Titre: Title:	Problèmes de contrôle optimal stochastique avec applications à la théorie des files d'attente et à la fiabilité
Auteur: Author:	Roozbeh Yaghoubi
Date:	2025
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Yaghoubi, R. (2025). Problèmes de contrôle optimal stochastique avec applications à la théorie des files d'attente et à la fiabilité [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/71711/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/71711/
Directeurs de recherche: Advisors:	Mario Lefebvre
Programme: Program:	Doctorat en génie industriel

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Problèmes de contrôle optimal stochastique avec applications à la théorie des files d'attente et à la fiabilité

ROOZBEH YAGHOUBI
Département de mathématiques et de génie industriel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie industriel

Novembre 2025

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Problèmes de contrôle optimal stochastique avec applications à la théorie des files d'attente et à la fiabilité

présentée par **Roozbeh YAGHOUBI**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Samira KEIVANPOUR, présidente
Mario LEFEBVRE, membre et directeur de recherche
Marzieh GHIYASINASAB, membre
Georges ABDUL-NOUR, membre externe

DÉDICACE

À ma famille, pour leur soutien inconditionnel.

REMERCIEMENTS

Je tiens à remercier mon directeur de recherche à Polytechnique Montréal, le professeur Mario Lefebvre, qui a eu la gentillesse de m'accepter comme doctorant, m'a guidé tout au long de mon parcours doctoral, a contribué de manière considérable à l'avancement de mes études doctorales, et m'a conduit avec patience vers la direction la plus élevée possible, tant pour la profondeur de ma recherche que pour le développement de ma trajectoire professionnelle future.

Je souhaite exprimer ma reconnaissance aux professeurs et mentors suivants pour avoir contribué à l'achèvement de mes études à Montréal en transmettant leur savoir et en développant la confiance nécessaire à la poursuite de mes ambitions académiques : le Dr Sorin Voiculescu, M. Messaoudi Darragi et le professeur Georges Abdul-Nour, pour leurs idées créatives et leur soutien considérable, tant sur le plan académique que dans la manière de traduire les connaissances théoriques en pratique industrielle grâce à leurs riches expériences. Je souhaite également remercier la Dre Soumaya Yacout, la Dre Samira Keyvanpour, la Dre Marzieh Ghiyasinasab et le Dr Ardalan Sabamehr pour leurs encouragements particulièrement marquants et leurs enseignements éclairants.

Je suis également reconnaissant envers le regretté Dr Mir Bahador Aryanezhad, le Dr Farhad Hosseinzadeh Lofi, le Dr Shervin Asadzadeh, le Dr Hamidreza Mostafaei et le Dr Shahnam Javadi pour avoir cru en mon potentiel à poursuivre des études supérieures en génie industriel, m'aidant ainsi à faire le lien entre ce domaine et ma formation de premier cycle en mathématiques appliquées. J'apprécie profondément tous mes professeurs pour leur inspiration et leur dévouement tout au long de mes nombreuses et belles années académiques.

Je tiens également à mentionner l'utilisation de DeepL et d'outils de traduction assistée par l'IA, qui ont été utilisés pour soutenir et améliorer cette version française de la thèse, initialement rédigée en anglais

Je souhaite exprimer ma profonde gratitude à mon directeur durant les vingt années ayant précédé mon immigration au Canada, M. Zia Momenzadeh, pour la richesse de sa formation professionnelle ainsi que pour ses leçons de vie amicales et pleines de sagesse.

Enfin, je remercie ma mère, mon père et toute ma famille pour leur espoir constant, leur patience et leur soutien ininterrompu au fil des années, jusqu'à l'obtention de mon diplôme de doctorat au Canada.

RÉSUMÉ

La thèse s'articule autour de plusieurs axes thématiques complémentaires, présentés ci-dessous.

Contrôle optimal d'un système de files d'attente

Dans le modèle classique de files d'attente $M/M/k$, le système est décrit de telle manière que, selon un processus de Poisson de taux λ, les clients arrivent dans le système et sont servis par l'un des k serveurs disponibles. Les temps de service fournis par chaque serveur sont distribués exponentiellement avec le paramètre μ. Ces temps de service sont indépendants et ne dépendent pas des temps interarrivées des clients. Le système fonctionne selon une discipline de service FIFO (premier entré, premier sorti) et a une capacité totale c_p, qui peut être infinie.

Des stratégies de commande optimale pour les systèmes de files d'attente ont été explorées dans de nombreux articles, souvent sous l'hypothèse que les paramètres de service peuvent être ajustés par le contrôleur, par exemple, en modifiant le nombre de serveurs actifs ou le taux de service. Certains auteurs ont également envisagé la possibilité de contrôler les paramètres d'arrivée. L'optimisation de ces paramètres varie, incluant des stratégies conçues pour servir différents types de clients et minimiser les coûts totaux sur un horizon temporel fini ou infini, sur la base de critères de coût.

Dans cette recherche, un modèle de file d'attente $M/M/k$ modifié est considéré, dans lequel le nombre de serveurs actifs peut être contrôlé dynamiquement dans le système à tout moment. Supposons qu'à un moment donné, $k+l$ clients attendent un service dans le système. Afin de réduire ce nombre de clients à $k+r$, où $0 \leq r < l$, il s'agit de déterminer le nombre optimal de serveurs à déployer le plus rapidement possible, tout en tenant compte des coûts de contrôle associés. La fonction de valeur qui satisfait l'équation de programmation dynamique est dérivée dans le cas général, et des solutions sont proposées pour des cas particuliers du problème.

Distribution optimale des temps de service dans une file d'attente $M/G/1$

Un système de files d'attente $M/G/1$ est considéré dans lequel le serveur a la flexibilité de choisir entre deux distributions de temps de service distinctes. L'objectif est de déterminer le choix qui minimise la valeur espérée d'un critère de coût, qui prend en compte à la fois le coût d'exploitation à un taux de service plus élevé et le temps nécessaire pour vider la file d'attente. Le temps final aléatoire est défini comme le premier instant où aucun client ne reste

en ligne en attente de service. La programmation dynamique peut être appliquée pour obtenir la solution optimale lorsque les temps de service suivent une distribution exponentielle. Dans le cas plus général, la probabilité conditionnelle est utilisée pour analyser le comportement du système. En outre, des solutions explicites sont fournies pour des cas particuliers dans lesquels la capacité du système est finie.

Pour déterminer les distributions optimales du temps de service d'un serveur unique actif opérant en continu sur un intervalle de temps aléatoire, une méthodologie novatrice est proposée dans ce travail. Ce travail apporte une solution à un problème non résolu concernant la prédiction de la vitesse de service optimale des modules actifs à un instant donné, permettant ainsi un contrôle efficace du système dans un cadre temporel aléatoire continu. L'accent est mis sur un cadre où un serveur peut fonctionner à deux vitesses distinctes, et la question centrale est de déterminer quelle vitesse est optimale selon les conditions.

La méthodologie proposée intègre la programmation dynamique avec un conditionnement sur les taux d'arrivée et de service, garantissant une approche structurée pour optimiser la dynamique de service. Le modèle de file d'attente est formulé sur la base d'un système d'équations différentielles qui analyse le comportement du système dans ce cadre. Une formulation modifiée de commande optimale est appliquée afin de minimiser la valeur espérée de la fonction de coût et de déterminer la distribution optimale du temps de service à un instant donné.

La méthodologie proposée peut être adaptée à diverses applications réelles, lesquelles sont abordées ici comme des cas particuliers avec capacité finie du système. Elle peut aussi être utilisée pour déterminer la politique optimale en fonction des préférences du système, telles que les contraintes de capacité et le nombre initial de clients au moment de la planification. Les résultats sont illustrés graphiquement afin de montrer l'impact significatif de la solution proposée sur l'optimisation des performances du système pendant les périodes analysées.

Comparée à d'autres travaux, la méthodologie actuelle se distingue par sa capacité à contrôler le processus via une programmation dynamique et une modélisation conditionnelle sur un horizon temporel indéterminé. Cela apporte un niveau supérieur de sophistication aux modèles de commande optimale, en abordant des aspects encore inexplorés. La validation des modèles proposés montre leur aptitude à déterminer efficacement les politiques de commande optimales et à assurer la stabilité du système selon ses paramètres. Les résultats démontrent que la programmation dynamique et la modélisation conditionnelle constituent un cadre systématique pour optimiser les coûts totaux dans des conditions opérationnelles variables.

Politique optimale d'inspection et de maintenance : Intégration d'une chaîne de Markov en temps continu

L'état d'une machine est modélisé comme une chaîne de Markov à temps continu contrôlée, où les temps de service sont aléatoires. En intégrant les coûts de maintenance, l'objectif est de maximiser le temps de fonctionnement avant qu'une réparation soit nécessaire. Cette recherche dérive l'équation de programmation dynamique associée à la fonction de valeur, permettant une prise de décision optimale concernant les taux d'inspection. Cette méthodologie résout explicitement des problèmes visant à minimiser à la fois les coûts directs de maintenance et les dépenses liées aux défaillances, offrant ainsi un cadre robuste pour identifier des fréquences d'inspection optimales dans un contexte de maintenance centrée sur la fiabilité.

Le développement d'une politique optimale d'inspection-maintenance pour les systèmes, en modélisant l'état de dégradation comme une chaîne de Markov à temps continu, est l'objectif principal de cette recherche. En particulier, l'ajustement dynamique du taux d'inspection u, qui détermine la fréquence ou l'intensité des tâches de maintenance préventive, constitue le concept central. Ce taux agit comme un paramètre de contrôle équilibrant les compromis entre les activités de maintenance et le risque de défaillances.

Un taux élevé u augmente les activités de maintenance, réduisant potentiellement le taux de défaillance, mais génère également des coûts directs plus élevés. À l'inverse, un taux faible u diminue les coûts immédiats mais peut engendrer davantage de défaillances et d'indisponibilités à long terme. Déterminer le taux d'inspection optimal u qui équilibre ces facteurs est donc l'enjeu central. L'objectif est d'appliquer la programmation dynamique pour obtenir la valeur optimale de u, minimisant les coûts espérés à long terme en prenant en compte la fréquence des interventions et les indisponibilités dues aux pannes.

La programmation dynamique est un cadre adapté pour évaluer diverses stratégies de maintenance sous incertitude. La nature dynamique du problème, où l'état du système et la politique optimale dépendent des états et décisions antérieures, exige un modèle capable de gérer la prise de décision séquentielle sous aléa. Cette recherche vise à développer une méthode intégrant la programmation dynamique dans une chaîne de Markov à temps continu, afin de déterminer le taux d'inspection-maintenance optimal. Cela permet une planification robuste et adaptable aux fluctuations en temps réel des conditions du système et des contraintes opérationnelles.

Dans cette méthodologie, un nouveau problème de type "homing" est proposé dans le contexte de la théorie de la fiabilité, caractérisé par une prise de décision dans des environnements stochastiques. La formulation et l'implémentation de l'équation de programmation dynamique présentent des défis théoriques que ce travail vise à relever. Représenter correctement les phénomènes de dégradation implique d'intégrer efficacement les coûts de maintenance et

d'indisponibilité, tout en assurant une résolution calculable des équations pour les systèmes complexes. Surmonter ces défis permettra de faire progresser les stratégies de maintenance, en fournissant des solutions flexibles et économiquement efficaces pour différents contextes opérationnels et profils de risque.

ABSTRACT

The thesis is structured around several complementary research themes, presented below.

Optimal control of a queueing system

In the classic $M/M/k$ queueing model, the system is described so that, according to a Poisson process with rate λ, customers arrive at the system and are served by one of the k available servers. The service times provided by each server are exponentially distributed with parameter μ. These service times are independent and do not depend on the interarrival times of customers. The system operates under a FIFO (first in, first out) service discipline and has a total capacity of c_p, which may be infinite.

Optimal control strategies for queueing systems have been explored in numerous papers, often under the assumption that service parameters can be adjusted by the controller, for example, through changes in the number of active servers or the service rate. Some authors have also considered the possibility of controlling the arrival parameters. The optimization of these parameters varies, including strategies designed to serve different types of customers and minimize total costs over a finite or infinite time horizon, with the optimizer's objective based on cost criteria.

In this research, a modified $M/M/k$ queueing model is considered, where we, as decision makers, can dynamically control the number of active servers in the system at any given time. Suppose that, at a particular time instant, $k + l$ customers are waiting for service in the system. In order to reduce this number of customers to $k + r$, where $0 \leq r < l$, our aim is to determine the optimal number of servers to be deployed as quickly as possible, while accounting for the associated control costs. The value function which satisfies the dynamic programming equation is derived in the general case, and solved solutions are proposed for specific instances of the problem.

Optimal service time distribution for an $M/G/1$ waiting queue

We consider an $M/G/1$ queueing system in which the server has the flexibility to choose between two distinct service-time distributions. The objective is to determine the choice that minimizes the expected value of a cost criterion, which considers both the cost of operating at a higher service rate and the time needed to empty the queue. The random final time is defined as the first instance in which no customer remains in the line waiting for service. Dynamic programming can be applied to obtain the optimal solution when service times follow an exponential distribution. In the more general case, conditional probability is employed to

analyze the system's behavior. In addition, explicit solutions are provided for particular cases in which the system has a finite capacity.

To determine the optimal service-time distributions of an active single server operating continuously over a random time interval, a novel methodology is proposed in this work. This work provides a solution to the previously unsolved problem of predicting the optimal service speed of active modules at any given time, thereby enabling effective control over the system within a continuous random time framework. The primary focus is on a setting where a server can operate at two distinct speeds, and the central concern lies in determining which speed is optimal under varying conditions. The proposed methodology integrates dynamic programming with conditioning on both arrival rates and service rates, ensuring a structured approach to optimize the service dynamics.

The proposed queueing model is formulated on the basis of a system of differential equations that analyzes the system's behavior within the dynamic programming framework. A modified optimal control formulation is applied to minimize the expected value of the cost function and to determine the optimal service time distribution at any given time instant, leveraging conditioning based on the comparison of arrival and service speeds.

Ultimately, the proposed methodology can be adapted to various real-world applications, which will be discussed in this investigation as particular cases where the system capacity is finite. Furthermore, the proposed methodology can be employed to determine the optimal policy in different preferences of the system, such as capacity constraints and the initial number of customers in the planning stage. Finally, the results are illustrated through graphical representations to demonstrate the significant impact of the proposed solution in optimizing system performance over the analyzed periods.

Compared to other studies, the key advantage of the current methodology lies in its ability to achieve process control by leveraging dynamic programming and condition-based modeling over an undetermined time horizon. This characteristic introduces a higher level of sophistication to optimal control models, addressing aspects that have not been explored in prior works. The validation of the proposed research models confirms their ability to accurately determine optimal control policies and to ensure the stability of the system based on the parameters of the system. The findings illustrate that both dynamic programming and conditioning-based modeling provide a systematic framework for optimizing total costs under varying operating conditions. The results are further supported by specific problem scenarios that are explicitly solved using the proposed equations.

Preventive maintenance and inspection policy

The state of a machine is modeled as a controlled continuous-time Markov chain, where the service times are random. By incorporating maintenance costs, our objective is to maximize operational time before the machine requires repair. The research derives the dynamic programming equation associated with the value function, enabling optimal decision making regarding inspection rates. This methodology explicitly solves specific problems that aim to minimize both direct maintenance costs and potential failure-related expenses, thereby establishing a robust framework to identify optimal inspection frequencies within a reliability-centered maintenance context. These findings improve the development of maintenance strategies and provide explicit solutions for particular scenarios, which provide valuable insight for application in reliability engineering.

Developing an optimal inspection-maintenance policy for systems by modeling the degradation state as a continuous-time Markov chain is the main focus of this research. Specifically, the dynamic adjustment of the inspection rate u, which dictates the frequency or intensity of preventive maintenance tasks, is the central concept. The inspection rate acts as a control parameter that balances the trade-off between maintenance tasks and the risk of system failures.

A higher inspection rate u increases maintenance activities, potentially reducing failure rates ; however, it also contributes to higher direct maintenance costs. In contrast, assigning a lower inspection rate u reduces immediate maintenance costs, but can lead to long-term increases due to the increased number of failures and prolonged downtimes. Determining the optimal inspection rate u that effectively balances these opposing factors is the central challenge of this research. As mentioned above, our objective is to apply dynamic programming to obtain the optimal value of the inspection rate u that minimizes expected long-term costs. These costs account for both the frequency of maintenance activities and the potential unavailability of the system resulting from failures.

A well-suited framework for this problem is dynamic programming, as it enables the systematic evaluation of various maintenance strategies under uncertainty. The intrinsically dynamic nature of the problem, where the state of the system and the selection of the optimal maintenance policy are influenced by historical states and actions, demands a model capable of handling sequential decision-making under randomness. The objective of this research is to develop a method that integrates dynamic programming within a continuous-time Markov chain framework to determine the optimal inspection-maintenance rate. This approach offers valuable information on the optimal scheduling of maintenance tasks based on the current state of the system. These insights contribute to robust maintenance planning, enhancing responsiveness to real-time fluctuations in system conditions and operational constraints.

In this methodology, a novel type of homing problem is proposed in the context of reliability theory, characterized by decision making in stochastic environments. The formulation and implementation of the dynamic programming equation involve significant theoretical challenges, which this research aims to address. Accurately representing degradation phenomena is a key aspect of these challenges, which requires effective integration of maintenance and downtime costs, while ensuring computational tractability in solving the dynamic programming equations for complex systems. The advancement of maintenance planning strategies will be enabled by overcoming these difficulties, providing cost-effective and flexible solutions across a diverse range of operational settings and risk preferences.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

Symboles latins

k	Nombre total de serveurs disponibles dans le système
c_p	Capacité totale du système de files d'attente (en service + en file)
l	Nombre actuel de clients en file d'attente, au-delà des k en service
r	Nombre souhaité de clients à maintenir, au-delà des k en service
t	Variable temporelle continue
$F(t_0, t_1, l, r)$	Fonction de valeur
$F'(t_0)$	Dérivée temporelle de la fonction de valeur
$n(t)$	Nombre de serveurs en fonctionnement à l'instant t
$m[n(t)]$	Revenu par unité de temps lorsque $n(t)$ serveurs sont actifs
$G(t_0)$	Fonction auxiliaire combinant des fonctions de valeur à t_0
$H(t_0)$	Fonction auxiliaire basée sur les fonctions de valeur à t_0
e	Nombre d'Euler
$\mathrm{Exp}(\cdot)$	Fonction exponentielle de base e
$G_i(t_0)$	Valeur de la fonction objectif lorsque $n(t_0) = i$
q_0	Taux de revenu par serveur et par unité de temps
T	Temps de premier passage
$J(K)$	Fonction de coût total sous une politique de contrôle k
$\mathbb{E}[S(t)]$	Durée de service espérée à l'instant t
$K(T(k))$	Fonction de coût terminal au temps d'arrêt $T(k)$
inf	Infimum (borne inférieure la plus grande)
$g[u(t)]$	Compromis entre le coût de contrôle et le bénéfice du système
U	Ensemble des valeurs de contrôle admissibles
A	Temps d'arrivée d'un nouveau client
D	Temps de départ d'un client
G	Distribution générale du temps de service
u	Taux d'inspection pour la maintenance préventive
L_q	Longueur moyenne de la file d'attente
$X(t)$	Nombre de clients dans le système à l'instant t
$X_q(t)$	Nombre de clients en attente dans la file à l'instant t
N	Nombre de clients dans le système à l'état stationnaire
m	Nombre maximal de clients considéré
$f[u(t)]$	Fonction de coût de contrôle

s	Nombre de serveurs dans un système M/M/s
Q	Temps d'attente en file avant le début du service
i	Indice d'un système de files d'attente
k	Nombre total de systèmes de files d'attente (ou nœuds) dans le réseau
$p_{i,j}$	Probabilité qu'un client se dirige vers le système j après le service dans i
p	Probabilité
p_i	Probabilité qu'un client quitte le réseau après le service dans i
C	Constante positive
$C(t_p)$	Coût moyen par unité de temps sous remplacement à intervalles fixes
C_p	Coût de la maintenance préventive
$\mathbb{E}[N(t_p)]$	Nombre espéré de pannes durant l'intervalle $[0, t_p]$
C_f	Coût de défaillance (action corrective après une panne)
t_p	Temps planifié de remplacement
$R(t_p)$	Fonction de fiabilité (probabilité de non-défaillance jusqu'à t_p)
$F(t_p)$	Fonction de distribution cumulative des défaillances
$f(t)$	Densité de probabilité du temps de défaillance
$r(n, i, k)$	Revenu à l'étape n, état i, action k
$t(n, i, k)$	Fonction de transition d'état issue de l'état i, action k, à l'étape n
$f(n, i)$	Fonction de valeur à partir de l'état i à l'étape n
d	Période d'un état de Markov
E	Espace d'états (ensemble de tous les états du système)

Symboles grecs

μ	Taux de service de chaque serveur (moyenne de clients servis par unité de temps)
λ	Taux d'arrivée des clients dans un processus de Poisson
ν	Taux de transition depuis l'état i dans une chaîne de Markov en temps continu
ρ	Taux d'utilisation du système (proportion du temps où le système est occupé)
σ	Premier instant de commutation auquel la politique de contrôle change
τ	Deuxième instant de commutation auquel la politique de contrôle change
α	Niveau de service cible ou pourcentage de performance de la file d'attente

π_n	Probabilité stationnaire d'avoir n clients
θ_i	Taux d'arrivée externe dans le système de files d'attente i
λ_e	Taux d'arrivée externe total dans tout le réseau
Θ	Paramètre d'échelle (Weibull)
β	Paramètre de forme (Weibull)
η	Nombre de serveurs en activité à l'instant t
Δt	Petit incrément de temps
$O(\Delta t)$	Erreur proportionnelle à Δt
∂	Dérivée partielle

Abbreviations

EDO	Équation différentielle ordinaire
$M/M/k$	Arrivées exponentielles, services exponentiels, k serveurs
$M/G/1$	Arrivées exponentielles, services généraux, 1 serveur
$GI/M/1$	Arrivées générales indépendantes, services exponentiels, 1 serveur
$M/M/1/m$	Arrivées exponentielles, services exponentiels, capacité finie m
PDF	Fonction de densité de probabilité
CDF	Fonction de répartition
Exp(μ)	Loi exponentielle de paramètre μ
$G(\alpha, \beta)$	Loi gamma de paramètres α, β
FIFO	Discipline Premier Entré, Premier Sorti
IoT	Internet des objets
CBM	Maintenance conditionnelle
RUL	Durée de vie utile restante
VAR	Modèle vectoriel autorégressif
PHM	Pronostic et gestion de la santé des systèmes
DPE	Équation de programmation dynamique
LQ	Linéaire quadratique
LQG	Linéaire quadratique gaussien
CP	Capacité du système
POMDP	Processus de décision de Markov partiellement observable

CHAPITRE 1 INTRODUCTION

1.1 Systèmes de files d'attente

La plupart des installations de services sont capables d'absorber davantage de clients à long terme ; cependant, en dehors des périodes de pointe, les systèmes deviennent souvent sous-utilisés, laissant les serveurs inactifs, ce qui augmente le coût par service fourni. Par conséquent, les concepteurs du système de files d'attente doivent équilibrer le coût des services offerts avec le coût du temps d'attente moyen des clients pour recevoir les services. La planification et l'analyse de la capacité de service sont basées sur la théorie des files d'attente, qui est une technique mathématique permettant d'étudier le temps d'attente et le temps de réponse dans lequel la file d'attente est généralement créée par l'arrivée des clients.

Les systèmes de files d'attente sont des processus stochastiques qui impliquent des entités entrant dans les systèmes à un taux d'arrivée donné et attendant en ligne pour recevoir des services. Ils sont utilisés dans divers domaines tels que les télécommunications, les systèmes de demande en ligne, les banques, etc. Pour concevoir des systèmes de files d'attente efficaces, la théorie de la commande optimale est appliquée afin de déterminer la meilleure allocation des ressources et d'optimiser la capacité de service. Cela peut être réalisé par l'optimisation en formulant un problème de programmation mathématique qui vise à trouver les conditions optimales, y compris le taux des services fournis et la capacité du système de files d'attente afin d'optimiser la fonction objective. Le terme taille maximale de la file, également connu sous le nom de capacité du système, fait référence au plus grand nombre possible de clients pouvant attendre dans une file en plus de ceux étant actuellement servis. De plus, la stabilité fait référence à la propriété du système de files d'attente dans lequel il y a un nombre fini de clients dans le système à tout moment donné.

La fonction objective contient le temps total d'attente, le temps de service et les coûts liés à la capacité du système. Ces facteurs permettent d'évaluer les performances du système et aident à établir des politiques de commande optimales en équilibrant les compromis entre des objectifs concurrents. Les fondements de la théorie moderne des files d'attente ont été initiés par un ingénieur danois, Agner Krarup Erlang, au début du vingtième siècle. Peu d'études ont été menées sur le sujet avant la Seconde Guerre mondiale, bien que la théorie mentionnée ait été mise en œuvre dans un large éventail de pratiques.

1.1.1 Modélisation de la commande optimale des files avec un *temps aléatoire*

Le modèle de file d'attente $M/M/k$ (qui décrit un système de file d'attente avec k serveurs, des arrivées de type Poisson et des temps de service exponentiels) est modifié en permettant d'ajuster le nombre de serveurs à tout moment. Le scénario envisagé à un instant donné est celui dans lequel il y a $k + l$ clients dans le système, tous en attente de service. L'objectif est de déterminer le nombre de serveurs nécessaires pour réduire le nombre de clients à $k+r$, où r est une valeur comprise entre 0 et l ($0 \leq r < l$), aussi rapidement que possible, tout en tenant compte du coût du contrôle des serveurs. L'équation de programmation dynamique satisfaite par la fonction de valeur est obtenue dans le cas général et des problèmes spécifiques sont résolus explicitement à l'aide de cette équation.

Il est essentiel de considérer que les files d'attente à temps aléatoire peuvent être plus complexes à modéliser que les systèmes déterministes. Cependant, les perspectives issues de cette recherche pourraient avoir des effets considérables sur la conception et la gestion de divers systèmes de services. Dans les systèmes de files d'attente à temps aléatoire, les intervalles entre les arrivées des clients dans le système ou le temps nécessaire pour fournir un service à un client sont aléatoires. L'application du temps aléatoire est utile pour modéliser des situations réelles où les temps ne sont pas supposés déterministes, mais suivent une distribution de probabilité spécifique. Cela est crucial pour reconnaître le comportement stochastique des files et faire des prévisions plus optimisées sur leur fonctionnement.

1.1.2 Objectifs de l'analyse de file d'attente

Le contenu des sections 1.1.2 à 1.1.4 est adapté de Aryanezhad et al. [1], et résumé par l'auteur pour cette thèse. Un objectif dans l'analyse des files est la minimisation des coûts, soit le temps perdu par les clients en attente, soit les coûts associés à la prestation des services. En raison de la complexité du calcul du coût d'attente, celui-ci est généralement remplacé par la taille de la file, pour laquelle des informations pertinentes sont insérées dans le modèle par des experts comme valeur par défaut. Un objectif secondaire dans l'analyse des systèmes de files est l'équilibre entre les services fournis et le temps total perdu des clients dans la file ; cette notion est illustrée à la figure 1.1.

Une file est une fonction de la variation de l'entrée et des services fournis ; dans de nombreux cas, cette variation peut être modélisée par une distribution connue. Les schémas d'entrée de nombreux systèmes sont caractérisés par la distribution de Poisson, et les temps de service sont définis par une distribution exponentielle. La distribution exponentielle modélise le temps entre les événements dans un processus de Poisson, en supposant que les événe-

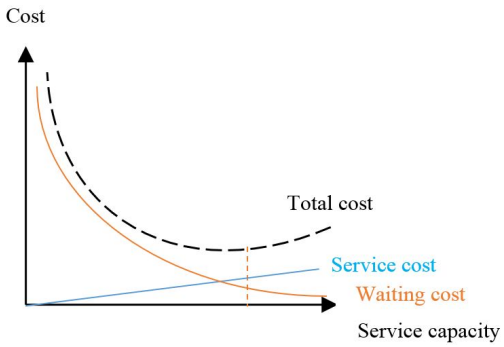

Figure 1.1 Coûts du système

ments se produisent indépendamment à un taux constant et sans schéma prédéterminé dans le processus de Poisson. Dans la distribution de Poisson, le taux d'arrivée des clients dans le système est constant et bien que les distributions de Poisson et exponentielle soient liées, la première décrit le nombre d'arrivées dans un intervalle de temps fixe, tandis que la seconde modélise les temps inter-arrivées.

Les événements quotidiens et expériences dans les systèmes impliquent des files d'attente, comme attendre dans une file à une agence bancaire ou utiliser des services électroniques. Ces événements sont souvent soumis à l'incertitude et à la variation, ce qui peut rendre l'optimisation des performances une tâche difficile.

La théorie de la commande optimale est une méthode pour créer des systèmes de files efficaces en analysant l'interaction entre les paramètres d'entrée et de sortie. L'application de la théorie de la commande optimale dans les files peut aider à réduire les temps d'attente, à optimiser les performances du système et à allouer les ressources de manière efficace. En conséquence, la théorie de la commande optimale fournit un outil précieux pour améliorer la conception et la planification des systèmes de files dans divers domaines, y compris la formulation de politiques de maintenance et la conception de réseaux de systèmes dans de multiples installations. Il a toujours été supposé que la fenêtre temporelle dans laquelle le système entier est examiné tend vers l'infini $t \in [0, \infty)$ ou que les plages horaires et les intervalles de temps sont définis à l'avance dans $t \in [0, T]$, ce qui serait irréaliste, car avec l'émergence de nouvelles technologies connectées, telles que l'IoT (Internet des objets), l'informatique en nuage et dans le domaine de l'Industrie 4.0, la définition du temps devient dépendante du comportement dynamique de divers paramètres, selon la perspective de l'efficacité du système ; il serait donc judicieux d'initier des recherches sur la stochasticité de l'horizon temporel.

1.1.3 Définition de la fonction objective

1- L'utilisation du système, qui définit le pourcentage d'utilisation du système, notée ρ.

2- La relation entre le coût d'utilisation à un niveau donné d'utilisation du système et la longueur de file résultante.

$$\rho = \frac{\lambda}{k\mu} \tag{1.1}$$

ρ : Coefficient d'utilisation du système

λ : Taux d'arrivée des clients

k : Nombre de serveurs

μ : Taux de service

$\frac{\lambda}{\mu}$: Indicateur de flux du système

Si le flux du système (rapport entre le taux d'arrivée et le taux de service) dépasse 1 et qu'il n'existe qu'un seul serveur dans le système, alors le processus ne pourra pas gérer tous les clients. Le calcul et l'interprétation du concept d'utilisation nécessitent une attention particulière, car il définit le pourcentage de temps pendant lequel le système est occupé par rapport à l'état d'inactivité. Par conséquent, le système global devrait être conçu de manière à ce que les coûts combinés du temps d'attente des clients et du fonctionnement des serveurs (équipements) soient minimisés.

1.1.4 Hypothèses

- Le taux d'entrée suit une distribution de Poisson ; par conséquent, le taux d'arrivée des clients dans le système est constant.
- Le système fonctionne dans des conditions stables, ce qui entraîne des valeurs moyennes constantes pour les temps d'arrivée et de réponse.
- Le nombre moyen de clients en attente dans la file est connu et noté L_q.

Les modèles couramment utilisés dans la littérature sont listés dans la section suivante.

a. Modèles de systèmes de files d'attente avec population infinie

1- Système à serveur unique avec temps de service exponentiel ($M/M/1$), où les arrivées suivent une distribution de Poisson (voir Figure 1.2).

La Figure 1.3 illustre le système à serveur unique avec plusieurs étapes de service.

Figure 1.2 Serveur unique, étape unique

Figure 1.3 Serveur unique, étapes multiples

2- Système à serveur unique avec taux de service constant $(M/D/1)$, où les variations dans le taux d'arrivée et le taux de service sont atténuées ou négligeables.

3- Système à serveurs multiples avec temps de service exponentiel $(M/M/k)$, où deux serveurs ou plus fournissent indépendamment des services avec les hypothèses suivantes (voir Figure 1.4) :

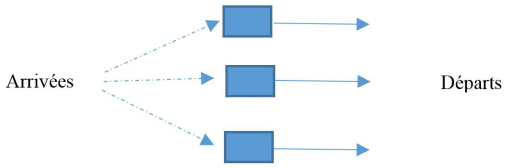

Figure 1.4 Serveurs multiples, étape unique

3.1- Les arrivées suivent une distribution de Poisson, et les temps de service sont distribués exponentiellement.

3.2- Tous les serveurs ont un taux de service identique.

3.3- Tous les clients forment une seule file d'attente et une méthode de service selon la priorité du premier arrivé, premier servi est utilisée.

Un autre élément discuté dans la conception de la capacité serait la détermination de la taille de la salle d'attente, qui devrait être définie de manière à ce que la probabilité que la longueur de la file dépasse un certain niveau reste en dessous d'un seuil prédéfini. Par exemple, le concepteur peut viser à garantir que la taille de la file soit adéquate dans 98%

des scénarios. La longueur estimée de la file résulte de la résolution des équations suivantes :

$$n = \frac{\log k}{\log \rho} \tag{1.2}$$

$$\text{où :} \quad k = \frac{1-\alpha}{L_q(1-p)} \tag{1.3}$$

ρ : Coefficient d'utilisation du système

α : Pourcentage attendu

L_q : Nombre moyen de clients dans la file

b. Modèles de systèmes de files d'attente avec population finie

Ces modèles sont utilisés lorsque la population interagissant avec le système est limitée (par exemple, lorsqu'on reçoit un soutien de ressources externes ou qu'on dispose d'un petit nombre de clients) ; l'un des facteurs distinctifs entre les modèles de files à population finie et infinie est le taux d'arrivée des clients dans le système : le taux d'entrée dans la file est constant dans les modèles à population infinie, tandis que dans les modèles à population finie, le taux d'arrivée dépend du nombre de clients actuellement présents dans le système (plus la taille est grande, plus le taux d'arrivée diminue, car la population génératrice de la file décroît).

1.1.5 Processus stochastiques en temps continu et à états discrets

Le contenu de cette section est adapté et résumé à partir de Lefebvre [2], initialement présenté dans le contexte des processus stochastiques appliqués, afin de convenir au cadre de modélisation utilisé dans cette thèse. Ce modèle mathématique prend en compte tous les clients dans le système, y compris ceux en attente et ceux en cours de service.

La distribution conditionnelle de la variable aléatoire N, qui représente le nombre de clients dans le système à l'état stationnaire, sachant que $N \leq m$, est donnée par :

$$P_N(n \mid N \leq m) = \frac{\pi_n}{\sum_{k=0}^{m} \pi_k} = \frac{(\lambda/\mu)^n}{\sum_{k=0}^{m} (\lambda/\mu)^k}, \quad \text{pour } n = 0, 1, \ldots, m \tag{1.4}$$

Il est important de noter que cette condition n'implique pas que la capacité du système est égale à m, mais plutôt qu'à un moment donné à l'équilibre, il y avait au plus m clients dans le système.

Note 1 : L'étude examine également une modification du système de file $M/M/1$, dans lequel

un client retourne en fin de file avec une probabilité $p \in (0, 1)$, c'est-à-dire qu'un client peut retourner dans la file un nombre illimité de fois. Les équations d'équilibre du système sont présentées comme suit :

$$\text{État } j: \quad \text{Taux de départ de l'état } j = \text{Taux d'arrivée à l'état } j \tag{1.5}$$

$$\text{État } 0: \quad \lambda\pi_0 = \mu(1-p)\pi_1 \tag{1.6}$$

$$\text{État } n\ (\geq 1): \quad [\lambda + \mu(1-p)]\pi_n = \lambda\pi_{n-1} + \mu(1-p)\pi_{n+1} \tag{1.7}$$

Note 2 : De plus, le modèle $M/M/s$ est discuté, qui est une généralisation du modèle $M/M/1$. Ce modèle suppose qu'il y a s serveurs dans le système, avec un taux exponentiel μ. Les clients arrivent selon un processus de Poisson avec un taux λ. Le système est supposé avoir une capacité infinie et la politique de service par défaut est premier arrivé, premier servi.

Comme les clients arrivent un par un et sont servis un par un, le processus $\{X(t),\ t \geq 0\}$, où $X(t)$ représente le nombre de clients dans le système de file d'attente au temps t, est un processus de naissance et de mort, un processus stochastique dans lequel la taille de la population peut augmenter ou diminuer d'une unité à la fois. Les équations d'équilibre de ce système sont exprimées comme suit :

$$\text{État } j: \quad \text{Taux de départ de } j = \text{Taux d'arrivée à } j \tag{1.8}$$

$$\text{État } 0: \quad \lambda\pi_0 = \mu\pi_1 \tag{1.9}$$

$$0 < k < s: \quad (\lambda + k\mu)\pi_k = (k+1)\mu\pi_{k+1} + \lambda\pi_{k-1} \tag{1.10}$$

$$k \geq s: \quad (\lambda + s\mu)\pi_k = s\mu\pi_{k+1} + \lambda\pi_{k-1} \tag{1.11}$$

Figure 1.5 illustre le modèle $M/M/2$:

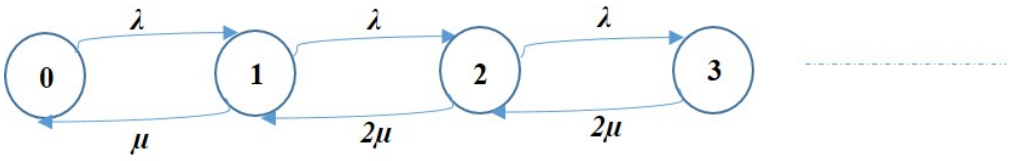

Figure 1.5 Diagramme de transition d'état pour le modèle $M/M/2$

Note 3 : Pour calculer la distribution du temps total qu'un client passe dans un système $M/M/S$ à l'équilibre, on doit trouver le produit de convolution de la fonction de densité de probabilité de la variable aléatoire Q et $S \sim \text{Exp}(\mu)$, qui désigne une variable aléatoire suivant une loi exponentielle de paramètre μ.

Note 4 : La formule suivante donne les probabilités limites valides pour un modèle plus général connu sous le nom de système de perte d'Erlang ou $M/G/s/s$, avec un temps de service arbitraire non négatif. Ce résultat permet de résoudre les problèmes où le temps de service ne suit pas une distribution exponentielle.

$$\pi_k = \frac{(\lambda \mathbb{E}[s])^k / k!}{\sum_{j=0}^{s} (\lambda \mathbb{E}[s])^j / j!}, \quad \text{pour } k = 0, 1, \ldots, s \tag{1.12}$$

De plus, la Note 4 précise que dans le système $M/M/s$, le taux d'arrivée des clients doit être inférieur au taux total de service fourni par les s serveurs, ce qui est représenté par la condition $\lambda < s\mu$, afin que le système atteigne une distribution stationnaire ; sinon, la longueur de la file continuera de croître et aucun équilibre ne sera atteint. Toutefois, en pratique, certains clients peuvent choisir de ne pas entrer dans le système s'ils voient que la file est trop longue, et d'autres peuvent partir avant de recevoir un service.

1.2 Maintenance préventive et politique d'inspection

Dans le contexte de l'ingénierie de la fiabilité, de la maintenance et de l'optimisation de la disponibilité, la maintenance préventive et la politique d'inspection se sont révélées être des éléments clés pour obtenir des systèmes fiables et efficaces. La planification de la maintenance implique des tâches planifiées visant à garantir un fonctionnement ininterrompu et à prévenir les pannes par la détection précoce et le contrôle des défaillances potentielles. Modéliser la dégradation par des processus stochastiques, tout en tenant compte de l'aléa inhérent au comportement du système, rend l'objectif d'optimisation de ces stratégies de plus en plus complexe. Pour modéliser efficacement les transitions probabilistes entre les états de la machine pendant la période de maintenance planifiée, un modèle mathématique largement adopté est la chaîne de Markov en temps continu.

L'équilibre entre les coûts associés aux tâches de maintenance et les pannes potentielles du système est un élément crucial de l'optimisation. Dans une stratégie de maintenance efficace, il est nécessaire de prendre en compte à la fois les coûts directs liés à la mise en œuvre des actions de maintenance et les coûts indirects associés aux arrêts de système, aux pertes de production et aux réparations. Cet ajustement optimal est particulièrement essentiel dans les secteurs où une grande fiabilité et disponibilité des systèmes sont exigées pour répondre aux besoins de sécurité opérationnelle et aux contraintes de coûts des processus.

La commande optimale du taux d'inspection, en tant qu'élément essentiel de la maintenance préventive, joue un rôle clé dans la réalisation de cet équilibre. La fréquence des inspec-

tions détermine la fréquence à laquelle l'état d'un système est évalué, influençant ainsi la planification de la maintenance et la nécessité d'interventions.

Bien qu'effectuer des inspections trop fréquemment entraîne des coûts de maintenance excessifs, des inspections trop espacées peuvent augmenter la probabilité de défaillances et les coûts d'arrêt associés. Par conséquent, identifier le taux d'inspection optimal est crucial pour minimiser les coûts globaux de maintenance tout en améliorant la fiabilité du système.

1.2.1 Maintenance conditionnelle (CBM)

Comme introduit dans des ouvrages tels que ceux de Demitri Kececioglu [3], les distributions statistiques couramment utilisées pour modéliser les temps de défaillance comprennent les distributions de Weibull, exponentielle, uniforme, Rayleigh, normale et log-normale.

