MAINTENANCE OF DETERIORATING NON-EXPONENTIAL SINGLE SERVER QUEUE

KOH SIEW KHEW

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2013

ABSTRACT

Consider the single server queue in which the system capacity is infinite and the customers are served on a first come, first served basis. The case of a system without deterioration is first studied. The stationary queue length distribution and the stationary waiting time distribution are derived for the system in which the service time and interarrival time distributions are assumed to have constant asymptotic rates. The results found are verified by using simulation. Next consider a system in which the server would deteriorate due to random shocks and the seriously affected server will be sent for repair. A similar method is applied for deriving the stationary queue length distribution in a system in which the interarrival time distribution (or service time) is assumed to have a constant asymptotic rate while the service time (or interarrival time) remains exponentially distributed. From the stationary queue length distribution, a number of other characteristics can be derived. These include the sojourn time distribution of a customer who arrives when the queue is in a stationary state, and the expected length of the duration between two successive repair completions. From these distributions and expected length, the value of the specified maintenance level is found such that the long run average cost is minimized.

ABSTRAK

Pertimbangkan giliran satu-pelayan di mana muatan sistem adalah tak terhingga dan pelanggan dilayan berasaskan siapa yang datang dulu akan dilayan dulu. Kes yang mana sistem tidak akan merosot dikaji terlebih dahulu. Taburan panjang giliran pegun dan taburan masa menunggu pegun diterbitkan untuk sistem di mana taburan untuk masa layanan dan lat ketibaan dianggap mempunyai kadar asimptot yang malar. Hasil yang didapati disahkan dengan menggunakan simulasi. Sistem yang dipertimbangkan seterusnya ialah sistem di mana pelayan akan merosot disebabkan kejutan rawak dan pelayan yang terjejas teruk akan dihantar untuk dibaiki. Kaedah yang serupa digunakan untuk menerbitkan taburan panjang giliran pegun bagi sistem di mana taburan lat ketibaan (atau masa layanan) dianggap mempunyai kadar asimptot yang malar manakala masa layanan (atau lat ketibaan) kekal bertaburan eksponen. Daripada taburan panjang giliran pegun, beberapa ciri lain boleh diterbitkan. Ciri tersebut termasuk taburan masa persinggahan pelanggan yang sampai ke giliran dalam keadaan pegun, dan panjang jangkaan untuk tempoh di antara dua pembaikan yang berjaya. Daripada taburan dan panjang jangkaan tersebut, nilai tahap penyelenggaraan ditentukan supaya kos purata jangka panjang adalah minimum.

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisors, Professor Pooi Ah Hin and Dr. Ng Kok Haur, for their continuous support and guidance throughout the course of my research.

I would like to give special thanks to Dr. Tan Yi Fei, lecturer from Multimedia University (MMU), who had been encouraging me and gave me advice to solve some of the research problems. Thanks are also extended to MMU for offering me the position of contract research officer at the beginning of the research period.

I also wish to thank my family and friends for helping me get through the difficult times, and for all the unconditional emotional support, camaraderie and caring they provided. I would also like to thank Institute of Mathematical Sciences, University of Malaya and the staff in the institution for the administrative support and equipment provided.

Lastly, once again, I wish to express my sincere appreciation to Professor Pooi Ah Hin, a lecturer full of kindness, compassion and tolerance, who continuously offered me advice in a humble manner and gave me the confidence and encouragement to solve the research problems.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Literature Review

Diverse field of applications in queueing theory has aroused interest of many researchers to study this topic. In reality, the phenomena of queue exist in our daily life, and in areas such as, telecommunication, manufacturing, and computing. Queueing Theory was first applied to telephone traffic in the early 20^{th} century. One of the most influential persons in this field of study is Erlang [1], who applied the theory of probability to problems of telephone traffic and published in 1909 his first work on this subject, entitled "The Theory of Probabilities and Telephone Conversations". The works in queueing theory had developed initially rather slowly, but the pace has quickened since the mid of 20^{th} century when the computing machinery had advanced, and the applications were extended beyond the scope of telephone system. Despite the slow momentum of growth in the early days, there were a few significant contributions from the researchers who laid the foundation for further dynamic development in the field. In 1953, Kendall [2] introduced the well-known Kendall's notation to classify different types of queueing systems. The notation is described by A/B/m where A indicates the interarrival time distribution, B the service time distribution, and m the number of parallel service channels. The notation is later extended by Lee [3] to five-part descriptor A/B/m/Y/Z, where Y is the maximum number of customers the system can accommodate and Z is the queue discipline. For example, the M/M/1 denotes a single server queue of which both the interarrival time and service time have an exponential distribution, M/G/1 denotes a single server queue with exponential input, and a general service time distribution, and GI/G/c denotes a c-server queue with general interarrival time and service time distributions.

Different queueing systems may be represented by different models. Steady-state queue length and waiting time distributions are basic performance measures in the analysis of queueing systems. These distributions are of paramount importance for further study to optimize the production, reduce the manufacturing cost, avoid excessive waiting time, etc. The earliest queueing systems are modeled based on the assumption that the service time is exponentially distributed and the customers arrive according to a Poisson process in a single service channel (M/M/1) with the first come first served (FCFS) queueing characteristic. The memoryless property of the exponential distribution allows the M/M/1 system to be modeled as a Markov process which satisfies the Chapman-Kolmogorov Equation and the steady-state derivation is not an arduous process. When the service time or interarrival time distribution is non-exponential, we may no longer have a Markov process since the memoryless property does not hold. However, the imbedded Markov chains identified in the queueing systems (for example, M/G/1 and GI/M/c) allow a probabilistic approach to be applied for analyzing the system.

With the relaxation of the exponential assumption in both the service time and interarrival time distributions, we may find difficulty in analyzing the model for the queueing process. When the interarrival and service times are discrete random variables, the GI/G/1 queue is referred to as a discrete-time queue. In a discrete-time queue, the time axis is segmented into equidistant time units of length Δt, called slots. Several authors determined the steady-state waiting time distribution in the discrete-time GI/G/1 queue by numerical evaluations of the Wiener-Hopf factorization ([4–9]) and the matrix-analytic method ([10] and [11]). Steady-state waiting time distribution in the continuous-time GI/G/1 queues can be found by numerical approximations based on the theory of Fredholm integral equations [12]. In [13], a discrete-time version of the distributional Little's law was established. Based on this law, the queue length distribution for the discrete-time GI/G/1 queue may be obtained from its waiting-time distribution. Neuts [14] approximated the general distributions of service time and arrival time with phase-type distributions, and solved for the steady-state distribution using the matrix-geometric approach.

Works of all the aforementioned authors focused on the queueing systems without deterioration. However, some of the queueing systems may deteriorate or fail due to different causes, including age, usage and catastrophe. White & Christie [15] were the first to consider an M/M/1 queueing system with the service station subject to exponentially distributed interruptions. Soon after White & Christie paper, several papers related to server with interruptions were published [16–19]. The study of server with interruptions triggered investigations on maintenance of queueing systems subject to breakdowns. Maintenance can be categorized into preventive and corrective maintenance. Preventive maintenance is the maintenance carried out to prevent the systems from failing during operation. Corrective maintenance is the task performed to rectify and restore the systems back to operational condition when the systems fail. In 1960, Barlow & Hunter [20] initiated a simple periodic replacement model with minimum repair at failure. In their model, the system after repair is restored to its prior state before failure. Further investigations and extensions of the original minimal repair model have been proposed [21–25]. Various policies have been developed to provide maintenance for different queueing problems (see for example, [26–36]). Reviews in this area can be found in [26, 37–42] .

Besides maintenance policies, shock models have also been studied extensively. A general shock model is composed of two components X_n and Y_n, where X_n is the magnitude of the n-th shock and Y_n the interarrival time between two consecutive shocks. Shock models may be categorized into three distinct types: cumulative shock model, extreme shock model and δ-shock model. In the cumulative shock model, the system breaks down when the cumulative shock magnitude exceeds the given threshold [34−36, 43−46]. The extreme shock model is one in which the system fails as soon as the magnitude of an individual shock goes into some critical region [27, 47−49]. When the time lag between the two successive shocks falls into some critical region defined by a parameter δ, we get the δ-shock model [30, 50−54]. Some extensions have been made to the traditional shock models. For instance, Igaki et al. [55] extended the general shock models in [46] and [49] to a trivariate stochastic process $\{X_n, Y_n, J_n\}_{n=0}^{\infty}$ where (X_n, Y_n) is a correlated pair of renewal sequences, J_n a Markov chain formed by the external system states, and (X_{n+1}, Y_{n+1}, J_{n+1}) depends on (X_i, Y_i, J_i) for $0 \le i \le n$ through J_n only. Gut [56] presented a mixed shock model in which the system may break down either due to a large shock or an accumulation of many small shocks, depending on which reaches its critical level first. In 2005, Gut & Hüsler [45] extended this model to a framework in which the critical boundary for fatal shock decreases when there is an arriving non-fatal shock. A current literature review on shock models can be found in [57].

When a system subject to failure could only presume two operational states, namely perfect functioning state and complete failure state, it is called a two-state deteriorating system. Considering the existence of intermediate states between the above two operational states, the research can be extended to multi-state deteriorating system. For example, for a multi-state system which goes from the current operating state to the next inferior state, replacement policies [58−60] and inspection policies [61−64] have been developed. When random shock could occur and deteriorate the system, the corresponding systems have also been examined [35−36, 65−71].

Some authors assumed that the service rate may be reduced in the multi-state deteriorating system. For example, Kaufman & Lewis [72] considered a multi-state single server whose state may deteriorates from a state s to $s-1$ after a random amount of time and the service rate at state $s-1$ is less than that in state s. They analyzed the maintenance policies in the repair model and the replacement model using a semi-Markov decision process. Yang et al. [35] studied a model given by an unreliable M/M/1 queue with a multi-state server whose service rate deteriorates due to the shocks which occur randomly with random magnitudes. They derived the system size distribution, sojourn time distribution and expected length of the duration between two successive repair completions by using the matrix-geometric method of Neuts [14]. Based on the above characteristics of the system, they derived the long run average cost of the system and found the optimal strategy which minimized the cost. Yang et al. [36] modified their previous model by assuming that the system may also deteriorate whenever it produces an item. Chakravarthy [73] changed the arrival process in the model in [35] to a Markovian arrival process and studied the resulting unreliable MAP/M/1 queue.

1.2 Introduction to the Thesis

The present thesis considers a distribution of which the rate tends to a constant as the time t tends to infinity. Abbreviating **c**onstant **a**symptotic **r**ate to CAR, we may refer to the distribution as the CAR distribution. The requirement for the distribution to have a constant asymptotic rate is not a great restriction since in practice many distributions such as exponential, Erlang, hyperexponential, gamma, etc. satisfy this requirement. A numerical method is proposed to find the stationary queue length distribution and waiting time distribution in a one-server queue of which the interarrival time and service time distributions are CAR distributions. The numerical method proposed is adapted to investigate the model given in [35] with the distribution of the interarrival time or service time changed to a CAR distribution. The resulting queue, denoted as a CAR/M/1 or M/CAR/1 queue, cannot be represented as a continuous-time Markov chain. Hence, to analyze the queue, we may either explore the possibility of applying the matrix-geometric method to a Markov chain imbedded within the CAR/M/1 or M/CAR/1 queue, or use the proposed numerical method which is applicable for a non-Markovian process. In this thesis, the latter is chosen. For each of the above non-reliable CAR/M/1 and M/CAR/1 queues, its basic characteristics are derived.

1.3 Layout of the Thesis

In Chapter 2, a numerical method is proposed to find the stationary queue length and waiting time distributions of a CAR/CAR/1 queue.

In Chapter 3, the model given in [35] is studied. The interarrival time distribution in the model is changed to a CAR distribution. The numerical method proposed in Chapter 2 is adapted to find the queue length distribution, sojourn time distribution and the expected length of the duration between two successive repair completions when the queue is in a stationary state. The results thus found are used to find an optimal maintenance policy that minimizes the long run average cost.

The multi-state M/M/1 queue studied in [35] is considered again in Chapter 4. In this chapter, the distribution of the service time is instead changed to a CAR distribution. The model is then analyzed.

The thesis is concluded by some concluding remarks.

CHAPTER 2

QUEUE LENGTH AND WAITING TIME DISTRIBUTIONS IN A SINGLE SERVER QUEUE

2.1 Introduction

Consider the single server queue in which the system capacity is infinite and the customers are served on a first come, first served basis. Suppose the probability density function $f(t)$ and the cumulative distribution function $F(t)$ of the interarrival time are such that the rate $f(t)/[1-F(t)]$ tends to a constant as $t \to \infty$, and the rate computed from the distribution of the service time tends to another constant. Distributions of interarrival time and service time with the above constant asymptotic rates have been referred to in Chapter 1 as CAR distributions. We may denote the resulting queue as a CAR/CAR/1 queue. When the queue is in a stationary state, a set of equations for the stationary probabilities of the queue length and the states of the arrival and service processes is derived. Approximate results for the stationary probabilities can be obtained by solving the equations. Each probability may be found more accurately by an extrapolation of the probability on the values of Δt. The stationary probabilities obtained can be used to find the stationary queue length distribution and the waiting time distribution of a customer who arrives when the queue is in the stationary state.

2.2 Derivation of Equations for the Stationary Probabilities

A set of equations for the stationary probabilities of the queue length and the states of the arrival and service processes in the discretized CAR/CAR/1 queue is derived. First let

$g(t)$ be the probability density function (pdf) of the service time and τ_k the interval $((k-1)\Delta t, k\Delta t]$ for $k = 1, 2, 3, \ldots$. Furthermore let

$$\mu_k = \frac{g(k\Delta t)}{\int_{k\Delta t}^{\infty} g(u)du}, \qquad 1 \le k \le I,$$

where I is large enough such that

$$\mu_I \cong \lim_{k\to\infty} \mu_k .$$

Suppose a service starts at time $t = 0$. Then the probability that the service will be completed in the interval τ_1 is approximately $\mu_1\Delta t$, and given that the service is not completed in $\tau_1, \tau_2, \ldots, \tau_{k-1}$, the probability that the service will be completed in τ_k is approximately $\mu_k\Delta t$, $k = 2, 3, 4, \ldots$ where $\mu_k = \mu_I$ for $k \ge I$.

For the arrival process, let $f(t)$ be the pdf of the arrival time. Furthermore let

$$\lambda_k = \frac{f(k\Delta t)}{\int_{k\Delta t}^{\infty} f(u)du}, \qquad 1 \le k \le J,$$

where J is large enough such that

$$\lambda_J \cong \lim_{k\to\infty} \lambda_k .$$

Suppose a customer has arrived at time $t = 0$. Then the next customer will arrive in the interval τ_1 with an approximate probability $\lambda_1\Delta t$, and given that the next customer does not arrive in the interval $\tau_1, \tau_2, \ldots, \tau_{k-1}$, the probability that the next customer will arrive in τ_k will be approximately $\lambda_k\Delta t$ for $k = 2, 3, 4, \ldots$ where $\lambda_k = \lambda_J$ for $k \ge J$.

