

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE ABDERAHMANE MIRA DE BEJAIA  
FACULTE DES SCIENCES EXACTES  
DEPARTEMENT MATHEMATIQUES



## Mémoire de Master

En vue de l'obtention du diplôme de Master en Mathématiques

Option : Statistique et Analyse Décisionnelle

Thème

---

Conception et Réalisation d'un Simulateur de  
Systèmes d'Attente Non-Markoviens Sous Java

---

Présenté par :

**Mr. HOCINE ZAHIR**

*Devant le jury composé de :*

Président :	BOUALEM	MOHAMED	Professeur	U. A. Mira, Bejaïa.
Examineur:	HAMMOUDI	SAMIR	M.A.B.	Université de Bejaïa.
Rapporteur :	OURBIH	MEGDOUDA	Professeur	Université de Bejaïa

---

*Promotion 2013-2014*

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE  
UNIVERSITE ABDERAHMANE MIRA DE BEJAIA  
FACULTE DES SCIENCES EXACTES  
DEPARTEMENT MATHEMATIQUES



## Mémoire de Master

En vue de l'obtention du diplôme de Master en Mathématiques

Option : Statistique et Analyse Décisionnelle

Thème

---

Conception et Réalisation d'un Simulateur de  
Systèmes d'Attente Non-Markoviens Sous Java

---

Présenté par :

**Mr. HOCINE ZAHIR**

*Devant le jury composé de :*

Président :	BOUALEM	MOHAMED	Professeur	U. A. Mira, Bejaïa.
Examineur:	HAMMOUDI	SAMIR	M.A.B.	Université de Bejaïa.
Rapporteur :	OURBIH	MEGDOUDA	Professeur	Université de Bejaïa

---

*Promotion 2013-2014*

# Remerciements

---

Je remercie Dieu tout puissant de m'avoir donné la force, la santé, le courage et la patience de pouvoir accomplir ce modeste travail.

Je remercie énormément ma promotrice Madame MEGDOUDA OURBIH, professeur à l'université de Béjaia, d'avoir accepté de mettre toute sa confiance en moi pour bien mener ce mémoire.

Je remercie infiniment les membres de jury M<sup>r</sup> BOUALEM et M<sup>r</sup> HAMOUDI SAMIR d'avoir jugé et évaluer mon modeste travail, et je leur suis très reconnaissant pour l'intérêt de leurs critiques constructives.

Je remercie M<sup>r</sup> SALHI pour ces informations consistantes et très pertinentes dans les bases de données et le langage UML.

Je remercie ma chère famille et celle de mon frère OUEZGGANE REDOUANE de m'avoir soutenu tout au long de mon cursus.

Je remercie ma future épouse, B.S., que Dieu me réunis avec elle, pour sa compréhension et sa tendresse.

Je remercie sans exception tous mes amis.

# *Dédicace*

---

Je dédie le fruit de ce travail à ma chère famille

Je dédie ce modeste mémoire à la famille de mon frère Redouane

Je dédie ces simples idées que j'ai mises dans quelques papiers à tous mes amis

---

# TABLE DES MATIERES

---

<b>TABLE DES MATIERES.....</b>	<b>I</b>
<b>LISTE DES FIGURES.....</b>	<b>IV</b>
<b>INTRODUCTION GENERALE.....</b>	<b>1</b>
<b>CHAPITRE I : GENERALITES ET CONCEPTS DE BASE .....</b>	<b>3</b>
I.1. Introduction .....	3
I.2. Origine de la théorie des files d'attente.....	3
I.3. Intérêt des files d'attente .....	3
I.4. Une modélisation simple de la file d'attente.....	4
I.5. Analyse mathématique d'un système de files d'attente.....	6
I.6. Types de modèles .....	7
I.6.1. Modèles markoviens .....	7
I.6.2. Modèles non markoviens.....	7
I.7. Caractéristiques d'un système de files d'attente.....	8
I.8. Stabilité d'une file d'attente.....	10
I.9. Processus d'arrivée des clients dans une file.....	10
I.9.1. Le processus de Poisson.....	10
I.9.4. Utilisation de la chaîne de Markov .....	12
I.10. Utilisation de processus de naissance et de mort .....	20
I.11. Conclusion .....	21
<b>CHAPITRE II : SYSTEMES D'ATTENTE MARKOVIENS .....</b>	<b>22</b>
II.1. Introduction.....	22
II.2. La file d'attente M/M/1.....	22

II.3. La file d'attente à une source finie $M/M/1/K$ .....	26
II.4. La file d'attente $M/M/1/1$ .....	28
II.5. La file d'attente $M/M/s$ .....	30
II.6. La file d'attente à source finie $M/M/s/K$ .....	34
II.7. La file d'attente $M/M/s/s$ .....	37
II.8. La file d'attente $M/M/\infty$ .....	38
II.9. La file d'attente $M^{(x)}/M/1$ .....	39
II.10. La file d'attente $M/M^{(x)}/1$ .....	42
II.11. Conclusion.....	44
 <b>CHAPITRE III : SYSTEMES D'ATTENTE NON-MARKOVIENS</b> .....	<b>48</b>
III.1. Introduction.....	48
III.2. Le système d'attente $M/G/1$ .....	49
III.3. Le système d'attente $M/G/1/K$ .....	61
III.4. Le système d'attente $G/M/1$ .....	63
III.5. Le système d'attente $G/M/s$ .....	69
III.6. Le système d'attente $G/M/1/K$ .....	71
III.7. Le système d'attente $M/G^K/1$ .....	72
III.8. Le système d'attente $G^K/M/1$ .....	74
III.9. Le système d'attente $G/G/1$ .....	75
III.10. Conclusion.....	82
 <b>CHAPITRE IV : APPLICATION ET SIMULATION D'UN MODELE NON-MARKOVIEN</b> .....	<b>86</b>
IV.1. Introduction .....	86
IV.2. Phase d'Analyse des Besoins .....	86
IV.2.1. Description de l'Application .....	86
IV.2.2. Cahier des Charges.....	86
IV.2.3. Diagramme des cas d'utilisation.....	87
IV.2.4. Diagrammes de Séquence Globale.....	87

---

<b>IV.3. Phase de Conception</b> .....	<b>88</b>
IV.3.1. Diagramme de Package.....	88
IV.3.1. Diagramme de classes .....	90
<b>IV.4. Phase de Réalisation</b> .....	<b>91</b>
<b>IV.4.1. Choix Techniques</b> .....	<b>91</b>
<b>IV.4.2. Interfaces Homme-Machine</b> .....	<b>92</b>
<b>IV.6. Conclusion</b> .....	<b>95</b>
<b>CONCLUSION GENERALE</b> .....	<b>96</b>
<b>BIBLIOGRAPHIE</b> .....	<b>97</b>

---

# LISTE DES FIGURES

---

<b>Figure 1:</b> Modèle de File d'Attente .....	4
<b>Figure 2:</b> Caractéristiques d'un système de files d'attente .....	9
<b>Figure 3:</b> La représentation des flux d'entrée et de sortie .....	10
<b>Figure 4:</b> la représentation des variables aléatoires $N(t)$ et $T$ .....	11
<b>Figure 5:</b> La représentation d'une CMTD .....	13
<b>Figure 6:</b> la représentation du comportement d'une CMTC .....	17
<b>Figure 7:</b> La représentation d'un système d'attente M/G/1 .....	49
<b>Figure 8:</b> La représentation d'un système d'attente G/M/1 .....	63
<b>Figure 9:</b> Les processus de temps d'attente $W(t)$ et $W_n$ .....	76
<b>Figure 10 :</b> Diagramme de cas d'utilisation du QueueSimulator .....	88
<b>Figure 11 :</b> Diagramme de Séquence pour le cas « Créer un Modèle » .....	89
<b>Figure 12 :</b> Diagramme de Package de l'application QueueSimulator .....	90
<b>Figure 13 :</b> Diagramme de Classe de l'Application « QueueSimulator » .....	91
<b>Figure 14 :</b> Fenêtre Principale de « QueueSimulator » .....	92
<b>Figure 15 :</b> Modèle d'Attente sans paramètres .....	92
<b>Figure 16 :</b> Les deux boites de dialogue pour la définition de loi .....	93
<b>Figure 17 :</b> Modèle d'Attente avec paramètres des lois .....	93
<b>Figure 18 :</b> Paramètres de simulation .....	94
<b>Figure 19 :</b> Boite de dialogue pour les résultats de Simulation .....	94



INTRODUCTION  
GENERALE

# INTRODUCTION GENERALE

---

Les phénomènes d'attente, on les observe quotidiennement dans la vie courante. Quand nous nous rendons à la poste, à la gare, à la banque, bien souvent nous devons “faire la queue” pour obtenir des timbres, un billet, de l'argent. Néanmoins, l'étude formelle de la théorie des files d'attente a seulement commencé au début du vingtième siècle avec le travail de **A.K. Erlang[Er]**. Il a réalisé la première analyse mathématique d'un modèle de file d'attente téléphonique. Depuis lors, la théorie des files d'attente s'est développée et elle a une large application dans divers domaines.

La théorie des files d'attente joue un rôle très prépondérant dans plusieurs domaines d'application que d'autres diverses théories connues dans le monde des mathématiques mais elle aussi, elle s'appuie sur des bases mathématiques comme par exemple la théorie des probabilités ainsi que la théorie des processus stochastiques.

La file d'attente la plus simple est  $M/M/1/FIFO$  à une capacité illimitée modélisée par le processus de naissance et de mort qui est très facile à étudier et de voir aussi quelques indices de performances de ce système. La file la plus complexe dans l'analyse des systèmes d'attente est la file  $G/G/1/FIFO$  à une source infinie, c'est quand les arrivées suivent une loi générale et les temps de service sont aussi d'une distribution générale ce qui veut dire on n'a pas la propriété de Markov sans mémoire.

L'objectif de ce mémoire est d'analyser le comportement des systèmes d'attente non markoviens. Pour atteindre cet objectif, nous avons structuré ce mémoire en quatre chapitres, à savoir :

Le premier chapitre sera consacré à quelques concepts de base et introduit quelques généralités nous permettant de préparer les chapitres suivants.

Pour le second chapitre, nous allons voir les systèmes d'attente markoviens qui sont définis par des arrivées régis par la loi de Poisson et les durées de service qui sont exponentielles et donner quelques mesures de performances et la distribution stationnaire de chaque système.

Le chapitre suivant présentera une analyse bien détaillée sur les systèmes d'attente non markoviens qui se diffèrent complètement des systèmes markoviens. Cette analyse sera

faite par les chaînes de Markov induites. Et à la fin de ce chapitre, quelques caractéristiques statistiques seront données pour chaque système ainsi que leur distribution stationnaire dans un régime permanent.

Le dernier chapitre sera consacré à la conception et la réalisation d'un simulateur de systèmes d'attente non markoviens. L'objectif principal de notre travail est d'avoir un simulateur paramétrable et générique dans la mesure de pouvoir représenter le plus grand nombre de modèles de files d'attente et de prendre en considération la quasi-totalité des distributions de probabilités existante, à savoir : l'Exponentiel, l'Uniforme, l'Hyper-exponentiel, Pareto, Cox, Weibull, etc. Ceci nous permet d'avoir les performances opérationnelles de différents systèmes d'attente, comme par exemple, le nombre moyen de clients dans le système, le temps moyen de réponse (durée moyenne de séjour), le débit moyen d'entrée, le débit moyen de sortie et la moyenne des mesures de performances calculées après plusieurs itérations.

# CHAPITRE I

GENERALITES ET CONCEPTS DE BASE

# CHAPITRE I

## GENERALITES ET CONCEPTS DE BASE

---

### *1.1. Introduction*

Dans ce premier chapitre nous allons présenter quelques concepts et outils de base pour bien mener les autres chapitres par la suite.

Un phénomène d'attente peut être décrit comme un système composé d'un certain nombre (fini ou non) de places d'attente d'un ou plusieurs serveurs et de clients arrivant à des instants aléatoires. Les clients attendent, se font servir selon des règles spécifiées et quittent le système. Ils peuvent être des appels téléphoniques, des machines, ..., de même que les serveurs peuvent être un central téléphonique, un processeur,.... Quand les serveurs sont tous occupés, les clients doivent alors patienter dans un espace d'attente (s'il existe) jusqu'à ce qu'un serveur soit disponible.

L'identification des systèmes de files d'attente classiques se base principalement sur trois éléments : le processus stochastique décrivant l'arrivée des clients dans le système, le mécanisme de service (le nombre de serveurs et la loi probabiliste décrivant la durée des services) et la discipline d'attente.

### *1.2. Origine de la théorie des files d'attente*

- La théorie des files d'attente fut développée pour fournir des modèles permettant de prévoir le comportement de systèmes répondant à des demandes aléatoires.
- Les premiers problèmes étudiés concernaient la congestion du trafic téléphonique (Erlang, « the theory of probabilities and telephone conversations », 1909)
- Erlang constata qu'un système téléphonique peut être modélisé par des arrivées de clients poissonniennes et des temps de service exponentiels
- Molina, Pollaczek, Kolmogorov, Khintchine, Palm, Crommelin reprirent le flambeau.

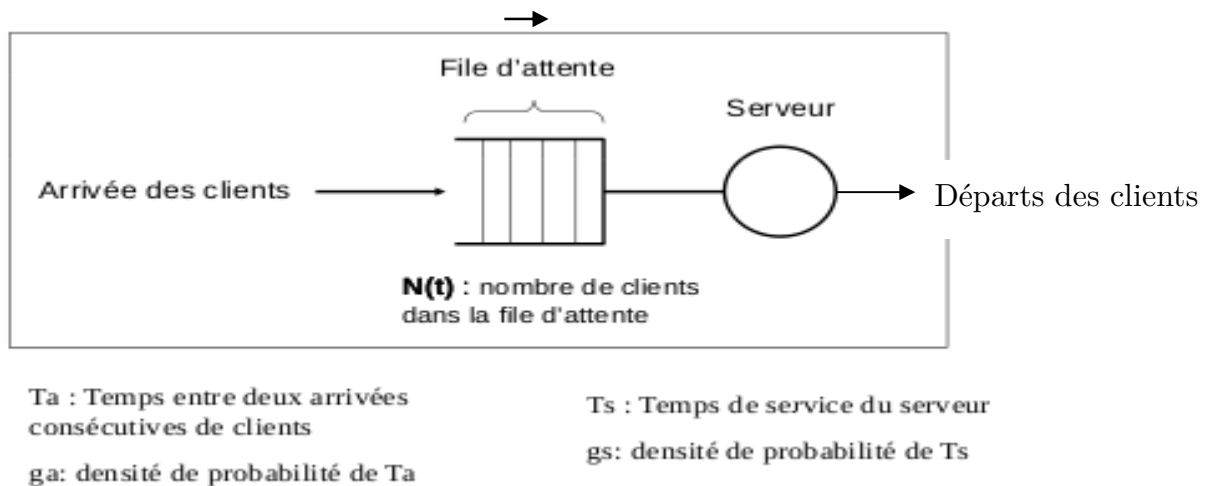
### *1.3. Intérêt des files d'attente*

La théorie des files d'attente a de nombreuses applications dans:

- La gestion de trafic (réseaux de communication, compagnies aériennes,

- embouteillages, ...)
- La planification (opérations sur des machines de production, programmes sur des ordinateurs, ...)
- Le dimensionnement d'infrastructures (usines, ...)

#### *I.4. Une modélisation simple de la file d'attente*



*Figure 1: Modèle de File d'Attente*

Le processus d'arrivée spécifie les instants auxquels les clients arrivent dans le système. Dans la théorie classique des files d'attente, on fait le plus souvent l'hypothèse que les clients arrivent de manière isolée et indépendamment les uns des autres. Sous ces hypothèses, les intervalles de temps entre deux arrivées successives forment une suite de variables aléatoires indépendantes et identiquement distribuées.

La liste qui suit résume les lois de probabilité les plus couramment rencontrées dans la modélisation des systèmes de files d'attente ainsi que les symboles associés.

- ✗ La lettre M désigne la loi exponentielle.
- ✗ La lettre D correspond à une loi déterministe.
- ✗ Le symbole  $E_k$  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  $k$ .
- ✗ La lettre G est utilisée lorsqu'aucune hypothèse particulière n'est faite sur le processus d'arrivée, ce dernier étant alors un processus de renouvellement quelconque.

Les temps de service nécessaire au traitement des clients sont supposés être des réalisations de variables aléatoires indépendantes et identiquement distribuées. La description du processus de service revient alors à préciser la loi de probabilité de ces variables aléatoires. Les symboles utilisés pour décrire les processus de service sont les mêmes que ceux introduits pour les processus d'arrivée.

Le nombre de serveurs correspond au nombre maximal de clients pouvant être traités simultanément. Tous les serveurs sont supposés identiques, en particulier les temps de service sont indépendants d'un serveur à l'autre et distribués selon une même loi de probabilité.

La capacité d'accueil d'un système de files d'attente correspond au nombre maximal de clients pouvant être présents dans le système à un instant quelconque. Il est égal à la somme du nombre de serveurs et du nombre de places d'attente disponibles. Si un client arrive dans un système ayant atteint sa capacité maximale d'accueil, il est refoulé et doit quitter le système sans avoir été servi.

Le plus souvent, le nombre de clients susceptibles d'accéder au service est supposé illimité et leur fréquence d'arrivée constante. Certaines situations sont, cependant, caractérisées par un nombre fixe et limité de clients. Chaque client présent dans le système diminue alors le nombre d'arrivées potentielles, le taux d'arrivée dans la file n'est donc plus constant mais dépend du nombre de clients présents dans le système. Dans un tel cas, le processus d'arrivée décrit le temps nécessaire à un client entre le moment où il quitte le système et celui où il y revient.

La discipline d'attente, est la règle de priorité déterminant l'ordre dans lequel les clients vont accéder à la ressource modélisé par le serveur. Les disciplines d'attente classiques, ainsi que leurs acronymes, sont :

- ✗ FIFO : first in first out "premier arrivé, premier servi", c'est la discipline de service employée le plus souvent et c'est celle qui sera admise par défaut ;
- ✗ LIFO: last in, first out "dernier arrivé, premier servi" ;
- ✗ SIRO : service dans un ordre aléatoire ;
- ✗ RR : round robin "les clients sont servi à tour de rôle pendant un intervalle de temps fixe, appelé quantum"
- ✗ PS : processor sharing "cas limite de la discipline RR lorsque le quantum tend vers zéro".

Certains modèles de files d'attente font usage de disciplines plus complexes reposant sur des priorités de traitement. Un système peut également accepter la préemption d'un service, c'est-à-dire l'interruption du service d'un client au profit d'un autre. Si un tel mode de fonctionnement est autorisé, le service interrompu peut être simplement terminé plus tard (preemptive resume) ou, alors, être recommencé depuis le début, le travail déjà effectué étant perdu (preemptive repeat).

Pour la classification des systèmes de files d'attente, on a recours à une notation symbolique introduite par Kendall, comprenant six symboles rangés dans l'ordre  $A/B/c/m/P - D$ , où :

- **A**: décrit la nature du processus d'arrivée,
- **B**: décrit la nature du processus de service,
- **c**: est le nombre de serveurs,
- **m**: est la capacité d'accueil d'un système,
- **P**: est la taille de la population,
- **D**: est la discipline de la file.

Dans sa version courte, seuls les trois premiers symboles  $A/B/c$  sont utilisés. Dans un tel cas, on suppose que la file est régie par une discipline FIFO et que le nombre de places d'attente ainsi que celui des clients susceptibles d'accéder au système sont illimités.

### ***1.5. Analyse mathématique d'un système de files d'attente***

L'étude mathématique d'un système de files d'attente se fait généralement par l'introduction d'un processus stochastique, défini de façon appropriée. 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 déterminer :

- ✱ Les probabilités d'état ( $P_n = P(N(t) = n)$ ), 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 qui est défini par :  $\pi(n) = \lim_{t \rightarrow \infty} P_n(t) = P(N(\infty) = n)$ ,  $n = 0, 1, 2, \dots$  et  $\{\pi(n), n \geq 0\}$  est appelée distribution stationnaire du processus  $\{N(t), t \geq 0\}$ . Le calcul explicite du régime transitoire s'avère généralement pénible, voire impossible, pour la plupart des modèles donnés. On se contente donc de déterminer le régime stationnaire.



---

## ***I.6. Types de modèles***

### **I.6.1. 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  $\{N(t), t \geq 0\}$  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.

### **I.6.2. Modèles non markoviens**

En l'absence de la propriété sans mémoire (la loi exponentielle) ou lorsque l'on s'écarte de l'hypothèse de cette dernière, 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 introduction de paramètres supplémentaires, on aboutit à un modèle non 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 :

#### ***a) Méthode des étapes d'Erlang***

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

#### ***b) Méthode de la chaîne de Markov induite***

Élaborée par *Kendall*, et souvent utilisée, elle consiste à choisir une suite d'instantanés  $1, 2, 3, \dots, n$  (déterministes ou aléatoires) tels que la chaîne induite  $\{N_n, n > 0\}$ , où  $N_n = N$ , soit markovienne et homogène.

**c) Méthode des variables supplémentaires**

Elle consiste à compléter l'information sur le processus  $\{N(t), t \geq 0\}$  de telle manière à lui donner le caractère markovien. Ainsi, on se ramène à l'étude du processus  $\{N(t), A(t_1), A(t_2), \dots, A(t_n), t \geq 0\}$ . Les variables  $A(t_k), k \in \{1, 2, \dots, n\}$ , sont dites supplémentaires.

**d) Méthode des événements fictifs**

Le principe est d'introduire des événements fictifs qui permettent de donner une interprétation probabiliste aux transformées de Laplace et aux variables aléatoires décrivant le système étudié.

**e) Simulation**

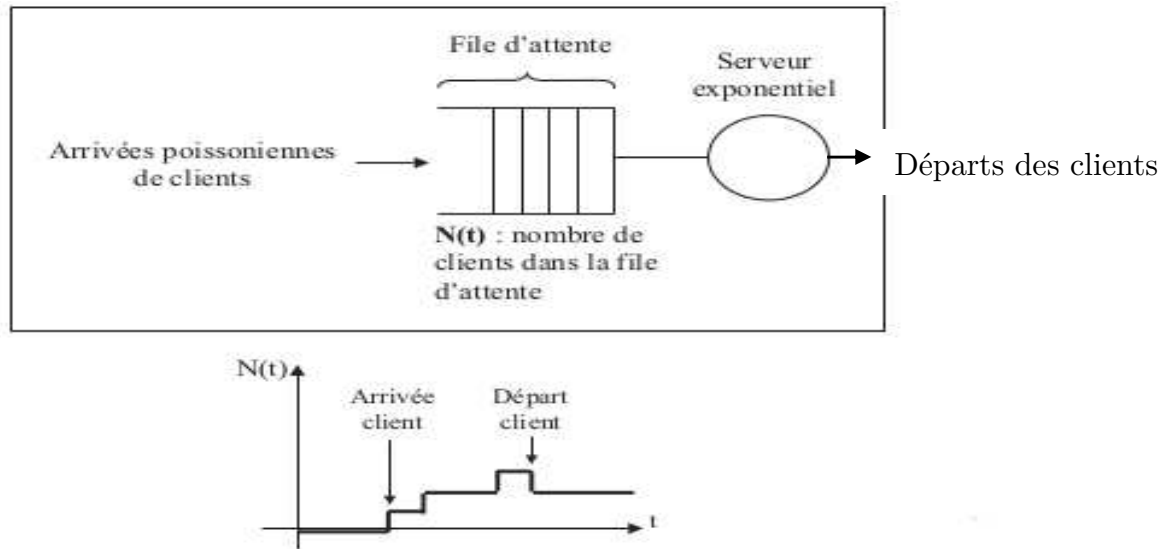
C'est un procédé d'imitation artificielle d'un processus réel effectué sur ordinateur. Elle nous permet d'étudier les systèmes les plus complexes, de prévoir leurs comportements et de calculer leurs caractéristiques. Les résultats obtenus ne sont qu'approximatifs, mais peuvent être utilisés avec une bonne précision. Cette technique se base sur la génération de variables aléatoires suivant les lois gouvernant le système.

**I.7. Caractéristiques d'un système de files d'attente**

Soit  $\lambda$  le taux d'arrivée des clients. Cela signifie que l'espérance mathématique de la durée séparant deux arrivées successives est :  $E[A] = \frac{1}{\lambda}$ .

Soit  $\mu$  le taux de service des clients. Cela signifie que l'espérance de la durée de service est :  $E[S_e] = \frac{1}{\mu}$ .

L'intensité du trafic s'exprime de la manière suivante :  $\rho = \frac{\lambda}{\mu} = E[S_e] = E\left[\frac{1}{\mu}\right]$



*Figure 2: Caractéristiques d'un système de files d'attente*

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 (le temps qu'un client passe dans la file d'attente), le temps de séjour d'un client dans le système (composé du temps d'attente et de la durée de service), le taux d'occupation des dispositifs de service, la durée de la période d'activité (l'intervalle de temps pendant lequel il y a toujours au moins un client dans le système) ; et les mesures de performance suivantes :

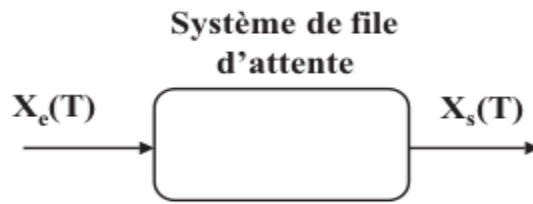
- $L$  : nombre moyen de client dans le système de files d'attente,
- $L_q$  : nombre moyen de clients dans la file d'attente,
- $W$  : temps moyen de séjour d'un client dans le système,
- $W_q$  : temps moyen d'attente d'un client dans la file d'attente.

Ces valeurs sont liées les unes aux autres par les relations suivantes :

- $L = \lambda W$  ,
- $L_q = \lambda W_q$  ,
- $L = L_q + \lambda / \mu$  ,
- $W = W_q + 1 / \mu$  .

Les deux premières formules sont appelées "formules de Little". Il est à noter que ces formules sont valables sous la vérification de la condition que le système est en régime stationnaire  $\rho = \lambda / \mu < 1$ .

### ***I.8. Stabilité d'une file d'attente***



**Figure 3:** La représentation des flux d'entrée et de sortie

#### **Définition**

Une file d'attente est stable si et seulement si :

$$\lim_{T \rightarrow \infty} X_s(T) = \lim_{T \rightarrow \infty} X_e(T)$$

- ✓  $X_e(T)$  : le flux d'entrée des clients dans le système à l'instant  $T$
- ✓  $X_s(T)$  : le flux de sortie des clients dans le système à l'instant  $T$

La stabilité entraîne :

$$\lim_{T \rightarrow \infty} \left( \frac{D(T)}{A(T)} \right) = 1$$

- ✓  $D(T)$  : le nombre des départs des clients à l'instant  $T$
- ✓  $A(T)$  : le nombre des arrivées des clients à l'instant  $T$

Dans une file stable, le nombre de clients reste fini.

### ***I.9. Processus d'arrivée des clients dans une file***

#### **I.9.1. Le processus de Poisson**

Le processus de Poisson occupe une place privilégiée pour d'écrire:

- l'arrivée des clients vers un guichet
- l'occurrence d'accidents dans une entreprise
- l'apparition de pannes dans un parc de machines
- l'arrivée de tâches dans l'unité centrale d'un ordinateur etc.

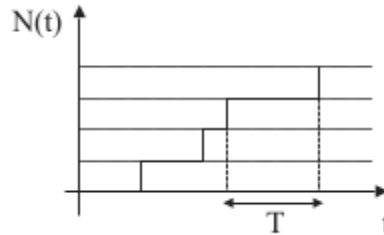
Il y a principalement deux variables aléatoires à considérer :

- Le nombre de clients  $\{N(t)\}$  arrivant dans la file pendant le temps  $t$  ; cette

variable est un entier positif ou nul.

- Le temps  $T$  qui s'écoule entre deux arrivées consécutives ; cette variable est un nombre réel positif ou nul.