Formules pour la maintenance préventive :

$$C(t_p) = \frac{C_p + \mathbb{E}[N(t_p)] \cdot C_f}{t_p}, \quad \text{pour un remplacement à intervalle fixe} \tag{1.13}$$

$$C(t_p) = \frac{C_p \cdot R(t_p) + C_f \cdot F(t_p)}{t_p \cdot R(t_p) + \int_0^{t_p} t \cdot f(t)\, dt}, \quad \text{pour un remplacement à âge fixe} \tag{1.14}$$

$C(t_p)$: Coût moyen par unité de temps sous la politique de remplacement

C_p : Coût de la maintenance préventive

C_f : Coût de la défaillance (action corrective après panne)

$N(t_p)$: Nombre de défaillances

$\mathbb{E}[N(t_p)]$: Nombre moyen de défaillances pendant l'intervalle $[0, t_p]$

t_p : Temps de remplacement planifié

$R(t_p)$: Fonction de fiabilité (probabilité de non-défaillance jusqu'à t_p)

$F(t_p)$: Fonction de répartition des défaillances, i.e., $F(t_p) = 1 - R(t_p)$

$f(t)$: Fonction de densité de probabilité du temps de défaillance

Ainsi, selon les formules ci-dessus pour la maintenance préventive, qu'il s'agisse de remplacements à intervalle fixe ou à âge fixe, le coût total de la maintenance préventive peut être déterminé en fonction du coût des défaillances, du coût des actions de maintenance, du nombre de défaillances et de la distribution de fiabilité associée. Cette estimation des coûts permet une planification plus approfondie et une optimisation des stratégies de maintenance,

favorisant la rentabilité.

Les activités attendues visent à orienter les décisions tant en matière de maintenance préventive que corrective afin de minimiser les défaillances et de réduire les temps de maintenance et de réparation, évitant ainsi à la fois la surconception et la sous-conception. Ces activités peuvent inclure la détermination des taux d'inspection optimaux, la planification des pièces de rechange et des tâches connexes.

1.2.2 Surveillance en ligne de l'état

En complément des méthodes précédemment utilisées, la surveillance en ligne de l'état est une approche adaptée pour diagnostiquer l'état des systèmes à partir de données en temps réel, permettant ainsi de prédire à l'avance les états futurs du système et d'estimer sa durée de vie (durée de vie restante). Par conséquent, l'évaluation du système est un processus continu qui nécessite une maintenance et une surveillance constantes pour garantir le respect durable des normes en vigueur.

1.2.3 Détermination de l'intervalle de maintenance et de la durée de vie restante (RUL)

Un autre aspect clé de la maintenance prédictive concerne la prévision de la durée de vie restante (RUL) des composants du système. Une combinaison de modélisation physique, de connaissances d'experts et d'approches basées sur les données est essentielle pour une estimation précise de la RUL et une planification efficace de la maintenance.

En conséquence, des études récentes se sont concentrées sur l'analyse et la préparation des données, menant à des résultats issus de l'application de techniques d'apprentissage automatique. L'objectif est de développer des modèles prédictifs capables de déterminer le moment optimal pour le remplacement des composants au sein des systèmes. La formulation d'un programme de maintenance préventive implique un équilibre entre le risque de gaspillage de durée de vie utile et la fréquence des arrêts imprévus dus aux pannes.

Le principal défi réside dans l'utilisation efficace des données de durée de vie pour produire des prévisions précises de la RUL, ce qui est vital pour la planification efficace de la maintenance. Diverses techniques, notamment la régression et la classification, ont été appliquées pour estimer la RUL. L'identification de la RUL demeure un élément critique de la planification de la maintenance préventive. L'objectif principal est d'exploiter les données sensibles au temps afin d'entraîner des modèles statistiques et d'apprentissage automatique capables de prédire la RUL avec précision, comme à travers l'utilisation du paramètre d'échelle (Θ) et

du paramètre de forme (β), exprimée par la formule suivante :

$$\mathrm{RUL}(t_r) = \mathbb{E}(T - t_r \mid T > t_r) = \frac{\int_{t_r}^{\infty}(t - t_r)f(t)\,dt}{R(t_r)} = \frac{\int_{t_r}^{\infty} R(t)\,dt}{R(t_r)} \tag{1.15}$$

1.3 Programmation dynamique

En tant que technique mathématique, la *programmation dynamique* est une méthode structurée utilisée pour prendre une séquence de décisions interdépendantes et résoudre des problèmes complexes d'optimisation. Il s'agit d'une approche polyvalente capable de traiter à la fois des problèmes linéaires et non linéaires. La programmation dynamique est généralement utilisée dans les cas où un problème peut être décomposé en plusieurs sous-problèmes interconnectés et où aucune technique analytique standard ne permet de le résoudre efficacement.

Bien que les principes fondamentaux de la programmation dynamique soient bien compris, leur application efficace à des problèmes variés et l'obtention de solutions optimales nécessitent une attention et une précision rigoureuses. Elle est qualifiée de "dynamique" car elle résout les problèmes de manière séquentielle, en ne considérant que les solutions réalisables à chaque étape. Cette décomposition systématique d'un problème en sous-problèmes plus petits et gérables — connue sous le nom de *staging du problème* — a constitué la base du développement de la programmation dynamique par Richard Bellman en 1957. La programmation dynamique est largement appliquée dans divers contextes de planification et d'optimisation, y compris la gestion des stocks, la planification de la production, la planification de la maintenance, la prise de décisions d'investissement, l'allocation des ressources, l'appariement de produits et les systèmes de contrôle.

Définition des concepts de base :

Dans le cadre de tout problème de programmation dynamique, les facteurs fondamentaux suivants doivent être clairement définis :

1. **Étape (notée n) :** Un problème de programmation dynamique est décomposé en sous-problèmes plus petits, appelés étapes, chacun nécessitant une décision.
2. **État (noté (n, i)) :** À chaque étape, le décideur doit déterminer l'état du système. Chaque étape comporte plusieurs états possibles, et une caractéristique fondamentale d'un état est qu'il sert de lien entre les étapes consécutives.
3. **Action (notée k) :** Dans chaque état, un ensemble d'actions (également appelées variables de décision) est disponible, parmi lesquelles une ou plusieurs doivent être choisies. Ces choix influencent directement les transitions entre états.

4. **Procédure :** Une procédure représente la stratégie choisie pour résoudre le problème. Elle spécifie quelle action doit être entreprise à chaque état, orientant ainsi le système de l'état actuel vers l'état final. La procédure optimale est la politique de décision qui donne le meilleur résultat possible.
5. **Revenu (noté $r(n, i, k)$) :** Cette fonction représente le résultat obtenu à chaque étape. Elle peut prendre diverses formes, telles que bénéfice, économie de coûts, revenu ou toute autre mesure de performance du système.
6. **Fonction de transition d'état (notée $j = t(n, i, k)$) :** Cette fonction détermine comment le système passe d'un état à un autre au fil des étapes. Elle est particulièrement importante dans les problèmes impliquant des espaces d'états continus, car elle régit l'évolution du système en fonction des décisions prises.
7. **Valeur d'un état (notée $f(n, i)$) :** La fonction de valeur d'un état représente la valeur espérée cumulée (revenu ou coût) si le système suit une procédure particulière à partir de cet état. Elle quantifie l'intérêt d'être dans un état spécifique selon une stratégie de décision optimale.
8. **Équation de récurrence :** Lors de la résolution d'un problème de programmation dynamique, une fois les éléments fondamentaux définis et les symboles appropriés attribués, l'équation de récurrence peut être formulée. Cette équation constitue l'outil de calcul principal permettant d'obtenir la solution optimale en évaluant récursivement la valeur de chaque état.

Avant d'introduire l'équation de récurrence, il est essentiel de souligner que la programmation dynamique repose fondamentalement sur la décomposition d'un problème en plusieurs étapes et la détermination de la solution optimale par la sélection de la meilleure décision à chaque étape. En général, connaître l'état actuel dans un problème de programmation dynamique fournit toutes les informations nécessaires pour dériver la solution optimale. En d'autres termes, pour un état donné, la procédure optimale pour les étapes restantes est indépendante des décisions prises lors des étapes antérieures. De plus, à chaque étape, les solutions optimales de toutes les étapes précédentes restent disponibles. L'équation de récurrence, qui est formulée en fonction de la valeur optimale d'un état tout en tenant compte de toutes les actions possibles durant les transitions, est définie comme suit :

$$f(n, i) = \min\left[r(n, i, k) + f(n-1, j)\right] \quad \text{où } j = t(n, i, k) \tag{1.16}$$

1.4 Chaînes de Markov

Chaîne de Markov. Considérons X_t comme une variable aléatoire dépendant du temps t. Une suite de variables aléatoires $X_1, X_2, \ldots, X_k$ forme un processus stochastique. En général, l'indice temporel t peut appartenir à un ensemble T, qui peut être continu ou discret. Dans la majorité des cas d'intérêt, T est associé à des temps discrets et est souvent représenté comme un ensemble d'entiers non négatifs : $T = \{0, 1, 2, \ldots\}$.

Dans sa forme générale, la distribution de probabilité de la variable aléatoire X_n peut dépendre de l'ensemble de l'historique des valeurs précédentes, ce qui s'exprime mathématiquement par :

$$P(X_n = x_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1}). \tag{1.17}$$

Un modèle qui capture pleinement une telle dépendance est souvent complexe et coûteux en termes de calcul. Cependant, dans de nombreux cas pratiques, on suppose la *propriété de Markov*, selon laquelle la valeur de X_n ne dépend que de sa valeur précédente immédiate X_{n-1}. Cette hypothèse simplifie l'expression de probabilité conditionnelle :

$$P(X_n = x_n \mid X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = P(X_n = x_n \mid X_{n-1} = x_{n-1}). \tag{1.18}$$

Classification des états

Tout comme une chaîne physique peut adopter différentes structures, une chaîne de Markov peut également présenter diverses caractéristiques structurelles, chacune correspondant à une classification particulière.

Définition 1 : Une chaîne de Markov dans laquelle tous les états communiquent entre eux, c'est-à-dire qu'il est possible d'atteindre n'importe quel état à partir de n'importe quel autre, est appelée une *chaîne de Markov irréductible*.

Au sein d'un processus de Markov à chaîne unique, les états individuels peuvent être classés en fonction de leur comportement à long terme, comme indiqué ci-dessous.

Définition 2 : Un état est dit *transitoire* s'il existe une probabilité non nulle que, une fois que le système quitte cet état, il n'y revienne jamais. Un état transitoire est ainsi nommé parce que, à long terme, le système ne le traverse qu'un nombre fini de fois.

Définition 3 : Un état qui n'est pas transitoire est dit *récurrent*, ce qui signifie que si le système entre dans cet état, il est certain qu'il y reviendra un jour.

Types d'états récurrents

Les états récurrents peuvent être classés en trois catégories :

1. États absorbants
2. États périodiques
3. États apériodiques (également appelés états ergodiques dans certains contextes)

Définition 4 : Un état est dit *absorbant* si, une fois que le système y entre, il est impossible d'en sortir. Autrement dit, la probabilité de quitter cet état est nulle.

Définition 5 : Un état est dit *périodique* si le système y revient à intervalles de temps fixes et réguliers. Il existe alors un entier $d > 1$ tel que le système ne peut revenir dans cet état qu'à des multiples de d unités de temps.

Définition 6 : Un état est dit *apériodique* s'il est certain que le système y reviendra, mais que le moment exact du retour est imprévisible, ce qui signifie qu'il ne suit pas un intervalle fixe. Ces états sont souvent appelés *états ergodiques*, en particulier dans les contextes d'analyse du comportement à long terme et de l'équilibre.

Propriétés et types de problèmes résolus par les chaînes de Markov

Les problèmes de programmation dynamique probabiliste formulés à l'aide de chaînes de Markov pour la prise de décision à long terme présentent les propriétés clés suivantes :

— L'horizon de planification est considéré comme infini, ce qui signifie que le système est analysé sur un nombre illimité d'étapes temporelles.
— L'espace des états du système est fini, ce qui signifie que le système peut occuper un nombre fixe d'états possibles. Cet ensemble d'états est noté : $E = \{1, 2, \ldots, N\}$.
— La fonction de transition d'état est définie à l'aide d'un vecteur de probabilités de transition, exprimé comme suit :

$$p_{i,j}(k) = P(X_{n+1} = j \mid X_n = i, d_n = k), \quad k \in D \tag{1.19}$$

où :
— X_n représente l'état du système à l'instant n.
— d_n est la décision ou action prise à l'instant n.
— La somme des probabilités de transition vers tous les états suivants doit être égale à un : $\sum_j p_{i,j}(k) = 1$.

— La récompense ou le gain associé à l'action k dans l'état i est donné par une fonction $r(i, k)$, qui définit le résultat — tel que le coût, le bénéfice ou l'utilité — résultant d'une décision particulière dans un état donné.

CHAPITRE 2 REVUE DE LA LITTÉRATURE

2.1 Revue de la littérature sur le contrôle optimal des files d'attente

Un aperçu des recherches antérieures et des études fondamentales liées au premier axe de recherche est présenté au début de ce chapitre :

La théorie des files d'attente fournit un cadre de modélisation mathématique pour étudier et analyser le flux de demandes au sein d'un système. L'objectif principal est d'examiner le comportement du système et d'obtenir les indicateurs clés de performance, tels que le temps d'attente moyen des clients, l'efficacité du système et le nombre moyen de clients dans le système.

Cette revue de littérature sur les systèmes de files d'attente avec des paramètres stochastiques dépendant du temps explore les recherches antérieures qui ont examiné le comportement des files d'attente sous l'aléa des arrivées et des temps de service. Ces études appliquent des techniques de modélisation mathématique pour évaluer comment la dynamique du système de files d'attente est affectée par différentes hypothèses de distribution concernant les arrivées de clients et les temps de service. De plus, ces travaux traitent des effets de différentes stratégies de contrôle sur la performance des systèmes de files d'attente, telles que les politiques d'ordonnancement ou l'allocation de ressources, qui fonctionnent dans un contexte de temporalité stochastique. En outre, cette revue explore les applications pratiques de la théorie des files d'attente dans divers domaines, y compris l'analyse des systèmes bancaires, des industries de services et d'autres contextes réels.

L'analyse des systèmes de files d'attente avec des intervalles de temps aléatoires reste un sujet de recherche actif dans les domaines de la recherche opérationnelle et de la théorie des files d'attente, avec plusieurs publications récentes mettant l'accent sur les stratégies de contrôle des files d'attente.

Dans les travaux existants, plusieurs modèles ont été introduits pour considérer la perspective générale de l'« horizon temporel infini », qui sont listés ci-dessous :

Cet ensemble d'études comprend la recherche de Wang [4] sur les files d'attente dont le serveur n'est pas fiable en raison de pannes, entre autres problèmes, et propose un modèle de coût pour la politique optimale ; dans un autre article, Kumar et al. [5] ont proposé une méthode numérique pour résoudre une classe de problèmes de contrôle stochastique. Ke [6] a étudié la file $M/G/1$ pour déterminer une commande optimale et les seuils minimisant le coût en tenant compte des vacances distribuées généralement et des pannes du serveur. Aksenova et al. [7]

ont considéré deux files FIFO et des marches aléatoires bidimensionnelles, et ont présenté des modèles mathématiques et des algorithmes optimaux pour contrôler les systèmes. Banik [8] a présenté une analyse pour une file à serveur unique ayant la caractéristique d'un processus d'arrivée markovien ; la politique optimale à coût minimal ainsi que les distributions de la longueur de file et les mesures de performance ont été obtenues. Özkan et al. [9] ont étudié un système de file markovienne à deux serveurs et ont conclu que le routage des clients vers le serveur le plus rapide serait toujours optimal ; pour les serveurs plus lents, une politique optimale à seuil existe selon la longueur de la file et l'état du serveur le plus rapide. Ahmed et al. [10] ont étudié les problèmes de contrôle de puissance stochastique sur un canal avec file de données et contraintes de délai ; des algorithmes de contrôle et leur efficacité ont été proposés et vérifiés par des simulations numériques.

Grishunina [11] a travaillé sur l'optimisation des systèmes de files $M/G/1/\infty$, dans lesquels il a été montré que si une politique optimale existe, alors elle serait de type seuil. Jiang et al. [12], afin de dériver la stratégie de commande optimale, ont proposé des méthodes d'apprentissage par renforcement pour la gestion active des files d'attente, et une détection précoce aléatoire est proposée ; la longueur moyenne de la file serait contrôlée en tenant compte de diverses charges du réseau. Asadzadeh et al. [13] ont contribué à la littérature en étudiant la performance d'une station-service ; pour optimiser le système de file d'attente, un modèle de régression du second ordre a été proposé à partir de données simulées, et le nombre optimal de pompes et d'opérateurs de pompes a été déterminé, en tenant compte des ventes de carburant et des coûts. Lefebvre [14] a modifié le modèle de file $M/M/m/n$ qui permet aux tâches de requérir jusqu'à m serveurs et d'avoir un temps d'attente aléatoire avant exécution. Luo et al. [15] ont proposé une politique de gestion de file tenant compte à la fois du nombre de clients et de la charge de travail du serveur ; la longueur de file et le coût attendu à long terme sont analysés, et les politiques optimales sont comparées à l'aide d'exemples numériques.

Dans le deuxième ensemble d'articles, d'autres modèles sont proposés sur la base d'horizons temporels prédéfinis, qui sont présentés comme suit :

Fan-Orzechowski [16] a étudié l'admission optimale de clients dans une file à capacité finie, contenant plusieurs types de clients ; une solution via la programmation linéaire est proposée, dans laquelle des politiques optimales sont caractérisées et une structure est établie à cette fin. Dans la recherche de Cao et al. [17], une politique optimale est examinée pour servir des clients dont la valeur évolue ; une politique heuristique et une étude numérique sont proposées, montrant que dans la plupart des cas cette politique est proche de la solution optimale. Lefebvre [18] discute d'un modèle de théorie des files d'attente, dans lequel le

débit d'une rivière est prédit pendant les périodes de hautes eaux ; en outre, il présente une application démontrant la pertinence du modèle de file $M/M/1/c$ lorsque le débit dépasse un certain seuil, utilisé pour prévoir une augmentation des taux d'événements. Wen [19] a proposé un modèle de problème optimal pour les clients à haute priorité, visant à minimiser le coût total d'attente en assignant des cibles de temps d'attente ; pour résoudre les problèmes à grande échelle, une politique de seuil est présentée comme solution de commande optimale, validée par des expériences numériques. Bhati [20] a introduit un problème de commande optimal, dans lequel l'âge moyen de l'information est minimisé en tenant compte de serveurs hétérogènes.

Le troisième groupe d'articles traite de l'objectif d'« horizon temporel infini » dans le cas distinct de « réseau de files », qui sont présentés dans les paragraphes suivants : Dans un article, une file d'attente exponentielle multi-serveur a été analysée par Efrosinin [21] et des méthodes pour calculer les probabilités à l'état stationnaire ont été démontrées ; les distributions des temps d'attente et de séjour ont été dérivées, et enfin les caractéristiques de performance du système de commande optimale ont été comparées aux politiques de contrôle heuristiques. De Santi [22] a introduit une politique de réseau pour trouver une gestion optimale des files d'attente ; cela est comparé à la détection précoce aléatoire et a montré de meilleures performances dans les réseaux à forte latence de bande passante. Une politique de contrôle de réseau a été proposée par Banirazi [23] afin de minimiser le délai réseau sous des conditions temporelles variables dans des réseaux sans fil stochastiques, laquelle est optimale et mise en œuvre sans connaissance préalable des statistiques du système. Un algorithme de gestion de file d'attente a été proposé par Ping et al. [24] pour traiter les systèmes Internet à grande latence, afin de prédire la longueur de la file d'attente et d'améliorer la stabilité et la robustesse du système. Afin de calculer les caractéristiques d'une file d'attente à source finie pour maximiser le revenu total, Ponomarov et al. [25] ont présenté une solution au problème en considérant des modèles de Markov bidimensionnels et tridimensionnels et des politiques de contrôle par seuil. Pour évaluer et classer les réseaux Internet, Chavari et al. [26] ont utilisé l'Analyse Enveloppante des Données ; en simulant un réseau de services et en introduisant un modèle de cross-efficacité pour le classement, le réseau le plus performant a été sélectionné. Zheng [27] a exploré la stabilité et la planification optimale de réseaux stochastiques dans des environnements aléatoires ainsi que les problèmes de contrôle de diffusions avec sauts. Afin de maximiser le taux de détection de données à long terme des dispositifs sans fil, Li et al. [28] ont proposé un algorithme pour un système mobile sans fil visant l'efficacité énergétique, sous la contrainte de stabilité de la file d'attente de données. Reticcioli et al. [29] ont utilisé des modèles prédictifs du comportement des dispositifs réseau pour permettre un contrôle plus flexible et plus facile du réseau ; les données historiques recueillies à partir du réseau

sont utilisées à cet effet. Petrović et al. [30] ont étudié le problème des péages autoroutiers et proposé une méthodologie de commande optimale en temps continu, dans laquelle le nombre optimal d'unités actives dans les stations est prédit à l'aide de réseaux neuronaux et de la théorie des files d'attente. Dans le but de mettre en œuvre un contrôle prédictif pour les bus, Gao et al. [31] ont utilisé un modèle pour estimer le temps d'inactivité et prédire les variations du trafic aux intersections à l'aide d'une architecture de cloud véhiculaire, afin de réaliser des économies d'énergie et de temps d'attente par rapport aux décisions humaines.

Enfin, Fu et al. [32] ont considéré les deux objectifs de « réseau de files » et de « temps prédéfini », et étudié le commande optimale dans des réseaux de files stochastiques ; afin de maximiser l'utilité totale obtenue à la fin d'un horizon temporel fini. Un ensemble de tâches reliées à un ensemble de serveurs parallèles avec files d'attente a été considéré pour proposer une politique de contrôle.

De nombreuses études dans la littérature révèlent un écart de recherche notable concernant l'intégration du temps aléatoire dans les problèmes de commande optimale, que nous explorerons plus en détail dans la section suivante. La diversité des méthodologies utilisées dans le domaine de la théorie des files d'attente est démontrée dans ces études, où des intervalles de temps aléatoires sont discutés. Ces méthodes incluent la modélisation analytique, l'analyse numérique et les techniques d'optimisation.

2.2 Revue de la littérature sur la distribution des temps de service dans les systèmes de files d'attente

Ensuite, ce chapitre se poursuit par une revue de la littérature relative aux travaux antérieurs liés à cette thématique.

La théorie des files d'attente intègre la modélisation mathématique des arrivées de clients et de leurs temps de service correspondants dans les systèmes de files d'attente et joue un rôle unique dans la prise de décision en fournissant des perspectives sur le fonctionnement et l'efficacité des systèmes. L'optimisation des décisions est un aspect essentiel de la théorie des files d'attente qui permet aux décideurs d'optimiser les mesures de performance en se concentrant sur l'examen des paramètres fondamentaux et des contraintes du modèle. Plusieurs configurations de files peuvent être évaluées, y compris la vitesse de service et le nombre de serveurs, afin d'établir la politique la plus rentable.

L'analyse de sensibilité constitue un autre aspect essentiel de la théorie des files d'attente dans le cadre du processus décisionnel. Elle permet aux décideurs de comparer divers scénarios en modélisant plusieurs configurations de files d'attente, leur permettant ainsi d'examiner les

effets des variations de paramètres dans le système, notamment en termes de coût associé aux changements dans les schémas de service des clients. De plus, l'analyse de sensibilité permet aux décideurs de comprendre les résultats potentiels résultant de différentes configurations du système. Cela facilite des actions éclairées basées sur la compréhension des dynamiques des systèmes, comme discuté dans Hosseinzadeh Lotfi et al. [33].

Une revue de littérature complète est effectuée pour observer les études qui traitent du problème de l'identification de la vitesse de service optimale dans les systèmes de files d'attente à tout instant donné, et donc du contrôle de l'efficacité du système dans un cadre de temps aléatoire continu.

De nombreux articles ont étudié des variantes de la vitesse de service afin de réduire les files d'attente. Al-Hawari et al. [34] ont examiné la position optimale du capteur et étudié son impact sur la performance du système. Dans cette étude, un système de production contrôlé est analysé à l'aide d'un modèle de simulation contenant trois stations, y compris un goulot d'étranglement. Srinivas et al. [35] ont utilisé des processus de Poisson pour examiner un modèle de file d'attente avec des taux de service dépendant de l'état et ont analysé les comportements transitoires. Les mesures de performance sont dérivées et les applications pratiques dans les réseaux de communication sont mises en évidence. Li et al. [36] ont examiné une file markovienne à serveur unique avec un taux de service variable. Les coûts d'attente, les récompenses et les stratégies d'équilibre pour tous les clients sont pris en compte. Chen et al. [37] ont contribué au domaine pour capturer des modèles similaires du trafic Internet. En se concentrant sur le modèle de trafic de mouvement brownien fractionnaire, les limitations du modèle et la vitesse de service sont explorées.

Lee et al. [38] se sont concentrés sur les applications aux centres de données pour optimiser une file multiserveur avec des taux d'arrivée et de service variables. Pour explorer les politiques de commande optimales, un processus de décision de Markov en temps continu est utilisé, basé sur la rétention des clients, les coûts de fonctionnement des serveurs et les récompenses de fonctionnement du système. Munoz et al. [39] ont fourni des analyses de système et amélioré la théorie floue des files d'attente en introduisant une méthode de simulation pour modéliser des arrivées et taux de service variables. L'impact des véhicules électriques sur les réseaux électriques est étudié par cette approche ; pour une évaluation précise de la performance du réseau, leurs demandes de recharge sont modélisées comme un problème de files d'attente floues. Delasay et al. [40] ont introduit un modèle de file avec des vitesses de serveur adaptatives, pouvant répondre aux périodes de forte charge du système. Le modèle met en évidence les décalages potentiels entre performance attendue et réelle, ce qui impacte les décisions d'affectation de personnel. Pour réduire les temps d'attente dans un modèle

$M/G/1$, une distribution de Poisson étendue a été introduite par Rodrigues et al. [41] pour modéliser la congestion d'une file finie, validée par simulation et données réelles.

Pour passer en revue les travaux connexes et comparer les méthodes récentes de contrôle des vitesses de service variables, les articles suivants ont été listés.

Dans le but d'optimiser les vitesses de service pour minimiser les coûts, Laxmi et al. [42] ont examiné une file en temps discret à l'aide de l'optimisation par essaim particulaire et exploré son application dans les centres de contact email entrants. Pour maximiser le bien-être en ajustant de manière optimale les prix d'admission et les vitesses de service, Chen et al. [43] ont présenté une méthode utilisant l'optimisation basée sur la sensibilité pour un commutateur de réglage de débit, avec des expériences numériques pour valider l'approche. En utilisant un modèle pour équilibrer la qualité de service et les coûts des ressources humaines, Dai et al. [44] ont décrit un nouveau modèle de service client en ligne dans lequel les agents peuvent aider simultanément plusieurs clients ; l'approche a ensuite été validée par simulations. Büke et al. [45] ont utilisé des processus stochastiques pour l'analyse, et pour comprendre ces processus sous différentes politiques de routage, les auteurs ont employé des martingales dans l'étude. Ainsi, des systèmes de files d'attente à plusieurs serveurs avec taux de service variables ont été explorés. Su et al. [46] ont établi la politique optimale en se concentrant sur des vitesses de service variables dans une file à serveur unique dans des réseaux de communication avec plusieurs types de clients.

Laxmi et al. [47] ont travaillé sur la stabilité du système et les mesures de performance pour optimiser la vitesse de service en analysant un système $M/M/2$ avec un deuxième service optionnel. Enfin, l'impact de divers paramètres du modèle a été illustré par des résultats numériques. Pour développer une politique optimale de vitesse de service et minimiser les coûts, Lakkumikanthan et al. [48] ont utilisé la théorie de la décision markovienne et la programmation linéaire pour un système de files d'attente-inventaire en temps discret. Ensuite, les vitesses de service sont ajustées en fonction des arrivées de clients et du réapprovisionnement de l'inventaire. Un modèle de files d'attente avec deux types de requêtes a été analysé par Dudin et al. [49] pour explorer le comportement du système à l'aide d'une chaîne de Markov : le type 1 avec un taux de service constant et le type 2 avec des taux flexibles et un partage de processeur.

Pour un aperçu d'autres études dans le domaine concerné, le lecteur est invité à consulter Chen et al. [50], dans lequel des politiques de commande optimales sont développées via une optimisation basée sur les événements afin d'explorer le contrôle de la vitesse de service. Un algorithme est introduit pour déterminer les vitesses de service et son efficacité est comparée à l'optimisation basée sur les états par une analyse numérique. Wu et al. [51] visent à réduire

les coûts et les temps d'attente en menant une analyse à l'état stationnaire pour l'optimisation du système. Une file $GI/M/1$ avec des vitesses de service variables est examinée pour contrôler l'entrée des clients. Singh et al. [52] analysent les changements dans la distribution de probabilité d'un processus stochastique pour détecter les points de changement dans la vitesse de service d'une file $M/M/1/m$. Une fonction de vraisemblance basée sur le nombre de clients observés aux départs est utilisée. Des simulations de Monte Carlo montrent l'efficacité et la validité de cette approche. Enfin, un système de file markovien avec un seul serveur offrant des vitesses de service variables est examiné au moyen de chaînes de Markov en temps continu pour modéliser et analyser comment les décisions des clients affectent la performance à l'équilibre de la file, dans l'étude de Tian et al. [53].

La littérature ci-dessus tend à illustrer les avancées fondamentales et à identifier les lacunes dans la littérature existante que cette recherche vise à combler.

2.3 Revue de la littérature sur l'inspection et la maintenance en contexte stochastique

Enfin, nous clôturons ce chapitre en présentant le contexte des études existantes liées à l'objectif principal de cet axe de recherche.

À travers une analyse approfondie des développements dans le domaine de la commande optimale des chaînes de Markov en temps continu, notamment dans le contexte de la théorie de la fiabilité, une attention significative à ce domaine est évidente, en particulier dans le cadre de la planification de la maintenance. Les études antérieures couvrent un large éventail de modèles et de méthodologies visant à obtenir des politiques de maintenance optimales pour améliorer la fiabilité, la disponibilité et la sécurité des machines tout en minimisant simultanément les coûts. Ces travaux indiquent les efforts continus en vue de l'amélioration des techniques de commande optimale dans ce domaine.

Les premiers développements en matière de commande optimale et de planification de maintenance ont été introduits par Parmigiani [54] pour identifier les meilleures fréquences de maintenance ; une chaîne de Markov en temps continu a été appliquée et un modèle de commande optimale pour un système de fabrication a été proposé. Un développement ultérieur de ce modèle a été effectué par Boukas et al. [55] ; le modèle se concentre sur l'amélioration de la fiabilité du système et discute du contrôle de maintenance en utilisant un processus de Markov en temps continu en appliquant la commande optimale des taux de maintenance préventive et corrective. Ensuite, en développant des modèles de politiques de maintenance et d'inspection, Wang et al. [56] ont contribué au domaine et développé une politique optimale

pour minimiser le coût total ; un modèle de chaîne de Markov en temps continu a été utilisé pour les intervalles d'inspection et la maintenance de la détérioration dans un système de production.

Suryadi et al. [57] ont élargi ce champ en fournissant une méthode de programmation mathématique et en utilisant des chaînes de Markov en temps continu pour modéliser la maintenance afin d'optimiser la planification de la maintenance dans les usines. Afin de permettre des réparations imparfaites avant le remplacement du composant, Liying et al. [58] ont optimisé le cycle d'inspection pour maximiser le revenu et le diagnostic de défaillance, et la politique d'inspection a été identifiée ; le processus vectoriel de Markov et un exemple numérique ont été utilisés pour valider la méthodologie. En prenant en compte la commande optimale dans des conditions aléatoires selon l'état de santé du patient, un modèle de processus décisionnel de Markov partiellement observable a été utilisé par Erenay [59] ; une stratégie de dépistage optimale a été fournie pour prévenir le cancer et surveiller, et le taux optimal de coloscopies a été déterminé.

Dans le but de trouver une politique de commande optimale pour maximiser la disponibilité du système, Jiang et al. [60] ont modélisé un processus de Markov en temps continu pour fournir une approche de prévision des défaillances dans un système observable avec pannes. Pour trouver une politique de commande optimale qui optimise les coûts de maintenance et d'inspection, et avec l'objectif de maximiser les profits en tenant compte de la dégradation du système sous contrôle, Ahmadi et al. [61] ont utilisé le processus de Markov pour fournir un modèle d'optimisation de la planification de maintenance et des stratégies d'inspection dans un système en dégradation.

Dans l'optimisation des politiques de maintenance, la maintenance conditionnelle (CBM) est apparue comme une approche très efficace qui dépend de l'état en temps réel de l'équipement plutôt que de suivre des intervalles prédéfinis. L'intégration de la maintenance conditionnelle et des politiques de commande optimale a suscité une attention notable de la part des chercheurs ; Golmakani et al. [62] ont utilisé le processus de Markov et proposé un modèle visant à optimiser les coûts de maintenance avec des intervalles d'inspection équilibrant les coûts de remplacement préventif et de défaillance ; une stratégie de remplacement optimale et une fréquence d'inspection ont été fournies pour la maintenance conditionnelle. D'autre part, dans le but de minimiser les coûts tout en maintenant la fiabilité par des politiques de maintenance et d'inspection efficaces, Wang [63] a cherché à optimiser les fréquences de surveillance à l'aide de la théorie bayésienne et d'un modèle non markovien. Une politique de commande optimale pour un système vieillissant a été proposée en mettant en évidence les stratégies optimales de maintenance préventive et de remplacement par Kim et al. [64] ; une chaîne de

Markov en temps continu a été utilisée dans leur travail, se concentrant sur la minimisation du coût total. Liu et al. [65] visaient à optimiser les intervalles d'inspection pour minimiser les coûts de maintenance et ont utilisé un modèle de chaîne de Markov en temps continu avec maintenance conditionnelle tandis que Kutzner et al. [66] ont pris en compte les coûts de réparation et la qualité de la production pour minimiser les coûts attendus en considérant les stratégies de maintenance et d'inspection et en évaluant la variabilité de la production et les fréquences d'inspection ; en mettant en œuvre un processus de Markov, un modèle de commande optimale a été proposé dans ce travail.

Par conséquent, pour ajuster les intervalles d'inspection selon les états de vieillissement et minimiser les coûts de maintenance, Naderkhani et al. [67] ont développé une stratégie optimale de maintenance conditionnelle en appliquant un processus de Markov caché en temps continu ; les taux d'inspection sont ajustés en fonction des états de dégradation des composants pour réduire les dépenses de maintenance. Une méthode de planification et des applications de l'optimisation de la maintenance dans le domaine de la santé pour l'inspection des stents ont été proposées par Arab et al. [68] ; des modèles de fiabilité et le problème de commande optimale ont été utilisés pour optimiser les plannings d'angiographie tout en équilibrant les risques pour la santé et les coûts totaux. En utilisant un modèle vectoriel autorégressif (VAR), des modèles de Markov cachés en temps continu et une méthode de contrôle bayésien optimal, Lin et al. [69] ont utilisé la prévision de la durée de vie restante et la maintenance conditionnelle ; une méthode de coût optimal pour la détection précoce de défauts dans des boîtes de vitesses à l'aide de données de capteurs a été développée dans cette étude.

En utilisant un modèle d'optimisation pour la maintenance conditionnelle et un cadre de processus décisionnel de Markov partiellement observable, Kim [70] a abordé le contrôle robuste pour les systèmes avec pannes dans l'industrie minière ; par conséquent, des stratégies qui restent efficaces en présence d'aléas dans le modèle ont été définies dans ce travail. La méthode de planification de maintenance optimale dans les systèmes de santé a été développée dans une thèse à l'aide de processus stochastiques par He [71] ; les événements avec des intervalles d'inspection imparfaits et une maintenance préventive non ponctuelle y sont abordés. Pour améliorer la fiabilité à l'aide d'un cadre markovien par des stratégies de maintenance optimales, Mousavi et al. [72] ont optimisé les politiques de réparation en maintenance conditionnelle ; les coûts de réparation, la probabilité de défaillance et les stratégies de remplacement préventif ont été intégrés pour minimiser les coûts attendus. Pour modéliser la dégradation afin de minimiser les coûts de maintenance, Li et al. [73] ont utilisé un processus semi-markovien caché en temps continu et une maintenance optimale pour la détection précoce de défauts dans une boîte de vitesses d'hélicoptère a été proposée.

De plus, Sun et al. [74] ont utilisé le processus de décision de Markov pour proposer des stratégies optimales d'inspection et de remplacement pour les systèmes vieillissants ; afin de modéliser la dégradation des composants et ainsi trouver les intervalles d'inspection et les choix de remplacement pour minimiser le coût opérationnel total, l'étude utilise le processus de Wiener. Dans l'objectif de la gestion de la santé et de la maintenance prévisionnelle (PHM), Liu et al. [75] ont présenté une stratégie de maintenance dynamique ; pour établir le moment optimal d'inspection, un modèle a été développé pour analyser la fiabilité du système, combinant les effets de la dégradation et du vieillissement. Dans un autre travail utilisant un processus de décision de Markov partiellement observable, Liu et al. [76] ont développé des politiques de maintenance qui intègrent l'inspection incomplète ; les politiques de maintenance optimales pour minimiser les coûts sont obtenues en tenant compte des états non observables du système selon les inspections incomplètes.

Pour gérer les incertitudes dans la prise de décision en matière de maintenance, des modèles plus avancés sont introduits dans les travaux récents, qui combinent les processus de décision de Markov partiellement observables et les réseaux bayésiens dynamiques. Andriotis et al. [77] ont développé un modèle de quantification de la valeur pour une structure de surveillance de la santé afin d'optimiser la durée de vie utile des systèmes en dégradation avec incertitude, basé sur un processus de décision de Markov partiellement observable.

L'utilisation de la chaîne de Markov en temps continu et de la programmation dynamique s'étend au-delà des pratiques industrielles classiques. Dans le but d'optimiser les procédures d'évaluation logicielle, un modèle de contrôle stochastique en temps continu a été appliqué par Cao et al. [78], visant à équilibrer les coûts de test par rapport au temps de mise sur le marché afin de réduire les coûts totaux. La programmation dynamique a été utilisée dans le modèle proposé pour déterminer le temps de test optimal. En utilisant l'optimisation stochastique et les méthodes de chaînes de Markov et en appliquant le processus de Wiener pour modéliser la dégradation, Sun et al. [79] ont développé une planification de maintenance pour gérer les composants vieillissants dans les systèmes en série ; la méthode optimise le remplacement préventif et minimise les coûts opérationnels totaux. En se concentrant sur la planification de l'inspection et de la maintenance, une approche de gestion de la santé et de la maintenance prévisionnelle (PHM) a été introduite par Mancuso et al. [80]. En outre, un processus de commande optimale ajustant à la fois la maintenance et les coûts a été étudié en utilisant un cadre de contrôle markovien ; la planification de la maintenance a été développée en utilisant les équations de Hamilton-Jacobi et un temps de défaillance basé sur le processus de Wiener par Ahmadi [81].

De plus, en utilisant un processus de décision de Markov en temps discret et en intégrant la

dégradation du système et la gestion des pièces de rechange, une stratégie de maintenance de commande optimale pour les systèmes de production a été développée par Gan et al. [82]. Wang et al. [83] ont utilisé le processus de décision de Markov pour explorer la commande optimale d'un système avec pièces de rechange limitées sur une période de temps finie afin de développer un plan de maintenance. Afin de minimiser les coûts moyens de maintenance sur une période de temps infinie, Roux et al. [84] ont contribué au domaine et, pour planifier la maintenance sous surveillance imprécise, un modèle efficace de processus de décision de Markov partiellement observable a été introduit dans leur travail ; un algorithme efficace a été développé pour optimiser les politiques de maintenance conditionnelle couvrant l'inspection parfaite, la maintenance préventive et la maintenance corrective. Pour résoudre les limites des méthodes heuristiques, optimiser la fiabilité des structures et minimiser les coûts par l'optimisation stochastique, les réseaux bayésiens dynamiques intégrés aux processus de décision de Markov partiellement observables ont été incorporés par Morato et al. [85] ; la planification optimale de l'inspection et de la maintenance de la détérioration structurelle a été explorée comme résultat de cette étude.

Des avancées plus récentes sur la commande optimale et la programmation dynamique en maintenance ont été proposées par Lefebvre et Pazhoheshfar [86] ; reconnaissant que le temps final sera une variable aléatoire, un problème de commande optimale pour la maintenance des machines est présenté. Dans cette étude, l'usure normale et aléatoire des systèmes est prise en compte afin de maximiser les revenus au cours de la durée de vie de la machine ; la méthodologie consiste à résoudre une équation de programmation dynamique pour identifier si des activités de maintenance doivent être réalisées à chaque unité de temps.

Asadzadeh et al. [87] ont formulé une politique optimale de maintenance conditionnelle pour les panneaux électriques en utilisant un modèle de risque proportionnel. Dans le but de minimiser les coûts attendus, leur modèle tient compte de l'interdépendance des défaillances des composants.

Dans une littérature plus récente, les applications des chaînes de Markov en temps continu ont été appliquées à des contextes complexes de maintenance ; en utilisant un processus de Markov déterministe par morceaux, un modèle de maintenance conditionnelle a été proposé par Wang et al. [88]. Dans le but d'identifier des stratégies de maintenance pour minimiser les coûts en présence de réparations imparfaites, le modèle proposé prend en compte à la fois la dégradation progressive et les chocs aléatoires en appliquant la théorie de la commande optimale. Leur travail est notable, en particulier dans les cas où les deux types de phénomènes de dégradation doivent être gérés efficacement. Enfin, pour obtenir des contrôleurs optimaux à temps fini et à temps infini, Wang et al. [89] abordent la commande optimale des systèmes

avec données échantillonnées et échantillonnage stochastique.

Bien qu'il y ait eu des avancées considérables dans la commande optimale des chaînes de Markov en temps continu dans le cadre de l'optimisation de la maintenance, des lacunes de recherche subsistent encore. De nombreux modèles existants sont formulés pour des applications spécifiques, telles que les systèmes de production ou les soins de santé, sans explorer de manière exhaustive le taux d'inspection optimal dans un cadre plus généralisé. De plus, l'intégration de la programmation dynamique pour obtenir des variables de décision continues, telles que les intervalles de test, n'a pas été suffisamment explorée, en particulier en ce qui concerne les problèmes de type homing dans la théorie de la fiabilité.