Let the interval before τ_1 as τ_0. Given that a service starts at a time in τ_0, we may define the state number ξ_k of the service process at the end of τ_k as

$$\xi_k = \begin{cases} 0, & \text{if} \\ & \bullet\ k = 0\text{; or} \\ & \bullet\ \text{the service ends in } \tau_k\text{, for } k \geq 1\text{; or} \\ & \bullet\ \text{the server is idle in } (0, k\Delta t]. \\ \min(k, I), & \text{if the service does not end in } \tau_k,\ k \geq 1. \end{cases}$$

Next, given that a customer arrives at a time in τ_0, we may define the state number ψ_k of the arrival process at the end of τ_k as

$$\psi_k = \begin{cases} 0, & \text{if } k = 0 \text{ or the next customer arrives in } \tau_k,\ k \geq 1. \\ \min(k, J), & \text{if the next customer does not arrive in } \tau_k,\ k \geq 1. \end{cases}$$

Let n_k be the queue length at the end of τ_k and $\boldsymbol{h}_k = (n_k, \xi_k , \psi_k)$. We may refer to $\boldsymbol{h}_k$ as the vector of characteristics of the queue at the end of τ_k.

Let $P_{nij}^{(k)}$ be the probability that at the end of τ_k, the number of customers in the system is n (including the customer that is being served), the service process is in state i and the arrival process is in state j, where $n \geq 0$, $i \in \{0, 1, 2,\ldots, I\}$ and $j \in \{0, 1, 2,\ldots, J\}$. Assume that

$$P_{nij} = \lim_{k \to \infty} P_{nij}^{(k)}$$

exists. To find the P_{nij}, we first make the following observations.

Suppose at the end of τ_{k-1}, the system is not empty, and the service and arrival processes are in state $i-1$ and $j-1$ at the end of τ_{k-1} respectively. Then only one of the following events can occur in the next time interval τ_k:

(a) A customer enters the system with the arrival rate λ_{j^*}, and at the end of τ_k, the vector of characteristics becomes $\boldsymbol{h}_k = (n + 1, i^*, 0)$;

(b) A customer leaves the system with the departure rate μ_{i^*}, and $\boldsymbol{h}_k = (n - 1, 0, j^*)$;

(c) No customers enter or leave the system, and $\boldsymbol{h}_k = (n, i^*, j^*)$;

where $i^* = \min(i, I)$ and $j^* = \min(j, J)$. However if the system is empty at the end of τ_{k-1}, and the arrival process is in state $j - 1$, then either one of the following events may occur in τ_k:

(d) A customer enters the system with arrival rate λ_{j^*}, and $\boldsymbol{h}_k = (1, 0, 0)$;

(e) No customers enter the system, and $\boldsymbol{h}_k = (0, 0, j^*)$.

Figures 2.2.1 to 2.2.5 illustrate the occurrence of the five events described above. In the figures,

1) the number inside the rectangle denotes the queue length at the end of indicated small time interval.
2) the number inside the ellipse denotes the state of the service process at the end of indicated small time interval.
3) the number inside the circle denotes the state of the arrival process at the end of indicated small time interval.
4) the symbol ‘x’ indicates that a customer enters the system at the indicated time.
5) the symbol ‘↓’ indicates that a customer leaves the system at the indicated time.

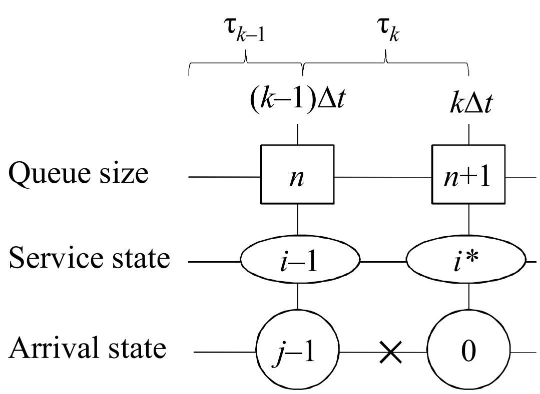

Figure 2.2.1 Transitions of queue length and states when Event (a) occurs.

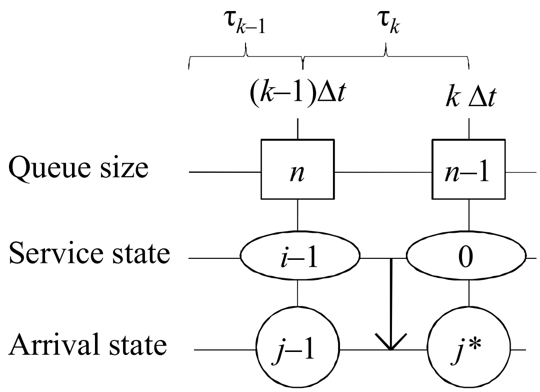

Figure 2.2.2 Transitions of queue length and states when Event (b) occurs.

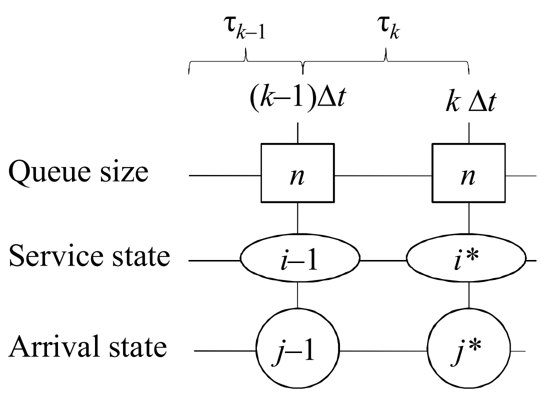

Figure 2.2.3 Transitions of queue length and states when Event (c) occurs.

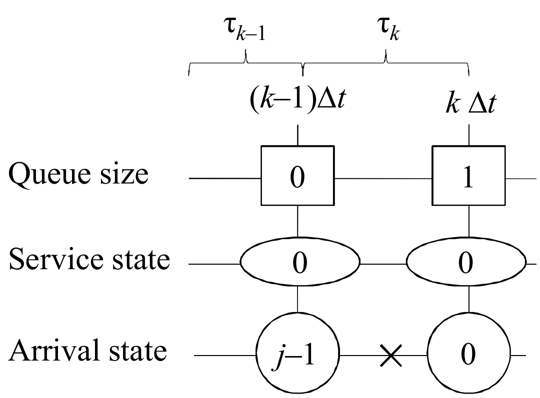

Figure 2.2.4 Transitions of queue length and states when Event (d) occurs.

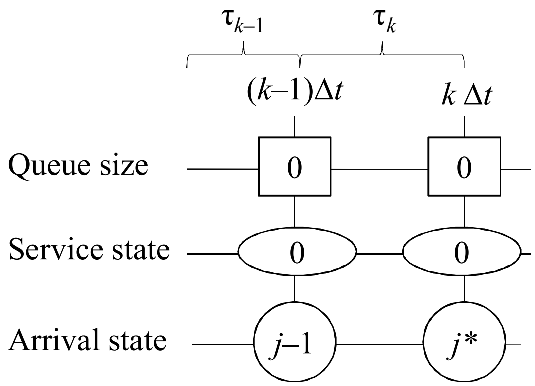

Figure 2.2.5 Transitions of queue length and states when Event (e) occurs.

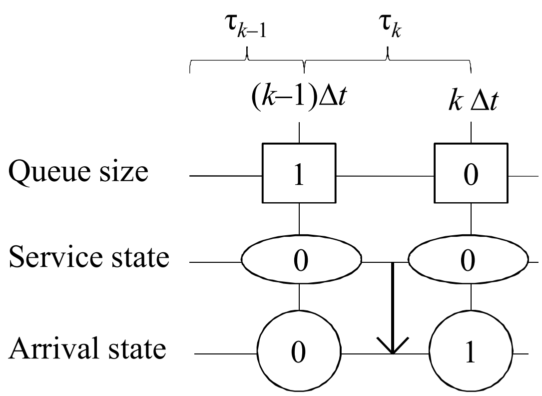

Figure 2.2.6 Transitions of queue length and states when Event (b) occurs in τ_k.

From Figure 2.2.6, it is easy to see that

$$P_{001}^{(k)} \cong P_{100}^{(k-1)}(\mu_1 \Delta t)(1-\lambda_1 \Delta t)\,. \tag{2.2.1}$$

When $k\to\infty$, (2.2.1) yields,

$$P_{001} \cong P_{100}(\mu_1 \Delta t)(1-\lambda_1 \Delta t)\,. \tag{2.2.2}$$

Similarly, with the aid of Figures 2.2.1–2.2.5, the following equations can be obtained.

$$P_{00j} \cong P_{00(j-1)}(1-\lambda_j \Delta t) + \sum_{i=0}^{j-1} P_{1i(j-1)}(\mu_{i+1}\Delta t)(1-\lambda_j \Delta t) \qquad \text{for } 2 \le j \le J-1, \tag{2.2.3}$$

$$P_{00J} \cong P_{00(J-1)}(1-\lambda_J\Delta t) + P_{00J}(1-\lambda_J\Delta t) + \sum_{i=0}^{J-1} P_{1i(J-1)}(\mu_{i+1}\Delta t)(1-\lambda_J\Delta t)$$
$$+ \sum_{i=0}^{J-1} P_{1iJ}(\mu_{i+1}\Delta t)(1-\lambda_J\Delta t) + P_{1IJ}(\mu_I\Delta t)(1-\lambda_J\Delta t) \quad . \qquad (2.2.4)$$

When the queue length is n = 1,

$$P_{n00} \cong \sum_{j=1}^{J-1} P_{00j}(\lambda_{j+1}\Delta t) + P_{00J}(\lambda_J\Delta t), \qquad (2.2.5)$$

$$P_{n01} \cong \sum_{i=1}^{I-1} P_{(n+1)i0}(\mu_{i+1}\Delta t)(1-\lambda_1\Delta t) + P_{(n+1)I0}(\mu_I\Delta t)(1-\lambda_1\Delta t), \qquad (2.2.6)$$

$$P_{n0j} \cong \sum_{i=0, i\neq j-1}^{I-1} P_{(n+1)i(j-1)}(\mu_{i+1}\Delta t)(1-\lambda_j\Delta t)$$
$$+ P_{(n+1)I(j-1)}(\mu_I\Delta t)(1-\lambda_j\Delta t) \qquad \text{for } 2 \le j \le J-1, \qquad (2.2.7)$$

$$P_{n0J} \cong \sum_{i=0, i\neq J-1}^{I-1} P_{(n+1)i(J-1)}(\mu_{i+1}\Delta t)(1-\lambda_J\Delta t) + P_{(n+1)I(J-1)}(\mu_I\Delta t)(1-\lambda_J\Delta t)$$
$$+ \sum_{i=0}^{I-1} P_{(n+1)iJ}(\mu_{i+1}\Delta t)(1-\lambda_J\Delta t) + P_{(n+1)IJ}(\mu_I\Delta t)(1-\lambda_J\Delta t) \quad , \qquad (2.2.8)$$

$$P_{nij} \cong P_{n(i-1)(j-1)}(1-\mu_i\Delta t)(1-\lambda_j\Delta t) \qquad \text{for } i \le j,\ 1 \le i \le I-1,\ 1 \le j \le J-1, \qquad (2.2.9)$$

$$P_{niJ} \cong P_{n(i-1)(J-1)}(1-\mu_i\Delta t)(1-\lambda_J\Delta t) + P_{n(i-1)J}(1-\mu_i\Delta t)(1-\lambda_J\Delta t) \qquad \text{for } 1 \le i \le I-1, \qquad (2.2.10)$$

$$P_{nIJ} \cong P_{n(I-1)(J-1)}(1-\mu_I\Delta t)(1-\lambda_J\Delta t) + P_{n(I-1)J}(1-\mu_I\Delta t)(1-\lambda_J\Delta t)$$
$$+ P_{nIJ}(1-\mu_I\Delta t)(1-\lambda_J\Delta t) \quad . \qquad (2.2.11)$$

When the queue length is n = 2, the expressions for P_{n0j}, $1 \le j \le J$ are the same as those given by (2.2.6), (2.2.7) and (2.2.8). Other P_{nij} when the queue length is n = 2, can be computed from the equations below:

$$P_{ni0} \cong \sum_{j \ge i-1}^{J-1} P_{(n-1)(i-1)j}(1-\mu_i\Delta t)(\lambda_{j+1}\Delta t)$$
$$+ P_{(n-1)(i-1)J}(1-\mu_i\Delta t)(\lambda_J\Delta t) \qquad \text{for } 1 \le i \le I-1, \qquad (2.2.12)$$

$$P_{nI0} \cong P_{(n-1)(I-1)(J-1)}(1-\mu_I \Delta t)(\lambda_J \Delta t) + P_{(n-1)(I-1)J}(1-\mu_I \Delta t)(\lambda_J \Delta t) + P_{(n-1)IJ}(1-\mu_I \Delta t)(\lambda_J \Delta t), \quad (2.2.13)$$

$$P_{nij} \cong P_{n(i-1)(j-1)}(1-\mu_i \Delta t)(1-\lambda_j \Delta t) \qquad \text{for } i \neq j,\ 1 \le i \le I-1,\ 1 \le j \le J-1, \quad (2.2.14)$$

$$P_{niJ} \cong P_{n(i-1)(J-1)}(1-\mu_i \Delta t)(1-\lambda_J \Delta t) + P_{n(i-1)J}(1-\mu_i \Delta t)(1-\lambda_J \Delta t) \quad \text{for } 1 \le i \le I-1, \quad (2.2.15)$$

$$P_{nIj} \cong P_{n(I-1)(j-1)}(1-\mu_I \Delta t)(1-\lambda_j \Delta t) + P_{nI(j-1)}(1-\mu_I \Delta t)(1-\lambda_j \Delta t) \quad \text{for } 1 \le j \le J-1, \quad (2.2.16)$$

$$P_{nIJ} \cong P_{n(I-1)J}(1-\mu_I \Delta t)(1-\lambda_J \Delta t) + P_{nI(J-1)}(1-\mu_I \Delta t)(1-\lambda_J \Delta t) + P_{nIJ}(1-\mu_I \Delta t)(1-\lambda_J \Delta t). \quad (2.2.17)$$