Ces deux variables aléatoires sont liées car si le nombre de clients  $\{N(t)\}$  qui arrivent dans la file est élevé, c'est que le temps entre deux arrivées successives est faible. De manière plus précise,



**Figure 4:** la représentation des variables aléatoires  $N(t)$  et  $T$

### Théorème

Les trois conditions suivantes sont équivalentes :

1. La probabilité qu'un client arrive dans la file pendant un intervalle de temps infiniment petit  $\varepsilon > 0$  vaut  $\lambda \varepsilon + 0(\varepsilon)$  avec  $\lambda$  constant
2. Le nombre de clients  $N(t)$  arrivant dans la file pendant un intervalle de temps quelconque  $T$  suit une loi de Poisson de moyenne  $\lambda t$ , autrement dit,

$$P\{N(t) = k\} = e^{-\lambda t} [(\lambda t)^k / k!]$$

3. Le temps  $T$  qui s'écoule entre deux arrivées consécutives obéit à une loi exponentielle de paramètre (moyenne  $(1/\lambda)$ ), autrement dit,

$$P\{T > t\} = e^{-\lambda t}$$

#### **I.9.1.1. Utilisation de la loi de Poisson $P(\lambda)$**

On dit que la variable aléatoire entière  $N$  suit une loi de Poisson de paramètre  $\lambda$  lorsque :

$$P\{N = k\} = e^{-\lambda} [\lambda^k / k!]$$

Il reste à justifier pourquoi le nombre de clients entrant dans une file d'attente pendant un temps  $t$  obéit raisonnablement à une loi de Poisson. Cette loi est appelée loi des événements rares, en raison du fait qu'elle s'obtient comme la limite d'une loi de Bernoulli quand le nombre d'épreuves tend vers l'infini en même temps que la probabilité

de succès d'une épreuve tend vers zéro.

Soit une population de  $n$  personnes susceptibles de rejoindre indépendamment la file pendant une unité de temps, chacune avec une probabilité  $p$ . Le nombre  $N$  de personnes qui arrivent dans la file suit une loi de Bernoulli  $B(n, p)$ .

$$P\{N = k\} = C_n^k p^k (1 - p)^{n-k}$$

$$P(\lambda) = \lim_{n \rightarrow \infty} B(n, \lambda/n)$$

Ainsi le choix d'une loi de Poisson pour compter le nombre de clients qui arrivent dans une file d'attente pendant un intervalle de temps donné est justifié par le fait que la population susceptible d'alimenter cette file est nombreuse et que les choix individuels sont pris de manière indépendante.

### *I.9.1.2. Propriétés des processus de Poisson*

- Probabilité pour qu'un client arrive pendant  $dt \approx \lambda dt$

$$P[N(t + dt) = k + 1 | N(t) = k] = \lambda t + 0(dt)$$

- Probabilité pour que 0 client arrive pendant  $1 - dt \approx \lambda dt$

$$P[N(t + dt) = K + 1 | N(t) = K] = 1 - \lambda t + 0(dt)$$

- Probabilité pour que plus d'un client arrive pendant  $dt \approx 0(dt)$

$$P[N(t + dt) = K + j | N(t) = K] = 0(dt), \quad j \geq 2$$

### **I.9.4. Utilisation de la chaîne de Markov**

Il y a deux types de chaînes de Markov à savoir:

#### *I.9.4.1. Chaînes de Markov à temps discret*

On dit qu'un processus stochastique à espace d'état à temps discret  $X_n$ ,  $n > 0$  est une chaîne de Markov à temps discret (CMTD) si seulement si :

$$\forall n \in N, \forall (j, i_0, i_1, i_2, \dots, i_n) \in E^{n+2}$$

$$P[X_{n+1} = j | X_n = i_n, \dots, X_0] = P[X_{n+1} = j | X_n = i_n] = P_{ij}(n)$$

Dans une CMTD, le comportement futur ne dépend que de l'état actuel et  $P_{ij}(n)$  est appelé probabilité de transition de l'état  $i$  à l'état  $j$  à l'instant  $n$ .

Une CMTD est homogène si et seulement si ses probabilités de transition ne dépendent pas de  $n$ .

$$P[X_{n+1} = j | X_n = i_n] = P_{ij}(n) = P_{ij}$$

### I.9.4.1.2. Matrice stochastique

Une matrice carrée est dite stochastique si et seulement si :

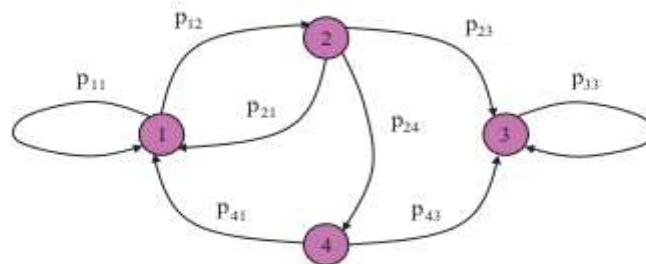
- Tous ses termes sont positifs ou nuls
- La somme des termes de chaque ligne vaut 1

#### Propriétés

- Une matrice de transition est une matrice stochastique
- Si  $P$  est stochastique, alors  $P^n$  est stochastique
- Les valeurs propres de  $P$  ne dépassent pas l'unité :  $|\lambda| \leq 1$
- Hypothèses :
  - Dans le reste du chapitre, le terme CMTD désignera (sauf mention contraire) une Chaîne de Markov :
  - à temps discret
  - définie sur un ensemble d'états  $E$  fini
  - homogène dans le temps

Notons que la plupart des résultats s'appliquent même si l'espace d'état n'est plus fini mais dénombrable.

### I.9.4.1.3. Représentation Graphique d'une CMTD



**Figure 5:** La représentation d'une CMTD

**I.9.4.1.4. Classification des États**

Soit  $f_{jj}$  la probabilité de revenir en  $j$  après l'avoir quitté, on dit alors que :

- Un état  $j$  est transitoire si seulement si  $f_{jj} < 1$
- Un état  $j$  est récurrent si seulement si  $f_{jj} = 1$
- Un état  $i$  est absorbant si seulement si  $P_{ii} = 1$

**I.9.4.1.5. Chaîne de Markov Irréductible**

Une CMTD est dite irréductible si et seulement si de tout état  $i$ , on peut atteindre tout état  $j$  en un nombre fini d'étape.

**I.9.4.1.6. Chaîne de Markov Périodique**

- Un état  $j$  est périodique si on ne peut y revenir qu'après un nombre d'étapes multiple de  $K > 1$
- La période d'une CMTD est égale au PGCD de la période de chacun de ses états
- Une CMTD est périodique si sa période est  $> 1$ , sinon elle est apériodique.

**I.9.4.1.7. Classification des sous-chaînes irréductibles**

- Une sous-chaîne est absorbante si on ne peut pas en sortir
- Sinon une sous-chaîne est transitoire

**I.9.4.1.8. Écriture canonique d'une matrice de transition**

$$P = \left[ \begin{array}{ccc|c} P_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & P_k & 0 \\ \hline R_1 & \dots & R_k & Q \end{array} \right]$$

- $Q$  : Transitions des sous-chaînes transitoires
- $P_i$  : Transitions entre états de la sous-chaîne absorbante  $i$
- $R_i$  : Transitions vers la sous-chaîne absorbante  $i$

**I.9.4.1.9. Temps de séjour dans un état**

- Soit  $T_i$  le temps de séjour dans l'état  $i$  :



$$P[T_i = n] = P_{ii}^{n-1}(1 - P_{ii})$$

- $T_i$  suit donc une loi géométrique

#### ***I.9.4.1.10. Probabilités de transition en $m$ étapes***

- La probabilité d'aller de  $i$  à  $j$  en  $m$  étapes est :

$$P_{ij}^m = P[X_{n+m} = j | X_n = i] = P[X_m = j | X_0 = i]$$

- On note  $P^m = P_{ij}^m$  la matrice de transition en  $m$  étapes

#### **Propriétés :**

-  $P^{(m)} = P^m$

- Équations de Chapman-Kolmogorov:  $P^{(l+m)} = P^{(l)}P^{(m)}$

- Ce qui s'écrit sous forme développée:  $P_{ij}^{(l+m)} = \sum_{k \in E} P_{ik}^l P_{kj}^{(m)}$

#### ***I.9.4.1.11. Probabilité d'aller de $i$ à $j$ en exactement $n$ étapes***

- $f_{ij}^n$  : probabilité, d'aller de  $i$  à  $j$  en exactement  $n$  étapes (sans passer par  $j$  de façon intermédiaire)
- $f_{ij}$  : probabilité d'aller de  $i$  à  $j$  en un nombre quelconque d'étapes

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n) = P_{ij} + \sum_{k \neq j} P_{ik} f_{kj}$$

#### ***I.9.4.1.12. Vecteur des probabilités d'état***

- Soit  $\pi_i(n)$  la probabilité d'être dans l'état  $i$  à l'étape  $n$  est  $\pi_i(n) = P[X_n = i]$
- Soit  $\pi(n) = (\pi_1(n), \pi_2(n), \dots)$  le vecteur des probabilités d'état à l'étape  $n$
- Le vecteur des probabilités  $\pi(n)$  dépend :
  - De la matrice de transition  $P$
  - Du vecteur des probabilités initiales
- Remarque : si l'état initial est connu avec certitude et est égale à  $i$ , on a simplement  $\pi_i(0) = 1$  et  $\pi_j(0) = 0, j \neq i$

#### ***I.9.4.1.12. Équations en régime transitoire***

- D'après la formule des probabilités totales, on a :

$$P[X_{n+1} = j] = P[X_{n+1} = j | X_n = i]P[X_n = i] \Leftrightarrow \pi_j(n+1) = \sum_{i \in E} P_{ij}\pi_i(n)$$

- Propriétés: soit  $P$  la matrice de transition d'une chaîne de Markov et  $\pi(0)$  la

distribution initiale. Pour tout  $n \geq 0$ , on a alors :

$$\pi(n+1) = \pi(n)P \text{ et } \pi(n) = \pi(0)P^n$$

#### ***I.9.4.1.14. Régime stationnaire d'une CMTD irréductible et apériodique:***

##### ***Théorème (admis) :***

Pour une CMTD irréductible et apériodique, le vecteur des probabilités  $\pi(0)$  tend vers un vecteur  $\pi$  qui ne dépend pas de  $\pi(0)$  et qui est la solution unique du système :

$$\left\{ \begin{array}{l} \pi = \pi P \\ \sum_{i=0}^{\infty} \pi_i \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \forall j \in E, \pi_j = \sum_{i \in E} \pi_i P_{ij} \\ \sum_{j \in E} \pi_j = 1 \end{array} \right.$$

Les probabilités  $\pi_j$  sont alors appelées probabilités stationnaires

#### ***I.9.4.1.15. Équilibre des flux***

- L'équation  $\pi_j = \sum_{i \in E} \pi_i P_{ij}$  peut aussi s'écrire :  $\sum_{i \in E} \pi_j P_{ji} = \sum_{i \in E} \pi_i P_{ij}$
- $\sum_{i \in E} \pi_j P_{ji}$  est le flux moyen de sortie de l'état  $j$
- $\sum_{i \in E} \pi_i P_{ij}$  est le flux moyen d'entrée dans l'état  $j$
- Interprétation : flux moyen sortant = flux moyen entrant

#### ***I.9.4.2. Chaînes de Markov à Temps Continu (CMTC)***

- Un processus stochastique  $X(t)$ ,  $t \geq 0$  à espace d'état discret et à temps continu est une Chaîne de Markov à Temps Continu (CMTC) si et seulement si :

$$P[X(t+s) = j | X(u), \quad 0 \leq u \leq s] = P[X(t+s) = j | X(s)], \quad \forall s, t, j$$

- Dans une CMTC, le comportement futur ne dépend que de l'état actuel

##### ***I.9.4.2.1. CMTC homogène***

- Une CMTC est homogène si et seulement si :

$$P[X(t+s) = j | X(s) = i] = P[X(t) = j | X(0) = i] = P_{ij}(s), \quad \forall s, t, i, j$$

- On considérera par la suite uniquement des CMTC homogènes

### I.9.4.2.2. Comportement d'une CMTC

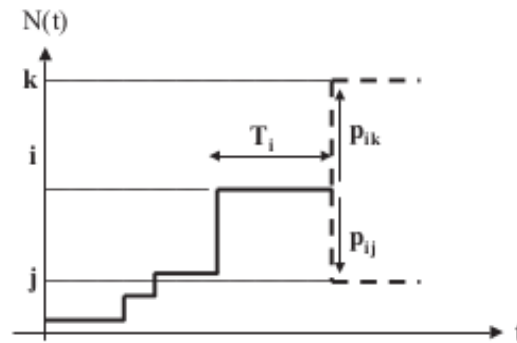


Figure 6: la représentation du comportement d'une CMTC

- $T_i$  = temps passé dans l'état  $i$  (variable aléatoire)
- $P_{ij}$  = probabilité de transition de  $i$  vers  $j$  lorsqu'il y a changement d'état

### I.9.4.2.3. Temps de séjour dans un état

- Notons  $T_i$  la variable aléatoire mesurant le temps passé dans l'état  $i$
- La propriété sans mémoire d'une CMTC homogène implique que :
 
$$P[T_i \geq t + x | T_i \geq t] = P[T_i \geq x], \forall t, \forall x$$
- La variable exponentielle est la seule variable aléatoire continue ayant la propriété précédente
- Le temps de séjour dans un état suit donc une loi exponentielle dans une chaîne de Markov à temps continu.

### I.9.4.2.4. Probabilités de transition

- Lorsque la CMTC quitte l'état  $i$ , elle se déplace dans l'état  $j$  avec une probabilité  $P_{ij}$   
 Cette probabilité est :
  - indépendante du temps car le processus est homogène.
  - indépendante du temps  $T_i$  passé dans l'état  $i$  car le processus est markovien (sans mémoire)

### I.9.4.2.5. 1<sup>ère</sup> Caractérisation d'une CMTC

- On peut entièrement caractériser une CMTC par les paramètres suivants:
  - $\{\mu_i\}_{i \in E}$  avec paramètre  $\mu_i$  de la variable aléatoire exponentielle  $T_i$
  - $\{P_{ij}\}_{i \neq j}$  avec  $P_{ij}$  probabilité de transition de  $i$  vers  $j$  lorsqu'il y a changement

d'état.

#### ***I.9.4.2.6. CMTD Incluse Associée à une CMTC***

- Soit une CMTC caractérisée par des taux de transition  $\{\mu_i\}_{i \in E}$  et par des probabilités de transition  $\{P_{ij}\}_{i,j \in E}$
- On appelle CMTD incluse la CMTD de matrice de transition  $\{P_{ij}\}_{i,j \in E}$

#### ***I.9.4.2.7. Classification d'une CMTC***

- On définit pour une CMTC des états transitoires, récurrents, absorbants de la même façon que pour une CMTD.

##### **Propriétés:**

- Un état  $i$  d'une CMTC est transitoire (respectivement récurrent, absorbant) si et seulement si l'état  $i$  de la CMTD incluse est transitoire (respectivement récurrent, absorbant).
- Une CMTC est irréductible si et seulement si sa CMTD incluse est irréductible

*Remarque : il n'y a pas de notion de périodicité dans une CMTC.*

#### ***I.9.4.2.8. 2<sup>ème</sup> caractérisation d'une CMTC***

- La CMTC se déplace de l'état  $i$  à l'état  $j$  après un temps  $T_{ij}$  variable aléatoire exponentielle de taux  $\mu_{ij}$
- $\mu_{ij}$  est appelé taux de transition de  $i$  vers  $j$

#### ***I.9.4.2.9. Probabilités d'état***

- On définit, comme pour les CMTD, des probabilités d'état définies par :

$$\pi_i(t) = P[X(t) = i]$$

- et un vecteur des probabilités d'état défini par:

$$\pi(t) = (\pi_1(t), \pi_2(t), \dots)$$

#### ***I.9.4.2.10. Probabilités d'état en régime transitoire***

- Le théorème des probabilités totales s'écrit:

$$P[X(t + dt) = j] = \sum_{i \in E} P[X(t + dt) = j | X(t) = i] P[X(t) = i]$$

Avec :

$$\begin{cases} P[X(t+dt) = j] = \pi_j(t+dt) \text{ et } P[X(t) = i] = \pi_i(t) \\ P[X(t+dt) = j|X(t) = i] = P[T_i \leq dt]P_{ij} = (1 - e^{-\mu_i dt})P_{ij} = \mu_i P_{ij} + 0(dt), \forall j \neq i \\ P[X(t+dt) = j|X(t) = j] = P[T_i \geq dt] = e^{-\mu_i dt} = 1 - \mu_j(dt) + 0(dt) \end{cases}$$

- On en déduit que :

$$\begin{aligned} \pi_j(t+dt) &= \sum_{i \neq j} \pi_i(t) (\mu_{ij} + 0(dt)) + [1 - \mu_j(dt) + 0(dt)]\pi_j(t) \Rightarrow \frac{\pi_j(t+dt) - \pi_j(t)}{dt} \\ &= \sum_{i \neq j} \mu_{ij}\pi_i(t) - \mu_j\pi_j(t) + \frac{0(dt)}{dt} \end{aligned}$$

- En faisant tendre  $dt$  vers 0, on obtient :

$$\frac{d\pi_j(t)}{dt} = \mu_j\pi_j(t) + \sum_{i \neq j} -\mu_{ij}\pi_i(t)$$

#### ***1.9.4.2.11. Générateur infinitésimal***

- Posons :  $\begin{cases} q_{ij} = \mu_{ij}, & j \neq i \\ q_{ii} = -\mu_i = \sum_{j \neq i} \mu_{ij} \end{cases}$
- La matrice  $Q = q_{ij}$  est appelée le générateur infinitésimal de la CMTC (Cette matrice ne peut être défini que si l'espace des états est fini)
- On a alors:

$$\frac{d\pi_j(t)}{dt} = \sum_i -q_{ij}\pi_i(t)$$

- Soit, sous forme matriciel :

$$\frac{d\pi(t)}{dt} = \pi(t)Q$$

#### ***1.9.4.2.12. Régime permanent d'une CMTC irréductible***

##### **Propriétés :**

- Dans une CMTC irréductible, le vecteur des probabilités stationnaires  $\pi = (\pi_1, \pi_2, \dots)$  existe toujours et est indépendant de la distribution initiale
- est une solution de :

$$\begin{cases} \pi Q = 0 \\ \sum_{i \in E} \pi_i = 1 \end{cases} \Leftrightarrow \begin{cases} \forall j \in E, \sum_{i \in E} \pi_i q_{ij} = 0 \\ \sum_{j \in E} \pi_j = 1 \end{cases}$$

**I.9.4.2.13. Équilibre des Flux**

- L'équation  $\sum_{i \in E} \pi_i q_{ij} = 0$  peut aussi s'écrire :  $\sum_{i \neq j} \pi_j \mu_{ji} = \sum_{i \neq j} \pi_i \mu_{ij}$
- Avec :  $\begin{cases} \sum_{i \neq j} \pi_j \mu_{ji} \rightarrow \text{flux moyen sortant de l'état } j \\ \sum_{i \neq j} \pi_i \mu_{ij} \rightarrow \text{flux moyen entrant dans l'état } j \end{cases}$
- Interprétation : flux moyen sortant = flux moyen entrant

**I.9.4.2.14. Notion d'ergodicité**

- « Un système est ergodique si les performances stationnaires du système (résultant de l'analyse stochastique) sont égales aux performances de n'importe quelle réalisation particulière du système, observée sur une période suffisamment longue »
- Exemple : Les CMTC irréductibles, apériodiques et finies sont ergodiques.
- Sauf mention contraire, tous les systèmes que nous considérerons seront ergodiques

**I.10. Utilisation de processus de naissance et de mort****Définition :**

On appelle processus de naissance et mort une chaîne de Markov homogène à temps continu  $(X_t)_{t \geq 0}$  à valeurs dans  $N$  telle que :

$$\forall i, j \in N, \forall h > 0$$

$$P_{ij}(h) = P[X_{t+h} = j | X_t = i] = \begin{cases} \lambda_i(h) + 0(h) & \text{si } j = i + 1 \\ \mu_i(h) + 0(h) & \text{si } j = i - 1 \\ 1 - (\lambda_i + \mu_i) + 0(h) & \text{si } j = i \\ 0(h) & \text{sinon} \end{cases}$$

Autrement dit, le générateur d'un processus de naissance et mort de taux de naissance  $\lambda_i, i \in N$  et  $\mu_i, i \in N$  respectivement le taux de naissance et de taux de mort est

donné par une matrice tri-diagonale  $G$  :

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots & \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 \\ \vdots & 0 & & \ddots & \ddots \\ & \ddots & & \ddots & \ddots \end{pmatrix}$$

- Si tous les  $\lambda_i$  (taux de croissance) sont nuls, on parle de processus de mort.
- Si tous les  $\mu_i$  (taux de décroissance) sont nuls, on parle de processus de naissance.
- $\forall i \in N, \lambda_i = \lambda_i + a$  avec  $\lambda > 0$  et  $a > 0$  (processus avec immigration) ou  $a=0$  (processus sans immigration).
- $\forall i \in N^*, \mu_i = \mu_i + b$  avec  $\mu > 0$  et  $b > 0$  (processus avec émigration) ou  $b=0$  (processus sans émigration).

### ***I.11. Conclusion***

Nous avons vu dans ce chapitre que des concepts de base qui sont utiles dans la théorie de files d'attente car dans le deuxième chapitre, on verra l'importance de processus de naissance et de mort dans la modélisation des files d'attente markoviennes ainsi que le processus de Poisson dans l'extraction des équations de Chapman-Kolmogorov or les équations de balance et dans le troisième chapitre, on verra l'utilité des chaînes de Markov et leurs caractéristiques car la résolution de problème des systèmes non markoviens demande des chaînes de Markov induites.

# CHAPITRE II

SYSTEMES D'ATTENTE MARKOVIENS



# CHAPITRE II

## SYSTEMES D'ATTENTE MARKOVIENS

---

### *II.1. Introduction*

Nous abordons maintenant l'analyse proprement dite des files d'attente. Cette analyse va consister à chercher le régime permanent ou stationnaire du système, c'est-à-dire le régime, lorsqu'il existe, vers lequel le système tend naturellement après un certain temps d'utilisation. La détermination du régime permanent permet ensuite de calculer un certain nombre de mesures de performance du système dans ce régime. L'analyse repose essentiellement sur la théorie des chaînes de Markov en temps continu, et notamment son application aux processus de naissance et de mort.

Nous allons voir dans ce chapitre les techniques les plus classiques de résolution des files d'attente simples, c'est-à-dire des files d'attente prises individuellement, et présentant un caractère markovien, au moins sur l'un des deux processus des arrivées ou de service. Nous commençons par les files les plus simples, les files dites markoviennes, dont le processus du nombre de clients dans le système à l'instant  $t$ ,  $\{X(t), t \geq 0\}$  est un processus de naissance et de mort.

### *II.2. La file d'attente M/M/1*

Soit une file d'attente ayant le comportement suivant :

- Les clients arrivent selon un processus de Poisson de taux  $\lambda$
- chaque client reçoit une quantité de service dont la distribution est exponentielle de paramètre  $m$  ; en particulier, la durée moyenne d'un service vaut  $1/m$ .
- il y a un serveur ;
- les clients sont servis un par un dans l'ordre de leur arrivée ;
- la capacité de la file d'attente est infinie : il y a toujours de la place pour les clients en attente.
- Selon la notation de Kendall cette file est connue sous l'abréviation

–  $M/M/1$ .

Le système d'attente  $M/M/1$  est le modèle le plus simple dans toute la théorie des files d'attente ainsi que dans la pratique.  $N(t)$  qui arrivent durant un intervalle de temps  $(t, t + \Delta t)$  suit la distribution de Poisson:

$$P[N(t) = j] = e^{-\lambda t} \frac{(\lambda t)^j}{j!}, \quad j = 0, 1, 2, \dots$$

C'est à dire les temps des inter-arrivées suivent la distribution exponentielle avec la densité de probabilité :

$$a(x) = \lambda e^{-\lambda x}, \quad x > 0$$

On suppose que le temps de service suit une distribution exponentielle avec la densité de probabilité :

$$b(x) = \mu e^{-\mu x}, \quad x > 0$$

Avec ces suppositions, on a:

$$E[\text{inter-arrivées}] = 1/\lambda = \frac{1}{\text{taux des arrivées}}$$

$$E[\text{temps de service}] = 1/\mu = \frac{1}{\text{taux de service}}$$

La proportion de taux des arrivées sur le taux de service joue un rôle significatif dans la mesure des performances dans la théorie de files d'attente. On a :

$$\rho = \text{intensité de trafic} = \lambda/\mu = \frac{\text{taux des arrivées}}{\text{taux de service}}$$

Dans la file d'attente  $M/M/1$ ,

Ce système est modélisé par un processus de naissance et de mort tel que :

$$\begin{cases} \lambda_n = \lambda \\ \mu_n = \mu \end{cases}$$

La matrice de transition correspondante  $A$  adéquate à cette file est donnée :

$$\mathbf{A} = \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & \mu & -(\lambda + \mu) & \lambda & \\ & & & \ddots & \ddots \\ & & & & \ddots \end{bmatrix}.$$

$$(2.1)$$

Les équations de Kolmogorov correspondantes pour  $\{P_n(t), n = 0, 1, 2, \dots\}$  sont :

$$\begin{aligned} P_0'(t) &= \lambda P_0(t) + \mu P_1(t) \\ P_n'(t) &= -(\lambda + \mu)P_n(t) + \lambda P_{n-1}(t) + \mu P_{n+1}(t), \quad n = 1, 2, \dots \end{aligned} \quad (2.2)$$

Pour la solution complète de différentes équations différentielles, on a besoin d'utiliser de la fonction génératrice et la transformée de Laplace. D'où  $P_n(t)$  est obtenue en utilisant des formules inverses.

### *La distribution limite*

On a la probabilité limite,  $\lim_{n \rightarrow \infty} P_n(t) = P_n$  l'état des équations de balance :

$$\begin{aligned} \lambda P_0(t) &= \mu P_1(t) \\ (\lambda + \mu)P_n(t) &= \lambda P_{n-1}(t) + \mu P_{n+1}(t), \quad n = 1, 2, \dots \end{aligned} \quad (2.3)$$

La résolution de ces équations avec la condition :  $\sum_{n=0}^{\infty} P_n = 1$ , nous donne :

$$P_n = (1 - \rho)\rho^n, \quad n = 0, 1, 2, \dots \quad (2.4)$$

Où

La probabilité que le serveur est occupé est qui est une mesure de performance du système. En effet, ce facteur d'utilisation = l'intensité du trafic =  $1 - P_0 = \rho$ .

On a défini  $Q(t)$  est le nombre des clients dans le système.

Écrire  $Q(\infty) = Q$  et, on a :  $Q_n$  est le nombre des clients dans la file sans compter un dans le service.

Maintenant, on peut avoir deux valeurs moyennes qui sont appelées des caractéristiques statistiques ou des mesures de performances à savoir :

- $L = E(Q)$ : le nombre moyen des clients dans le système ;
- $L_q = E(Q_q)$  : le nombre moyen des clients dans la file.

D'après la distribution de  $P_n$  dans la formule (2.4), on a :

$$L = \sum_{n=1}^{\infty} n(1 - \rho)\rho^n = \frac{\rho}{1 - \rho}$$

Sera écrit comme suit:

$$L = \frac{\lambda}{\mu - \lambda} \quad (2.5)$$

Pour  $L_q$ , on a :

$$L_q = \sum_{n=1}^{\infty} (n-1)P_n = L - \rho = \frac{\rho^2}{1-\rho} = \frac{\lambda^2}{\mu(\mu-\lambda)} \quad (2.6)$$

A partir de la formule (2.4), on obtient la variance des clients dans le système comme suit :

$$V(Q) = \frac{\rho}{(1-\rho)^2} = \frac{\mu\lambda}{(\mu-\lambda)^2} \quad (2.7)$$

### *Le temps d'attente d'un client*

- $T$  : variable aléatoire qui représente le temps total dans le système ;
- $T_q$  : variable aléatoire qui représente le temps passé dans la file.

Quand on a :  $n$  clients dans le système, le temps de service est exponentiellement distribué avec le paramètre  $\mu$ , le temps total de service de  $n$  clients suit la loi d'Erlang avec la densité de probabilité:

$$f_n = e^{-\mu x} \frac{(\mu x)^{(n-1)}}{(n-1)!} \quad (2.8)$$

On a :  $F_q(t) = P(T_q \leq t)$ , la fonction de répartition de temps d'attente  $T_q$ .

En outre,

$$F_q(0) = P(T_q = 0) = P(Q = 0) = 1 - \rho \quad (2.9)$$

Le temps de service d'un client est exponentiel avec le même paramètre  $\mu$ .

$$dF_q(t) = P(t < T_q \leq t + dt)$$

$$dF_q = \sum_{n=1}^{\infty} P_n e^{-\mu t} \frac{([\mu t])^{(n-1)}}{(n-1)!} dt = (1 - \rho) \sum_{n=1}^{\infty} \rho^n e^{-\mu t} \frac{([\mu t])^{(n-1)}}{(n-1)!} dt$$

Pour  $t > 0$ , on a :

Cela avec la simplification nous donne:

$$dF_n = \lambda(1 - \rho)e^{-\mu(1-\rho)t} \quad (2.10)$$

A cause de la discontinuité de la distribution de  $T_q$  en 0, on aura:

$$F_q(t) = P(T_t = 0) + \int_0^t dF_q(t) = 1 - \rho e^{-\mu(1-\rho)t} \quad (2.11)$$

Où on a combiné des résultats de (2.9) et (2.10). On a :  $E(T_q) = W_q$  et  $E(T) = W$ .