Cette étude vise à combler les lacunes existantes en fournissant un cadre complet permettant d'obtenir un taux optimal de maintenance préventive et d'inspection des systèmes en utilisant des chaînes de Markov en temps continu combinées à la programmation dynamique, comme discuté dans Lefebvre et Yaghoubi [90]. Le modèle proposé dans notre étude suggérera une solution robuste au problème de type homing dans l'optimisation en théorie de la fiabilité en intégrant des techniques avancées de contrôle stochastique pour formuler des politiques de maintenance optimales. De plus, en intégrant les dernières avancées en matière d'optimisation de la planification de maintenance avec les approches de programmation dynamique, cette recherche contribue à faire progresser la théorie de la fiabilité et à promouvoir la rentabilité dans les systèmes d'ingénierie. Elle propose des politiques applicables aux industries où la fiabilité des composants est hautement critique.

CHAPITRE 3 OBJECTIFS

Objectifs principaux

Dans le cadre de la théorie des files d'attente, notre objectif principal est de traiter le problème de homing. Cela consiste, par exemple, à contrôler la durée espérée pendant laquelle le processus contrôlé reste dans un intervalle $[a; b]$, soit en prolongeant ou en réduisant cette durée, soit en forçant le processus à sortir de l'intervalle par une extrémité prédéterminée. Un système de file d'attente est considéré sans déterminer explicitement l'horizon temporel d'optimisation, c'est-à-dire en supprimant l'hypothèse conventionnelle d'un horizon temporel infini $(0, \infty)$ ou d'une période prédéterminée $(0, T)$. Dans ce cas, le temps de contrôle est considéré comme une variable aléatoire et l'objectif est de modéliser le problème dans le but d'optimiser l'utilisation du système, où la fonction de coût associée est étudiée sur une période allant de 0 à un instant terminal aléatoire. Cet intervalle de temps représente la durée pendant laquelle la capacité du système est contrôlée ou la période assignée pour servir les clients.

Sous l'hypothèse ci-dessus, dans un premier temps, on a examiné la dynamique des files d'attente en optimisant le nombre de serveurs, comme présenté dans la première partie de la thèse.

Deuxièmement, une analyse de sensibilité du système de file d'attente a été réalisée en ce qui concerne les variations de la vitesse de service, comme discuté en détail dans la partie consacrée à cette thématique.

Par la suite, la recherche a été complétée par l'application des chaînes de Markov en temps continu, en tenant compte des mêmes caractéristiques stochastiques que celles considérées précédemment. L'objectif est d'identifier une politique efficace de maintenance et d'inspection pour le système analysé.

Objectifs dérivés

1. Étudier les techniques d'optimisation des entrées et sorties du système, telles que le nombre de serveurs et la vitesse de service, pour minimiser les temps d'attente et optimiser la performance du service.
2. Évaluer la performance des systèmes de files d'attente avec plusieurs serveurs, dans des conditions où le temps présente une nature aléatoire dans l'ensemble du système.
3. Construire des modèles qui intègrent des intervalles de temps aléatoires pour les pro-

cessus d'arrivée et de service, et explorer leur impact sur la performance du système de file d'attente.

4. Étudier l'influence de différentes distributions de probabilité des temps de service pour en observer les effets sur la performance du système.
5. Développer des solutions à des problèmes pratiques en appliquant la programmation dynamique et les processus stochastiques, tels que le problème de homing et les chaînes de Markov, plutôt que de se limiter à des exemples mathématiques.

Notre méthodologie de recherche étend les modèles existants. La Figure 3.1 illustre les étapes de recherche pour les modèles de files d'attente étudiés dans les premières parties de la thèse.

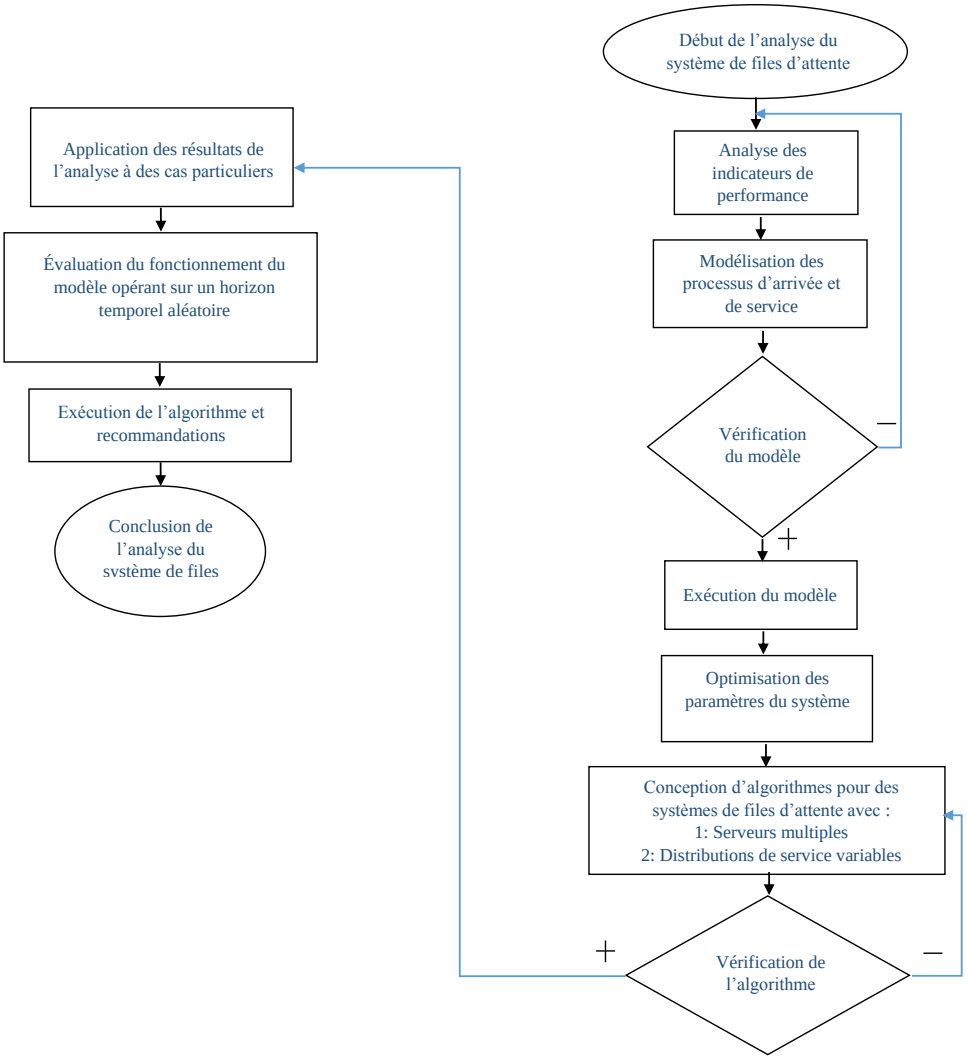

Figure 3.1 Schéma des étapes de recherche des modèles de files d'attente

La Figure 3.2 illustre le schéma détaillé des principales étapes méthodologiques de la recherche appliquées à la politique optimale de maintenance et d'inspection.

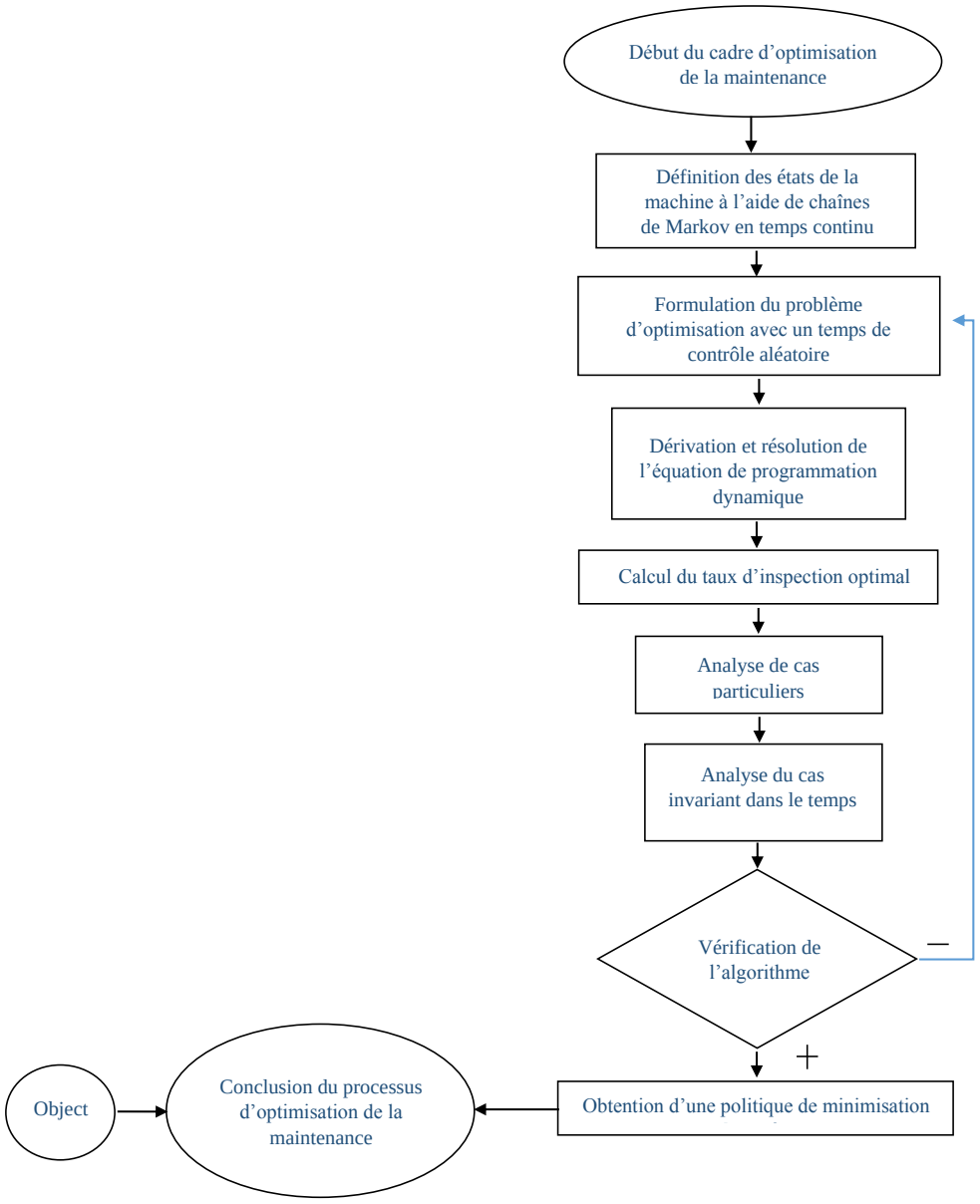

Figure 3.2 Schéma pour la politique optimale de maintenance et d'inspection

CHAPITRE 4 APPROCHE

Interconnexion des études de recherche

La connexion et la cohérence de ma recherche doctorale sont mises en évidence dans ce chapitre, montrant comment chaque axe de recherche a contribué à une mise en œuvre détaillée de la commande optimale dans les processus stochastiques en reliant la théorie des files d'attente à la maintenance et à l'ingénierie de la fiabilité à travers des méthodologies communes, faisant ainsi progresser à la fois les fondements théoriques et les techniques appliquées.

La thèse est structurée autour de trois axes de recherche principaux, tous centrés sur le contrôle optimal et la prise de décision dans des processus stochastiques. Chaque article traite d'un domaine d'application distinct, à savoir le contrôle du nombre de serveurs dans les files d'attente, l'évaluation des distributions des temps de service dans les systèmes de files d'attente, et enfin le développement de politiques de maintenance et d'inspection. En utilisant des cadres mathématiques tels que la programmation dynamique, leur adaptabilité et leur efficacité sont mises en évidence dans plusieurs domaines. En particulier, notre recherche fournit une approche méthodologique pour relever les défis en théorie des files d'attente, en planification de la maintenance et en ingénierie de la fiabilité.

La programmation dynamique constitue le principe fondamental des trois articles. Ce cadre mathématique optimise les processus de prise de décision séquentielle dans des environnements incertains. Dans l'ensemble des travaux présentés, la fonction de valeur joue un rôle fondamental dans l'obtention de la politique optimale. Ce cadre commun relie les trois problèmes de recherche distincts, proposant une base robuste pour l'étude des processus dans divers domaines d'application.

Commande optimale d'un système de file d'attente

L'optimisation de l'allocation des serveurs dans des files $M/M/k$ est abordée dans la première partie de la thèse. Déterminer le nombre optimal de serveurs actifs dans le système, à tout instant donné, est l'objectif central afin de minimiser la longueur de la file dans le temps le plus court possible tout en tenant compte du coût total de contrôle. Le compromis dynamique entre le coût et la performance du système est exploré en détail dans cette partie de la thèse.

Pertinence par rapport aux autres parties de la thèse :

— La gestion des serveurs dans les files est parallèle à l'optimisation des distributions

des temps de service dans la partie consacrée à la distribution des temps de service et à la détermination des taux d'inspection dans la partie relative à la maintenance et à l'inspection

— L'utilisation du cadre de programmation dynamique pour dériver la fonction de valeur constitue un lien méthodologique entre es différents axes de recherche.

Distribution optimale des temps de service pour une file d'attente $M/G/1$

Dans la partie consacrée à cette thématique, le champ d'optimisation du système de files d'attente est élargi en se concentrant sur un système $M/G/1$, dans lequel les serveurs peuvent choisir parmi plusieurs distributions de temps de service. Minimiser la fonction de coût est ici notre objectif, de manière à équilibrer le taux de service, le temps d'attente dans la file et les coûts totaux sur la durée. Dans cette recherche, la programmation dynamique et la probabilité conditionnelle sont intégrées pour choisir la distribution de temps de service optimale parmi les options possibles.

Pertinence par rapport aux autres parties de la thèse :

— En cohérence avec la première partie de la thèse, cette étude se concentre sur la prise de décision optimale dans les files d'attente tout en ajoutant une seconde couche de complexité par la sélection de la meilleure distribution du temps de service.

— L'objectif de minimisation des coûts, tout en atteignant simultanément une performance optimale, reflète celui de la partie relative à la maintenance et à l'inspection, où l'obtention de taux d'inspection optimaux vise à équilibrer les coûts totaux de maintenance et la fiabilité du système.

Politique optimale d'inspection et de maintenance

Le domaine dans cette partie de la thèse passe des systèmes de files d'attente à celui de la maintenance et de la fiabilité. Une machine représentée par une chaîne de Markov en temps continu est analysée, avec l'objectif d'optimiser les taux d'inspection et les coûts de maintenance afin de minimiser les temps d'arrêt opérationnels. Ce travail progresse en intégrant le cadre de programmation dynamique aux processus de Markov en temps continu afin de fournir des solutions explicites pour différents cas particuliers.

Pertinence par rapport aux autres parties de la thèse :

— Cette étude est en cohérence avec les parties précédentes de la thèse en se concentrant sur la minimisation des coûts, l'incorporation d'un cadre de programmation dynamique pour dériver des solutions optimales, et enfin l'adoption de stratégies de

contrôle dépendant de l'état. Cette stratégie dépend de l'état actuel et s'adapte dynamiquement à l'évolution du système, plutôt que d'adopter des stratégies fixes et non personnalisées.

— Le modèle de chaîne de Markov en temps continu utilisé dans cette étude est cohérent avec la modélisation par processus stochastiques dans les parties précédentes de la thèse, démontrant la polyvalence et la cohérence de ces techniques mathématiques dans divers domaines d'application.

La Figure 4.1 synthétise la cohérence méthodologique entre les trois études.

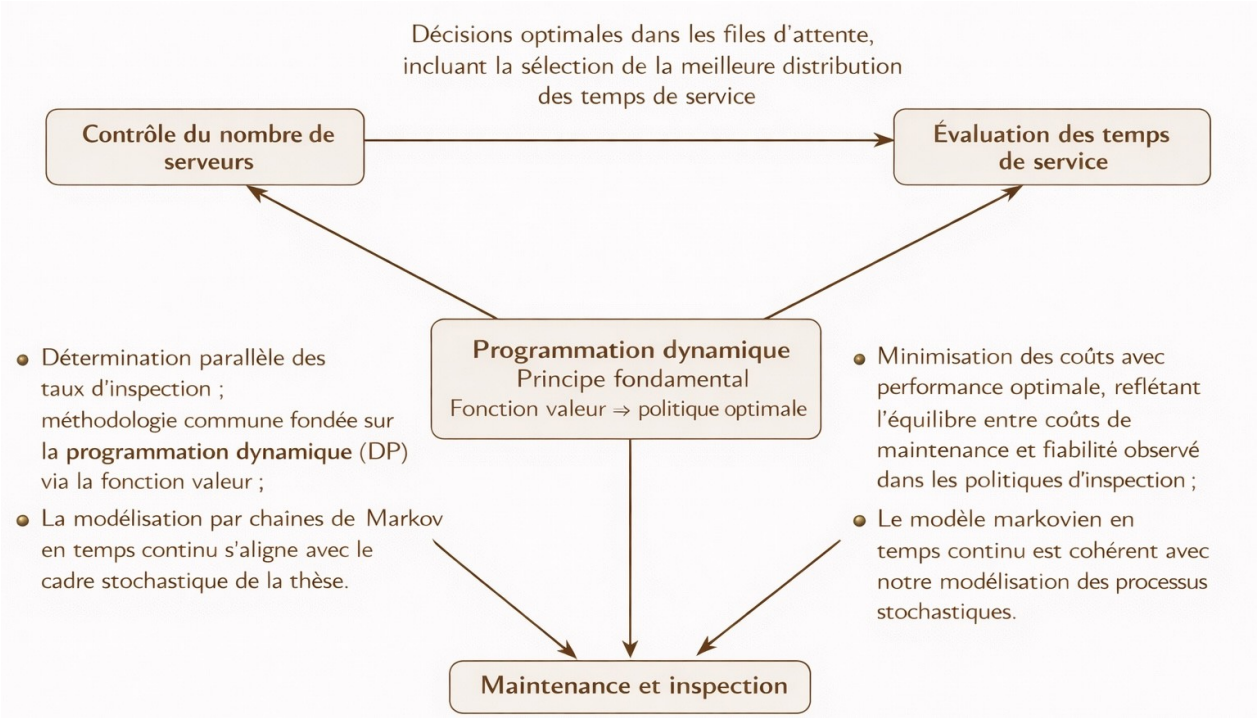

Figure 4.1 Organisation et interconnexions de la thèse.

CHAPITRE 5 CONTRÔLE OPTIMAL D'UN SYSTÈME DE FILES D'ATTENTE

Ce chapitre s'inscrit dans des travaux de recherche en contrôle optimal stochastique appliqué aux systèmes de files d'attente [91], dont les résultats sont ici repris et étendus dans un cadre de modélisation unifié.

5.1 Introduction

Dans cette partie, la présente section étudie une version modifiée du modèle de file d'attente $M/M/k$ dans laquelle le nombre de serveurs actifs peut être modifié dynamiquement au cours de la horizon de contrôle. À un moment donné, supposons qu'il y ait $k + l$ clients présents dans le système de files d'attente, tous en attente de service. L'objectif de cette étude est d'identifier le nombre optimal de serveurs pouvant être activés afin de réduire le nombre de clients à $k + r$, où $0 \leq r < l$, aussi rapidement que possible, tout en tenant compte des coûts de contrôle associés induits dans le système par ces décisions. L'équation de programmation dynamique satisfaite par la fonction de valeur est dérivée dans le cas général, et enfin des cas particuliers sont résolus explicitement pour démontrer l'applicabilité de la méthodologie proposée.

Les arrivées des clients dans le modèle classique $M/M/k$ se font selon un processus de Poisson de taux λ, et l'un des k serveurs disponibles sert le client nouvellement arrivé. Les temps de service suivent une loi exponentielle de taux μ, sont indépendants les uns des autres et sont indépendants des temps entre les arrivées. La capacité du système, qui peut être infinie, est notée c_p, et le système fonctionne selon une discipline de service FIFO (premier arrivé, premier servi).

Plusieurs études dans la littérature ont traité de la commande optimale des systèmes de files d'attente. Typiquement, on suppose que les paramètres liés au service peuvent être modifiés par le décideur, comme le nombre de serveurs actifs et/ou le taux de service. De plus, certaines études considèrent la possibilité de contrôler les paramètres d'arrivée. En outre, les modèles peuvent intégrer plusieurs classes de clients, en essayant d'obtenir la stratégie de service optimale pour chaque type de client. La commande optimale peut être effectuée sur un horizon de temps fini ou infini, avec des objectifs selon des critères de coût moyen ou de coût actualisé.

Des travaux antérieurs sur le contrôle des files d'attente incluent l'étude de Wen *et al.* [19],

qui se concentre sur l'insertion optimale de clients à haute priorité en intégrant des objectifs de temps d'attente associés à leurs niveaux de priorité. L'objectif principal est de minimiser le temps d'attente des clients réguliers, tout en maintenant la qualité de service pour les clients à haute priorité. Un problème de commande optimale est formulé par Bhati *et al.* [20], qui implique des serveurs hétérogènes, avec pour objectif de minimiser l'âge moyen de l'information. Asadzadeh *et al.* [13] se concentrent sur la performance opérationnelle d'une station-service, afin d'optimiser le nombre de pompes à essence et les opérateurs en fonction des ventes de carburant et des coûts associés. Une politique de gestion des files d'attente est développée par Luo *et al.* [15] en tenant compte à la fois du nombre de clients et de la charge de travail des serveurs ; la distribution de la longueur de file, la longueur de file espérée et le coût moyen à long terme par unité de temps sont examinés dans leur travail, et leur approche est validée par des comparaisons numériques de politiques optimales. Xinzhe et Modiano [32] abordent la commande optimale dans les réseaux de files d'attente stochastiques, en supposant que les tâches sont associées à des files en parallèle, et proposent une stratégie de contrôle qui maximise l'utilité cumulée sur un horizon temporel fini. Enfin, un cadre de commande optimale en temps continu est introduit par Petrović *et al.* [30] pour prédire le nombre optimal d'unités de péage actives dans les opérations de péage autoroutier, utilisant des réseaux de neurones combinés avec la théorie des files d'attente.

Dans notre travail, il est supposé que le nombre de serveurs actifs peut être contrôlé à tout moment. Soit $X(t)$ le nombre de clients dans le système à l'instant t, et soit $n(t) \in \{1, \ldots, k\}$ le nombre de serveurs en activité à ce moment. Le système est supposé démarrer à l'instant t_0, avec $k + l$ clients initialement présents, où $l \geq 1$.

Remarque. La constante t_0 peut représenter l'heure d'ouverture d'un magasin spécifique. En conséquence, il est supposé que $k + l$ clients attendent déjà que le magasin commence à fonctionner.

L'objectif est de déterminer le nombre optimal de serveurs à engager afin de réduire le nombre initial de clients $X(t_0) = k + l$ à $k + r$, où $0 \leq r < l$, le plus rapidement possible, tout en minimisant l'espérance de la fonction de coût (interprétée comme une *récompense* lorsqu'elle est négative)

$$J(t_0, t_1, l, r) := \int_{t_0}^{T(t_0,t_1,l,r)} \{q_0 n(t) - m[n(t)]\} \, \mathrm{d}t, \tag{5.1}$$

où

$$T(t_0, t_1, l, r) := \inf\{t > t_0 : X(t) = k + r \text{ ou } t = t_1 (> t_0) \mid X(t_0) = k + l\}. \tag{5.2}$$

$m[n(t)]$ est la somme gagnée par unité de temps par le système lorsque $n(t)$ serveurs sont en service, et q_0 est une constante positive. La variable aléatoire T est appelée *instant de premier*

passage en théorie des probabilités. Ce type de problème, dans lequel l'optimiseur contrôle un processus stochastique jusqu'à ce qu'un certain événement se produise, est désigné comme *problème de « homing »*. Whittle [92] a étudié les problèmes de homing pour les processus de diffusion n-dimensionnels ; on pourra se référer au travail de Lefebvre et Pazhoheshfar [86] pour une extension de ces problèmes au contexte de la maintenance optimale.

Les problèmes de contrôle stochastique tels que les problèmes de homing sont étroitement liés à la théorie du *contrôle robuste* telle que discutée par Zhou et Doyle [93]. Dans une contribution récente à ce domaine, Uğurlu [94] propose une mesure de risque multivariée quantifiant les risques conjoints provenant de plusieurs sources, tout en conservant les propriétés d'une mesure de risque cohérente. Cette approche permet une plus grande flexibilité pour la gestion des risques dans les modèles de mélanges gaussiens et les structures dépendantes, généralisant la moyenne à risque univariée à un cadre multivarié, notamment dans les contextes financiers et actuariels. De plus, l'application de mesures de probabilité déformées pour gérer le risque de portefeuille en situation d'incertitude de marché est examinée par Uğurlu et Brzeczek [95]. Dans cette étude, une version dynamique de ces opérateurs est développée pour soutenir les politiques de commande optimale et la programmation dynamique, et la méthode est mise en œuvre sur un exemple numérique d'optimisation de portefeuille.

Soit $F(t_0, t_1, l, r)$ la fonction de valeur :

$$F(t_0, t_1, l, r) := \inf_{\substack{n(t) \\ t \in [t_0, T)}} E[J(t_0, t_1, l, r)]. \tag{5.3}$$

La condition aux limites est $F(t_0, t_1, l, r) = 0$ si $l \leq r$, et on suppose que

$$F(t_0, t_1, l, r) = F_1 > 0 \quad \text{si } T(t_0, t_1, l, r) = t_1. \tag{5.4}$$

L'équation de programmation dynamique que satisfait la fonction de valeur est dérivée dans la section suivante.

5.2 Programmation dynamique

Soit $n(t_0) = \eta \in \{1, 2, \ldots, k\}$. On note par A et D le temps nécessaire pour qu'un nouveau client arrive et pour qu'un client quitte le système, respectivement, après le temps t_0. On a $A \sim \text{Exp}(\lambda)$ et $D \sim \text{Exp}(\eta\mu)$. Il en résulte que

$$P[A \leq \Delta t] = 1 - e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t) \tag{5.5}$$

Cette équation exprime la probabilité qu'une arrivée de client se produise pendant l'intervalle de temps infinitésimal $(t_0, t_0 + \Delta t]$, sur la base de l'hypothèse que les temps entre les arrivées suivent une distribution exponentielle de taux λ. La fonction de répartition (CDF) de la loi exponentielle est donnée par :

$$P[A \leq \Delta t] = 1 - e^{-\lambda \Delta t}. \tag{5.6}$$

En appliquant le développement de Taylor d'ordre un de la fonction exponentielle autour de $\Delta t = 0$, on obtient :

$$e^{-\lambda \Delta t} = 1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{2!} - \cdots . \tag{5.7}$$

En substituant l'équation (5.7) dans l'équation (5.6), on obtient :

$$1 - e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t). \tag{5.8}$$

Par conséquent, pour un Δt petit, la probabilité d'une seule arrivée est approximativement proportionnelle à $\lambda \Delta t$, les termes d'ordre supérieur étant négligeables.

De plus, on a :

$$P[D \leq \Delta t] = \eta \mu \Delta t + o(\Delta t). \tag{5.9}$$

L'équation ci-dessus donne la probabilité qu'un service se termine (départ) pendant l'intervalle $(t_0, t_0 + \Delta t]$, en supposant que η serveurs sont en activité et que chacun opère de manière indépendante avec un taux exponentiel μ. Le temps jusqu'à la première fin de service suit une loi exponentielle de taux $\eta\mu$. Ainsi, la fonction de répartition est :

$$P[D \leq \Delta t] = 1 - e^{-\eta \mu \Delta t}. \tag{5.10}$$

En utilisant le même développement de Taylor que dans l'équation (5.7), en remplaçant λ par $\eta\mu$, on obtient :

$$e^{-\eta \mu \Delta t} = 1 - \eta \mu \Delta t + \frac{(\eta \mu \Delta t)^2}{2!} - \cdots . \tag{5.11}$$

En substituant l'équation (5.11) dans l'équation (5.10), on obtient :

$$1 - e^{-\eta \mu \Delta t} = \eta \mu \Delta t + o(\Delta t). \tag{5.12}$$

Ainsi, pour un Δt petit, la probabilité qu'un client quitte le système de files d'attente est approximativement $\eta\mu\Delta t$, où $o(\Delta t)$ représente les termes d'ordre supérieur négligeables.

De plus, par indépendance,

$$P[A \leq \Delta t, D \leq \Delta t] = o(\Delta t). \tag{5.13}$$

Cette équation indique que la probabilité qu'une arrivée et une fin de service aient lieu dans le même petit intervalle $(t_0, t_0 + \Delta t]$ est négligeable. En supposant l'indépendance entre le temps d'arrivée A et le temps de départ D, on peut écrire :

$$P[A \leq \Delta t,\ D \leq \Delta t] = P[A \leq \Delta t] \cdot P[D \leq \Delta t]. \tag{5.14}$$

En utilisant les approximations d'ordre un dérivées précédemment dans les équations (5.8) et (5.12), on remplace :

$$P[A \leq \Delta t] = \lambda \Delta t + o(\Delta t), \qquad P[D \leq \Delta t] = \eta\mu \Delta t + o(\Delta t). \tag{5.15}$$

En multipliant les deux expressions de l'équation (5.15), on obtient :

$$P[A \leq \Delta t,\ D \leq \Delta t] = (\lambda \Delta t + o(\Delta t))(\eta\mu \Delta t + o(\Delta t)) = \lambda\eta\mu \Delta t^2 + o(\Delta t). \tag{5.16}$$

Puisque $\Delta t^2 = o(\Delta t)$ lorsque $\Delta t \to 0$, on conclut que :

$$P[A \leq \Delta t,\ D \leq \Delta t] = o(\Delta t). \tag{5.17}$$

Ce résultat implique que la probabilité que les deux événements (arrivée et départ) aient lieu simultanément pendant l'intervalle $(t_0, t_0 + \Delta t]$ est négligeable (égale à $o(\Delta t)$) par rapport aux probabilités des événements individuels. Ainsi, le nombre $X(t_0 + \Delta t) := k + L$ de clients dans le système à l'instant $t_0 + \Delta t$ sera égal à

$$\begin{cases} k + l + 1 & \text{avec une probabilité de } \lambda \Delta t + o(\Delta t), \\ k + l - 1 & \text{avec une probabilité de } \eta\mu \Delta t + o(\Delta t), \\ k + l & \text{avec une probabilité de } 1 - \lambda \Delta t - \eta\mu \Delta t + o(\Delta t). \end{cases} \tag{5.18}$$

Explication mathématique : Toutes les transitions possibles dans le nombre de clients dans le système de files d'attente pendant un petit intervalle de temps $(t_0, t_0 + \Delta t]$ sont décrites dans cette équation. Soit $X(t)$ le nombre de clients à l'instant t, et supposons que $X(t_0) = k + l$. Selon les résultats précédents, seuls trois événements mutuellement exclusifs peuvent se produire dans cet intervalle :

— Une arrivée (sans départ) : augmente la taille de la file à $k+l+1$, avec une probabilité $\lambda\Delta t + o(\Delta t)$.
— Un départ (sans arrivée) : diminue la taille de la file à $k+l-1$, avec une probabilité $\eta\mu\Delta t + o(\Delta t)$.
— Aucune arrivée ni départ : la taille de la file reste $k+l$, avec une probabilité $1-\lambda\Delta t - \eta\mu\Delta t + o(\Delta t)$.

Cette formulation par cas est nécessaire pour formuler la fonction de valeur espérée dans l'équation de programmation dynamique. Elle garantit qu'un seul événement (arrivée ou départ) affecte l'état du système pendant un intervalle de temps Δt suffisamment petit.

En appliquant le principe d'optimalité de Bellman, on peut écrire (en omettant la dépendance de F en t_1 et r pour alléger la notation) :

$$\begin{aligned}
F(t_0,l) \;=\; & \inf_{\substack{n(t)\\ t\in[t_0,T)}} \left\{ E\left[\int_{t_0}^{t_0+\Delta t} \{q_0\, n(t) - m[n(t)]\}\, \mathrm{d}t \right.\right. \\
& \qquad \left.\left. +\; \int_{t_0+\Delta t}^{T} \{q_0\, n(t) - m[n(t)]\}\, \mathrm{d}t \right] \right\} \\
=\; & \inf_{\substack{n(t)\\ t\in[t_0,t_0+\Delta t]}} \{[q_0\,\eta - m(\eta)]\,\Delta t + E[F(t_0+\Delta t, L)]\} + o(\Delta t) \\
=\; & \inf_{\substack{n(t)\\ t\in[t_0,t_0+\Delta t]}} \{[q_0\,\eta - m(\eta)]\,\Delta t + F(t_0+\Delta t, l+1)\,\lambda\Delta t \\
& + F(t_0+\Delta t, l-1)\,\eta\mu\Delta t + F(t_0+\Delta t, l)\,(1-\lambda\Delta t - \eta\mu\Delta t)\} \\
& + o(\Delta t). \qquad (5.19)
\end{aligned}$$

Explication mathématique : Le principe d'optimalité de Bellman stipule que la fonction de valeur optimale à l'instant t_0 peut être décomposée en la somme du coût espéré immédiat sur l'intervalle $[t_0, t_0+\Delta t]$ et de la valeur espérée du processus à partir de $t_0+\Delta t$. Soit $F(t_0,l)$ le coût espéré optimal quand il y a l clients dans la file à l'instant t_0. Alors :

$$
\begin{aligned}
F(t_0, l) &= \inf_{\substack{n(t) \\ t \in [t_0, T)}} \left\{ E\left[\int_{t_0}^{t_0+\Delta t} (q_0 n(t) - m[n(t)])\, \mathrm{d}t \right.\right. \\
&\qquad \left.\left. + \int_{t_0+\Delta t}^{T} (q_0 n(t) - m[n(t)])\, \mathrm{d}t \right] \right\} \\
&= \inf_{\substack{n(t) \\ t \in [t_0, t_0+\Delta t]}} \{(q_0\eta - m(\eta))\,\Delta t + E[F(t_0+\Delta t, L)]\} + o(\Delta t). \qquad (5.20)
\end{aligned}
$$

Dans l'équation (5.20), le contrôle est constant sur $[t_0, t_0 + \Delta t]$ à la valeur $\eta = n(t)$, et L est le nombre aléatoire de clients dans le système à $t_0 + \Delta t$. Le premier terme correspond au coût sur l'intervalle court, et le second à la valeur future espérée. En utilisant les probabilités de l'équation (5.18), la valeur future espérée est :

$$
\begin{aligned}
E[F(t_0+\Delta t, L)] &= F(t_0+\Delta t, l+1)\cdot \lambda\Delta t + F(t_0+\Delta t, l-1)\cdot \eta\mu\Delta t \\
&\quad + F(t_0+\Delta t, l)\cdot(1 - \lambda\Delta t - \eta\mu\Delta t) + o(\Delta t). \qquad (5.21)
\end{aligned}
$$

En remplaçant l'équation (5.21) dans l'équation (5.20), on obtient l'expression récursive complète de $F(t_0, l)$ sur l'intervalle $[t_0, t_0 + \Delta t]$.

En considérant cette formulation récursive, on soustrait $F(t_0 + \Delta t, l)$ des deux côtés de l'équation pour isoler le changement de la fonction de valeur pendant Δt, soit :

$$
\begin{aligned}
F(t_0, l) - F(t_0+\Delta t, l) &= \inf_{\substack{n(t) \\ t \in [t_0, t_0+\Delta t]}} \{(q_0\eta - m(\eta))\,\Delta t \\
&\quad + F(t_0+\Delta t, l+1)\lambda\Delta t + F(t_0+\Delta t, l-1)\eta\mu\Delta t \\
&\quad - F(t_0+\Delta t, l)(\lambda\Delta t + \eta\mu\Delta t)\} + o(\Delta t). \qquad (5.22)
\end{aligned}
$$

Cette équation représente la différence d'ordre un de la fonction de valeur sur l'intervalle Δt, exprimée en fonction du coût immédiat et des transitions espérées.

Développement de Taylor et passage à la limite : Pour dériver l'équation de programmation dynamique en temps continu, on applique un développement de Taylor d'ordre un aux expressions :

$$
F(t_0+\Delta t, l) = F(t_0, l) + \Delta t \cdot \frac{\partial F}{\partial t_0}(t_0, l) + o(\Delta t), \qquad (5.23)
$$

$$F(t_0 + \Delta t, l + 1) = F(t_0, l + 1) + \Delta t \cdot \frac{\partial F}{\partial t_0}(t_0, l + 1) + o(\Delta t), \tag{5.24}$$

$$F(t_0 + \Delta t, l - 1) = F(t_0, l - 1) + \Delta t \cdot \frac{\partial F}{\partial t_0}(t_0, l - 1) + o(\Delta t). \tag{5.25}$$

En remplaçant les équations (5.23)–(5.25) dans l'équation (5.22) et en divisant les deux côtés par Δt, on obtient :

$$\begin{aligned} \frac{F(t_0, l) - F(t_0 + \Delta t, l)}{\Delta t} &= \inf_{\eta} \{q_0\eta - m(\eta) + \lambda F(t_0, l + 1) + \eta\mu F(t_0, l - 1) \\ &\quad -(\lambda + \eta\mu)F(t_0, l)\} + \frac{o(\Delta t)}{\Delta t}. \end{aligned} \tag{5.26}$$

En divisant à nouveau par Δt et en laissant $\Delta t \to 0$, on obtient la proposition suivante :

Proposition 5.2.1. *La fonction de valeur satisfait l'*équation de programmation dynamique *(EPD)*

$$-\frac{\partial F}{\partial t_0}(t_0, l) = \inf_{\eta} \{q_0\eta - m(\eta) + \lambda F(t_0, l + 1) + \eta\mu F(t_0, l - 1) - (\lambda + \eta\mu)F(t_0, l)\}. \tag{5.27}$$

L'équation (5.27) est l'équation de programmation dynamique qui régit l'évolution optimale de la fonction de coût $F(t_0, l)$ sous le contrôle η. Elle équilibre le coût instantané et les valeurs attendues des transitions futures :

$$\frac{\partial F(t_0, l)}{\partial t_0} = -\left(\textit{coût instantané} + \textit{variation attendue du coût due aux transitions de file}\right). \tag{5.28}$$

De plus, les conditions aux limites sont :

$$F(t_0, l) = \begin{cases} 0 & \text{si } l \leq r, \\ F_1 & \text{si } t_0 = t_1. \end{cases} \tag{5.29}$$

Remarque. Si la capacité du système c_p est finie et qu'un nouveau client arrive alors que le système est plein, on fixe la fonction de valeur F égale à zéro.

Explication mathématique des conditions aux limites : Les conditions aux limites appropriées doivent être spécifiées pour résoudre l'équation de programmation dynamique (5.27). Ces conditions garantissent que le problème est bien posé et dépendent du comportement du système dans les états critiques. La première condition aux limites dans nos problèmes de

contrôle est :

$$F(t_0, l) = 0 \quad \text{si} \quad l \leq r, \tag{5.30}$$

Cela reflète le fait que lorsqu'on a $l \leq r$, il n'y a plus de coût supplémentaire à considérer et le processus se termine ou reste inactif. Typiquement, on suppose que le contrôleur cesse d'agir lorsque le système est vide. Une condition aux limites spécifiée pour l'instant terminal t_1 est :

$$F(t_0, l) = F_1 \quad \text{si } t_0 = t_1. \tag{5.31}$$

où F_1 représente un coût terminal. Ces deux conditions — d'état terminal et d'état absorbant — sont essentielles pour l'intégration en arrière de l'EPD, permettant la résolution (numérique ou analytique) de la fonction de valeur $F(t_0, l)$ pour tout $t_0 < t_1$.

5.3 Exemples numériques

Cas 1. Le premier cas particulier considéré est celui pour lequel $k = 2$ et la capacité du système est $c_p = 3$. De plus, on suppose $l = 1$, $r = 0$, $\lambda = \mu = 1$, $q_0 = 1$ et $m[n(t)] = cn(t)$, où $c > 0$. Ainsi, il existe deux possibilités : soit $n(t_0) = 1$, soit $n(t_0) = 2$.

1 : On précise ci-après une explication détaillée des paramètres de notre modèle afin de faciliter la compréhension de leur influence et de leur importance dans le système.

Le choix de ces valeurs de paramètres permet d'analyser les dynamiques du système et son comportement avec des taux d'arrivée et de service équilibrés, ce qui fournit une intuition de l'efficacité d'utilisation des serveurs. Cette hypothèse permet d'illustrer les dynamiques fondamentales de la file d'attente dans des conditions simples mais représentatives.

Dans le cas 1, un scénario du modèle de files d'attente est examiné, dans lequel :

- $k = 2$: Indique le nombre maximal de serveurs qui peuvent travailler activement à un instant t. Le système limite le fonctionnement à deux serveurs actifs en parallèle.
- $\lambda = \mu = 1$: Le modèle de file $M/M/k$ dans lequel les clients arrivent selon un processus de Poisson de taux λ, et sont servis par l'un des k serveurs disponibles. Chaque serveur a un temps de service exponentiel avec paramètre μ. Les temps de service sont des variables aléatoires indépendantes et indépendantes des temps inter-arrivées. Ici, λ et μ sont tous deux égaux à 1, ce qui signifie qu'en moyenne, un client arrive et un serveur traite un client par unité de temps.
- $c_p = 3$: Ce paramètre signifie que la capacité du système est de c_p (ici finie), et une politique de service FIFO (premier arrivé, premier servi) est utilisée ; indiquant que le système peut contenir jusqu'à trois clients simultanément avant que les nouvelles

arrivées soient bloquées ou rejetées.

— $l = 1$, $r = 0$: Ces paramètres spécifient le nombre initial et cible de clients excédant le nombre de serveurs actifs. En particulier, $l = 1$ signifie qu'il y a initialement un client de plus que le nombre de serveurs en activité, et $r = 0$ signifie que l'objectif est de réduire la file à seulement k clients. L'objectif du modèle est donc de trouver une stratégie optimale minimisant à la fois le temps et le coût nécessaires pour ramener la file à un niveau égal au nombre de serveurs.

— $q_0 = 1$ et $m[n(t)] = cn(t)$, avec $c > 0$: q_0 est une constante représentant le coût par serveur et par unité de temps, ici fixé à 1. La fonction $m[n(t)]$ représente le revenu généré par le système lorsqu'il y a $n(t)$ serveurs actifs.