When the queue length is $n \ge 3$, the values of all the P_{nij} (except P_{ni0}) can be computed using (2.2.14) to (2.2.17) and (2.2.6) to (2.2.8), whereas those of P_{ni0} can be computed using the following equations:

$$P_{n10} \cong \sum_{j=1}^{J-1} P_{(n-1)0j}(1-\mu_1 \Delta t)(\lambda_{j+1} \Delta t) + P_{(n-1)0J}(1-\mu_1 \Delta t)(\lambda_J \Delta t), \quad (2.2.18)$$

$$P_{ni0} \cong \sum_{j=0,\, j \neq i-1}^{J-1} P_{(n-1)(i-1)j}(1-\mu_i \Delta t)(\lambda_{j+1} \Delta t) + P_{(n-1)(i-1)J}(1-\mu_i \Delta t)(\lambda_J \Delta t) \qquad \text{for } 2 \le i \le I-1, \quad (2.2.19)$$

$$P_{nI0} \cong \sum_{j=0,\, j \neq I-1}^{J-1} P_{(n-1)(I-1)j}(1-\mu_I \Delta t)(\lambda_{j+1} \Delta t) + P_{(n-1)(I-1)J}(1-\mu_I \Delta t)(\lambda_J \Delta t) + \sum_{j=0}^{J-1} P_{(n-1)Ij}(1-\mu_I \Delta t)(\lambda_{j+1} \Delta t) + P_{(n-1)IJ}(1-\mu_I \Delta t)(\lambda_J \Delta t). \quad (2.2.20)$$

2.3 Stationary Queue Length Distribution

Before solving (2.2.2) to (2.2.20) in Section 2.2 to obtain the stationary queue length distribution, we may first let c_{ij}, d_{ij}, e_{ij}, f_j and g_{ij} be constants and introduce the following notations:

(a) $P_{n**} = \{ P_{nij}: 0 \le i \le I,\ 0 \le j \le J \}$;

(b) $(P_{m**}, P_{(m+1)**}, P_{(m+2)**})$ denotes the set of equations of the form

$$\sum_{i=0}^{I}\sum_{j=0}^{J} c_{ij} P_{mij} + \sum_{i=0}^{I}\sum_{j=0}^{J} d_{ij} P_{(m+1)ij} + \sum_{i=0}^{I}\sum_{j=0}^{J} e_{ij} P_{(m+2)ij} \cong 0;$$

(c) $(P_{mij} \mid P_{0**}, P_{(m+1)**})$ denotes the equation of the form

$$P_{mij} \cong \sum_{j=0}^{J} f_j P_{00j} + \sum_{i=0}^{I}\sum_{j=0}^{J} g_{ij} P_{(m+1)ij} .$$

With the above notations, (2.2.5) to (2.2.11) in the case when $n = 1$ can be represented as

$(P_{0**}, P_{1**}, P_{2**})$, (2.3.1)

and (2.2.12) to (2.2.17) together with (2.2.6) to (2.2.8) in the case when $n = 2$ may be represented as

$(P_{1**}, P_{2**}, P_{3**})$. (2.3.2)

Furthermore (2.2.18) to (2.2.20) together with (2.2.14) to (2.2.17) and (2.2.6) to (2.2.8) in the case when $n \geq 3$ may be represented as

$(P_{(n-1)**}, P_{n**}, P_{(n+1)**})$. (2.3.3)

It can be shown that from the set of equations given by (2.3.1), we can get

$(P_{1ij} \mid P_{0**}, P_{2**})$ for $0 \leq i \leq I$, $0 \leq j \leq J$. (2.3.4)

By substituting the expression of P_{1ij} given by (2.3.4) into (2.3.2), and solving for P_{2ij}, we get

$(P_{2ij} \mid P_{0**}, P_{3**})$ for $0 \leq i \leq I$, $0 \leq j \leq J$. (2.3.5)

By substituting the expression of P_{2ij} given by (2.3.5) into (2.3.3) when $n = 3$ and solving for P_{3ij}, we get

$(P_{3ij} \mid P_{0**}, P_{4**})$ for $0 \leq i \leq I$, $0 \leq j \leq J$. (2.3.6)

Next for $n \geq 4$, repeat the process of substituting the expression of $P_{(n-1)ij}$ given by

$(P_{(n-1)ij} \mid P_{0**}, P_{n**})$ (2.3.7)

into (2.3.3) and solving for P_{nij} to get

$(P_{nij} \mid P_{0**}, P_{(n+1)**})$ for $0 \leq i \leq I,\ 0 \leq j \leq J.$ (2.3.8)

When $n = N$ is large enough, we may set all the $P_{(n+1)ij}$ in (2.3.8) to be zero and obtain

$(P_{Nij} \mid P_{0**})$ for $0 \leq i \leq I,\ 0 \leq j \leq J.$ (2.3.9)

Substituting (2.3.9) into (2.3.8) when $n = N - 1$, we get

$(P_{(N-1)ij} \mid P_{0**}, P_{N**}) \cong (P_{(N-1)ij} \mid P_{0**})$ for $0 \leq i \leq I,\ 0 \leq j \leq J.$ (2.3.10)

Similarly, for $n = N - 2, N - 3, \ldots, 1$, we may perform the substitution of $(P_{(n+1)ij} \mid P_{0**})$ into (2.3.8) and obtain

$(P_{nij} \mid P_{0**})$ for $0 \leq i \leq I,\ 0 \leq j \leq J.$ (2.3.11)

When $n = 1$, (2.3.11) yields $(P_{1ij} \mid P_{0**})$. By using the results given by $(P_{1ij} \mid P_{0**})$ and (2.2.2) to (2.2.4), we get the following system of J equations:

$(P_{00j} \mid P_{0**})$ for $0 \leq j \leq J.$ (2.3.12)

An inspection of (2.3.12) reveals that among the J equations, only $J - 1$ of them are linearly independent. Hence, we need to include another linearly independent equation so that the resulting system of J equations has a unique solution. Equating the sum of the left sides of the equations given by (2.3.11) to the sum of the right sides of (2.3.11), we get an equation of the form,

$$\sum_{n=1}^{N}\sum_{i}\sum_{j} P_{nij} = \sum_{j} k_j P_{00j} \qquad (2.3.13)$$

where the k_j are constants.

As $\sum_{n=0}^{N}\sum_{i}\sum_{j} P_{nij} \cong 1$, we get from (2.3.13) an equation involving only P_{00j}, $1 \leq j \leq J$. The equation derived from (2.3.13), and $J - 1$ equations chosen from (2.3.12), constitute a system of J equations which can be solved to yield numerical answers for P_{00j}, $1 \leq j \leq J$. Then using (2.3.11), we can get numerical answers for P_{nij} where $n \geq 1$, $0 \leq i \leq I$ and $0 \leq j \leq J$. The stationary probability that the queue length is n can then be obtained as

$$P_n = \sum_{i=0}^{I} \sum_{j=0}^{J} P_{nij} \, . \tag{2.3.14}$$

2.4 Waiting Time Distribution

Suppose a customer arrives at the system which is in the stationary state. Let the time of arrival of the customer be denoted as $t = 0$. Furthermore, let W_q be the time the customer needs to wait before being served and

$$W_q(t) = \mathrm{P}(W_q \leq t),$$

the cumulative distribution function (cdf) of the waiting time W_q.

To find the waiting time distribution $W_q(t)$, we first note that when the system is in the stationary state, an arrival of a customer at time $t = 0$ which is inside an interval τ of length Δt may occur with an approximate probability $\lambda_{j+1}\Delta t$ if the arrival process is in state j at the beginning of the interval τ. Meanwhile at the beginning of τ, the service process may be in state i where $0 \leq i \leq I$. Thus the probability that

(i) the queue length at the beginning of τ is n;

(ii) the service process is in state i at the beginning of τ;

(iii) the arrival process is in state j at the beginning of τ;

(iv) a customer arrives in τ;

is given approximately by

$$P_{nij} \, \lambda_{j+1}\Delta t. \tag{2.4.1}$$

Let $g_{i+1}(t)$ be the pdf of the service time given that the service process is in state i at the beginning of τ, and $g^{(n-1)}(t)$ the $(n-1)$-fold convolution of $g(t)$. The customer who arrives in τ (see (iv)) under the conditions given by (i), (ii) and (iii) above will have a waiting time of zero if $n = 0$, and a waiting time of which the pdf is given by the convolution $g_{i+1}(t) * g^{(n-1)}(t)$ if $n \geq 1$. The cdf $W_q(t)$ is then given approximately by

$$W_q(t) \cong \frac{\sum_{j=1}^{J} P_{00j}\lambda_{j+1}\Delta t + \sum_{n=1}^{N}\sum_{i=0}^{I}\sum_{j=0}^{J} P_{nij}\lambda_{j+1}\Delta t \int_{u=0}^{t} g_{i+1}(u) * g^{(n-1)}(u)du}{\sum_{j=1}^{J} P_{00j}\lambda_{j+1}\Delta t + \sum_{n=1}^{N}\sum_{i=0}^{I}\sum_{j=0}^{J} P_{nij}\lambda_{j+1}\Delta t}. \qquad (2.4.2)$$

The cdf $W_q(t)$ may also be computed approximately by a simulation procedure described below.

Suppose a customer arrives at time $t = 0$ and the next m-th customer arrives at time $t = \sum_{k=1}^{m} A_k$ where $A_1, A_2, \ldots$ are independent and identically distributed with pdf $f(t)$. Next let the service time of the next m-th customer be B_m of which $B_0, B_1, \ldots$ are independent and identically distributed with pdf $g(t)$. For a chosen large integer M, the value of $\underset{\sim}{v} = \{(0, B_0), (A_1, B_1), \ldots, (A_M, B_M)\}$ is generated and the following waiting times are obtained:

$$W_{q,0} = 0,$$

$$W_{q,m} = \begin{cases} 0, & \text{if } W_{q,m-1} + B_{m-1} < A_m, \\ W_{q,m-1} + B_{m-1} - A_m, & \text{if } W_{q,m-1} + B_{m-1} \geq A_m,\ 1 \leq m \leq M-1, \end{cases}$$

where $W_{q,m}$ is the waiting time of the m-th customer. Then

$W_q(t) \cong$ (Number of the $W_{q,m}$ which are less than t)/ M.

2.5 Numerical Examples

Let $\text{Gamma}(\kappa, \theta)$ denote a gamma distribution of which κ is the shape parameter and θ the scale parameter. The related probability density function is then given by $f(x;\kappa,\theta) = \left(x^{\kappa-1}e^{-x/\theta}\right) \big/ \left(\theta^{\kappa}\Gamma(\kappa)\right)$. Consider an example in which the service time (S) has a gamma distribution with the parameter vector $(\kappa_1, \theta_1) = (1.5, 2)$, and the interarrival time (T) has another gamma distribution with the parameter vector $(\kappa_2, \theta_2) = (3.1, 2)$. The utilization factor will then be $\rho = \text{E}(S)/\text{E}(T) = 0.48$. The reason for considering gamma distribution (κ, θ) with fractional values of the shape parameter κ is that the term $t^{\kappa-1}$ appearing in the pdf $f(t)$ and $g(t)$ will usually make the existing analytical methods for finding queue length distribution fail. The reason behind such failure is that when we set $t = x + y$, the function $(x+y)^{\kappa-1}$ cannot be expressed as a finite sum of products of a function of x alone and a function of y alone.

The following is a procedure to find the values of Δt and I (or J). Initially we find the value of T such that the rates at time $t \geq T$ exhibit small variations. A small fractional value (for example, 0.1 or 0.05) is assigned to Δt and I is then obtained as the integer which is approximately equal to $T/\Delta t$. If I is very large (for example, $I > 1000$), then a bigger unit is chosen for t until $I \leq 1000$. It can be shown that when $\Delta t = 0.04$, suitable values of I and J are respectively $I = 550$ and $J = 550$. By using the proposed numerical method, the stationary queue length distribution is found. The stationary queue length distribution may also be computed using the simulation procedure in the software "QtsPlus" (accompanying software for Gross and Harris [74]) when the number of runs is $M_1 = 10^7$. The results obtained are shown in Table 2.5.1.

Table 2.5.1

Comparison of stationary queue length distribution computed from the proposed numerical method, and those obtained from the software "QtsPlus"
[$(\kappa_1, \theta_1) = (1.5, 2)$, $(\kappa_2, \theta_2) = (3.1, 2)$, $\Delta t = 0.04$].

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation (QtsPlus)
0	0.518854	0.516033
1	0.366244	0.366475
2	0.089790	0.091326
3	0.019748	0.020488
4	0.004227	0.004444
5	8.97E-04	9.48E-04
6	1.90E-04	2.15E-04
7	4.01E-05	5.44E-05
8	8.47E-06	1.33E-05
9	1.79E-06	3.43E-06
10	3.78E-07	1.79E-07
…	…	…
20	6.67E-14	0

From Table 2.5.1, we see that the stationary queue length distribution obtained using the proposed numerical method is close to that obtained from the software "QtsPlus".

Figures 2.5.1–2.5.8 show the stationary queue length probabilities found by the numerical method using various other values of Δt. The dotted lines in the figures give the extrapolated values based on polynomials of low degrees fitted to the values (represented by the symbol "•") of $(P_n, \Delta t)$. The y-values in the dotted lines when the x-values are zero will represent the final results based on the numerical method for the queue length probabilities. The plots given in Figures 2.5.1, 2.5.3 and 2.5.4 show that the final results based on the numerical method agree quite well with the results based on "QtsPlus". Meanwhile the plot given in Figure 2.5.2 indicates that the result based on numerical method would be more accurate than that found by simulation. The plots for P_n against Δt for $n = 4$, 5, 6 and 7 (Figures 2.5.5, 2.5.6, 2.5.7 and 2.5.8) indicate that only the final result for P_4 based on simulation agrees quite well with that based on the numerical method.

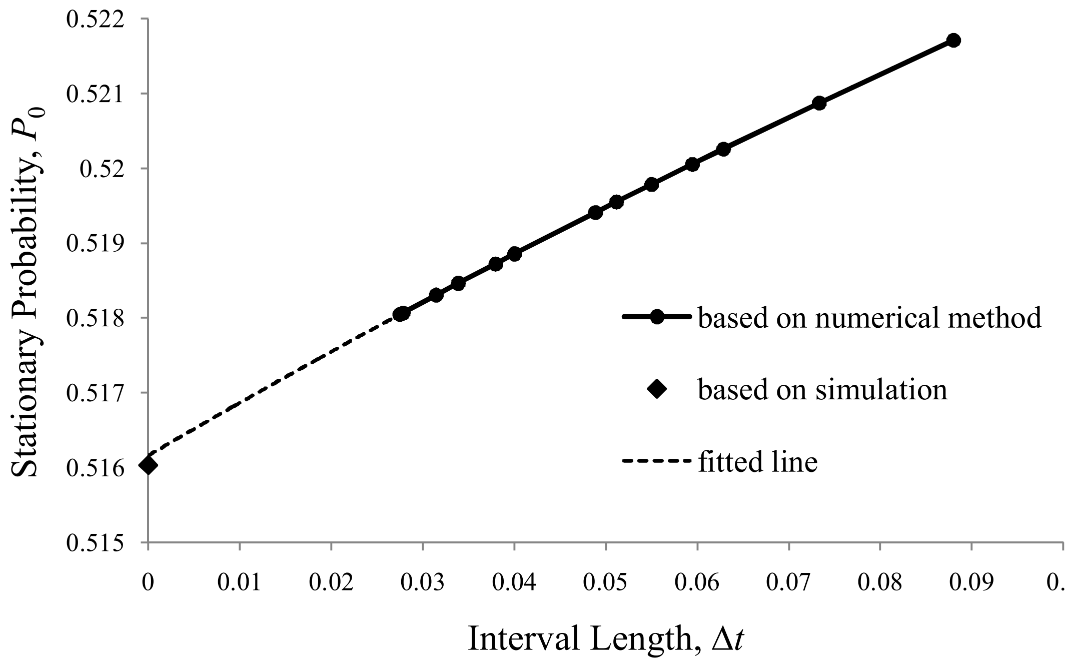

Figure 2.5.1 Stationary probability that queue length is $n = 0$ [(κ_1, θ_1) = (1.5, 2), (κ_2, θ_2) = (3.1, 2)].