A partir de la formule (2.11), on peut facilement avoir:

$$E(T_q) = W_q = \frac{\rho}{\mu(1-\rho)} = \frac{\lambda}{\mu(\mu-\lambda)} \quad (2.12)$$

Ainsi que,

$$V(Q) = \frac{\rho(2-\rho)}{\mu^2(1-\rho)^2} \quad (2.13)$$

Rappelant que le temps total dans le système  $T$  est la somme de  $T_q$  et le temps de service, on aura:

$$W = E(T) = \frac{\lambda}{\mu(\mu-\lambda)} + 1/\mu = \frac{1}{(\mu-\lambda)} \quad (2.14)$$

Comparant le résultat (4.2.14) avec (4.2.5), on note la relation suivante:

$$L = \lambda W \quad (2.15)$$

Avec la même comparaison entre les résultats (2.6) et (2.12), on établie la relation suivante:

$$L_q = \lambda W_q \quad (2.16)$$

### ***II.3. La file d'attente à une source finie $M/M/1/K$***

Soit une file d'attente ayant le comportement suivant :

- les clients arrivent selon un processus de Poisson de paramètre  $\lambda > 0$
- chaque client reçoit une quantité de service dont la distribution est exponentielle de paramètre  $\mu > 0$
- les clients sont servis un par un dans l'ordre de leur arrivée,
- il y a au plus  $K$  clients en attente ; si un nouveau client se présente alors que la file d'attente est pleine, il est rejeté.
- Soit  $N(t)$  le nombre de clients dans la file à l'instant  $t$ .

La matrice des probabilités de transition  $Q$  associée à ce modèle est :

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

Le fonctionnement est donc le même que la file  $M/M/1$ , sauf en ce qui concerne le rejet des clients quand la longueur de la file d'attente atteint  $K$ . Dans la notation de Kendall, c'est la file d'attente  $M/M/1/K$ .

Les résultats seront donnés directement par la spécialisations  $s = 1$ ,  $\alpha = \rho = \mu/\lambda$ , dans le cas du système d'attente  $M/M/s/K$ .

$$P_0 = \begin{cases} \frac{1-\rho}{1-\rho^{(K+1)}}, & \rho \neq 1 \\ \frac{1}{K+1}, & \rho = 1 \end{cases} \quad (2.17)$$

$$P_n = \begin{cases} \frac{(1-\rho)\rho^n}{1-\rho^{(K+1)}}, & \rho \neq 1 \\ \frac{1}{K+1}, & \rho = 1 \end{cases} \quad (2.18)$$

Et,

$$1 - P_K = \begin{cases} \frac{(1-\rho^K)}{1-\rho^{(K+1)}}, & \rho \neq 1 \\ \frac{K}{K+1}, & \rho = 1 \end{cases}$$

$$F_q(t) = \begin{cases} 1 - \frac{1-\rho}{1-\rho^K} \sum_{n=1}^{K-1} \rho^n \sum_{r=1}^{n-1} e^{-\mu t} \frac{(\mu t)^r}{r!}, & \rho \neq 1 \\ 1 - \frac{1}{K} \sum_{n=1}^{K-1} \sum_{r=1}^{n-1} e^{-\mu t} \frac{(\mu t)^r}{r!}, & \rho = 1 \end{cases} \quad (2.19)$$

$$W_q = \begin{cases} \frac{1}{\mu} \left[ \frac{\rho}{1-\rho} - \frac{K\rho^K}{1-\rho^K} \right], & \rho \neq 1 \\ \frac{1}{2\mu} (K-1), & \rho = 1 \end{cases} \quad (2.20)$$

$$W = \begin{cases} \frac{1}{\mu} \left[ \frac{\rho}{1-\rho} - \frac{K\rho^K}{1-\rho^K} \right], & \rho \neq 1 \\ \frac{1}{2\mu} (K+1), & \rho = 1 \end{cases} \quad (2.21)$$

$$L_q = \begin{cases} \frac{\rho}{1-\rho} - \frac{\rho(1+K\rho^K)}{1-\rho^{(K+1)}}, & \rho \neq 1 \\ \frac{K(K-1)}{2(K+1)}, & \rho = 1 \end{cases} \quad (2.22)$$

$$L = \begin{cases} \frac{\rho(1-\rho^K)}{(1-\rho)(1-\rho^{(K+1)})} - \frac{(K\rho^K)}{1-\rho^{(K+1)}}, & \rho \neq 1 \\ \frac{K}{2}, & \rho = 1 \end{cases} \quad (2.23)$$

On note que les simplifications montrent quelques résultats donnés, on a utilisé cette formule :

$$\sum_{n=1}^{K-1} n\rho^{(n-1)} = \frac{d}{d\rho} \left( \frac{1-\rho^K}{1-\rho} \right)$$

#### II.4. La file d'attente $M/M/1/1$

Soit une file d'attente ayant le comportement suivant :

- les clients arrivent selon un processus de Poisson de paramètre  $\lambda > 0$
- chaque client reçoit une quantité de service dont la distribution est exponentielle de paramètre  $\mu > 0$
- les clients sont servis un par un dans l'ordre de leur arrivée,
- il y a au plus un client en attente ; si un nouveau client se présente alors que la file d'attente est pleine, il est rejeté.
- $\{N(t)\}$  le nombre de clients dans la file à l'instant  $t$ .

Soit  $Q(t)$  est le nombre des clients dans le système à l'instant  $t$  et

$$\lim_{t \rightarrow \infty} Q(t) = Q$$

La variable aléatoire  $Q$  peut contenir deux valeurs (0, 1),

En plus, on a :

$$P(Q = n) = P_n, \quad n = 0, 1$$

En effet,  $Q(t)$ ,  $t \in T$  est un processus de Markov avec la matrice de transition  $A$  correspondante :

$$A = \begin{matrix} 0 \\ 1 \end{matrix} \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}.$$

Avec les probabilités de transitions  $P_{ij} = P[Q(t) = j | Q(0) = i]$ , ( $i, j = 0, 1$ ), on aura des équations de Kolmogorov qui s'écrivent de cette manière:

$$\begin{aligned} P'_{i0}(t) &= -\lambda P'_{i0}(t) + \mu P'_{i1}(t) \\ P'_{i1}(t) &= -\mu P'_{i1}(t) + \lambda P'_{i0}(t) \end{aligned} \quad (2.24)$$

Si on note que  $P'_{i0}(t) + P'_{i1}(t) = 1$ , les deux équations dans le système d'équations (2.24) donnent une seule équation différentielle linéaire d'ordre 1:

$$P'_{i0}(t) = \mu - (\mu + \lambda)P_{i0}(t) \quad (2.25)$$

Utilisation de la condition initiale:

$$P_{i0}(0) = \begin{cases} 1 & \text{si } i = 0 \\ 0 & \text{si } i = 1 \end{cases}$$

L'équation (2.25) est résolue à travers des techniques standards pour donner:

$$\begin{aligned} P_{00}(t) &= \frac{\mu}{\mu + \lambda} + \frac{\lambda}{\mu + \lambda} e^{-(\mu + \lambda)t} \\ P_{10}(t) &= \frac{\mu}{\mu + \lambda} - \frac{\lambda}{\mu + \lambda} e^{-(\mu + \lambda)t} \end{aligned} \quad (2.26)$$

On a, aussi:

$$\begin{aligned} P_{01}(t) &= 1 - P_{00}(t) = \frac{\lambda}{\mu + \lambda} - \frac{\lambda}{\mu + \lambda} e^{-(\mu + \lambda)t} \\ P_{11}(t) &= 1 - P_{10}(t) = \frac{\lambda}{\mu + \lambda} + \frac{\mu}{\mu + \lambda} e^{-(\mu + \lambda)t} \end{aligned}$$

Les probabilités limites  $\{P_n, n = 0, 1\}$  peuvent être déterminées si on fait tendre  $t \rightarrow \infty$  dans la formule (2.26) ou bien avec la résolution de l'équation de balance :

$$\lambda P_0 = P_1 \mu \quad (2.27)$$

Et avec la normalisation de la condition  $P_0 + P_1 = 1$ , on aura:

$$P_0 = \frac{\mu}{\mu + \lambda}, \quad P_1 = \frac{\lambda}{\mu + \lambda} \quad (2.28)$$

Ces probabilités peuvent avoir une expression en termes de la moyenne des périodes libres et occupées. Divisant le numérateur et le dénominateur des expressions pour  $P_0$  et  $P_1$  sur  $(\lambda\mu)$ .

On aura :



$$P_0 = \frac{1/\lambda}{1/\lambda + 1/\mu}, \quad P_1 = \frac{1/\mu}{1/\lambda + 1/\mu} \quad (2.28)$$

On note  $1/\lambda$  est la moyenne des périodes où le serveur est libre (inoccupé) et  $1/\mu$  est la moyenne des périodes où le serveur est occupé. La généralisation de ce concept pour un processus à deux états 0 et 1 représentée par deux variables aléatoires indépendantes  $X$  et  $Y$  et les probabilités que le processus génère dans ces deux états 0 et 1 sont données:

$$P_0 = \frac{E(X)}{E(X)+E(Y)}, \quad P_1 = \frac{E(Y)}{E(X)+E(Y)} \quad (2.29)$$

### II.5. La file d'attente $M/M/s$

Par hypothèse :

- les clients arrivent selon un processus de Poisson de paramètre  $\lambda > 0$
- le temps de service est une loi exponentielle de paramètre  $\mu > 0$
- il y a  $s$  serveurs montés en parallèle,
- la file d'attente peut accueillir un nombre quelconque de clients.

Ce système est modélisé par un processus de naissance et de mort tel que :

$$\begin{cases} \lambda_n = \lambda, & n = 0, 1, 2, \dots \\ \mu_n = \begin{cases} n\mu, & 0 < n \leq s \\ s\mu, & n \geq s \end{cases} \end{cases} \quad (2.30)$$

La matrice de transition  $\mathbf{A}$  est donnée comme suit:

$$\mathbf{A} = \begin{matrix} \begin{matrix} \rightarrow 0 \\ 1 \\ \vdots \\ s \\ s+1 \\ \vdots \end{matrix} & \begin{bmatrix} -\lambda & \lambda & & & & & \\ \mu & -(\lambda + \mu) & \lambda & & & & \\ & & \cdot & \cdot & & & \\ & & & & s\mu & -(\lambda + s\mu) & \lambda \\ & & & & & s\mu & -(\lambda + s\mu) & \lambda \\ & & & & & & \ddots & \end{bmatrix} \end{matrix} \quad (2.31)$$

Les équations de Kolmogorov de  $P_n(t)$  s'écrivent comme suit avec la probabilité limite :

$$P_n = \lim_{t \rightarrow \infty} P_n(t)$$

On a :

$$\begin{aligned} \lambda P_0 &= \mu P_1 \\ (\lambda + n\mu)P_n &= \lambda P_{n-1} + (n+1)\mu P_{n+1}, \quad 0 < n < s \\ (\lambda + s\mu)P_n &= \lambda P_{n-1} + s\mu P_{n+1}, \quad s \leq n < \infty \end{aligned} \quad (2.32)$$

Une procédure récursive dans les lignes utilisées dans le cas de  $M/M/1$  donne la solution suivante:

$$\begin{aligned} n\mu P_n &= \lambda P_{n-1}, \quad n = 1, 2, \dots, s \\ s\mu P_n &= \lambda P_{n-1}, \quad n = s+1, s+2, \dots \end{aligned}$$

Par conséquent,

$$P_n = \begin{cases} \frac{1}{n!} (\lambda/\mu)^n P_0, & 0 \leq n \leq s \\ (\lambda/s\mu)^{n-s} P_s, & n = s, s+1, \dots \end{cases} \quad (2.33)$$

$$P_{s+r} = (\lambda/s\mu)^r P_s, \quad r = 0, 1, 2, \dots$$

Écrire  $\rho = \lambda/s\mu$  et avec les simplifications, on aura :

$$P_n = \begin{cases} \frac{1}{n!} (\rho)^n P_0, & 0 \leq n \leq s \\ \frac{1}{s!} (n\rho)^s (\rho)^{n-s} P_0, & s \leq n < \infty \end{cases} \quad (2.34)$$

L'utilisation de la condition  $\sum_{n=0}^{\infty} P_n = 1$  et la formule (2.34) nous donnent :

$$\begin{aligned} P_0 &= \left[ \sum_{r=0}^{s-1} \frac{(s\rho)^r}{r!} + \frac{(s\rho)^s}{s! (1-\rho)} \right]^{-1} \\ P_n &= \begin{cases} \frac{(n\rho)^n}{n!} P_0, & 0 \leq n \leq s \\ \frac{s^s \rho^n}{s!} P_0, & s \leq n < \infty \end{cases} \end{aligned} \quad (2.35)$$

A condition  $\rho = \lambda/s\mu$ . Avec  $s\mu$  est le taux de service maximal et on considère est l'intensité du trafic pour le système. Écrire la dernière équation dans la formule (2.32) tel que:

$$P_s = \rho^{n-s} P_s, \quad n \geq s \quad (2.36)$$

On peut dire que le nombre des clients dans le système est  $\geq s$ , le système apparaîtra comme le  $M/M/1$  avec un taux de service  $s$ . On peut aussi écrire  $\alpha = \lambda/\mu$  et  $\rho = \alpha/s$ . En remplaçant dans la formule (2.34) avec l'utilisation de  $\alpha$ , on aura:

$$P_0 = \left[ \sum_{r=0}^{s-1} \frac{(\alpha)^r}{r!} + \frac{(\alpha)^s}{s!} \left(1 - \frac{\alpha}{s}\right)^{-1} \right]^{-1}$$

$$P_n = \begin{cases} \frac{(\alpha)^n}{n!} P_0, & 0 \leq n \leq s \\ \frac{\alpha^s}{s!} \left(\frac{\alpha}{s}\right)^{n-s} P_0, & s \leq n < \infty \end{cases} \quad (2.37)$$

On note que les clients vont attendre pour le service seulement si le nombre de clients dans le système est  $\geq s$ . La probabilité de cet événement est donnée par :  $\sum_{n=s}^{\infty} P_n$ .

Et c'est pour cette raison que :

$$P(\text{délai d'un client}) = C(s, \alpha) = P_0 = \frac{\alpha^s}{s!} (1 - \frac{\alpha}{s})^{-1} [\sum_{r=0}^{s-1} \frac{(\alpha)^r}{r!} + \frac{(\alpha)^s}{s!} (1 - \frac{\alpha}{s})^{-1}]^{-1} \quad (2.38)$$

La formule  $C(s, \alpha)$  est connue dans la littérature comme la formule de délai d'Erlang ou bien la seconde formule d'Erlang, et elle est notée  $E_{2,s}(\alpha)$ .

Écrire  $L$  et  $L_q$  comme nombre moyen de clients dans le système et dans la queue respectivement, on peut les développer de cette manière en utilisant des expressions de la formule (2.34),

On aura avec  $s\rho = \alpha$ :

$$\begin{aligned} \sum_{n=1}^{\infty} n P_n &= P_0 \left[ \sum_{n=1}^s n \frac{\alpha^n}{n!} + \sum_{n=s+1}^{\infty} n \rho^{n-s} \frac{\alpha^s}{s!} \right] = P_0 \left[ \sum_{n=1}^s \alpha \frac{\alpha^{n-1}}{(n-1)!} + \frac{\alpha^s}{s!} \sum_{n=s+1}^{\infty} n \rho^{n-s} \right] \\ &= P_0 \left[ \sum_{r=0}^s \alpha \frac{\alpha^r}{r!} + \frac{\alpha^s}{s!} \sum_{r=1}^{\infty} (r+s) \rho^r \right] = P_0 \left[ \sum_{r=0}^s \alpha \frac{\alpha^r}{r!} + \frac{\alpha^s}{s!} \left( \frac{\rho}{(1-\rho)^2} + \frac{s\rho}{1-\rho} \right) \right] \\ &= \frac{\rho \alpha^s P_0}{s! (1-\rho)^2} + \left[ \sum_{r=0}^s \alpha P_0 \frac{\alpha^r}{r!} + \frac{\alpha^s}{s!} (1-\rho) \right] \end{aligned}$$

Par conséquent,

$$L = \alpha + \frac{\rho \alpha^s P_0}{s! (1-\rho)^2} \quad (2.39)$$

D'où on peut aussi écrire:

$$L = \alpha + \frac{\rho P_s}{(1-\rho)^2} \quad (2.40)$$

Pour avoir  $L_q$ , on écrit:

$$L_s = \sum_{n=s+1}^{\infty} (n-s) \frac{\alpha^s}{s!} \rho^{n-s} P_0 = \frac{\alpha^s}{s!} P_0 \sum_{n=s+1}^{\infty} (n-s) \rho^{n-s} = \frac{\alpha^s}{s!} P_0 \sum_{r=1}^{\infty} r \rho^r = \frac{\rho \alpha^s P_0}{s! (1-\rho)^2} \quad (2.41)$$

D'où on peut aussi écrire:

$$L = \alpha + \frac{\rho P_s}{(1-\rho)^2} \quad (2.42)$$

L'expression de la variance de nombre de clients dans le système est longue, et c'est pour cela qu'on ne peut pas la présenter ici. Comparant les expressions de  $L$  et  $L_q$ . Cela aussi peut être déterminé comme la contribution du facteur d'utilisation correspondant à  $s$ . Par exemple, on peut écrire:

$$\text{utilisation d'un serveur individuel} = \sum_{n=1}^{s-1} \frac{n}{s} P_n + \sum_{n=s}^{\infty} P_n \quad (2.43)$$

Utilisation des expressions pour  $P_n$  de la formule (2.34) dans la formule (2.43) et avec les simplifications, on trouve le facteur d'utilisation d'un serveur individuel est, c.-à-d., au long de déroulement, la probabilité qu'un serveur qu'il soit occupé est  $\rho$ .

### Le temps d'attente:

Or la discussion du temps d'attente des clients, on suppose qu'ils sont servis avec la discipline FCFS. Quand le nombre des clients dans le système est  $s$ , les durées des inter-départs sont exponentielles avec un taux  $s$ . Soit  $T_q$  est le temps d'attente d'un client dans la file, on aura:

$$F_q(t) = P[T_q \leq t]$$

En effet,

$$F_q(t) = P[T_q = 0] = P(Q < s) = \sum_{n=0}^{s-1} P_n = P_0 \sum_{n=0}^{s-1} \frac{\alpha^n}{n!}$$

De la première équation dans la formule (2.37), on a :

$$\sum_{n=0}^{s-1} \frac{\alpha^n}{n!} = \frac{1}{P_0} - \frac{\alpha^s}{s!} (1-\rho)^{-1}$$

Nous donne :

$$F_q(0) = 1 - \frac{\alpha^s P_0}{s!(1-\rho)} \quad (2.44)$$

Ainsi, les arguments suivants trouvés dans la formule (2.10) pour le système  $M/M/1$ , dans le cas de multiple-serveurs, on a (l'utilisation de (2.35) avec les simplifications) :

$$dF_q(t) = \sum_{n=s}^{\infty} P_n e^{-s\mu t} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt = P_s e^{-s\mu t} \sum_{n=s}^{\infty} \rho^{n-s} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt = s\mu P_s e^{-s\mu(1-\rho)t} dt \quad (2.45)$$

D'où on peut aussi écrire :

$$dF_q(q) = \frac{s\mu}{s!} P_0 e^{-s\mu(1-\rho)t} dt \quad (2.46)$$

A partir de (2.44), on a :

$$W_q = \int_0^\infty t dF_q(t) = \int_0^\infty s\mu P_s e^{-s\mu(1-\rho)t} dt = \frac{P_s}{s\mu(1-\rho)^2} \quad (2.47)$$

L'utilisation de  $P_0$  à la place de  $P_s$ , on peut aussi écrire:

$$W_q = \frac{\alpha^s P_0}{s! s\mu(1-\rho)^2} \quad (2.48)$$

La Comparaison de la formule (2.42) avec (2.47) (or (2.41) avec (2.48)), on peut encore vérifier la formule de Little  $L_q = \lambda W_q$ . La fonction de la distribution du temps d'attente peut maintenant être obtenue de (2.44) et (2.46):

$$\begin{aligned} F_q(t) &= F_q(0) + \int_0^t \frac{s\mu\alpha^s}{s!} P_0 e^{-s\mu(1-\rho)x} dx \\ &= 1 - \frac{\alpha^s P_0}{s!(1-\rho)} + \frac{\alpha^s P_0}{s!(1-\rho)} \int_0^t s\mu(1-\rho) e^{-s\mu(1-\rho)x} dx \\ &= 1 - \frac{\alpha^s P_0}{s!(1-\rho)} e^{-s\mu(1-\rho)t} \quad (2.49) \end{aligned}$$

## II.6. La file d'attente à source finie M/M/s/K

Soit une file d'attente ayant le comportement suivant :

- les clients arrivent selon un processus de Poisson de paramètre  $\lambda > 0$
- chaque client reçoit une quantité de service dont la distribution est exponentielle de paramètre  $\mu > 0$
- les clients sont servis un par un dans l'ordre de leur arrivée,
- il y a s serveurs montés en parallèle,
- il y a au plus K clients en attente ; si un nouveau client se présente alors que la file d'attente est pleine, il est rejeté.
- Soit  $N(t)$  le nombre de clients dans la file à l'instant  $t$ .

Ce système est modélisé par un processus de naissance et de mort tel que :

$$\begin{cases} \lambda_n = \lambda, & n = 0, 1, 2, \dots, K-1 \\ \mu_n = \begin{cases} n\mu, & n = 1, 2, \dots, s-1 \\ s\mu, & n = s, s+1, \dots, K \end{cases} \end{cases} \quad (2.50)$$

La matrice des probabilités de transition  $\mathbf{A}$  appropriée à ce modèle est :

$$\mathbf{A} = \begin{matrix} 0 \\ 1 \\ \vdots \\ \vdots \\ K-1 \\ K \end{matrix} \begin{bmatrix} -\lambda & \lambda & & & & \\ \mu & -(\lambda + \mu) & \lambda & & & \\ & \ddots & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & s\mu & -(\mu + s\mu) & \lambda \\ & & & & s\mu & & -s\mu \end{bmatrix}.$$

(2.51)

Pour les probabilités limites  $\{P_n, n = 0, 1, 2, \dots, K\}$  les équations de balance sont écrites de la même façon que celles des formules (2.32). La solution correspondante de (2.35) est donnée comme suit:

$$P_n = \begin{cases} \frac{1}{n!} (\lambda/\mu)^n P_0, & 0 \leq n \leq s \\ \frac{1}{s!} (\lambda/\mu)^s (\lambda/s\mu)^{n-s} P_0, & s \leq n \leq K \end{cases}$$

Écrire  $\rho = \lambda/s\mu$  et  $\alpha = \lambda/\mu$ ,  $P_0$  peut être obtenue en utilisant la condition :

$$\sum_{n=0}^{\infty} P_n = 1$$

On aura:

$$P_0 = \left[ \sum_{r=0}^{s-1} \frac{\alpha^r}{r!} + \frac{\alpha^s}{s!} \sum_{r=s}^K \rho^{n-s} \right]^{-1}$$

Avec la seconde somme de l'expression  $P_0$  qui est une somme finie, On a pas besoin d'imposer la condition  $\rho < 1$  pour la solution avec  $P_0 > 0$  et par conséquent, on a:

$$P_0 = \begin{cases} \left[ \sum_{r=0}^{s-1} \frac{\alpha^r}{r!} + \frac{\alpha^s}{s!} \frac{1 - \rho^{K-s+1}}{1 - \rho} \right]^{-1}, & \rho \neq 1 \\ \left[ \sum_{r=0}^{s-1} \frac{\alpha^r}{r!} + \frac{\alpha^s}{s!} (K - s + 1) \right]^{-1}, & \rho = 1 \end{cases}$$

$$P_n = \begin{cases} \frac{\alpha^n}{n!} P_0, & 0 \leq n \leq s \\ \frac{\alpha^s}{s!} \rho^{n-s} P_0, & s \leq n \leq K \end{cases} \quad (2.52)$$

A cause de difficulté de la nature des expressions de nombre moyen de clients dans le système  $L$  et dans la file  $L_q$ , qu'on ne peut pas les présenter ici. La procédure de les avoir, on commence avec la distribution limite donnée par la formule (2.52).

Dans la discussion des caractéristiques du temps d'attente des clients dans une file finie, on a besoin d'inclure une possibilité d'une arrivée d'un client qui rentre dans le

système.

Quand le système est à l'état stationnaire, la probabilité qu'un client qui arrive et ne rentre pas dans le système est :  $P_K$ . Et maintenant, quand il y a  $n$  clients dans le système et  $n < K$ , la probabilité que l'arrivée d'un client et rentre dans le système est donnée par :  $P_n / (1 - P_K)$ .

Par conséquent, avec la notation utilisée avant pour la distribution de temps d'attente, on a :

$$F_q(t) = F_q(0) + P(0 < W \leq t)$$

Où,

$$F_q(0) = \sum_{n=0}^{s-1} \frac{P_n}{1 - P_K}$$

Ainsi,

$$dF_q(t) = \sum_{n=s}^{K-1} \frac{P_n}{1 - P_K} e^{-\mu t} \frac{(s\mu t)^{n-s}}{(n-s)!} s\mu dt \quad (2.53)$$

$$F_q(q) = \begin{cases} F_q(0) + (1 - \rho) \sum_{n=s}^{K-1} \rho^n \int_0^t \frac{1}{1 - P_K} e^{-\mu t} \frac{(s\mu t)^{n-1}}{(n-s)!} s\mu dt \\ F_q(0) + \frac{1}{1 - P_K} \sum_{n=s}^{K-1} \rho^n (1 - \int_t^\infty e^{-\mu t} \frac{(s\mu t)^{n-1}}{(n-s)!} s\mu dt) \end{cases}$$

Avec la simplification de l'expression, on note que :

$$F_q(0) = F_0(0) + \frac{1}{1 - P_K} \sum_{n=s}^{K-1} \rho^n = 1$$

Ainsi,

$$\int_1^\infty e^{-s\mu t} \frac{(s\mu t)^{(n-s)}}{(n-s)!} s\mu dt = \sum_{r=0}^{n-s} e^{-s\mu t} \frac{(s\mu t)^r}{r!}$$

Et on aura :

$$F_q(t) = 1 - \frac{1}{1 - P_K} \sum_{n=s}^{K-1} P_n \sum_{n=s}^{K-1} e^{-s\mu t} \frac{(s\mu t)^r}{r!}$$

On a :

$$W_q = \int_0^\infty t dF_q(t) = \sum_{n=s}^{K-1} \frac{P^n}{1-P_K} \int_0^\infty e^{-\mu t} \frac{(s\mu t)^{(n-1)}}{(n-s)!} s\mu t dt = \frac{1}{s\mu(1-P_K)} \sum_{n=s}^{K-1} (n-s+1) P_n \quad (2.55)$$

Le temps dans le système est obtenu par :

$$W = W_q + 1/\mu \quad (2.56)$$

Le nombre de clients dans la file et dans le système sont obtenus par les deux formules suivantes :

$$L = \lambda(1 - P_K)W \quad (2.57)$$

$$L_q = \lambda(1 - P_K)W_q \quad (2.58)$$

Notant que le taux des arrivées est  $\lambda(1 - P_K)$

### II.7. La file d'attente M/M/s/s

Par hypothèse :

- les clients arrivent selon un processus de Poisson de paramètre  $\lambda > 0$
- le temps de service est une loi exponentielle de paramètre  $\mu > 0$
- il y a s serveurs montés en parallèle,
- il n'y a pas de file d'attente.