Désignons la fonction $F(t_0, l)$ par $F_{i,j}(t_0)$ si $n(t_0) = i$ et $n_q(t_0) = j$, où $n_q(t_0)$ est le nombre de clients en attente de service lorsque $n(t_0) = i$. On déduit de la proposition 1 que si l'optimiseur choisit toujours $n(t) = i$ pour tout $t \in [t_0, T)$, alors :

$$-F'_{1,2}(t_0) = 1 - c + F_{1,3}(t_0) + F_{1,1}(t_0) - 2F_{1,2}(t_0), \tag{5.32}$$

$$-F'_{2,1}(t_0) = 2 - 2c + F_{2,2}(t_0) + 2F_{2,0}(t_0) - 3F_{2,1}(t_0). \tag{5.33}$$

Puisque $c_p = 3$, il convient de poser $F_{1,3}(t_0) = 0$ et $F_{2,2}(t_0) = 0$. De plus, la condition aux limites $F(t_0, 0) = 0$ implique que $F_{2,0}(t_0) = 0$. Il en résulte que le système d'équations différentielles ordinaires (EDO) se réduit à :

$$-F'_{1,2}(t_0) = 1 - c + F_{1,1}(t_0) - 2F_{1,2}(t_0), \tag{5.34}$$

$$-F'_{2,1}(t_0) = 2 - 2c - 3F_{2,1}(t_0), \tag{5.35}$$

où la fonction $F_{1,1}(t_0)$ satisfait l'équation :

$$\begin{aligned} -F'_{1,1}(t_0) &= 1 - c + F_{1,2}(t_0) + F_{1,0}(t_0) - 2F_{1,1}(t_0) \\ &= 1 - c + F_{1,2}(t_0) - 2F_{1,1}(t_0). \end{aligned} \tag{5.36}$$

La solution de l'équation (5.35) qui satisfait la condition aux limites $F_{2,1}(t_1) = F_1$ est

$$F_{2,1}(t_0) = \frac{2}{3}(1 - c) + e^{3(t_0 - t_1)} \left(F_1 - \frac{2}{3}(1 - c) \right). \tag{5.37}$$

Dérivation mathématique de la solution :

Réécrivons l'équation (5.35) :

$$F'_{2,1}(t_0) + 3F_{2,1}(t_0) = 2 - 2c. \tag{5.38}$$

C'est une équation différentielle ordinaire (EDO) linéaire du premier ordre non homogène de la forme :

$$F'(t) + p(t)F(t) = q(t), \tag{5.39}$$

avec des coefficients constants $p(t_0) = 3$ et $q(t_0) = 2 - 2c$.

Étape 1 : Facteur intégrant. Le facteur intégrant est

$$\mu(t_0) = e^{\int 3\, dt_0} = e^{3t_0}. \tag{5.40}$$

Multiplions les deux côtés de l'équation (5.38) par $\mu(t_0) = e^{3t_0}$:

$$e^{3t_0}F'_{2,1}(t_0) + 3e^{3t_0}F_{2,1}(t_0) = (2 - 2c)e^{3t_0}. \tag{5.41}$$

Le membre de gauche devient la dérivée d'un produit :

$$\frac{d}{dt_0}\left(e^{3t_0}F_{2,1}(t_0)\right) = (2 - 2c)e^{3t_0}. \tag{5.42}$$

Étape 2 : Intégrons les deux membres.

$$\int \frac{d}{dt_0}\left(e^{3t_0}F_{2,1}(t_0)\right) dt_0 = \int (2 - 2c)e^{3t_0} dt_0, \tag{5.43}$$

$$e^{3t_0}F_{2,1}(t_0) = \frac{2 - 2c}{3}e^{3t_0} + C. \tag{5.44}$$

Divisons les deux côtés par e^{3t_0} :

$$F_{2,1}(t_0) = \frac{2}{3}(1 - c) + Ce^{-3t_0}. \tag{5.45}$$

Étape 3 : Appliquons la condition aux limites $F_{2,1}(t_1) = F_1$.

Substituons dans l'équation (5.45) :

$$F_1 = \frac{2}{3}(1 - c) + Ce^{-3t_1}, \tag{5.46}$$

$$C = \left(F_1 - \frac{2}{3}(1-c)\right) e^{3t_1}. \tag{5.47}$$

Étape 4 : Substituons C dans la solution générale.

$$F_{2,1}(t_0) = \frac{2}{3}(1-c) + \left(F_1 - \frac{2}{3}(1-c)\right) e^{3(t_0-t_1)}. \tag{5.48}$$

C'est la solution explicite de l'EDO (5.35) qui satisfait la condition aux limites à l'instant t_1.

De plus, la solution du système (5.34), (5.36) tel que $F_{1,1}(t_1) = F_{1,2}(t_1) = F_1$ est :

$$F_{1,1}(t_0) = F_{1,2}(t_0) = 1 - c + e^{t_0-t_1}\left(F_1 + c - 1\right). \tag{5.49}$$

Dérivation mathématique de la solution du système (5.36) et (5.34) :

On considère le système :

$$-F'_{1,1}(t_0) = 1 - c + F_{1,2}(t_0) - 2F_{1,1}(t_0), \tag{5.50}$$
$$-F'_{1,2}(t_0) = 1 - c + F_{1,1}(t_0) - 2F_{1,2}(t_0), \tag{5.51}$$

avec les conditions aux limites : $F_{1,1}(t_1) = F_{1,2}(t_1) = F_1$.

Supposons que $F_{1,1}(t_0) = F_{1,2}(t_0) =: D(t_0)$. Alors les deux équations se réduisent à la même EDO linéaire du premier ordre :

$$-D'(t_0) = 1 - c - D(t_0). \tag{5.52}$$

Mise sous forme standard :

$$D'(t_0) + D(t_0) = 1 - c. \tag{5.53}$$

Étape 1 : Calcul du facteur intégrant. Le facteur intégrant est :

$$\mu(t_0) = e^{\int 1\, dt_0} = e^{t_0}. \tag{5.54}$$

Multiplions les deux membres de l'équation (5.53) par e^{t_0} :

$$e^{t_0}D'(t_0) + e^{t_0}D(t_0) = (1-c)e^{t_0}, \tag{5.55}$$

$$\frac{d}{dt_0}\left(e^{t_0}D(t_0)\right) = (1-c)e^{t_0}. \tag{5.56}$$

Étape 2 : Intégrons les deux côtés :

$$e^{t_0} D(t_0) = \int (1-c) e^{t_0}\, dt_0 = (1-c)e^{t_0} + C. \tag{5.57}$$

En isolant $D(t_0)$, on obtient :

$$D(t_0) = 1 - c + Ce^{-t_0}. \tag{5.58}$$

Étape 3 : Appliquons la condition aux limites $D(t_1) = F_1$:

$$F_1 = 1 - c + Ce^{-t_1} \quad \Rightarrow \quad C = (F_1 + c - 1)e^{t_1}. \tag{5.59}$$

Étape 4 : Remplaçons dans la solution générale :

$$D(t_0) = 1 - c + e^{t_0 - t_1}(F_1 + c - 1). \tag{5.60}$$

Ainsi, la solution du système (5.50)–(5.51) est :

$$F_{1,1}(t_0) = F_{1,2}(t_0) = 1 - c + e^{t_0 - t_1}(F_1 + c - 1), \tag{5.61}$$

ce qui satisfait les deux équations et les conditions aux limites.

On définit ensuite :

$$G(t_0) = 1 - c + \min\{F_{1,1}(t_0), F_{2,0}(t_0)\} - 2\min\{F_{1,2}(t_0), F_{2,1}(t_0)\} \tag{5.62}$$

et

$$H(t_0) = 2 - 2c + 2\min\{F_{1,1}(t_0), F_{2,0}(t_0)\} - 3\min\{F_{1,2}(t_0), F_{2,1}(t_0)\}. \tag{5.63}$$

C'est-à-dire que les fonctions $G(t_0)$ et $H(t_0)$ correspondent au membre de droite de l'équation (5.27) si l'optimiseur choisit respectivement $n(t_0) = 1$ et $n(t_0) = 2$. Pour déterminer la commande optimale $n^*(t_0)$, on peut comparer ces deux fonctions.

Supposons que $t_1 = 1$, $c = 2$ et $F_1 = 1$, alors :

$$F_{2,1}(t_0) = -\frac{2}{3} + \frac{5}{3} e^{3(t_0 - 1)} \tag{5.64}$$

et

$$F_{1,1}(t_0) = F_{1,2}(t_0) = -1 + 2e^{t_0 - 1}. \tag{5.65}$$

Puisque $F_{2,0}(t_0) = 0$, on peut écrire :

$$\min\{F_{1,1}(t_0), F_{2,0}(t_0)\} = \min\{-1 + 2e^{t_0-1}, 0\}, \tag{5.66}$$

et

$$\min\{F_{1,2}(t_0), F_{2,1}(t_0)\} = \min\left\{-1 + 2e^{t_0-1}, -\frac{2}{3} + \frac{5}{3}e^{3(t_0-1)}\right\}. \tag{5.67}$$

Les fonctions $G(t_0)$ et $H(t_0)$ sont représentées dans la figure 5.1. On observe que la solution optimale consiste à prendre $n(t_0) = 2$ pour tout $t_0 \in [0, 1)$.

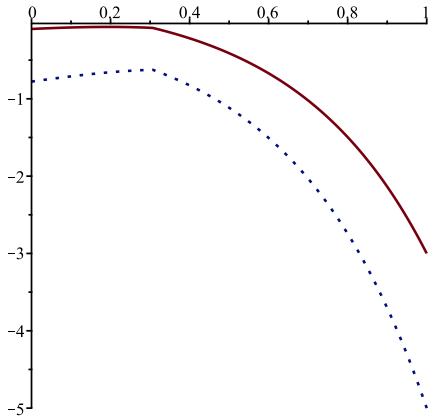

Figure 5.1 Fonctions $G(t_0)$ (ligne pleine) et $H(t_0)$ pour $t_0 \in [0, 1]$ lorsque $t_1 = 1$, $c = 2$ et $F_1 = 1$.

Supposons ensuite que $t_1 = 2$, $c = 1/2$ et $F_1 = 1$. On obtient alors :

$$F_{2,1}(t_0) = \frac{1}{3} + \frac{2}{3}e^{3(t_0-2)} \tag{5.68}$$

et

$$F_{1,1}(t_0) = F_{1,2}(t_0) = \frac{1}{2} + \frac{1}{2}e^{t_0-2}. \tag{5.69}$$

Les fonctions correspondantes $G(t_0)$ et $H(t_0)$ sont illustrées dans la figure 5.2. Les deux courbes se croisent en $\tau \approx 1{,}538$. On peut donc écrire que la commande optimale est donnée par :

$$n^*(t_0) = \begin{cases} 1 & \text{si } 0 \leq t_0 \leq \tau, \\ 2 & \text{si } \tau \leq t_0 < 2. \end{cases} \tag{5.70}$$

De plus, la fonction de valeur $F(t_0, l = 1)$ est représentée dans la figure 5.3. Elle est obtenue

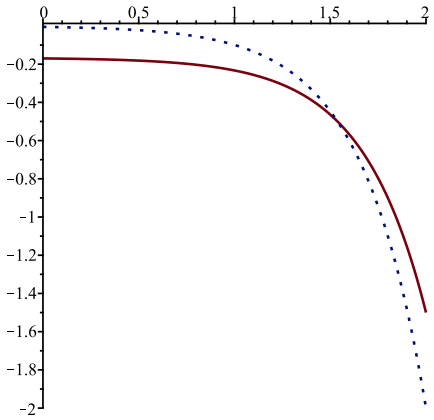

Figure 5.2 Fonctions $G(t_0)$ (ligne pleine) et $H(t_0)$ pour $t_0 \in [0, 2]$ lorsque $t_1 = 2$, $c = 1/2$ et $F_1 = 1$.

en utilisant d'abord l'EDO :

$$-F'(t_0, l = 1) = H(t_0) \quad \text{pour } t_0 \in [\tau, 2) \tag{5.71}$$

avec la condition aux limites $F(2, l = 1) = 1$. Ensuite, on utilise l'EDO :

$$-F'(t_0, l = 1) = G(t_0) \quad \text{pour } t_0 \in [0, \tau] \tag{5.72}$$

et le fait que la fonction est continue en $t_0 = \tau$.

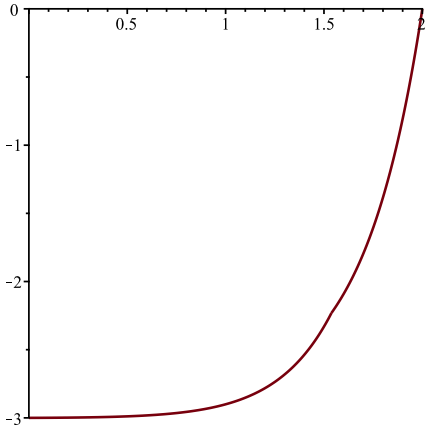

Figure 5.3 Fonction de valeur $F(t_0, l = 1)$ sur l'intervalle $[0, 2]$ lorsque $t_1 = 2$, $c = 1/2$ et $F_1 = 1$.

Remarques. (i) Dans cet exemple, on a :

$$\min\{F_{1,1}(t_0), F_{2,0}(t_0)\} = \min\left\{\frac{1}{2} + \frac{1}{2}e^{t_0-2}, 0\right\} \equiv 0. \tag{5.73}$$

(ii) Ici, il n'est pas possible de poser $c = 1$, sinon $J(t_0, t_1, l, r) \equiv 0$, ce qui rend la valeur de $n(t_0)$ sans importance.

(iii) Supposons que si un serveur vient de terminer de servir un client et qu'un autre client attend, alors il/elle le servira immédiatement. La fonction $H(t_0)$ définie dans l'équation (5.63) devient :

$$H(t_0) = 2 - 2c - 3\min\{F_{1,2}(t_0), F_{2,1}(t_0)\}. \tag{5.74}$$

Considérons $t_1 = 10$, $c = 2$ et $F_1 = 1$. On calcule :

$$F_{2,1}(t_0) = -\frac{2}{3} + \frac{5}{3}e^{3(t_0-10)} \tag{5.75}$$

et

$$F_{1,1}(t_0) = F_{1,2}(t_0) = -1 + 2e^{t_0-10}. \tag{5.76}$$

Les fonctions $G(t_0)$ et $H(t_0)$ sont représentées dans les figures 5.4 et 5.5 pour $t_0 \in [0, 10]$ et $t_0 \in [8, 10]$, respectivement. Il apparaît que la solution optimale consiste à prendre $n(t_0) = 1$ si $t_0 \in [0, \sigma]$, où $\sigma \approx 8{,}825$, et $n(t_0) = 2$ si $t_0 \in [\sigma, 10)$.

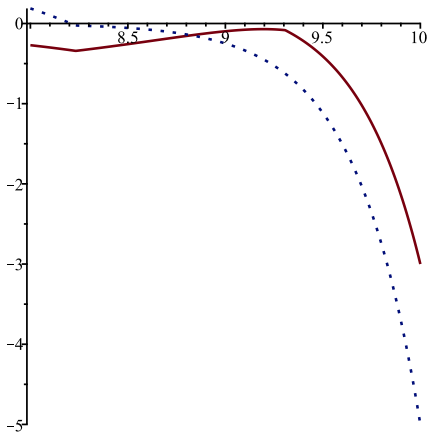

Figure 5.4 Fonctions $G(t_0)$ (ligne pleine) et $H(t_0)$ pour $t_0 \in [0, 10]$ lorsque $t_1 = 10$, $c = 2$ et $F_1 = 1$.

La fonction de valeur $F(t_0, l = 1)$ est présentée dans la figure 5.6. Comme précédemment,

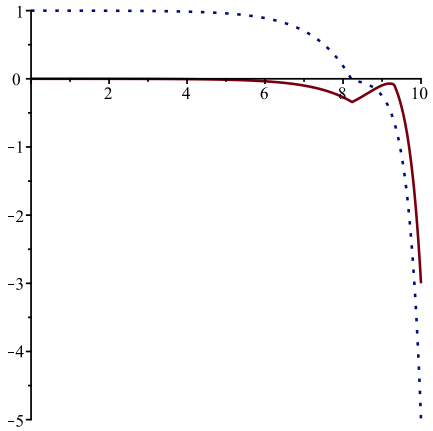

Figure 5.5 Fonctions $G(t_0)$ (ligne pleine) et $H(t_0)$ pour $t_0 \in [8, 10]$ lorsque $t_1 = 10$, $c = 2$ et $F_1 = 1$.

elle est obtenue en utilisant d'abord l'EDO :

$$-F'(t_0, l = 1) = H(t_0) \quad \text{pour } t_0 \in [\sigma, 10] \tag{5.77}$$

et la condition aux limites $F(10, l = 1) = 1$. Ensuite, on utilise l'EDO :

$$-F'(t_0, l = 1) = G(t_0) \quad \text{pour } t_0 \in [0, \sigma] \tag{5.78}$$

ainsi que la continuité de la fonction $F(t_0, l = 1)$ en $t_0 = \sigma$. On doit également diviser l'intervalle $[0, \sigma]$ en deux parties : $[0, s]$ et $[s, \sigma]$, où s est la valeur de t_0 pour laquelle $F_{1,2}(t_0) = F_{2,1}(t_0)$, à savoir :

$$s = 10 + \ln\left(\frac{3\sqrt{5}}{10} - \frac{1}{2}\right) \approx 8{,}233.$$

Cas 2. Supposons maintenant que le nombre de serveurs soit $k = 3$, la capacité du système $c_p = 5$ et $l = 2$. Comme dans le Cas 1, on prend $r = 0$, $\lambda = \mu = q_0 = 1$ et $m[n(t)] = cn(t)$, où $c > 0$. Il y a trois possibilités : $n(t_0) = 1, 2$ ou 3. Les EDO correspondantes sont :

$$-F'_{1,4}(t_0) = 1 - c + F_{1,3}(t_0) - 2F_{1,4}(t_0), \tag{5.79}$$

$$-F'_{2,3}(t_0) = 2 - 2c + 2F_{2,2}(t_0) - 3F_{2,3}(t_0), \tag{5.80}$$

$$-F'_{3,2}(t_0) = 3 - 3c + 3F_{3,1}(t_0) - 4F_{3,2}(t_0). \tag{5.81}$$

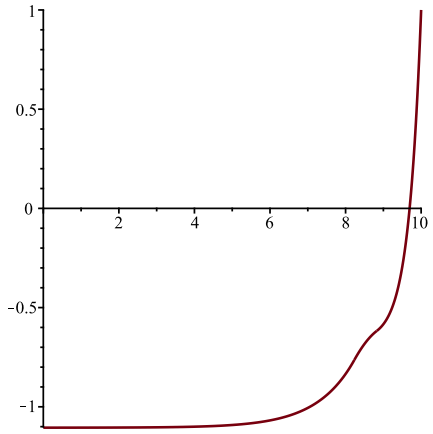

Figure 5.6 Fonction de valeur $F(t_0, l = 1)$ sur l'intervalle $[0, 10]$ lorsque $t_1 = 10$, $c = 2$ et $F_1 = 1$.

De plus, on a besoin des EDO suivantes :

$$
\begin{aligned}
-F'_{1,1}(t_0) &= 1 - c + F_{1,2}(t_0) - 2F_{1,1}(t_0), && (5.82)\\
-F'_{1,2}(t_0) &= 1 - c + F_{1,3}(t_0) + F_{1,1}(t_0) - 2F_{1,2}(t_0), && (5.83)\\
-F'_{1,3}(t_0) &= 1 - c + F_{1,4}(t_0) + F_{1,2}(t_0) - 2F_{1,3}(t_0), && (5.84)\\
-F'_{2,1}(t_0) &= 2 - 2c + F_{2,2}(t_0) - 3F_{2,1}(t_0), && (5.85)\\
-F'_{2,2}(t_0) &= 2 - 2c + F_{2,3}(t_0) + 2F_{2,1}(t_0) - 3F_{2,2}(t_0), && (5.86)\\
-F'_{3,1}(t_0) &= 3 - 3c + F_{3,2}(t_0) - 4F_{3,1}(t_0). && (5.87)
\end{aligned}
$$

On résout les équations ci-dessus car :

- Elles fournissent les fonctions de valeur $F_{\eta,l}(t_0)$ pour toutes les combinaisons de contrôles sur les serveurs $\eta \in \{1, 2\}$ et les longueurs de file $l \in \{0, 1, 2\}$.
- Ces équations sont essentielles pour construire la politique optimale en comparant le coût attendu sous chaque contrôle.
- Les décisions de changement de politique peuvent être identifiées à l'aide de ces équations.

Définissons ensuite la fonction $G_i(t_0)$ qui correspond au membre de droite de l'équation (5.27) si l'optimiseur choisit $n(t_0) = i$, pour $i = 1, 2, 3$:

$$
\begin{aligned}
G_i(t_0) &= i - ic + i \min\{F_{1,3}(t_0), F_{2,2}(t_0), F_{3,1}(t_0)\}\\
&\quad - (1 + i) \min\{F_{1,4}(t_0), F_{2,3}(t_0), F_{3,2}(t_0)\}. && (5.88)
\end{aligned}
$$

Les fonctions $G_1(t_0)$, $G_2(t_0)$ et $G_3(t_0)$ sont représentées dans la figure 5.7 dans le cas où $t_1 = 2$, $c = 1/2$ et $F_1 = 10$.

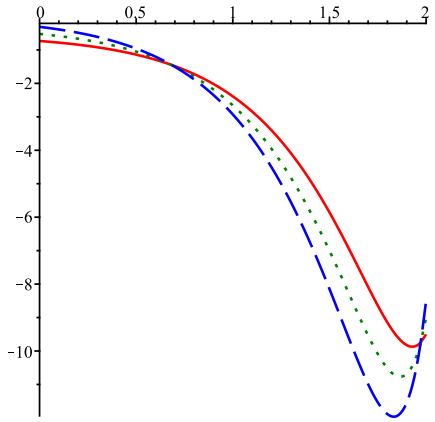

Figure 5.7 Fonctions $G_1(t_0)$ (ligne pleine), $G_2(t_0)$ (pointillée) et $G_3(t_0)$ (pointillée longue) lorsque $t_0 \in [0, 2]$.

Les trois courbes se croisent en deux points : $s_1 \approx 0{,}683$ et $s_2 \approx 1{,}972$. On en déduit à partir de la figure 5.7 que la commande optimale est donné par

$$n^*(t_0) = \begin{cases} 1 & \text{si } 0 \leq t_0 \leq s_1, \\ 3 & \text{si } s_1 \leq t_0 \leq s_2, \\ 1 & \text{si } s_2 \leq t_0 < 2. \end{cases} \tag{5.89}$$

Enfin, la fonction de valeur est présentée dans la figure 5.8.

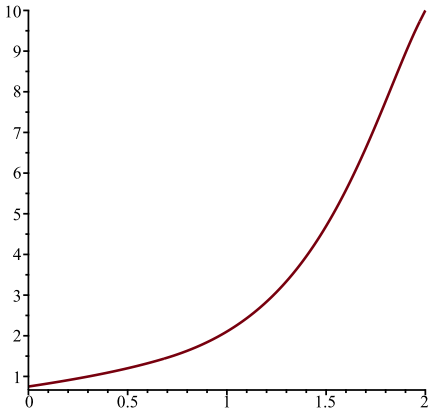

Figure 5.8 Fonction de valeur $F(t_0, l = 2)$ sur l'intervalle $[0, 2]$.

5.4 Conclusion

Le problème de la réduction rapide de la longueur de file est examiné dans le premier travail présenté dans ce chapitre, dans un système $M/M/k$ modifié, où le contrôleur interrompt le contrôle de la file une fois que l'objectif est atteint ou qu'un instant terminal prédéfini est atteint.

L'équation de programmation dynamique (DPE) prend la forme d'une équation différentielle-différence, et après avoir établi la DPE satisfaite par la fonction de valeur, on a résolu deux cas spécifiques de manière explicite pour démontrer l'applicabilité de notre méthodologie. La première étape pour résoudre un problème consiste à résoudre un système d'équations différentielles ordinaires (EDO) du premier ordre en supposant qu'un nombre fixe de serveurs est utilisé sur l'intervalle $[0, T)$. Ensuite, les fonctions représentant le membre de droite de la DPE sont définies pour chaque décision initiale possible prise par l'optimiseur. En comparant ces fonctions, le décideur peut identifier la politique de commande optimale à l'instant initial t_0, et par conséquent à tout instant futur.

L'élément clé est de souligner que l'approche par programmation dynamique dépend généralement de l'hypothèse selon laquelle les temps d'arrivée et de service suivent des lois exponentielles. Si l'équation (5.5) ou l'équation (5.9) n'est pas satisfaite, des complications peuvent survenir lors de la prise de la limite lorsque $\Delta t \to 0$ dans la dérivation de l'équation de programmation dynamique.

CHAPITRE 6 DISTRIBUTION OPTIMALE DES TEMPS DE SERVICE DANS UNE FILE D'ATTENTE $M/G/1$

Ce chapitre s'appuie sur les résultats développés en recherche sur des temps de service dans les systèmes de files d'attente [96], qui sont ici reformulés et étendus dans une formulation unifiée.

6.1 Introduction

Dans un système de files d'attente $M/G/1$, il est supposé que deux distributions de temps de service distinctes peuvent être sélectionnées par le prestataire de service. L'objectif de ce travail est de minimiser la valeur espérée d'une fonction de coût en déterminant le type de service qui prend en compte à la fois le coût opérationnel d'un service plus rapide et le temps requis pour vider la file d'attente. Dans cette étude, le temps terminal aléatoire est défini comme le premier instant où il ne reste plus de clients dans le système de files d'attente. La méthode de programmation dynamique peut être intégrée pour dériver la politique optimale lorsque les temps de service suivent des distributions exponentielles, mais des techniques de probabilité conditionnelle seront employées pour notre analyse dans le cas d'autres distributions de temps de service. Enfin, des solutions explicites seront fournies pour des cas particuliers où le système de files d'attente a une capacité finie.

La théorie des files d'attente fournit une modélisation mathématique des arrivées de clients et des temps de service associés, et joue un rôle essentiel dans la prise de décision en fournissant des informations sur le fonctionnement des systèmes de files d'attente. L'optimisation des décisions analyse les paramètres et contraintes du modèle pour soutenir les décideurs dans l'optimisation des indicateurs de performance, ce qui constitue un élément fondamental de la théorie des files d'attente. Plusieurs structures de files d'attente, telles que la vitesse de service variable ou le nombre de serveurs, peuvent être évaluées pour déterminer une stratégie opérationnelle efficace et rentable.

L'analyse de sensibilité est une autre contribution de la théorie des files d'attente à la prise de décision, qui permet, par la simulation de différents paramètres du système, aux décideurs d'examiner un éventail de scénarios. Cette approche permet d'évaluer les effets des changements dans les paramètres du système, tels que les coûts associés à des modifications dans les stratégies de service client. De plus, elle permet de mieux comprendre le comportement du système et favorise des décisions plus informées et proactives ; voir [33].

La théorie des files d'attente est appliquée dans un large éventail de domaines, tels que la transmission de données mobiles dans les réseaux sans fil, comme l'ont examiné Li *et al.* [28], la modélisation du débit fluvial en période d'inondation dans une étude de Lefebvre [18], et la planification des rendez-vous dans les cliniques, comme étudié par Aziati et Hamdan [97], entre autres.

Un défi central dans la gestion des files d'attente est la réduction du nombre de clients dans le système, une préoccupation fréquente dans de nombreux systèmes de service. Ce défi a été abordé dans diverses études, telles que la perte de productivité causée par les retards de circulation (voir De et Rajbongshi [98]) et les effets néfastes sur la santé (Requia *et al.* [99]), qui soulignent l'effet inefficace des systèmes de files d'attente.

Augmenter le nombre de serveurs dans le système est une décision simple pour réduire la congestion des clients dans les files d'attente. Cependant, chaque serveur supplémentaire engendre des coûts liés au service. Ce compromis a été étudié dans la littérature par plusieurs chercheurs, afin de déterminer le nombre optimal de prestataires de service en incluant la fonction de coût dans l'objectif, comme discuté par Lefebvre et Yaghoubi [91]. Ainsi, un compromis doit être trouvé entre la réduction du temps d'attente des clients et le coût global de la gestion de l'accélération du service, que ce soit par une augmentation de la vitesse de service ou du nombre de serveurs. Déterminer le taux de service optimal représente donc un défi pratique pour les décideurs.

Les ajustements du taux de service ont été largement étudiés pour réduire les files d'attente. Avec le développement de l'intelligence artificielle, de nombreuses techniques de calcul ont émergé dans le but d'optimiser les performances. Par exemple, afin d'identifier la vitesse de service optimale qui minimise le coût total, un modèle de file d'attente en temps discret est examiné par Laxmi et Jyothsna [42] à l'aide de l'algorithme d'optimisation par essaims particulaires dans le contexte des centres de contact par courriel.

Les orientations récentes de la théorie des files d'attente incluent de nouveaux modèles de programmation mathématique, comme dans une étude de Tian *et al.* [53] qui analyse un système de file d'attente à serveur unique avec des taux de service variables pour observer l'impact du comportement des clients sur la performance à l'équilibre du système, en utilisant des chaînes de Markov en temps continu pour la modélisation et l'analyse. Une méthode de détection de points de changement dans le taux de service d'une file $M/M/1/m$ a été proposée par Singh *et al.* [52], fondée sur une fonction de vraisemblance dérivée du nombre de clients observés lors des départs.

Un système de file d'attente $M/M/2$ a été développé par Laxmi *et al.* [47], dans lequel un service secondaire alternatif est intégré à l'étude et applique un cadre d'évaluation des perfor-

mances et de stabilisation du système. Wu *et al.* [51] ont étudié un système de files d'attente à vitesses de service variables afin de contrôler les arrivées de clients, dans le but d'optimiser la performance du système en minimisant les coûts et les temps d'attente. Chen *et al.* [50] ont utilisé des techniques d'optimisation basées sur les événements pour explorer le contrôle du taux de service. Leur étude propose un algorithme pour déterminer les taux de service optimaux, et sa performance est évaluée en comparaison avec des approches d'optimisation basées sur l'état.

Un autre modèle de file d'attente a été étudié par Dudin *et al.* [49], impliquant deux catégories de demandes : un taux de service de type fixe et un taux de service adaptable. Ils ont examiné le comportement du système à l'aide d'un cadre de chaînes de Markov, en se concentrant sur les indicateurs de performance. Lakkumikanthan et Balasubramanian [48] ont exploré une stratégie de taux de service optimal pour une file en temps discret avec tampon fini, en exploitant la théorie des décisions de Markov et la programmation linéaire pour optimiser les coûts en ajustant les taux de service en fonction des arrivées de clients et du réapprovisionnement des stocks.

Su et Li [46] ont analysé un système de files d'attente à serveur unique dans les réseaux de communication utilisant des taux de service variables et divers types de clients. En fonction des différents paramètres du système, différents niveaux de contrôle sont comparés dans leur analyse, permettant de déterminer une politique optimale lorsqu'il existe plusieurs politiques optimales. Büke et Qin [45] ont exploré des files multi-serveurs avec des taux de service variables et ont caractérisé la dynamique du système sous différentes politiques de service, en utilisant la théorie des processus stochastiques pour l'évaluation des performances.

Pour une revue approfondie des contributions associées, le lecteur est invité à consulter les travaux de Dai *et al.* [44], Chen et Xia [43], et Rodrigues *et al.* [41].

Dans cette recherche, on suppose que le temps de service aléatoire est soit S_1, soit S_2. De plus, dans la section 2, le serveur peut passer de S_1 à S_2 à tout moment.

6.2 Description du problème

On note respectivement par $X(t)$ et $X_q(t)$ le nombre total de clients dans le système et le nombre de clients en file d'attente à l'instant t. Il est supposé qu'au temps initial t_0, $X(t_0) = X_q(t_0) \in \{2, 3, \ldots, c_p\}$, où c_p est la capacité du système, qui peut être infinie. Le serveur commence donc son service à l'instant t_0, qui pourrait être, par exemple, l'heure d'ouverture d'un magasin.

On définit *l'instant de premier passage* $T(k)$ $[= T(k, t_0, t_1)]$ comme suit :

$$T(k) = \inf\{t > t_0 : X_q(t) = 0 \text{ ou } t = t_1 (> t_0) \mid X(t_0) = k\}. \tag{6.1}$$

Définition (Instant de Premier Passage). L'instant de premier passage $T(k)$ est défini comme l'instant aléatoire auquel le nombre de clients en file d'attente devient nul, ou bien lorsqu'on atteint un temps terminal fixé t_1, en partant de k clients à l'instant t_0. où :

- $X_q(t)$ désigne le nombre de clients en file d'attente à l'instant t ;
- $X(t_0) = k$ indique le nombre initial de clients dans le système ;
- t_1 est un instant terminal fixé agissant comme une échéance.

Ce temps d'arrêt détermine l'horizon temporel aléatoire du problème de contrôle de la file d'attente en analyse.

Notre objectif est de déterminer si le serveur doit choisir le temps de service S_1 ou S_2 à l'instant t pour minimiser la valeur espérée de la fonction critère

$$J(k) := \int_{t_0}^{T(k)} \left\{ \frac{q_0}{\left(E[S(t)]\right)^2} + C \right\} \mathrm{d}t + K[T(k)], \tag{6.2}$$

où $S(t) = S_1$ ou S_2, q_0 et C sont des constantes positives, et $K(\cdot)$ est la fonction de coût terminal. Si $E[S_1] > E[S_2]$, alors le coût instantané (ou courant) $q_0/(E[S(t)])^2$ sera plus faible. Cependant, il faudra plus de temps pour réduire à zéro le nombre de clients en attente, ce qui entraîne une pénalité plus élevée (puisque $C > 0$). L'optimiseur doit également prendre en compte le coût terminal $K[T(k)]$.

Remarque. Le terme $(E[S(t)])^2$ dans la fonction de coût $J(k)$ correspond au terme quadratique de contrôle $u^2(t)$ dans les modèles quadratiques linéaires (LQ) et quadratiques linéaires gaussiens (LQG), où l'effort de contrôle est pénalisé afin d'équilibrer performance et coût. De plus, lorsque le temps de service $S(t)$ reste constant dans le temps (indépendant de t), et donc $E[T(k)] = E[S]$, comme considéré dans la section 4, l'emploi d'un terme de coût de la forme $q_0/E[S]$ mène à l'annulation du temps de service espéré. Dans ce cas, la durée moyenne du service n'influence pas le critère d'optimisation.

Le problème de commande optimale traité ici est un cas particulier de « homing problem » ; voir Whittle [92]. Dans ces problèmes, l'optimiseur cherche à trouver un contrôle d'un processus stochastique qui minimise la valeur espérée d'une fonction de coût donnée entre un temps initial et l'instant où un certain événement survient. Whittle [92] a traité le cas où le processus stochastique est un processus de diffusion n-dimensionnel. Rishel [100] a formulé de tels problèmes pour des processus de diffusion n-dimensionnels pouvant modéliser l'usure

des dispositifs. Lefebvre a rédigé de nombreux articles sur les problèmes de retour. Dans un article récent ([101]), il a étendu les problèmes de retour au cas des processus autorégressifs ; voir aussi Lefebvre et Pazhoheshfar [86] pour un problème optimal de maintenance.

Des problèmes de retour pour des systèmes de files d'attente $M/G/k$ modifiés ont été traités par Lefebvre et Yaghoubi [91]. Dans ces articles, la variable de contrôle était le nombre de serveurs actifs à un instant donné t.

Remarque. Il existe de nombreuses applications de la théorie des files d'attente. Dans de nombreuses situations, le temps $T(k)$ est en fait une variable aléatoire. Par exemple, supposons que les « clients » dans les systèmes sont des commandes reçues par une entreprise. Le temps nécessaire pour satisfaire les commandes n'est généralement pas fixe. De plus, la constante t_1 pourrait représenter l'échéance à laquelle les commandes doivent être satisfaites, sous peine d'une pénalité imposée à l'entreprise.

La section 4 présente une application aéronautique. Dans ce cas, les clients sont des aéronefs amenés en atelier pour maintenance. L'objectif est bien sûr de remettre les avions en service aussi rapidement que possible. Une fois de plus, le temps requis pour effectuer la maintenance est aléatoire, car des problèmes peuvent être détectés lors de l'inspection de base.

La *fonction de valeur* est ensuite définie

$$F(k,t_0) = \inf_{\substack{S(t) \\ t \in [t_0, T(k))}} E[J(k)]. \tag{6.3}$$

C'est-à-dire que $F(k,t_0)$ représente le coût espéré encouru si l'optimiseur choisit la commande optimale entre l'instant initial t_0 et le temps final aléatoire $T(k)$, lorsque le système contient initialement k clients à l'instant t_0, et que le serveur sélectionne la distribution de temps de service $S(t)$ sur l'intervalle $[t_0, T(k))$. Cette fonction de valeur est définie pour capturer le compromis entre un service accéléré (pour réduire la file d'attente) et les coûts plus élevés résultant de cette opération, évalués sur un horizon temporel aléatoire se terminant lorsqu'il n'y a plus de client dans la file ou lorsqu'un temps terminal fixé t_1 est atteint.

Par ailleurs, on suppose que

$$K[T(k)] = \begin{cases} 0 & \text{si } T(k) \in [t_0, t_1), \\ F_1 & \text{si } T(k) = t_1, \end{cases} \tag{6.4}$$

où $F_1 > 0$.

Cette définition spécifie la forme de la fonction de coût terminal $K[T(k)]$ dans la fonction de

valeur. Elle attribue un coût nul si la file d'attente est vide avant le temps terminal fixé t_1, c'est-à-dire si $T(k) \in [t_0, t_1)$, et une pénalité $F_1 > 0$ si le système n'est pas capable de servir tous les clients avant l'échéance, c'est-à-dire si $T(k) = t_1$. Cette forme reflète une préférence pour le traitement des clients avant l'échéance et pénalise les retards au-delà de l'intervalle de temps prévu.

Il s'ensuit que la fonction de valeur satisfait les conditions aux limites

$$F(k, t_0) = \begin{cases} 0 & \text{si } k = 1, \\ F_1 & \text{si } (k > 1 \text{ et}) \ t_0 = t_1. \end{cases} \tag{6.5}$$

Remarque. Ceci signifie que s'il n'y a qu'un seul client dans le système ($k = 1$), la file est considérée comme immédiatement vide (puisqu'il s'agit d'une file à serveur unique), donc le coût est nul. En revanche, si le temps actuel t_0 atteint le temps terminal t_1 et qu'il reste plus d'un client dans le système ($k > 1$), alors le contrôle n'est plus appliqué au problème, et un coût terminal prédéterminé F_1 est imposé.

Dans la section 3, il est utilisé le fait que lorsque $S_i \sim \text{Exp}(\mu_i)$, c'est-à-dire lorsque S_i suit une loi exponentielle de paramètre μ_i, pour $i = 1, 2$, il est possible d'utiliser la programmation dynamique pour déterminer la commande optimale. Dans le cas général, en supposant que l'optimiseur choisit soit S_1, soit S_2 pour tout t dans $[t_0, T(k))$, la probabilité conditionnelle sera utilisée pour trouver la solution optimale. Ceci sera fait dans la section 4. Enfin, la section 5 conclut par quelques remarques.

6.3 Commande optimale de la file $M/M/1$ modifiée

Lorsque $S_i \sim \text{Exp}(\mu_i)$, il s'agit d'un modèle de file d'attente $M/M/1$ modifié, dans lequel la distribution des temps de service n'est pas fixe. La fonction de coût $J(k)$ s'écrit alors comme suit :

$$J(k) = \int_{t_0}^{T(k)} \left\{ q_0 \mu^2(t) + C \right\} \mathrm{d}t + K[T(k)], \tag{6.6}$$

où $\mu(t) = \mu_1$ ou μ_2.

Soit A le temps mis par un nouveau client pour arriver après l'instant t_0. On a $A \sim \text{Exp}(\lambda)$, donc

$$P[A \leq \Delta t] = 1 - e^{-\lambda \Delta t} = \lambda \Delta t + o(\Delta t) \tag{6.7}$$

Explication. L'équation (6.7) définit la probabilité qu'une arrivée de client survienne dans un court intervalle de temps Δt après l'instant t_0, en supposant que les temps d'arrivée suivent

une distribution exponentielle de taux λ. Soit $A \sim \text{Exp}(\lambda)$ le temps avant la prochaine arrivée. La fonction de répartition (CDF) de la loi exponentielle est donnée par

$$P[A \leq \Delta t] = 1 - e^{-\lambda \Delta t}. \tag{6.8}$$

En appliquant le développement de Taylor à l'ordre un autour de zéro :

$$e^{-\lambda \Delta t} = 1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{2} - \cdots, \tag{6.9}$$

on obtient

$$P[A \leq \Delta t] = 1 - \left(1 - \lambda \Delta t + \frac{(\lambda \Delta t)^2}{2} - \cdots\right) = \lambda \Delta t + o(\Delta t), \tag{6.10}$$

où le terme $o(\Delta t)$ désigne une fonction telle que

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0. \tag{6.11}$$

En d'autres termes, $o(\Delta t)$ englobe tous les termes d'ordre supérieur qui deviennent négligeables par rapport à Δt lorsque $\Delta t \to 0$.

De manière similaire,

$$P[S_i \leq \Delta t] = \mu_i \Delta t + o(\Delta t). \tag{6.12}$$

Explication. L'équation (6.12) donne la probabilité qu'un service soit terminé dans un petit intervalle de temps Δt, en supposant que les temps de service S_i suivent une distribution exponentielle de taux μ_i. Étant donné que $S_i \sim \text{Exp}(\mu_i)$, on a

$$P[S_i \leq \Delta t] = 1 - e^{-\mu_i \Delta t}. \tag{6.13}$$

En utilisant le même développement de Taylor que dans l'équation (6.9), on obtient

$$P[S_i \leq \Delta t] = \mu_i \Delta t + o(\Delta t), \tag{6.14}$$

ce qui exprime l'approximation linéaire de la probabilité de fin de service pour des Δt très petits.