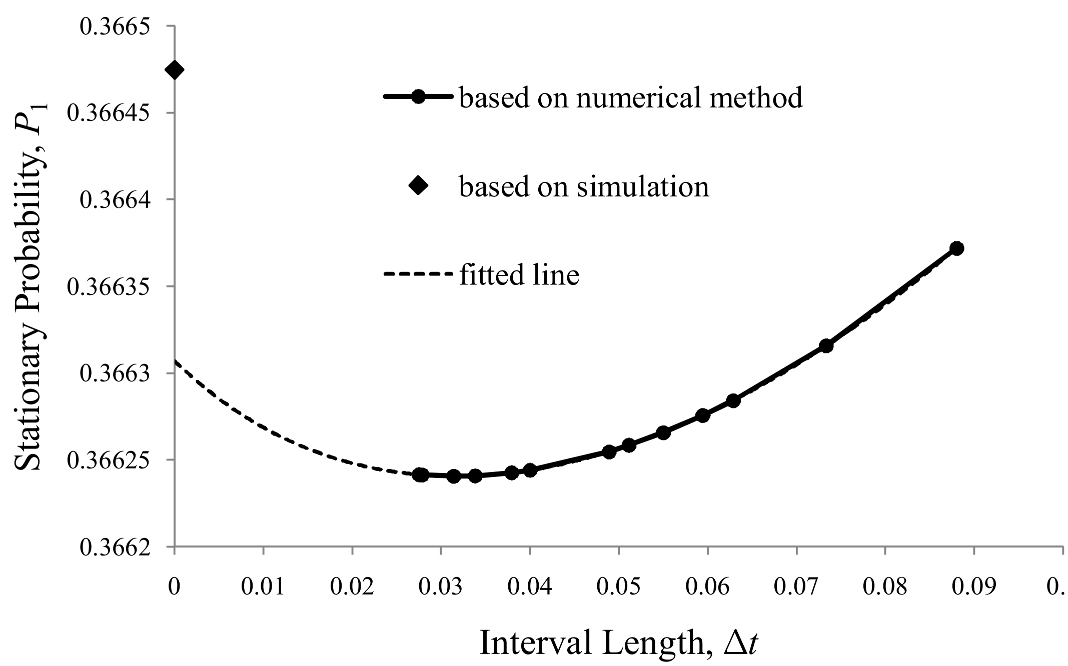

Figure 2.5.2 Stationary probability that queue length is $n = 1$ [(κ_1, θ_1) = (1.5, 2), (κ_2, θ_2) = (3.1, 2)].

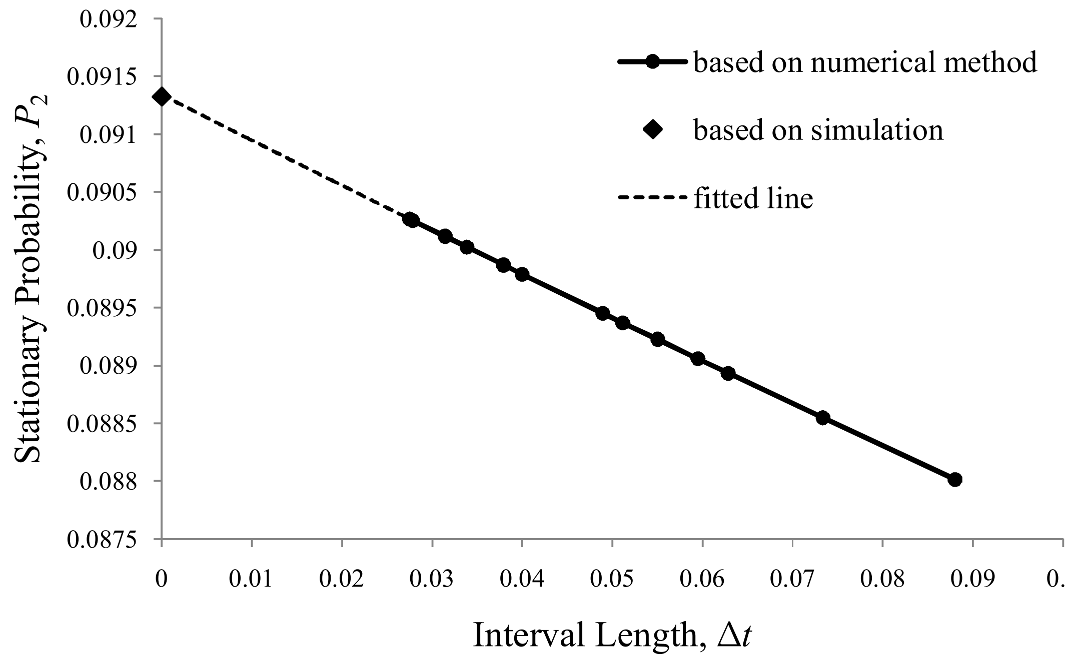

Figure 2.5.3 Stationary probability that queue length is $n = 2$ $[(\kappa_1, \theta_1) = (1.5, 2), (\kappa_2, \theta_2) = (3.1, 2)]$.

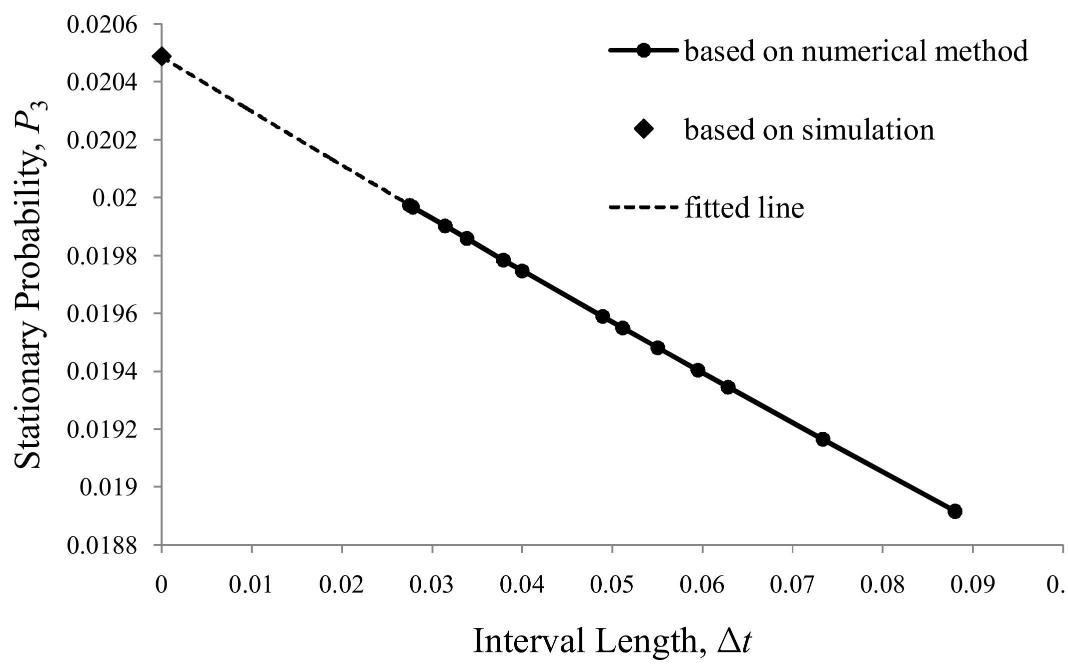

Figure 2.5.4 Stationary probability that queue length is $n = 3$ $[(\kappa_1, \theta_1) = (1.5, 2), (\kappa_2, \theta_2) = (3.1, 2)]$.

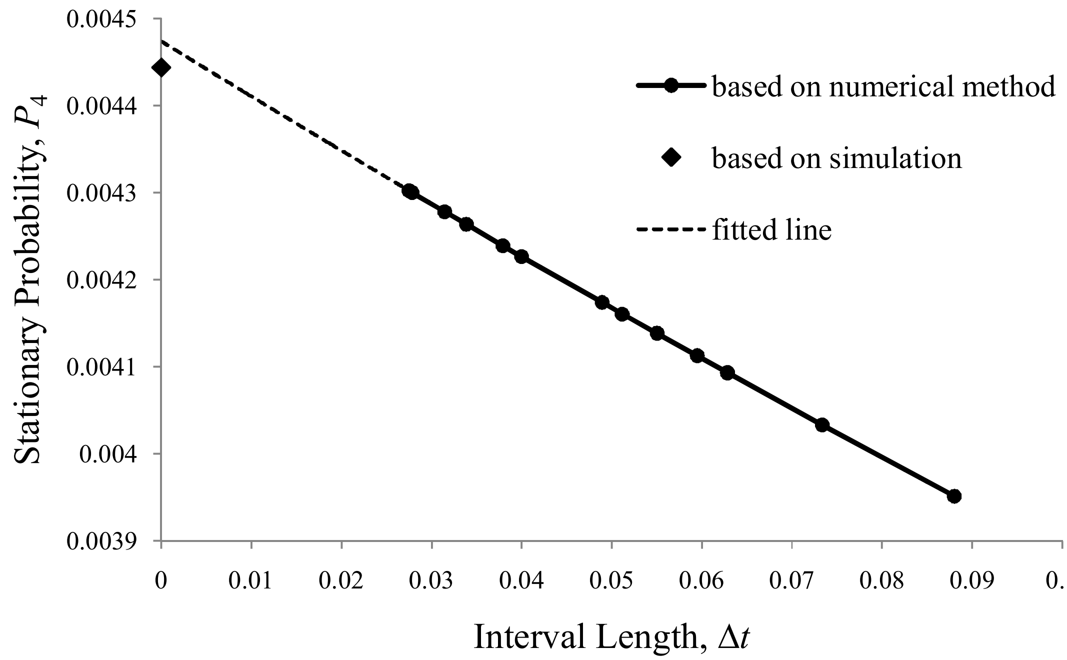

Figure 2.5.5 Stationary probability that queue length is $n = 4$ $[(\kappa_1, \theta_1) = (1.5, 2), (\kappa_2, \theta_2) = (3.1, 2)]$.

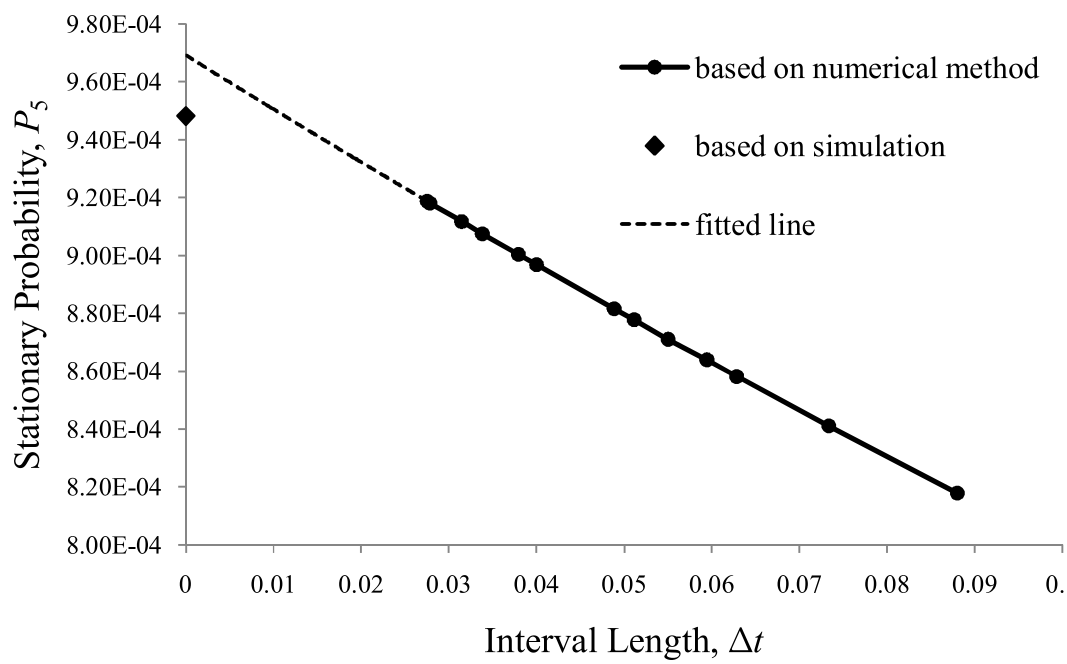

Figure 2.5.6 Stationary probability that queue length is $n = 5$ $[(\kappa_1, \theta_1) = (1.5, 2), (\kappa_2, \theta_2) = (3.1, 2)]$.

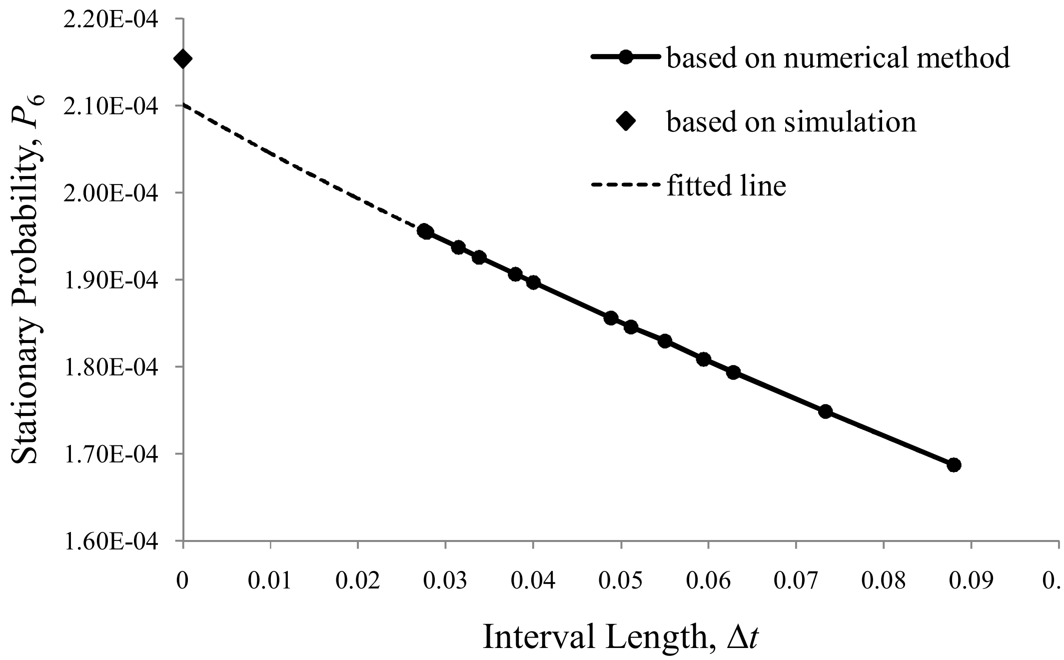

Figure 2.5.7 Stationary probability that queue length is $n = 6$ $[(\kappa_1, \theta_1) = (1.5, 2), (\kappa_2, \theta_2) = (3.1, 2)]$.

