{N(t), R(t), t \geq 0\}$ est à nouveau du type markovien, cependant, le calcul de son régime transitoire ferait intervenir des équations aux dérivées partielles. Par conséquent, on choisit une autre méthode qui ramène l'étude du processus non markovien

$\{N(t), t \geq 0\}$ à celle d'une chaîne de Markov à temps discret associée au processus considéré dont elle permet de calculer le régime stationnaire.

Description du modèle :

Le système d'attente M/G/1 est un système d'une file d'attente capacité illimitée de discipline FIFO et d'un seule serveur. Les clients arrivent dans le système selon un processus de Poisson ($\lambda > 0$). Les durées de service sont des variables aléatoires positives mutuellement indépendantes noté Se, et distribuées selon une loi générale de fonction de répartition $B(x)$, de moyenne finie $\mathbb{E}(Se) = \frac{1}{\mu}$ et de $\mathbb{E}(Se^2)$. Les durées entre deux arrivées consécutives et les durées de service sont également mutuellement indépendantes.

Analyse du modèle :

Soit $\{N(t), t \geq 0\}$. Montrons que $\{N(t)_{(t \geq 0)}\}$ ne définit pas une chaîne de Markov. Soient t_d et t_f les dates de début et de fin d'un service, t_a l'instant d'arrivée d'un nouveau client. Si $t_d < t_a < t_f$, la probabilité qu'un départ s'effectue dans $]t_a, t_a + \Delta t]$, ne dépend pas seulement de Δt, mais de la date t_d à la quelle le service en cours a commencé. Comme le temps résiduel du service $(t_f - t_a)$ dépend du passé, alors la chaîne $\{N(t)_{(t \geq 0)}\}$ n'est pas markovienne. Par conséquent, on utilise la méthode de **la chaîne de Markov induite**. A cet effet, on considère $N(t)$ aux instants $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ où les clients terminent leur service et quittent le système. On définit ainsi un processus stochastique à temps discret

$$\{N_n = N(\xi_n), n \geq 1\}. \tag{2.4}$$

Pour vérifier que cette suite de variables aléatoires est une chaîne de Markov à temps discret, on considère le nombre A_n de clients qui entrent dans le système pendant que le *n-ème* client est servi. Les variables A_n sont indépendantes entre elles, leur distribution commune est

$\mathbb{P}(A_n = k) = a_k = \displaystyle\int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t)$, où $a_k > 0$ et $k > 0$. Alors

$$N_{n+1} = \begin{cases} N_n - 1 + A_{n+1} & N_n \geq 1 \\ A_{n+1} & N_n = 0 \end{cases} , n \geq 1.$$

L'équation fondamentale de la chaîne vaut donc

$$N_{n+1} = N_n - \delta_n + A_{n+1}, \tag{2.5}$$

où $\delta_n = \begin{cases} 1 & N_n > 0 \\ 0 & N_n = 0 \end{cases}$.

N_{n+1} ne dépend que de N_n et de A_{n+1} et non pas des valeurs prises par $N_{n-1}, N_{n-2}, \ldots$ La suite $\{N_n, n \geq 1\}$ est une chaîne de Markov induite du processus $\{N(t), t \geq 0\}$. Ses probabilités de transition $p_{ij} = \mathbb{P}(N_{n+1} = j / N_n = i)$ se calcule par

$$\begin{cases} p_{0j} = a_j & \text{pour } j \geq 0 \\ p_{ij} = a_{j-i+1} & \text{pour } 1 \leq i \leq j+1 \cdot \\ p_{ij} = 0 & \text{ailleurs} \end{cases}$$

La matrice des transitions est

$$M = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \ldots \\ a_0 & a_1 & a_2 & a_3 & \ldots \\ 0 & a_0 & a_1 & a_2 & \ldots \\ 0 & 0 & a_0 & a_1 & \ldots \\ . & . & . & . & \ldots \\ . & . & . & . & \ldots \end{pmatrix}$$

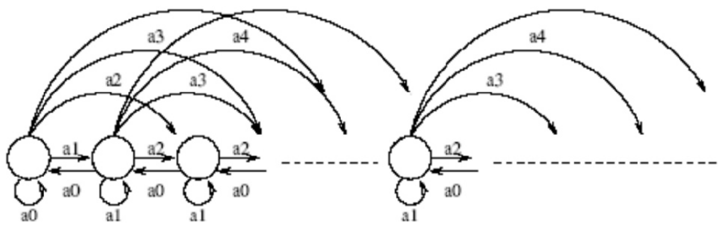


FIGURE 2.6 – Diagramme de transition d'état M/G/1.

Vu qu'on peut passer de chaque état vers n'importe quel autre état, il s'agit d'une chaîne de Markov irréductible. De plus, la matrice n'est pas décomposable (est apériodique). La chaîne est donc ergodique. La distribution stationnaire de la chaîne existe si $\rho = \frac{\lambda}{\mu} < 1$.

Pour les variables aléatoires A_n, nous disposons de quelques résultats importants :

$\mathbb{E}(A_n) = \lambda \mathbb{E}(Se) = \frac{\lambda}{\mu} = \rho$.

La fonction génératrice

$$\begin{aligned} A(z) = \sum_{k=0}^{\infty} a_k z^k &= \sum_{k=0}^{\infty} z^k \int_0^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} dB(t) \\ &= \int_0^{\infty} e^{-\lambda t} \left(\sum_{k=0}^{\infty} \frac{(\lambda t z)^k}{k!} \right) dB(t) \\ &= \int_0^{\infty} e^{-\lambda t} e^{\lambda t z} dB(t) \\ &= \int_0^{\infty} e^{-(\lambda - \lambda z)t} dB(t). \end{aligned}$$

Soit $\tilde{B}(s) = \displaystyle\int_0^{\infty} e^{-st} dB(t)$. Alors $A(z) = \tilde{B}(\lambda - \lambda z)$. Encore, la série $A(z)$ converge pour $|z| \leq 1$:

1) $|z| \leq 1$; $0 < a_k < 1 \; \forall k$, on a $|a_k z^k| < |z^k|$;

2) $|z| = 1$; $A(1) = 1$.

Remarque 2.2.0.1. *1. Théorème des probabilités totales :*

Cas discret : $\mathbb{P}(A) = \displaystyle\sum_k \mathbb{P}(A/Y = y_k)\mathbb{P}(Y = y_k)$.

Cas continu : $\mathbb{P}(A) = \int \mathbb{P}(A/Y = y) g(y) dy$.