Ce système est modélisé par un processus de naissance et de mort tel que :

$$\begin{cases} \lambda_n = \lambda, & 0 \leq n \leq s \\ \mu_n = n\mu, & 0 < n \leq s \end{cases}$$

La matrice de transition correspondante  $\mathbf{A}$  est donnée par :

$$\mathbf{A} = \begin{matrix} 0 \\ 1 \\ \vdots \\ s-1 \\ s \end{matrix} \begin{bmatrix} -\lambda & \lambda & & & & \\ \mu & -(\lambda + \mu) & & & & \\ & \vdots & & & & \\ & & & (s-1)\mu & -[\lambda + (s-1)\mu] & \lambda \\ & & & \mu & s\mu & -s\mu \end{bmatrix}.$$

$$(2.59)$$



Les probabilités limites sont obtenues par l'utilisation des équations de balance,

$$\begin{aligned}\lambda P_0 &= \mu P_1 \\ (\lambda + n\mu)P_n &= \lambda P_{n-1} + (n+1)\mu P_{n+1}, \quad 1 \leq n < s \\ \mu P_s &= \lambda P_{s-1} \quad (2.60)\end{aligned}$$

Écrire  $\alpha = \lambda/\mu$ , la formule (2.60) peut être résolue d'une manière récursive pour avoir :

$$\begin{aligned}P_0 &= [1 + \alpha + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^s}{s!}]^{-1} \\ P_n &= \frac{\alpha^n}{n!}, \quad n = 0, 1, 2, \dots, s \quad (2.61)\end{aligned}$$

Cela nous donne :

$$P_s = \frac{\alpha^s/s!}{[1 + \alpha + \frac{\alpha^2}{2!} + \dots + \frac{\alpha^s}{s!}]^{-1}} \quad (2.62)$$

Dont, elle est la probabilité qu'un client est bloqué. Ainsi,  $P_s$  nous donne le nombre de clients qui seront bloqués d'y rentrer dans le système. La formule (2.62) est connue comme la formule d'Erlang de perte ou la première formule d'Erlang et notée par  $E_{1,s}(\alpha)$  ou  $B(s, \alpha)$

### II.8. La file d'attente $M/M/\infty$

Par hypothèse :

- les clients arrivent selon un processus de Poisson de paramètre  $\lambda > 0$
- le temps de service est une loi exponentielle de paramètre  $\mu > 0$
- il y a une infinité de serveurs.

Il est évident qu'aucune file d'attente ne se forme puisque chaque client est servi dès son arrivée. Ce système a un intérêt théorique car il permet une étude approximative d'un système d'attente comportant un grand nombre de serveurs.

Ce système est modélisé par un processus de naissance et de mort tel que :

$$\begin{cases} \lambda_n = \lambda, & n \geq 0 \\ \mu_n = n\mu, & n \geq 1 \end{cases} \quad (2.63)$$

La matrice de transition  $\mathbf{A}$  est la suivante :

$$A = \begin{matrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \begin{bmatrix} -\lambda & \lambda & & & \\ \mu & -(\lambda + \mu) & \lambda & & \\ & 2\mu & -(\lambda + 2\mu) & \lambda & \dots \\ & & \vdots & & \vdots \end{bmatrix} \end{matrix}.$$

(2.64)

Les équations de balance pour les probabilités limites  $\{P_n, n = 0, 1, 2, \dots\}$  prennent la forme suivant:

$$\begin{aligned} \lambda P_0 &= \mu P_1 \\ (\lambda + n\mu)P_n &= \lambda P_{n-1} + (n+1)\mu P_{n+1}, \quad n = 0, 1, 2, \dots \end{aligned} \quad (2.65)$$

Ces équations avec la condition normalisée  $\sum_{n=0}^{\infty} P_n = 1$  nous mèneront à la solution suivante :

$$P_n = e^{-\lambda/\mu} \frac{(\lambda/\mu)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (2.66)$$

Cette probabilité  $P_n$  suit la distribution de Poisson de paramètre

On en déduit que le nombre  $L$  de clients en cours de service a pour moyenne et la variance (loi de Poisson):  $E(L) = V(L) = \lambda/\mu$ .

### II.9. La file d'attente $M^{(X)}/M/1$

Les systèmes d'attente discutés avant se sont des systèmes dont les clients arrivent un par un dans une unité de temps. Il y a des situations où les clients arrivent par groupe et le meilleur exemple est l'arrivée des clients dans un restaurant, et le trafic des données partagées en paquets dans un système de communication. On a les clients qui arrivent en groupe dans le système avec une taille  $X$ , où en général,  $X$  est une variable aléatoire supposée qu'elle prenne des valeurs entières supérieures à zéro. Nous avons le groupe de clients qui arrivent avec un processus de Poisson de paramètre  $\lambda$  et passe dans le service un par un et le temps du service est exponentiel avec un taux  $\mu$ . On utilise la notation symbolique :  $M^{(X)}/M/1$  pour signifier ce système.

Soit  $d_K = P(X = K)$ ,  $K = 1, 2, \dots$  est la distribution de la taille de groupe des clients qui arrivent. On suppose que la taille de groupe est indépendante complètement des autres caractéristiques du système. Par conséquent, une fois qu'une arrivée a eu lieu, le nombre de

clients dans le système s'accroît avec la taille de groupe.

Soit  $Q(t)$  est le nombre de clients dans le système à l'instant  $t$ , et  $Q$  représente  $Q(t)$  quand  $t \rightarrow \infty$ . A cause de l'accroissement de  $Q(t)$  par la taille de groupe qui arrive aux instants d'arrivées,  $\{Q(t)\}$  est un processus de naissance et de mort.

Soit  $P_n = P(Q = n)$ ,  $n = 0, 1, 2, \dots$  et faire des modifications appropriées dans les équations de balance de  $M/M/1$  données dans les équations (2.3), on aura:

$$\begin{aligned} \lambda P_0 &= \mu P_0 \\ (\lambda + \mu)P_n &= \lambda \sum_K d_K P_{n-K} + \mu P_{n+1}, \quad n = 1, 2, \dots \end{aligned} \quad (2.67)$$

On note que le premier terme dans la seconde équation (2.67) existe si seulement si  $(n - K) \geq 0$ .

Malheureusement, les équations (2.67) ne seront pas résolues par l'utilisation des méthodes récursives comme il a été fait dans le cas du modèle  $M/M/1$ . A la place, on utilise la PGF pour simplifier ces équations.

Soit :

$$P(z) = \sum_{n=0}^{\infty} P_n z^n, \quad \delta(z) = \sum_{K=1}^{\infty} d_K z^K, \quad |z| = 1$$

Multiplier les équations dans les équations (2.67) par  $z$  à la puissance  $n$ , on aura:

$$\begin{aligned} \lambda P_0 &= \mu P_1 \\ (\lambda + \mu) \sum_{n=1}^{\infty} P_n z^n &= \lambda \sum_{n=1}^{\infty} z^n \sum_{d=1}^{\infty} d_K P_{n-K} + \mu \sum_{n=1}^{\infty} P_{n+1} z^n \end{aligned} \quad (2.68)$$

Intervertir les sommes des équations (2.68) et simplifier, on obtient:

$$\begin{aligned} (\lambda + \mu)P(z) - \mu P_0 &= \sum_{K=0}^{\infty} d_K z^K \sum_{n=K}^{\infty} d_K P_{n-K} \\ &+ \mu \sum_{n=0}^{\infty} P_{n+1} z^n = \lambda \delta(z) P(z) + \left(\frac{\mu}{z}\right) \sum_{m=1}^{\infty} z^m P_m = \lambda \delta(z) P(z) + \left(\frac{\mu}{z}\right) [P(z) - P_0] \end{aligned}$$

Réarranger les termes et simplifier:

$$P(z) = \frac{\mu P_0 (1-z)}{\mu(1-z) - \lambda z [1 - \delta(z)]} \quad (2.69)$$

Pour déterminer  $P_0$ , on utilise la condition  $\sum_{n=0}^{\infty} P_n = 1$  et on note

que  $\lim_{z \rightarrow \infty} P(z) = 1$ . Passer à la limite dans la formule (2.69) et l'utilisation de la règle d'Hôpital, on aura:

$$\lim_{z \rightarrow 1} P(z) = \frac{\lim_{z \rightarrow 1} \mu P_0 (1 - z)}{\lim_{z \rightarrow \infty} (\mu(1 - z) - \lambda z [1 - \delta(z)])}$$

Nous donne:

$$1 = \frac{\mu P_0}{\mu + \lambda(1 - \delta'(1))} \Rightarrow P_0 = 1 - (\lambda \delta'(1) / \mu) \quad (2.70)$$

On note que:

$$\delta'(1) = \lim_{z \rightarrow 1} \sum_{K=1}^{\infty} K d_K z^{K-1} = E(X) = d$$

Avec la moyenne de la taille de groupe  $d$ , on note ainsi que ( $\rho = \lambda d / \mu$ ) l'intensité du trafic. Cela nous donne le résultat :  $\rho = 1 - P_0$  (2.71)

Et,

$$P(z) = \frac{\mu(1-z)(1-\rho)}{\mu(1-z) - \lambda z [1 - \delta(z)]} \quad (2.72)$$

Malheureusement, avec les formes simples de la distribution  $\{d_n\}$ , invertir les PGF de (2.72) n'est pas simple. Malgré, la formule (2.72) est facilement utilisé pour la détermination de la valeur moyenne  $Q$  tel que  $t \rightarrow \infty$  en notant que  $E(Q) = \lim_{z \rightarrow 1} P'(z)$ .

A cause des termes  $(1 - z)$  dans le numérateur et  $(1 - \delta(z))$  dans le dénominateur de (2.72), on utilise la règle d'Hôpital en prenant la limite dans  $P'(z)$ . Après les simplifications, on aura:

$$E(Q) = \lim_{z \rightarrow 1} P'(z) = \frac{2\rho + \left(\frac{\lambda}{\mu}\right) \delta''(1)}{2(1-\rho)} \quad (2.73)$$

Mais  $\delta''(1) = E(X^2) - E(X)$ . Avec  $d$  comme la taille moyenne de groupe, la formule (2.73) avec les simplifications:

$$L = E(Q) = \frac{\rho + \left(\frac{\lambda}{\mu}\right) E(X^2)}{2(1-\rho)} \quad (2.74)$$

La variance de  $Q$  peut être déterminée, il suffit de faire tendre  $z \rightarrow 1$  dans  $P''(z)$  et notant que  $V(Q) = P''(1) + P'(1) - [P'(1)]^2$ . Quand les tailles de groupe sont constante  $K$ , le nombre moyen des clients dans le système est donné par (2.74) avec les simplifications:

$$L = E(Q) = \frac{\rho + (\lambda/\mu) K^2}{2(1-\rho)} = \frac{K+1}{2} \frac{\rho}{1-\rho} \quad (2.75)$$

Avec  $\rho = \lambda K/\mu$

### II.10. La file d'attente $M/M^{(X)}/1$

Dans le modèle  $M/M^{(X)}/1$ , les clients arrivent un par un à l'instant  $t$ , mais ils sont servis par groupe d'une taille  $X$ . Tout simplement, on suppose que  $X$  est une constante  $K$ . Quand le service est en groupe, il y a deux facteurs de la discipline de la file que l'analyse peut être compliquée :

- Si oui ou non le serveur attend l'arrivée d'un client quand le nombre de clients est inférieur à  $K$  dans la file à l'instant  $t$  à la fin de service ;
- Si le serveur commence avec moins que  $K$  clients dans le groupe et les nouveaux clients qui arrivent sont permis de joindre le service en cours ou ils sont requis d'attendre la nouvelle série. On met une supposition que le serveur commence son service quand le lot est complet.

Donc les clients arrivent un par un à l'instant  $t$  avec un processus de Poisson de paramètre  $\lambda$  et ils sont servis par groupe d'une taille  $K$  si on a  $K$  ou plus clients dans la queue à la fin de service ou si on a moins que  $K$  clients, le serveur attend jusqu'à ce que l'achèvement du service de lot  $K$ . La distribution de temps de service est exponentielle avec un paramètre  $\mu$ . Avec ces suppositions, la distribution limite  $\{P_n, n = 0, 1, 2, \dots\}$  de nombre de clients dans le système à l'instant  $t \rightarrow \infty$ , sera déterminée à partir les équations de balance qui sont présentées comme suit:

$$\begin{aligned} \lambda P_0 &= \mu P_K \\ \lambda P_n &= \lambda P_{n-1} + \mu P_{n+K}, \quad n = 1, 2, \dots, K-1 \\ (\mu + \lambda)P_n &= \lambda P_{n-1} + \mu P_{n+K}, \quad n = K, K+1, \dots \end{aligned} \quad (2.76)$$

La méthode qu'on utilise pour résoudre ces équations est la PGF. Multiplier les deux cotés des équations (2.76) par des puissances appropriées de  $z$  et on fait la somme. On aura :

$$\begin{aligned} \lambda P_0 + \lambda \sum_{n=1}^{K-1} P_n z^n \\ + (\lambda + \mu) \sum_{n=K}^{\infty} P_n z^n \\ = \mu P_K + \lambda \sum_{n=1}^{K-1} P_{n-1} z^n + \mu \sum_{n=1}^{K-1} P_{n+K} z^n + \lambda \sum_{n=K}^{\infty} P_{n-1} z^n + \mu \sum_{n=K}^{\infty} P_{n+K} z^n \end{aligned}$$

Notant que  $\sum_{n=0}^{\infty} P_n z^n = P(z)$  et faire des simplifications appropriées, on peut écrire:

$$\begin{aligned} & (\lambda + \mu)P \\ - \mu \sum_{n=0}^{K-1} P_n z^n &= \lambda z \sum_{n=0}^{\infty} P_{n-1} z^n + \frac{\mu}{z^K} \sum_{n=0}^{\infty} P_{n+K} z^{n+K} \\ &= \lambda z P(z) + \frac{\mu}{z^K} [P(z) - \sum_{m=0}^{K-1} P_m z^m] \end{aligned}$$

Réarranger les termes et multiplier par  $z^n$ , on aura:

$$[(\lambda + \mu)z^K - \lambda z^{K+1} - \mu]P(z) = \mu(z^n - 1) \sum_{n=0}^{K-1} P_n z^n \Rightarrow P(z) = \frac{(1-z^K) \sum_{n=0}^{K-1} P_n z^n}{(\lambda/\mu)z^{K+1} - (\lambda/\mu+1)z^K + 1} \quad (2.77)$$

Pour une complète détermination de la PGF de  $P(z)$ , on a besoin de déterminer  $\sum_{n=0}^{K-1} P_n z^n$  dans le numérateur. Pour cela on a utilisé le théorème de Rouché dans la théorie des variables complexes.  $P(z)$  doit se converger à l'intérieur du cercle d'unité pour qu'elle soit une PGF.

Le dénominateur de (2.77) a  $(K+1)$  zéros. Donc pour  $P(z)$  d'être une propre PGF, le numérateur de (2.77) doit faire disparaître ces  $(K+1)$  zéros. Il est facile de voir que  $(z=1)$  est un zéro sur le numérateur que le dénominateur. On peut montrer exactement  $(K-1)$  zéros de dénominateur sont à l'intérieur du cercle d'unité, laissant un zéro à l'extérieur. On a  $(z_0 > 1)$  est la racine de l'équation :

$$\left(\frac{\lambda}{\mu}\right) z^{K+1} - \left(\frac{\lambda}{\mu} + 1\right) z^K + 1 = 0 \quad (2.78)$$

Effectivement, si on divise le dénominateur de la formule (2.77) par  $(z-1)(z-z_0)$ , on a quitté avec a polynomial avec  $(K-1)$  racines à l'intérieur du cercle d'unité. La portion du numérateur avec des zéros à l'intérieur du cercle d'unité est  $\sum_{n=0}^{K-1} P_n z^n$  par conséquent, on aura:

$$\sum_{n=0}^{K-1} P_n z^n = c \frac{(\lambda/\mu)z^{K+1} - (\frac{\lambda}{\mu}+1)z^K + 1}{(z-1)(z-z_0)} \quad (2.79)$$

Substituant ce résultat dans la formule (2.77), on a:

$$P(z) = \frac{c(1-z^K)}{(z-1)(z-z_0)} = \frac{c}{z_0-z} \sum_{n=0}^{K-1} z^n \quad (2.81)$$

Avec  $P(1)=1$ , et mettant  $(z=1)$  dans la formule (2.81), on aura:

$$C = \frac{z_0 + 1}{K}$$

Et,

$$P(z) = \frac{(z_0-1)\sum_{n=0}^{K-1} z^n}{K(z_0-z)} \quad (2.82)$$

A l'intérieur de la formule (2.82) peut être transformée à des séries en  $x$  pour déterminer  $\{P_n, n = 0, 1, 2, \dots\}$  explicitement comme suit :

$$P(z) = \frac{z_0-1}{Kz_0} (\sum_{s=0}^{K-1} z^s) (\sum_{r=0}^{\infty} (\frac{z}{z_0})^r) \quad (2.83)$$

$$P_n = \begin{cases} \frac{z_0-1}{Kz_0} \sum_{r=0}^n (\frac{1}{z_0})^r, & n < K \\ \frac{z_0-1}{Kz_0^{n-K+1}} \sum_{r=0}^{K-1} (\frac{1}{z_0})^r, & n \geq K \end{cases} \quad (2.84)$$

Notant que:

$$\sum_{r=0}^n (\frac{1}{z_0})^r = \frac{1 - (1/z_0)^{n+1}}{1 - (1/z_0)}$$

On peut présenter la formule (2.84) comme :

$$P_n = \begin{cases} \frac{z_0^{n+1}-1}{Kz_0^{n+1}}, & n < K \\ \frac{z_0^K-1}{Kz_0^{n+1}}, & n \geq K \end{cases} \quad (2.85)$$

Trouver la racine  $z_0$  à l'extérieur du cercle d'unité de la formule (2.78) est essentiel pour la détermination de la distribution limite. C'est le même problème dans l'analyse des systèmes de ce type, et il y a des algorithmes pour trouver des racines spécifiquement applicables dans chaque cas.

### ***II.11. Conclusion***

Dans ce chapitre, nous avons vu quelques systèmes d'attente markoviens classiques et nous avons présenté pour chaque modèle, quelques mesures de performances dans un régime stationnaire avec la discipline FCFS (First Come First Served).

# CHAPITRE III

SYSTEMES D'ATTENTE NON MARKOVIENS



# CHAPITRE III

## SYSTEMES D'ATTENTE NON-MARKOVIENS

---

### *III.1. Introduction*

Dans le système d'attente  $M/G/1$ , les clients arrivent avec un processus de Poisson et ils sont servis avec un seul serveur et on suppose que les durées de service sont indépendantes et identiquement distribuées avec une distribution générale. Soit  $Q(t)$  est le nombre de clients dans le système à l'instant  $t$ . Et pour une description complète de l'état de système à l'instant  $t$ , on a besoin de la valeur de  $Q(t)$  comme étant une information sur le reste de la durée de service des clients dans le service, si un est servi à cet instant. On a  $R(t)$  est le reste de temps de service de chaque client.

Maintenant le vecteur  $[Q(t), R(t)]$  est un vecteur de processus de Markov avec ces composants tel que les temps des arrivées et de service sont complètement spécifiés. La première investigation pour analyser ce processus vectoriel a été faite et améliorée par Cox (1955), et qui a utilisé l'information  $R(t)$  comme variable supplémentaire pour la construction des équations Kolmogorov.

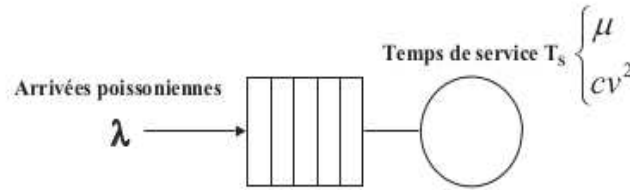
Après c'est Kendall (1951, 1953) qui a développé la procédure de convertir la longueur des processus de la file dans  $M/G/1$  et  $G/M/s$  à des chaînes de Markov. (Dans la file  $G/M/s$ , la durée de service vérifie la propriété sans mémoire. Par conséquent, dans le processus vectoriel  $[Q(t), R(t)]$ , où  $R(t)$  représente maintenant la durée jusqu'à une nouvelle arrivée). La stratégie ici est de considérer les instants des départs dans la file  $M/G/1$  et les instants des arrivées dans la file  $G/M/s$ .

Soit  $\{t_0 = 0, t_1, t_2, \dots\}$  sont des instants de départ des clients dans le modèle  $M/G/1$  et définir  $Q = Q(t_n + 0)$ . Par conséquent  $Q_n$  est définie comme étant la valeur de  $Q(t)$  après les départs. Dans ces instants  $\{t_n, n = 0, 1, 2, \dots\}$ ,  $R(t)$  est nul, et donc  $Q_n$  peut être étudiée sans référence à la variable aléatoire  $R(t)$ . A cause de la propriété de Markov de la distribution de Poisson que le processus  $\{Q_n, n = 0, 1, 2, \dots\}$  est une chaîne de Markov avec un paramètre discret et espace d'états. A cause de la nature de processus induit qui est connu comme une chaîne de Markov induite. Dans la file d'attente  $G/M/s$ , les instants des arrivées génèrent la chaîne de Markov induite. On verra après une analyse

détaillée et discussion profonde sur les deux systèmes  $M/G/1$  et  $G/M/1$ .

Les chaînes de Markov induites peuvent être ainsi utilisées dans l'analyse des temps d'attente dans le système  $G/G/1$  [1].

### III.2. Le système d'attente $M/G/1$



**Figure 7:** La représentation d'un système d'attente  $M/G/1$

Soit les clients qui arrivent avec un processus de Poisson du paramètre  $\lambda$  et ils sont servis par un seul serveur. Soit les durées de service de ces clients sont des variables aléatoires qui sont indépendantes et identiquement distribuées  $\{S_n, n = 0, 1, 2, \dots\}$  avec :

$$P(S_n \leq x) = B(x), \quad x \geq 0; \quad E(S_n) = b; \quad V(S_n) = \sigma_n^2$$

On suppose que  $S_n$  est la durée de service de  $n^{\text{ième}}$  client. Soit  $Q(t)$  est le nombre de clients dans le système à l'instant  $t$  et identifier  $\{t_0 = 0, t_1, t_2, \dots\}$  comme des instants de départs des clients.

Soit  $Q_n = Q(t_n + 0)$  est le nombre de clients dans le système après le  $n^{\text{ième}}$  départ. On peut montrer que  $\{Q_n, n = 0, 1, 2, \dots\}$  est une chaîne de Markov.

Soit aussi  $X_n$  est le nombre de clients qui arrivent durant la durée de service  $S_n$ .

Avec la supposition que le processus des arrivées est un processus de Poisson, on a :

$$K_j = P(X_n = j) = \int_0^\infty P(S_n = j | S_n) P(t < S_n \leq t + dt) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^j}{j!} dB(t), \quad j = 0, 1, 2, \dots \quad (3.1.1)$$

Dans l'écriture de la formule (3.1.1), on utilise la notation de Stieltjes en ordre de contenir ou de loger des distributions discrètes, continues et mixtes [1].

Considérant la relation entre  $Q_n$  et  $Q_{n+1}$ :

$$Q_{n+1} = \begin{cases} Q_n + X_{n+1} - 1 & \text{si } Q_n > 0 \\ X_{n+1} & \text{si } Q_n = 0 \end{cases} \quad (3.1.2)$$

La première expression pour  $Q_{n+1}$  est claire. La seconde expression (c-à-d,  $X_{n+1}$  si  $Q_n=1$ ) est un résultat dans le fait que  $t_{n+1}$  est un instant de départ d'un client qui est arrivé après  $t_n$ . En fait,  $Q_{n+1} = 1 - 1 + X_{n+1}$ .

Comme on peut voir à l'intérieur de la formule (3.1.2) que  $Q_{n+1}$  s'exprime en termes de  $Q_n$  et la variable aléatoire  $X_{n+1}$ , dont elle ne dépend d'aucun événement avant  $t_n$ . Avec  $X_{n+1}$  est indépendante et identiquement distribuée, elle ne dépend pas de  $Q_n$ .

D'où  $\{Q_n, n = 0, 1, 2, \dots\}$  est une chaîne de Markov et son espace de paramètre est construit à partir des points de départ, et l'espace d'états  $S$  est le nombre de clients dans le système;  $S = \{0, 1, 2, \dots\}$ . A cause de la nature induite de l'espace des paramètres qui forme une chaîne de Markov induite [1].

Soit :

$$P_{ij}^n = P(Q_n = j | Q_0 = i), \quad i, j \in S \quad (3.1.3)$$

Et écrire  $P_{ij}^1 = P_{ij}$

D'après la relation (3.1.2) et la définition de  $K_j$  dans (3.1.1), on peut écrire:

$$\begin{aligned} P_{ij} &= P(Q_{n+1} = j | Q_n = i); \quad i, j \in S \\ &= \begin{cases} P(i + X_{n+1} - 1 = j) & \text{si } i > 0 \\ P(X_{n+1} = j) & \text{si } i = 0 \end{cases} = \begin{cases} K_{j-i+1} & \text{si } i > 0 \\ K_j & \text{si } i = 0 \end{cases} \end{aligned} \quad (3.1.4)$$

La matrice des probabilités de transition  $\mathbf{P}$  est :

$$\mathbf{P} = \begin{matrix} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \dots \\ \mathbf{0} & \left[ \begin{array}{cccc} k_0 & k_1 & k_2 & \dots \\ k_0 & k_1 & k_2 & \dots \\ & k_0 & k_1 & \dots \\ & & k_0 & \dots \\ \dots & & & \dots \end{array} \right] \\ \mathbf{1} & \\ \mathbf{2} & \\ \mathbf{3} & \\ \dots & \end{matrix}.$$

Pour que la chaîne de Markov soit irréductible (l'espace d'état a une seule classe d'équivalence) les deux conditions suivantes doivent contenir :  $K_0 > 0$  et  $K_0 + K_1 = 1$ . Il est facile de voir que si  $K_0=0$ , avec un ou plus de clients qui arrivent pour chaque départ, il n'y

aura aucune possibilité que le système qu'il atteigne la stabilité, et le nombre de clients dans le système va seulement augmenter avec le temps. Si  $K_0 + K_1 = 1$  seulement deux états  $\{0, 1\}$  qui sont possible dans le système [1].

En outre la classification des états dépend de  $E(X_n)$ , le nombre de clients qui arrivent durant la durée de service.

On définit la transformée de Laplace–Stieltjes de la distribution de la durée de service :

$$\psi(\theta) = \int_0^{\infty} e^{-\theta t} dB(t), \quad \operatorname{Re}(\theta) > 0$$

Et la PGF de nombre de clients qui arrivent durant la durée de service,

$$K(z) = \sum_{j=0}^{\infty} K_j z^j, \quad |z| \leq 1$$

Les résultats au-dessous sont dérivés de propriétés connues de la transformée de Laplace–Stieltjes et la PGF:

$$E(S_n) = b = -\psi(0), \quad E(S_n^2) = \psi''(0), \quad E(X_n) = K'(1), \quad E(X_n^2) = K''(1) + K'(1) \quad (3.1.8)$$

A partir de (5.1.1) on aura :

$$K(z) = \sum_{j=0}^{\infty} \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!} z^j dB(t)$$

D'où :

$$\begin{aligned} K'(z) &= -\lambda \psi'(\lambda - \lambda z) \\ K'(1) &= -\lambda \psi'(0) = \lambda b = \rho \end{aligned} \quad (3.1.9)$$

Soit  $b\lambda = (\text{taux des arrivées}) \times (\text{le temps moyen de service})$ . Cette quantité est appelée l'intensité du trafic de système d'attente notée par  $\rho$ . La valeur de  $\rho$  indique est-ce que le système est stable (qu'il atteint l'état récurrent) quand le paramètre  $n$  de temps  $t_n$  tend vers l'infini.

On peut montrer que  $\rho < 1$ , et cela si seulement si la chaîne de Markov est récurrente positive c.-à-d., le processus retourne à n'importe quel état avec une probabilité  $P=1$  et le temps moyen de retour est fini  $< \infty$ ; quand  $\rho = 1$ , la chaîne est récurrente nulle (c.-à-d., le processus retourne à n'importe quel état avec une probabilité  $P=1$ , mais le temps moyen de retour n'est pas fini  $= \infty$ ; et quand  $\rho > 1$ , la chaîne est transitoire c.-à-

d., le processus ne peut pas retourner dans tout les états finis; et la probabilité que le processus sera trouvé dans l'un des états finis est nulle  $P=0$ .

Néanmoins, ces propriétés sont faciles à comprendre si on a compris réellement la signification la valeur de l'intensité du trafic [1].

Les probabilités de la  $n^{\text{ième}}$  transition  $P_{ij}^{(n)}$ , ( $i, j = 0, 1, 2, \dots$ ) de la chaîne de Markov  $Q_n$  sont obtenues comme étant des éléments de la  $n^{\text{ième}}$  puissance de la  $1^{\text{ière}}$  étape de probabilité de transition de la matrice  $P$ . En considérant  $P^n$  dans des systèmes réels, les deux observations suivantes seront utiles :

1. Comme  $n$  s'accroît dans  $P^n$ , le bloc des éléments non nuls dans les sous-matrices représente des classes d'équivalence récurrentes.
2. Dans une chaîne de Markov apériodique, irréductible et récurrente positive et comme  $n$  dans  $P^n$  s'accroît, les éléments de chaque colonne tendent vers une valeur intermédiaire.