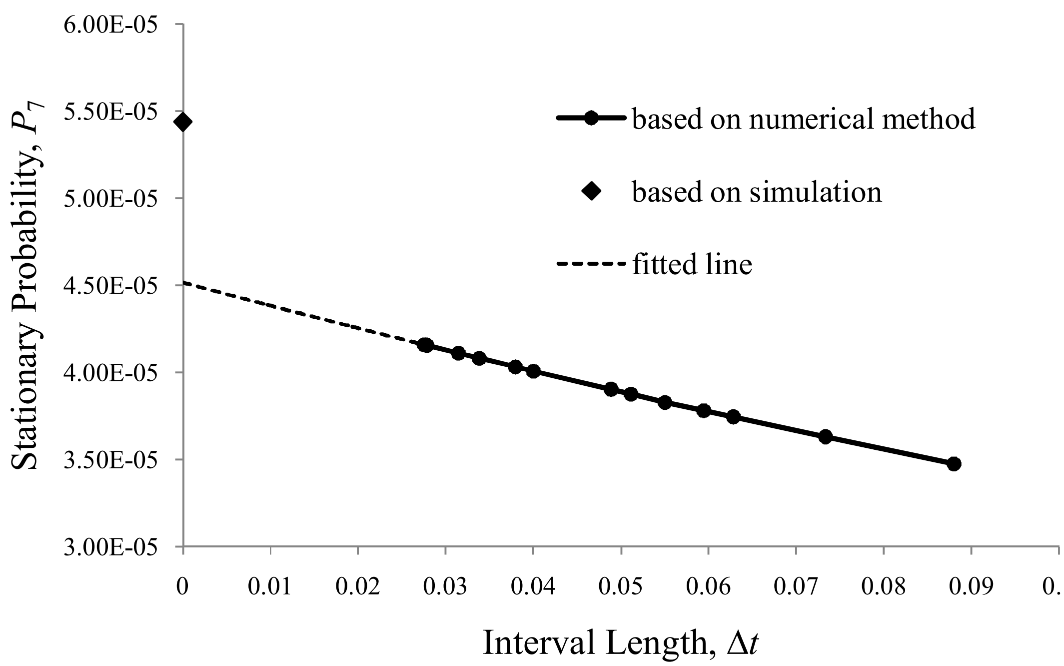

Figure 2.5.8 Stationary probability that queue length is $n = 7$ $[(\kappa_1, \theta_1) = (1.5, 2), (\kappa_2, \theta_2) = (3.1, 2)]$.

Table 2.5.2 shows that the stationary waiting time distribution obtained by using the numerical method in Section 2.4 is close to that obtained by the simulation procedure.

Table 2.5.2

Comparison of stationary waiting time distribution computed respectively by using the proposed numerical method, and the simulation procedure (Δt = 0.04).

Time, t	$P(W_q \le t)$	
	Numerical method	Simulation
0	0.716771	0.717751
0.04	0.720011	0.719061
0.08	0.723321	0.721774
0.12	0.726602	0.725983
0.16	0.729799	0.728072
0.20	0.732959	0.731658
0.24	0.735939	0.734215
0.28	0.738872	0.738393
0.32	0.741849	0.740451
0.36	0.744914	0.742415
…	…	…
20	0.999523	0.999638

2.6 Discrete Time GI/G/1 Queue

The stationary queue length and stationary waiting time distributions of a discrete time GI/G/1 can also be found by using the proposed numerical method in Sections 2.2 to 2.4 after some modifications of the equations for the stationary probabilities given in Section 2.2. An explanation of why the above modifications are necessary is as follows.

First we note that the values of the μ_k (or λ_k) for the discrete service time (or arrival time) distribution is such that all the μ_k (or λ_k) are zero except for the cases when $k\Delta t$ coincides with the service time (or arrival time) which has a nonzero probability of occurrence. Let the values of such k be denoted by k_1, k_2,…, k_d. The value of a typical λ_{k_i} will be such that $\lambda_{k_i}\Delta t$ is a constant. This means that when Δt is made very small, the value of λ_{k_i} will have to be inflated correspondingly. Thus, if the system is not empty at time t, the simultaneous occurrence of the events that

(A) a customer arrives within the interval $(t, t + \Delta t]$; and

(B) a service is completed within the interval $(t, t + \Delta t]$;

may not tend to zero when Δt tends to zero. Thus the equations for stationary probabilities given in Section 2.2 need to be modified by taking into account of the simultaneous occurrence of events (A) and (B). The modified version of the equations in Section 2.2 is as follows.

When the queue length is $n = 0$, the values of the P_{nij} can be found from (2.2.2)–(2.2.4). When the queue length is $n = 1$, the expressions for P_{nij}, $1 \le j \le J$, $i \le j$ are the same as those given by (2.2.9)–(2.2.11), whereas P_{n0j} can be computed from the equations below:

$$\begin{aligned} P_{n00} &\cong \sum_{j=1}^{J-1} P_{00j}(\lambda_{j+1}\Delta t) + P_{00J}(\lambda_J \Delta t) + \sum_{i=0}^{I-1}\sum_{j\ge i}^{J-1} P_{nij}(\mu_{i+1}\Delta t)(\lambda_{j+1}\Delta t) \\ &+ \sum_{i=0}^{I-1} P_{niJ}(\mu_{i+1}\Delta t)(\lambda_J \Delta t) + P_{nIJ}(\mu_I \Delta t)(\lambda_J \Delta t) \end{aligned}, \qquad (2.6.1)$$

$$\begin{aligned} P_{n0j} &\cong \sum_{i=0}^{I-1} P_{(n+1)i(j-1)}(\mu_{i+1}\Delta t)(1-\lambda_j \Delta t) \\ &+ P_{(n+1)I(j-1)}(\mu_I \Delta t)(1-\lambda_j \Delta t) \end{aligned} \qquad \text{for } 1 \le j \le J-1, \qquad (2.6.2)$$

$$\begin{aligned} P_{n0J} &\cong \sum_{i=0}^{I-1} P_{(n+1)i(J-1)}(\mu_{i+1}\Delta t)(1-\lambda_J \Delta t) + P_{(n+1)I(J-1)}(\mu_I \Delta t)(1-\lambda_J \Delta t) \\ &+ \sum_{i=0}^{I-1} P_{(n+1)iJ}(\mu_{i+1}\Delta t)(1-\lambda_J \Delta t) + P_{(n+1)IJ}(\mu_I \Delta t)(1-\lambda_J \Delta t) \end{aligned}. \qquad (2.6.3)$$

When $n = 2$, the values of the P_{n0j}, $1 \le j \le J$ can be computed using (2.6.2)–(2.6.3), while (2.2.12)–(2.2.13) can be used to find the values of P_{ni0}, $1 \le i \le I$. The values of the P_{niJ} and P_{nIj} can be obtained from (2.2.15) and (2.2.16) respectively. All the other values of P_{nij} can be computed using the following equations.

$$P_{n00} \cong \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} P_{nij}(\mu_{i+1}\Delta t)(\lambda_{j+1}\Delta t) + \sum_{i=0}^{I-1} P_{niJ}(\mu_{i+1}\Delta t)(\lambda_J \Delta t)$$
$$+ \sum_{j=0}^{J-1} P_{nIj}(\mu_I \Delta t)(\lambda_{j+1}\Delta t) + P_{nIJ}(\mu_I \Delta t)(\lambda_J \Delta t) \quad , \qquad (2.6.4)$$

$$P_{nij} \cong P_{n(i-1)(j-1)}(1-\mu_i \Delta t)(1-\lambda_j \Delta t) \qquad \text{for } 1 \le i \le I\text{–}1,\ 1 \le j \le J\text{–}1, \qquad (2.6.5)$$

$$P_{nIJ} \cong P_{n(I-1)(J-1)}(1-\mu_I \Delta t)(1-\lambda_J \Delta t) + P_{n(I-1)J}(1-\mu_I \Delta t)(1-\lambda_J \Delta t)$$
$$+ P_{nI(J-1)}(1-\mu_I \Delta t)(1-\lambda_J \Delta t) + P_{nIJ}(1-\mu_I \Delta t)(1-\lambda_J \Delta t) \quad . \qquad (2.6.6)$$

For $n \ge 3$, the values of all the P_{nij} (except P_{ni0}) can be computed using (2.2.15)–(2.2.16) and (2.6.2)–(2.6.6), whereas those of P_{ni0} can be computed using the following equations.

$$P_{ni0} \cong \sum_{j=0}^{J-1} P_{(n-1)(i-1)j}(1-\mu_i \Delta t)(\lambda_{j+1}\Delta t) + P_{(n-1)(i-1)J}(1-\mu_i \Delta t)(\lambda_J \Delta t) \qquad \text{for } 1 \le i \le I\text{–}1, \qquad (2.6.7)$$

$$P_{nI0} \cong \sum_{j=0}^{J-1} P_{(n-1)(I-1)j}(1-\mu_I \Delta t)(\lambda_{j+1}\Delta t) + P_{(n-1)(I-1)J}(1-\mu_I \Delta t)(\lambda_J \Delta t)$$
$$+ \sum_{j=0}^{J-1} P_{(n-1)Ij}(1-\mu_I \Delta t)(\lambda_{j+1}\Delta t) + P_{(n-1)IJ}(1-\mu_I \Delta t)(\lambda_J \Delta t) \quad . \qquad (2.6.8)$$

We may solve the above equations by using the proposed numerical method in Section 2.3 to obtain all the values of P_{nij} and hence the stationary queue length distribution. From the values of the stationary probabilities, we can find the stationary waiting time distribution by using the method proposed in Section 2.4. The cdf $W_q(t)$ for the discrete time GI/G/1 queue is now given approximately by

$$W_q(t) \cong \frac{\left(\begin{array}{l} \sum_{j=1}^{J} P_{00j}\lambda_{j+1}\Delta t + \sum_{n=1}^{N}\sum_{i=0}^{I}\sum_{j=0}^{J} P_{nij}(\lambda_{j+1}\Delta t)(1-\mu_{i+1}\Delta t)\int_{u=0}^{t} g_{i+1}(u) * g^{(n-1)}(u)\,du \\ + \sum_{n=1}^{N}\sum_{i=0}^{I}\sum_{j=0}^{J} P_{nij}(\lambda_{j+1}\Delta t)(\mu_{i+1}\Delta t)\int_{u=0}^{t} g^{(n-1)}(u)\,du \end{array} \right)}{\sum_{j=1}^{J} P_{00j}\lambda_{j+1}\Delta t + \sum_{n=1}^{N}\sum_{i=0}^{I}\sum_{j=0}^{J} P_{nij}\lambda_{j+1}\Delta t} \quad . \qquad (2.6.9)$$

Table 2.6.1 shows the results of the stationary queue length distribution computed using the proposed numerical method. The table also shows the results given in [13] where the authors found the stationary queue length distribution from the sojourn time distribution using the distributional Little's law. The functions $T(z)$ and $S(z)$ appearing in Tables 2.6.1–2.6.3 and 2.6.5–2.6.6 are respectively the probability generating functions of the discrete service time and interarrival time.

Table 2.6.1

Comparison of stationary queue length distribution computed using the proposed numerical method, and that given in Kim & Chaudhry [13].

Queue Length, n	Example 1		
	P(Queue Length = n) [$T(z) = z/10 + 3z^2/10 + 2z^3/5 + z^4/5$, $S(z) = 3z/10 + 3z^2/5 + z^3/10$]		
	Numerical method ($\Delta t = 1.0$)	Numerical method ($\Delta t = 0.1$)	Kim & Chaudhry [13]
0	0.333333	0.333333	0.333333
1	0.596799	0.596799	0.596799
2	0.067034	0.067034	0.067034
3	0.002728	0.002728	0.002728
4	0.000101	0.000101	0.000101
5	3.81E-06	3.81E-06	0.000004
6	1.43E-07	1.43E-07	0
7	5.38E-09	5.38E-09	0
8	2.02E-10	2.02E-10	0
9	7.60E-12	7.60E-12	0
10	2.86E-13	2.86E-13	0
…	…	…	…

From Table 2.6.1, we can see that the queue length probabilites obtained by using the proposed numerical method is close to that given in [13]. When $\Delta t \leq 1$, the values of the μ_k (or λ_k) are able to capture all the details of the discrete distribution of the service time (or arrival time). Thus the results given in columns 2 and 3 in Table 2.6.1 are identical.

Tables 2.6.2 and 2.6.3 show the stationary queue length distribution in three other examples of discrete queue.

Table 2.6.2

Comparison of stationary queue length distribution computed using the proposed numerical method, and that given in Kim & Chaudhry [13] ($\Delta t = 1.0$).

Queue Length, n	Example 2 P(Queue Length = n) [$T(z) = z(z/2 + 1/2)^{38}$, $S(z) = (z + z^2 + \ldots + z^{35})/35$] Numerical method	Kim & Chaudhry [13]
0	0.100000	0.100000
1	0.323533	0.323533
2	0.291648	0.291648
3	0.146160	0.146160
4	0.071163	0.071163
5	0.034640	0.034641
6	0.016862	0.016862
7	0.008208	0.008208
8	0.003995	0.003995
9	0.001945	0.001945
10	9.47E-04	9.47E-04
…	…	…

Table 2.6.3

Comparison of stationary queue length distributions computed using the proposed numerical method, and that given in Kim & Chaudhry [13] ($\Delta t = 1.0$).

Queue Length, n	Example 3 P(Queue Length = n) [$T(z) = z(z/2 + 1/2)^{38}$, $S(z) = (z + z^2 + z^3)/3$]		Example 4 P(Queue Length = n) [$T(z) = z(z/2 + 1/2)^{38}$, $S(z) = (z + z^2 + \ldots + z^{19})/19$]	
	Numerical method	Kim & Chaudhry [13]	Numerical method	Kim & Chaudhry [13]
0	0.900000	0.900000	0.500000	0.500000
1	0.100000	0.100000	0.494811	0.494811
2	2.49E-12	2.49E-12	0.005189	0.005189
3	2.21E-25	2.21E-25	1.53E-08	1.53E-08
4	1.52E-46	0	2.19E-14	2.19E-14
5	3.54E-60	0	3.13E-20	2.75E-20
6	…	…	…	…

Tables 2.6.2 and 2.6.3 show that the results obtained by using the proposed numerical method are very close to those given in [13].