2. Probabilité que le nombre d'événements N qui ont lieu pendant un intervalle $U = u$ dont la densité de probabilité $f(u)$ est connue, est égal à n :

$\mathbb{P}(N = n/U = u) = e^{-\lambda u} \frac{(\lambda u)^n}{n!}$.

D'où

$\mathbb{P}(N = n) = \displaystyle\int_0^{\infty} \mathbb{P}(N = n/U = u) f(u) du = \frac{1}{n!} \int_0^{\infty} e^{-\lambda u} (\lambda u)^n f(u) du$.

$\mathbb{E}(N) = \lambda \mathbb{E}(U)$; $Var(N) = \lambda^2 Var(U) + \lambda \mathbb{E}(U)$.

Supposons que $\rho < 1$. Le système se trouve dans un régime stationnaire. Soit $\Pi = [\pi_0, \pi_1, \ldots]$ la distribution stationnaire de la chaîne de Markov induite ($\pi_j = \lim_{n \to \infty} \mathbb{P}(N(\xi_n) = j)$). Par conséquent, $\Pi = \Pi \times M$, ou $\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, \quad j \geq 0.$

$$\pi_j = a_j \pi_0 + \sum_{i=1}^{j+1} a_{j-i+1}\pi_i = a_j\pi_0 + \sum_{i=0}^{j+1} a_{j-i+1}\pi_i - a_{j+1}\pi_0, \quad j \geq 0.$$

A présent, on applique la méthode des fonctions génératrices. En effet,

$$\sum_{j=0}^{\infty} \pi_j z^j = \pi_0 \sum_{j=0}^{\infty} a_j z^j + \frac{1}{z}\sum_{j=0}^{\infty} c_{j+1} z^{j+1} - \frac{\pi_0}{z}\sum_{j=0}^{\infty} a_{j+1} z^{j+1},$$

où $c_{j+1} = \sum_{i=0}^{j+1} a_{j-i+1}\pi_i$.

On introduit les fonctions génératrices suivantes :

$\Pi(z) = \sum_{i=0}^{\infty} \pi_i z^i\,; \quad A(z) = \sum_{i=0}^{\infty} a_i z^i\,; \quad C(z) = \sum_{j=0}^{\infty} c_j z^j = \Pi(z)A(z).$

Finalement, on obtient

$$\Pi(z) = \pi_0 A(z) + \frac{1}{z}(C(z) - c_0) - \frac{\pi_0}{z}(A(z) - a_0),$$

ou bien

$$\Pi(z) = \tfrac{\pi_0 A(z)(z-1)}{z-A(z)} \text{ pour } |z| < 1 \text{ et } |z| \neq 0.$$

On a que $\Pi(1) = 1$, Cependant, $\Pi(1) = \lim_{z\to 1} \Pi(z) = 0/0$. En appliquant la règle de l'Hôpital, on obtient $\frac{\pi_0}{1-A'(1)} = 1$. Alors $\pi_0 = 1 - A'(1) = 1 - \lambda\mathbb{E}(Se) = 1 - \rho$.

Le résultat final est la première équation de Pollaczek-Khintchine pour le nombre de clients dans le système :

$$\Pi(z) = \frac{(1-\rho)A(z)(z-1)}{z - A(z)} = \frac{(1-\rho)\tilde{B}(\lambda - \lambda z)(z-1)}{z - \tilde{B}(\lambda - \lambda z)}. \tag{2.6}$$

La condition d'existence d'un régime stationnaire est $\rho = \frac{\lambda}{\mu} < 1$.

Remarque 2.2.0.2. *La probabilité π_0 peut être trouvée d'une autre manière. De l'équation fondamentale de la chaîne de Markov induite (2.5), on a $\mathbb{E}(N_{n+1}) = \mathbb{E}(N_n) - \mathbb{E}(\delta_n) +$*

$\mathbb{E}(A_{n+1})$. *Vu que* $\mathbb{E}(N_{n+1}) = \mathbb{E}(N_n), \mathbb{E}(A_{n+1}) = \mathbb{E}(\delta_n) = \mathbb{P}(\delta_n > 0) = \mathbb{P}(N_n > 0) = 1 - \mathbb{P}(N_n = 0)$. *D'où* $\pi_0 = 1 - \rho$.

Considérons les probabilités suivantes :

$p_j = \lim_{t\to\infty} \mathbb{P}(N(t) = j),\ j \geq 0\,;$

$\pi_j = \lim_{n\to\infty} \mathbb{P}(N(\xi_n) = j),\ j \geq 0\,;$

$r_j = \lim_{n\to\infty} \mathbb{P}(N(\varsigma_n) = j),\ j \geq 0\,;$ ς_n est l'instant d'arrivée de n-*ème* client.

Vu que le processus des arrivées est poissonien, et le nombre $N(t)$ subit des changements discontinus de taille 1 (± 1), on obtient $p_j = r_j = \pi_j$. Comme suite logique, la distribution stationnaire du processus à temps continu $\{N(t), t \geq 0\}$ est identique à celle de la chaîne de Markov induite. Par conséquent $Q(z) = \sum_{j=0}^{\infty} p_j z^j = \Pi(z)$.

Caractéristiques du système M/G/1 :

► Le nombre moyen de clients dans le système $\overline{n}$

Formule de Pollaczek-Khintchine pour le nombre moyen de clients dans le système :

Considérons l'équation fondamentale (2.5). Vu que $\delta_n^2 = \delta_n$ et $\delta_n N_n = N_n$, on trouve

$$N_{n+1}^2 = N_n^2 + \delta_n + A_{n+1}^2 - 2N_n - 2\delta_n A_{n+1} + 2N_n A_{n+1}.$$

On a que : A_{n+1} est indépendante de N_n et de δ_n.

$\mathbb{E}(N_{n+1}^2) = \mathbb{E}(N_n^2)\,; \quad \mathbb{E}(A_n) = \rho = \frac{\lambda}{\mu}$.

Alors,

$$\mathbb{E}(N_{n+1}^2) = \mathbb{E}(N_n^2) + \mathbb{E}(\delta_n) + \mathbb{E}(A_{n+1}^2) - 2\mathbb{E}(N_n) + 2\mathbb{E}(A_{n+1})\mathbb{E}(N_n - \delta_n),$$

ou bien :

$2\mathbb{E}(N_n) = \rho + \mathbb{E}(A_{n+1}^2) + 2\rho(\mathbb{E}(N_n) - \rho)$.