La probabilité  $P_{ij}^{(n)}$ , ( $j = 0, 1, 2, \dots$ ) et  $n$  fini donne le temps dépendant du comportement du processus de la longueur de la file  $Q_n$ . Il y a des techniques analytiques pour dériver ces probabilités. Par exemple, voir Takács (1962)[3], qui a utilisé les PGFs pour simplifier les relations récursives générées par les relations de Chapman–Kolmogorov pour  $P_{ij}$ .

Prabhu et Bhat (1963)[4] ont cherché les transitions de  $Q_n$  comme quelques problèmes de premier passage et en utilisant des méthodes combinatoires pour les résoudre. (Voir ainsi Prabhu (1965)[2]). Dans la pratique, comme la puissance des ordinateurs ne cesse d'augmenter dans les opérations du calculs matriciels d'où on peut avec des simples multiplications de  $P$  pour atteindre sa  $n^{\text{ième}}$  puissance. Quand l'espace d'état n'est pas fini, les observations données avant peuvent être utilisées pour le limiter sans perdre la signification de l'ensemble de l'information [1].

### ***La distribution limite***

La deuxième observation donnée en plus, provienne de la propriété de la chaîne de Markov apériodique, récurrente positive et irréductible, dont les résultats dans  $\lim_{n \rightarrow \infty} P^n$  deviendront une matrice avec des rangées identiques. Cette propriété peut être validée par l'obtention des puissances successives de  $P^n$ ; comme  $n$  s'accroît, les éléments dans la colonne de la matrice tendront vers une valeur constante intermédiaire. Le comportement de la chaîne de Markov est codifié dans le théorème suivant et ces corollaires sont donnés

sans preuve ici:

**Théorème :**

✓ Soit  $i$  est l'état qui appartient à une classe d'équivalence récurrente apériodique. Soit  $P_{ij}^{(n)}$  est la probabilité de la  $n^{\text{ième}}$  étape de transition de  $i \rightarrow i$ , et  $\mu_i$  est son temps moyen de récurrence. Alors  $\lim_{n \rightarrow \infty} P_{ii}^{(n)}$  existe et donnée par:

$$\lim_{n \rightarrow \infty} P_{ii}^{(n)} = \frac{1}{\mu_i} = \pi_i$$

✓ Soit  $j$  un autre état qui appartient à la même classe d'équivalence et  $P_{ji}^{(n)}$  est la probabilité de la  $n^{\text{ième}}$  étape de transition de  $j \rightarrow i$ . Et

$$\lim_{n \rightarrow \infty} P_{ji}^{(n)} = \lim_{n \rightarrow \infty} P_{ii}^{(n)} = \pi_i$$

**Corollaire :**

Si  $i$  est récurrent positif,  $\pi_i > 0$ , et si  $i$  est récurrent nul,  $\pi_i = 0$ .

Notant que le terme du temps de récurrence dans le théorème signifie le nombre d'étapes que la chaîne de Markov prenne pour retourner vers l'état initial pour le premier temps (L'état que le processus a commencé pour la première fois).

Le théorème s'applique pour toutes les chaînes de Markov si leur espace d'état fini où infiniment dénombrable. Pour l'espace d'état  $s = \{0, 1, 2, \dots\}$  et soit  $(\pi_0, \pi_1, \pi_2, \dots)$  est le vecteur de probabilité limite où  $\pi_i = \lim_{n \rightarrow \infty} P_{ji}^{(n)}$ ,  $i, j \in S$ . Soit la matrice avec identiques rangées  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$ . Maintenant, on utilise les relations de Chapman-Kolmogorov, on peut écrire  $P^n = P^{n-1}P$ .

Avec l'application du théorème pour  $P^n$  et  $P^{n-1}$ , il est facile d'écrire :  $\Pi = \Pi P$

Or,

$$\pi = \pi P \quad (3.1.10)$$

En plus, on multiplie les deux cotés de (3.1.10) par  $P$ , on peut ainsi établir que:

$$\Pi = \Pi P = \Pi P^2$$

$$\Pi = \Pi P = \Pi P^n \quad (3.1.11)$$

La dernière équation montre que si on utilise la distribution comme distribution initiale de l'état la chaîne de Markov irréductible, apériodique, et récurrente positive, la distribution de l'état après  $n$  transitions ( $n=1, 2, 3, \dots$ ) est ainsi donnée par la même

distribution limite. Cette propriété est connue sous le nom, la distribution stationnaire.

Le théorème suivant résume ces résultats et nous fournit la procédure à qui la distribution limite peut être déterminée.

***Théorème :***

✓ Dans une chaîne de Markov irréductible, apériodique et récurrente positive, les probabilités limites  $\{\pi_i, i = 0, 1, 2, \dots\}$  satisfaisant les équations suivantes:

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j = 0, 1, 2, \dots$$

$$\sum_{j=0}^{\infty} \pi_j = 1 \quad (3.1.12)$$

La distribution limite est stationnaire.

✓ Chaque solution de ces équations,

$$\sum_{i=0}^{\infty} x_i P_{ij}, \quad j = 0, 1, 2, \dots \quad (3.1.13)$$

représente un scalaire multiple de  $\{\pi_i, i = 0, 1, 2, \dots\}$  à condition que:

$$\sum |x_i| \leq \infty$$

Par conséquent, la distribution limite de la chaîne de Markov est obtenue par la résolution de l'ensemble des équations simultanées (3.1.12) et normaliser la solution en utilisant la seconde équation  $\sum_{j=0}^{\infty} \pi_j = 1$ . On note que, à cause des sommes de la rangée de la chaîne de Markov sont 0 ou égal à 1, et la formule (3.1.12) par elle-même qui nous donne une solution sous forme d'une constante multiplicative.

La normalisation de la condition est par conséquent essentielle dans la détermination de la distribution limite. Avec ce contexte de la théorie générale des chaînes de Markov, on est maintenant dans la position de déterminer la distribution limite de la chaîne de Markov induite de la file  $M/G/1$ .

Soit  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$  est la distribution limite de la chaîne de Markov induite. L'utilisation de la matrice des probabilités de transition (3.1.5) dans l'équation  $\pi = \pi P$  (Dont est (3.1.12)), on a:

$$k_0 \pi_0 + k_0 \pi_1 = \pi_0$$

$$k_1 \pi_0 + k_1 \pi_1 + k_0 \pi_2 = \pi_1$$

$$k_2 \pi_0 + k_2 \pi_1 + k_1 \pi_2 + k_0 \pi_3 = \pi_2$$

$$\begin{aligned} & \vdots \\ (3.1.14) \end{aligned}$$

Le chemin qui convient à la résolution de ces équations est de définir d'abord :  $v_0 = 1$  et  $v_i = \pi_i/\pi_0$  et réécrire (3.1.14) avec les termes de  $\{v_i, i = 0, 1, 2, \dots\}$  tel que :

$$\begin{aligned} v_1 &= \frac{(1 - k_0)}{k_0} \\ v_2 &= \frac{(1 - k_1)}{k_0} v_1 - \frac{k_1}{k_0} \\ & \vdots \\ & \vdots \\ v_j &= \frac{(1 - k_1)}{k_0} v_{j-1} - \frac{k_2}{k_2} v_{j-2} - \dots - \frac{k_{j-1}}{k_0} v_1 - \frac{k_{j-1}}{k_0} \\ & \vdots \\ & \vdots \end{aligned}$$

(3.1.15)

Ces équations sont résolues d'une manière récursive pour déterminer  $\{v_i, i = 0, 1, 2, \dots\}$ . Les probabilités limites  $(\pi_0, \pi_1, \pi_2, \dots)$  sont connues pour être monotones et concaves, et donc pour des valeurs larges de  $n$ , elles deviennent extrêmement petites. Effectivement,  $v_i = \pi_i/\pi_0$  aura ainsi les mêmes propriétés, dont il est facile d'établir une valeur limite pour la taille de l'espace d'état.

Dans l'ordre de récupérer  $\pi_i$  de  $v_i$ , on note que:

$$\sum_{i=0}^{\infty} v_i = 1 + \sum_{i=1}^{\infty} \frac{\pi_i}{\pi_0} = \sum_{i=1}^{\infty} \frac{\pi_i}{\pi_0} = \frac{1}{\pi_0}$$

Ici on a incorporé la condition normalisée  $\sum_{i=0}^{\infty} \pi_i = 1$  et par conséquent, on aura:

$$\pi_0 = \left(1 + \sum_{i=1}^{\infty} v_i\right)^{-1}$$

Et,

$$\pi_i = v_i \left(1 + \sum_{i=1}^{\infty} v_i\right)^{-1}$$



Analytiquement, la distribution limite  $(\pi_0, \pi_1, \pi_2, \dots)$  peut être déterminée par la résolution des équations (3.1.14) en utilisant effectivement les fonctions génératrices. Malheureusement, les expressions explicites dérivées pour les probabilités ont besoin d'inverser le résultat de la PGF.

Cependant, on peut obtenir la moyenne et la variance de la distribution en utilisant des techniques standards. On définit :

$$\Pi(z) = \sum_{j=0}^{\infty} \pi_j z^j, \quad |z| \leq 1$$

Et,

$$K(z) = \sum_{j=0}^{\infty} K_j z^j, \quad |z| \leq 1$$

Multiplier les équations de (3.2.14) avec des puissances appropriées de  $z$  et faire la somme, on aura :

$$\begin{aligned} \Pi(z) &= \pi_0 K(z) + \pi_1 K(z) + \pi_2 z K(z) + \dots \\ &= \pi_0 K(z) + \frac{K(z)}{z} (\pi_1 z + \pi_2 z^2 + \dots) \\ &= \pi_0 K(z) + \frac{K(z)}{z} [\Pi(z) - \pi_0] \end{aligned}$$

Réarranger les termes,

$$\Pi(z) \left[1 - \frac{K(z)}{z}\right] = \pi_0 K(z) \left[1 - \frac{1}{z}\right] \Rightarrow \Pi(z) = \frac{\pi_0 K(z)(z-1)}{z-K(z)} \quad (3.1.17)$$

La quantité inconnue  $\pi_0$  à l'intérieur de l'expression  $\Pi(z)$  dans (3.1.17) peut être déterminée en utilisant la condition normalisée  $\sum_{j=0}^{\infty} \pi_j = 1$ . On a :  $\Pi(1) = \sum_{j=0}^{\infty} \pi_j = 1$

Faire tendre  $z \rightarrow \infty$  dans (3.1.17), on aura (application de la règle d'Hôpital):

$$1 = \frac{\lim_{z \rightarrow \infty} \pi_0 [K(z) + (z-1)K'(1)]}{\lim_{z \rightarrow \infty} [1 - K'(z)]}$$

Rappelant que  $K(1)$  et  $K(1) = \rho$  (à partir de (3.1.9)), on a:

$$1 = \frac{\pi_0}{1-\rho} \Rightarrow \pi_0 = 1 - \rho \quad (3.1.18)$$

Par conséquent, on aura :

$$\Pi(z) = \frac{(1-\rho)(z-1)K(z)}{z-K(z)} \quad (3.1.19)$$

Les expressions explicites pour les probabilités  $\{\pi_i, i = 0, 1, 2, \dots\}$  peuvent être obtenues par développement de  $\Pi(z)$  dans des cas spéciaux. Une forme alternative de  $\Pi(z)$  s'entraîne à être facile pour le développement.

On peut écrire :

$$\Pi(z) = \frac{(1-\rho)K(z)}{\frac{z-K(z)}{z-1}} = \frac{(1-\rho)K(z)}{1-\frac{z-K(z)}{z-1}} \quad (3.1.20)$$

Notant que  $\sum_{j=0}^{\infty} z^j (K_{j+1} + K_{j+2} + \dots)$  peut être simplifiée à :

$$c(z) = \frac{1-K(z)}{1-z}$$

(Voir aussi les simplifications algébriques trouvées à partir de (3.1.17)). Pour  $|z| \leq 1$ ,

$$|c(z)| = \left| \frac{1-K(z)}{1-z} \right| < 1 \quad \text{si } \rho < 1 \quad (3.1.21)$$

Maintenant, en utilisant le développement des séries géométriques, on peut écrire:

$$\Pi(z) = (1-\rho)K(z) = \sum_{j=0}^{\infty} [c(z)]^j \quad (3.1.22)$$

L'expression explicite de  $\pi_j$  est obtenue par développement de (3.1.22).

Dans les systèmes d'attente, le processus de la longueur de la file  $Q(t)$  peut être considéré avec trois instants différents :

- Quand  $t$  est juste avant l'instant d'une arrivée,
- Quand  $t$  est après l'instant d'un départ,
- Quand  $t$  est un instant arbitraire dans le temps. En général, la distribution de  $Q(t)$  avec la référence à ces trois points de temps ne peut être la même.

Cependant, quand le processus des arrivées est un processus de Poisson, on peut montrer que les distributions limites de  $Q(t)$  dans les trois cas sont les mêmes. Voir Wolff (1982)[5], qui a forgé l'acronyme PASTA (Poisson arrivals see time averages). Pour les preuves de cette propriété, voir ainsi Cooper (1981)[6] and Gross and Harris (1998)[7].

Par conséquent, la PGF  $\Pi(z)$  est dérivée à partir de la formule (3.2.19) et elle nous donne ainsi la distribution limite  $\lim_{t \rightarrow \infty} Q(t)$ . Il y a pas mal de littérature sur la distribution de transition de  $Q(t)$  et pour un  $t$  fini, parmi les articles, ceux de Prabhu and Bhat (1963)[8] and Bhat (1968)[9], qui a obtenu la distribution de transition en utilisant les méthodes récursives et la théorie de renouvellement des arguments. L'expression explicite pour la distribution limite  $Q(t)$  (et la distribution limite de  $Q_n$  dans le cas de la

chaîne de Markov induite) est donnée par:

$$\pi_0 = 1 - \rho$$

$$\pi_j = (1 - \rho) \int_0^\infty e^{-\lambda t} \sum_{n=0}^\infty \left[ \frac{(\lambda t)^{n+j-1}}{(n+j-1)!} - \frac{(\lambda t)^{n+j}}{(n+j)!} \right] dB_n(t) \quad (3.1.23)$$

Pour  $\rho < 1$ , où  $B_n(t)$  est le  $n$ -produit de convolution de  $B(t)$ . (Prabhu and Bhat [4][8](1963), Bhat (1968)[9]).

La moyenne et la variance de  $\lim_{n \rightarrow \infty} Q_n$  peut être déterminée à partir de la PGF (5.2.19) à travers des techniques standards. Écrire  $Q^* = \lim_{n \rightarrow \infty} Q_n$ , on a :  $L = E(Q^*) = \Pi'(1)$

$$V(Q^*) = \Pi''(1) + \Pi'(1) - [\Pi'(1)]^2 \quad (3.1.24)$$

Différentiation de  $(z)$  par rapport à  $z$ , on aura:

$$\Pi'(z) = \frac{1 - \rho}{[z - K(z)]^2} \{ [z - K(z)][K(z) + (z - 1)K'(z)] - (z - 1)[1 - K'(z)]K(z) \}$$

L'utilisation de la règle d'Hôpital deux fois quand on passe à la limite de  $z$ , on aura

:

$$\Pi'(1) = \frac{2K'(1)[1 - K'(1)] + K''(1)}{2[1 - K'(1)]} \quad (3.1.25)$$

Mais notant que, à partir de (3.1.9),  $K'(1) = \rho$  et,

$$K''(1) = \lambda^2 \psi''(0) = \lambda^2 E(S^2) \quad (3.1.26)$$

Où on a utilisé la notation générique pour le temps de service. La substitution de la formule (3.1.26) dans (3.1.25) on aura après les simplifications:

$$L = E(Q^*) = \rho + \frac{\lambda^2 E(S^2)}{2(1 - \rho)} \quad (3.1.27)$$

Dont on fait référence à la formule de Pollaczek–Khintchine. On note que  $\rho$  est le nombre de clients dans le service (Dont elle est la même probabilité que le serveur est occupé dans une file d'attente à un seul serveur), et  $L_q$  qui est le nombre moyen de clients dans la file est obtenu par :

$$L_q = \frac{\lambda^2 E(S^2)}{2(1 - \rho)} \quad (3.1.28)$$

On fait étendre la différentiation pour avoir  $\Pi''(z)$ , et on prend la limite de  $z$ . Avec l'utilisation de multiple de la règle d'Hôpital pour avoir  $\Pi''(1)$  qui nous donne :

$$V(Q^*) = \rho(1 - \rho) + \frac{\lambda^2 E(S^2)}{2(1-\rho)} \left[ 3 - 2\rho + \frac{\lambda^2 E(S^2)}{2(1-\rho)} \right] + \frac{\lambda^3 E(S^3)}{3(1-\rho)} \quad (3.1.29)$$

Rappelant que  $\sigma_S^2$  est la variance de la distribution de la durée de service. Alors  $\sigma_S^2 = E(S^2) - [E(S)]^2$ . En utilisant cette expression dans (3.1.27) et notant que  $\lambda E(S) = \rho$ , on aura une forme alternative pour  $E(Q^*)$  :

$$E(Q^*) = \rho + \frac{\rho^2}{2(1-\rho)} + \frac{\lambda^2 \sigma_S^2}{2(1-\rho)}$$

Dont il est clair que la longueur moyenne de la file s'accroît avec la variance de la distribution de la durée de service. Et pour l'instant, quand  $\sigma_S^2 = 0$ , c-à-d, quand la durée de service est constante ( Dans la file  $M/D/1$ )

$$E(Q^*) = \rho + \frac{\rho^2}{2(1-\rho)} = \frac{\rho}{1-\rho} \left( 1 - \frac{\rho}{2} \right) \quad (3.1.31)$$

Quand la distribution de la durée de service est une distribution d'Erlang avec la moyenne  $1/\mu$  et le paramètre de mesure  $K$ , on aura:  $\sigma_S^2 = 1/K$  et,

$$E(Q^*) = \rho + \frac{\rho^2}{2(1-\rho)} + \frac{\rho^2}{2K(1-\rho)} = \rho + \frac{\rho^2(1+K)}{2K(1-\rho)} \quad (3.1.32)$$

Quand on remplace  $K=1$ , on aura :  $E(Q^*)$  dans  $M/M/1$  tel :  $E(Q^*) = \rho/(1 - \rho)$

### **Temps d'attente :**

Le concept du temps d'attente est utilisé avant dans le contexte de la file  $M/M/1$ . On a explicitement la distribution de la longueur de la queue et on est capable de déterminer la distribution du temps d'attente. Conséquemment, on approche ce problème indirectement en utilisant la PGF  $\Pi(z)$  de la longueur de la queue [1].

Supposons que la discipline de la file est une discipline (FIFO) et soit  $T$  est le temps total passé par un client dans le système c'est à dire le temps dans l'attente et le service, et soit  $T_q$  est le temps d'attente dans la file, tel que  $t \rightarrow \infty$ , on a :  $E(T) = W$  et  $E(T_q) = W_q$ . Ainsi, soit  $F(.)$  est la fonction de la distribution de  $T$  et sa la transformée de Laplace-Stieltjes est donnée par :

$$\phi(\theta) = \int_0^\infty e^{-\theta t} dF(t), \quad R(\theta) > 0$$

Considérant le départ d'un client dans le système sachant qu'il a passé un temps total de  $T$ . On suppose que le départ de ce client laisse derrière lui  $n$  clients; effectivement, ces derniers sont arrivés durant le temps  $T$  dans le système, et on a :

$$P(Q^* = n) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dF(t), \quad n \geq 0 \quad (3.1.33)$$

En utilisant des fonctions génératrices, on aura :

$$\begin{aligned}\Pi(z) &= \sum_{n=0}^{\infty} P(Q^* = n) z^n = \sum_{n=0}^{\infty} z^n \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} dF(t) \\ &= \int_0^{\infty} e^{-\lambda t} \sum_{n=0}^{\infty} z^n \frac{(\lambda t)^n}{n!} dF(t) = \phi(\lambda - \lambda z) \quad (3.1.34)\end{aligned}$$

Comparant (3.2.19) avec (3.2.34), on a :

$$\frac{(1-\rho)(1-z)K(z)}{z-K(z)} = \phi(\lambda - \lambda z) \quad (3.1.35)$$

Rappelant que :

$$K(z) = \psi(\lambda - \lambda z)$$

Et faire la substitution dans la formule (3.1.35),

$$\phi(\lambda - \lambda z) = \frac{(1-\rho)(z-1)\psi(\lambda - \lambda z)}{z - \psi(\lambda - \lambda z)}$$

Écrire  $\lambda - \lambda z = \theta$ , on aura :  $z = 1 - \frac{\theta}{\lambda}$

$$\phi(\theta) = \frac{(1-\rho)\frac{\theta}{\lambda}\psi(\theta)}{\psi(\theta) - (\lambda - \theta)/\lambda} = \frac{(1-\rho)\theta\psi(\theta)}{\theta - \lambda[1 - \psi(\theta)]} \quad (3.1.36)$$

Avec le temps  $T$  dans le système qui est la somme des temps d'attente  $T_q$  et le temps dans le service  $S$ , on définit la transformée de Laplace–Stieltjes de la distribution de  $T_q$  qui est  $\phi_q(\theta)$ , on a :

$$\phi(\theta) = \phi_q(\theta)\psi(\theta) \quad (3.1.37)$$

Comparant (3.1.36) et (3.1.37), on écrit :

$$\phi_q(\theta) = \frac{(1-\rho)\theta}{\theta - \lambda[1 - \psi(\theta)]} \quad (3.1.38)$$

Dont elle peut être exprimée comme suit :

$$\phi_q(\theta) = \frac{(1-\rho)\theta}{1 - \frac{\lambda}{\theta}[1 - \psi(\theta)]} = (1-\rho) \sum_{n=0}^{\infty} \left[ \frac{\lambda}{\theta} (1 - \psi(\theta)) \right]^n \quad (3.1.39)$$

Avec l'utilisation des séries géométriques pour (5.1.39), on peut montrer que

$$\left| \left[ \frac{\lambda}{\theta} (1 - \psi(\theta)) \right] \right| \leq 1 \quad \text{pour } \rho < 1.$$

Dans ce chapitre on a introduit le processus de renouvellement qui est défini comme séquence de variables aléatoires indépendantes et identiquement distribuées.

On suppose que :  $t_{n+1} - t_n = Z_n$  est le  $n^{\text{ième}}$  membre de chaque séquence. Soit  $t$  est

l'instant tel que :  $t_n < t \leq t_{n+1}$ . Et  $t_{n+1} - t_n = R(t)$  est connu comme étant l'instant de récurrence avancé. Si  $B(\cdot)$  est la fonction de distribution de la variable  $Z_n$ , il est possible de montrer que quand  $t \rightarrow \infty$ , que  $r_t(x)$  est la fonction de la densité de  $R(t)$ , qui est donnée comme suit :

$$\lim_{t \rightarrow \infty} r_t(x) = [1 - B(x)] \frac{1}{E(Z_n)} \quad (3.1.40)$$

En utilisant ce concept, on peut invertir (3.1.39) pour donner la fonction de la distribution de  $T_q$  de la sorte :

$$F_q(t) = (1 - \rho) \sum_{n=\hat{a}}^{\infty} \rho^n R^n(t) \quad (3.1.41)$$

Où  $R^n(t)$  est le  $n$ -produit de convolution de la distribution du temps de service restant  $R(t)$ .

Comme il est mentionné dans le premier et le deuxième chapitre que les formules de Little  $L = W$  sont appliquées pour des systèmes d'attente en général qui ont une structure particulière et la discipline de ces systèmes [1].

Alors en utilisant la formule de Little dans (3.1.27) et (3.1.28), on aura :

$$W = E(S) + \frac{\lambda E(S^2)}{2(1-\rho)} \quad (3.1.42)$$

$$W_q = \frac{\lambda E(S^2)}{2(1-\rho)} \quad (3.1.43)$$

Ces moyennes sont déterminées à partir des transformations de  $\phi(\theta)$  et  $\phi_q(\theta)$  par exemple, on a :

$$\begin{aligned} W &= E(T) = \phi'(0) \\ \sigma_T^2 &= V(T) = \phi''(0)[\phi'(0)]^2 \end{aligned}$$

Et les mêmes expressions pour  $W_q$  et  $\sigma_{T_q}^2$ .

Le résultat suivant dérivé de cette manière, peut être utile dans quelques applications :

$$\begin{aligned} \sigma_{T_q}^2 = V(T_q) &= \frac{\lambda E(S^3)}{3(1-\rho)} + \frac{\lambda^2 [E(S^2)]^2}{4(1-\rho)^2} \\ & \quad (3.1.44) \end{aligned}$$

### III.3. Le système d'attente $M/G/1/K$

Considérant la file d'attente  $M/G/1$  décrite avant, avec la restriction que la capacité pour le nombre de clients qui rentrent dans le système est limitée à  $K$ . Avec l'espace d'état

pour la chaîne de Markov induite qui est évidemment le nombre de clients dans le système après le départ d'un client, et  $K$  ne sera pas inclus ou figuré dans cet espace d'état  $S = \{0, 1, 2, \dots, K-1\}$  [1].

Par conséquent, la correspondance à (3.1.2), on a la relation suivante :

$$Q_{n+1} = \begin{cases} \min(Q_n + X_{n+1} - 1, K - 1) & \text{si } Q_n > 0 \\ \min(X_{n+1}, K - 1) & \text{si } Q_n = 0 \end{cases} \quad (3.1.53)$$

En utilisant la distribution de probabilité  $\{K_j, j = 0, 1, 2, \dots\}$  définie dans la formule (3.2.1), on aura la matrice des probabilités de transition :

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \dots & K-1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ K-1 \end{matrix} & \begin{bmatrix} k_0 & k_1 & k_2 & \dots & 1 - \sum_0^{k-2} k_j \\ k_0 & k_1 & k_2 & \dots & 1 - \sum_0^{k-2} k_j \\ & k_0 & k_1 & \dots & 1 - \sum_0^{k-3} k_j \\ & & & \ddots & \\ & & & & 1 - k_0 \end{bmatrix} \end{matrix}.$$

Soit  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$  est la distribution limite de la chaîne de Markov. Ces probabilités sont déterminées par la résolution des équations suivantes:

$$\begin{aligned} \pi_j &= \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j = 0, 1, 2, \dots, K-1 \\ \sum_{j=0}^{\infty} \pi_j &= 1 \end{aligned} \quad (3.1.55)$$

Les premières  $(K-1)$  équations sont identiques à celles de  $M/G/1$  avec la restriction de la capacité illimitée.

On note ici que l'une des  $K$  équations simultanées dans (3.1.55) est redondante à cause la structure des coefficients de la chaîne de Markov et à la place, on utilise la normalisation de la condition  $\sum_{j=0}^{K-1} \pi_j = 1$  pour avoir la solution.

On peut ainsi noter qu'à partir de la même technique dans le cas infini qu'on peut obtenir la solution en utilisant le même  $\{v_i, i = 0, 1, 2, \dots, K-1\}$  dans le cas fini et on détermine :

$$\pi_0 = [\sum_{i=0}^{K-1} v_i]^{-1}, \quad \pi_i = \pi_0 v_i \quad (3.1.56)$$

La discussion de la distribution du temps d'attente est compliquée dans  $M/G/1/K$ ,

et pour la chaîne de Markov, son espace d'état est seulement  $\{0, 1, 2, \dots, K - 1\}$ , quand notre arrivée trouve  $K$  clients dans le système (avant le départ). Par conséquent, de point de vue d'une arrivée, on a besoin d'une distribution limite à un point arbitraire dans le temps.

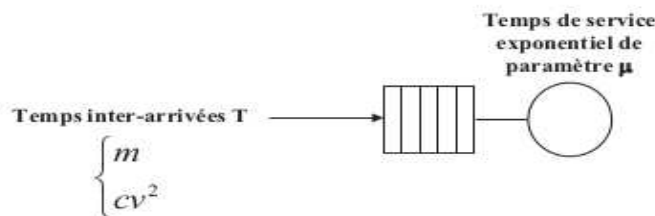
L'analyse de la période occupée est donnée avant pour la file  $M/G/1$  ne peut être modifiée d'une manière facile dans le cas d'une capacité finie. La meilleure approche paraît être de considérer la période occupée comme un problème de premier passage dans la chaîne de Markov irréductible à partir de l'état 1 à l'état 0. Ceci peut être fait par la transformation de l'état 0 à un état absorbant et en utilisant le concept de la matrice fondamentale dans la détermination de nombre de transitions pour le premier passage de transition [1].

### III.4. Le système d'attente $G/M/1$

Soit les clients qui arrivent aux instants  $\{t_0 = 0, t_1, t_2, \dots\}$  et ils sont servis par un seul serveur. Soit aussi  $\{Z_n = t_{n+1} - t_n, n = 0, 1, 2, \dots\}$ , se sont des variables aléatoires indépendantes et identiquement distribuées avec une fonction de distribution  $A(\cdot)$  de moyenne  $a$ .