From the stationary probabilities, the stationary waiting time distributions can be obtained using (2.6.9), the results obtained are shown in Table 2.6.4.

Table 2.6.4

Stationary waiting time distributions computed by using the proposed numerical method ($\Delta t = 1.0$).

Time, t	$P(W_q \leq t)$			
	Example 1	Example 2	Example 3	Example 4
0	0.847762	0.242019	1	0.955516
1	0.964654	0.259177	1	0.972644
2	0.992963	0.276887	1	0.984262
3	0.998538	0.294817	1	0.991529
4	0.999699	0.313157	1	0.995754
5	0.999938	0.331907	1	0.998002
6	0.999987	0.350904	1	0.999113
7	0.999997	0.370697	1	0.999619
8	0.999999	0.390268	1	0.999842
9	1	0.410278	1	0.999933
…	…	…	…	…

For a customer who arrives at time $t = 0$, his sojourn time is equal to the sum of his waiting time and service time. Thus from the waiting time and service time distributions of the incoming customer, we can compute his sojourn time distribution. Tables 2.6.5 and 2.6.6 show the results of the stationary sojourn time distribution computed using the proposed numerical method and those given in [13].

Table 2.6.5

Comparison of stationary sojourn time distributions computed by using the proposed numerical method, and that given in Kim & Chaudhry [13] (Δt = 1.0).

Time, t	Example 1: Sojourn Time Distribution [$T(z) = z/10 + 3z^2/10 + 2z^3/5 + z^4/5$, $S(z) = 3z/10 + 3z^2/5 + z^3/10$]		Example 2: Sojourn Time Distribution [$T(z) = z(z/2 + 1/2)^{38}$, $S(z) = (z + z^2 + \ldots + z^{35})/35$]	
	Numerical method	Kim & Chaudhry [13]	Numerical method	Kim & Chaudhry [13]
0	0	0	0	0
1	0.254329	0.254329	0.006915	0.006915
2	0.543725	0.543708	0.007410	0.007408
3	0.163404	0.163375	0.007911	0.007913
4	0.030347	0.030357	0.008432	0.008429
5	0.006524	0.006540	0.008968	0.008956
6	0.001326	0.001339	0.009503	0.009493
7	2.74E-04	0.000279	0.010059	0.010039
8	5.66E-05	0.000058	0.010610	0.010595
9	1.16E-05	0.000012	0.011178	0.011160
10	2.40E-06	0.000002	0.011746	0.011731
…	…	…	…	…

Table 2.6.6

Comparison of stationary sojourn time distributions computed using the proposed numerical method, and that given in Kim & Chaudhry [13] (Δt = 1.0).

Time, t	Example 3: Sojourn Time Distribution [$T(z) = z(z/2 + 1/2)^{38}$, $S(z) = (z + z^2 + z^3)/3$]		Example 4: Sojourn Time Distribution [$T(z) = z(z/2 + 1/2)^{38}$, $S(z) = (z + z^2 + \ldots + z^{19})/19$]	
	Numerical method	Kim & Chaudhry [13]	Numerical method	Kim & Chaudhry [13]
0	0	0	0	0
1	0.333333	0.333333	0.050290	0.050290
2	0.333333	0.333333	0.051192	0.051193
3	0.333333	0.333333	0.051803	0.051804
4	1.62E-11	1.62E-11	0.052186	0.052187
5	3.97E-13	4.04E-13	0.052408	0.052408
6	…	…	…	…

From Tables 2.6.5 and 2.6.6, we can see that the stationary sojourn time distributions computed by using the proposed numerical method are very close to those given in [13].

2.7 Conclusion

Most of the existing methods in the literatures find the queue length distribution in the GI/G/1 queue via the waiting time distribution. On the contrary, the present proposed method finds the queue length distribution directly for the CAR/CAR/1 queue. The accuracy of the numerical results for the distribution can be greatly improved by an extrapolation process. Furthermore the queue length distribution thus found can later be used to find the waiting time distribution. The main drawback of the proposed method is that we may encounter dimensionality problem when I(or J) is very large.

The method proposed in this chapter may also be applied to other queueing models. For example, in Chapters 3 and 4, it is applied to the queueing systems which are deteriorated by random shocks.

CHAPTER 3

MAINTENANCE OF A DETERIORATING QUEUE WITH NON-POISSON ARRIVALS

3.1 Introduction

Consider the model in [35] in which the service rate of a multi-state M/M/1 queue would deteriorate due to random shocks. In their model, it is assumed that the shocks arrive at the system according to a Poisson process with random magnitudes. The server is repaired when its state is above a specified maintenance level. In this chapter, the distribution of the customer's interarrival time in their model is changed to a CAR distribution while the service time remains exponentially distributed. The numerical method proposed in Section 2.2 is adapted for deriving the set of equations for the stationary probabilities of the queue length and the states of the arrival, service and repair processes. The stationary probabilities obtained can be used to find

(A) the sojourn time distribution of a customer who arrives when the queue is in a stationary state; and

(B) the expected length of the duration between two successive repair completions when the queue is in a stationary state.

The results in (A) and (B) can next be used to compute the average cost of the system and find the maintenance level such that the average cost is minimized.

3.2 Notations and Assumptions

The following notations are used throughout Chapter 3:

β largest possible service state

α maintenance level for the system, $\alpha \le \beta$

μ_i service rate in state i of the service process

δ_r repair rate in state r of the repair process

λ_j arrival rate in state j of the arrival process

γ shock rate

n number of customers in the system

g_x probability that the random amount of the shock is x

τ_k interval given by $((k-1)\Delta t, k\Delta t]$, $k = 0, 1, 2, \ldots$

n_k queue length of the system at the end of τ_k

ξ_k state number of the service process at the end of τ_k, $\xi_k \in \{1, 2, 3, ..., \beta\}$

φ_k state number of the repair process at the end of τ_k, $\varphi_k \in \{0, \alpha, \alpha + 1, ..., \beta\}$

ψ_k state number of the arrival process at the end of τ_k, $\psi_k \in \{0, 1, 2, ..., J\}$

$P_{nirj}^{(k)}$ the probability that at the end of τ_k,

(a) the number of customers in the system is n (including the customer that is being served);

(b) the service process is in state i;

(c) the repair process is in state r; and

(d) the arrival process is in state j.

Assumptions:

1. Service state indexes are ordered. State 1 is the best state with the largest service rate, state β is the worst.
2. Repair is performed immediately on the system when the service state exceeds $\alpha - 1$.
3. Each successful repair brings the service state back to state 1.
4. $P_{nirj} = \lim_{k \to \infty} P_{nirj}^{(k)}$ exists.

3.3 A Model for Deteriorating Single Server Queue

Consider the following multi-state M/M/1 queue studied in [35]. The server would deteriorate when the system is subject to random shocks. When the service state is i, the service rate is denoted as μ_i where $1 \le i \le \beta$, with $\mu_i > \mu_j$ for $i < j$. The server is initially in state 1. It is assumed that shocks arrive at the system according to a Poisson process with rate γ. A shock increases the current service state i to a new value given by $\min(i + x, \beta)$ where x is random and having a probability distribution given by g_x, $x = 1, 2, 3, \ldots$. The adopted preventive maintenance policy requires the server to be repaired when the service state i exceeds $\alpha - 1$ where $\alpha \le \beta$, and it is assumed that repair rate is δ_r where $r = i$ is defined to be the repair state. The server does not provide service to the customers during a repair.

In this chapter, the distribution of the interarrival time is changed to one which has a constant asymptotic rate (CAR), and the resulting queue is denoted as a CAR/M/1 queue. With the change in the interarrival time distribution, the model can be applied to the system where the assumption of the Poisson arrival process is violated. Yang et al. [35] used the matrix-geometric approach developed by Neuts [14] to derive the basic characteristics of the multi-state M/M/1 queue. In this chapter, in which the interarrival time has a CAR distribution while the service time still has an exponential distribution, a numerical procedure is used instead to derive the basic characteristics.

3.4 Derivation of Equations for the Stationary Probabilities

Let $f(t)$ be the probability density function (pdf) of the interarrival time of the customers. The rate of the interarrival time distribution evaluated at $t = k\Delta t$ is then given by

$$\lambda_k = \frac{f(k\Delta t)}{\int_{k\Delta t}^{\infty} f(u)du}.$$

When the interarrival time has a CAR distribution, we may assume that there is a large positive integer J such that

$$\lambda_J \cong \lim_{k\to\infty} \lambda_k .$$

Suppose a customer has arrived at time $t = 0$. Then the next customer will arrive in the interval τ_1 with an approximate probability $\lambda_1\Delta t$, and given that the next customer does not arrive in the intervals $\tau_1, \tau_2, \ldots, \tau_{k-1}$, the probability that he/she will arrive in τ_k will be approximately $\lambda_k\Delta t$ for $k = 2, 3, 4, \ldots$ where $\lambda_k = \lambda_J$ for $k \geq J$.

Given that a customer arrives at a time in the interval $\tau_0 = (-\Delta t, 0]$, we may define the state number ψ_k of the arrival process at the end of the interval $\tau_k = ((k-1)\Delta t, k\Delta t]$ as

$$\psi_k = \begin{cases} 0, & \text{if } k = 0 \text{ or the next customer arrives in } \tau_k,\ k \geq 1. \\ \min(k, J), & \text{if the next customer does not arrive in } \tau_k,\ k \geq 1. \end{cases}$$

We next define the state number of the service process at the end of τ_k as

$$\xi_k = \begin{cases} 1, & \text{if } k = 0 \text{ or a repair is completed in } \tau_k,\ k \geq 2. \\ i, & \text{if the service state is } i \text{ at the end of } \tau_{k-1} \text{ for } 1 \leq i < \alpha, \text{ and no shocks occur in } \tau_k,\ k \geq 1. \\ r, & \text{if the service state is } r \text{ at the end of } \tau_{k-1} \text{ for } \alpha \leq r \leq \beta, \text{ and no repair completions occur in } \tau_k,\ k \geq 1. \\ \min(i + x, \beta), & \text{if the service state is } i \text{ at the end of } \tau_{k-1} \text{ for } 1 \leq i < \alpha, \text{ and a shock with magnitude } x \text{ occurs in } \tau_k,\ k \geq 1. \end{cases}$$

The state number of the repair process at the end of τ_k is defined as

$$\varphi_k = \begin{cases} 0, & \text{if} \\ & \bullet\ k = 0; \text{ or} \\ & \bullet\ \xi_{k-1} < \alpha \text{ at the end of } \tau_{k-1}, \text{ and no shocks occur in } \tau_k,\ k \geq 1; \text{ or} \\ & \bullet\ \text{the service state is } i \text{ at the end of } \tau_{k-1} \text{ for } 1 \leq i < \alpha - 1, \text{ and a shock with magnitude } x \text{ occurs in } \tau_k \text{ for } i + x < \alpha,\ k \geq 1. \\ \min(i + x, \beta), & \text{if the service state is } i \text{ at the end of interval } \tau_{k-1} \text{ for } 1 \leq i < \alpha, \text{ and a shock with magnitude } x \text{ occurs in } \tau_k \text{ for } i + x \geq \alpha,\ k \geq 1. \\ r, & \text{if the repair state is } r \text{ at the end of } \tau_{k-1} \text{ for } \alpha \leq r \leq \beta, \text{ and no repair completions occur in } \tau_k,\ k \geq 2. \end{cases}$$

Let n_k be the queue length at the end of τ_k and $\boldsymbol{h}_k = (n_k, \xi_k, \varphi_k, \psi_k)$. We may refer to $\boldsymbol{h}_k$ as the vector of characteristics of the queue at the end of τ_k.

Let $P_{nirj}^{(k)}$ be the probability that at the end of τ_k, the number of customers in the system is n (including the customer that is being served), the service process is in state i, the repair process is in state r and the arrival process is in state j. Assume that

$$P_{nirj} = \lim_{k\to\infty} P_{nirj}^{(k)}$$

exists. To find the P_{nirj}, we first make the following observations.

Suppose at the end of τ_{k-1}, the queue length n is not empty (i.e. $n_{k-1} = n \geq 1$), the server is in state $i < \alpha$ (i.e. $\xi_{k-1} = i < \alpha$) and the arrival process is in state $j - 1$ (i.e. $\psi_{k-1} = j - 1$). In this case the server is still active and we define the repair state number to be zero (i.e. $\varphi_{k-1} = r = 0$). This means the vector of characteristics at the end of τ_{k-1} is given by $\boldsymbol{h}_{k-1} = (n, i, 0, j - 1)$. With this value of $\boldsymbol{h}_{k-1}$, only one of the following events can occur in τ_k:

(a) A customer enters the system with the arrival rate λ_{j^*}, and at the end of τ_k, the vector of characteristics becomes $\boldsymbol{h}_k = (n + 1, i, 0, 0)$;

(b) A customer leaves the system with the departure rate μ_i, and $\boldsymbol{h}_k = (n - 1, i, 0, j^*)$;

(c) A shock with magnitude x occurs and deteriorates the service state to $i^* = \min(i + x, \beta)$, yielding $\boldsymbol{h}_k = (n, i^*, r^*, j^*)$;

(d) No customers enter or leave the system, and no shocks arrive, yielding $\boldsymbol{h}_k = (n, i, 0, j^*)$;

where $j^* = \min(j, J)$, and

$$r^* = \begin{cases} 0, & \text{if } 1 < i + x < \alpha, \\ \min(i + x, \beta), & \text{if } i + x \geq \alpha, \end{cases} \qquad \text{for } x \geq 1.$$

However if at the end of τ_{k-1}, the system is empty (i.e. $n_{k-1} = 0$), the state number i of the idle server is less than α and the arrival process is in state $j - 1$, then one of the following events can occur in τ_k:

(e) A customer enters the system with arrival rate λ_{j^*} and $\boldsymbol{h}_k = (1, i, 0, 0)$;

(f) A shock with magnitude x occurs and deteriorates the service state to

$i^* = \min(i + x, \beta)$, yielding $\boldsymbol{h}_k = (0, i^*, r^*, j^*)$;

(g) No customers enter the system and no shocks arrive, yielding $\boldsymbol{h}_k = (0, i, 0, j^*)$.

Suppose at the end of τ_{k-1}, the queue length is $n_{k-1} = n \geq 0$, the arrival process is in state $j-1$, the repair process is in state $\varphi_{k-1} = r \geq \alpha$ and the service process is in state $\xi_{k-1} = i = r$. Then one of the following events can occur in τ_k:

(h) A customer enters the system with arrival rate λ_{j^*}, and $\boldsymbol{h}_k = (n + 1, r, r, 0)$;

(i) A completion of repair occurs with the repair rate δ_r, bringing the service state back to state 1, and yielding $\boldsymbol{h}_k = (n, 1, 0, j^*)$;

(j) No customers enter the system and no completion of repair occurs, yielding $\boldsymbol{h}_k = (n, r, r, j^*)$.