D'où

$$\mathbb{E}(N_n) = \frac{\rho + \mathbb{E}(A_{n+1}^2) + 2\rho^2}{2(1-\rho)}. \tag{2.7}$$

Pour trouver $\mathbb{E}(A_{n+1}^2)$, considérons le régime stationnaire.

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}(A_{n+1}^2) &= \mathbb{E}(A^2) = \int_0^{\infty} \mathbb{E}(A^2/T = t)dB(t) \\
&= \lambda \int_0^{\infty} t dB(t) + \lambda^2 \int_0^{\infty} t^2 dB(t) \\
&= \tfrac{\lambda}{\mu} + \lambda^2 (Var(Se) + (1/\mu)^2).
\end{aligned}$$

Enfin, la formule (2.7) devient

$$\overline{n} = \lim_{n\to\infty} \mathbb{E}(N_n) = \mathbb{E}(N) = \rho + \frac{\rho^2 + \lambda^2 Var(Se)}{2(1-\rho)}.$$

Le nombre moyen de clients dans le système peut être également trouvé à partir de la fonction génératrice $\Pi(z)$: $\mathbb{E}(N) = \overline{n} = \lim_{z\to 1} \Pi'(z)$. Ici, le calcul de la limite donne une indétermination. Par conséquent, il est nécessaire d'appliquer la règle de l'Hôpital deux fois.

▶ **Le nombre moyen de client dans la file d'attente** :

$$\overline{n}_f = \overline{n} - \rho = \frac{\rho^2 + \lambda^2 Var(Se)}{2(1-\rho)}.$$

▶ **Temps moyen de séjour d'un client dans le système** :

$$\overline{W}_s = \frac{1}{\mu} + \frac{\lambda(Var(Se) + \frac{1}{\mu^2})}{2(1-\rho)}.$$

▶ **Temps moyen d'attente d'un client** :

$$\overline{W} = \frac{\lambda(Var(Se) + \frac{1}{\mu^2})}{2(1-\rho)}.$$

Période d'activité

Soit J la durée de la période d'activité du système M/G/1 (l'intervalle de temps pendant lequel le dispositif de service est continuellement occupée). Admettons que pendant une longue durée t, le système d'attente passe par n cycles d'exploitation complets dont chacun est composé d'une période d'activité J et d'une période d'inactivité V. Pour les grandes valeurs de t ($t \to \infty$), on a $t \approx n(\mathbb{E}(J) + \mathbb{E}(V))$. D'autre part, la probabilité que le système soit vide est
$\pi_0 = p_0 = \frac{\mathbb{E}(V)}{\mathbb{E}(J)+\mathbb{E}(V)}$. Mais $p_0 = 1 - \rho$ et $\mathbb{E}(V) = \frac{1}{\lambda}$. Il en résulte que $\mathbb{E}(J) = \frac{1}{\mu-\lambda}$, si $\lambda < \mu$. Ce résultat est valable et pour le système de files d'attente M/M/1.

Chapitre 3

Système de files d'attente avec distribution générale du temps de service, vacances du serveur et clients impatients

Dans ce chapitre nous avons étudié le système M/G/1 avec vacances du serveur et clients impatients, et le cas des temps d'impatience à distribution exponentielle.

Vacances : dans un contexte de file d'attente représente une période pendant laquelle le serveur est absent ou indisponible pour offrir un service. Les situations qui conduisent à des vacances sont diverses, à savoir les pannes du système, la maintenance du système ou uniquement pour une pause. Au cours des dernières décennies, les modèles de files d'attente de vacances ont été largement étudiés, soit pour résoudre des problèmes particuliers dans de nombreuses situations pratiques, telles que les centres d'appels, les ordinateurs, les industries en croissance, les services Web, etc.
L'impatience : est une caractéristique très importante de la théorie des files d'attente. Les modèles de files d'attente de vacances avec l'impatience des clients sont considérés comme des outils très appropriés pour analyser divers systèmes de services complexes et industries importantes. Dans la littérature traditionnelle sur les files d'attente de vacances avec des

clients impatients, les études sur le comportement des clients ont toujours été basées sur l'hypothèse que l'impatience des clients ne se produit que lorsque le serveur est en vacances. C'est le cas où les clients peuvent voir l'état du serveur. Cependant, dans de nombreuses situations réelles, y compris les centres d'appels et les systèmes de production, il peut ne pas être possible d'obtenir des informations sur l'état du serveur. De plus, une longue attente dans la file d'attente est un autre facteur qui conduit à l'impatience des clients quel que soit l'état du système (actif ou en vacances).

3.1 Description et notations du modèle

Dans cette section, nous considérons le cas des temps de services généralement distribués, c'est-à-dire que le processus sous-jacent est la file d'attente M/G/1 avec vacances multiples de serveur.

Le processus d'arrivée est Poissonnien avec taux λ. Les temps de service sont des variables aléatoires i.i.d, tous distribués comme B, ayant le premier moment $\mathbb{E}(B)$, deuxième moment $\mathbb{E}(B^2)$ et Transformée de Laplace-Stieltjes (T L-S) $B^*(s) = \mathbb{E}(e^{-sB})$.

A la fin d'une période occupé le serveur prend des vacances U, ayant le premier moment $\mathbb{E}(U)$, et deuxième moment $\mathbb{E}(U^2)$ et Transformée de Laplace-Stieltjes (T L-S) $U^*(s) = \mathbb{E}(e^{-sU})$.

Si le système est vide à la fin des vacances, le serveur prend de nouvelles vacances. S'il y a $n \geq 1$ client à la fin des vacances, le serveur démarre immédiatement une période d'activité. Lorsque le serveur est en vacances et n'est pas disponible pour le service, les clients qui arrivent sont *impatients*. Une arrivée qui constate que le serveur est en vacances, active un " minuteur d'impatience ", T. Si le temps impatience T expire le client abandonne le système. Chaque client active sa propre minuterie et les T_i sont i.i.d. variables aléatoires, indépendantes du nombre de clients en attente.

Soit $t = 0$ l'instant de début des vacances. Ensuite, une observation clé est que, au sein de U, l'évolution du système est la même que celle d'une file $M/G/\infty$ avec des temps de service tous distribués en T. Pour le temps $t \leq U$, il est bien connu [16] (Takacs, 1962) que le nombre de clients dans le système a une distribution de Poisson avec le paramètre :

$$\Lambda(t) = \lambda \int_0^t (1 - \mathbb{P}(T \leq y))dy, \quad t \leq U. \tag{3.1}$$

3.2 Caractéristiques du modèle

3.2.1 Durée d'une période de vacances τ

Considérons l'instant $t = 0$ lorsque le serveur part pour la première fois en vacances de durée U_1. Si à l'instant $t = U_1$ la file d'attente est vide, le serveur prend une autre vacance U_2, et bientôt. Cette séquence d'événements se termine au premier instant lorsque le serveur retourne et trouve un système non vide. Nous appelons cette durée entière, τ. UNE PÉRIODE DE VACANCES.