Par ailleurs, la probabilité que deux événements ou plus (arrivées ou départs) se produisent dans un intervalle de longueur Δt est égale à $o(\Delta t)$.

$$P[A \leq \Delta t, D \leq \Delta t] = o(\Delta t). \tag{6.15}$$

Cette équation énonce que la probabilité qu'une arrivée et une fin de service aient lieu dans le même petit intervalle $(t_0, t_0 + \Delta t]$ est négligeable. Par hypothèse d'indépendance entre le temps d'arrivée A et le temps de départ D, on peut écrire :

$$P[A \leq \Delta t,\, D \leq \Delta t] = P[A \leq \Delta t] \cdot P[D \leq \Delta t]. \tag{6.16}$$

En utilisant les approximations du premier ordre précédemment établies aux équations (6.10) et (6.12), on substitue :

$$P[A \leq \Delta t] = \lambda \Delta t + o(\Delta t), \qquad P[D \leq \Delta t] = \eta\mu \Delta t + o(\Delta t). \tag{6.17}$$

En multipliant les deux expressions de l'équation (6.17), on obtient :

$$P[A \leq \Delta t,\, D \leq \Delta t] = (\lambda \Delta t + o(\Delta t))(\eta\mu \Delta t + o(\Delta t)) = \lambda\eta\mu \Delta t^2 + o(\Delta t). \tag{6.18}$$

Puisque $\Delta t^2 = o(\Delta t)$ lorsque $\Delta t \to 0$, on en déduit :

$$P[A \leq \Delta t,\, D \leq \Delta t] = o(\Delta t). \tag{6.19}$$

Puisque, par indépendance, $\min\{A, S_i\} \sim \text{Exp}(\lambda + \mu_i)$, le nombre aléatoire de clients $X(t_0 + \Delta t)$ dans le système à l'instant $t_0 + \Delta t$ sera :

$$X(t_0 + \Delta t) = \begin{cases} k+1 & \text{avec une probabilité } \lambda \Delta t + o(\Delta t), \\ k-1 & \text{avec une probabilité } \mu_i \Delta t + o(\Delta t), \\ k & \text{avec une probabilité } 1 - (\lambda + \mu_i)\Delta t + o(\Delta t). \end{cases} \tag{6.20}$$

Explication. L'équation (6.20) décrit comment le nombre de clients dans le système peut évoluer pendant un très court intervalle de temps Δt après l'instant t_0. Cela signifie :

- un nouveau client peut arriver avec probabilité $\lambda \Delta t + o(\Delta t)$;
- un client peut terminer son service avec probabilité $\mu_i \Delta t + o(\Delta t)$;
- si aucun événement ne se produit, le nombre de clients reste inchangé, avec une probabilité égale à 1 moins la somme des autres probabilités.

Ensuite, à l'aide du principe d'optimalité de Bellman, on peut écrire que

$$\begin{aligned}
F(k,t_0) \;=\; & \inf_{\substack{\mu(t)\\ t\in[t_0,T(k))}} \Bigg\{ E\left[\int_{t_0}^{t_0+\Delta t} \left\{q_0\mu^2(t)+C\right\}\mathrm{d}t \right. \\
& \qquad \left. + \int_{t_0+\Delta t}^{T(k)} \left\{q_0\mu^2(t)+C\right\}\mathrm{d}t\right]\Bigg\} \\
= \; & \inf_{\substack{\mu(t)\\ t\in[t_0,t_0+\Delta t]}} \Big\{ [q_0\mu^2(t_0)+C]\Delta t + E[F(X(t_0+\Delta t),t_0+\Delta t)] \Big\} \\
& \qquad + o(\Delta t) \\
= \; & \inf_{\substack{\mu(t)\\ t\in[t_0,t_0+\Delta t]}} \Big\{ [q_0\mu^2(t_0)+C]\Delta t + F(k+1,t_0+\Delta t)\,\lambda\Delta t \\
& \qquad + F(k-1,t_0+\Delta t)\,\mu(t_0)\,\Delta t \\
& \qquad + F(k,t_0+\Delta t)\big(1-[\lambda+\mu(t_0)]\Delta t\big)\Big\} \\
& \qquad + o(\Delta t). \qquad (6.21)
\end{aligned}$$

Explication. L'équation (6.21) est construite en appliquant le principe d'optimalité de Bellman sur un court intervalle de temps Δt, en partant de l'instant actuel t_0 avec k clients présents dans le système. L'objectif est de déterminer le coût total espéré $F(k,t_0)$ comme la somme de :

1. le coût immédiat espéré encouru pendant l'intervalle $[t_0, t_0+\Delta t]$;
2. le coût futur espéré à partir du temps $t_0+\Delta t$ jusqu'au temps terminal aléatoire $T(k)$, conditionné aux événements pendant Δt.

Le coût immédiat sur l'intervalle $[t_0, t_0+\Delta t]$ est donné par :

$$[q_0\mu^2(t_0)+C]\Delta t, \qquad (6.22)$$

où :

— $q_0\mu^2(t_0)$ correspond à l'effort de contrôle (service plus rapide),

— C représente une pénalité fixe (ex. coût de congestion ou de retard).

Ensuite, on considère le coût futur espéré à partir de l'instant $t_0+\Delta t$, dépendant de l'état du système à ce moment, conditionné aux événements. En utilisant les probabilités de transition

de l'équation (6.20), le coût futur espéré est :

$$\begin{aligned}\mathbb{E}[F(X(t_0+\Delta t), t_0+\Delta t)] = {} & F(k+1, t_0+\Delta t)\cdot(\lambda\Delta t + o(\Delta t)) \\ & + F(k-1, t_0+\Delta t)\cdot(\mu(t_0)\Delta t + o(\Delta t)) \\ & + F(k, t_0+\Delta t)\cdot(1-(\lambda+\mu(t_0))\Delta t + o(\Delta t)). \end{aligned} \tag{6.23}$$

En combinant le coût immédiat de l'équation (6.22) et le coût futur espéré de l'équation (6.23), on obtient :

$$F(k,t_0) = \inf_{\mu(t)\in\{\mu_1,\mu_2\}} \Big\{[q_0\mu^2(t_0)+C]\Delta t + \mathbb{E}[F(X(t_0+\Delta t), t_0+\Delta t)]\Big\} + o(\Delta t), \tag{6.24}$$

qui est identique à l'équation (6.21) dans notre modèle. Cette structure récursive servira de base pour dériver l'équation de programmation dynamique en temps continu dans la limite $\Delta t \to 0$, laquelle sera discutée dans cette étude.

On a, en posant $\mu_0 := \mu(t_0)$,

$$\begin{aligned} F(k,t_0) - F(k,t_0+\Delta t) = \inf_{\mu_0} \Bigg\{ & \left(q_0\mu_0^2 + C\right)\Delta t \\ & + F(k+1, t_0+\Delta t)\,\lambda\,\Delta t \\ & + F(k-1, t_0+\Delta t)\,\mu_0\,\Delta t \\ & - F(k, t_0+\Delta t)\,(\lambda+\mu_0)\,\Delta t \Bigg\} \\ & + o(\Delta t). \end{aligned} \tag{6.25}$$

Explication. L'équation (6.25), où $\mu_0 := \mu(t_0)$ est introduit pour simplifier la notation, est obtenue en réorganisant le membre droit de l'équation (6.21) afin de mettre en évidence la différence entre la fonction de valeur actuelle $F(k,t_0)$ et sa valeur future $F(k,t_0+\Delta t)$.

Plus précisément, on soustrait $F(k,t_0+\Delta t)$ des deux côtés (ce terme ayant été mis en facteur dans la moyenne pondérée du membre droit), pour montrer comment le changement dans la fonction de valeur sur le petit intervalle Δt peut être exprimé en termes de :

— le coût immédiat de contrôle $q_0\mu_0^2$ et le coût de pénalité C,

— les transitions possibles du nombre de clients (vers $k+1$, $k-1$ ou restant à k).

Par ailleurs, en utilisant le développement de Taylor, on peut écrire :

$$F(\cdot, t_0+\Delta t) = F(\cdot,t_0) + \Delta t\, F'(\cdot,t_0) + o(\Delta t). \tag{6.26}$$

Explication. En appliquant un développement de Taylor du premier ordre à la fonction de valeur $F(k, t_0 + \Delta t)$ autour du point t_0, en supposant que F est différentiable par rapport au temps, on obtient :

$$F(k, t_0 + \Delta t) = F(k, t_0) + \Delta t \cdot \frac{\partial F}{\partial t_0}(k, t_0) + o(\Delta t), \tag{6.27}$$

où :

- $F(k, t_0)$ est la fonction de valeur à l'instant courant,
- $\frac{\partial F}{\partial t_0}(k, t_0)$ est la dérivée partielle de F par rapport à la variable temporelle t_0,
- $o(\Delta t)$ désigne les termes d'ordre supérieur qui tendent vers zéro plus rapidement que Δt lorsque $\Delta t \to 0$.

Il en résulte que :

$$F(k+1, t_0 + \Delta t)\,\lambda \Delta t = F(k+1, t_0)\,\lambda \Delta t + o(\Delta t). \tag{6.28}$$

Explication. L'équation (6.28) est dérivée en appliquant un développement de Taylor du premier ordre au terme $F(k+1, t_0 + \Delta t)$ autour de t_0. En supposant la différentiabilité de $F(k, t)$ par rapport au temps, on écrit :

$$F(k+1, t_0 + \Delta t) = F(k+1, t_0) + \Delta t \cdot \frac{\partial F}{\partial t_0}(k+1, t_0) + o(\Delta t). \tag{6.29}$$

Cependant, dans le développement de Bellman ci-dessus (équation (6.25)), le terme $F(k+1, t_0 + \Delta t)$ est multiplié par le taux d'arrivée λ et par Δt, ce qui donne la forme :

$$F(k+1, t_0 + \Delta t) \cdot \lambda \cdot \Delta t.$$

En substituant le développement de Taylor, on obtient :

$$\begin{aligned} F(k+1, t_0 + \Delta t) \cdot \lambda \cdot \Delta t &= \left(F(k+1, t_0) + \Delta t \cdot \frac{\partial F}{\partial t_0}(k+1, t_0) + o(\Delta t) \right) \cdot \lambda \cdot \Delta t \\ &= \lambda \cdot F(k+1, t_0) \cdot \Delta t + o(\Delta t), \end{aligned} \tag{6.30}$$

car le terme du second ordre $\lambda \cdot \frac{\partial F}{\partial t_0}(k+1, t_0) \cdot \Delta t^2$ est d'ordre $o(\Delta t)$ et donc négligeable dans la limite.

Par conséquent, on peut approcher :

$$F(k+1, t_0 + \Delta t) \cdot \lambda \cdot \Delta t = \lambda \cdot F(k+1, t_0) \cdot \Delta t + o(\Delta t), \tag{6.31}$$

ce qui justifie l'équation (6.28). La même méthode d'approximation s'applique aux termes contenant $F(k-1, t_0+\Delta t)$ et $F(k, t_0+\Delta t)$ dans l'équation (6.28), chacun étant multiplié respectivement par $\mu_0 \cdot \Delta t$ et $(\lambda+\mu_0) \cdot \Delta t$.

Ainsi, en appliquant également la formule de Taylor aux termes $F(k-1, t_0+\Delta t)\mu_0\Delta t$ et $F(k, t_0+\Delta t)(\lambda+\mu_0)\Delta t$, on obtient :

$$\begin{aligned} F(k,t_0) - F(k,t_0+\Delta t) = \inf_{\mu_0} \Big\{ & \left(q_0\mu_0^2 + C\right)\Delta t \\ & + F(k+1,t_0)\lambda\Delta t \\ & + F(k-1,t_0)\mu_0\Delta t \\ & - F(k,t_0)(\lambda+\mu_0)\Delta t \Big\} \\ & + o(\Delta t). \end{aligned} \tag{6.32}$$

Enfin, en divisant les deux membres de l'équation précédente par Δt et en faisant tendre Δt vers zéro, on obtient la proposition suivante.

Proposition 1. *Si $S_i \sim Exp(\mu_i)$, pour $i = 1, 2$, la fonction de valeur $F(k, t_0)$ satisfait l'*équation de programmation dynamique *(EPD)*

$$\begin{aligned} -\frac{\partial}{\partial t_0}F(k,t_0) = \inf_{\mu_0} \big\{ & q_0\mu_0^2 + C + F(k+1,t_0)\lambda + F(k-1,t_0)\mu_0 \\ & - F(k,t_0)(\lambda+\mu_0) \big\}. \end{aligned} \tag{6.33}$$

De plus, les conditions aux limites sont données par l'équation (6.5).

Remarque. L'équation (6.33) est valide pour tout $k \in \{2, 3, \ldots\}$ si la capacité du système est infinie. Dans le cas où $c_p < \infty$, l'équation (6.20) pour $k = c_p$ se réduit à :

$$\begin{cases} k-1 & \text{avec probabilité } \mu_i\Delta t + o(\Delta t), \\ k & \text{avec probabilité } 1 - \mu_i\Delta t + o(\Delta t). \end{cases} \tag{6.34}$$

Explication. L'équation (6.34) montre la forme simplifiée des probabilités des événements futurs lorsque la capacité maximale du système est atteinte, c'est-à-dire lorsque $k = c_p$ (le nombre maximal de clients autorisés dans le système). Dans ce cas, la file est bloquée, ce qui signifie qu'aucune nouvelle arrivée n'est possible.

Rappelons à partir de l'équation (6.20) que, en général, le système peut passer de l'état k vers :

— $k+1$ avec une probabilité $\lambda\Delta t + o(\Delta t)$ (arrivée d'un nouveau client),
— $k-1$ avec une probabilité $\mu_i\Delta t + o(\Delta t)$ (fin de service d'un client),
— k avec une probabilité $1-(\lambda+\mu_i)\Delta t + o(\Delta t)$ (aucun événement).

Cependant, lorsque $k = c_p$, aucune arrivée supplémentaire n'est autorisée. Ainsi, la probabilité de transition vers l'état $k+1$ doit être nulle, car le système ne peut pas accueillir plus de c_p clients. Les événements possibles se réduisent donc à :

$$X(t_0+\Delta t) = \begin{cases} c_p - 1 & \text{avec une probabilité } \mu_i\Delta t + o(\Delta t), \\ c_p & \text{avec une probabilité } 1-\mu_i\Delta t + o(\Delta t). \end{cases} \tag{6.35}$$

Il s'ensuit que l'équation de programmation dynamique devient :

$$-\frac{\partial}{\partial t_0}F(k,t_0) = \inf_{\mu_0}\left\{q_0\mu_0^2 + C + F(k-1,t_0)\,\mu_0 - F(k,t_0)\,\mu_0\right\}. \tag{6.36}$$

pour $k = c_p$. On pourrait également supposer, par exemple, que si un client arrive alors que le système est déjà plein, le problème de contrôle est arrêté et l'on pose $F(c_p+1,t_0) = F_2 \geq 0$. On aurait alors :

$$-\frac{\partial}{\partial t_0}F(k,t_0) = \inf_{\mu_0}\left\{q_0\mu_0^2 + C + F_2\,\lambda + F(k-1,t_0)\,\mu_0 - F(k,t_0)\,(\lambda+\mu_0)\right\} \tag{6.37}$$

pour $k = c_p$. Dans cette étude, il est supposé que l'équation (6.36) est celle qui s'applique lorsque $k = c_p$.

Ensuite, comme $\mu_0 = \mu_1$ ou μ_2, l'équation (6.33) peut être réécrite comme suit :

$$\begin{aligned} -\frac{\partial}{\partial t_0}F(k,t_0) = \min_{\mu_0\in\{\mu_1,\mu_2\}} \big\{ & q_0\mu_0^2 + C + F(k+1,t_0)\,\lambda + F(k-1,t_0)\,\mu_0 \\ & - F(k,t_0)\,(\lambda+\mu_0)\big\}. \end{aligned} \tag{6.38}$$

Explication. L'équation (6.38) est une forme particulière de l'équation générale de programmation dynamique présentée précédemment dans l'équation (6.33). Ce qui distingue les deux équations, c'est le domaine de choix de la variable de contrôle μ_0.

Dans l'équation (6.33), l'optimisation est effectuée sur tous les μ_0 possibles (c'est-à-dire toutes les valeurs non négatives, puisque μ_0 représente un taux de service) :

$$-\frac{\partial F}{\partial t_0}(k,t_0) = \inf_{\mu_0}\left\{q_0\mu_0^2 + C + F(k+1,t_0)\lambda + F(k-1,t_0)\mu_0 - F(k,t_0)(\lambda+\mu_0)\right\}, \tag{6.39}$$

ce qui permet un contrôle continu du taux de service.

En revanche, l'équation (6.40) limite le choix de μ_0 à un ensemble discret de taux de service, à savoir μ_1 et μ_2 :

$$-\frac{\partial F}{\partial t_0}(k,t_0) = \min_{\mu_0 \in \{\mu_1,\mu_2\}} \left\{ q_0\mu_0^2 + C + F(k+1,t_0)\lambda + F(k-1,t_0)\mu_0 - F(k,t_0)(\lambda+\mu_0) \right\}. \tag{6.40}$$

La structure de notre problème initial est capturée dans cette formulation, où le serveur peut uniquement basculer entre deux taux de service prédéterminés. Le décideur remplace donc l'infimum par un minimum, et la comparaison est effectuée en évaluant la fonction objectif pour les deux cas μ_1 et μ_2.

Pour déterminer la valeur de $F(k,t_0)$, on peut considérer l'équation différentielle issue de l'équation (6.40) lorsque l'optimiseur choisit $\mu_0 = \mu_i$ pour toute valeur de k et tout t_0, à savoir :

$$-F_i'(k,t_0) = q_0\mu_i^2 + C + F_i(k+1,t_0)\,\lambda + F_i(k-1,t_0)\,\mu_i - F_i(k,t_0)\,(\lambda+\mu_i), \tag{6.41}$$

sous les conditions données dans l'équation (6.5). Pour obtenir $F_i(k,t_0)$, il faut en général résoudre un système d'équations différentielles linéaires du premier ordre.

Une fois les fonctions $F_i(k,t_0)$ calculées pour $i = 1,2$ et pour tout k, on définit :

$$\begin{aligned} G_i(k,t_0) \;=\; & q_0\mu_i^2 + C + \min\{F_1(k+1,t_0), F_2(k+1,t_0)\}\,\lambda \\ & +\; \min\{F_1(k-1,t_0), F_2(k-1,t_0)\}\,\mu_i \\ & -\; \min\{F_1(k,t_0), F_2(k,t_0)\}\,(\lambda+\mu_i) \end{aligned} \tag{6.42}$$

pour $i = 1,2$.

Explication. La fonction $G_i(k,t_0)$ est définie pour représenter le membre de droite de l'équation de programmation dynamique (6.33) lorsque le contrôle est fixé à $\mu_i \in \{\mu_1,\mu_2\}$. C'est-à-dire,

$$G_i(k,t_0) = q_0\mu_i^2 + C + F(k+1,t_0)\lambda + F(k-1,t_0)\mu_i - F(k,t_0)(\lambda+\mu_i), \tag{6.43}$$

où :

- $F(k+1,t_0)$, $F(k,t_0)$, et $F(k-1,t_0)$ sont les évaluations de la fonction de valeur,
- μ_i est le taux de service sélectionné (soit μ_1, soit μ_2),
- λ est le taux d'arrivée, et

— C est le coût de pénalité associé à l'attente des clients.

En pratique, puisque la commande optimale dépend du temps, la fonction de valeur est évaluée pour les deux politiques μ_1 et μ_2, ce qui génère deux fonctions de valeur candidates de la forme $F_1(k, t_0)$ et $F_2(k, t_0)$, respectivement. Ces fonctions sont les solutions des équations différentielles :

$$-\frac{\partial F_i}{\partial t_0}(k, t_0) = q_0\mu_i^2 + C + F_i(k+1, t_0)\lambda + F_i(k-1, t_0)\mu_i - F_i(k, t_0)(\lambda + \mu_i), \tag{6.44}$$

pour $i = 1, 2$, sous les conditions aux limites spécifiées dans l'équation (6.5).

Une fois que $F_1(k, t_0)$ et $F_2(k, t_0)$ sont obtenues, une comparaison peut être effectuée à chaque instant pour choisir le contrôle préféré. Dans ce contexte, l'équation (6.43) devient :

$$\begin{aligned} G_i(k, t_0) = q_0\mu_i^2 + C &+ \min\{F_1(k+1, t_0), F_2(k+1, t_0)\} \cdot \lambda \\ &+ \min\{F_1(k-1, t_0), F_2(k-1, t_0)\} \cdot \mu_i \\ &- \min\{F_1(k, t_0), F_2(k, t_0)\} \cdot (\lambda + \mu_i), \end{aligned} \tag{6.45}$$

où chaque $G_i(k, t_0)$ est calculé en supposant que le taux de service choisi est μ_i, et les quantités G_1 et G_2 seront comparées pour obtenir la commande optimale à l'instant $t_0 < t_1$.

Enfin, pour obtenir la fonction de valeur $F(k, t_0)$, on peut résoudre l'équation :

$$-F'(k, t_0) = G_i(k, t_0) \tag{6.46}$$

dans tout intervalle où la commande optimale est $\mu_0 = \mu_i$, en utilisant la condition limite $F(k, t_1) = F_1$ ainsi que la continuité de la fonction $F(k, t_0)$.

Remarque. Il est important de noter que la programmation dynamique ne fonctionne que pour le modèle $M/M/1$ (à l'exception de quelques cas particuliers considérés dans des travaux récents de M.Lefebvre). En effet, les résultats donnés dans les équations (6.8) et (6.12) ne sont pas valides pour des distributions générales. Si $P[A \leq \Delta t]$ ou $P[S_i \leq \Delta t]$ n'est pas proportionnel à Δt, un problème surgit lorsqu'on divise les deux membres de l'équation (6.21) par Δt et que l'on fait tendre Δt vers zéro. Supposer que A et S_i suivent des lois exponentielles est une hypothèse simplificatrice. Cependant, il est raisonnable de considérer que le modèle $M/M/1$ peut servir de modèle approximatif dans de nombreuses applications.

On considère maintenant un exemple qui sera résolu explicitement.

6.3.1 Un exemple

Le problème le plus simple que l'on puisse considérer est celui où $k = c_p = 2$. On a alors l'équation suivante [voir l'équation (6.36)] :

$$-F_i'(2, t_0) = q_0 \mu_i^2 + C - F_i(2, t_0)\mu_i. \tag{6.47}$$

Explication. L'équation (6.47) décrit l'équation de programmation dynamique lorsque le système est plein.

Dans cet état, les clients ne sont plus autorisés à entrer dans le système ; celui-ci ne peut donc passer à $k = 1$ qu'en servant un client. Ainsi, les probabilités de transition se simplifient comme indiqué dans l'équation (6.34), et l'équation générale de programmation dynamique (voir l'équation (6.39)) se réduit à :

$$-\frac{\partial F_i}{\partial t_0}(2, t_0) = q_0 \mu_i^2 + C + F_i(1, t_0) \cdot \mu_i - F_i(2, t_0) \cdot \mu_i. \tag{6.48}$$

D'après la condition limite dans l'équation (6.5), on a $F_i(1, t_0) = 0$, car lorsqu'il ne reste qu'un client dans le système, la file est considérée comme vide. En remplaçant cela dans l'équation, on obtient :

$$-\frac{\partial F_i}{\partial t_0}(2, t_0) = q_0 \mu_i^2 + C - F_i(2, t_0) \cdot \mu_i, \tag{6.49}$$

ce qui donne à nouveau l'équation (6.47).

La solution qui satisfait la condition limite $F_i(2, t_1) = F_1$ est donnée par :

$$F_i(2, t_0) = \frac{q_0 \mu_i^2 + C}{\mu_i} \left[1 - e^{-\mu_i(t_1 - t_0)}\right] + F_1 e^{-\mu_i(t_1 - t_0)}. \tag{6.50}$$

Explication. On réécrit l'équation (6.49) sous la forme standard d'une EDO linéaire :

$$\frac{\partial F_i}{\partial t_0}(2, t_0) - \mu_i F_i(2, t_0) = -(q_0 \mu_i^2 + C). \tag{6.51}$$

Pour résoudre cette EDO, on utilise le facteur intégrant :

$$I(t_0) = e^{-\mu_i t_0}. \tag{6.52}$$

En multipliant les deux membres de l'équation (6.51) par $I(t_0)$, on obtient :

$$e^{-\mu_i t_0}\frac{\partial F_i}{\partial t_0}(2,t_0) - \mu_i e^{-\mu_i t_0}F_i(2,t_0) = -(q_0\mu_i^2 + C)e^{-\mu_i t_0}, \tag{6.53}$$

ce qui se simplifie en :

$$\frac{d}{dt_0}\left(e^{-\mu_i t_0}F_i(2,t_0)\right) = -(q_0\mu_i^2 + C)e^{-\mu_i t_0}. \tag{6.54}$$

Les deux membres de l'équation sont maintenant intégrés entre t_0 et t_1, en utilisant la condition terminale $F_i(2,t_1) = F_1$:

$$e^{-\mu_i t_1}F_1 - e^{-\mu_i t_0}F_i(2,t_0) = \int_{t_0}^{t_1}(q_0\mu_i^2 + C)e^{-\mu_i s}ds. \tag{6.55}$$

Le calcul de l'intégrale donne :

$$\int_{t_0}^{t_1}(q_0\mu_i^2 + C)e^{-\mu_i s}ds = (q_0\mu_i^2 + C)\cdot\frac{e^{-\mu_i t_0} - e^{-\mu_i t_1}}{\mu_i}. \tag{6.56}$$

En isolant $F_i(2,t_0)$, on obtient :

$$F_i(2,t_0) = e^{\mu_i t_0}\left[e^{-\mu_i t_1}F_1 + \frac{q_0\mu_i^2 + C}{\mu_i}\left(e^{-\mu_i t_0} - e^{-\mu_i t_1}\right)\right], \tag{6.57}$$

ce qui correspond à l'équation (6.50).

On considère $C = 1$, $\mu_1 = 2$, $\mu_2 = 3$, $q_0 = 1$, $F_1 = 20$ et $t_1 = 10$. On a alors :

$$F_1(2,t_0) = \frac{5}{2} + \frac{35}{2}e^{-2(10-t_0)} \tag{6.58}$$

et

$$F_2(2,t_0) = \frac{10}{3} + \frac{50}{3}e^{-3(10-t_0)} \tag{6.59}$$

pour $t_0 \leq 10$. Les fonctions $G_1(2,t_0)$ et $G_2(2,t_0)$ peuvent maintenant être calculées. Ces fonctions sont représentées dans les figures 6.1 et 6.2. Les courbes se croisent lorsque $t_0 = s := 10 - \frac{1}{3}\ln(10)$. Il en résulte que la commande optimale s'écrit :

$$\mu_0^* = \begin{cases} 2 & \text{si } 0 \leq t_0 \leq s, \\ 3 & \text{si } s \leq t_0 < 10. \end{cases} \tag{6.60}$$

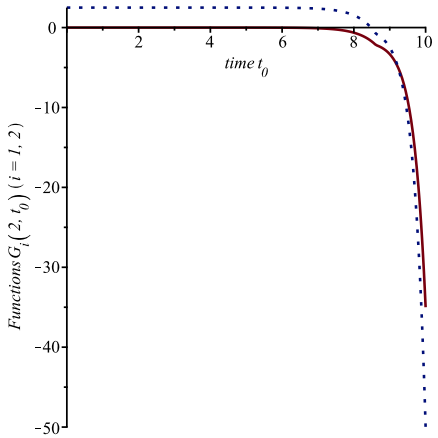

Figure 6.1 La figure présente les fonctions $G_1(2, t_0)$ et $G_2(2, t_0)$ sur l'intervalle $[0, 10]$. La ligne pleine représente G_1 et la ligne pointillée représente G_2. Ces deux fonctions sont calculées à partir des paramètres $C = 1$, $\mu_1 = 2$, $\mu_2 = 3$, $q_0 = 1$, $F_1 = 20$ et $t_1 = 10$. Le point d'intersection de ces deux courbes est $s := 10 - 1/3 \ln(10)$, ce qui suggère le moment de bascule optimal entre μ_1 et μ_2 pour le contrôle du temps de service.

Finalement, en appliquant la démarche précédente, La fonction de valeur $F(2, t_0)$ peut être calculée. Elle est illustrée dans la figure 6.3.

Explication de la figure : Pour $i \in \{1, 2\}$, la fonction $G_i(k, t_0)$ est définie comme suit :

$$G_i(k, t_0) = q_0\mu_i^2 + C + F(k + 1, t_0)\lambda + F(k - 1, t_0)\mu_i - F(k, t_0)(\lambda + \mu_i), \tag{6.61}$$

qui représente le membre de droite de l'équation de programmation dynamique.

Pour la figure 6.1 (où $k = 2$) :

Les expressions ont été évaluées pour $G_1(2, t_0)$ et $G_2(2, t_0)$ en utilisant la définition donnée dans l'équation (6.61). Ces expressions correspondent au membre de droite de l'équation de programmation dynamique lorsque le contrôle est fixé à μ_1 et μ_2, respectivement :

$$G_1(2, t_0) = q_0\mu_1^2 + C + F_1(3, t_0) \cdot \lambda + F_1(1, t_0) \cdot \mu_1 - F_1(2, t_0)(\lambda + \mu_1), \tag{6.62}$$

$$G_2(2, t_0) = q_0\mu_2^2 + C + F_2(3, t_0) \cdot \lambda + F_2(1, t_0) \cdot \mu_2 - F_2(2, t_0)(\lambda + \mu_2). \tag{6.63}$$

Comme le système a une capacité finie $c_p = 2$, les états au-delà de $k = 2$ ne sont pas permis. Ainsi, on pose : $F_1(3, t_0) = F_2(3, t_0) = 0$, et selon les conditions aux limites, $F_1(1, t_0) =$

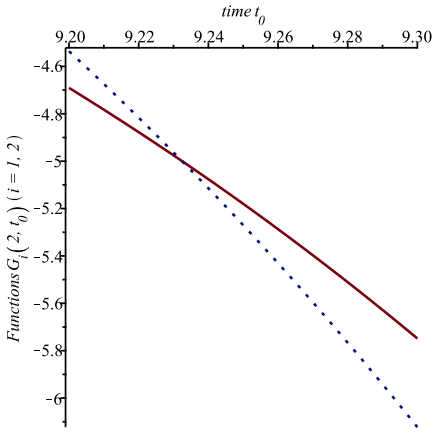

Figure 6.2 Les fonctions $G_1(2, t_0)$ et $G_2(2, t_0)$ sont zoomées sur l'intervalle critique $[9.2, 9.3]$. On observe sur la figure le moment exact de l'intersection, point à partir duquel le contrôle doit être modifié, validant ainsi le point de bascule optimal identifié dans la figure 6.1.

$F_2(1, t_0) = 0$. En substituant ces valeurs, les équations se simplifient comme suit :

$$G_1(2, t_0) = q_0\mu_1^2 + C - F_1(2, t_0)(\lambda + \mu_1), \tag{6.64}$$

$$G_2(2, t_0) = q_0\mu_2^2 + C - F_2(2, t_0)(\lambda + \mu_2). \tag{6.65}$$

Ces expressions simplifiées sont utilisées pour générer les courbes présentées dans la figure 6.1.

Dans la section suivante, Le cas est considéré où les temps de service ne sont pas nécessairement exponentiels.

6.4 Commande optimale de la file d'attente $M/G/1$ modifiée

Supposons que l'optimiseur doive choisir soit le temps de service aléatoire S_1, soit S_2, à partir du temps initial t_0 jusqu'au temps final $T(k)$, et que la fonction $K(\cdot)$ soit identiquement nulle. Alors, la valeur espérée de la fonction de coût $J(k)$ définie dans l'équation (6.2) devient (en posant $t_0 = 0$) :

$$E[J(k)] = E\left[\int_0^{T(k)} \left\{\frac{q_0}{(E[S])^2} + C\right\} \mathrm{d}t\right] = \left(\frac{q_0}{(E[S])^2} + C\right) E[T(k)], \tag{6.66}$$

où $S = S_1$ ou S_2.

Explication. Dans l'équation (6.66), l'intégrale du membre de gauche est éliminée dans l'ex-

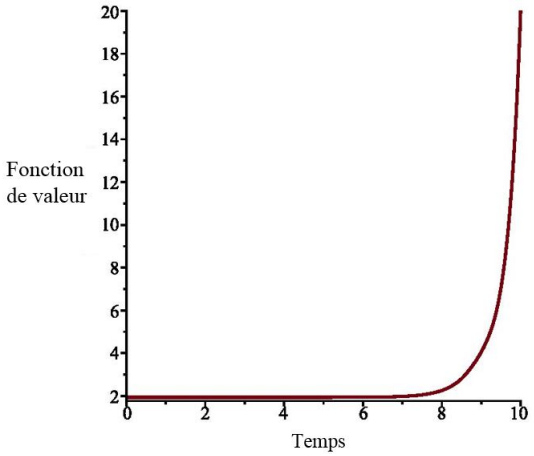

Figure 6.3 Cette figure illustre la fonction de valeur $F(2, t_0)$ sur l'intervalle $[0, 10]$. Les paramètres utilisés sont les mêmes que dans les figure 6.1 et figure 6.2 ($C = 1$, $\mu_1 = 2$, $\mu_2 = 3$, $q_0 = 1$, $F_1 = 20$, et $t_1 = 10$). La fonction de valeur $F(2, t_0)$ exprime le coût minimal attendu sous la stratégie de commande optimale obtenue précédemment.

pression finale parce que l'intégrande est constant par rapport au temps. Plus précisément, le terme :

$$\left\{ \frac{q_0}{(E[S])^2} + C \right\} \tag{6.67}$$

ne dépend pas de t, et peut donc être sorti de l'intégrale. Ensuite, par linéarité de l'espérance, on a :

$$\mathbb{E}\left[\int_0^{T(k)} \left\{ \frac{q_0}{(E[S])^2} + C \right\} dt \right] = \left(\frac{q_0}{(E[S])^2} + C \right) \cdot \mathbb{E}[T(k)], \tag{6.68}$$

ce qui donne l'expression simplifiée finale dans l'équation (6.66).

La section suivante examine plusieurs cas particuliers.

Cas 1. Supposons que $k = c_p = 2$. Comme le système est plein, aucun nouveau client ne peut entrer entre t_0 et $T(2)$. De plus, on peut écrire que $T(2)$ est simplement égal à S. Il en résulte que

$$E[J(2)] = \left(\frac{q_0}{(E[S])^2} + C \right) E[S]. \tag{6.69}$$

Explication. Dans le Cas 1, puisque le système est plein, le seul événement possible est le service d'un client. Une fois que le premier client est servi, le système passe à l'état $k = 1$, auquel cas la fonction de valeur devient nulle grâce à la condition aux limites $F(1, t_0) = 0$.

Par conséquent, le coût s'accumule uniquement jusqu'à ce que le service du premier client soit terminé. Cette durée correspond à une réalisation unique du temps de service aléatoire S, choisi selon la distribution de service sélectionnée. Ainsi, dans ce cadre de modélisation,

on a $T(2) = S$.

Dans le cas particulier où $S_i \sim \text{Exp}(\mu_i)$, pour $i = 1, 2$, l'équation ci-dessus devient

$$E[J(2)] = \left(q_0 \mu^2 + C\right) \frac{1}{\mu}, \tag{6.70}$$

ce qui donne

$$F(2, t_0) = \min_{\mu \in \{\mu_1, \mu_2\}} \frac{q_0 \mu^2 + C}{\mu}. \tag{6.71}$$

Supposons que $\mu_1 = 1$, $\mu_2 = 2$ et $C = 1$. Alors, on trouve que la solution optimale est $\mu^* = \mu_1$ si et seulement si $q_0 \geq 1/2$.

Explication. En substituant $\mu_1 = 1$, $\mu_2 = 2$ et $C = 1$, on évalue les deux options :

$$\text{Pour } \mu = \mu_1 : \quad \frac{q_0 + 1}{1}, \qquad \text{Pour } \mu = \mu_2 : \quad \frac{4q_0 + 1}{2}. \tag{6.72}$$

Pour déterminer quand μ_1 est optimale, comparons :

$$q_0 + 1 \leq \frac{4q_0 + 1}{2}, \tag{6.73}$$

$$2q_0 + 2 \leq 4q_0 + 1, \tag{6.74}$$

$$-2q_0 + 1 \leq 0 \quad \Rightarrow \quad q_0 \geq \frac{1}{2}. \tag{6.75}$$

Remarque. Si l'on fait tendre t_1 vers l'infini dans l'équation (6.57), on obtient :

$$F_i(2, t_0) = \frac{q_0 \mu_i^2 + C}{\mu_i}, \tag{6.76}$$

ce qui est en accord avec l'équation (6.71).

Cas 2. Soit $k = 2$ et $c_p = 3$. On peut alors écrire :

$$T(2) = \begin{cases} S_I & \text{si } S_I \leq A, \\ S_I + T(2) & \text{si } A < S_I, \end{cases} \tag{6.77}$$

où S_I est le temps nécessaire pour servir le premier client, et $A \sim \text{Exp}(\lambda)$ est le temps entre deux arrivées successives, supposé indépendant des temps de service.

Explication. L'équation (6.77) caractérise le temps de premier passage $T(2)$. Elle distingue deux scénarios mutuellement exclusifs :

— **Si** $S_I \leq A$ **:** Aucune arrivée ne se produit avant la fin du service pour le premier client. Par conséquent, un client sera servi, sans qu'un nouveau client n'entre dans le système. Le nombre de clients passe de 2 à 1, et le système n'est plus saturé. Dans ce cas, le temps total pour vider la file est simplement égal à la durée du temps de service : $T(2) = S_I$.

— **Si** $A < S_I$ **:** Un nouveau client arrive dans le système avant que le premier service ne soit terminé, augmentant ainsi le nombre de clients de 2 à 3. Le système devient saturé (c'est-à-dire atteint sa capacité $c_p = 3$). Ensuite, après que le premier client a reçu le service, un client quitte le système et celui-ci revient à l'état avec $k = 2$ clients (le système suit donc la séquence $2 \to 3 \to 2$). Dans ce cas, le processus redémarre essentiellement à partir de la configuration initiale de manière récursive : le temps total pour vider le système devient : temps de service actuel S_I + un autre cycle complet $T(2)$, donc du point de vue du décideur, le processus redémarre à l'état initial : $k = 2$.

Ainsi, l'équation (6.77) met en évidence la structure récursive du temps de premier passage selon les deux événements stochastiques.

Il en résulte que :

$$E[T(2)] = \frac{E[S_I]}{1 - P[A < S_I]}. \tag{6.78}$$

Explication. Soit $p := P(A < S_I)$. En appliquant la loi de l'espérance totale :

$$E[T(2)] = E[T(2) \mid S_I \leq A] \cdot P(S_I \leq A) + E[T(2) \mid A < S_I] \cdot P(A < S_I). \tag{6.79}$$

D'après

$$T(2) = \begin{cases} S_I, & \text{si } S_I \leq A, \\ S_I + T(2), & \text{si } A < S_I, \end{cases} \tag{6.80}$$

on obtient :

$$E[T(2)] = E[S_I] \cdot (1 - p) + (E[S_I] + E[T(2)]) \cdot p. \tag{6.81}$$

On résout :

$$E[T(2)] = E[S_I] + p \cdot E[T(2)] \quad \Rightarrow \quad E[T(2)](1 - p) = E[S_I], \tag{6.82}$$

ce qui donne :

$$E[T(2)] = \frac{E[S_I]}{1 - P(A < S_I)}. \tag{6.83}$$

C'est le résultat donné dans l'équation (6.78).

Si $S_I \sim \text{Exp}(\mu)$, on a :

$$1 - P[A < S_I] = P[S_I \leq A] = \frac{\mu}{\mu + \lambda}. \tag{6.84}$$

Ainsi,

$$E[T(2)] = \frac{\mu + \lambda}{\mu^2}. \tag{6.85}$$

Pour obtenir la solution optimale, Il convient de comparer la valeur de $E[J(2)]$ déduite de l'équation (6.66) et celle de la formule ci-dessus, pour $\mu = \mu_1$ et $\mu = \mu_2$.

Remarque. Soit $Y_1, Y_2, \ldots$ une suite de variables aléatoires indépendantes suivant la loi $\text{Exp}(\mu)$, et soit N une variable aléatoire géométrique de paramètre :

$$p := P[Y_1 < A] = \frac{\mu}{\mu + \lambda}. \tag{6.86}$$

On peut écrire :

$$E[T(2)] = E\left[\sum_{i=1}^{N} Y_i\right] = E[N]\,E[Y_i] = \frac{\mu + \lambda}{\mu}\frac{1}{\mu} = \frac{\mu + \lambda}{\mu^2}. \tag{6.87}$$

Ainsi, on retrouve la formule de l'équation (6.85).

Pour les cas suivants, on note $T(k)$ par $T(k, c_p)$.

Cas 3. Lorsque $k = 3$ et $c_p = 3$, aucun nouveau client ne peut entrer dans le système pendant le service du premier client. Par conséquent :

$$T(3,3) = S_I + T(2,3) \quad \Longrightarrow \quad E[T(3,3)] = \frac{1}{\mu} + \frac{E[S_I]}{1 - P[A < S_I]}. \tag{6.88}$$

Si $S_I \sim \text{Exp}(\mu)$, alors :

$$E[T(3,3)] = \frac{1}{\mu} + \frac{\mu + \lambda}{\mu^2}. \tag{6.89}$$

Remarque. Dans la section 2, pour obtenir la fonction $F_1(3, t_0)$ lorsque $c_p = 3$, on doit résoudre le système d'équations différentielles linéaires du premier ordre :

$$-F_1'(3,t_0) = q_0\mu_1^2 + C + F_1(2,t_0)\mu_1 - F_1(3,t_0)\mu_1, \tag{6.90}$$

$$-F_1'(2,t_0) = q_0\mu_1^2 + C + F_1(3,t_0)\lambda - F_1(2,t_0)(\lambda + \mu_1). \tag{6.91}$$

La solution de ce système, satisfaisant les conditions de l'équation (6.5), est facile à obtenir. En posant $C = 1$ et $\mu_1 = 3$, on trouve que la limite lorsque $t_1 \to \infty$ de la solution est donnée

par :

$$F_1(2,t_0) = \frac{4}{9}(9q_0+1) \quad \text{et} \quad F_1(3,t_0) = \frac{7}{9}(9q_0+1). \tag{6.92}$$

On remarque que $4/9 = E[T(2,3)]$ et $7/9 = E[T(3,3)]$ [voir respectivement les équations (6.85) et (6.89)].