Figures 3.4.1 to 3.4.10 illustrate the occurrence of events (a)–(j) described above. In the figures,

1) the number inside the rectangle denotes the queue length at the end of indicated small time interval.
2) the number inside the ellipse denotes the state of the service process at the end of indicated small time interval.
3) the number inside the triangle denotes the state of the repair process at the end of indicated small time interval.
4) the number inside the circle denotes the state of the arrival process at the end of indicated small time interval.
5) the symbol 'x' indicates that a customer enters the system at the indicated time.
6) the symbol '↓' indicates that a customer leaves the system at the indicated time.
7) the symbol '⇓' indicates that a repair is completed at the indicated time.
8) the symbol '↑' indicates that a shock deteriorates the system at the indicated time.

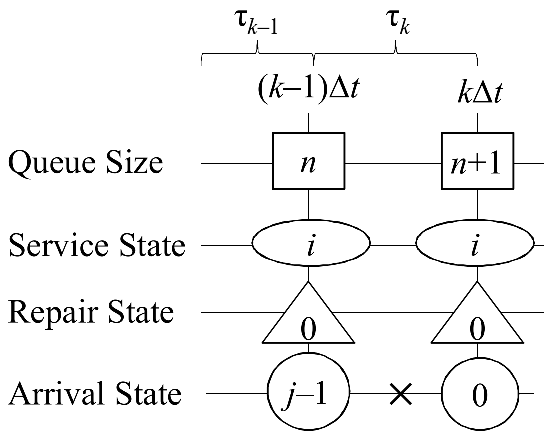

Figure 3.4.1 Transitions of queue length and states when Event (a) occurs.

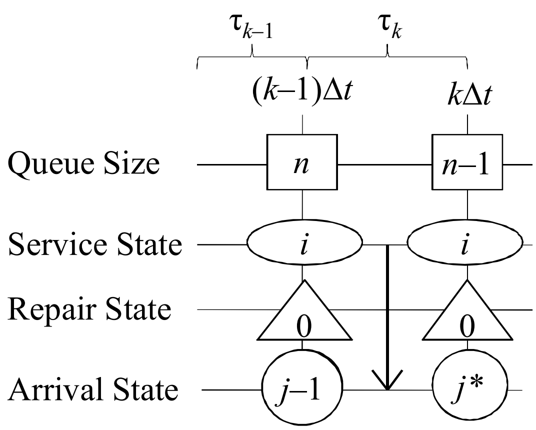

Figure 3.4.2 Transitions of queue length and states when Event (b) occurs.

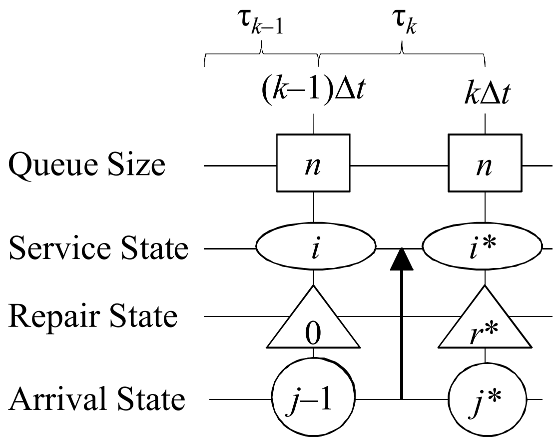

Figure 3.4.3 Transitions of queue length and states when Event (c) occurs.

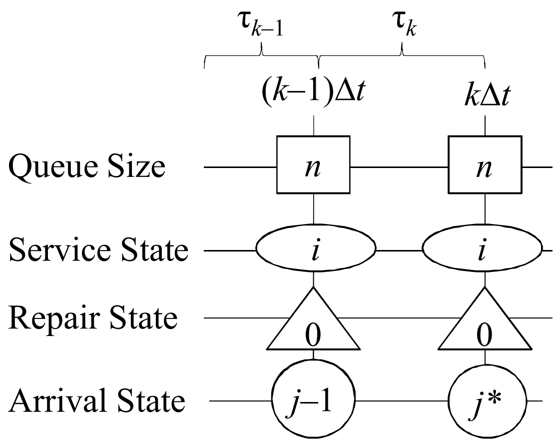

Figure 3.4.4 Transitions of queue length and states when Event (d) occurs.

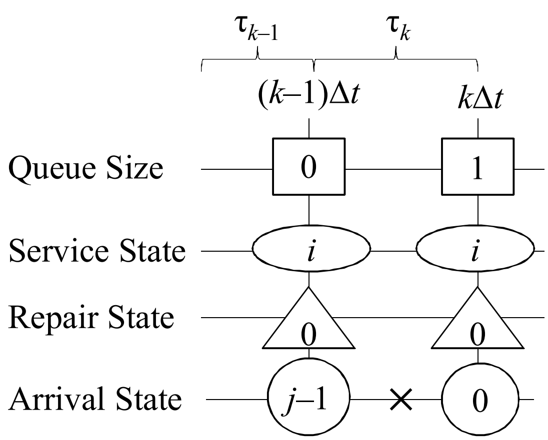

Figure 3.4.5 Transitions of queue length and states when Event (e) occurs.

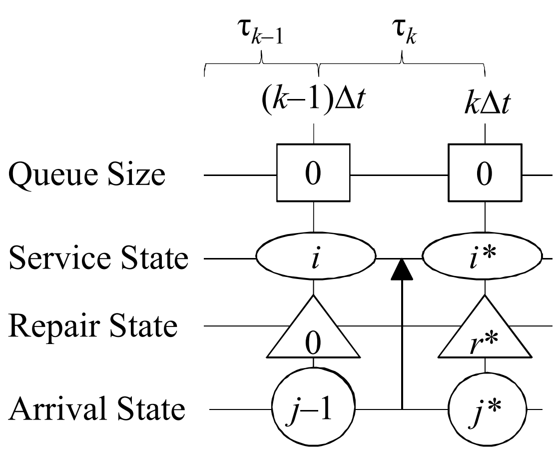

Figure 3.4.6 Transitions of queue length and states when Event (f) occurs.

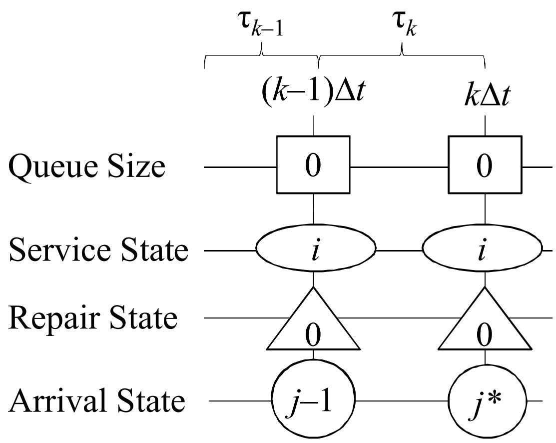

Figure 3.4.7 Transitions of queue length and states when Event (g) occurs.

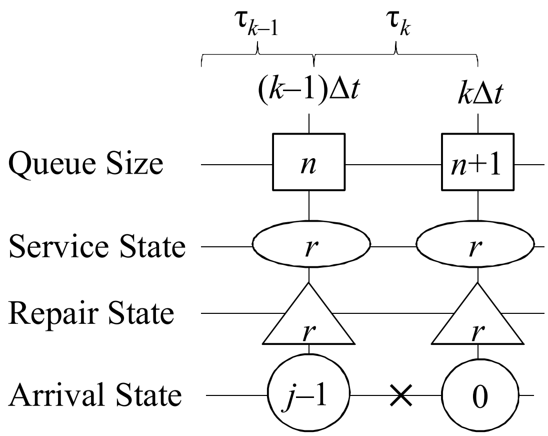

Figure 3.4.8 Transitions of queue length and states when Event (h) occurs.

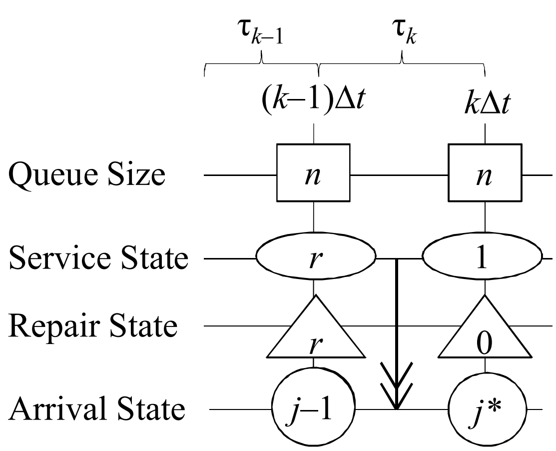

Figure 3.4.9 Transitions of queue length and states when Event (i) occurs.

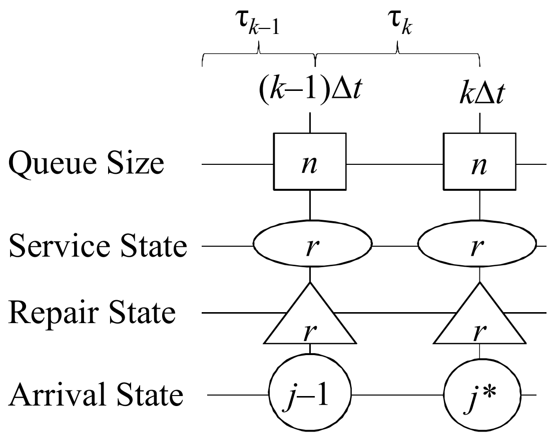

Figure 3.4.10 Transitions of queue length and states when Event (j) occurs.

By setting $n_{k-1} = 1$, $\xi_{k-1} = 1$, $\varphi_{k-1} = 0$ and $\psi_{k-1} = 0$ and letting Event (b) occur in τ_k, we get

$$P_{0101}^{(k)} \cong P_{1100}^{(k-1)}(1-\gamma\Delta t)(\mu_1\Delta t)\,. \tag{3.4.1}$$

When $k \to \infty$, we get from (3.4.1),

$$P_{0101} \cong P_{1100}(1-\gamma\Delta t)(\mu_1\Delta t)\,. \tag{3.4.2}$$

In general, for a given value of $\boldsymbol{h}_k$, we can likewise find the combinations of $\boldsymbol{h}_{k-1}$ and the event in τ_k which lead to $\boldsymbol{h}_k$, and obtain an equation similar to (3.4.2). The following equations can thus be obtained.

$$P_{0i01} \cong P_{1i00}(1-\gamma\Delta t)(\mu_i\Delta t) \qquad \text{for } 1 \le i < \alpha, \tag{3.4.3}$$

$$P_{0102} \cong P_{0101}(1-\gamma\Delta t)(1-\lambda_2\Delta t) + P_{1101}(1-\gamma\Delta t)(\mu_1\Delta t)\,, \tag{3.4.4}$$

$$\begin{aligned} P_{010j} \cong\ & P_{010(j-1)}(1-\gamma\Delta t)(1-\lambda_j\Delta t) + \sum_{m=\alpha}^{\beta} P_{0mm(j-1)}(\delta_m\Delta t) \\ & + P_{110(j-1)}(1-\gamma\Delta t)(\mu_1\Delta t) \end{aligned} \qquad \text{for } 3 \le j < J, \tag{3.4.5}$$

$$P_{010J} \cong P_{010(J-1)}(1-\gamma\Delta t)(1-\lambda_J\Delta t) + P_{010J}(1-\gamma\Delta t)(1-\lambda_J\Delta t)$$
$$+ \sum_{m=\alpha}^{\beta} P_{0mm(J-1)}(\delta_m\Delta t) + \sum_{m=\alpha}^{\beta} P_{0mmJ}(\delta_m\Delta t) + P_{110(J-1)}(1-\gamma\Delta t)(\mu_1\Delta t), \quad (3.4.6)$$
$$+ P_{110J}(1-\gamma\Delta t)(\mu_1\Delta t)$$

$$P_{0i0j} \cong \sum_{m=1}^{i-1} P_{0m0(j-1)}(\gamma\Delta t)(g_{i-m}) + P_{0i0(j-1)}(1-\gamma\Delta t)(1-\lambda_j\Delta t) + P_{1i0(j-1)}(1-\gamma\Delta t)(\mu_i\Delta t) \quad \text{for } 2 \le i < \alpha,\ 2 \le j < J, \quad (3.4.7)$$

$$P_{0i0J} \cong \sum_{m=1}^{i-1} P_{0m0(J-1)}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{0m0J}(\gamma\Delta t)(g_{i-m})$$
$$+ P_{0i0(J-1)}(1-\gamma\Delta t)(1-\lambda_J\Delta t) + P_{0i0J}(1-\gamma\Delta t)(1-\lambda_J\Delta t) \quad \text{for } 2 \le i < \alpha, \quad (3.4.8)$$
$$+ P_{1i0(J-1)}(1-\gamma\Delta t)(\mu_i\Delta t) + P_{1i0J}(1-\gamma\Delta t)(\mu_i\Delta t)$$

$$P_{0rr2} \cong \sum_{m=1}^{\alpha-1} P_{0m01}(\gamma\Delta t)(g_{r-m}) \quad \text{for } \alpha \le r < \beta, \quad (3.4.9)$$

$$P_{0\beta\beta 2} \cong \sum_{m=1}^{\alpha-1} P_{0m01}(\gamma\Delta t)\left(1 - \sum_{u=1}^{\beta-m-1} g_u\right). \quad (3.4.10)$$

For $n \ge 0$ and $\alpha \le r \le \beta$,

$$P_{nrrj} \cong \sum_{m=1}^{\alpha-1} P_{nm0(j-1)}(\gamma\Delta t)(g_{r-m}) + P_{nrr(j-1)}(1-\delta_r\Delta t)(1-\lambda_j\Delta t) \quad \text{for } \alpha \le r < \beta,\ 3 \le j < J \text{ if } n = 0,\ 1 \le j < J \text{ if } n \ge 1, \quad (3.4.11)$$

$$P_{nrrJ} \cong \sum_{m=1}^{\alpha-1} P_{nm0(J-1)}(\gamma\Delta t)(g_{r-m}) + \sum_{m=1}^{\alpha-1} P_{nm0J}(\gamma\Delta t)(g_{r-m})$$
$$+ P_{nrr(J-1)}(1-\delta_r\Delta t)(1-\lambda_J\Delta t) + P_{nrrJ}(1-\delta_r\Delta t)(1-\lambda_J\Delta t) \quad \text{for } \alpha \le r < \beta, \quad (3.4.12)$$

$$P_{n\beta\beta j} \cong \sum_{m=1}^{\alpha-1} P_{nm0(j-1)}(\gamma\Delta t)\left(1 - \sum_{u=1}^{\beta-m-1} g_u\right) + P_{n\beta\beta(j-1)}(1-\delta_\beta\Delta t)(1-\lambda_j\Delta t) \quad \text{for } 3 \le j < J \text{ if } n = 0,\ 1 \le j < J \text{ if } n \ge 1, \quad (3.4.13)$$

$$P_{n\beta\beta J} \cong \sum_{m=1}^{\alpha-1} P_{nm0(J-1)}(\gamma\Delta t)\left(1 - \sum_{u=1}^{\beta-m-1} g_u\right) + \sum_{m=1}^{\alpha-1} P_{nm0J}(\gamma\Delta t)\left(1 - \sum_{u=1}^{\beta-m-1} g_u\right).$$
$$+ P_{n\beta\beta(J-1)}(1-\delta_\beta\Delta t)(1-\lambda_J\Delta t) + P_{n\beta\beta J}(1-\delta_\beta\Delta t)(1-\lambda_J\Delta t) \quad (3.4.14)$$