En utilisant l'analogie $M/G/\infty$, la probabilité d'un système vide au temps U est $e^{-\Lambda(U)}$. Ainsi,

$\tau = \sum_{i=1}^{k} U_i + U_{k+1}$ avec probabilité $(\prod_{i=1}^{k} e^{-\Lambda(U_i)})(1 - e^{-\Lambda(U_{k+1})})$.

Par conséquent, le TLS, $\widetilde{\tau}(s)$, de la Période de vacances est donné par :

$$\begin{aligned}
\tau^*(s) &= \mathbb{E}(e^{-s\tau}) \\
&= \sum_{k=0}^{\infty} \mathbb{E}(e^{-s(\sum_{i=1}^{k} U_i)} e^{-sU_{k+1}} (e^{-\sum_{i=1}^{k} \Lambda(U_i)})(1 - e^{-\Lambda(U_{k+1})})) \\
&= \sum_{k=0}^{\infty} (\mathbb{E}(e^{-(sU+\Lambda(U))})^k (\mathbb{E}(e^{-sU}) - \mathbb{E}(e^{-(sU+\Lambda(U))})) \\
&= \frac{U^*(s) - \mathbb{E}(e^{-(sU+\Lambda(U))})}{1 - \mathbb{E}(e^{-(sU+\Lambda(U))})}.
\end{aligned} \tag{3.2}$$

Donc la durée moyenne d'une période de vacances est :

$$\mathbb{E}(\tau) = \frac{\mathbb{E}(U)}{1 - \mathbb{E}(e^{-\Lambda(U)})}. \tag{3.3}$$

3.2.2 Nombre de clients en début de période occupé

Une période occupé commence avec $N(\tau) \geq 1$ client. Nous dérivons maintenant la fonction génératrice de probabilité (FGP) de $N(\tau)$. Il est à noter que $N(\tau)$ n'est pas distribué comme une variable de Poisson de paramètre $\Lambda(\cdot)$. Il s'ensuit que les dernières vacances U en τ (dans lesquelles il y a au moins une arrivée) ne sont pas régulières. En effet

$$\begin{aligned} U^*(s)|_{N(U)\geq 1} &= \mathbb{E}(e^{-sU}|N(U)\geq 1) \\ &= \frac{\mathbb{E}(e^{-sU}\mathbb{1}_{\{N(U)\geq 1\}})}{\mathbb{E}(\mathbb{P}(N(U)\geq 1)} \\ &= \frac{U^*(s)-\mathbb{E}(e^{-(sU+\Lambda(U))})}{1-\mathbb{E}(e^{-\Lambda(U)})}. \end{aligned}$$

Cela résulte dans

$$\mathbb{E}(U|_{N(U)\geq 1}) = \frac{\mathbb{E}(U)-\mathbb{E}(Ue^{-\Lambda(U)})}{1-\mathbb{E}(e^{-\Lambda(U)})}.$$

Nous écrivons $N(\tau) = \begin{cases} N(U_1) & \text{si } N(U_1)\geq 1, \\ N'(\tau') & \text{si } N(U_1)=0. \end{cases}$

Où $N'(\tau')$ et τ' sont i.i.d, répliques de $N(\tau)$ et τ, respectivement. Alors, le FGP de $N(\tau)$ est donné par

$$\begin{aligned} G_{N(\tau)}(z) &= \mathbb{E}(z^{N(\tau)}) \\ &= \mathbb{E}(\mathbb{E}(z^{N(U)}|N(U)\geq 1)\mathbb{P}(N(U)\geq 1)) + \mathbb{E}(\mathbb{E}(z^{N(\tau)}|N(U)=0)\mathbb{P}(N(U)=0)) \\ &= \mathbb{E}(\sum_{n=1}^{\infty} z^n e^{-\Lambda(U)}\frac{(\Lambda(U))^n}{n!}) + \mathbb{E}(z^{N(\tau)})\mathbb{E}(e^{-\Lambda(U)}). \end{aligned}$$

Ainsi,

$$\begin{aligned} G_{N(\tau)}(z) &= \frac{\mathbb{E}(e^{-(1-z)\Lambda(U)})-\mathbb{E}(e^{-\Lambda(U)})}{1-\mathbb{E}(e^{-\Lambda(U)})} \\ &= \frac{\sum_{n=1}^{\infty}\frac{1}{n!}\mathbb{E}(e^{-\Lambda(U)}(\Lambda(U)^n)z^n}{1-\mathbb{E}(e^{-\Lambda(U)})}. \end{aligned} \tag{3.4}$$

On trouve facilement

$$\mathbb{P}(N(\tau)=n) = \frac{\frac{1}{n!}\mathbb{E}(e^{-\Lambda(U)}(\Lambda(U)^n))}{1-\mathbb{E}(e^{-\Lambda(U)})}, \quad (n=1,2,3,\ldots) \tag{3.5}$$

et

$$\mathbb{E}(N(\tau)) = \frac{\mathbb{E}(\Lambda(U))}{1-\mathbb{E}(e^{-\Lambda(U)})}. \tag{3.6}$$

3.2.3 Période d'occupation

Soit Γ la durée d'une période d'occupation.