Cas 4. Supposons enfin que $c_p = 4$ et que $k = 2$ ou $k = 3$. On définit l'événement B_i : exactement i clients arrivent pendant le temps de service S_I du premier client, pour $i = 0, 1, 2$, et on note A_I (resp. A_{II}) le temps nécessaire pour l'arrivée du premier (resp. second) nouveau client pendant S_I. On peut écrire :

$$T(3,4) = \begin{cases} S_I + T(2,4) & \text{si } B_0 \text{ a lieu,} \\ S_I + T(3,4) & \text{si } B_1 \text{ a lieu.} \end{cases} \tag{6.93}$$

De même,

$$T(2,4) = \begin{cases} S_I & \text{si } B_0 \text{ a lieu,} \\ S_I + T(2,4) & \text{si } B_1 \text{ a lieu,} \\ S_I + T(3,4) & \text{si } B_2 \text{ a lieu.} \end{cases} \tag{6.94}$$

Explication. Les équations (6.93) et (6.94) représentent le processus récursif pour déterminer le temps total nécessaire à vider le système à partir de l'état $k = 3$ ou $k = 2$ lorsque la capacité est $c_p = 4$.

— Dans l'équation (6.93), si aucun client n'arrive pendant le premier service (B_0), la file diminue à $k = 2$, et le temps restant est $T(2,4)$. Si exactement un client arrive (B_1), la file reste à 3 clients, redémarrant le processus : $T(3,4)$.

— Dans l'équation (6.94), si B_0 a lieu, la file passe à 1 et se vide après un service. Si B_1 a lieu, la file retourne à 2, et le processus se répète. Si B_2 a lieu, la file passe à 3, et le temps restant devient $T(3,4)$.

Ainsi, on obtient :

$$E[T(3,4)] = E[S_I] + E[T(2,4)]\,P[S_I \le A_I] + E[T(3,4)]\,P[S_I > A_I] \tag{6.95}$$

et

$$\begin{aligned} E[T(2,4)] \;&=\; E[S_I] + E[T(2,4)]\,P[A_I < S_I, A_I + A_{II} \ge S_I] \\ &\quad + E[T(3,4)]\,P[A_I + A_{II} < S_I]. \end{aligned} \tag{6.96}$$

Explication. Les équations (6.95) et (6.96) décrivent les espérances du temps pour vider le

système à partir des états $k = 3$ ou $k = 2$, respectivement, dans une file à capacité finie $c_p = 4$.

— Dans l'équation (6.95), le temps moyen nécessaire pour vider le système à partir de l'état $k = 3$ comprend le temps de service initial $E[S_I]$, auquel s'ajoute le temps moyen de continuation selon le nombre d'arrivées pendant S_I. Si aucun client n'arrive (B_0), la file descend à $k = 2$ et le temps restant est $E[T(2, 4)]$; sinon, avec une arrivée (B_1), le système reste à l'état $k = 3$, menant à un temps supplémentaire $E[T(3, 4)]$.
Ici, la condition $P[S_I > A_I]$ correspond au scénario dans lequel le prochain client arrive avant la fin du service en cours, maintenant ainsi le système à l'état $k = 3$. En revanche, $P[S_I \leq A_I]$ indique que le service est terminé avant l'arrivée d'un nouveau client, permettant au système de passer à l'état $k = 2$.

— Dans l'équation (6.96), pour $k = 2$, l'espérance prend en compte trois scénarios : aucune arrivée (B_0), une arrivée (B_1), ou deux arrivées (B_2) pendant S_I. Ces événements entraînent respectivement la fin du processus, un retour à l'état 2, ou une transition vers l'état 3.
Dans cette expression, la condition $P[A_I < S_I, A_I + A_{II} \geq S_I]$ indique le cas où la première arrivée a lieu pendant S_I, mais la seconde non, entraînant une seule arrivée et maintenant le système à $k = 2$. En revanche, $P[A_I + A_{II} < S_I]$ signifie que deux clients arrivent avant la fin du service, ce qui augmente la taille de la file à $k = 3$.
Pour obtenir $E[T(2, 4)]$ et $E[T(3, 4)]$, on peut résoudre le système ci-dessus de deux équations linéaires. Les diverses probabilités présentes dans le système peuvent au moins être évaluées numériquement dans un cas particulier.

6.4.1 Problème particulier

Un problème particulier est considéré dans le cas 4. Supposons que $S_1 \sim \text{Exp}(1)$ et $S_2 \sim \text{G}(2, 2)$, c'est-à-dire que S_2 suit une loi gamma avec les deux paramètres égaux à 2, donc :

$$f_{S_2}(t) = 4te^{-2t} \quad \text{pour } t \geq 0. \tag{6.97}$$

Explication. La forme générale de la loi gamma est donnée par :

$$f(t) = \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \quad \text{pour } t \geq 0, \tag{6.98}$$

où $\Gamma(\alpha)$ est la fonction gamma. Pour $\alpha = 2$ et $\beta = 2$, et sachant que $\Gamma(2) = 1$, on obtient :

$$f_{S_2}(t) = \frac{2^2 t^1 e^{-2t}}{1} = 4te^{-2t}, \quad \text{pour } t \geq 0. \tag{6.99}$$

Pour traiter un cas plus pertinent, la distribution gamma est ici considérée, car elle permet de modéliser un temps de service caractérisé par un délai initial suivi d'une probabilité croissante d'achèvement. La densité de probabilité est asymétrique à droite et est souvent utilisée pour modéliser des durées comme des opérations de maintenance ou des services en plusieurs étapes.

On remarque que $E[S_1] = E[S_2] = 1$.

Explication.

— Pour la distribution exponentielle $S_1 \sim \text{Exp}(1)$, la moyenne est :

$$E[S_1] = \frac{1}{\mu} = \frac{1}{1} = 1. \tag{6.100}$$

— Pour la distribution gamma $S_2 \sim \text{Gamma}(\alpha = 2, \beta = 2)$, la moyenne est :

$$E[S_2] = \frac{\alpha}{\beta} = \frac{2}{2} = 1. \tag{6.101}$$

Il en découle que la commande optimale dépendra uniquement de la valeur de $E[T(k, 4)]$, pour $k = 3$ et $k = 4$.

Posons :

$$p_{1i} := P[S_{Ii} \leq A_I], \tag{6.102}$$

$$p_{2i} := P[A_I + A_{II} < S_{Ii}] \tag{6.103}$$

et

$$p_{3i} := P[A_I < S_{Ii}, A_I + A_{II} \geq S_{Ii}], \tag{6.104}$$

où S_{Ii} est la variable aléatoire S_I si $S = S_i$, pour $i = 1, 2$. On trouve que :

$$p_{11} = \frac{1}{\lambda + 1}, \quad p_{12} = \frac{4}{(\lambda + 2)^2}, \tag{6.105}$$

$$p_{21} = \frac{\lambda^2}{(\lambda + 1)^2}, \quad p_{22} = \frac{\lambda^2(\lambda + 6)}{(\lambda + 2)^3}, \tag{6.106}$$

$$p_{31} = \frac{\lambda}{(\lambda + 1)^2}, \quad p_{32} = \frac{8\lambda}{(\lambda + 2)^3}. \tag{6.107}$$

Explication. Les équations (6.105)-(6.107) représentent les probabilités de trois événements mutuellement exclusifs liés aux arrivées de clients pendant une période de service S_{Ii}, où $i = 1, 2$ correspond à la distribution exponentielle ou gamma.

— **L'équation** (6.105) donne $p_{1i} = P[S_{Ii} \leq A_I]$, la probabilité que le service se termine avant la première arrivée.

— Pour $i = 1$, $S_{I1} \sim \text{Exp}(1)$, donc :

$$p_{11} = \int_0^\infty (1 - e^{-a}) \cdot \lambda e^{-\lambda a}\, da = \frac{1}{\lambda + 1}. \tag{6.108}$$

— Pour $i = 2$, $S_{I2} \sim \text{Gamma}(2, 2)$, et en utilisant la fonction de répartition :

$$p_{12} = \int_0^\infty \left(1 - e^{-2a}(1 + 2a)\right) \cdot \lambda e^{-\lambda a}\, da = \frac{4}{(\lambda + 2)^2}. \tag{6.109}$$

— **L'équation** (6.106) donne $p_{2i} = P[A_I + A_{II} < S_{Ii}]$, la probabilité que deux clients arrivent avant la fin du service.

— Comme $A_I + A_{II} \sim \text{Gamma}(2, \lambda)$, pour $i = 1$, on a :

$$p_{21} = \int_0^\infty e^{-a} \cdot \lambda^2 a e^{-\lambda a}\, da = \frac{\lambda^2}{(\lambda + 1)^2}. \tag{6.110}$$

— Pour $i = 2$, en intégrant la convolution des densités gamma, on obtient :

$$p_{22} = \frac{\lambda^2(\lambda + 6)}{(\lambda + 2)^3}. \tag{6.111}$$

— **L'équation** (6.107) donne $p_{3i} = P[A_I < S_{Ii}, A_I + A_{II} \geq S_{Ii}]$, représentant le cas où un client arrive pendant le service, mais pas un deuxième. Elle s'obtient par soustraction :

$$p_{3i} = 1 - p_{1i} - p_{2i}. \tag{6.112}$$

— Pour $i = 1$:

$$p_{31} = \frac{\lambda}{(\lambda + 1)^2} \tag{6.113}$$

— Pour $i = 2$:

$$p_{32} = \frac{8\lambda}{(\lambda + 2)^3} \tag{6.114}$$

Remarquons que $\sum_{j=1}^{3} p_{ji} = 1$ pour $i = 1, 2$, comme requis.

Soit $E[T_i(k, 4)]$ la valeur de $E[T(k, 4)]$ si $S = S_i$, pour $i = 1, 2$. Pour obtenir $E[T_i(2, 4)]$ et

$E[T_i(3,4)]$, il faut résoudre le système d'équations linéaires suivant :

$$E[T_i(2,4)] = 1 + E[T_i(2,4)]p_{3i} + E[T_i(3,4)]p_{2i}, \tag{6.115}$$

$$E[T_i(3,4)] = 1 + E[T_i(2,4)]p_{1i} + E[T(3,4)](1 - p_{1i}). \tag{6.116}$$

On trouve alors :

$$E[T_1(2,4)] = \lambda^2 + \lambda + 1, \quad E[T_2(2,4)] = \frac{1}{16}\lambda^4 + \frac{1}{2}\lambda^3 + \lambda^2 + \lambda + 1, \tag{6.117}$$

$$E[T_1(3,4)] = \lambda^2 + 2\lambda + 2, \quad E[T_2(3,4)] = \frac{1}{16}\lambda^4 + \frac{1}{2}\lambda^3 + \frac{5}{4}\lambda^2 + 2\lambda + 2. \tag{6.118}$$

Ces fonctions sont illustrées dans les figures 6.4 et 6.5, respectivement.

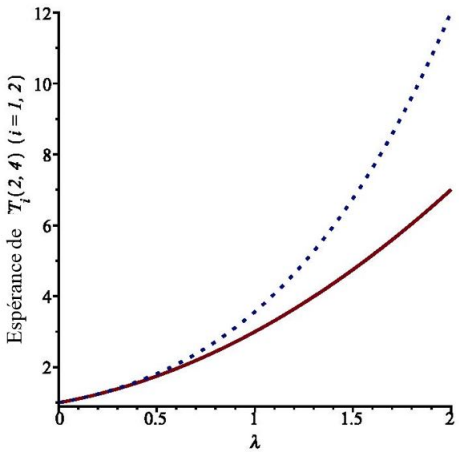

Figure 6.4 Cette figure montre les deux fonctions de temps moyen $E[T_1(2,4)]$ et $E[T_2(2,4)]$ pour $\lambda \in [0,2]$. La ligne continue représente $E[T_1(2,4)]$ et la ligne pointillée $E[T_2(2,4)]$.

On observe que $E[T_1(2,4)]$ et $E[T_2(2,4)]$ sont presque égales lorsque λ est très petit. Toutefois, lorsque λ augmente, $E[T_1(2,4)]$ devient nettement plus petit que $E[T_2(2,4)]$. En fait, on montre que $E[T_1(2,4)] < E[T_2(2,4)]$ pour tout $\lambda > 0$. Ainsi, la solution optimale est toujours de choisir $S = S_1$. Les mêmes conclusions s'appliquent à $E[T_1(3,4)]$ et $E[T_2(3,4)]$.

6.4.2 Une application pratique

Le problème du choix d'une politique de maintenance optimale pour les aéronefs est examiné. En pratique, les programmes de maintenance doivent être optimisés afin de minimiser le temps d'immobilisation et de garantir le niveau de sécurité attendu, deux éléments fondamentaux dans l'industrie aéronautique. Dans notre formulation, il s'agit de trouver un équilibre entre

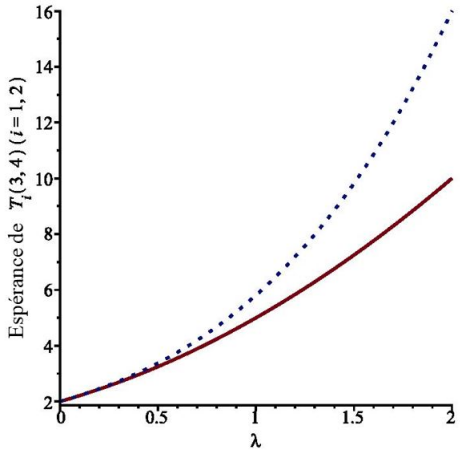

Figure 6.5 Comme la figure 6.4, cette figure illustre les deux fonctions $E[T_1(3,4)]$ et $E[T_2(3,4)]$ sur le même intervalle $[0,2]$ pour λ. La ligne continue représente $E[T_1(3,4)]$ et la ligne pointillée $E[T_2(3,4)]$. Ces résultats mettent en évidence la différence entre les stratégies de gestion de files d'attente dans le cas de distributions de temps de service variables.

les coûts de maintenance et la nécessité de remettre l'appareil en service le plus rapidement possible.

Supposons que S_1 suive une loi de Rayleigh, de densité :

$$f_{S_1}(s) = \frac{s}{\sigma^2} \exp\left\{-\frac{s^2}{2\sigma^2}\right\} \quad \text{pour } s \geq 0, \tag{6.119}$$

où σ est une constante strictement positive, et que

$$f_{S_2}(s) = \theta e^{-\theta s} \quad \text{pour } s \geq 0. \tag{6.120}$$

C'est-à-dire, S_2 suit une loi exponentielle de paramètre θ. La distribution de Rayleigh, avec sa queue plus lourde, est plus adaptée aux activités de maintenance complexes, tandis que la distribution exponentielle peut être utilisée pour les tâches de maintenance courantes.

Les coûts espérés $E[J(3,4)]$ et $E[J(2,4)]$ associés à l'exécution des tâches de maintenance ont été calculés sous chacune des deux hypothèses sur la distribution des temps de service, pour différents taux d'arrivée d'aéronefs. Ces coûts incluent à la fois les coûts liés à la rapidité de l'exécution (dépendant de la distribution de S) et les coûts liés au temps d'immobilisation de l'aéronef (influencés par le coût de pénalité C).

La figure 6.6 présente les coûts espérés $E[J(3,4)]$ et $E[J(2,4)]$ pour $\sigma = 3$, $\theta = 1$, $q_0 = 50$ et $\lambda \in [0.1, 5]$. À la lecture de cette figure, on peut conclure que pour les inspections de routine et les opérations prévisibles, l'usage d'une distribution exponentielle pour le temps de service

serait plus rentable, notamment lorsque le taux d'arrivée est élevé et que des interventions rapides sont requises.

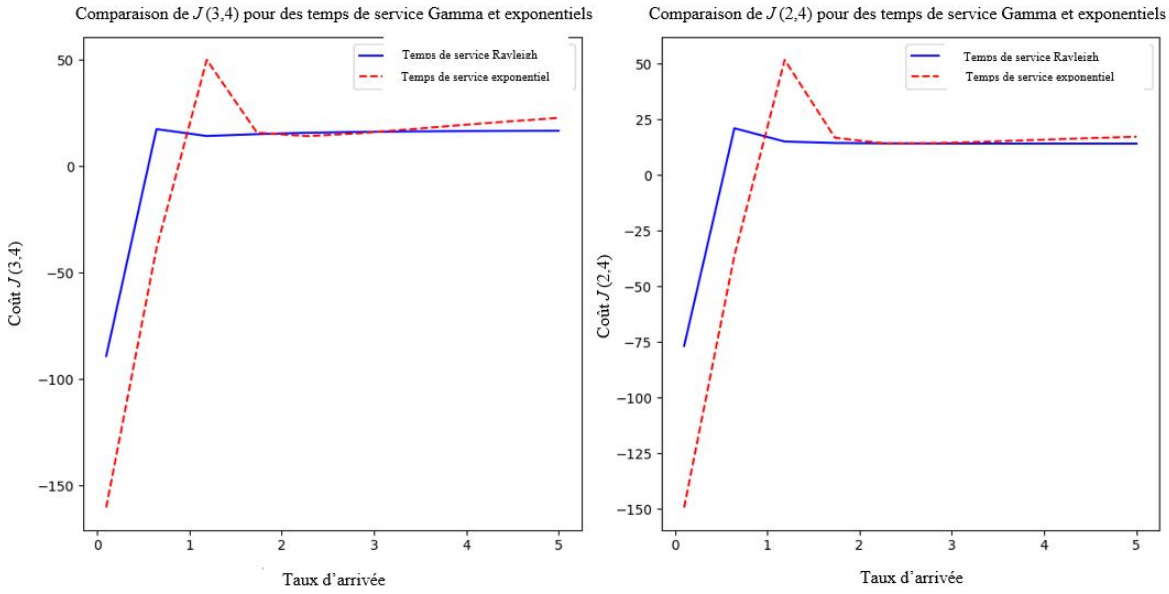

Figure 6.6 Coûts espérés $E[J(3,4)]$ (graphique gauche) et $E[J(2,4)]$ (graphique droite) lorsque S_1 suit une loi de Rayleigh de paramètre $\sigma = 3$ (ligne continue) et S_2 une loi exponentielle de paramètre $\theta = 1$ (ligne pointillée), avec $q_0 = 50$ et $\lambda \in [0.1, 5]$.

Explication – Analyse de sensibilité : Une analyse de sensibilité a été réalisée en supposant que le temps de service suit une distribution de Rayleigh, en comparaison avec la distribution exponentielle, pour différentes valeurs des paramètres. Les résultats ont été recalculés et représentés dans deux graphiques séparés, chacun pour la fonction de coût $J(3,4)$ et $J(2,4)$, sur un intervalle de valeurs de λ allant de 0.1 à 5.

Les résultats de ces analyses mettent en évidence l'effet de λ sur les temps d'attente et les coûts correspondants dans le système de files d'attente, éléments essentiels pour améliorer la performance et l'efficacité économique.

Remarquons que les valeurs des paramètres des distributions sont : $\sigma = 3$ (Rayleigh) ; $\mu = 1$ (Exponentielle).

Pour résoudre le système d'équations avec les probabilités indiquées et mener une analyse pour différentes valeurs de λ, on suit les étapes suivantes :

1. Substitution des probabilités dans les équations.
2. Fixation d'une valeur pour q_0 (ex. $q_0 = 50$).
3. Résolution du système pour $T(2,4)$ et $T(3,4)$.

4. Analyse de sensibilité pour différents λ.
5. Création de graphiques et tableaux pour $E[J(3)]$ et $E[J(2)]$ selon les valeurs de λ.
6. Intégration de l'hypothèse alternative (temps de service exponentiel) dans l'analyse de sensibilité, calcul des coûts, et comparaison graphique entre $J(3,4)$ et $J(2,4)$ sous les deux distributions.

Tout d'abord, les valeurs des probabilités P_2 et P_3 en fonction des différentes valeurs du taux d'arrivée λ sont indiquées dans le tableau 6.1 :

Tableau 6.1 Relation entre les probabilités P_2 et P_3 et le taux d'arrivée λ.

Lambda	P3_rayleigh	P3_Expo
0.1	0.238	0.083
0.6	0.220	0.238
1.2	0.109	0.248
1.7	0.061	0.232
2.3	0.038	0.212
2.8	0.026	0.193
3.4	0.019	0.177
3.9	0.014	0.162
4.5	0.011	0.150
5.0	0.009	0.139

Lambda	P2_rayleigh	P2_Expo
0.1	0.937	0.992
0.6	0.385	0.846
1.2	0.174	0.705
1.7	0.094	0.598
2.3	0.058	0.517
2.8	0.039	0.455
3.4	0.028	0.406
3.9	0.021	0.366
4.5	0.016	0.333
5.0	0.013	0.306

Ensuite, les graphes comparatifs permettant d'évaluer la performance (fonction de coût) des deux distributions, Rayleigh et exponentielle, en fonction des variations du taux d'arrivée λ, sont présentés dans la figure 6.6.

Enfin, un tableau regroupant les valeurs numériques des fonctions de coût associées aux graphes ci-dessus est présenté dans le tableau 6.2.

Discussion

Pour $J(3,4)$, qui représente une fonction de coût dépendant du temps d'attente dans la file, les remarques suivantes peuvent être formulées :

Lorsqu'on utilise la distribution de Rayleigh pour modéliser le temps de service, le coût $J(3,4)$ décroît initialement à mesure que λ augmente dans l'intervalle inférieur à 1. Cela traduit l'effet classique dans les files d'attente : une augmentation du taux d'arrivée (ou du service) réduit le temps d'attente moyen, et donc les coûts associés. Cependant, au-delà de $\lambda = 1$, une nouvelle tendance se manifeste : la fonction de coût commence à croître à cause des dépenses supplémentaires induites par des taux de service plus élevés.

Tableau 6.2 Valeurs de la fonction de coût pour les distributions Rayleigh et exponentielle selon différentes valeurs de λ.

Lambda	J(3,4)_Rayleigh	J(3,4)_Expo
0.1	-89	-160
0.6	17	-39
1.2	14	50
1.7	15	16
2.3	16	14
2.8	16	15
3.4	16	17
3.9	17	19
4.5	17	21
5.0	17	23

Lambda	J(2,4)_Rayleigh	J(2,4)_Expo
0.1	-77	-149
0.6	21	-36
1.2	15	52
1.7	14	17
2.3	14	14
2.8	14	14
3.4	14	15
3.9	14	16
4.5	14	17
5.0	14	17

La distribution exponentielle montre un comportement similaire mais génère des coûts systématiquement plus faibles pour différentes valeurs de λ, ce qui montre qu'elle est plus adaptée pour les tâches de maintenance de routine où la prédictibilité est essentielle. En effet, pour des taux de service identiques, la distribution exponentielle conduit à des coûts d'attente espérés inférieurs à ceux associés à la distribution de Rayleigh, et ce, pour la majorité des valeurs de λ. Ce phénomène peut être attribué à la queue plus lourde de la distribution de Rayleigh, qui reflète une plus grande probabilité de durées de service prolongées – un inconvénient majeur lorsque les tâches de maintenance sont peu fréquentes mais longues. À l'inverse, la distribution exponentielle, grâce à sa propriété de mémoire nulle, s'avère plus performante dans ces cas.

Concernant $J(2,4)$, une fonction de coût affectée par la probabilité que le système soit dans un état donné :

il apparaît que le coût $J(2,4)$ est généralement plus faible pour la distribution exponentielle que pour la distribution de Rayleigh aux faibles valeurs de λ, mais que les coûts convergent vers des valeurs similaires à mesure que λ augmente. Cette tendance, observée dans les deux cas, confirme l'hypothèse selon laquelle un service plus efficace permet de réduire les coûts associés à l'état du système.

Les résultats obtenus offrent des pistes intéressantes pour les travaux de modélisation en théorie des files d'attente et en optimisation des systèmes de service, car ils montrent dans quelle mesure la loi de distribution du temps de service influe sur les fonctions de coût globales. Ces éléments sont essentiels dans la prise de décision pour la conception de systèmes et l'allocation optimale de ressources limitées.

Dans les cas d'activités complexes ou fortement variables (par exemple dues à des imprévus

dans les opérations de maintenance), la distribution de Rayleigh peut représenter un choix de modélisation plus réaliste, bien qu'elle génère un coût plus élevé. Ces résultats sont fondamentaux pour planifier efficacement la disponibilité du personnel et des pièces de rechange.

Les centres de maintenance et compagnies aériennes peuvent ajuster leurs politiques de maintenance en fonction de la fréquence historique des interventions (reflétée par λ) et de la complexité des tâches (complexes ou standards). Par exemple, pendant les périodes de forte activité, il pourrait être plus judicieux de privilégier la réduction du temps d'immobilisation (via un temps de service exponentiel).

6.5 Conclusion

Ce travail étudie un problème de commande optimale dans un modèle de file d'attente avec un seul serveur. Ce modèle constitue une extension du modèle classique $M/G/1$, dans lequel le serveur peut choisir entre deux distributions différentes pour le temps de service. L'objectif est de réduire à zéro le nombre de clients en attente dans la file, tout en prenant en compte un coût de contrôle quadratique. Le temps final du problème est aléatoire et correspond au premier instant où il ne reste plus aucun client dans le système.

Dans la section 6.3, il est supposé que le temps de service suit une loi exponentielle, de paramètre μ_1 ou μ_2. Cela a permis d'utiliser la programmation dynamique pour dériver l'équation vérifiée par la fonction de valeur, laquelle donne le coût espéré lorsque la stratégie optimale est appliquée depuis l'instant initial jusqu'au temps final aléatoire.

Par la suite, le problème est généralisé dans la section 6.4 au cas où le temps de service est une variable aléatoire positive quelconque. Dans ce contexte général, la programmation dynamique n'est plus applicable directement. Plusieurs cas particuliers sont analysés en supposant une capacité finie pour le système. En théorie, tout problème particulier peut être traité, mais lorsque la capacité du système devient importante, la méthode de résolution devient rapidement complexe.

Il est possible d'envisager d'autres modifications du modèle de base $M/G/1$. Par exemple, on pourrait supposer qu'il y a plusieurs serveurs. L'objectif pourrait alors être de contrôler le taux d'entrée des clients dans le système. Si toutes les variables aléatoires du modèle suivent des lois exponentielles, on peut espérer utiliser les principes de la programmation dynamique pour déterminer la solution optimale.

CHAPITRE 7 POLITIQUE OPTIMALE D'INSPECTION ET DE MAINTENANCE : INTÉGRATION D'UNE CHAÎNE DE MARKOV EN TEMPS CONTINU

Ce chapitre s'appuie sur des travaux de recherche en inspection et maintenance fondées sur des chaînes de Markov en temps continu [90], dont les résultats sont ici réintroduits et étendus dans une approche unifiée.

7.1 Introduction

Dans la troisième partie de cette thèse, on a représenté les conditions opérationnelles d'une machine par une chaîne de Markov en temps continu contrôlée, avec des activités de service se produisant à des moments aléatoires. L'objectif principal est de prolonger la période de fonctionnement avant que la machine ne nécessite une réparation, tout en tenant compte des coûts de maintenance associés. À cette fin, l'équation de programmation dynamique satisfaite par la fonction de valeur est dérivée, ce qui permet une prise de décision optimale concernant les fréquences d'inspection, suivie de problèmes spécifiques résolus explicitement. L'approche proposée permet de minimiser les coûts directs de maintenance ainsi que les coûts potentiels de défaillance, établissant ainsi une méthodologie robuste pour déterminer les fréquences d'inspection dans le cadre de la planification de maintenance centrée sur la fiabilité.

La maintenance préventive constitue un élément essentiel de l'ingénierie de la fiabilité, de l'optimisation de la disponibilité et de l'amélioration de l'efficacité opérationnelle. Elle implique des interventions programmées pour éviter les pannes du système et garantir la continuité des opérations. La complexité de l'optimisation de telles stratégies de maintenance augmente lorsque l'aléa intrinsèque du comportement du système est explicitement pris en compte dans la modélisation des phénomènes de dégradation. L'un des cadres mathématiques les plus couramment utilisés pour capturer les processus probabilistes est la chaîne de Markov en temps continu, qui représente les transitions d'état du système sur un horizon de planification déterminé.

Un élément clé de l'optimisation de la maintenance est l'ajustement des coûts associés aux activités de maintenance et aux défaillances potentielles du système. Une politique de maintenance efficace doit prendre en compte à la fois les coûts directs liés à l'exécution des tâches de maintenance et les coûts indirects associés aux arrêts de système, à la perte de production et aux réparations dues aux défaillances. Cet ajustement optimal est essentiel dans les do-

maines où des exigences élevées sont imposées à la fiabilité et à la disponibilité du système, en raison des enjeux de sécurité opérationnelle et de coûts de processus.

Dans le cadre de la maintenance préventive, la commande optimale de la fréquence d'inspection joue un rôle fondamental dans l'atteinte Des objectifs de coût et de fiabilité, lesquels sont examinés en détail dans les sections suivantes de cette thèse. La fréquence d'inspection détermine la fréquence à laquelle l'opérateur doit inspecter l'état du système ; elle affecte donc directement la planification de la mise en œuvre de la maintenance. Un taux d'inspection plus élevé entraîne davantage de tâches d'inspection, ce qui peut réduire le taux de défaillance, mais augmente simultanément les coûts directs d'inspection-maintenance. À l'inverse, un taux plus faible réduit les dépenses immédiates de maintenance, mais peut entraîner des coûts à long terme plus élevés en raison de l'augmentation des défaillances et des temps d'arrêt. Le problème revient à identifier le taux optimal qui équilibre les facteurs contradictoires mentionnés. Ainsi, l'identification d'une politique optimale pour le taux d'inspection est cruciale pour réduire les coûts totaux liés à la maintenance tout en assurant une fiabilité élevée du système.

La présente étude propose un plan optimal d'inspection et de maintenance en modélisant les états de dégradation d'un système comme une chaîne de Markov en temps continu et en ajustant le taux d'inspection u, variable de contrôle spécifiant l'intensité des tâches de maintenance préventive. Ce taux d'inspection u sert de paramètre de contrôle équilibrant le compromis entre les actions de maintenance et le risque de défaillance du système. L'objectif est d'appliquer la programmation dynamique pour déterminer la valeur optimale de u, en minimisant les coûts à long terme associés aux opérations d'inspection et aux temps d'arrêt potentiels du système dus aux défaillances.

La programmation dynamique constitue un cadre naturel pour évaluer des politiques de maintenance dans un contexte d'incertitude. La nature dynamique du problème, dans lequel l'état du système et le choix de la stratégie de maintenance optimale sont influencés par les états et décisions passés, nécessite un cadre capable de traiter la prise de décision séquentielle sous aléa. En appliquant la programmation dynamique dans un cadre de chaîne de Markov en temps continu, l'étude définit un taux optimal d'inspection-maintenance en fonction de l'état du système, ce qui éclaire le positionnement temporel des activités d'inspection-maintenance. Ces perspectives peuvent contribuer à modéliser une planification robuste d'inspection-maintenance adaptable aux modifications en temps réel des conditions du système et des exigences opérationnelles.

Cette méthodologie développe un nouveau type de "homing problem" appliqué à la théorie de la fiabilité, caractérisé par une prise de décision sous aléa. La dérivation et la mise en œuvre

de l'équation de programmation dynamique associée présentent des défis théoriques notables abordés dans cette recherche. Ceux-ci incluent la modélisation précise des phénomènes de dégradation, l'intégration appropriée des coûts d'inspection-maintenance et de temps d'arrêt, ainsi que l'assurance de la faisabilité computationnelle lors de la résolution des équations de programmation dynamique pour des systèmes complexes.

7.2 Revue de la littérature

Une revue approfondie des développements récents en matière de commande optimale des chaînes de Markov en temps continu, en particulier dans le cadre de la théorie de la fiabilité, révèle que le sujet a été largement étudié dans ce domaine, notamment en ce qui concerne la planification de l'inspection et de la maintenance. Les travaux antérieurs couvrent une large gamme de modèles et de stratégies pour établir des stratégies optimales de test et de maintenance visant à améliorer la fiabilité, la disponibilité et la sécurité des machines tout en réduisant les coûts. Ces explorations témoignent des efforts continus pour améliorer les méthodes de commande optimale dans ce domaine. Ce qui suit met en évidence une série chronologique de contributions clés ayant façonné ce domaine.

Pour commencer, G. Parmigiani [54] présente un modèle de planification optimale d'inspections faillibles dans un système de production. L'étude applique une chaîne de Markov en temps continu pour contrôler de manière optimale les fréquences de maintenance afin d'améliorer la fiabilité du système. À la suite de ce travail, E. K. Boukas et Z. K. Liu [55] abordent le contrôle de la maintenance à l'aide d'un processus de Markov en temps continu. Le modèle se concentre sur la commande optimale des taux de maintenance préventive et corrective pour améliorer la fiabilité du système. En s'appuyant sur ces fondations, C.-H. Wang et S. H. Sheu [56] prennent en compte les erreurs d'inspection et développent une politique optimale pour les intervalles d'inspection et la maintenance d'un système de production se dégradant, en utilisant un modèle de chaîne de Markov en temps continu pour minimiser le coût total, en tenant compte de la fiabilité du système. Dans une direction parallèle, H. Suryadi et L. G. Papageorgiou [57] proposent une méthode de programmation mathématique pour optimiser la planification de la maintenance dans les usines, utilisant des chaînes de Markov en temps continu pour modéliser la maintenance afin d'améliorer la fiabilité du système.

De plus, W. Liying et al. [58] identifient une politique de diagnostic des pannes et d'inspection et optimisent le cycle d'inspection pour maximiser les revenus, permettant des réparations imparfaites avant le remplacement des composants. Un processus vectoriel de Markov et un exemple numérique sont utilisés pour validation. D'autres contributions apparaissent dans les travaux de H. R. Golmakani et F. Fattahipour [62], qui utilisent un processus de Markov

pour proposer une stratégie optimale de remplacement et une fréquence d'inspection dans le cadre d'une maintenance conditionnelle. Le modèle vise à optimiser les coûts de maintenance, les intervalles d'inspection équilibrant les coûts de remplacement préventif et de défaillance, améliorant ainsi la fiabilité du système. En élargissant les capacités de modélisation, F. Naderkhani et V. Makis [67] développent une stratégie de maintenance conditionnelle optimale à l'aide d'un processus de Markov caché en temps continu. Elle ajuste les intervalles d'inspection selon les états de vieillissement afin de minimiser les coûts de maintenance et d'améliorer la fiabilité.

Dans une application novatrice, K. He [71] développe une méthode de planification optimale de la maintenance dans les systèmes de santé, en se concentrant sur les événements avec intervalles d'inspection imparfaits et maintenance préventive non ponctuelle, en utilisant des processus stochastiques. Le domaine a également connu des développements dans l'analyse de systèmes multiunités, comme le montrent Q. Sun et al. [74], qui proposent des stratégies optimales d'inspection et de remplacement pour les systèmes vieillissants en utilisant un cadre de processus de décision markovien. Le processus de Wiener est utilisé pour modéliser la dégradation des composants et déterminer les intervalles d'inspection et les choix de remplacement afin de minimiser le coût opérationnel total tout en maintenant la fiabilité du système. Parallèlement, P. Cao et al. [78] proposent une sélection optimale pour les tests logiciels à l'aide d'un contrôle stochastique en temps continu. Il prend en compte le compromis entre les coûts de test et le temps de disponibilité pour minimiser les coûts totaux. Le modèle proposé utilise la programmation dynamique pour déterminer quand tester, livrer ou rejeter un logiciel, en tenant compte de la fiabilité du logiciel.

Également, en abordant la fiabilité structurelle, Q. Sun, Z.-S. Ye, et X. Zhu [79] examinent le contrôle du vieillissement des composants dans les systèmes en série en optimisant la politique de remplacement préventif et de réaffectation, en utilisant l'optimisation stochastique et la chaîne de Markov. Dans le domaine de la surveillance de la santé structurelle, C. P. Andriotis et al. [77] se concentrent sur la commande optimale du cycle de vie des systèmes vieillissants dans un environnement incertain. Un processus de décision markovien partiellement observable est utilisé pour trouver des stratégies optimales d'intervention et de surveillance. Le contrôle de la maintenance dans les systèmes discrets est également exploré par S. Gan et al. [82], qui développent une stratégie de maintenance pour un système de production avec plusieurs états de traitement, intégrant l'âge des machines et la gestion des pièces de rechange pour optimiser les tâches de maintenance. La méthode utilise un processus de décision markovien en temps discret pour identifier les activités de maintenance optimales dans le but de minimiser les coûts globaux à long terme. Le contexte plus large de la maintenance durable est capté par P. Vrignat et al. [102], dont la revue de littérature aborde la fabrication du-

rable en se concentrant sur les politiques de maintenance, la gestion de la santé des systèmes (HMS) et les pronostics. Elle discute de l'intégration de l'industrie 4.0 et de l'e-maintenance pour faire évoluer les stratégies de maintenance proactive compatibles avec les objectifs de durabilité.

Pour aborder la planification complexe de l'inspection des structures, P. G. Morato et al. [85] introduisent un schéma intégrant les réseaux bayésiens dynamiques et les processus de décision markoviens partiellement observables pour explorer la planification optimale de l'inspection et de la maintenance de la détérioration structurelle. L'étude met en évidence les limites de la méthode heuristique pour optimiser la fiabilité structurelle et minimiser les coûts via l'optimisation stochastique. En complément, M. Roux et al. [84] introduisent un modèle efficace de processus de décision markovien partiellement observable pour la planification de la maintenance en présence d'observations imprécises.

En ce qui concerne les modèles en temps discret, M. Lefebvre et P. Pazhoheshfar [86] développent un problème de commande optimale pour la maintenance des machines à l'aide d'un modèle de processus stochastique en temps discret. La méthodologie consiste à résoudre une équation de programmation dynamique pour déterminer si des activités de maintenance doivent être effectuées à chaque unité de temps, sachant que le temps final sera une variable aléatoire. Pour les systèmes disposant de pièces de rechange limitées, Y. Wang et Y. Li [83] développent un schéma de planification de la maintenance avec un processus de décision markovien afin d'explorer la commande optimale d'un système disposant de pièces de rechange limitées sur une période de temps finie.

Une étude récente de S. Nasersarraf et al. [87] développe une politique de maintenance optimale fondée sur l'état pour des panneaux électriques dans des systèmes parallèles, en utilisant un modèle de risques proportionnels pour tenir compte de la dépendance de défaillance entre les composants. Le modèle vise à minimiser les coûts totaux espérés tout en maintenant la fiabilité du système, en déterminant les intervalles d'inspection et les stratégies de remplacement préventif optimaux. D'autres innovations apparaissent chez W. Wang et X. Chen [88], qui proposent un modèle de maintenance conditionnelle utilisant un processus de Markov déterministe par morceaux. Celui-ci intègre la maintenance corrective et imparfaite. L'article applique la théorie de la commande optimale pour explorer la politique de maintenance optimale, dans le but de minimiser le coût total du système. Enfin, G. Wang et Z. Zhu [89] étudient la commande optimale d'un système avec des données échantillonnées et un échantillonnage stochastique. L'article présente plusieurs fonctions de coût, y compris une pour la fréquence d'échantillonnage, et implémente la programmation dynamique pour obtenir des contrôleurs optimaux en temps fini et infini.

Désormais, les chaînes de Markov en temps continu et la programmation dynamique sont de plus en plus appliquées au-delà des contextes industriels classiques, soutenant des stratégies avancées en fiabilité et en optimisation, comme le montrent les études récentes (voir, par exemple, Liu *et al.* [75] et Mancuso *et al.* [80]) dans le contexte de la gestion de la santé des systèmes (PHM).

Malgré des avancées remarquables dans la commande optimale des chaînes de Markov en temps continu appliquées à l'optimisation de la maintenance, il subsiste encore plusieurs lacunes dans cette littérature. De nombreux modèles existants sont adaptés à certaines applications spécifiques, telles que les systèmes de production ou les soins de santé, avec une évaluation incomplète du taux d'inspection optimal dans un cadre plus généralisé. En outre, l'utilisation de la programmation dynamique pour extraire des variables de décision continues, comme les intervalles de test, n'a pas encore été explorée, notamment dans le contexte des problèmes de retour en théorie de la fiabilité.

Cette recherche vise à combler ces lacunes en proposant un cadre étendu pour définir le taux optimal de maintenance préventive et d'inspection d'un système en utilisant les chaînes de Markov en temps continu et la programmation dynamique. Le modèle présenté offrira une approche robuste du problème de retour en optimisation de la fiabilité, qui intègre des techniques avancées de contrôle stochastique pour formuler une planification optimale de la maintenance. De plus, en synthétisant les dernières avancées en matière d'optimisation de la planification inspection-maintenance et de programmation dynamique, cette étude contribue à l'objectif plus large de faire progresser la théorie de la fiabilité et la rentabilité des systèmes d'ingénierie, en proposant des politiques pratiques pour les industries où la fiabilité des composants est d'une grande importance.

Cette recherche traite de la question des moments d'inspection optimaux pour la maintenance dans un cadre généralisé en améliorant l'applicabilité des chaînes de Markov en temps continu dans un contexte plus global. La section 4 de l'étude traite des taux d'inspection obtenus par programmation dynamique, ce qui permet de proposer une méthodologie polyvalente pour déterminer le taux d'inspection optimal dans un système, sans limiter sa pertinence à des domaines spécialisés. Cette méthodologie est appuyée par l'équation (7.35), dans laquelle la formulation généralisée du calcul du taux d'inspection optimal est décrite, améliorant la flexibilité du modèle pour divers contextes d'application en maintenance.