Alors, soit la distribution de la durée de service est exponentielle avec la même moyenne  $1/\mu$ . On note que ce système est traditionnellement représenté par le symbole  $GI/M/1$  ( $GI$  = générale indépendante).

On utilise la représentation symbolique  $G/M/1$  pour la symétrie avec le système  $M/G/1$ . (Ainsi que, le  $I$  en  $GI$  n'ajoute rien réellement aucune information de plus).



**Figure 8:** La représentation d'un système d'attente  $G/M/1$

Soit  $Q(t)$  est le nombre de clients dans le système à l'instant  $t$  et on définit  $\{Q(t_n - 0) = Q_n, n = 1, 2, \dots\}$ , et par conséquent  $Q_n$  est le nombre de clients dans le



système juste avant la  $n^{\text{ième}}$  arrivée.

On définit  $X_n$  comme étant le nombre des fins de service durant la période des inter-arrivées  $Z_n$ . Soit  $\{b_j, j = 0, 1, 2, \dots\}$  est la distribution de  $X_n$ . On a :

$$b_j = P(X_n = j) = \int_0^\infty e^{-\mu t} \frac{(\mu t)^j}{j!} dA(t) \quad (3.2.1)$$

Maintenant considérant la relation entre  $Q_n$  et  $Q_{n+1}$  :

$$Q_{n+1} = \begin{cases} Q_n + 1 - X_{n+1} & \text{si } Q_n + 1 - X_{n+1} > 0 \\ 0 & \text{si } Q_n + 1 - X_{n+1} \leq 0 \end{cases} \quad (3.2.2)$$

On note qu'avec  $X_{n+1}$  qui est définie comme étant le nombre des départs et la quantité  $Q_n + 1 - X_{n+1}$  peut être inférieure à zéro. Effectivement,  $Q_{n+1}$  ne dépend pas d'aucune variable aléatoire avec un paramètre d'indice inférieur à  $n$  ; et pour cette raison que  $\{Q_n, n = 0, 1, 2, \dots\}$  est considérée comme étant une chaîne de Markov induite du processus de la longueur de la file.

A partir de (3.2.2), On aura les probabilités de transition:

$$P_{ij} = P(Q_{n+1} = j | Q_n = i) = \begin{cases} P(X_{n+1} = i + j - 1) & \text{si } j > 0 \\ P(X_{n+1} = i + 1) & \text{si } j = 0 \end{cases}$$

Nous donne,

$$P_{ij} = b_{i-j+1}, \quad j > 0$$

$$P_{i0} = \sum_{r=i+1}^\infty b_r \quad (3.2.3)$$

La matrice des probabilités de transition  $\mathbf{P}$  prend la forme suivante :

$$\mathbf{P} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & \dots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \left[ \begin{array}{cccc} \sum_1^\infty b_r & b_0 & & \\ \sum_2^\infty b_r & b_1 & b_0 & \\ \sum_3^\infty b_r & b_2 & b_1 & b_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{array} \right] \end{matrix}.$$

$$(3.2.4)$$

La chaîne de Markov est irréductible pour  $b_0 > 0$  et  $b_0 + b_1 < 1$ . On peut facilement déterminer que la chaîne de Markov est apériodique.

Soit :

$$\varphi(\theta) = \int_0^{\infty} e^{-\theta t} dA(t), \quad R(\theta) > 0$$

Est la transformée de Laplace–Stieltjes de  $A(\cdot)$ . En utilisant la PGF de  $b_j$  qui est obtenue par :

$$\beta(z) = \sum_{j=0}^{\infty} b_j z^j = \int_0^{\infty} e^{-(\mu - \mu z)t} dA(t) = \varphi(\mu - \mu z)$$

Les définitions données dans (3.1.8), on aura (en utilisant les symboles génériques  $X$  et  $Z$  pour  $X_n$  et  $Z_n$  respectivement) :

$$E(z) = \beta'(1) = -\varphi'(0) = a\mu \quad (3.2.5)$$

On définit l'intensité du trafic C'est à partir de la formule (5.3.5) qu'on aura :

$$\rho = \frac{1}{a\mu} \quad (3.2.6)$$

On peut montrer que la chaîne de Markov est récurrente positive si seulement si  $\rho < 1$ , récurrente nulle quand  $\rho = 1$ , et transitoire quand  $\rho > 1$ . (Voir la discussion selon  $M/G/1$  des implications de ces propriétés. Ainsi, la preuve sera donnée après en (3.2.31)).

La  $n^{\text{ième}}$  étape des probabilités de transition  $P_{ij}^n$ , ( $i, j = 0, 1, 2, \dots$ ) de la chaîne de Markov  $\{Q_n\}$  sont obtenues comme étant des éléments de  $n^{\text{ième}}$  puissance de  $\mathbf{P}$ .

Pour les expressions analytiques de  $P_{ij}^n$ , il faut voir Takács (1962)[3], Prabhu and Bhat (1963)[4], and Prabhu (1965)[2].

### ***La distribution limite***

Soit  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$  est les probabilités limites définies par  $\pi_j = \lim_{n \rightarrow \infty} P_{ij}^n$ . On se base sur le théorème (3.1.1), la distribution limite existe si seulement si la chaîne de Markov est irréductible, apériodique et récurrente positive, c-à-d, quand  $\rho < 1$

Le théorème (3.1.2) nous montre la méthode pour déterminer cette distribution limite. Donc à partir de (3.1.12), on a les équations suivantes:

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad j = 0, 1, 2, \dots$$

$$\sum_{j=0}^{\infty} \pi_j = 1$$

En utilisant  $P_{ij}$  de la formule (3.2.4), on aura:

$$\begin{aligned}
\pi_0 &= \sum_{r=0}^{\infty} \pi_i \left( \sum_{r=i+1}^{\infty} b_r \right) \\
\pi_1 &= \pi_0 b_0 + \pi_1 b_1 + \pi_2 b_2 + \dots, \\
\pi_2 &= \pi_1 b_0 + \pi_2 b_1 + \pi_3 b_2 + \dots, \\
&\vdots \\
&\vdots \\
\pi_j &= \sum_{r=0}^{\infty} \pi_{j+r-r} b_r, \quad (j \geq 1)
\end{aligned}
\tag{3.2.7}$$

La meilleure méthode pour la détermination de la distribution limite semble être tout simplement la multiplication matricielle directe pour avoir  $P^n$  et faire accroître les valeurs de  $n$  jusqu'à ce que les rangées puissent être considérées raisonnablement identiques.

La technique suggérée pour  $M/G/1$  (voir (3.1.15)) ne marche pas à cause la structure triangulaire de la matrice  $P$ . On verra après dans la discussion de la file d'attente finie  $G/M/1/K$ , sauf si on commence avec un  $K$  assez large, et limitant l'espace d'état pour que la valeur finie retouche la dernière rangée de la matrice.

Dans ce cas, néanmoins le système d'équations (3.2.7) est facilement à résoudre par l'utilisation des méthodes de différence finie.

On définit l'opérateur de différence finie  $D$  comme suit :

$$D\pi_i = \pi_{i+1} \tag{3.2.8}$$

En utilisant (3.2.8), l'équation (3.2.7) peut être écrite de la manière suivante :

$$\pi_{j-1}(D - b_0 - Db_1 - D^2b_2 - D^3b_3 - \dots) = 0 \tag{3.2.9}$$

L'utilisation des méthodes de différence finie, une solution non-triviale de (3.2.9) est obtenue par la résolution de son équation caractéristique :

$$\begin{aligned}
(D - b_0 - Db_1 - D^2b_2 - \dots) &= 0 \\
D = \sum_{j=0}^{\infty} b_j D^j &\Leftrightarrow D = \beta(D) \tag{3.2.10}
\end{aligned}$$

D'où la solution de (3.2.10) satisfera l'équation fonctionnelle :

$$z = \beta(z) \tag{5.3.11}$$

Dans les formules (3.2.10) et (3.2.11), on a utilisé le fait que  $\beta(z)$  est la PGF de

$\{b_j, j = 0, 1, 2, \dots\}$

Pour obtenir les racines de (3.2.11), on considère les deux équations  $y=z$  et  $y = \beta(z)$   
Les intersections de ces deux équations nous donnent les racines requises.

On a ainsi les propriétés suivantes:

1.  $\beta(0) = b_0 > 0, \beta(1) = \sum_{j=0}^{\infty} b_j = 1, \beta'(1) = \frac{1}{\rho}$
2.  $\beta''(z) = 2b_2 + 6b_3z + \dots > 0, z > 0$

Par conséquent,  $\beta'(z)$  est monotone, croissante et par conséquent  $\beta(z)$  est convexe.

D'après les deux équations,  $y=z$  est la ligne droite passant par 0 et avec  $\beta(0) = 0$ ,  $\beta(1) = 1$  et  $\beta(z)$  est convexe, les équations  $y=z$  et  $y = \beta(z)$  font l'intersection à plus de deux fois, une à  $z=1$ . On pose  $\zeta_s$  la seconde racine. Si  $\zeta_s$  se situe à gauche où à droite de 1 est dépendant de la valeur de l'intensité du trafic  $\rho$ .

- Cas 1 : Si  $\rho < 1$ ,  $\beta'(1) > 1$ ; alors  $y = \beta(z)$  intersection avec  $y=z$  approchant au dessus de  $z=1$ . Mais  $b_0 > 0$ . Donc  $\zeta_s < 1$
- Cas 2 : Si  $\rho > 1$ ,  $\beta'(1) < 1$  alors  $y = \beta(z)$  intersection avec  $y=z$  approchant au dessus de  $z=1$ . Donc  $\zeta_s > 1$
- Cas 3 : Si  $\rho = 1$ , dans ce cas  $\beta'(1) = 1$  et  $y=z$  est la tangente de  $y = \beta(z)$  à  $z=1$ . Cela veut dire que  $\zeta_s = 1$

Soit  $\zeta_s$  est la racine inférieure positive. On a :  $\zeta_s < 1$  si  $\rho = 1$  et  $\zeta_s = 1$  si  $\rho \geq 1$ .

Cette solution est utilisée dans la solution de l'équation de différence finie (3.2.9).

On retourne en arrière vers l'équation de différence (3.2.9), on peut dire que:

$$\pi_j = c\zeta^j \quad (3.2.12)$$

Est une solution. Avec  $\zeta < 1$  et  $\sum_{j=0}^{\infty} \pi_j = 1$ , on aura :

$$\sum_{j=0}^{\infty} \pi_j = c \sum_{j=0}^{\infty} \zeta^j = \frac{c}{1-\zeta}$$

Nous donne :

$$c = 1 - \zeta$$

En fait la substitution de cette dernière dans la formule (3.2.12), on aura:

$$\pi_j = (1 - \zeta)\zeta^j, j = 0, 1, 2, \dots \quad (3.2.13)$$

Est la distribution limite de l'état du système dans la file d'attente  $G/M/1$ .

On note que  $\zeta$  est la racine de l'équation suivante:

$$z = \varphi(\mu - \mu z) \quad (3.2.14)$$

Dans plusieurs cas, la racine  $\zeta$  de (3.2.14) est déterminée en utilisant des techniques numériques. Avec la structure géométrique de la distribution limite de la formule de (3.2.13), la moyenne et la variance de nombre de clients dans le système, dit  $Q^A$ , sont facilement à obtenir. On a (l'indice  $A$  représente la limite de l'instant d'une arrivée)

$$L^A = E(Q^A) = \frac{\zeta}{1-\zeta}, \quad L_q^A = \frac{\zeta^2}{1-\zeta}, \quad v(Q^A) = \frac{\zeta}{(1-\zeta)^2} \quad (3.2.15)$$

Il est important de noter que l'analyse avec la chaîne de Markov induite nous donne les propriétés de nombre de clients dans le système aux instants des arrivées. Selon la discussion de la file  $M/G/1$ , les distributions limites du nombre de clients dans le système aux instants d'arrivées, aux instants de départs, et des instants arbitraires dans le temps sont les mêmes seulement quand l'occurrence des arrivées suit un processus de Poisson. Prabhu (1965a)[2] et Bhat (1968)[9] [1].

Écrire  $P_j = \lim_{t \rightarrow \infty} P[Q(t) = j]$ , où  $Q(t)$  est le nombre de clients à un instant arbitraire  $t$ , ces auteurs sont arrivés à l'expression explicite suivante de la distribution limite :

$$\{P_j, \quad j = 0, 1, 2, \dots\}$$

Quand :

$$P_0 = 1 - \rho \quad , \quad P_j = \rho(1 - \zeta)\zeta^{j-1}, \quad j \geq 1 \quad (3.2.16)$$

A partir de (5.3.16), les résultats au-dessous découlent la moyenne de la longueur de la file:

$$L = \frac{\rho}{1-\zeta} \quad , \quad L_q = \frac{\rho\zeta}{1-\zeta} \quad (3.2.17)$$

Comme exemple, en considère la file  $M/M/1$ . On pose :  $A(t) = 1 - e^{-\lambda t}, t \geq 0$ .

Donc, on a :

$$\varphi(\theta) = \frac{\lambda}{\lambda + \theta}, \quad \varphi(\mu - \mu z) = \frac{\lambda}{\lambda + \mu - \mu z}$$

Maintenant l'équation fonctionnelle (3.2.11) prend la forme suivante :

$$z = \frac{\lambda}{\lambda + \mu - \mu z} \Rightarrow -\mu z^2 + (\mu + \lambda)z - \lambda = 0 \quad (3.2.18)$$

Cette équation quadratique a deux racines, 1 et  $\rho = \lambda/\mu$ , en substituant  $\rho$  à la

place de  $\zeta$  dans (3.2.16) et (3.2.17), on a la distribution limite et les valeurs moyennes pour la queue  $M/M/1$ .

### *Temps d'attente*

Déterminer la distribution limite du temps d'attente d'un client, on a besoin de la distribution du nombre de clients dans le système à l'instant de l'arrivée d'un client.

La distribution limite trouvée dans la formule (3.2.13) est, en fait, la distribution d'un instant d'arrivée dans le système  $G/M/1$ . En outre, sa structure est la même que la distribution géométrique qu'on a pour  $M/M/1$ , avec  $\zeta$  qui prenne la place de  $\rho$  dans le résultat de la file d'attente  $M/M/1$ .

Les durées de service des clients dans le système sont exponentielles avec  $\mu$ , ainsi pareil aux durées de la queue  $M/M/1$ . Par conséquent les résultats du temps d'attente pour  $G/M/1$  ont les mêmes formes que ceux de  $M/M/1$  avec  $\zeta$  à la place de  $\rho$ . Sans y rentrer dans les détails de leur dérivation, on peut écrire :

$$F_q(t) = P(T_q \leq t) = 1 - \zeta e^{-\mu(1-\zeta)t}, \quad W_q = E(T_q) = \frac{\zeta}{\mu(1-\zeta)}, \quad V[T_q] = \frac{\zeta(2-\zeta)}{\mu^2} \quad (3.2.19)$$

Le temps  $T$  passé par un client dans le système est obtenu par l'ajout de la durée de service à  $T_q$ . On aura:

$$W = E(T) = E(T_q) + \frac{1}{\mu} = \frac{1}{\mu(1-\zeta)}, \quad V(T) = V(T_q + S) = \frac{1}{\mu^2(1-\zeta)^2} \quad (3.2.20)$$

### **III.5. Le système d'attente $G/M/s$**

L'analyse avec la chaîne de Markov induite pour le système d'attente  $G/M/1$  peut être facilement étendue à un système à multiple serveurs  $G/M/s$ . Avec la chaîne de Markov qui est définie aux instants des arrivées, la structure de processus est la même que celle de  $G/M/1$ , sauf à l'exception les probabilités de transition. Retenons les mêmes notations, pour la relation entre  $Q_n$  et  $Q_{n+1}$ , on aura :

$$Q_{n+1} = \begin{cases} Q_n + 1 - X_{n+1} & \text{si } Q_n + 1 - X_{n+1} > 0 \\ 0 & \text{si } Q_n + 1 - X_{n+1} \leq 0 \end{cases}$$

Où  $X_{n+1}$  est le nombre total des clients qui peuvent être servis par  $s$  serveurs durant la durée des inter-arrivées avec la distribution  $A(\cdot)$ .

Pour déterminer les probabilités de transition  $\{P_{ij}, i, j = 0, 1, 2, \dots\}$ , on a qu'à

considérer trois cas pour la valeur initiale  $i$  et la valeur finale  $j$ :  $i + 1 \geq J \geq s$ ,  $i + 1 \leq s$  et  $j \leq s$  et  $i + 1 > s$  mais  $j < s$ .

On note que, quand  $Q_n = i$ , la transition commence avec  $i+1$  et toujours  $j \leq i + 1$ . Avec les durées de service sont exponentiellement distribuées avec la densité

$(\mu e^{-\mu x}, x > 0)$ , la probabilité que le serveur complètera son service durant l'intervalle  $[0, t]$  est  $(1 - \mu e^{-\mu x})$  et la probabilité que le service continuera au-delà de  $t$  est  $\mu e^{-\mu x}$ .

En incorporant ces concepts avec les suppositions telles que les serveurs travaillent indépendamment entre eux, on aura les expressions suivantes de  $P_{ij}$  [1].

- **1<sup>er</sup> Cas** :  $i + 1 \geq J \geq s$

Cela représente  $i+1-j$  des fins de service durant une période d'inter-arrivée, quand tout les serveurs sont occupés. Voir la discussion selon  $M/M/s$  ; pour justifier que le taux de service est  $s$  :

$$P_{ij} = \int_0^\infty e^{-s\mu t} \frac{(s\mu t)^{i+1-j}}{(i+1-j)!} dA(t) \quad (3.2.21)$$

- **2<sup>ième</sup> Cas** :  $i + 1 \leq s$  et  $j \leq s$

Cette expression prene en compte l'événement dont  $i+1-j$  sur  $i+1$  clients qui complètent leurs service durant l'intervalle  $[0, t]$  si les  $j$  clients sont entrains de servir. A cause de l'indépendance des serveurs entre eux que chaque service est considéré comme étant un processus de Bernoulli et le résultat a une distribution binomiale avec la probabilité de succès qui est égale à :  $1 - \mu e^{-\mu x}$ .

$$P_{ij} = \binom{i+1}{i+1-j} \int_0^\infty (1 - e^{-\mu t})^{i+1-j} e^{-j\mu t} dA(t) \quad (3.2.22)$$

- **3<sup>ième</sup> Cas**:  $i + 1 > s$  mais  $j < s$

$$P_{ij} = \int_{t=0}^\infty \int_{\tau=0}^t e^{-s\mu\tau} \frac{(s\mu\tau)^{i-s}}{(i-s)!} s\mu \binom{s}{s-j} (1 - e^{-\mu(t-\tau)})^{s-j} e^{-j\mu(t-\tau)} d\tau dA(t) \quad (3.2.23)$$

Initialement,  $i+1-s$  clients complètent leurs service avec un taux  $s$ , et  $s-j$  en dehors de reste de  $s$  clients qui complètent leurs service indépendamment entre eux.

La matrice des probabilités de transition de la chaîne de Markov induite a la même structure que celle de la formule (3.2.4). A cause de la structure des valeurs de  $P_{ij}$  selon les cas 2 et 3, la solution de l'équation de différence finie donnée avant pour la distribution limite a besoin des modifications majeures [1].

### III.6. Le système d'attente $G/M/1/K$

Considérant le système d'attente  $G/M/1$  décrit avant avec capacité finie et que le système peut contenir seulement  $K$  clients à l'instant  $t$ . Étant donné que la chaîne induite est définie juste avant l'instant d'arrivée, le nombre de clients dans le système après l'instant d'arrivée est  $K$ , s'il y a  $K$  ou  $K-1$  avant cet instant  $t$ . S'il y a  $K$  avant l'arrivée d'un client qui n'est pas admis dans le système. Par conséquent à la place de (3.2.2), on a la relation :

$$Q_{n+1} = \begin{cases} \min(Q_n + 1 - X_{n+1}, K) & \text{si } Q_n + 1 - X_{n+1} > 0 \\ 0 & \text{si } Q_n + 1 - X_{n+1} \leq 0 \end{cases} \quad (3.2.24)$$

En utilisant les probabilités  $\{b_j, j = 0, 1, 2, \dots\}$  définies dans la formule (3.2.1), la matrice de probabilités de transition est donnée comme suit:

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \dots & K-1 & K \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ K-1 \\ K \end{matrix} & \begin{bmatrix} \sum_1^\infty b_r & b_0 & 0 & & & \\ \sum_2^\infty b_r & b_1 & b_0 & & & \\ \vdots & \vdots & \vdots & & & \\ \sum_K^\infty b_r & b_{K-1} & b_{K-2} & \dots & b_1 & b_0 \\ \sum_K^\infty b_r & b_{K-1} & b_{K-2} & \dots & b_1 & b_0 \end{bmatrix} \end{matrix}.$$

On note que les deux dernières rangées de la matrice  $P$  sont identiques car la chaîne de Markov effectivement commence avec  $K$  clients dans chaque des états  $K-1$  et  $K$ .

Soit  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$  est la distribution limite de la chaîne de Markov induite.

Écrire:  $\pi_j = \sum_{i=0}^K \pi_i P_{ij}$ , on a:

$$\begin{aligned} \pi_0 &= \sum_{i=0}^{\infty} \pi_i \left( \sum_{r=i+1}^{\infty} b_r \right) \\ \pi_1 &= \pi_0 b_0 + \pi_1 b_1 + \pi_2 b_2 + \dots + \pi_{K-1} b_{K-1} + \pi_K b_{K-1} \\ \pi_2 &= \pi_1 b_0 + \pi_2 b_1 + \pi_3 b_2 + \dots + \pi_{K-1} b_{K-2} + \pi_K b_{K-2} \\ &\vdots \\ &\vdots \\ \pi_{K-1} &= \pi_{K-2} b_0 + \pi_{K-1} b_1 + \pi_K b_1 \\ \pi_K &= \pi_{K-1} b_0 + \pi_K b_0 \end{aligned} \quad (3.2.26)$$



Si la valeur de  $K$  n'est pas aussi large, la résolution de ces équations simultanées en  $\{\pi_j, j = 0, 1, 2, \dots, K\}$ , avec la condition normalisée  $\sum_{j=0}^K \pi_j = 1$  est directement pratique dans les calculs.

Or, l'obtention de  $P^n$  pour les valeurs croissantes de  $n$  jusqu'à ce que les éléments de la rangée seront identiques, nous donne ainsi la distribution limite selon ces circonstances.

Faire en sorte de commencer avec la dernière équation de (3.2.26) et définir :

$$v_i = \frac{\pi_i}{\pi_{i-1}}, \quad i = 1, 2, \dots, K$$

On a :

$$\begin{aligned} \pi_i &= v_i \pi_{i-1} \\ \pi_i &= v_i v_{i-1} \pi_{i-2} \\ \pi_i &= v_i v_{i-1} \dots v_1 \pi_0 \end{aligned} \quad (3.2.27)$$

A partir de la dernière équation dans (3.2.26), on aura :

$$v_K = b_0 + v_K b_0 \Rightarrow v_K = \frac{b_0}{1 - b_0}$$

A partir de l'après dernière équation dans (3.2.26), on aura :

$$v_{K-1} = b_0 + v_{K-1} v_1 + v_K v_{K-1} b_1 \Rightarrow v_{K-1} = \frac{b_0}{1 - b_1 - v_K b_1}$$

Et ainsi de suite. Avec  $\sum_{j=0}^K \pi_j = 1$  et à partir de la formule (3.2.27), on aura:

$$(1 + v_1 + v_1 v_2 + \dots + v_1 v_2 \dots v_K) \pi_0 = 1$$

Et donc:

$$\pi_0 = [1 + \sum_{i=1}^K \prod_{r=1}^i v_r]^{-1} \quad (3.2.28)$$

Substituant ces dernières dans (3.2.27), on aura :  $\{\pi_j, j = 0, 1, 2, \dots, K\}$ .

### III.7. Le système d'attente $M/G^K/1$

On considère la file d'attente  $M/G^K/1$  où les clients passent au service par groupe avec la description suivante : Les clients arrivent un à un instant  $t$  avec un processus de Poisson de taux  $\lambda$ . Il y a un seul serveur qui assure un service à un groupe exactement de  $K$  clients à un instant  $t$ .

Les durées de service ont une distribution générale  $B(\cdot)$ . S'il y a moins de  $K$  clients qui attendaient dans la queue à la fin du service, le serveur va attendre jusqu'à  $t + K$  avoir

exactement le nombre  $K$  pour commencer son service.

On note qu'on introduit cette politique de supposition pour que l'analyse aille nous arranger énormément en termes de facilité d'étude de ce genre de système.

Les modifications de cette politique, comme tel le commencement du service avec un nombre spécifié de clients inférieur à  $K$ , vont porter des changements appropriés sur les expressions.

Soit  $\{Q_n, n = 0, 1, 2, \dots\}$  est le nombre de clients dans le système après le départ de  $n^{i\text{ème}}$  groupe. Soit  $X_n$  est le nombre de clients qui arrivent durant le  $n^{i\text{ème}}$  service. Et pour la distribution  $\{K_j, j = 0, 1, 2, \dots\}$  de  $X_n$  est donnée :

$$K_j = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^j}{j!} dB(t), \quad j = 0, 1, 2 \dots \quad (3.2.29)$$

On a ainsi la relation entre les variables aléatoires  $Q_n$  et  $Q_{n+1}$ ,

$$Q_{n+1} = \begin{cases} Q_n + X_{n+1} - K & \text{si } Q_n > K \\ X_{n+1} & \text{si } Q_n \leq K \end{cases} \quad (3.2.30)$$

La justification pour cette relation est exactement la même que celle qui a été donnée en (3.1.2) à l'exception que maintenant, on a besoin de  $K$  clients pour commencer le service au lieu d'un seul client. Pour  $P_{ij} = P(Q_{n+1} = j | Q_n = i)$ .

La formule (3.2.30) nous donne :

$$P_{ij} = \begin{cases} P(i + X_{n+1} - K = j) & \text{si } i > K \\ P(X_{n+1} = j) & \text{si } i \leq K \end{cases} = \begin{cases} K_{j-i+K} & \text{si } i > 0 \\ K_j & \text{si } i \leq 0 \end{cases} \quad (3.2.31)$$

On affiche ces probabilités sous la forme matricielle, on aura la matrice des probabilités de transition  $\mathbf{P}$  propre à la chaîne de Markov induite :

$$\mathbf{P} = \begin{matrix} & \mathbf{0} & \mathbf{1} & \dots \\ \mathbf{0} & \left[ \begin{array}{ccc} k_0 & k_1 & \dots \\ k_0 & k_1 & \dots \\ k_0 & k_1 & \dots \\ \vdots & \vdots & \vdots \\ k_0 & k_1 & \dots \\ k_0 & k_1 & \dots \\ \vdots & \vdots & \vdots \end{array} \right] \\ \mathbf{1} & & & \\ \mathbf{2} & & & \\ \vdots & & & \\ \mathbf{K} & & & \\ \mathbf{K} + \mathbf{1} & & & \\ \vdots & & & \end{matrix}.$$

On compare les deux formules (3.2.32) avec (3.1.5) et on note que (3.2.32) a  $K+1$  rangées identiques au lieu de deux comme dans la formule (3.1.5). Si on s'intéresse à l'expression mathématique de la distribution limite de la chaîne de Markov, on procède de

la même façon qu'on a procédé dans la file  $M/G/1$ .

Dans l'ordre de spécifier complètement la PGF de la distribution, il est nécessaire de spécifier les zéros de son dénominateur en utilisant le théorème de Rouché. Cependant, comme approche pratique, dans l'âge des ordinateurs, on peut utiliser la matrice  $P$  elle-même.

On note que les éléments  $K_j$  sont connus (peuvent être déterminés d'une manière numérique) et la distribution limite est donnée par  $\lim_{n \rightarrow \infty} P^n$ . Ainsi, rappelant que la matrice limite a des rangées identiques.

### III.8. Le système d'attente $G^K/M/1$

L'analyse avec la chaîne de Markov induite du processus de la longueur de la queue dans la file  $G^K/M/1$  suit la même méthode décrite avant par la considération de nombre de clients dans le système juste avant les instants d'arrivées. Soit  $A(\cdot)$  est la fonction de la distribution des instants des inter-arrivées et  $f(x) = \mu e^{-\mu x}$ ,  $x > 0$  est la distribution de la durée de service [1].

On suppose que les clients arrivent en groupe de taille  $K$ . (Si on suppose que les tailles de groupe sont variables, on a qu'à incorporer la distribution de la taille de groupe dans notre analyse).