Une période occupé commence avec $N(\tau) \geq 1$ clients, est donc égal à la somme de θ_1,θ_2,...,$\theta_{N(\tau)}$, périodes régulières $M/G/1$, sont i.i.d, tous distribués comme θ, où $\theta^*(s) = B^*(s + \Lambda(1 - \theta^*(s)))$. Ainsi, le (TL-S) de Γ est donné par :

$$\Gamma^*(s) = \mathbb{E}(e^{-s\Gamma}) = \mathbb{E}(e^{-s(\sum_{i=1}^{N(\tau)} \theta_i)})$$

$$= \mathbb{E}((\theta^*(s))^{N(\tau)}) = G_{N(\tau)}(\theta^*(s)).$$

En utilisant (3.4) on obtient

$$\Gamma^*(s) = \frac{\mathbb{E}(e^{-(1-\theta^*(s))\Lambda(U)}) - \mathbb{E}(e^{-\Lambda(U)})}{1 - \mathbb{E}(e^{-\Lambda(U)})}. \tag{3.7}$$

Avec $\rho = \lambda\mathbb{E}(B)$

$$\mathbb{E}(\Gamma) = \mathbb{E}(N(\tau))\mathbb{E}(\theta) = \frac{\mathbb{E}(\Lambda(U))}{1 - \mathbb{E}(e^{-\Lambda(U)})} \frac{\mathbb{E}(B)}{1 - \rho}. \tag{3.8}$$

Maintenant, la proportion de temps pendant laquelle le serveur est occupé, $\mathbb{P}_{(occup)}$, est donnée par

$$\mathbb{P}_{(occup)} = \frac{\mathbb{E}(\Gamma)}{\mathbb{E}(\Gamma) + \mathbb{E}(\tau)} = \frac{\mathbb{E}(\Lambda(U))\mathbb{E}(B)}{\mathbb{E}(\Lambda(U))\mathbb{E}(B) + (1 - \rho)\mathbb{E}(U)}. \tag{3.9}$$

3.2.4 Probabilité que le système est vide et le serveur en vacences

Soit D la somme des intervalles de temps, dans τ, où le système est vide. C'est-à-dire

$$D = \int_0^\tau \mathbb{1}_{\{N(t)=0\}} dt. \tag{3.10}$$

En raison de la propriété régénérative du système, nous pouvons écrire

$$D = \int_0^{U_1} \mathbb{1}_{\{N(t)=0\}} dt + D' \mathbb{1}_{\{N(U_1)=0\}}$$

où D' a la même distribution que D. Puisque $\mathbb{E}(\mathbb{1}_{\{N(t)=0\}}) = e^{-\Lambda(t)}$ on a

$$\mathbb{E}(D) = \frac{\mathbb{E}(\int_0^U e^{-\Lambda(t)} dt)}{1 - \mathbb{E}(e^{-\Lambda(U)})}. \tag{3.11}$$

Maintenant, puisque $\mathbb{P}_{00}$ est la fraction de temps pendant laquelle à la fois le système est vide et le serveur est en vacances, nous avons

$$\mathbb{P}_{00} = \frac{\mathbb{E}(D)}{\mathbb{E}(\Gamma) + \mathbb{E}(\tau)}. \tag{3.12}$$

En utilisant (3.11), (3.8) et (3.3) nous obtenons finalement

$$\mathbb{P}_{00} = \frac{\mathbb{E}(\int_0^U e^{-\Lambda(t)}dt)}{\mathbb{E}(\Lambda(U))\mathbb{E}(\theta) + \mathbb{E}(U)}$$

$$\mathbb{P}_{00} = (1-\rho)\frac{\mathbb{E}(\int_0^U e^{-\Lambda(t)}dt)}{\mathbb{E}(\Lambda(U))\mathbb{E}(B) + (1-\rho)\mathbb{E}(U)}. \tag{3.13}$$

Pour le cas où la variable d'impatience T est distribuée exponentiellement avec le paramètre ξ.

$$\mathbb{P}_{00} = \frac{(1-\rho)\mathbb{E}(\int_0^U e^{-\frac{\lambda}{\xi}(1-\xi t)}dt)}{\frac{\rho}{\xi}(1-U^*(\xi)) + (1-\rho)\mathbb{E}(U)}. \tag{3.14}$$

Notez que lorsque $\xi \to 0$ nous obtenons

$$\mathbb{P}_{00} = \frac{1-\rho}{\mathbb{E}(U)}\frac{1-U^*(\lambda)}{\lambda}.$$

Si U est distribué exponentiellement avec le paramètre γ alors cela se simplifie en

$$\mathbb{P}_{00} = \frac{1-\rho}{\mathbb{E}(U)(\lambda+\gamma)} = (1-\rho)\frac{\gamma}{\lambda+\gamma}.$$

3.2.5 Temps de vacances et d'impatience distribués de façon exponentielle

En supposant que $U \sim Exp(\gamma)$, l'équation (3.14) donne

$$\mathbb{P}_{00} = \frac{(1-\rho)\int_{u=0}^{\infty}\gamma e^{-\gamma u}(\int_{t=0}^{u} e^{-\frac{\lambda}{\xi}(1-e^{-\xi t})}dt)du}{\frac{\rho}{\xi}(1-\frac{\gamma}{\gamma+\xi}) + (1-\rho)\frac{1}{\gamma}}.$$

En changeant l'ordre d'intégration et en appliquant le changement de variable :$s = 1 - e^{-\xi t}$ dans le numérateur ci-dessus, on obtient

$$\int_{s=0}^{1} e^{-\frac{\lambda}{\xi}s}(1-s)^{\gamma/\xi}\frac{ds}{\xi(1-s)} = \int_{s=0}^{1}\frac{1}{\xi}(1-s)^{\frac{\gamma}{\xi}-1}e^{-\frac{\lambda}{\xi}s}ds = \frac{K}{\xi},$$

(pour définir K, voir [16])

où la dernière égalité provient de l'équation (2.10)(voir le [16]). Ainsi,

$$\mathbb{P}_{00} = \frac{(1-\rho)\frac{\rho}{\xi}}{\frac{\rho}{\gamma+\xi} + \frac{1-\rho}{\gamma}} = \frac{\gamma K}{\xi} \frac{(\gamma+\xi)(\mu-\lambda)}{\gamma\lambda + (\gamma+\xi)(\mu-\lambda)},$$

qui est l'expression de $\mathbb{P}_{00}$ donnée par l'éq (2.19)(voir le [16]).

3.2.6 Nombre de clients à un instant de fin de service

Soit X_n le nombre de clients présents après l'achèvement du *n-ème* service. On a

$$X_{n+1} =_d \begin{cases} X_n - 1 + A(B) & \text{si } X_n \geq 1, \\ N(\tau) - 1 + A(B) & \text{si } X_n = 0 \end{cases}$$

où $A(t)$ est le nombre d'arrivées de Poisson dans $(0,t]$. Le symbole $=_d$ signifie "égal en distribution". Par conséquent, en régime permanent, le FGP de $X = \lim_{n\to\infty} X_n$ est donné par

$$\begin{aligned}\hat{X}(z) = \mathbb{E}(z^X) &= \mathbb{E}(z^X | X > 0) z^{-1} \mathbb{E}(z^{A(B)})(1-\mathbb{P}_0) + \mathbb{E}(z^{N(\tau)}) z^{-1} \mathbb{E}(z^{A(B)}) \mathbb{P}_0 \\ &= z^{-1} B^*(\lambda(1-z))((\hat{X}(z) - \mathbb{P}_0) + \mathbb{E}(z^{N(\tau)})\mathbb{P}_0)\end{aligned}$$

où $\mathbb{P}_0 = (X = 0)$. Ainsi,

$$\hat{X}(z) = \mathbb{P}_0 \frac{\mathbb{E}(z^{N(\tau)}) - 1}{z - B^*(\lambda(1-z))} B^*(\lambda(1-z)) \tag{3.15}$$

où $\mathbb{E}(z^{N(\tau)}) = G_{N(\tau)}(z)$ est donné par (3.4).