La lacune est en outre comblée par l'intégration de la programmation dynamique pour déterminer des variables de décision continues, comme le taux d'inspection, dans le domaine de la maintenance et de la fiabilité et dans le contexte du problème de retour. Les résultats, en particulier dans les sections 5 et 6, illustrent comment la détermination de l'inspection op-

timale est soutenue par la programmation dynamique, notamment lorsque des temps finaux aléatoires sont pris en compte dans la fonction de coût, ce qui constitue une mise en œuvre novatrice du problème de retour dans le domaine d'application abordé. En se concentrant sur les variables de décision continues, le modèle proposé comble cette lacune en proposant une approche robuste pour le calcul des fréquences d'inspection qui équilibre les coûts et la fiabilité découlant des états de fonctionnement-maintenance du système, comme le soulignent les résultats.

7.3 Formulation mathématique du problème

Soit $X(t)$ l'état d'une machine au temps t. On suppose que le processus stochastique $\{X(t),\ t \geq t_0\}$ est une chaîne de Markov en temps continu ayant les états suivants :

0 : la machine est en fonctionnement,
1 : la machine fait l'objet d'une maintenance de routine,
2 : la machine fait l'objet d'une maintenance prolongée,
3 : la machine est en cours de réparation.

Le diagramme de transition de la chaîne de Markov est présenté dans la figure 7.1. Cette figure offre une illustration visuelle des transitions d'état pendant le cycle de vie opérationnel de la machine, qui est modélisé comme une chaîne de Markov en temps continu. Les flèches indiquent les transitions possibles entre les états définis, chacune avec des probabilités de transition spécifiées. La machine peut suivre différents chemins au cours de son fonctionnement dans l'état « En fonctionnement ». Elle peut passer à l'état « Maintenance de routine » (avec une probabilité de transition $p_{0,1}$), ou à l'état « Maintenance prolongée » (avec une probabilité de transition $p_{0,2}$), ou à l'état « Réparation » (avec une probabilité de transition $p_{0,3}$). Chacune de ces transitions est déclenchée par des besoins particuliers en maintenance ou par des événements de défaillance. À la fin de la « Maintenance de routine » ou de la « Maintenance prolongée », la machine revient à l'état « En fonctionnement », comme indiqué par les transitions $p_{1,0} = 1$ et $p_{2,0} = 1$, respectivement. Cela indique qu'une fois la maintenance terminée, la machine fonctionne à nouveau à pleine capacité (comme neuve). De même, la transition $p_{3,0} = 1$ indique qu'après une « Réparation », la machine retourne à l'état « En fonctionnement », retrouvant ainsi toute sa fonctionnalité (comme si elle était neuve).

Le temps passé par la chaîne de Markov dans l'état i est une variable aléatoire exponentielle T_i de paramètre ν_i, pour $i = 0, 1, 2, 3$; de plus, l'inspection est effectuée à des instants aléatoires. On suppose que le temps écoulé entre deux opérations d'inspection est une variable aléatoire

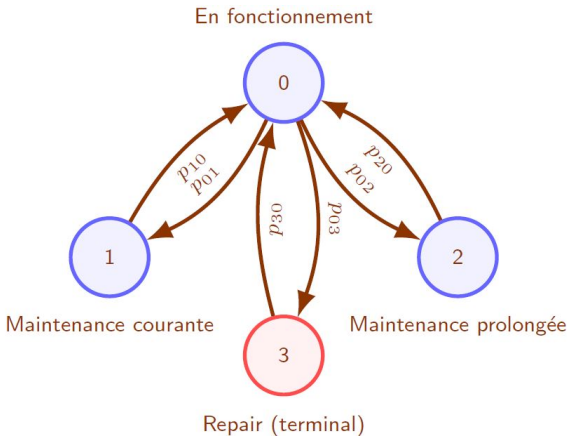

Figure 7.1 Diagramme de transition de la chaîne de Markov.

exponentielle de paramètre $u > 0$. Lors de l'entretien de la machine, il y a une probabilité égale à p qu'aucun problème sérieux ne soit détecté (et une probabilité égale à $1 - p$ qu'au moins un problème important soit découvert).

Soit $p_{i,j}$ la probabilité que la chaîne de Markov passe à l'état j lorsqu'elle quitte l'état i, pour $i \neq j \in \{0, 1, 2, 3\}$. On peut écrire (voir, par exemple, Lefebvre [2]) que $p_{j,0} = 1$ pour $j = 1, 2, 3$ et

$$p_{0,1} = \frac{p\,u}{u + \nu_0}, \quad p_{0,2} = \frac{(1-p)\,u}{u + \nu_0} \quad \text{et} \quad p_{0,3} = \frac{\nu_0}{u + \nu_0}. \tag{7.1}$$

Explication.

À l'état 0 (représentant l'état de fonctionnement), on observe la présence de deux déclencheurs de transition indépendants :

1. un mécanisme de défaillance agissant avec un taux ν_0,
2. un mécanisme d'inspection agissant avec un taux u.

Dans un tel cadre, le temps jusqu'au prochain événement (défaillance ou inspection) est gouverné par le minimum de deux processus exponentiels, ce qui donne un nouveau temps de séjour exponentiel de paramètre $u + \nu_0$, qui définit le taux réel auquel le système quitte l'état 0.

Lors du départ de l'état 0, la destination est sélectionnée en fonction de l'événement qui s'est produit en premier :

— Avec une probabilité $\nu_0/(u + \nu_0)$, une défaillance survient et le système passe à l'état 3 (réparation),

— Avec une probabilité $u/(u + \nu_0)$, une inspection est effectuée. L'issue de l'inspection

détermine :

$$\begin{cases} \text{état } 1 & \text{avec une probabilité } p, \\ \text{état } 2 & \text{avec une probabilité } 1-p. \end{cases} \tag{7.2}$$

Par conséquent, les probabilités de transition totales depuis l'état 0 sont celles données dans l'équation (7.1).

Ensuite, le *temps de premier passage* est défini

$$\tau(i) = \inf\{t \geq t_0 : X(t) = 3 \text{ ou } t = t_1 \ (> t_0) \mid X(t_0) = i\} \tag{7.3}$$

pour $i = 0, 1, 2$ (avec $\tau(3) = t_0$).

Interprétation de l'équation (7.3) **: Instant de premier passage**

L'équation (7.3) introduit le concept du *temps de premier passage* $\tau(i)$ pour le processus de Markov en temps continu $\{X(t), t \geq t_0\}$, commençant dans l'état $i \in \{0, 1, 2\}$, où $t_1 > t_0$ est un horizon temporel fini fixé, et $X(t)$ désigne l'état du système au temps t.

Cette expression représente le premier instant auquel :

— le processus atteint l'état 3, correspondant à un scénario de réparation, ou

— le temps final prédéterminé t_1 est atteint,

selon celui qui survient en premier. Ainsi, $\tau(i)$ définit le moment de terminaison du problème de commande optimale.

La définition du temps de premier passage est essentielle dans la formulation de la fonction de coût à l'équation (7.4), puisqu'elle délimite l'intervalle d'intégration sur lequel les coûts d'inspection et de maintenance sont accumulés.

On suppose que l'optimiseur peut choisir à tout instant une valeur quelconque de la constante $u \in U$ (le taux d'inspection), de sorte que $u = u(t)$ soit la variable de contrôle et que $p_{i,j}$ devienne $p_{i,j}(t)$. On cherche la valeur $u^*(t) \geq 0$ de $u(t)$ qui minimise la valeur espérée du critère de coût :

$$J(i) = \int_{t_0}^{\tau(i)} \{f[u(t)] - \lambda\} \, \mathrm{d}t + K \, I_F, \tag{7.4}$$

où $f(0) = 0$, $f[u(t)] > 0$ pour $u(t) > 0$, $\lambda > 0$, $K > 0$, et I_F est la fonction indicatrice de l'événement F : l'optimiseur choisit $u(t) \equiv 0$. Le but est donc de maximiser le temps espéré avant que la machine n'ait besoin d'être réparée, tout en tenant compte des coûts de contrôle $f[u(t)]$.

Explication. Interprétation de la fonction indicatrice I_F :

Dans la fonction de coût définie par l'équation (7.4), le terme KI_F introduit une pénalité fixe appliquée sous une stratégie de contrôle particulière. Le symbole I_F désigne la fonction indicatrice de l'événement F, où :

$$F: \quad u(t) \equiv 0 \quad \text{pour tout } t \in [t_0, \tau(i)]. \tag{7.5}$$

En d'autres termes, l'événement F correspond à une politique de contrôle dégénérée où l'optimiseur ne procède à aucune inspection ni maintenance préventive à aucun moment. La fonction indicatrice prend la valeur :

$$I_F = \begin{cases} 1, & \text{si } u(t) \equiv 0, \\ 0, & \text{sinon.} \end{cases} \tag{7.6}$$

En incluant le terme KI_F dans la fonction de coût, une pénalité est introduite par le modèle pour toute stratégie omettant complètement l'inspection et la maintenance. En appliquant ce terme, on s'assure que de telles politiques de non-intervention ne sont choisies que si elles représentent réellement le compromis optimal entre fiabilité et coûts. Le coût fixe de pénalité $K > 0$ agit comme un désincitatif économique à éviter les actions préventives, et met l'accent sur l'importance de maintenir la fonctionnalité du système pendant l'horizon de planification.

Remarques. (i) Le coût final K est imposé si l'optimiseur décide de ne pas effectuer de maintenance, afin de refléter le fait que lorsque la machine doit être réparée (état 3), les coûts de réparation sont plus élevés.

(ii) Le problème ci-dessus constitue un cas particulier de "homing problem", dans lequel l'optimiseur contrôle un processus stochastique jusqu'à ce qu'un certain événement survienne ; voir Whittle [92] et [103]. Les auteurs ont récemment étudié des problèmes de retour pour des systèmes de files d'attente ; voir Lefebvre et Yaghoubi [96] et [91].

(iii) À la connaissance de l'auteur, c'est la première fois qu'un problème de retour est traité pour une chaîne de Markov en temps continu qui n'est pas un modèle de file d'attente. De nombreux articles ont été publiés dans le domaine de la fiabilité, dans lesquels des chaînes de Markov en temps continu ont été utilisées comme modèles. De plus, comme l'a montré la revue de la littérature, les problèmes de commande optimale pour ces modèles ont également fait l'objet de nombreuses publications. Dans ce problème, plutôt que de contrôler un processus stochastique jusqu'à un instant final fixé ou infini, le contrôle du processus cesse à l'instant où la machine doit être réparée (ou lorsqu'un certain temps s'est écoulé). En réalité, cet instant est effectivement une variable aléatoire.

(iv) Pour valider le modèle formulé, on a appliqué une série d'actions visant à évaluer sa robustesse et sa fiabilité. En premier lieu, une analyse de sensibilité est effectuée pour examiner l'effet de la variation des paramètres clés sur les résultats du modèle. Cette analyse a permis de vérifier que le modèle reste fiable et produit des résultats cohérents dans une certaine plage de scénarios hypothétiques. En outre, on a analysé la réponse du modèle à plusieurs conditions et paramètres initiaux afin de simuler différents comportements du système. Ces étapes de vérification démontrent l'utilité du modèle pour la prise de décision en matière de maintenance.

La *fonction de valeur* est maintenant définie par

$$F(i, t_0) = \inf_{\substack{u(t) \\ t \in [t_0, \tau(i))}} E[J(i)] \tag{7.7}$$

pour $i = 0, 1, 2, 3$. De plus, on peut écrire que $F(3, t_0) = 0$.

Explication. Dans l'équation (7.7), la fonction de valeur $F(i, t_0)$ désigne l'infimum de la fonction de coût espérée $J(i)$ sur l'ensemble des fonctions de contrôle admissibles $u(t)$ définies sur l'intervalle $[t_0, \tau(i))$, à partir de l'état i au temps t_0.

Lorsque le processus commence dans l'état 3, il est déjà en état de réparation, ce qui est considéré comme un état terminal dans le modèle. Par définition, le temps de premier passage dans ce cas est :

$$\tau(3) = t_0. \tag{7.8}$$

En substituant cela dans la fonction de coût définie à l'équation (7.4), on observe que le terme intégral s'annule puisque l'intervalle d'intégration est de longueur nulle :

$$\int_{t_0}^{\tau(3)} \{f[u(t)] - \lambda\}\, dt = \int_{t_0}^{t_0} (\cdot)\, dt = 0. \tag{7.9}$$

Par ailleurs, aucune action de contrôle n'est exercée au-delà de ce point, et le coût terminal n'est pas engagé (ou est implicitement nul lorsqu'on commence en état 3), le coût total espéré est donc :

$$J(3) = 0 \quad \Rightarrow \quad F(3, t_0) = \inf_{u(t)} E[J(3)] = 0. \tag{7.10}$$

Cela reflète le fait qu'une fois le système en état de réparation, la période de contrôle se termine immédiatement et aucun coût supplémentaire n'est accumulé.

Il en résulte que

$$F(i,t_0) = -\lambda E[T_i] + F(0,t_0) = -\frac{\lambda}{\nu_i} + F(0,t_0) \quad \text{pour } i = 1,2. \tag{7.11}$$

Explication. Interprétation de l'équation (7.11) :

L'équation (7.11) exprime la fonction de valeur $F(i,t_0)$ lorsque le système commence dans l'état $i \in \{1,2\}$, ce qui correspond aux états de maintenance de routine et de maintenance prolongée, respectivement.

Tant que le processus demeure dans l'état 1 ou 2, le taux d'inspection est supposé nul, c'est-à-dire $u(t) = 0$. En conséquence, la fonction de coût se réduit au terme $-\lambda$, et aucun coût lié à l'inspection n'est encouru, ce qui signifie que la fonction de coût de contrôle satisfait $f[0] = 0$, donc seul le terme $-\lambda$ contribue à la fonction de coût.

Le temps de séjour dans l'état i suit une loi exponentielle de taux ν_i, donc le temps moyen passé est :

$$\mathbb{E}[T_i] = \frac{1}{\nu_i}. \tag{7.12}$$

Pendant cet intervalle, la seule contribution à la fonction de coût est le terme $-\lambda$. Par conséquent, le coût accumulé espéré dans l'état i est :

$$\int_{t_0}^{t_0+T_i} (-\lambda)\, dt = -\lambda \cdot \mathbb{E}[T_i] = -\frac{\lambda}{\nu_i}. \tag{7.13}$$

Une fois la maintenance terminée (état 1 ou 2), le système passe de manière déterministe à l'état de fonctionnement (état 0), d'où le coût espéré restant est $F(0,t_0)$. En combinant les deux, on obtient :

$$F(i,t_0) = -\frac{\lambda}{\nu_i} + F(0,t_0), \quad \text{pour } i = 1,2. \tag{7.14}$$

Remarque. Notons $F_0(0,t_0)$ la valeur espérée de $J(0)$ si l'optimiseur n'utilise aucun contrôle, donc $u(t) \equiv 0$. Étant donné que $f(0) = 0$, on peut écrire :

$$F_0(0,t_0) = -\lambda E[T_0] + K = -\frac{\lambda}{\nu_0} + K. \tag{7.15}$$

Ainsi, la fonction de valeur $F(0,t_0)$ doit nécessairement être inférieure ou égale à $-\frac{\lambda}{\nu_0} + K$.

La section suivante dérive l'équation de programmation dynamique satisfaite par la fonction $F(0,t_0)$. Ensuite, dans la section 4, un cas particulier du problème ci-dessus sera résolu explicitement. La section 5 traite le cas où la constante t_1 dans la définition de $\tau(i)$ tend vers

l'infini. La section 6 conclut l'étude par quelques remarques.

7.4 Équation de programmation dynamique

Supposons que la machine soit en fonctionnement au temps t_0, donc $X(t_0) = 0$. On a :

$$P[T_0 < \Delta t] = 1 - e^{-\nu_0 \Delta t} = \nu_0 \Delta t + o(\Delta t). \tag{7.16}$$

Explication.

Soit T_0 la variable aléatoire représentant le temps de séjour dans l'état 0, supposée suivre une loi exponentielle de paramètre ν_0. Autrement dit, $T_0 \sim \text{Exp}(\nu_0)$.

Alors, la probabilité que la machine quitte l'état 0 pendant un court intervalle de temps Δt est donnée par :

$$P[T_0 < \Delta t] = 1 - e^{-\nu_0 \Delta t}. \tag{7.17}$$

En utilisant un développement de Taylor d'ordre un de la fonction exponentielle autour de $\Delta t = 0$, on obtient :

$$1 - e^{-\nu_0 \Delta t} = \nu_0 \Delta t + o(\Delta t) \tag{7.18}$$

où $o(\Delta t)$ désigne les infinitésimaux d'ordre supérieur satisfaisant $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$.

Ainsi, pour Δt suffisamment petit, la probabilité qu'une transition se produise dans l'intervalle $[t_0, t_0 + \Delta t]$ est approximativement linéaire en Δt, et l'on écrit :

$$P[T_0 < \Delta t] = \nu_0 \, \Delta t + o(\Delta t). \tag{7.19}$$

Il en résulte que l'on peut écrire :

$$X(t_0 + \Delta t) = \begin{cases} 1 & \text{avec une probabilité } \nu_0 \Delta t p_{0,1}(t_0) \; [+o(\Delta t)], \\ 2 & \text{avec une probabilité } \nu_0 \Delta t p_{0,2}(t_0), \\ 3 & \text{avec une probabilité } \nu_0 \Delta t p_{0,3}(t_0), \\ 0 & \text{avec une probabilité } 1 - \nu_0 \Delta t. \end{cases} \tag{7.20}$$

Explication.

L'équation (7.20) décrit le comportement probabiliste du processus $X(t)$ immédiatement après le temps t_0, en supposant que le système est en état de fonctionnement au temps t_0, c'est-à-dire $X(t_0) = 0$.

Pour un petit incrément de temps Δt, la probabilité de quitter l'état 0 est approximativement $\nu_0 \Delta t + o(\Delta t)$, comme indiqué dans l'équation (7.19), où ν_0 est le taux de transition exponentielle hors de l'état 0. Les probabilités de transition vers les autres états sont pondérées par les probabilités conditionnelles $p_{0,j}(t_0)$ pour $j = 1, 2, 3$. Ainsi, les probabilités totales de transition sur l'intervalle $[t_0, t_0 + \Delta t]$ sont données par l'équation (7.20).

Remarque. Le fait que les différentes probabilités dans l'équation ci-dessus soient toutes proportionnelles à Δt est essentiel pour pouvoir utiliser la programmation dynamique.

Soit

$$g[u(t)] := f[u(t)] - \lambda. \tag{7.21}$$

On a

$$\begin{aligned} F(0, t_0) &= \inf_{\substack{u(t) \\ t \in [t_0, \tau(0))}} E\left[\int_{t_0}^{t_0+\Delta t} g[u(t)]\,\mathrm{d}t + \int_{t_0+\Delta t}^{\tau(0)} g[u(t)]\,\mathrm{d}t\right] \\ &= \inf_{\substack{u(t) \\ t \in [t_0, \tau(0))}} E\left\{g[u(t_0)]\,\Delta t + \int_{t_0+\Delta t}^{\tau(0)} g[u(t)]\,\mathrm{d}t + o(\Delta t)\right\}. \end{aligned} \tag{7.22}$$

Explication.

Pour simplifier la dérivation de la programmation dynamique, on définit : $g[u(t)] := f[u(t)] - \lambda$, et la fonction de valeur $F(0, t_0)$ représente le coût espéré minimal à partir de l'état de fonctionnement (état 0) au temps t_0, tel qu'exprimé dans l'équation (7.22).

Dans cette décomposition :

- Le premier terme intégral rend compte du coût accumulé pendant le petit intervalle $[t_0, t_0 + \Delta t]$.
- Le second terme intégral représente le coût restant à partir du temps $t_0 + \Delta t$ jusqu'au temps d'arrêt $\tau(0)$.
- Le terme $o(\Delta t)$ prend en compte les contributions infinitésimales d'ordre supérieur qui tendent vers zéro plus rapidement que Δt.

De plus,

$$E\left[\int_{t_0+\Delta t}^{\tau(0)} g[u(t)]\,\mathrm{d}t\right] = E\left[E\left[\int_{t_0+\Delta t}^{\tau(0)} g[u(t)]\,\mathrm{d}t \;\middle|\; X(t_0 + \Delta t)\right]\right]. \tag{7.23}$$

Explication.

Évaluation conditionnelle et décomposition par état :

L'équation (7.23) applique la *loi de l'espérance totale* pour réécrire le coût futur espéré après le court intervalle $[t_0, t_0 + \Delta t]$.

À ce stade, on évalue le coût espéré restant à partir de $t_0 + \Delta t$ jusqu'au temps terminal $\tau(0)$. Ce coût dépend de l'état aléatoire $X(t_0 + \Delta t)$ dans lequel le système se retrouve après ce court intervalle.

Au temps $t_0 + \Delta t$, le système peut passer dans l'un des états $\{0, 1, 2, 3\}$, avec des probabilités déterminées précédemment dans l'équation (7.20). Puisque le coût futur dépend de l'état atteint, on applique la loi de l'espérance totale pour décomposer le coût attendu global en composants conditionnels, comme exprimé dans l'équation (7.23).

Ainsi, en appliquant le principe d'optimalité de Bellman, on peut écrire :

$$\inf_{\substack{u(t) \\ t \in [t_0 + \Delta t, \tau(0))}} E\left[\int_{t_0+\Delta t}^{\tau(0)} g[u(t)]\,\mathrm{d}t\right] = E\Big[F\left[X(t_0 + \Delta t), t_0 + \Delta t\right]\Big]. \tag{7.24}$$

Explication.

Dans cette étape, on a appliqué le *principe d'optimalité de Bellman*, qui stipule qu'une politique de commande optimale doit rester optimale à chaque étape du processus de décision, à partir de l'état que le système atteint au moment suivant, et ce, indépendamment de l'état initial et des décisions de contrôle initiales.

Dans ce contexte, une fois que le système passe à un nouvel état au temps $t_0 + \Delta t$, le coût espéré minimal encouru à partir de cet instant, optimisé sur l'ensemble des fonctions de contrôle admissibles $u(t)$, est donné par la fonction de valeur évaluée au nouvel état et au nouveau temps. Par conséquent, le coût de continuation à partir de $t_0+\Delta t$ dépend uniquement du prochain état du système $X(t_0 + \Delta t)$ et de la stratégie de contrôle appliquée à partir de ce moment. Ainsi, le coût espéré restant, minimisé sur la trajectoire de contrôle $u(t)$ définie sur l'intervalle $[t_0 + \Delta t, \tau(0))$, peut être exprimé comme indiqué dans l'équation (7.24).

Par ailleurs, on déduit de l'équation (7.20) que [puisque $X(t_0) = 0$]

$$\begin{aligned} E\Big[F\left[X(t_0 + \Delta t), t_0 + \Delta t\right]\Big] \;&=\; \sum_{j=1}^{2} F(j, t_0 + \Delta t)\,\nu_0 \Delta t\, p_{0,j}(t_0) \\ &\quad + F(0, t_0 + \Delta t)\,(1 - \nu_0 \Delta t). \end{aligned} \tag{7.25}$$

Explication. Évaluation de la fonction de valeur espérée :

L'espérance de la fonction $F(X(t_0 + \Delta t), t_0 + \Delta t)$ est maintenant calculée, qui représente

le coût futur espéré à partir de l'état atteint au temps $t_0 + \Delta t$. Cette formulation découle de la loi de l'espérance totale, où les états possibles sont pondérés par leur probabilité de transition.

Puisque le système est initialement dans l'état $X(t_0) = 0$, les probabilités de transition vers les autres états sur l'intervalle Δt sont données par l'équation (7.20).

Pour calculer la valeur espérée de la fonction $F(X(t_0 + \Delta t), t_0 + \Delta t)$, on effectue une somme pondérée de $F(j, t_0 + \Delta t)$ sur tous les états j, chacun étant multiplié par sa probabilité d'occurrence sur l'intervalle Δt.

Dans cette expression :

- Les termes pour $j = 1$ et $j = 2$, correspondant aux transitions vers les états 1 et 2, apparaissent sous la forme $\nu_0 \Delta t \cdot p_{0,j}(t_0)$, où ν_0 est le taux de sortie de l'état 0 et $p_{0,j}(t_0)$ est la probabilité de transition.
- Le terme correspondant à rester dans l'état 0 est approximé par la probabilité $1-\nu_0 \Delta t$, ce qui reflète le fait qu'aucune transition ne se produit pendant Δt.
- La transition vers l'état 3 est exclue du calcul car la fonction de valeur dans cet état est définie comme nulle, i.e., $F(3, t) = 0$. Cela reflète le fait qu'entrer dans l'état 3 correspond à une défaillance du système et à la fin immédiate du processus de contrôle, sans coûts supplémentaires.

En supposant maintenant que la fonction $F(0, t_0)$ est dérivable par rapport à t_0, il découle du théorème de Taylor que

$$F(\cdot, t_0 + \Delta t) = F(\cdot, t_0) + \Delta t F'(\cdot, t_0) + o(\Delta t). \tag{7.26}$$

Explication. Développement temporel via le théorème de Taylor :

On suppose que, pour chaque état, la fonction de valeur $F(i, t)$ est dérivable par rapport au temps t. $F'(\cdot, t_0)$ désigne la dérivée partielle de F par rapport au temps, évaluée en t_0, et le terme $o(\Delta t)$ comprend les termes d'ordre supérieur qui tendent plus rapidement vers zéro que Δt lorsque $\Delta t \to 0$.

Ce développement permet de remplacer les valeurs futures de la fonction F au temps $t_0 + \Delta t$ par une approximation en fonction des valeurs connues et des dérivées temporelles évaluées au temps actuel t_0.

Il s'ensuit que

$$F(\cdot, t_0 + \Delta t)(1 - \nu_0 \Delta t) = (1 - \nu_0 \Delta t) F(\cdot, t_0) + \Delta t F'(\cdot, t_0) + o(\Delta t). \tag{7.27}$$

Explication.

À ce stade, on multiplie le développement de Taylor de $F(\cdot, t_0 + \Delta t)$ par le terme $1 - \nu_0 \Delta t$, qui représente la probabilité que le système reste dans l'état 0 pendant l'intervalle de temps Δt. Cette étape permet de simplifier l'expression de l'espérance formulée précédemment dans l'équation (7.25).

En utilisant les résultats précédents, on obtient :

$$\begin{aligned} 0 \;=\; \inf_{u(t_0)} \Big\{ & g[u(t_0)]\,\Delta t + \nu_0\,\Delta t \sum_{j=1}^{2} p_{0,j}(t_0)\,F(j,t_0) \\ & - \nu_0\,\Delta t\,F(0,t_0) + \Delta t F'(0,t_0) + o(\Delta t) \Big\}. \end{aligned} \tag{7.28}$$

Explication.

On part de la fonction de valeur à l'instant t_0 dans l'état 0 :

$$F(0,t_0) = \inf_{u(t)} E\left[\int_{t_0}^{t_0+\Delta t} g[u(t)]\,dt + F(X(t_0+\Delta t), t_0+\Delta t) \right], \tag{7.29}$$

où $g[u(t)] = f[u(t)] - \lambda$ est le coût instantané.

Le premier terme est approché par un développement du premier ordre :

$$\int_{t_0}^{t_0+\Delta t} g[u(t)]\,dt \approx g[u(t_0)]\Delta t. \tag{7.30}$$

Le second terme est évalué à l'aide de la loi de l'espérance totale, comme à l'équation (7.25) :

$$E[F(X(t_0+\Delta t), t_0+\Delta t)] = \sum_{j=1}^{2} F(j, t_0+\Delta t)\nu_0\Delta t\,p_{0,j}(t_0) + F(0, t_0+\Delta t)(1-\nu_0\Delta t). \tag{7.31}$$

On développe la fonction de valeur dans le temps via le théorème de Taylor :

$$F(0, t_0+\Delta t) = F(0,t_0) + \Delta t\,F'(0,t_0) + o(\Delta t), \tag{7.32}$$

et de même, $F(j, t_0+\Delta t) = F(j,t_0) + o(\Delta t)$ pour $j = 1, 2$.

En remplaçant (7.30) et (7.31) dans (7.29), on obtient :

$$F(0,t_0) = \inf_{u(t_0)} \Big\{ g[u(t_0)]\Delta t + \nu_0 \Delta t \sum_{j=1}^{2} p_{0,j}(t_0) F(j,t_0) \\ + (1-\nu_0\Delta t)\left[F(0,t_0) + \Delta t\, F'(0,t_0)\right] + o(\Delta t) \Big\}. \tag{7.33}$$

En soustrayant $F(0,t_0)$ des deux côtés et en réarrangeant, on obtient l'équation (7.28).

Par ailleurs, on déduit de l'équation (7.11) que

$$0 = \inf_{u(t_0)} \Big\{ g[u(t_0)]\,\Delta t + \nu_0\,\Delta t \sum_{j=1}^{2} p_{0,j}(t_0) \left\{ -\frac{\lambda}{\nu_j} + F(0,t_0) \right\} \\ - \nu_0\,\Delta t\, F(0,t_0) + \Delta t F'(0,t_0) + o(\Delta t) \Big\}. \tag{7.34}$$

Finalement, en divisant chaque côté de l'équation précédente par Δt et en prenant la limite lorsque Δt tend vers zéro, on peut énoncer la proposition suivante.

Proposition 1. *La fonction de valeur $F(0,t_0)$ satisfait l'équation de programmation dynamique (EPD)*

$$0 = \inf_{u(t_0)} \left\{ g[u(t_0)] + \nu_0 \sum_{j=1}^{2} p_{0,j}(t_0) \left[-\frac{\lambda}{\nu_j} + F(0,t_0) \right] - \nu_0\, F(0,t_0) + F'(0,t_0) \right\}. \tag{7.35}$$

De plus, on a la condition limite suivante : $F(0,t_1) = 0$ si $u(t_0) > 0$.

Posons $u_0 := u(t_0)$, et comme mentionné dans l'équation (7.1), on a :

$$p_{0,1}(t_0) = \frac{p\,u_0}{u_0 + \nu_0} \quad \text{et} \quad p_{0,2}(t_0) = \frac{(1-p)\,u_0}{u_0 + \nu_0}. \tag{7.36}$$

En outre, la fonction $f(\cdot)$ est souvent choisie sous la forme :

$$f[u(t)] = \frac{1}{2} q_0\, u^2(t), \tag{7.37}$$

où q_0 est une constante strictement positive. Alors, si $u(t)$ est une fonction continue de t, on peut obtenir la commande optimale u_0^* en fonction de $F(0,t_0,t_1)$ en dérivant l'expression entre accolades dans l'équation (7.35) par rapport à u_0 et en résolvant l'équation obtenue pour u_0.

Si l'on parvient à déterminer une expression explicite de u_0^*, on la remplace dans l'équation (7.35) et l'on tente de résoudre l'équation différentielle résultante, sous la condition au bord $F(0, t_1) = 0$.

La section suivante considère le cas le plus simple possible, à savoir celui où $u(t) \equiv k_1$ ou k_2.

7.5 Cas particulier

Supposons que l'ensemble U ne contienne que deux valeurs, notées k_1 et k_2. Il découle de l'équation (7.35) que l'équation suivante doit être considérée :

$$\begin{aligned} 0 = g(k_i) &+ \nu_0 \left(\frac{p k_i}{k_i + \nu_0} \right) \left(-\frac{\lambda}{\nu_1} + F(0, t_0) \right) \\ &+ \nu_0 \left(\frac{(1-p) k_i}{k_i + \nu_0} \right) \left(-\frac{\lambda}{\nu_2} + F(0, t_0) \right) - \nu_0 F(0, t_0) + F'(0, t_0) \end{aligned} \tag{7.38}$$

pour $i = 1, 2$.

Choisissons la fonction $f[u(t)]$ définie dans l'équation (7.37) avec $q_0 = 2$. De plus, on prend $\lambda = \nu_i = 1$ pour $i = 0, 1, 2, 3$, $p = 1/2$, $k_1 = 1$ et $k_2 = 0.9$.

Explication.

Dans ce cas particulier, les paramètres du modèle sont choisis afin de faciliter la résolution analytique de l'équation de programmation dynamique. En fixant le coefficient de coût $q_0 = 2$ dans l'équation (7.37), on obtient une fonction de coût quadratique de la forme $f[u(t)] = u^2(t)$, indiquant une augmentation proportionnelle au carré du taux d'inspection.

Les valeurs $\nu_0 = \nu_1 = \nu_2 = \nu_3 = 1$ représentent les taux de transition à partir de chaque état défini dans le processus de Markov. Leur affectation à 1 implique que la durée moyenne de séjour dans chaque état est d'une unité de temps.

Le facteur $\lambda = 1$ reflète l'importance relative accordée aux coûts futurs dans la fonction de coût.

Une probabilité égale de transition vers l'état 1 ou 2 après une inspection est donnée par le paramètre $p = \frac{1}{2}$, ce qui reflète une incertitude symétrique concernant le comportement de dégradation du système après inspection.

Enfin, l'ensemble de contrôle $U = \{k_1, k_2\}$ est restreint à deux taux d'inspection constants. Ici, $k_1 = 1$ correspond à une fréquence d'inspection élevée, tandis que $k_2 = 0.9$ représente une intensité d'inspection réduite. Ces valeurs fixes permettent une évaluation comparative directe entre deux politiques d'inspection définies dans le même cadre de modélisation.

Avec ces valeurs, il apparaît nécessaire de résoudre des équations différentielles linéaires du premier ordre (en notant la fonction $F(0,t_0)$ par $F_i(t_0)$ si $u_0 = k_i$, pour $i = 1, 2$)

$$1 = -F_1(t_0) + 2\,F_1'(t_0) \tag{7.39}$$

et

$$0 = 639 + 1000\,F_2(t_0) + 1900\,F_2'(t_0). \tag{7.40}$$

Explication.

Les équations différentielles (7.39) et (7.40) résultent de la substitution des taux d'inspection constants $u_0 = k_1 = 1$ et $u_0 = k_2 = 0.9$ dans l'équation de programmation dynamique (EPD), telle que présentée précédemment :

$$\begin{aligned} 0 = {} & g(k_i) + \nu_0 \left(\frac{p k_i}{k_i + \nu_0}\right) \left(-\frac{\lambda}{\nu_1} + F(0,t_0)\right) \\ & + \nu_0 \left(\frac{(1-p)\,k_i}{k_i + \nu_0}\right) \left(-\frac{\lambda}{\nu_2} + F(0,t_0)\right) - \nu_0 F(0,t_0) + F'(0,t_0) \end{aligned} \tag{7.41}$$

La fonction de coût de contrôle est donnée par :

$$g[u_0] = f[u_0] - \lambda = u_0^2 - 1 \tag{7.42}$$

Les probabilités de transition deviennent :

$$p_{0,1}(t_0) = \frac{p u_0}{u_0 + \nu_0}, \quad p_{0,2}(t_0) = \frac{(1-p)u_0}{u_0 + \nu_0} \tag{7.43}$$

En substituant les valeurs des paramètres : $p = \frac{1}{2}, \quad \nu_0 = \nu_1 = \nu_2 = 1, \quad \lambda = 1$

L'équation (7.39) est dérivée comme suit : pour $u_0 = k_1 = 1$

D'après l'équation (7.42) :

$$g[1] = 1^2 - 1 = 0 \tag{7.44}$$

D'après l'équation (7.43), on obtient :

$$p_{0,1} = \frac{1/2 \cdot 1}{1+1} = \frac{1}{4}, \quad p_{0,2} = \frac{1/2 \cdot 1}{1+1} = \frac{1}{4} \tag{7.45}$$

En substituant dans l'équation (7.41), on obtient :

$$0 = 0 + 1\left[\frac{1}{4}\left(-1 + F_1(t_0)\right) + \frac{1}{4}\left(-1 + F_1(t_0)\right)\right] - 1 \cdot F_1(t_0) + F_1'(t_0) \tag{7.46}$$

En simplifiant :

$$0 = -\frac{1}{2} - \frac{1}{2}F_1(t_0) + F_1'(t_0) \Rightarrow 1 = -F_1(t_0) + 2F_1'(t_0) \tag{7.47}$$

L'équation (7.40) est dérivée comme suit : pour $u_0 = k_2 = 0.9$

On calcule maintenant :

$$g[0.9] = (0.9)^2 - 1 = 0.81 - 1 = -0.19 \tag{7.48}$$

D'après l'équation (7.43), on obtient :

$$p_{0,1} = \frac{1/2 \cdot 0.9}{0.9 + 1} = \frac{0.45}{1.9}, \quad p_{0,2} = \frac{0.45}{1.9} \tag{7.49}$$

En substituant tout dans l'équation (7.41), on a :

$$0 = -0.19 + 1\left[\frac{0.45}{1.9}(-1 + F_2(t_0)) + \frac{0.45}{1.9}(-1 + F_2(t_0))\right] - F_2(t_0) + F_2'(t_0) \tag{7.50}$$

Ainsi :

$$0 = -\frac{19}{100} + \frac{9}{19}(-1 + F_2(t_0)) - F_2(t_0) + F_2'(t_0) \tag{7.51}$$

On multiplie les deux membres par 1900 pour éliminer les fractions, et après simplification, on obtient le résultat indiqué dans les équations (7.40) : $0 = 639 + 1000\ F_2(t_0) + 1900\ F_2'(t_0)$.

Les solutions des équations ci-dessus qui satisfont la condition aux limites $F_i(t_1) = 0$, pour $i = 1, 2$, sont

$$F_1(t_0) = -1 + e^{(t_0 - t_1)/2} \tag{7.52}$$

et

$$F_2(t_0) = \frac{639}{1000}\left\{1 - e^{10(t_1 - t_0)/19}\right\}. \tag{7.53}$$

Explication.

Les équations différentielles (7.39) et (7.40) sont maintenant dérivées, sous la condition aux limites $F_i(t_1) = 0$, pour $i = 1, 2$.

Solution de l'équation (7.39) :

$$1 = -F_1(t_0) + 2F_1'(t_0) \quad \Longrightarrow \quad F_1'(t_0) - \frac{1}{2}F_1(t_0) = \frac{1}{2} \tag{7.54}$$

En utilisant le facteur intégrant $\mu(t_0) = e^{-t_0/2}$ pour cette EDO linéaire du premier ordre, on obtient :

$$\frac{d}{dt_0}\left(e^{-t_0/2}F_1(t_0)\right) = \frac{1}{2}e^{-t_0/2} \tag{7.55}$$

En intégrant les deux membres :

$$e^{-t_0/2}F_1(t_0) = -e^{-t_0/2} + C \quad \Longrightarrow \quad F_1(t_0) = -1 + Ce^{t_0/2} \tag{7.56}$$

En appliquant la condition aux limites $F_1(t_1) = 0$, on obtient $C = e^{-t_1/2}$, donc le résultat est identique à celui présenté dans l'équation (7.52) : $F_1(t_0) = -1 + e^{(t_0-t_1)/2}$.

Solution de l'équation (7.40)

$$0 = 639 + 1000F_2(t_0) + 1900F_2'(t_0) \quad \Longrightarrow \quad F_2'(t_0) + \frac{1000}{1900}F_2(t_0) = -\frac{639}{1900} \tag{7.57}$$

En posant le facteur intégrant $\mu(t_0) = e^{\frac{10}{19}t_0}$, on écrit :

$$\frac{d}{dt_0}\left(e^{\frac{10}{19}t_0}F_2(t_0)\right) = -\frac{639}{1900}e^{\frac{10}{19}t_0} \tag{7.58}$$

L'intégration des deux membres donne :

$$e^{\frac{10}{19}t_0}F_2(t_0) = -\frac{639}{1000}e^{\frac{10}{19}t_0} + C \quad \Rightarrow \quad F_2(t_0) = -\frac{639}{1000} + Ce^{-\frac{10}{19}t_0} \tag{7.59}$$

En utilisant la condition limite $F_2(t_1) = 0$, on obtient $C = \frac{639}{1000}e^{\frac{10}{19}t_1}$, ce qui donne le même résultat que celui indiqué dans l'équation (7.53) : $F_2(t_0) = \frac{639}{1000}\left(1 - e^{\frac{10}{19}(t_1-t_0)}\right)$.

Les deux fonctions sont représentées dans la figure 7.2 dans le cas où $t_1 = 2$. On observe que la solution optimale dans cet exemple est de choisir

$$u_0 = \begin{cases} 0.9 & \text{si } 0 \le t_0 \le t^*, \\ 1 & \text{si } t^* \le t_0 \le 2, \end{cases} \tag{7.60}$$

où $t^* \approx 1.2285$.

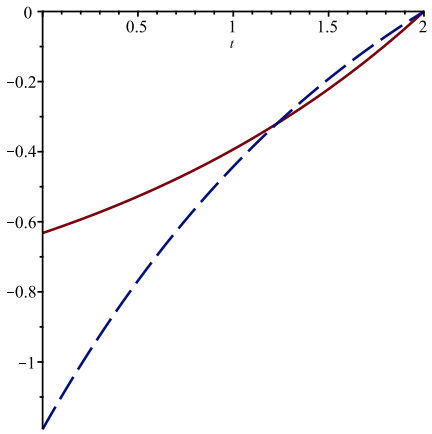

Figure 7.2 Fonctions $F_1(t_0)$ (ligne pleine) et $F_2(t_0)$ définies respectivement par les équations (7.52) et (7.53) pour $t \in [0, 2]$, avec $t_1 = 2$.

Remarque. La valeur de u_0 donnée dans l'équation (7.60) est en fait la solution optimale si l'on suppose que l'optimiseur doit choisir $u(t)$ égal à $k_1 = 1$ ou $k_2 = 0{,}9$ pour tout $t \in [t_0, t_1)$. Cependant, si l'on admet la possibilité de choisir $u(t) \equiv 0$, alors ce résultat est correct au temps t_0 si, et seulement si, la fonction de valeur correspondante est inférieure ou égale à $K-1$ (voir l'équation (7.15)). Lorsque $K = 0$, on constate que l'optimiseur ne devrait exercer aucun contrôle pour $t_0 \in [0.21, 2)$ (approximativement).

La section suivante considère le cas où t_1 tend vers l'infini.