Pour des raisons expliquées suivant la formule (3.2.2), on définit  $X_{n+1}$  comme étant le nombre des départs durant la  $(n+1)^{ième}$  période des inter-arrivées.

Soit  $\{Q_n, n = 0, 1, 2, \dots\}$  est le nombre de clients dans le système juste avant l'arrivée de  $n^{ième}$  groupe.

De la manière analogue à la formule (3.2.2), on a :

$$Q_{n+1} = \begin{cases} Q_n + K - X_{n+1} & \text{si } Q_n + K - X_{n+1} > 0 \\ 0 & \text{si } Q_n + K - X_{n+1} \leq 0 \end{cases} \quad (3.2.33)$$

Soit  $\{P(X_n = j) = b_j, j = 0, 1, 2, \dots\}$  étant donnée dans la formule (3.2.1) et pour les probabilités de transition  $P_{ij} = P(Q_{n+1} = j | Q_n = i)$ , à partir de la formule (3.2.33)

On aura :

$$P_{ij} = \begin{cases} P(i + K - X_{n+1} = j) & \text{si } j > 0 \\ P(i + K - X_{n+1} \leq 0) & \text{si } j = 0 \end{cases} = \begin{cases} b_{i+K-j} & \text{si } j > 0 \\ \sum_{r=i+K}^{\infty} b_r & \text{si } j = 0 \end{cases} \quad (3.2.34)$$

La matrice des probabilités de transition  $\mathbf{P}$  est la suivant:

$$\mathbf{P} = \begin{array}{c} \begin{array}{cccc} & 0 & 1 & 2 & \dots \\ 0 & \sum_{K}^{\infty} b_r & b_{K-1} & b_{K-2} & \dots \\ 1 & \sum_{K+1}^{\infty} b_r & b_K & b_{K-1} & \dots \\ & \vdots & & & \\ K & \sum_{2K}^{\infty} b_r & b_{2K-1} & b_{2K-2} \dots & \\ & \vdots & & & \vdots \end{array} \end{array} \Bigg[ \begin{array}{cccc} & & & \\ & & & \\ & & & \\ & & & \\ & & & \end{array} \Bigg].$$

Et pour avoir la distribution limite  $\pi = (\pi_0, \pi_1, \pi_2, \dots)$ , on peut utiliser la procédure standard de la résolution de l'équation  $\pi = \pi P$  et  $\sum_{j=0}^{\infty} \pi_j = 1$  avec les mêmes lignes illustrées dans la file  $G/M/1$ . Mais comme question pratique, les éléments de  $\mathbf{P}$  peuvent être déterminés numériquement à partir de la formule (3.2.1), en obtenant une approximation de  $\lim_{n \rightarrow \infty} P^n$  par la multiplication matricielle qui est vraisemblablement simple [1].

### III.9. Le système d'attente $G/G/1$

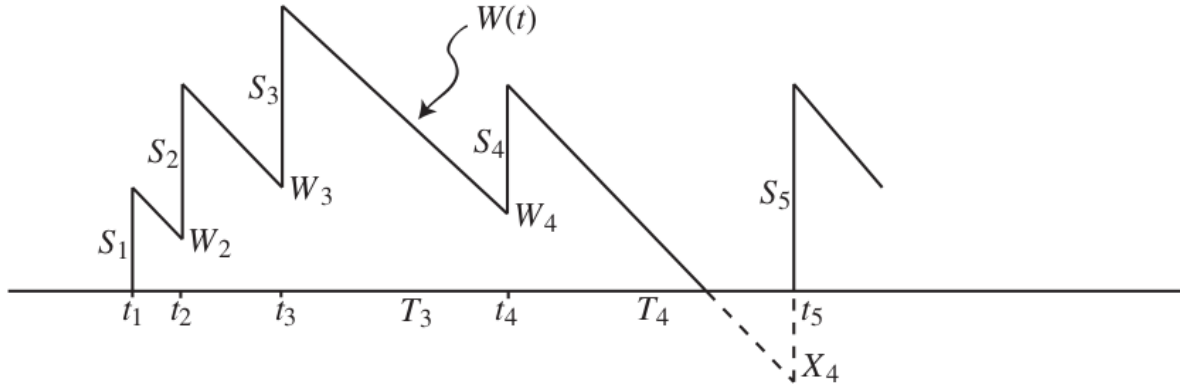
On considère la file générale  $G/G/1$  (Ainsi connu tel  $GI/G/1$  dans la littérature) avec la description suivante : les clients arrivent aux instants  $\{t_n, n = 0, 1, 2, \dots\}$ , et soient les durées des inter-arrivées  $\{T_n = t_{n+1} - t_n\}$  qui représentent des variables aléatoires indépendantes et identiquement distribuées avec la fonction de la distribution  $A(\cdot)$ . Soit aussi la durée de service du  $n^{\text{ième}}$  client notée par  $S_n$ , tel que  $\{S_n, n = 0, 1, 2, \dots\}$  sont des variables aléatoires indépendantes et identiquement distribuées avec la fonction de distribution  $B(\cdot)$ [1].

On représente les moyennes et les variances de ces variables aléatoires comme suit:

$$\begin{aligned} E(T_n) &= \frac{1}{\lambda}, & E(S_n) &= 1/\mu \\ V(T_n) &= \sigma_A^2, & V(S_n) &= \sigma_B^2 \end{aligned} \tag{3.3.1}$$

On note que pour les moyennes des temps des inter-arrivées et les temps de service, on a utilisé la même notation  $1/\lambda$  et  $1/\mu$  comme dans la queue  $M/M/1$ . On ainsi définit l'intensité du trafic .

Soit  $\{W_n, n = 0, 1, 2, \dots\}$  est le temps d'attente de  $n^{\text{ième}}$  client et  $W(t)$  est le temps d'attente d'un client à l'instant  $t$ . Avec l'arrivée ou non d'un client à l'instant  $t$ , le processus  $\{W(t)\}$  est un processus de temps d'attente virtuel. Cependant  $\{W_n, n = 0, 1, 2, \dots\}$  se sont que des temps d'attente actuels des arrivées aux instants  $(t_1, t_2, \dots)$ , et le processus  $W_n$  est un sous-ensemble de processus  $W(t)$ . Cela est illustré graphiquement dans la figure ci-dessous :



**Figure 9:** Les processus de temps d'attente  $W(t)$  et  $W_n$

Tel est montré dans la figure, on peut écrire les relations suivantes :

$$\begin{aligned}
 W_1 &= 0 \\
 W_2 &= W_1 + S_1 - T_1 \\
 W_3 &= W_2 + S_2 - T_2 \\
 W_4 &= W_3 + S_3 - T_3 \\
 W_5 &= 0 = W_4 + S_4 - T_4 + X_4 \quad (3.3.2)
 \end{aligned}$$

Dans l'écriture de ces relations, on a utilisé le fait que, entre deux arrivées successives, le processus  $W(t)$  décroît d'une unité de taux a cause d'un service offert à un client. Il sera clair si on interprète  $W_n$  comme un taux service du système juste avant une arrivée à l'instant  $t_n$ , et pour un service offert, le taux  $W_n + S_n$  sera réduit d'une unité jusqu'à une arrivée à l'instant  $t_{n+1}$ , quand sa valeur est égale à  $W_n + S_n - T_n$ . Quand cette quantité devienne négative (par une grandeur  $X_n$ ), pour rendre  $W_{n+1} = 0$ , on écrit:  $W_n + S_n - T_n + X_n$ .

D'où, la généralisation de (3.3.2) :

$$W_{n+1} = \begin{cases} W_n + S_n - T_n & \text{si } W_n + S_n - T_n > 0 \\ 0 & \text{sinon} \end{cases} \quad (3.3.3)$$

Ou bien :

$$W_{n+1} = W_n + S_n - T_n + X_n \quad (3.3.4)$$

Tel-que  $X_n$  est définie comme suit:

$$X_n = -\min(0, W_n + S_n - T_n) . \quad (3.3.5)$$

En utilisant la relation (3.3.3), on peut écrire:

$$\begin{aligned} P(W_{n+1} \leq t) &= P(W_{n+1} = 0) + P(0 < W_{n+1} \leq t) \\ &= P(W_n + S_n - T_n \leq t) + P(0 < W_n + S_n - T_n \leq t) \\ &= P(W_n + S_n - T_n \leq t) \end{aligned} \quad (3.3.6)$$

On définit :  $F_n(t) = P(W_n \leq t)$ ,  $S_n - T_n = U_n$ , et  $U_n(t) = P(U_n \leq t)$ .

Avec cette notation, (3.3.6) peut être écrite de cette manière :

$$F_{n+1}(t) = \int_{-\infty}^t F_n(t-x) dU_x, \quad 0 \leq t < \infty \quad (9.1.7)$$

Pour l'existence d'un état récurrent, il faut que l'intensité du trafic  $\rho < 1$ . C'est la même chose que  $E(U_n) = E(S_n) - E(T_n) < 0$ . Selon cette condition, on enlève les indices notés dans la formule (3.3.7),

On aura:

$$F_{n+1}(t) = \int_{-\infty}^t F_n(t-x) dU_x \quad (3.3.8)$$

Où

$$U(x) = \int_x^{\infty} B(y) dA(y-x) \quad (3.3.9)$$

L'équation (3.3.8) est la première établie par Lindley (1952)[14][1]. Elle est l'une des équations fondamentales dans la théorie des files d'attente. Malheureusement, sa solution demande d'utiliser la méthode de Wiener-Hopf, dont elle est bien illustrée dans Kleinrock (1975)[10][1].

Au lieu de la description de  $W_n$ , on regarde maintenant à sa moyenne. Quand  $t \rightarrow \infty$ , on peut écrire  $E(W_{n+1}) = E(W_n)$ . On enlève les indices et prendre les suppositions de (3.3.4), on aura :

$$E(S) - E(T) = E(U) = -E(X)$$

Avec  $X$  qui est la longueur de la période libre dite I, qui finira avec une arrivée dont elle trouve le système est vide, d'où on peut écrire:

$$P \text{ (une arrivée trouve le système est vide)} = a_0 \quad (3.3.11)$$

$E(X) = E(I) * P$  (une arrivée trouve le système est vide).

On note que la probabilité de (3.3.11) est  $a_0$ . Et on a :

$$E(I) = \frac{E(X)}{a_0} = -\frac{E(U)}{a_0} = \frac{1-\rho}{\lambda a_0} \quad (3.3.12)$$

Retourner à la formule (3.3.4) et réécrire, on a :

$$W_{n+1} - X_{n+1} = W_n + U_n \quad (3.3.13)$$

Mettre l'égalité de (3.3.13) au carré et prendre les suppositions suivantes :

$$E(W_{n+1}^2) + E(X_{n+1}^2) - 2E(W_{n+1}X_n) = E(W_n^2) + E(U_n^2) + 2E(W_nU_n)$$

On observe que  $E(X_{n+1}^2) = E(W_n^2)$  quand  $n \rightarrow \infty$ ,  $W_n$  et  $U_n$  sont indépendantes, et  $X_nW_{n+1} = 0$ . Par conséquent, quand  $n \rightarrow \infty$ , on a :

$$E(X^2) = E(U^2) + 2E(W)E(U) \Rightarrow E(W) = \frac{E(X^2) - E(U^2)}{2E(U)} \quad (3.3.14)$$

On définit  $E(X^2)$  de la même façon que (3.3.11), et on peut écrire :  $E(X^2) = a_0 E(I^2)$ .

A partir de la formule (3.3.10), on aura ainsi :

$$\begin{aligned} E(U) &= \frac{1}{\mu} - \frac{1}{\lambda} \\ [E(U)]^2 &= \frac{(1-\rho)^2}{\lambda^2} \\ V(U) &= \sigma_A^2 + \sigma_B^2 \\ E(U^2) &= V(U) + [E(U)]^2 = \sigma_A^2 + \sigma_B^2 + \frac{(1-\rho)^2}{\lambda^2} \end{aligned} \quad (3.3.15)$$

Réécrire la formule (3.3.14) comme suit :

$$E(W) = \frac{E(X^2) - E(U^2)}{2E(U)}$$

Et en utilisant les formules (3.3.12) et (3.3.15), on aura :

$$E(W) = \frac{a_0 E(I^2)}{2[-a_0 E(I)]} - \frac{\sigma_A^2 + \sigma_B^2 + \frac{(1-\rho)^2}{\lambda^2}}{2\left(\frac{1}{\mu} - \frac{1}{\lambda}\right)} \Rightarrow E(W) = \frac{\lambda^2(\sigma_A^2 + \sigma_B^2) + (1-\rho)^2}{2\lambda(1-\rho)} - \frac{E(I^2)}{2E(I)} \quad (3.3.16)$$

Ce résultat nous mène à une borne supérieure de  $E(W)$  dans la file générale  $G/G/1$ .

L'expression (3.3.16) pour  $E(W)$  inclue  $E(I^2)$  et  $E(I)$ , dont on ne peut la déterminer sans faire une analyse complète de système. Néanmoins, pour obtenir une borne inférieure

de  $\frac{E(I^2)}{2E(I)}$  (dans l'ordre d'avoir une borne supérieure de  $E(W)$ , on procède de la manière suivante :

Mettre  $a_0 = 1$  dans cette égalité  $E(I) = -\frac{E(S-T)}{a_0}$  de la formule (3.3.12), nous donnera :

$$E(I) > \frac{1}{\mu} - \frac{1}{\lambda} \quad (3.3.17)$$

Ainsi,

$$E(I^2) = V(I) + [E(I)]^2$$

Avec  $V(I)$  est une quantité positive,

$$E(I^2) \geq [E(I)]^2 \quad (3.3.18)$$

En utilisant ces deux résultats dans la formule (3.3.16), on aura:

$$E(W) = \frac{\lambda^2(\sigma_A^2 + \sigma_B^2)}{2\lambda(1-\rho)} + \frac{1-\rho}{2\lambda} - \frac{E(I^2)}{2E(I)} \leq \frac{\lambda^2(\sigma_A^2 + \sigma_B^2)}{2\lambda(1-\rho)} + \frac{(1-\rho)}{2\lambda} - \frac{E(I)}{2}$$

Nous donne :

$$E(W) \leq \frac{\lambda^2(\sigma_A^2 + \sigma_B^2)}{2\lambda(1-\rho)} \leq \frac{\lambda(\sigma_A^2 + \sigma_B^2)}{2(1-\rho)} \quad (9.1.19)$$

Ces résultats sont dus à Kingman (1962) [11] et Marshall (1968)[12]. Ils ont donné ainsi des bornes inférieures. Malheureusement, les bornes inférieures ont été données par ces auteurs ne sont pas faciles de les obtenir. Une simple borne inférieure a été donnée par Marchal (1978) [13] comme suit :

$$E(W) \geq \frac{\rho^2 + \lambda^2 \sigma_B^2 - 2\rho}{2\lambda(1-\rho)} \quad (9.1.20)$$

Dans le cas de la file  $G/G/s$  à multiple serveurs, les bornes de  $E(W)$  deviennent très compliquées. Le seul résultat mentionné ici est celui de Kingman (1962), dont on a la forme suivante:

$$E(W) \leq \frac{\lambda(\sigma_A^2 + \sigma_B^2) + (s-1)\rho/\mu}{2s(1-\rho)} \quad (9.1.21)$$

Où  $\rho = \lambda/s\mu$  est l'intensité du trafic.

La relation de la formule (9.1.3) entre  $W_n$  et  $W_{n+1}$  établie la propriété du processus de Markov  $\{W_n, n = 1, 2, \dots\}$

C'est un processus de Markov à état continu et à un temps discret et toutes les



techniques appliquées pour les processus de Markov peuvent être utilisées pour son analyse.

Dans l'approximation d'un processus sous-jacent, on peut utiliser un simple processus pour faciliter l'analyse si les propriétés originales sont maintenues le plus possible. Un exemple d'application est l'approximation du trafic dense dans la file générale G/G/1.

La relation (9.1.3) entre  $W_{n+1}$  et  $W_n$  peut être située comme suit :

$$W_{n+1} = \max(0, W_n + S_n - T_n) = \max(0, W_n + U_n) \quad (9.3.1)$$

Où  $U_n = S_n - T_n$ ,  $n = 0, 1, 2, \dots$  et pour  $n = 0, 1, 2, \dots$ , on a :

$$W_1 = \max(0, W_0 + U_0)$$

$$W_2 = \max(0, W_1 + U_1)$$

$$W_3 = \max(0, W_2 + U_2)$$

Par conséquent, on a  $W_2 > 0$  seulement si  $U_1 > 0$  (Notant que on a supposé que le premier client entre dans une file vide; autre, il faut que  $W_1 + U_1 > 0$ ;  $W_3 > 0$  si seulement si  $W_2 + U_2 > 0$  ; et ainsi de suite. Quand le trafic est dense, on peut supposer que le taux des arrivées et le taux de service sont égaux dans le voisinage.

Soit  $\left(\frac{1}{\lambda} - \frac{1}{\mu} = \alpha\right)$  et  $\sigma_A^2 + \sigma_B^2 = \sigma$  en donnant  $E(U) = E(U_n) = -\alpha$  et  $V(U) = V(U_n) = \sigma^2$ .

Selon le trafic dense, on peut écrire (En utilisant  $\sim$  pour indiquer une équivalence approximative)

$$W_2 \approx U_1$$

$$W_2 \approx U_1 + U_2$$

$$\vdots$$

$$W_{n+1} \approx \sum_{r=1}^n U_r = U^{(n)} \quad (9.3.2)$$

On a :  $\{U^{(n)}, n = 1, 2, \dots\}$  sont connues comme des sommes partielles de  $\{U_n\}$ , et on

a :

$$E[U^{(n)}] = -n\alpha, \quad V[U^{(n)}] = n\sigma^2 \quad (9.3.3)$$

Avec  $\{U_n, n = 1, 2, \dots\}$  qui sont des variables aléatoires indépendantes et identiquement distribuées, pour  $n$  est large, en utilisant le théorème de centrale limite d'où on peut écrire :

$$\frac{U^{(n)} + n\alpha}{\sigma\sqrt{n}} \rightsquigarrow N(0, 1) \quad (9.3.4)$$

On indique que à gauche de l'équation (9.3.4) a une distribution normale avec une moyenne zéro et une variance un.

Quand  $\alpha/\sigma$  est petit et quand  $n \rightarrow \infty$ , le temps d'attente  $W_n$  a approximativement une distribution exponentielle avec une moyenne  $\sigma^2/2\alpha$ .

Mais :

$$\frac{\sigma^2}{2\alpha} = \frac{1}{2} \left( \frac{V(U)}{-E(U)} \right) = \frac{1}{2} \left[ \frac{V(T)+V(S)}{E(T)-E(S)} \right] \quad (9.3.5)$$

On se réfère à la formule (9.1.19), on note que ce résultat est exactement la borne supérieure pour  $E(W)$ . Nous devons mettre l'accent sur le fait que le résultat (9.3.5) est une approximation de la durée moyenne d'attente dans la file  $G/G/1$  avec un trafic dense. Et ce résultat est uniquement utile pour des grandes valeurs de l'intensité du trafic.

Dans le cas de la file d'attente  $G/G/s$  à multiple serveurs, Kingman suggère deux d'approximations possibles pour le temps moyen d'attente. Le premier résultat approximatif est obtenu par l'extension de l'approximation au temps moyen d'attente selon le trafic dense dans la file  $G/M/s$  :

$$E(W) \approx \frac{V(T)+V(S/s)}{2[E(T)-E(S/s)]} \quad (9.3.6)$$

Ce résultat peut être ainsi obtenu par considération la performance de la file d'attente  $G/G/s$  dans le trafic dense comme étant approximativement le même que celui de  $G/G/1$ .

La seconde approximation est obtenue par considération de la performance de la file  $G/G/s$  dans le trafic dense comme étant la même que l'ensemble de  $s$  files  $G/G/1$  en parallèle qui sont nourri par un processus d'arrivée avec le temps moyen d'inter-arrivées  $sE(T) = s/\lambda$

$$E(W) \approx \frac{sV(T)+V(S)}{2[sE(T)-E(S)]} = \frac{V(T)+sV(S/s)}{2[E(T)-E(S/s)]} \quad (9.3.7)$$

Effectivement, la formule (9.3.7) est large que la formule (9.3.6) par  $(s-1)V(S/s)$ , et par conséquent, ces résultats doivent être utilisés avec caution.

Cela nous ramène à prendre en considération deux types de trafic lorsque on traite une file de type  $G/G/1$  : le trafic au voisinage de la saturation (heavy traffic) lorsque tend vers 1 et le trafic normal ou le bas trafic (low or normal traffic) lorsque est

largement inférieur à 1.

Les résultats déterminés pour cette file représentent des résultats globaux et valables pour toutes les autres files ( $M/M/1$ ,  $M/G/1$ ,  $M/G/k$ , . . .). Toutefois, ces résultats sont moins précis que les résultats originaux propres à ces files.

### ***III.10. Conclusion***

Dans ce chapitre, nous avons étudié des systèmes d'attente non markoviens et nous avons présenté quelques caractéristiques statistiques ou on se contente juste d'une distribution limite de chaque système en utilisant une analyse bien détaillée avec les chaînes de Markov induites et nous sommes arrivés à présenter une approximation avec une borne inférieure et supérieure du temps d'attente d'un client dans le système  $G/G/1$ .

# CHAPITRE IV

APPLICATION ET SIMULATION D'UN MODELE  
NON-MARKOVIAN

# CHAPITRE IV

## APPLICATION ET SIMULATION D'UN MODELE NON-MARKOVIEN

---

### *IV.1. Introduction*

Après l'étude des systèmes d'attente non-markoviens, nous passons maintenant à la conception et la réalisation d'une application simulant ce type de systèmes. Pour Cela, nous avons adopté l'approche de programmation orientée objet, permettant une abstraction plus réaliste du domaine d'application et une maintenance (corrective et évolutive) du programme.

### *IV.2. Phase d'Analyse des Besoins*

#### **IV.2.1. Description de l'Application**

L'application à développer, dite *QueueSimulator*, représente un simulateur des systèmes de files d'attente, que ce soit markoviens ou non markoviens, et ceci grâce à une interface graphique intuitive et ergonomique permettant de créer un modèle de file d'attente par de simples clics, de paramétrer les arrivées et les services par des lois de probabilités paramétrables, et de régler les paramètres de simulation (durée de la simulation, le nombre d'itérations, etc.)

#### **IV.2.2. Cahier des Charges**

##### *IV.2.2.1. Besoins fonctionnels*

L'application à développer doit répondre aux besoins fonctionnels suivant :

- Création d'un modèle de files d'attente, soit pour les stations simples, stations à plusieurs serveurs ou bien un réseau de files d'attentes
- Pouvoir enregistrer un modèle de file d'attente et de pouvoir l'ouvrir pour d'éventuelles modifications ou lancement de simulation
- Modification de paramètres de simulation : la durée maximale de simulation,

le nombre d'itérations, la vitesse de simulation.

- Affichage des résultats d'une simulation par itération ainsi que la moyenne des toutes les itérations.
- Offrir un moyen pour affecter des lois de probabilités aux flux d'arrivée et aux serveurs afin de calculer les dates d'arrivée et de départs.
- Offrir un moyen pour paramétrer une loi de probabilité.
- Affectation des lois de probabilités pour les arrivées et les services.

#### ***IV.2.2.2. Besoins non-fonctionnels***

Les besoins non-fonctionnels représentent les contraintes techniques et non-techniques nécessaires afin de réaliser les besoins fonctionnel :

- Réaliser les différentes fonctionnalités avec un temps de latence raisonnable
- Utiliser un format de représentation de données général, comme par exemple le format XML.
- Fonctionnement sous-réseaux permettant un accès distant vers des modèles créer par autrui, et permettre ainsi un travail collaboratif.

#### **IV.2.3. Diagramme des cas d'utilisation**

Un diagramme de cas d'utilisation<sup>1</sup> permet de mettre en évidence les différents acteurs de notre application ainsi que les fonctionnalités offertes pour chaque acteur. Pour notre cas, ce diagramme est illustré par la *figure 10*.

#### **IV.2.4. Diagrammes de Séquence Globale**

Dans la phase d'analyse des besoins, pour chaque cas d'utilisation on donne un ou plusieurs digrammes de séquence<sup>2</sup> globaux dans lesquels on représente l'application comme une boîte noire et on représente les communications entre l'acteur et le système par des messages.

Ces diagrammes de séquence, seront le point de départ pour la conception, en établissant des diagrammes de séquence détaillés permettant l'établissement des diagrammes de classes partielles.

En guise de simplification, nous allons présenter quelques diagrammes de séquence

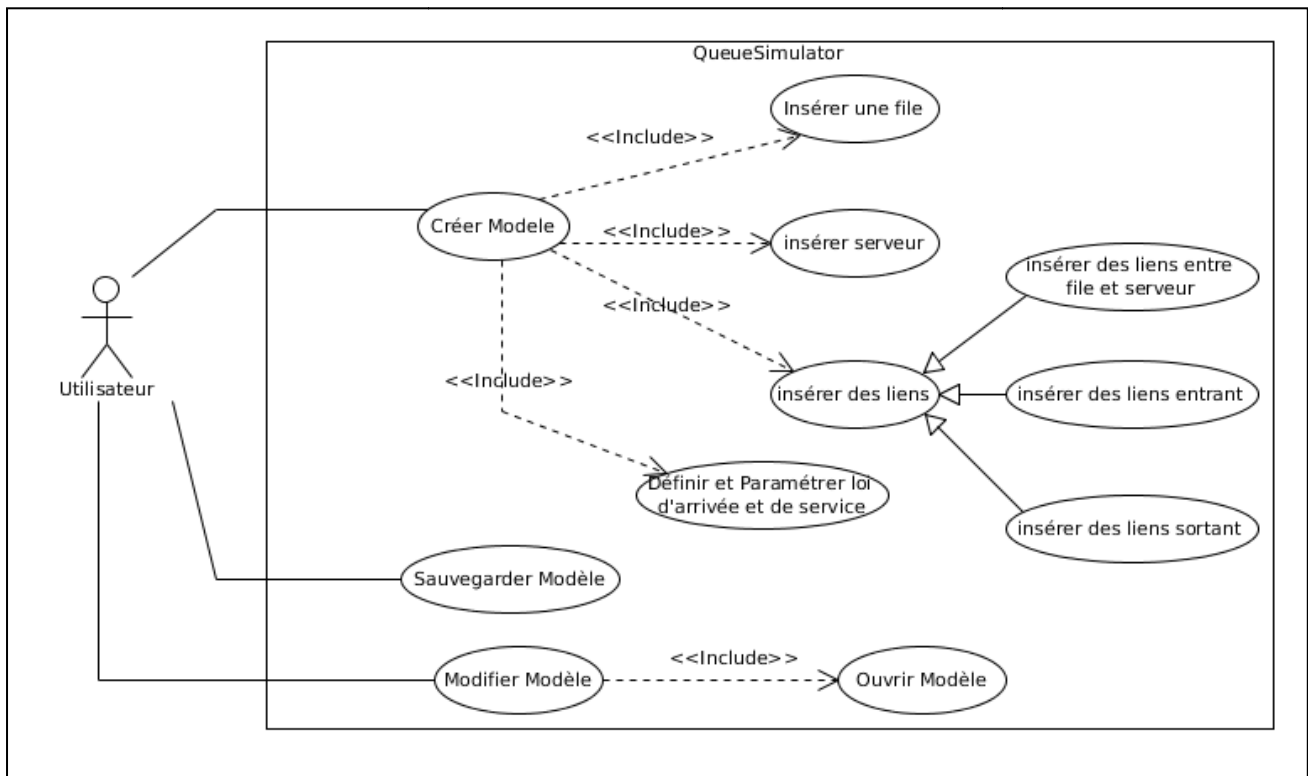
---

<sup>1</sup> Un cas d'utilisation représente une fonctionnalité à fournir à un acteur de l'application

<sup>2</sup> Un diagramme de séquence permet de décrire un ou plusieurs scénarios

---

englobant plusieurs cas d'utilisation en utilisant les différentes constructions syntaxiques comme : *opt*, *loop*, *alt*, *etc.*



*Figure 10 : Diagramme de cas d'utilisation du QueueSimulator.*

### IV.3. Phase de Conception

#### IV.3.1. Diagramme de Package

Le diagramme de package illustre le regroupement de classes dans des ensembles en respectant des relations logiques. Les classes appartenant à un même package sont logiquement reliées d'une façon à coopérer pour réaliser des fonctionnalités.

Pour notre application, le diagramme de package que nous avons adapté est illustré par la *figure 12*.