Pour calculer $\mathbb{P}_0$, nous substituons $z = 1$ dans (3.15) et appliquons la règle de L'Hôpital pour obtenir

$$\mathbb{P}_0 = \mathbb{P}(X = 0) = \frac{1-\rho}{\mathbb{E}(N(\tau))} = (1-\rho) \frac{1 - \mathbb{E}(e^{-\Lambda(U)})}{\mathbb{E}(\Lambda(U))}. \tag{3.16}$$

Où $\rho = \lambda \mathbb{E}(B)$ et $\mathbb{E}(N(\tau))$ est donné par (3.6).

Notez que $\mathbb{P}_0 \neq \mathbb{P}_{00}$ puisque $\mathbb{P}_{00}$ est la fraction de temps pendant laquelle le système est vide, tandis que $\mathbb{P}_0$ est la fréquence relative des occurrences, parmi les instants de fin de service, lorsque le système devient vide.

Pour terminer

$$\hat{X}(z) = (1-\rho) \frac{(G_{N(\tau)}(z) - 1) B^*(\lambda(1-z))}{\mathbb{E}(N(\tau))(z - B^*(\lambda(1-z)))}$$

$$= (1-\rho)\frac{(\mathbb{E}(e^{(1-z)\lambda(U)}) - 1)B^*(\lambda(1-z))}{\mathbb{E}(\Lambda(U))(z - B^*(\lambda(1-z)))}. \tag{3.17}$$

Lorsque $T = \infty$, le processus se transforme en file d'attente $M/G/1$ avec plusieurs vacances de serveur. Puis $\Lambda(U) = \lambda U$ et $(G_{N(\tau)}(z) - 1)/\mathbb{E}(N(\tau))$=$(U^*(\lambda(1-z)) - 1)/(\lambda\mathbb{E}(U))$. L'équation (3.17) conduit alors à l'expression connue (voir Levy et Yechiali (1975)[16] et Boxma[9])

$$\hat{X}(z) = (1-\rho)\frac{U^*(\lambda(1-z)) - 1}{\lambda\mathbb{E}(U)(z - B^*(\lambda(1-z)))}B^*(\lambda(1-z)). \tag{3.18}$$

3.2.7 Décomposition stochastique

L'équation (3.17) peut être écrite sous une forme de décomposition, c'est-à-dire,

$$\hat{X}(z) = \hat{L}_{M/G/1}(z)\frac{G_{N(\tau)}(z) - 1}{(z-1)\mathbb{E}(N(\tau))} \tag{3.19}$$

où $\hat{L}_{M/G/1}(z)$ est le (FGP) de l'état du système (occupation) à un moment arbitraire dans la file d'attente régulière M/G/1 correspondante, donné par

$$\hat{L}_{M/G/1}(z) = (1-\rho)\frac{(z-1)B^*(\lambda(1-z))}{z - B^*(\lambda(1-z))}. \tag{3.20}$$

Autrement dit, X est la somme de deux variables aléatoires indépendantes, $L_{M/G/1}$ et Y, où le FPG de Y est donné par

$$\hat{Y}(z) = \frac{G_{N(\tau)}(z) - 1}{(z-1)\mathbb{E}(N(\tau))}.$$

Maintenant, à partir de (3.19) et en utilisant (3.4), il s'ensuit que

$$\mathbb{E}(X) = \mathbb{E}(L_{M/G/1}) + \mathbb{E}(Y)$$

$$= \left(\frac{\lambda^2\mathbb{E}(B^2)}{2(1-\rho)} + \rho\right) + \frac{\mathbb{E}(\Lambda(U)^2)}{2\mathbb{E}(\Lambda(U))}. \tag{3.21}$$

Lorsque $T \to \infty$, $\frac{\mathbb{E}(\Lambda(U)^2)}{2\mathbb{E}(\Lambda(U))} = \frac{\lambda\mathbb{E}(U^2)}{2\mathbb{E}(U)}$=$\Lambda\mathbb{E}(R_U)$, où R_U est la durée de vie résiduelle de U.

3.2.8 Nombre moyen de clients dans le système

Nous calculons maintenant $\mathbb{E}(L)$, $(\overline{n} = \mathbb{E}(L))$, le nombre moyen de clients dans le système. Nous écrivons

$$\mathbb{E}(L) = \mathbb{E}(L|P.V)(1 - \mathbb{P}_{(occup)}) + \mathbb{E}(L|occup)\mathbb{P}_{(occup)}. \tag{3.22}$$

(P.V : Période de Vacances)

Considérez la période de vacances τ. Soit $N(t)$ le nombre de clients dans le système au temps $t \in [0, \tau]$. Soit $\Delta = \displaystyle\int_0^{\tau} N(t)dt$. Puis,

$$\Delta = \int_0^{U_1} N(t)dt + \Delta' \mathbb{1}_{\{N(U_1)=0\}}$$

où Δ' a la même distribution que Δ, et $N(t)$ a une distribution de Poisson avec le paramètre $\Lambda(t)$. On calculé l'espérance on trouve :

$$\mathbb{E}(\Delta) = \frac{\mathbb{E}(\int_0^U N(t)dt)}{1 - \mathbb{E}(e^{-\Lambda(U)})}.$$

Ensuite, en utilisant (3.3),

$$\mathbb{E}(L|P.V) = \frac{\mathbb{E}(\Delta)}{\mathbb{E}(\tau)} = \frac{\mathbb{E}(\int_0^U N(t)dt)}{\mathbb{E}(U)}.$$

Maintenant, $\mathbb{E}(\int_0^U N(t)dt) = \mathbb{E}(\Lambda(t))$, ce qui implique que :

$$\mathbb{E}(L|P.V) = \frac{\mathbb{E}(\int_0^U \Lambda(t)dt)}{\mathbb{E}(U)}. \tag{3.23}$$