7.6 Le cas invariant dans le temps

Une fois la fonction de valeur calculée dans la section précédente, il est possible de considérer la limite lorsque t_1 tend vers l'infini. En réalité, comme on peut le voir dans l'exemple présenté, la fonction de valeur devient alors constante. Cela est dû au fait que lorsque $t_1 \to \infty$, le problème devient *stationnaire*, de sorte que $F'(0, t_0) = 0$.

Explication.
Cette section formule des observations essentielles sur le comportement de la fonction de valeur lorsque la borne de temps t_1 tend vers l'infini. Dans un cadre de commande optimale à horizon fini, la fonction de valeur $F(0, t_0)$ dépend explicitement du temps restant jusqu'à l'instant terminal t_1. Toutefois, lorsque $t_1 \to \infty$, l'horizon de planification devient illimité. Dans ce cas, la fonction de valeur devient indépendante du temps t_0 :

$$\lim_{t_1 \to \infty} F(0, t_0) = F(0), \tag{7.61}$$

où $F(0)$ est une constante représentant le coût à long terme en régime stationnaire.

En conséquence, la dérivée de la fonction de valeur par rapport au temps s'annule :

$$F'(0,t_0) = \frac{d}{dt_0} F(0,t_0) = 0. \tag{7.62}$$

Cette condition exprime le fait que, dans les problèmes de commande optimale stationnaire, les décisions optimales dépendent uniquement de l'état du système et non du temps absolu.

Il en résulte que l'équation de programmation dynamique (7.35) se réduit à :

$$0 = \inf_{u(t_0)} \left\{ g[u(t_0)] + \nu_0 \sum_{j=1}^{2} p_{0,j}(t_0) \left[-\frac{\lambda}{\nu_j} + F(0) \right] - \nu_0 F(0) \right\}. \tag{7.63}$$

Un cas particulier pouvant être résolu explicitement est maintenant présenté. Supposons que la fonction $f(\cdot)$ apparaissant dans le coût $J(i)$ défini dans l'équation (7.4) soit donnée par :

$$f(u_0) = \frac{c u_0^2}{(u_0 + \nu_0)^2}, \tag{7.64}$$

où c est une constante positive. On remarque que $f(0) = 0$ et que $f(u_0) > 0$ si $u_0 > 0$, comme requis. De plus, $f'(u_0) > 0$, ce qui indique que les coûts de maintenance augmentent avec u_0, ce qui est logique.

Supposons ensuite que $\nu_1 = \nu_2$ (les taux de transition depuis les états dégradés 1 et 2). En dérivant l'équation (7.63) par rapport à u_0, on obtient, après simplification :

$$0 = \frac{2c u_0}{u_0 + \nu_0} - \frac{\nu_0}{\nu_1} + \nu_0 F(0). \tag{7.65}$$

Il en résulte que la commande optimale est donnée par :

$$u_0^* = \frac{\nu_0^2 \left[1 - \nu_1 F(0)\right]}{2c\nu_1 - \nu_0 \left[1 - \nu_1 F(0)\right]}. \tag{7.66}$$

Explication.
Pour dériver l'expression de l'équation (7.65), on part de la forme stationnaire de l'équation de programmation dynamique donnée en (7.63).

On suppose :

$$\begin{cases} \nu_1 = \nu_2, \\ p = \dfrac{1}{2} \Rightarrow p_{0,1}(u_0) = \dfrac{u_0}{2(u_0+\nu_0)}, \quad p_{0,2}(u_0) = \dfrac{u_0}{2(u_0+\nu_0)}, \\ f(u_0) = \dfrac{cu_0^2}{(u_0+\nu_0)^2} \Rightarrow g(u_0) = \dfrac{cu_0^2}{(u_0+\nu_0)^2} - \lambda. \end{cases} \tag{7.67}$$

Ainsi, le terme de somme dans l'équation (7.63) se simplifie comme suit :

$$\sum_{j=1}^{2} p_{0,j}(u_0) \left(-\frac{\lambda}{\nu_j} + F(0,t_0) \right) = \frac{u_0}{u_0+\nu_0} \left(-\frac{\lambda}{\nu_1} + F(0,t_0) \right). \tag{7.68}$$

En remplaçant dans l'équation (7.63), on obtient :

$$0 = \frac{cu_0^2}{(u_0+\nu_0)^2} - \lambda + \nu_0 \cdot \frac{u_0}{u_0+\nu_0} \left(-\frac{\lambda}{\nu_1} + F(0,t_0) \right) - \nu_0 F(0,t_0). \tag{7.69}$$

Pour trouver la commande optimale u_0^*, dérivons le membre de droite de l'équation (7.69) par rapport à u_0 :

$$\frac{d}{du_0} \left(\frac{cu_0^2}{(u_0+\nu_0)^2} \right) = c \cdot \frac{2u_0(u_0+\nu_0)^2 - u_0^2 \cdot 2(u_0+\nu_0)}{(u_0+\nu_0)^4} = \frac{2cu_0\nu_0}{(u_0+\nu_0)^3}. \tag{7.70}$$

Également :

$$\frac{d}{du_0} \left(\frac{u_0}{u_0+\nu_0} \right) = \frac{(u_0+\nu_0)\cdot 1 - u_0 \cdot 1}{(u_0+\nu_0)^2} = \frac{\nu_0}{(u_0+\nu_0)^2}. \tag{7.71}$$

Ainsi, la dérivée de l'équation (7.69) devient :

$$\begin{aligned} &\frac{d}{du_0} \left\{ \frac{cu_0^2}{(u_0+\nu_0)^2} + \nu_0 \cdot \frac{u_0}{u_0+\nu_0} \left(-\frac{\lambda}{\nu_1} + F(0,t_0) \right) \right\} \\ &= \frac{2cu_0\nu_0}{(u_0+\nu_0)^3} + \nu_0 \cdot \left(-\frac{\lambda}{\nu_1} + F(0,t_0) \right) \cdot \frac{\nu_0}{(u_0+\nu_0)^2}. \end{aligned} \tag{7.72}$$

Multipliant les deux termes par $(u_0+\nu_0)^2$ pour éliminer les dénominateurs :

$$0 = \frac{2cu_0\nu_0}{(u_0+\nu_0)^3}(u_0+\nu_0)^2 + \nu_0 \left(-\frac{\lambda}{\nu_1} + F(0,t_0) \right) \cdot \frac{\nu_0}{(u_0+\nu_0)^2}(u_0+\nu_0)^2. \tag{7.73}$$

En simplifiant et en posant $\lambda = 1$ comme dans le cas particulier de la section précédente, on

obtient :

$$0 = \frac{2cu_0\nu_0}{u_0+\nu_0} - \frac{\nu_0^2}{\nu_1} + \nu_0^2 F(0,t_0). \tag{7.74}$$

En divisant par ν_0 :

$$0 = \frac{2cu_0}{u_0+\nu_0} - \frac{\nu_0}{\nu_1} + \nu_0 F(0,t_0). \tag{7.75}$$

En réarrangeant cette équation pour isoler u_0, on obtient :

$$\begin{aligned} \frac{2cu_0}{u_0+\nu_0} = \frac{\nu_0}{\nu_1} - \nu_0 F(0,t_0) \Rightarrow 2cu_0 &= \left(\frac{\nu_0}{\nu_1} - \nu_0 F(0,t_0)\right)(u_0+\nu_0) \\ \Rightarrow \quad 2cu_0 &= A(u_0+\nu_0) \end{aligned} \tag{7.76}$$

$$\text{où } A := \frac{\nu_0}{\nu_1} - \nu_0 F(0,t_0)$$

$$\Rightarrow 2cu_0 = Au_0 + A\nu_0 \Rightarrow u_0(2c-A) = A\nu_0 \Rightarrow u_0 = \frac{A\nu_0}{2c-A} \tag{7.77}$$

En résolvant pour u_0, on obtient l'expression fermée :

$$A = \frac{\nu_0}{\nu_1} - \nu_0 F(0,t_0) = \frac{\nu_0(1-\nu_1 F(0,t_0))}{\nu_1} \Rightarrow u_0^* = \frac{\nu_0^2(1-\nu_1 F(0,t_0))}{2c\nu_1 - \nu_0(1-\nu_1 F(0,t_0))} \tag{7.78}$$

Comme u_0^* représente un taux de contrôle (taux d'inspection), il doit satisfaire $u_0^* \geq 0$.

Prenons maintenant $\nu_0 = \nu_1 = 1$. Alors, u_0^* devient :

$$u_0^* = \frac{1-F(0,t_0)}{2c-1+F(0,t_0)}. \tag{7.79}$$

En remplaçant cette expression dans l'équation (7.63), on obtient, si $p = 1/2$, que la fonction de valeur $F(0,t_0)$ satisfait l'équation algébrique :

$$0 = \frac{[1-F(0,t_0)]^2}{4c} - \lambda - \lambda\frac{[1-F(0,t_0)]}{2c} - \frac{[2c-1-F(0,t_0)]}{2c}F(0,t_0). \tag{7.80}$$

Explication.
Étant donné que $\nu_0 = \nu_1 = 1$, l'expression de la commande optimale dans l'équation (7.66) se simplifie en (7.79) :

$$u_0^* = \frac{\nu_0^2[1-\nu_1 F(0,t_0)]}{2c\nu_1 - \nu_0[1-\nu_1 F(0,t_0)]} \Rightarrow u_0^* = \frac{1-F(0,t_0)}{2c-1+F(0,t_0)}. \tag{7.81}$$

Cette expression est maintenant insérée dans l'équation stationnaire (7.69) :

$$0 = f(u_0) - \lambda + \frac{u_0}{u_0+1}(-\lambda + F(0,t_0)) - F(0,t_0), \tag{7.82}$$

où $f(u_0) = \dfrac{cu_0^2}{(u_0+1)^2}$ et $\lambda = 1$.

$$\begin{aligned}
&\text{Étant donné } u_0^* = \frac{1-F(0,t_0)}{2c-1+F(0,t_0)},\\
&\text{alors } f(u_0^*) = \frac{c(u_0^*)^2}{(u_0^*+1)^2} = \frac{(1-F(0,t_0))^2}{4c},\\
&\text{Également, } \frac{u_0^*}{u_0^*+1} = \frac{1-F(0,t_0)}{2c},
\end{aligned}$$

En remplaçant dans l'équation DPE :

$$\begin{aligned}
0 &= f(u_0^*) - \lambda + \left(\frac{u_0^*}{u_0^*+1}\right)(-1+F(0,t_0)) - F(0,t_0)\\
&= \frac{(1-F(0,t_0))^2}{4c} - \lambda + \left(\frac{1-F(0,t_0)}{2c}\right)(-1+F(0,t_0)) - F(0,t_0).
\end{aligned} \tag{7.83}$$

ce qui se simplifie pour donner l'équation (7.80).

On trouve que :

$$3F(0,t_0) = 2c - \lambda - \sqrt{4c^2+8c\lambda+\lambda^2+6\lambda-3}, \tag{7.84}$$

d'où il découle que la commande optimale est donnée par :

$$u_0^* = \frac{-3+2c-\lambda-\sqrt{\lambda^2+(8c+6)\lambda+4c^2-3}}{3-8c+\lambda+\sqrt{\lambda^2+(8c+6)\lambda+4c^2-3}}. \tag{7.85}$$

Explication.

Pour déterminer la fonction de valeur $F(0,t_0)$, on résout l'équation algébrique obtenue dans le cas stationnaire (voir l'équation (7.80)), qui prend la forme d'une équation quadratique. En utilisant la formule quadratique, on obtient deux racines possibles. Parmi celles-ci, seule celle qui satisfait la condition $F(0,t_0) \geq 0$ est retenue comme solution.

L'expression explicite pour la commande optimale u_0^* est obtenue en remplaçant cette solution

de $F(0,t_0)$ dans la relation précédente donnée par (7.79) : $u_0^* = \dfrac{1-F(0,t_0)}{2c-1+F(0,t_0)}$.

En utilisant cette solution pour $F(0,t_0)$, à savoir :

$$F(0,t_0) = \frac{1}{3}\left(2c-\lambda-\sqrt{\lambda^2+(8c+6)\lambda+4c^2-3}\right), \tag{7.86}$$

et en la substituant dans l'expression de u_0^*, puis en procédant à une simplification algébrique, on parvient à exprimer la commande optimale de manière entièrement explicite, comme cela est présenté dans l'équation (7.85) : $u_0^* = \dfrac{-3+2c-\lambda-\sqrt{\lambda^2+(8c+6)\lambda+4c^2-3}}{3-8c+\lambda+\sqrt{\lambda^2+(8c+6)\lambda+4c^2-3}}$.

Il apparaît clairement que la commande optimale doit être $u_0^* = 0$ si $K = 0$ et $\lambda \leq 1$. Prenons $\lambda = 10$. On trouve alors que la constante c doit vérifier :

$$c > \frac{1}{5}\left(12+\sqrt{139}\right) \simeq 4.76. \tag{7.87}$$

Par exemple, si l'on prend $c = 5$, alors $F(0,t_0) = -\sqrt{657}/3 \simeq -8.544$ et

$$u_0^* = \frac{-1-\sqrt{73}}{1-\sqrt{73}} \simeq 20.93. \tag{7.88}$$

La commande optimale est représenté dans la figure 7.3 pour $c \in [5,10]$.

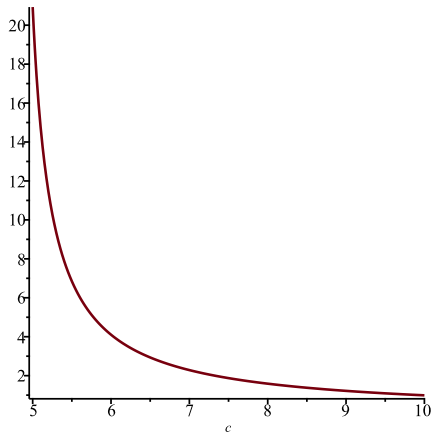

Figure 7.3 Commande optimale u_0^* donnée par l'équation (7.85) lorsque $\lambda = 10$ et $c \in [5,10]$.

Remarque. Si $u(t) \equiv 0$, alors $F_0(0,t_0) = K - 10$. Par conséquent, la commande optimale vaut bien approximativement 20.93 lorsque $\lambda = 10$ et $c = 5$ si et seulement si $K > 1.456$.

7.7 Conclusion

Dans le cadre d'une chaîne de Markov à temps continu, cette étude traite d'un problème de commande optimale appelé problème de retour à l'état initial (homing problem), représentant l'état de fonctionnement d'une machine. L'objectif principal était de déterminer le taux d'inspection optimal qui maximise le temps de fonctionnement avant que la réparation ne devienne nécessaire, tout en tenant compte des coûts de maintenance.

Le modèle peut être étendu pour inclure d'autres états, tels qu'un état absorbant représentant une panne irréversible de la machine. Si aucun contrôle n'est appliqué, c'est-à-dire en l'absence de maintenance, la chaîne évoluera très probablement de l'état de fonctionnement vers cet état de défaillance terminale.

Des solutions analytiques exactes ont été obtenues pour deux problèmes particuliers dans cette recherche. Pour des cas plus généraux, des techniques numériques ou des simulations peuvent être nécessaires afin d'évaluer la fonction de valeur et de dériver la politique de commande optimale correspondante.

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS

8.1 Conclusion

Le premier ensemble d'objectifs de notre travail était le développement d'une méthodologie pour réguler efficacement le flux de clients dans un système de files d'attente en déterminant le nombre optimal de serveurs et en identifiant la distribution optimale des temps de service pour servir les clients. On a mené cette recherche afin d'atteindre des objectifs tels que la minimisation des temps d'attente dans la file, la maximisation du débit, ou la réduction des coûts totaux, tout en respectant les contraintes imposées, telles que des capacités limitées et des temps d'arrivée et de service stochastiques. Le second ensemble d'objectifs était de proposer une politique optimale de maintenance et d'inspection, de manière à satisfaire les objectifs dérivés et les contraintes.

Lors de la modélisation du temps comme variable aléatoire, notre souci est de construire des modèles tenant compte de l'aléa dans l'horizon temporel des systèmes étudiés et de développer des politiques de contrôle intégrant cette incertitude pour améliorer l'efficacité des opérations. Cette approche implique l'utilisation de la modélisation probabiliste, des méthodes d'optimisation stochastique et des politiques de contrôle pour prendre des décisions éclairées sur la base des informations disponibles concernant le système de file d'attente.

Chacun des trois articles issus de notre thèse examine comment les temps de premier passage dans les processus stochastiques peuvent être utilisés pour déclencher des actions spécifiques lorsqu'un système atteint un seuil prédéterminé. Certaines orientations clés de cette étude sont listées ci-dessous.

1. **Analyse des indicateurs fonctionnels :**
   En intégrant le temps aléatoire, notre analyse se concentre sur la simulation ou la modélisation des effets des horizons temporels stochastiques sur les indicateurs de performance du système, dans un contexte où les temps ne peuvent pas être supposés déterministes mais suivent plutôt une caractéristique aléatoire. Cela est essentiel pour comprendre la nature stochastique et formuler des prévisions précises concernant leur comportement, tant dans les systèmes de files d'attente que dans l'amélioration de la fiabilité des systèmes par l'optimisation de la politique de maintenance et d'inspection.
2. **Modélisation des pratiques d'arrivée et de service :**
   La conception d'un système de service est envisagée dans un horizon temporel allant de 0 à un instant terminal aléatoire. L'objectif est d'établir une condition d'équilibre

entre la capacité du système de files d'attente et le coût global de l'attente des clients, qui est intégré dans les formulations de programmation mathématique.

Progression de la recherche :

1. **Gestion du nombre de serveurs dans les systèmes de files d'attente**
   Notre thèse s'est initialement concentrée sur l'optimisation de l'affectation des serveurs dans les files $M/M/k$, ce qui a permis d'établir une méthodologie de contrôle dépendant de l'état.
2. **Choix de la distribution des temps de service**
   En s'appuyant sur cette base, notre travail s'est poursuivi avec un accent mis sur l'optimisation du temps de service dans les files $M/G/1$, dans lesquelles la distribution du temps de service est considérée comme variable de décision principale.
3. **Politique de maintenance et d'inspection**
   La thèse a ensuite évolué vers la planification de la maintenance et l'ingénierie de la fiabilité, en appliquant les chaînes de Markov en temps continu et en intégrant la programmation dynamique pour la prise de décision dans des environnements stochastiques, dans le but d'optimiser les taux d'inspection.

8.2 Suggestions pour les travaux futurs

1. **Optimisation des paramètres dans un réseau de files et systèmes en maintenance :**
   Lorsqu'on considère le temps comme une variable aléatoire, l'optimisation des paramètres du système peut devenir difficile dans le contexte d'un réseau de files d'attente ou de systèmes intégrés en maintenance, notamment dans les scénarios où les unités opérationnelles sont en activité. Ainsi, la méthodologie pour obtenir les résultats attendus dans de telles conditions représente une voie prometteuse pour des recherches futures.
2. **Extension de l'optimisation des files multi-serveurs $M/M/k$ avec modélisation prédictive et ajustements dynamiques :**
   En plus du fait que l'on peut décider du nombre de serveurs actifs à chaque instant, d'autres modifications du modèle de file $M/M/k$ peuvent être explorées. Par exemple, les serveurs pourraient être autorisés à servir plusieurs clients simultanément, ou la discipline de la file pourrait être modifiée. De plus, l'intégration de techniques d'apprentissage automatique ou de modèles autorégressifs pour prédire la dynamique des files dans des conditions variables pourrait améliorer la prise de décision en temps réel et

la planification de l'optimisation, et peut être proposée comme une extension précieuse de ce travail.

3. **Extension de la file $M/G/1$ pour l'optimisation de la prise de décision :**
   Diverses modifications du modèle de file $M/G/1$ peuvent être proposées, telles que l'introduction de plusieurs serveurs. L'objectif pourrait être de mettre en œuvre des mécanismes de contrôle qui améliorent le taux d'entrée des clients dans le système de file. L'exploration d'extensions du modèle actuel à des systèmes plus complexes, tels que les files multi-serveurs et les systèmes interconnectés, fait actuellement partie de nos intérêts, car cela serait plus représentatif des systèmes réels.
   Pour surmonter les limites des solutions explicites, le recours à des simulations numériques est suggéré ; les travaux futurs pourraient inclure des simulations pour évaluer l'applicabilité de nos résultats dans des contextes plus vastes. Cela pourrait aider à identifier le comportement des stratégies de contrôle dans différents scénarios et sous diverses distributions de temps de service.
4. **Modélisation des états de défaillance de machines : de l'exploitation à la panne irréparable (dans le modèle de défaillance des machines) :**
   Le modèle peut être étendu en introduisant des états supplémentaires pour capturer un comportement de dégradation plus réaliste. Par exemple, un état peut être ajouté pour représenter un scénario dans lequel une panne de machine devient irréparable. En l'absence de pratiques de contrôle, c'est-à-dire si la machine n'est jamais entretenue, il peut exister une forte probabilité que le processus passe d'un état opérationnel à un état de panne irréversible.

RÉFÉRENCES

[1] M. B. Aryanezhad et S. J. Sajadi, *Operations Research 2.* Tehran : Amir Kabir Publications, 2008, in Persian.

[2] M. Lefebvre, *Applied stochastic processes.* Springer Science & Business Media, 2007.

[3] D. Kececioglu *et al.*, *Reliability engineering handbook.* DEStech Publications, Inc, 2002, vol. 1.

[4] K.-H. Wang, "Optimal control of a removable and non-reliable server in an m/m/1 queueing system with exponential startup time," *Mathematical Methods of Operations Research*, vol. 58, p. 29–39, 2003.

[5] S. Kumar et K. Muthuraman, "A numerical method for solving singular stochastic control problems," *Operations Research*, vol. 52, n 4, p. 563–582, 2004.

[6] J.-C. Ke, "Threshold control for a removable and un-reliable server with different type vacations and startup," *American Journal of Mathematical and Management Sciences*, vol. 25, n 1-2, p. 97–120, 2005.

[7] E. A. Aksenova, A. V. Sokolov *et al.*, "The optimal implementation of two fifo-queues in single-level memory," *Applied Mathematics*, vol. 2, n 10, p. 1297–1302, 2011.

[8] A. D. Banik, "Stationary distributions and optimal control of queues with batch markovian arrival process under multiple adaptive vacations," *Computers & Industrial Engineering*, vol. 65, n 3, p. 455–465, 2013.

[9] E. Özkan et J. P. Kharoufeh, "Optimal control of a two-server queueing system with failures," *Probability in the Engineering and Informational Sciences*, vol. 28, n 4, p. 489–527, 2014.

[10] I. Ahmed, K. T. Phan et T. Le-Ngoc, "Optimal stochastic power control for energy harvesting systems with statistical delay constraint," dans *2015 IEEE Global Communications Conference (GLOBECOM).* IEEE, 2015, p. 1–6.

[11] Y. B. Grishunina, "Optimal control of queue in the $m|g|1|\infty$ system with possibility of customer admission restriction," *Automation and Remote Control*, vol. 76, p. 433–445, 2015.

[12] F. Jiang, C. Feng, C. Zhu et Y. Sun, "Performance analysis of active queue management algorithm based on reinforcement learning." *Journal Européen des Systèmes Automatisés*, vol. 53, n 5, 2020.

[13] S. Asadzadeh, B. Akhavan et B. Akhavan, “Multi-objective optimization of gas station performance using response surface methodology,” *International Journal of Quality & Reliability Management*, vol. 38, n 2, p. 465–483, 2021.

[14] M. Lefebvre, “On a modified m/m/m/n queueing model,” *Computer Science Journal of Moldova*, vol. 85, n 1, p. 29–40, 2021.

[15] L. Luo, Y.-H. Tang, M.-M. Yu et W.-Q. Wu, “Optimal control policy of m/g/1 queueing system with delayed randomized multiple vacations under the modified min (n, d)-policy control,” *Journal of the Operations Research Society of China*, vol. 11, n 4, p. 857–874, 2023.

[16] X. Fan-Orzechowski et E. A. Feinberg, “Optimal admission control for a markovian queue under the quality of service constraint,” dans *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005, p. 1729–1734.

[17] P. Cao et J. Xie, “Optimal control of a multiclass queueing system when customers can change types,” *Queueing Systems*, vol. 82, n 3, p. 285–313, 2016.

[18] M. Lefebvre, “An application of queueing theory in hydrology,” *Atti della Accademia Peloritana dei Pericolanti-Classe di Scienze Fisiche, Matematiche e Naturali*, vol. 97, n S2, p. 18, 2019.

[19] J. Wen, N. Geng et X. Xie, “Optimal insertion of customers with waiting time targets,” *Computers & Operations Research*, vol. 122, p. 105001, 2020.

[20] A. Bhati, S. R. B. Pillai et R. Vaze, “On the age of information of a queuing system with heterogeneous servers,” dans *2021 National Conference on Communications (NCC)*. IEEE, 2021, p. 1–6.

[21] D. V. Efrosinin et V. V. Rykov, “On performance characteristics for queueing systems with heterogeneous servers,” *Automation and Remote Control*, vol. 69, n 1, p. 61–75, 2008.

[22] J. de Santi et N. L. da Fonseca, “Design of optimal active queue management controllers for hstcp in large bandwidth-delay product networks,” *Computer Networks*, vol. 55, n 12, p. 2772–2790, 2011.

[23] R. Banirazi, E. Jonckheere et B. Krishnamachari, “Minimum delay in class of throughput-optimal control policies on wireless networks,” dans *2014 American Control Conference*. IEEE, 2014, p. 2668–2675.

[24] P. Wang, D. Zhu et X. Lu, “Active queue management algorithm based on data-driven predictive control,” *Telecommunication Systems*, vol. 64, p. 103–111, 2017.

[25] V. Ponomarov et E. Lebedev, “Optimal control of retrial queues with finite population and state-dependent service rate,” dans *Advances in Computer Science for Engineering and Education 13*. Springer, 2019, p. 359–369.

[26] E. A. Chavari, M. Rostamy-Malkhalifeh et F. H. Lotfi, “An approach for selection of the most desirable internet network based on the cross-efficiency model in data envelopment analysis,” *Studies in Informatics and Control*, vol. 30, n 3, p. 99–108, 2021.

[27] Y. Zheng, *Stochastic Control and Stability for Queueing Networks in Random Environments*. The Pennsylvania State University, 2021.

[28] X. Li, S. Bi, Y. Zheng et H. Wang, “Energy-efficient online data sensing and processing in wireless powered edge computing systems,” *IEEE Transactions on Communications*, vol. 70, n 8, p. 5612–5628, 2022.

[29] E. Reticcioli, G. D. Di Girolamo, F. Smarra, A. Torzi, F. Graziosi et A. D’innocenzo, “Modeling and control of priority queueing in software defined networks via machine learning,” *IEEE Access*, vol. 10, p. 91 481–91 496, 2022.

[30] A. Petrović, M. Nikolić, U. Bugarić, B. Delibašić et P. Lio, “Controlling highway toll stations using deep learning, queuing theory, and differential evolution,” *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105683, 2023.

[31] B. Gao, Q. Chen, Y. Liu, K. Wan et K. Li, “Predictive cruise control under cloud control system for urban bus considering queue dissipation time,” *IEEE Transactions on Intelligent Vehicles*, vol. 8, n 4, p. 2639–2649, 2023.

[32] X. Fu et E. Modiano, “Joint learning and control in stochastic queueing networks with unknown utilities,” *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 6, n 3, p. 1–32, 2022.

[33] F. Hosseinzadeh Lotfi, T. Allahviranloo, W. Pedrycz, M. Shahriari, H. Sharafi et S. Razipour GhalehJough, “Foundations of decision,” dans *Fuzzy decision analysis : multi attribute decision making approach*. Springer, 2023, p. 1–56.

[34] T. Al-Hawari, F. Aqlan et O. Al-Araidah, “Performance analysis of an automated production system with queue length dependent service rates,” *International Journal of Simulation Modelling*, vol. 9, n 4, p. 184–194, 2010.

[35] K. S. Rao, M. GovindaRao et G. Rajkumar, “Transient analysis of interdependent queueing model with state dependent service rates,” *International Journal of Ultra Scientist of Physical Sciences*, vol. 23, n 3, 2011.

[36] L. Li, J. Wang et F. Zhang, “Customers’ equilibrium balking strategies in an m/m/1 queue with variable service rate,” dans *LISS 2012 : Proceedings of 2nd International Conference on Logistics, Informatics and Service Science*. Springer, 2012, p. 619–623.

[37] J. Chen, R. G. Addie et M. Zukerman, “Performance evaluation and service rate provisioning for a queue with fractional brownian input,” *Performance Evaluation*, vol. 70, n 11, p. 1028–1045, 2013.

[38] N. Lee et V. G. Kulkarni, “Optimal arrival rate and service rate control of multi-server queues,” *Queueing Systems*, vol. 76, p. 37–50, 2014.

[39] E. Munoz et E. H. Ruspini, “Simulation of fuzzy queueing systems with a variable number of servers, arrival rate, and service rate,” *IEEE Transactions on Fuzzy Systems*, vol. 22, n 4, p. 892–903, 2013.

[40] M. Delasay, A. Ingolfsson et B. Kolfal, “Modeling load and overwork effects in queueing systems with adaptive service rates,” *Operations Research*, vol. 64, n 4, p. 867–885, 2016.

[41] J. Rodrigues, S. M. Prado, N. Balakrishnan et F. Louzada, “Flexible m/g/1 queueing system with state dependent service rate,” *Operations Research Letters*, vol. 44, n 3, p. 383–389, 2016.

[42] P. V. Laxmi et K. Jyothsna, “Optimization of service rate in a discrete-time impatient customer queue using particle swarm optimization,” dans *International Conference on Distributed Computing and Internet Technology*. Springer, 2015, p. 38–42.

[43] S. Chen et L. Xia, “Optimal control of admission prices and service rates in open queueing networks,” *IFAC-PapersOnLine*, vol. 50, n 1, p. 928–933, 2017.

[44] T. Dai, T. Yu et X. Zhao, “Decision strategy of single server online customer service with variable service rates,” dans *2019 International Conference on Industrial Engineering and Systems Management (IESM)*. IEEE, 2019, p. 1–6.

[45] B. Büke et W. Qin, “Many-server queues with random service rates in the halfin-whitt regime : A measure-valued process approach,” *arXiv preprint arXiv :1905.04198*, 2019.

[46] Y. Su et J. Li, “Optimality of admission control in an m/ m/ 1/ n queue with varying services,” *Stochastic Models*, vol. 37, n 2, p. 317–334, 2021.

[47] P. V. Laxmi, E. G. Bhavani et K. Jyothsna, “Analysis of markovian queueing system with second optional service operating under the triadic policy,” *OPSEARCH*, vol. 60, n 1, p. 256–275, 2023.

[48] I. Lakkumikanthan et S. Balasubramanian, “Optimal control of service rates of discrete-time (s, s) queueing-inventory systems with finite buffer,” dans *2023 5th International Conference on Problems of Cybernetics and Informatics (PCI)*. IEEE, 2023, p. 1–4.

[49] A. Dudin, S. Dudin et O. Dudina, “Analysis of a queueing system with mixed service discipline,” *Methodology and Computing in Applied Probability*, vol. 25, n 2, p. 57, 2023.

[50] G. Chen, Z. Liu et L. Xia, “Event-based optimization of service rate control in retrial queues,” *Journal of the Operational Research Society*, vol. 74, n 3, p. 979–991, 2023.

[51] C.-H. Wu, D.-Y. Yang et C.-R. Yong, “Performance evaluation and bi-objective optimization for f-policy queue with alternating service rates.” *Journal of Industrial & Management Optimization*, vol. 19, n 5, 2023.

[52] S. K. Singh, G. M. Cruz et F. R. Cruz, “Change point estimation of service rate in m/m/1/m queues : A bayesian approach,” *Applied Mathematics and Computation*, vol. 465, p. 128423, 2024.

[53] R. Tian, S. Su et Z. G. Zhang, “Equilibrium and social optimality in queues with service rate and customers’ joining decisions,” *Quality Technology & Quantitative Management*, vol. 21, n 1, p. 1–34, 2024.

[54] G. Parmigiani, “Optimal scheduling of fallible inspections,” *Operations Research*, vol. 44, n 2, p. 360–367, 1996.

[55] E.-K. Boukas et Z. Liu, “Production and maintenance control for manufacturing systems,” *IEEE Transactions on Automatic Control*, vol. 46, n 9, p. 1455–1460, 2001.

[56] C.-H. Wang et S.-H. Sheu, “Determining the optimal production–maintenance policy with inspection errors : using a markov chain,” *Computers & Operations Research*, vol. 30, n 1, p. 1–17, 2003.

[57] H. Suryadi et L. Papageorgiou, “Optimal maintenance planning and crew allocation for multipurpose batch plants,” *International Journal of Production Research*, vol. 42, n 2, p. 355–377, 2004.

[58] W. Liying, F. Youtong, S. Liying et L. Baoyou, “On fault diagnosis and inspection policy for deteriorating system,” dans *2007 Chinese Control Conference*. IEEE, 2007, p. 477–481.

[59] F. S. Erenay, “Optimal screening policies for colorectal cancer prevention and surveillance,” Thèse de doctorat, University of Wisconsin–Madison, 2010.

[60] R. Jiang et V. Makis, “Availability maximization for a cbm model,” dans *IISE Annual Conference. Proceedings*. Institute of Industrial and Systems Engineers (IISE), 2011, p. 1.

[61] R. Ahmadi et M. Newby, “Maintenance scheduling of a manufacturing system subject to deterioration,” *Reliability Engineering & System Safety*, vol. 96, n 10, p. 1411–1420, 2011.

[62] H. R. Golmakani et F. Fattahipour, “Optimal replacement policy and inspection interval for condition-based maintenance,” *International Journal of Production Research*, vol. 49, n 17, p. 5153–5167, 2011.

[63] W. Wang, “A simulation-based multivariate bayesian control chart for real time condition-based maintenance of complex systems,” *European Journal of Operational Research*, vol. 218, n 3, p. 726–734, 2012.

[64] M. J. Kim et V. Makis, “Optimal control of a partially observable failing system with costly multivariate observations,” *Stochastic Models*, vol. 28, n 4, p. 584–608, 2012.

[65] L. Liu, M. Yu, Y. Ma et Y. Tu, “Economic and economic-statistical designs of an x control chart for two-unit series systems with condition-based maintenance,” *European Journal of Operational Research*, vol. 226, n 3, p. 491–499, 2013.

[66] S. C. Kutzner et G. P. Kiesmüller, “Optimal control of an inventory-production system with state-dependent random yield,” *European Journal of Operational Research*, vol. 227, n 3, p. 444–452, 2013.

[67] F. Naderkhani ZG et V. Makis, “Optimal condition-based maintenance policy for a partially observable system with two sampling intervals,” *The International Journal of Advanced Manufacturing Technology*, vol. 78, p. 795–805, 2015.

[68] A. Arab, S. K. Khator, Q. Feng et Z. Han, “Control theoretic scheduling of angiography for implanted stents in human arteries,” dans *IISE Annual Conference. Proceedings*. Institute of Industrial and Systems Engineers (IISE), 2015, p. 3161.

[69] C. Lin et V. Makis, “Optimal bayesian maintenance policy and early fault detection for a gearbox operating under varying load,” *Journal of Vibration and Control*, vol. 22, n 15, p. 3312–3325, 2016.

[70] M. J. Kim, “Robust control of partially observable failing systems,” *Operations Research*, vol. 64, n 4, p. 999–1014, 2016.

[71] K. He, “Optimal maintenance planning in novel settings,” Thèse de doctorat, University of Pittsburgh, 2017.

[72] S. M. Mousavi, H. Shams et S. Ahmadi, “Simultaneous optimization of repair and control-limit policy in condition-based maintenance,” *Journal of Intelligent Manufacturing*, vol. 28, p. 245–254, 2017.

[73] X. Li, J. Cai, H. Zuo et H. Li, “Optimal cost-effective maintenance policy for a helicopter gearbox early fault detection under varying load,” *Mathematical Problems in Engineering*, vol. 2017, n 1, p. 4682409, 2017.

[74] Q. Sun, Z.-S. Ye et N. Chen, “Optimal inspection and replacement policies for multi-unit systems subject to degradation,” *IEEE Transactions on Reliability*, vol. 67, n 1, p. 401–413, 2017.

[75] B. Liu, J. Lin, L. Zhang et M. Xie, “A dynamic maintenance strategy for prognostics and health management of degrading systems : Application in locomotive wheel-sets,” dans *2018 IEEE International Conference on Prognostics and Health Management (ICPHM)*. IEEE, 2018, p. 1–5.

[76] E. Liu, T. T. Allen et S. Roychowdhury, “Cyber vulnerability maintenance policies that address the incomplete nature of inspection,” *Applied Stochastic Models in Business and Industry*, vol. 35, n 6, p. 1390–1410, 2019.

[77] C. Andriotis, K. Papakonstantinou et E. Chatzi, “Value of structural health monitoring quantification in partially observable stochastic environments,” *arXiv preprint arXiv :1912.12534*, 2019.

[78] P. Cao, K. Yang et K. Liu, “Optimal selection and release problem in software testing process : a continuous time stochastic control approach,” *European Journal of Operational Research*, vol. 285, n 1, p. 211–222, 2020.

[79] Q. Sun, Z.-S. Ye et X. Zhu, “Managing component degradation in series systems for balancing degradation through reallocation and maintenance,” *IISE Transactions*, vol. 52, n 7, p. 797–810, 2020.

[80] A. Mancuso, M. Compare, A. Salo et E. Zio, “Optimal prognostics and health management-driven inspection and maintenance strategies for industrial systems,” *Reliability Engineering & System Safety*, vol. 210, p. 107536, 2021.

[81] R. Ahmadi, “A markovian control approach to maintenance optimization of a system with stochastic dependence and hidden failures,” *International Journal of Reliability, Quality and Safety Engineering*, vol. 28, n 04, p. 2150028, 2021.

[82] S. Gan, N. Yousefi et D. W. Coit, “Optimal control-limit maintenance policy for a production system with multiple process states,” *Computers & Industrial Engineering*, vol. 158, p. 107454, 2021.

[83] Y. Wang et Y. Li, “Replacement policy for a single-component machine with limited spares in a finite time horizon,” dans *12th International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE 2022)*, vol. 2022. IET, 2022, p. 1474–1481.

[84] M. Roux, Y.-P. Fang et A. Barros, “Maintenance planning under imperfect monitoring : An efficient pomdp model using interpolated value function,” *IFAC-PapersOnLine*, vol. 55, n 16, p. 128–135, 2022.

[85] P. G. Morato, K. G. Papakonstantinou, C. P. Andriotis, J. S. Nielsen et P. Rigo, “Optimal inspection and maintenance planning for deteriorating structural components

through dynamic bayesian networks and markov decision processes," *Structural Safety*, vol. 94, p. 102140, 2022.

[86] M. Lefebvre et P. Pazhoheshfar, "An optimal control problem for the maintenance of a machine," *International Journal of Systems Science*, vol. 53, n 15, p. 3364–3373, 2022.

[87] S. Nasersarraf, S. Asadzadeh et Y. Samimi, "Determining the optimal policy in condition-based maintenance for electrical panels," *Iranian Electric Industry Journal of Quality and Productivity*, vol. 12, n 4, p. 37–45, 2023.

[88] W. Wang et X. Chen, "Piecewise deterministic markov process for condition-based imperfect maintenance models," *Reliability Engineering & System Safety*, vol. 236, p. 109271, 2023.

[89] G. Wang et Z. Zhu, "Optimal control of sampled-data systems based on an optimized stochastic sampling," *International Journal of Robust and Nonlinear Control*, vol. 33, n 7, p. 4304–4325, 2023.

[90] M. Lefebvre et R. Yaghoubi, "Optimal inspection and maintenance policy : Integrating a continuous-time markov chain into a homing problem," *Machines*, vol. 12, n 11, 2024. [En ligne]. Disponible : https://www.mdpi.com/2075-1702/12/11/795

[91] M. Lefebvre et R. Y. and, "Optimal control of a queueing system," *Optimization*, p. 1–14, 2024.

[92] P. Whittle, *Optimization over Time*. John Wiley & Sons, Inc., 1982.

[93] K. Zhou et J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[94] K. Uğurlu, "A new coherent multivariate average-value-at-risk," *Optimization*, vol. 72, n 2, p. 493–519, 2023.

[95] K. Uğurlu et T. Brzeczek, "Distorted probability operator for dynamic portfolio optimization in times of socio-economic crisis," *Central European Journal of Operations Research*, vol. 31, n 4, p. 1043–1060, 2023.

[96] M. Lefebvre et R. Yaghoubi, "Optimal service time distribution for an m/g/1 waiting queue," *Axioms*, vol. 13, n 9, 2024.

[97] A. N. Aziati et N. S. B. Hamdan, "Application of queuing theory model and simulation to patient flow at the outpatient department," dans *Proceedings of the international conference on industrial engineering and operations management bandung, indonesia*, 2018, p. 3016–3028.

[98] U. K. De et G. Rajbongshi, "Statistical application for the analysis of traffic congestion and its impact in a hill city," *Int. Journal Stat. Sci*, vol. 20, p. 19–48, 2020.

[99] W. J. Requia, C. D. Higgins, M. D. Adams, M. Mohamed et P. Koutrakis, "The health impacts of weekday traffic : A health risk assessment of pm2. 5 emissions during congested periods," *Environment International*, vol. 111, p. 164–176, 2018.

[100] R. Rishel, "Controlled wear process : modeling optimal control," *IEEE Transactions on Automatic Control*, vol. 36, n 9, p. 1100–1102, 1991.

[101] M. Lefebvre, "The homing problem for autoregressive processes," *IMA Journal of Mathematical Control and Information*, vol. 39, n 1, p. 322–344, 2022.

[102] P. Vrignat, F. Kratz et M. Avila, "Sustainable manufacturing, maintenance policies, prognostics and health management : A literature review," *Reliability Engineering & System Safety*, vol. 218, p. 108140, 2022.

[103] P. Whittle, *Risk-sensitive Optimal Control*. John Wiley & Sons, 1990.