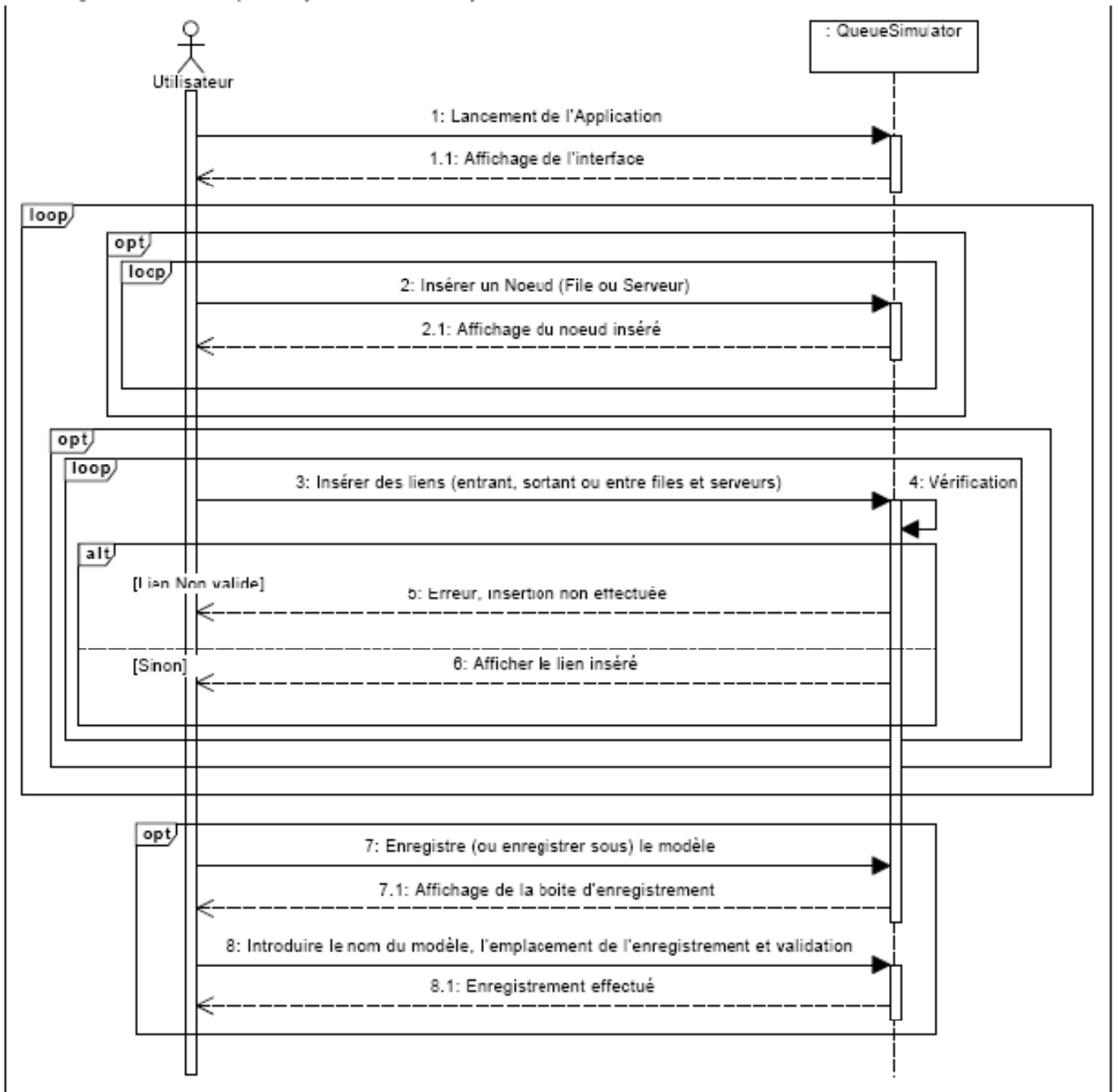


Figure 11 : Diagramme de Séquence pour le cas « Créer un Modèle »



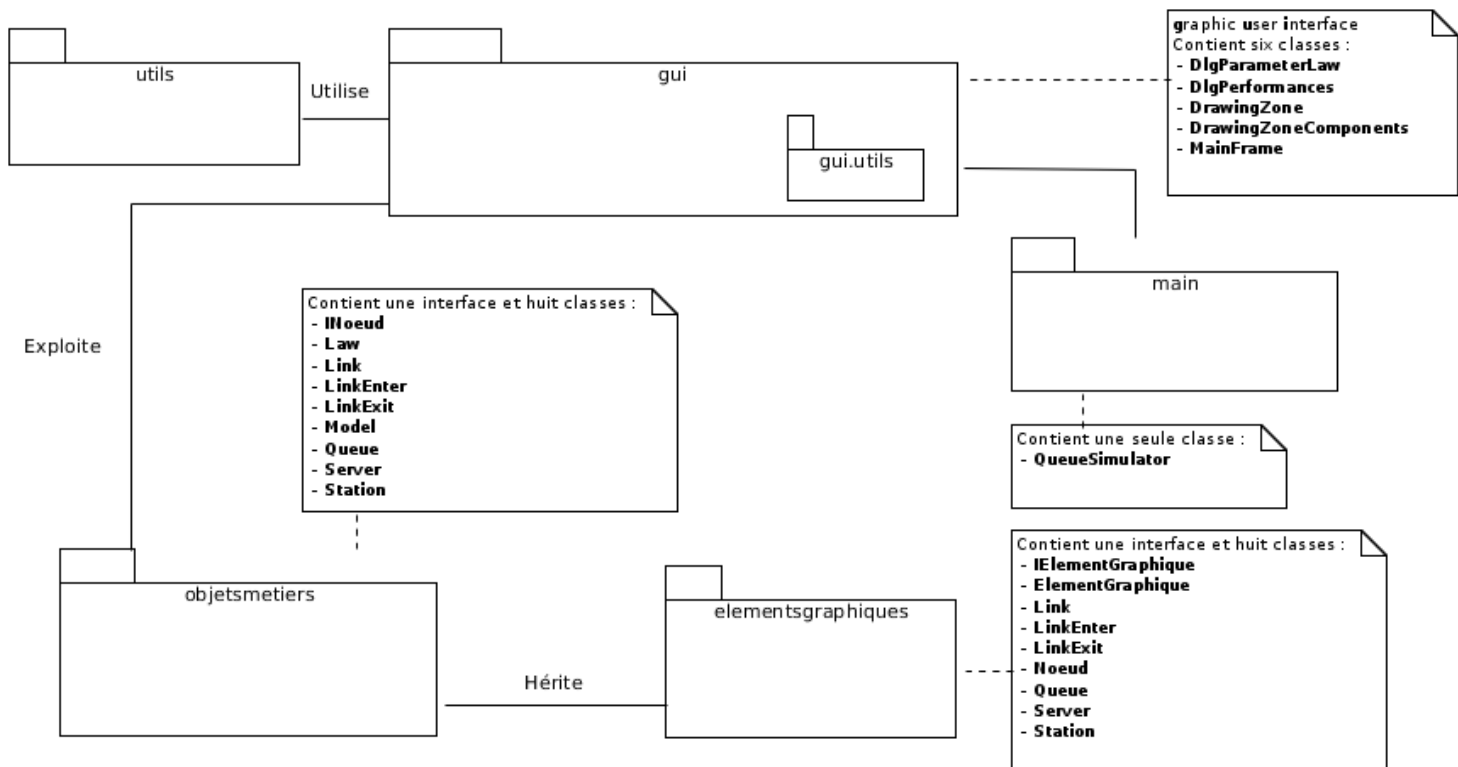


Figure 12 : Diagramme de Package de l'application QueueSimulator

### IV.3.1. Diagramme de classes

Chaque package contient une ou plusieurs classes logiquement liées. Nous nous focalisons sur le package *objetsmetiers* contenant uniquement les classes du domaine d'application à savoir les files d'attente. La figure dans la page suivante illustre les classes de ce package. Voici quelques classes de ce package :

- La classe *Model* : cette classe représente un modèle de files d'attente,
- La classe *Station* : représente une station de file d'attente (une file avec un ou plusieurs serveurs)
- La classe *Queue* : la file d'attente elle même.
- La classe *Server* : l'entité fournissant le service.
- La classe *Law* : représente une loi de probabilité
- La classe *Link* : représente un lien entre une file et un serveur.
- Les classes *LinkEnter* et *LinkExit* : respectivement, pour les arrivées et les départs.

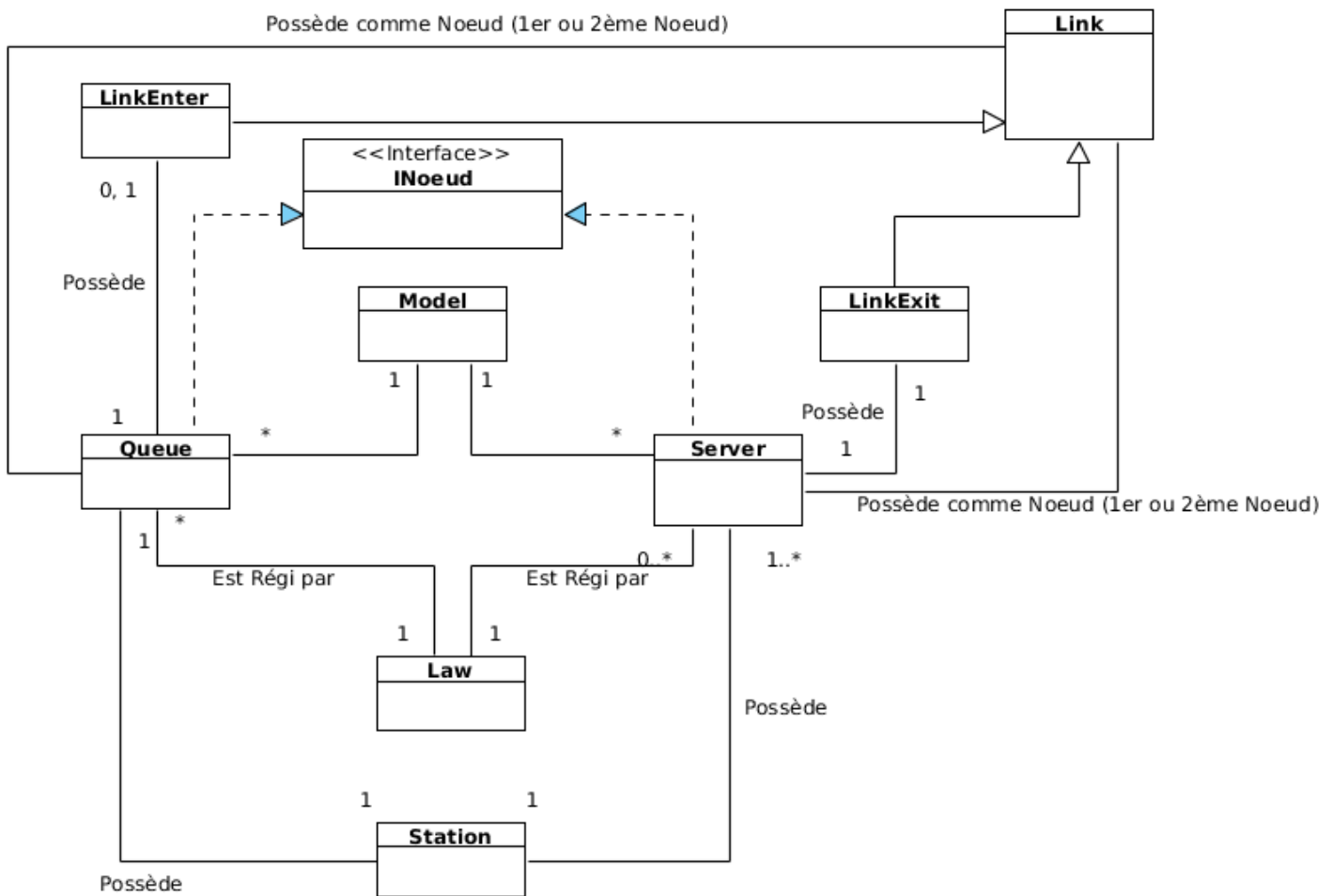


Figure 13 : Diagramme de Classe de l'Application « QueueSimulator »

## IV.4. Phase de Réalisation

### IV.4.1. Choix Techniques

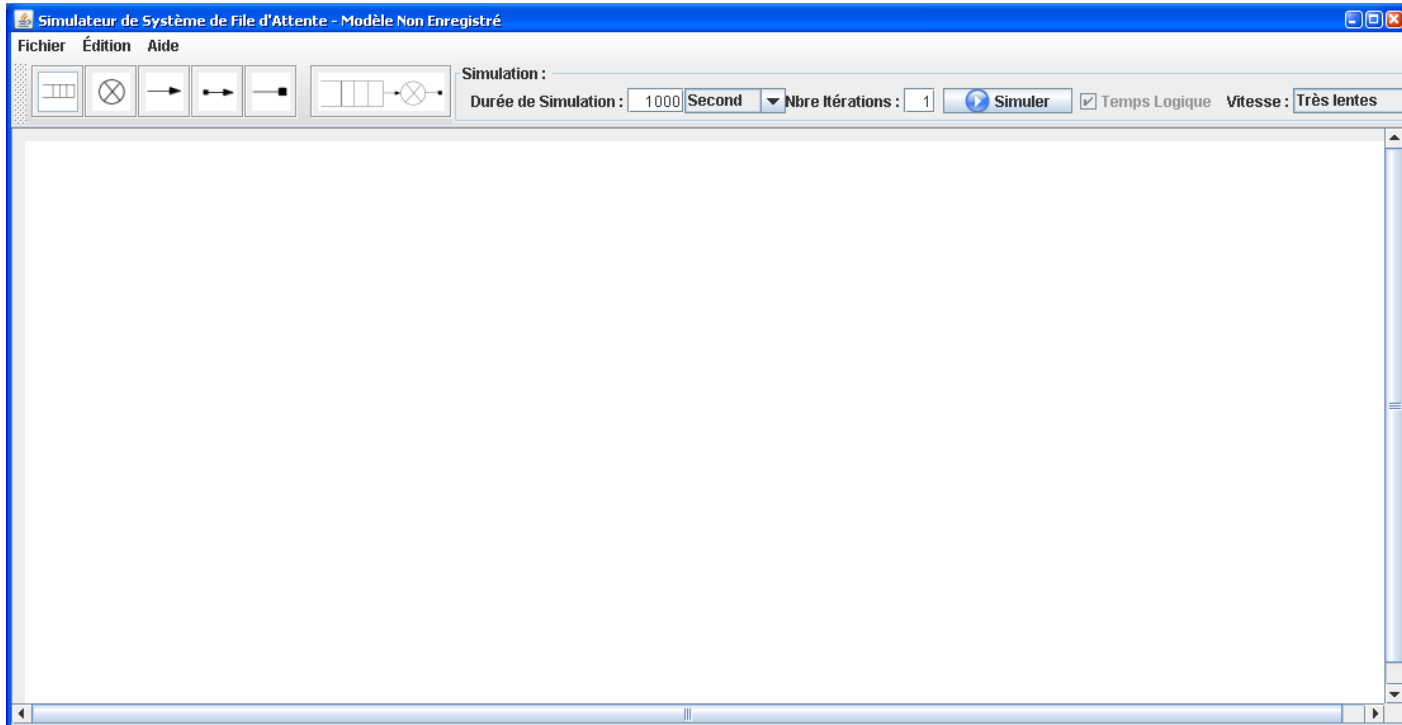
Pour le développement de l'application nous avons procédé aux choix techniques suivants :

- ✓ Environnement de Développement Eclipse ;
- ✓ Langage de programmation Java 2 (la version utilisée est 1.6) ;
- ✓ les interfaces graphiques avec Swing<sup>3</sup> ;
- ✓ langage XML pour la sauvegarde des modèles et des lois de probabilités ;

<sup>3</sup> Swing est une bibliothèque Java permettant la création des interfaces graphique (Fenêtres, Dialogues, boutons, zones de textes, etc.)

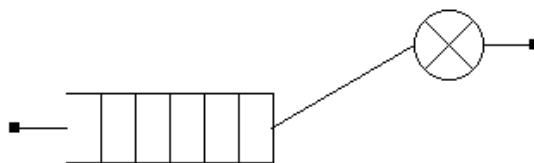
## IV.4.2. Interfaces Homme-Machine

### IV.4.2.1. Fenêtre Principale



**Figure 14** : Fenêtre Principale de « QueueSimulator »

A travers cette fenêtre, nous allons créer un modèle de file d'attente et le paramétrer. Par exemples, dans la figure suivante on voit un modèle simple de file d'attente :



**Figure 15** : Modèle d'Attente sans paramètres

### IV.4.2.2. Boite de Dialogue des lois de probabilité

Le modèle illustré dans la figure ci-dessus ne possède pas les lois régissant les arrivées et les départs. L'application offre une boite de dialogue pour pouvoir définir ces lois. Dans la *figure 16*, nous illustrons les deux boites de dialogue pour la sélection et le paramétrage de loi de probabilité.

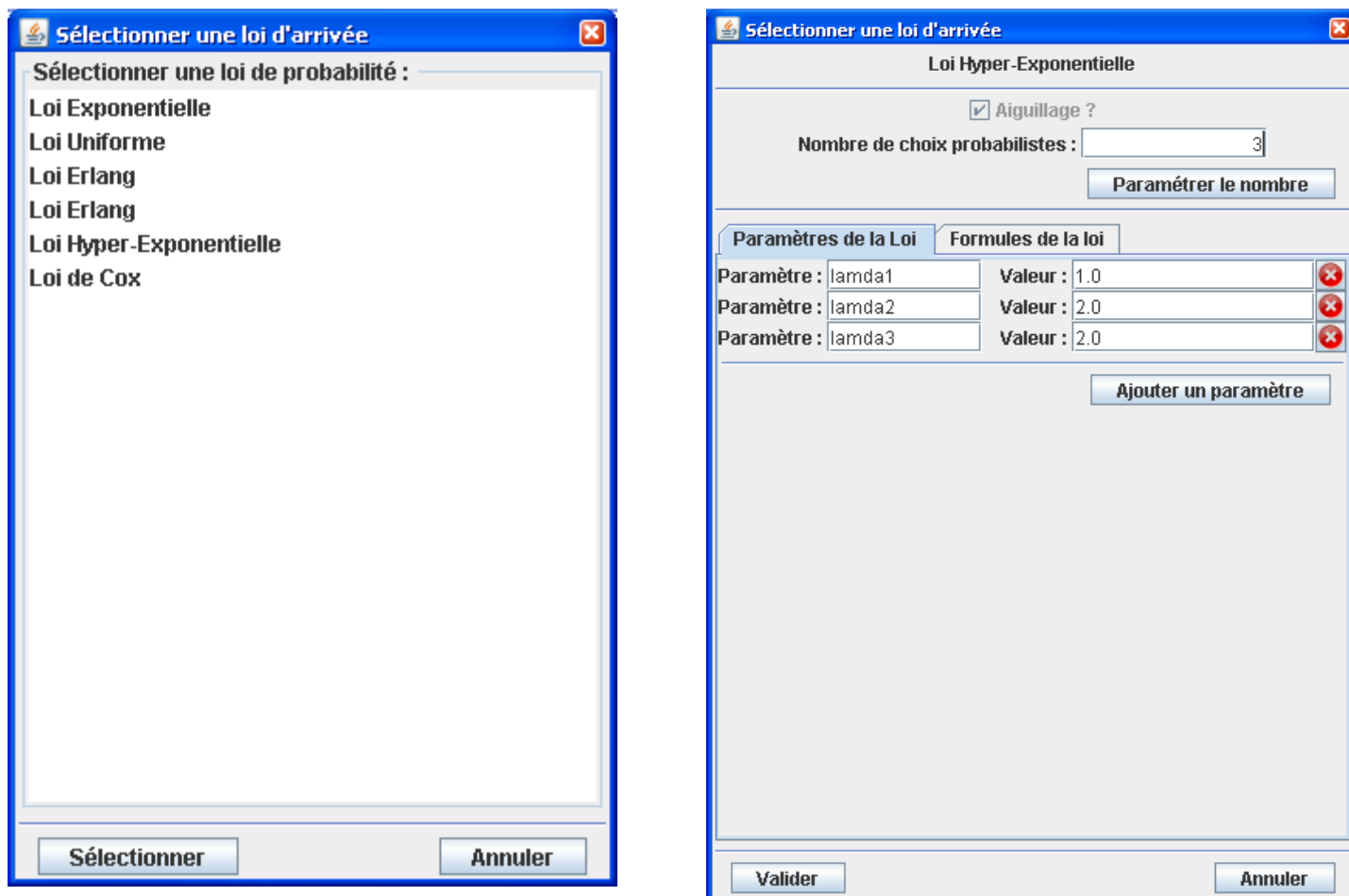


Figure 16 : Les deux boites de dialogue pour la définition de loi

Un fois les lois sont définies, l'application affiche le résultat suivant :

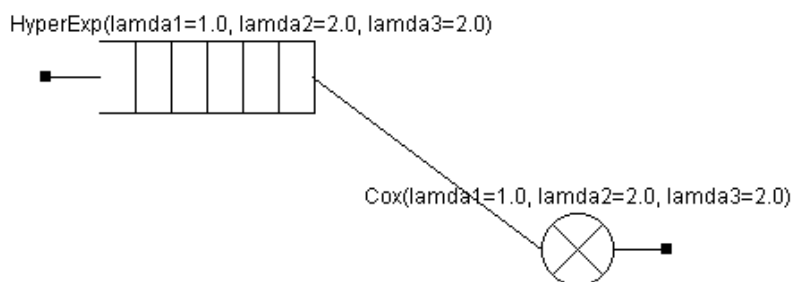
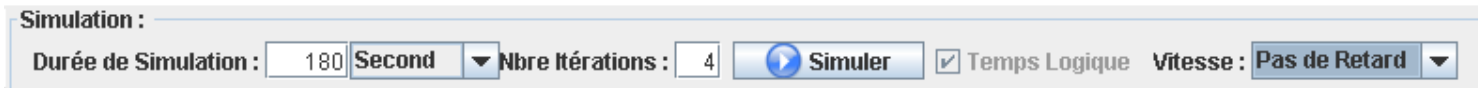


Figure 17 : Modèle d'Attente avec paramètres des lois

Bien évidemment, on peut à tout moment modifier soit les paramètres d'une loi de probabilité prudemment définie, ou bien définir carrément une autre loi.

### IV.4.2.3. Boîte de Dialogue des Résultats

Avant de lancer la simulation, on doit tout d'abord procéder au paramétrage de cette dernière, comme la durée maximale, le nombre d'itération, et sa vitesse, et ceci grâce à la barre d'outil de la fenêtre principale.



*Figure 18 : Paramètres de simulation.*

Après simulation, on aura la boîte dialogue suivante



*Figure 19 : Boîte de dialogue pour les résultats de Simulation.*

***IV.6. Conclusion***

Dans ce dernier chapitre, nous avons présenté la conception et la réalisation de notre application de simulation, dite QueueSimulator, de modèles non-markoviens. Ceci nous a permis de développer une interface graphique intuitive et interactive, permettant aux utilisateurs de créer des modèles de files d'attente d'une façon simplifiée et avec de simples clics. En plus de ça, le paramétrage des différentes lois de distribution, pour les arrivées et pour les services, nous permet d'obtenir différents résultats pour un même modèle.

# CONCLUSION GENERALE

# CONCLUSION GENERALE

---

Nous avons vu des concepts de base qui sont utiles dans la théorie de files d'attente, et l'importance de processus de naissance et de mort dans la modélisation de ces systèmes d'attente markoviens ainsi que l'importance du processus de Poisson dans l'extraction des équations de Chapman-Kolmogorov ou les équations de balance, comme nous avons vu l'utilité des chaînes de Markov dans l'analyse des systèmes d'attente non markoviens.

Nous avons vu quelques systèmes d'attente markoviens classiques et nous avons présenté pour chaque modèle, quelques mesures de performances dans un régime stationnaire avec la discipline FCFS (First Come First Served).

Nous avons présenté des systèmes d'attente non markoviens et nous avons donné quelques caractéristiques statistiques de chaque système en utilisant une analyse bien détaillée avec les chaînes de Markov induites et nous sommes arrivés à donner une approximation avec une borne inférieure et supérieure du temps d'attente d'un client dans le système G/G/1.

Nous avons présenté la conception et la réalisation de notre application de simulation de modèles non-markoviens avec JAVA qui est un langage orienté objet. Ceci nous a permis d'obtenir des résultats à priori valides. Néanmoins, ce programme nécessite des maintenances évolutives permettant de traiter des cas de modèles plus compliqués.

## *Perspectives :*

- ✓ Introduire d'autres distributions de probabilité restantes : La loi Géométrique, hypergéométrique, la Khidoux, Student, Bernoulli, la Binomiale, la Béta et la Gamma,
- ✓ Donner plus de performances pour une seule station et les généraliser pour plusieurs stations,
- ✓ Faire la comparaison entre les résultats obtenus par le simulateur et les résultats théoriques (analytiques).



# BIBLIOGRAPHIE

---

# BIBLIOGRAPHIE

---

- [1] U. Narayan. Bhat (2008), An Introduction to Queueing Theory , Modeling and Analysis in Applications . Statistical Science & Operations Research Southern Methodist University Dallas, TX 75275-0332 . USA
- [2] N. U. Prabhu (1965a), Queues and Inventories, Wiley, New York.
- [3] L. Takács (1962), Introduction to the Theory of Queues, Oxford University Press, New York.
- [4] N. U. Prabhu and U. N. Bhat (1963a), Some first passage problems and their application to queues, Sankhya A, 25, 281–292.
- [5] R. W. Wolff (1982), Poisson arrivals see time averages, Oper. Res., 30, 223–231.
- [6] R. B. Cooper (1981), Introduction to Queueing Theory, North-Holland, New York.
- [7] D. Gross and C. M. Harris (1998), Fundamentals of Queueing Theory, 3rd ed., Wiley, New York.
- [8] N. U. Prabhu and U. N. Bhat (1963b), Further results for the queue with Poisson arrivals, Oper. Res., 11, 380–386.
- [9] U. N. Bhat (1968), A Study of the Queueing Systems M/G/1 and GI /M/1, Lecture Notes in Operations Research and Mathematical Economics, Vol. 2, Springer-Verlag, New York.
- [10] Leonard Kleinrock (1975), Queueing systems vol. 1 : Queueing theory. Published simultaneously in Canada.
- [11] J. F. C. Kingman (1962 a, b), Some inequalities for the GI /G/1 queue, Biometrika, 49, 315–324 , on queues in heavy traffic, J. Roy. Statist. Soc. B, 24, 383–392.
- [12] K. T. Marshall (1968), Some inequalities in queueing, Oper. Res., 16, 651–665.
- [13] W. G. Marchal (1978), Some simpler bounds on the mean queueing time, Oper.

Res., 26, 1083–1088.

- [14] D. V. Lindley (1952), The theory of queues with a single server, Proc. Cambridge Philos. Soc. 48, 277–289.

## **Résumé**

L'influence et l'importance de la théorie des files d'attente dans plusieurs domaines divers et le recours à cette théorie pour la résolution des problèmes confrontés dans la réalité, comme les problèmes de la télécommunication et la gestion des stocks *etc.* Quand nous sommes en face d'un réel problème, et après la modélisation avec une file d'attente, on tombe généralement sur l'un des systèmes d'attente non markoviens et ce qui rend les choses très difficiles à analyser ce système et dégager des mesures de performances. Pour cela, nous avons procédé avec les chaînes de Markov induites, qui sont faciles et utiles dans l'extraction d'une distribution limite ; ce qui n'est pas le cas avec les techniques mathématiques vigoureuses. D'où on procède avec la simulation à événement discret, avec le langage orienté objet Java, pour obtenir quelques caractéristiques de performance, comme le nombre moyen de clients, le débit moyen d'entrée, le débit moyen de sortie et le temps moyen de réponse, *etc.*

### **Mots Clés**

Files d'Attente, chaîne de Markov induite, Orienté Objet, Java, Simulateur

## **Abstract**

The influence and the importance of the queuing theory in several domains and with it we can resolve many diverse problems that we can confronted in the reality like in the telecommunication and problems in stocks management *etc.* When we are face to any real problem, and after the establishment of a model with a queue waiting. Therefore we cannot find a classical queue Markovian, but a non-Markovian queue and it's very hard to get the performances measures with using all functions of Laplace and generating functions to extract them. So it's not evidently to calculate without using the imbedded chain of Markov to obtain the limit distribution, which is not evident with mathematical techniques. Hence, we must using the simulation method by a discrete event, using oriented object language Java, to get some performance criteria, such the mean number of clients in the system, the arrival mean debit, the departure mean debit, and the response mean time, *etc.*

### **Keywords**

Queuing, Imbeded Markov Chain, Object Oriented, Java, Simulator