En particulier, lorsque $T \sim Exp(\xi)$ et $U \sim Exp(\gamma)$, alors $\mathbb{E}(L|P.V) = \frac{\lambda}{\gamma+\xi}$. Considérons maintenant une période occupé (qui commence avec $N(\tau) \geq 1$ clients). Puis,

$$\mathbb{E}(L|occup) = \mathbb{E}_{N(\tau)}\left(\sum_{n=1}^{N(\tau)} (\mathbb{E}(L_{M/G/1}|occup) + (N(\tau) - n)\right)$$

$$= \mathbb{E}(N(\tau))\mathbb{E}(L_{M/G/1}|occup) + \frac{1}{2}\mathbb{E}_{N(\tau)}(N(\tau)(N(\tau) - 1)). \tag{3.24}$$

Clairement, $\mathbb{E}(L_{M/G/1}|occup) = \mathbb{E}(L_{M/G/1})/\rho$. En collectant les termes, on obtient

$$\mathbb{E}(L) = \mathbb{E}\left(\int_0^U \Delta(t)dt\right) \frac{1-\rho}{(1-\rho)\mathbb{E}(U) + \mathbb{E}(\Lambda(U))\mathbb{E}(B)}$$
$$+ (\mathbb{E}(N(\tau))\mathbb{E}(L_{M/G/1})/\rho + \frac{1}{2}\mathbb{E}((N(\tau)(N(\tau)-1))) \frac{\mathbb{E}(\Lambda(U))\mathbb{E}(B)}{(1-\rho)\mathbb{E}(U) + \mathbb{E}(\Lambda(U))\mathbb{E}(B)}. \tag{3.25}$$

Où $\mathbb{E}(N(\tau))$ est donné par (3.6), $\mathbb{E}(L_{M/G/1})$ est donné par le premier terme du membre de droite de (3.21) et

$$\mathbb{E}(N(\tau)(N(\tau)-1)) = \frac{\mathbb{E}(\Lambda(U)^2)}{1-\mathbb{E}(e^{-\lambda(U)})}, \tag{3.26}$$

qui est dérivé en différenciant (3.4) deux fois à $z = 1$. lorsque $T \sim Exp(\xi)$ et $U \sim Exp(\gamma)$. Alors

•(i) $\mathbb{E}(\Lambda(U)^2) = \frac{\lambda^2}{\xi^2}(1 - \frac{2\gamma}{\xi+\gamma} + \frac{\gamma}{2\xi+\gamma})$,

•(ii) $\mathbb{E}(e^{-\Lambda(U)}) = \frac{\gamma}{\xi}K$,

•(iii) $\mathbb{E}\left(\int_0^U \Delta(t)dt\right) = \frac{\lambda}{\xi}\left(\mathbb{E}(U) - \frac{1}{\xi} + \frac{\gamma}{\xi(\xi+\gamma)}\right) = \frac{\lambda}{\gamma(\xi+\gamma)}$.

La substitution de ce qui précède dans (3.25), avec $\mathbb{E}(\Lambda(U)) = \lambda/(\xi+\gamma)$ (que nous établirons en (3.28)), donne une solution explicite pour $\mathbb{E}(L)$ lorsque T et U sont distribué de façon exponentielle. Notons en outre qu'on peut facilement montrer que dans le cas exponentiel, et avec $\mathbb{E}(B) = 1/\mu$, le premier terme de (3.25) coïncide avec $\mathbb{E}(L_0)$ donné dans la section 2.4(voir [16]).

3.2.9 Proportion de clients servis

Une mesure de performance importante est la proportion de clients servis, notée $\mathbb{P}(servi)$. Nous pouvons écrire :

$\mathbb{P}(servi)$ =Nombre attendu de clients servis au cours d'un cycle/Nombre d'arrivées prévu au cours d'un cycle.

En utilisant (3.9) puis (3.3) et (3.6), on obtient

$$\mathbb{P}(servi) = \frac{\mathbb{E}(\Gamma)/\mathbb{E}(B)}{\lambda(\mathbb{E}(\Gamma)+\mathbb{E}(\tau))} = \frac{\mathbb{P}(occup)}{\rho}$$

$$= \frac{1}{\rho + \lambda(1-\rho)\frac{\mathbb{E}(U)}{\mathbb{E}(\Lambda(U))}}. \tag{3.27}$$

Clairement, $\mathbb{P}(servi) \to 1$ quand $\rho \to 1$. Pour des vacances et des temps d'impatience à distribution exponentielle, où $U \sim Exp(\gamma)$ et $T \sim Exp(\xi)$, on a :

$\lambda(U) = \lambda \int_{y=0}^{U} e^{-\xi y} dy = \frac{\lambda}{\xi}(1 - e^{-\xi U})$

menant à

$$\mathbb{E}(\Lambda(U)) = \int_{u=0}^{\infty} \frac{\lambda}{\xi}(1 - e^{-\xi u})\gamma e^{-\gamma u} du = \frac{\lambda}{\xi + \gamma}. \tag{3.28}$$

Donc

$$\mathbb{P}(servi) = \frac{1}{\rho + (1-\rho)\frac{\xi+\gamma}{\gamma}} = \frac{\gamma}{\gamma + (1-\rho)\xi}. \tag{3.29}$$

Conclusion

Dans ce travail, nous avons analysé d'un système de fils d'attente M/G/1 avec distribution générale du temps de service, vacances du serveur et clients impatients.
En premier lieu, nous avons passé un rappel des processus stochastiques qui sont un outil dans l'analysé de files d'attente, et introduit quelques notions de base de la théorie des files d'attente.
En deuxième lieu, nous avons étudié le système de files d'attentes classiques, et système M/G/1 avec distribution générale du temps de service, et le mesure de performance du système.
Ensuite, nous avons étudié le système de files d'attente M/G/1 avec distribution générale du temps de service, vacances multiples du serveur et client impatient. Nous obtenons les fonctions génératrices des diverses caractéristiques du système (Durée d'une période de vacances, Nombre de clients en début de période occupé, La période d'occupation,. . .).

Bibliographie

[1] A. Gomez-Corral and M.F. Ramalhoto. On the waiting time distribution and the busy period of a retrial queue with constant retrial rate. Stochastic Modelling and Applications, 3, 37-47, (2000).

[2] Abdel-Karim Aboul-Hassan, Sherif I. Rabia and Ahmed Kadry, Analytical study of a discrete time retrial queue with balking customers and early arrival scheme, Alexandria Engineering Journal, 44 (2005), No. 6, 911-91.

[3] Allen,A.O.1990.Probability, Statistics, and Queueing Theory with Computer Science Applications. Second edition, Academic Press, New York (First edition :1978).

[4] Agnès Lagnoux, Claudie Hassenforder :Processus stochastiques modélisation Responsable UE (lagnoux@univ-tlse2.fr) Conception polycopié.

[5] Anisimov, V. V, Zakusilo, O. K, and Donchenko, V. S. 1987. Elements of Queueing Theory and Asymptotic System Analysis. Vishcha Shkola, Kiev (in Russian).

[6] Baynat, B. Théorie des files d'attente-des chaînes de Markove aux réseaux à forme produit. Paris, Hermèse Science Publications , 2000.

[7] B. T. Doshi. Queueing systems with vacations a survey. Queueing Systems Theory and Applications, 1986.

[8] B. Doshi, "Single server queues with vacations," in Stochastic Analysis of Computer and Communication Systems, H. Takag, Ed., pp. 217-265, Elsevier, 1990.

[9] Boxma, O.J., Schlegel, S. and Yechiali, U., "A Note on the M/G/1 Queue with a Waiting Server, Timer and Vacations," American Mathematical Society Translations, Series 2, 207 (2002) 25 :35.

[10] Cheprasov, V. P. Elements of Queueing Theory. Kazan Aviation Institute (in Russian)1985.

[11] Chretienne. P. And Faure. R. Processus stochastique, leurs graphes, leurs usages. Gauthier villars, Paris, 1957.

[12] Claudie Hasseforder CHABRIAC , Eléments de Théorie des files d'attente, page05, Janvier 2008.

[13] Chabriac. C. Processus stochastiques et modélisation. Université de Toulouse le Mirail, Master 2, Année 2012-2013.

[14] D. Gross, J.F. Shortle, J.M. Thompson and C.M. Harris. Fundamentals of Queueing Theory. John Wiley and Sons, 2008.

[15] D. Baum and L. Breuer . An Introduction to Queueing Theory and Matrix-Analytic Methods. Springer, 2005.

[16] Eitan. Altman - Uri. Yechiali. Analysis of customers' impatience in queues with server vacations, Queueing Syst (2006) 52 :261-279.

[17] Florin Avran, "Processus de Markov, de Levy, Files d'attente, Actuariat et Fiabilité", 2009-2010.

[18] H. Takagi, Queueing Analysis : A Foundation of Performance Analysis, vol. 1 of Vacation and Priority Systems, part 1, Elsevier Science Publishers B.V., Amsterdam, The Netherlands, 1991.

[19] Khintchine, A. Y 1969. Mathematical Methods in the Theory of Queueing. Second edition, Hafner Publishing Company, New York (First edition : Griffin,London, 1960 ; Russian original : 1955). 859-866.

[20] Kashyap, B. R. K., and Chaudhry, M. L. An Introduction to Queueing Theory. Aarkay, Calcutta, India. 1988.

[21] Lionel. B. Processus stochastique : Processus de poisson et chaîne de Markov. 2004.

[22] M. Boualem, N. Djellab, D. Aissani, "An M/G/1 retrial queue with exhaustive service and server vacations", Journal of Communication and Computer 8 (2011) 720-726.

[23] Moshe Zukerman. Introduction to Queueing Theory and Stochastic Teletraffic Models. EE Department. City University of Hong Kong.

[24] Newell, G. F. 1982. Applications of Queueing Theory. Second edition, Chapman and Hall, London (First edition : 1971).

[25] Nicolas Fournier, d'après Philippe Bougerol. Processus de Sauts et Files d'Attente. Université Pierre et Marie Curie. Master 1 de Mathématiques.

[26] N. Tian and Z. G. Zhang, Vacation Queueing Models : Theory and Applications, Springer, New York, NY, USA, 2006.

[27] Pnilippe. N. Basic elements of queueing theory : application to medeling of computer systems. Tech report, The University of Massachusetts, 2004.

[28] P. Robert. Réseaux et Files d'Attente : Méthodes Probabilistes. Springer-Verlag Berlin Heidelberg, 2000.

[29] Philippe Malbos. Analyse Matricielle et Algèbre Linéaire Appliquée. université Claude Bernard Lyon 1.Licence Sciences, Technologies, Santé.

[30] Raphaël. Danchin. Equations différentielles L3 de Mathématiques. Année 2010 - 2011.

[31] Rugg. R. Processus stochastique. Presses Polytechniques Romandes, 1989.

[32] Ruegg.A : processus stochastique, Presses polytechnique romandes, Lausanne Suisse,(1989).

[33] S. Ross. Stochastic Processes. John-Wiley and Sons, New York, 2e éd. 1996.

[34] Sheldon M. Ross. Introduction to Probability Models Tenth Edition. University of Southern California Los Angeles, California.

[35] Sanjay K. Bose, "Analyzing the M/G/1 queue using the Methode of Supplementary Variables", 2002.

[36] Sébastien Loustau. Chaînes de Markov et Processus markoviens de sauts. Applications aux files d'attente. Ecole Centrale de Marseille, Année 2008-2009.

[37] T.L. Saaty. Elements of Queueing Theory. M.C. Graw-Hill, New York, 1961.

[38] Takacs, L. "Introduction to the Theory of Queues," Oxford University Press, New York (1962).

[39] U. Narayan Bhat. An Introduction to Queueing Theory, Modeling and Analysis in Applications. Second Edition.

[40] Watson, G. S. and Leadbetter, M. R. Hazard Analysis I. Biometrika, 51, 175-184 (1964 a).

[41] Watson, G. S. and Leadbetter, M. R. Hazard Analysis II. Sankhya Ser. A, 26, 101-116 (1964 b).

[42] W.Grassmann, Modeling Markovien Queues and Similar Process, Departement of Cumputer Science, University of Sakatchewan, 2000.

[43] Y. Levy and U. Yechiali, "Utilization of idle time in an M/G/1 queueing system," Management Science, vol. 22, no. 2, pp. 202- 211, 1975.

[44] Zakhar Kabluchko, Stochastic Processes (Stochastik II), University of Ulm Institute of Stochastics, (2013-2014).

[45] Zhang, M., Hou, Z. : Performance analysis of M/G/1 queue with working vacations and vacation interruption. J. Comput. Appl. Math. 234(10), 2977-2985 (2010).