When $n = 1$,

$$P_{1i00} \cong \sum_{m=1}^{J-1} P_{0i0m}(1-\gamma\Delta t)(\lambda_{m+1}\Delta t) + P_{0i0J}(1-\gamma\Delta t)(\lambda_J\Delta t) \qquad \text{for } 1 \le i < \alpha, \quad (3.4.15)$$

$$P_{1rr0} \cong \sum_{m=2}^{J-1} P_{0rrm}(1-\delta_r\Delta t)(\lambda_{m+1}\Delta t) + P_{0rrJ}(1-\delta_r\Delta t)(\lambda_J\Delta t) \qquad \text{for } \alpha \le r \le \beta. \quad (3.4.16)$$

When $n \ge 1$ and $1 \le i < \alpha$,

$$\begin{aligned} P_{n10j} &\cong P_{n10(j-1)}(1-\gamma\Delta t)(1-\mu_1\Delta t)(1-\lambda_j\Delta t) + \sum_{m=\alpha}^{\beta} P_{nmm(j-1)}(\delta_m\Delta t) \\ &\quad + P_{(n+1)10(j-1)}(1-\gamma\Delta t)(\mu_1\Delta t) \end{aligned} \qquad \text{for } 1 \le j < J, \quad (3.4.17)$$

$$\begin{aligned} P_{n10J} &\cong P_{n10(J-1)}(1-\gamma\Delta t)(1-\mu_1\Delta t)(1-\lambda_J\Delta t) + P_{n10J}(1-\gamma\Delta t)(1-\mu_1\Delta t)(1-\lambda_J\Delta t) \\ &\quad + \sum_{m=\alpha}^{\beta} P_{nmm(J-1)}(\delta_m\Delta t) + \sum_{m=\alpha}^{\beta} P_{nmmJ}(\delta_m\Delta t) \\ &\quad + P_{(n+1)10(J-1)}(1-\gamma\Delta t)(\mu_1\Delta t) + P_{(n+1)10J}(1-\gamma\Delta t)(\mu_1\Delta t) \end{aligned} \quad , \qquad (3.4.18)$$

$$\begin{aligned} P_{ni0j} &\cong \sum_{m=1}^{i-1} P_{nm0(j-1)}(\gamma\Delta t)(g_{i-m}) \\ &\quad + P_{ni0(j-1)}(1-\gamma\Delta t)(1-\mu_i\Delta t)(1-\lambda_j\Delta t) \\ &\quad + P_{(n+1)i0(j-1)}(1-\gamma\Delta t)(\mu_i\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha,\ 1 \le j < J, \quad (3.4.19)$$

$$\begin{aligned} P_{ni0J} &\cong \sum_{m=1}^{i-1} P_{nm0(J-1)}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{nm0J}(\gamma\Delta t)(g_{i-m}) \\ &\quad + P_{ni0(J-1)}(1-\gamma\Delta t)(1-\mu_i\Delta t)(1-\lambda_J\Delta t) \\ &\quad + P_{ni0J}(1-\gamma\Delta t)(1-\mu_i\Delta t)(1-\lambda_J\Delta t) \\ &\quad + P_{(n+1)i0(J-1)}(1-\gamma\Delta t)(\mu_i\Delta t) + P_{(n+1)i0J}(1-\gamma\Delta t)(\mu_i\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha. \quad (3.4.20)$$

When $n \ge 2$,

$$\begin{aligned} P_{ni00} &\cong \sum_{m=0}^{J-1} P_{(n-1)i0m}(1-\gamma\Delta t)(1-\mu_i\Delta t)(\lambda_{m+1}\Delta t) \\ &\quad + P_{(n-1)i0J}(1-\gamma\Delta t)(1-\mu_i\Delta t)(\lambda_J\Delta t) \end{aligned} \qquad \text{for } 1 \le i < \alpha, \quad (3.4.21)$$

and

$$P_{nrr0} \cong \sum_{m=0}^{J-1} P_{(n-1)rrm}(1-\delta_r\Delta t)(\lambda_{m+1}\Delta t) + P_{(n-1)rrJ}(1-\delta_r\Delta t)(\lambda_J\Delta t) \qquad \text{for } \alpha \le r \le \beta. \quad (3.4.22)$$

3.5 Stationary Queue Length Distribution

Before solving the equations in Section 3.4 to find the P_{nirj}, we may first let b_{ij}, c_{rj}, d_{ij}, e_{rj}, f_{ij}, h_{ij}, u_{rj} and v_{ij} be constants and introduce the following notations:

(a) $P_{n***} = \left\{ P_{nirj} : \left(1 \le i < \alpha, r = 0, 0 \le j \le J\right) \text{or} \left(i = r, \alpha \le r \le \beta, 0 \le j \le J\right) \right\}$;

(b) $P_{n*0*} = \{P_{ni0j}: 1 \le i < \alpha,\ 0 \le j \le J\}$;

(c) $(P_{m***}, P_{(m+1)***}, P_{(m+2)*0*})$ denotes the set of equations of the form

$$\sum_{i=1}^{\alpha-1}\sum_{j=0}^{J} b_{ij}P_{mi0j} + \sum_{r=\alpha}^{\beta}\sum_{j=0}^{J} c_{rj}P_{mrrj} + \sum_{i=1}^{\alpha-1}\sum_{j=0}^{J} d_{ij}P_{(m+1)i0j} + \sum_{r=\alpha}^{\beta}\sum_{j=0}^{J} e_{rj}P_{(m+1)rrj} + \sum_{i=1}^{\alpha-1}\sum_{j=0}^{J} f_{ij}P_{(m+2)i0j} \cong 0 \quad ;$$

(d) $(P_{mirj} \mid P_{0***}, P_{(m+1)*0*})$ denotes the equation of the form

$$P_{mirj} \cong \sum_{i=1}^{\alpha-1}\sum_{j=0}^{J} h_{ij}P_{0i0j} + \sum_{r=\alpha}^{\beta}\sum_{j=0}^{J} u_{rj}P_{0rrj} + \sum_{i=1}^{\alpha-1}\sum_{j=0}^{J} v_{ij}P_{(m+1)i0j} \,.$$

With the above notations, (3.4.11) to (3.4.20) in the case when $n = 1$ can be represented as

$$(P_{0***}, P_{1***}, P_{2*0*}), \tag{3.5.1}$$

and (3.4.17) to (3.4.22) together with (3.4.11) to (3.4.14) in the case when $n \ge 2$ can be represented as

$$(P_{(n-1)***}, P_{n***}, P_{(n+1)*0*}). \tag{3.5.2}$$

It can be shown that from the set of equations given by (3.5.1), we can get

$$(P_{1irj} \mid P_{0***}, P_{2*0*}) \qquad \text{for } 1 \le i < \alpha,\ r = 0,\ 0 \le j \le J \text{ or } i = r,\ \alpha \le r \le \beta,\ 0 \le j \le J. \tag{3.5.3}$$

By substituting the expression of P_{1irj} given by (3.5.3) into (3.5.2) when $n = 2$, and solving for P_{2irj}, we get

$$(P_{2irj} \mid P_{0***}, P_{3*0*}) \qquad \text{for } 1 \le i < \alpha,\ r = 0,\ 0 \le j \le J \text{ or } i = r,\ \alpha \le r \le \beta,\ 0 \le j \le J. \tag{3.5.4}$$

By substituting the expression of P_{2irj} given by (3.5.4) into (3.5.2) when $n = 3$ and solving for P_{3irj}, we get

$(P_{3irj} \mid P_{0***}, P_{4*0*})$ for $1 \le i < \alpha$, $r = 0$, $0 \le j \le J$ or $i = r$, $\alpha \le r \le \beta$, $0 \le j \le J$. (3.5.5)

Next for $n \ge 4$, we repeat the process of substituting the expression of $P_{(n-1)irj}$ given by

$(P_{(n-1)irj} \mid P_{0***}, P_{n*0*})$ for $1 \le i < \alpha$, $r = 0$, $0 \le j \le J$ or $i = r$, $\alpha \le r \le \beta$, $0 \le j \le J$ (3.5.6)

into (3.5.2) and solving for P_{nirj} to get

$(P_{nirj} \mid P_{0***}, P_{(n+1)*0*})$. (3.5.7)

When $n = N$ is large enough, we may set all the $P_{(n+1)*0*}$ in (3.5.7) to be zero and obtain

$(P_{Nirj} \mid P_{0***})$ for $1 \le i < \alpha$, $r = 0$, $0 \le j \le J$ or $i = r$, $\alpha \le r \le \beta$, $0 \le j \le J$. (3.5.8)

For $n = N - 1, N - 2, \ldots, 1$, we may perform the substitution of $(P_{(n+1)irj} \mid P_{0***})$ into (3.5.7) and obtain

$(P_{nirj} \mid P_{0***})$ for $1 \le i < \alpha$, $r = 0$, $0 \le j \le J$ or $i = r$, $\alpha \le r \le \beta$, $0 \le j \le J$. (3.5.9)

When $n = 1$, (3.5.9) yields $(P_{1irj} \mid P_{0***})$. By using the results given by $(P_{1irj} \mid P_{0***})$ and (3.4.3) to (3.4.14), we get the following system of $N_0 = \{(J \times \beta) - (\beta - \alpha + 1)\}$ equations:

$(P_{0irj} \mid P_{0***})$ for $1 \le i < \alpha$, $r = 0$, $0 \le j \le J$ or $i = r$, $\alpha \le r \le \beta$, $0 \le j \le J$. (3.5.10)

An inspection of (3.5.10) reveals that among the N_0 equations, only $N_0 – 1$ of them are linearly independent. Hence, we need to include another linearly independent equation so that the resulting system of N_0 equations has a unique solution. Equating the sum of the left sides of the equations given by (3.5.9) to the sum of the right sides of (3.5.9), we get an equation of the form,

$$\sum_{n=1}^{N}\sum_{i}\sum_{j} P_{nirj} = \sum_{i}\sum_{j} k_{ij} P_{0irj} \tag{3.5.11}$$

where the k_{ij} are constants, and the value of r depends on i.

As $\sum_{n=0}^{N}\sum_{i}\sum_{j} P_{nirj} \cong 1$, we get from (3.5.11) an equation involving only P_{0irj}, $1 \le i < \alpha$,

$r = 0$, $0 \leq j \leq J$ or $i = r$, $\alpha \leq r \leq \beta$, $0 \leq j \leq J$. This equation derived from (3.5.11), and $N_0 - 1$ equations chosen from (3.5.10), constitute a system of N_0 equations which can be solved to yield numerical answers for the P_{0irj}, $1 \leq i < \alpha, r = 0, 0 \leq j \leq J$ or $i = r, \alpha \leq r \leq \beta, 0 \leq j \leq J$. Then using (3.5.9), we can get numerical answers for the P_{nirj} where $n \geq 1$, $1 \leq i < \alpha$, $r = 0$, $0 \leq j \leq J$ or $i = r$, $\alpha \leq r \leq \beta$, $0 \leq j \leq J$. The stationary probability that the queue length is n is then given by the sum of the P_{nirj} over all i, r and j,

$$P_n = \sum_i \sum_j P_{nirj} \,. \tag{3.5.12}$$

In Equation (3.5.12), the sum over the value of r is not included as the value of r depends on i as summarized below:

$$r = \begin{cases} 0 & \text{for } 1 \leq i < \alpha \\ i & \text{for } \alpha \leq i \leq \beta \end{cases} .$$

3.6 Sojourn Time Distribution

Suppose the system is in the stationary state. Let $t = 0$ be a reference point in time under this condition of the system and assume that a customer arrives at $t = 0$. The sojourn time of the arriving customer is equal to the length of time between $t = 0$ and the time when the service given to the customer is completed. The sojourn time distribution will be derived in this section.

Let $P^{(k)}_{nir|n_0 i_0 r_0}$ be the probability that at the end of τ_k, the service state is i, the repair state is r and there are n customers in the queue formed by the customers who arrive before $t = 0$ and still remain in the system, given that a customer has arrived in τ_0, and at the end of τ_0, the queue length is n_0, the service state is i_0 and the repair state is r_0. When the system is in the stationary state, we note the probability of the event $E^{(0)}$ that

(a) the queue length at the beginning of τ_0 is $n_0 - 1$;

(b) the service process is in state i_0 at the beginning of τ_0;

(c) the repair process is in state r_0 at the beginning of τ_0; and

(d) a customer arrives in τ_0;

is given approximately by

$$\sum_{j=0}^{J-1} P_{(n_0-1)i_0 r_0 j}(\lambda_{j+1}\Delta t) + P_{(n_0-1)i_0 r_0 J}(\lambda_J \Delta t)\,. \tag{3.6.1}$$

When $E^{(0)}$ has occurred, the queue length, service state and repair state at the end of τ_0 will be n_0, i_0, and r_0, respectively. Thus we may denote the probability of $E^{(0)}$ by $P^{(0)}_{n_0 i_0 r_0}$. By using a method similar to that used in Section 3.4, it can be shown that

$$P^{(k)}_{0i0|n_0 i_0 r_0} \cong P^{(k-1)}_{1i0|n_0 i_0 r_0}(1-\gamma\Delta t)(\mu_i \Delta t) \qquad \text{for } 1 \le i < \alpha, \tag{3.6.2}$$

$$\begin{aligned} P^{(k)}_{n10|n_0 i_0 r_0} &\cong P^{(k-1)}_{n10|n_0 i_0 r_0}(1-\gamma\Delta t)(1-\mu_1\Delta t) + \sum_{m=\alpha}^{\beta} P^{(k-1)}_{nmm|n_0 i_0 r_0}(\delta_m \Delta t) \\ &\quad + P^{(k-1)}_{(n+1)10|n_0 i_0 r_0}(1-\gamma\Delta t)(\mu_1\Delta t) \end{aligned}, \tag{3.6.3}$$

$$\begin{aligned} P^{(k)}_{ni0|n_0 i_0 r_0} &\cong \sum_{m=1}^{i-1} P^{(k-1)}_{nm0|n_0 i_0 r_0}(\gamma\Delta t)(g_{i-m}) + P^{(k-1)}_{ni0|n_0 i_0 r_0}(1-\gamma\Delta t)(1-\mu_i\Delta t) \\ &\quad + P^{(k-1)}_{(n+1)i0|n_0 i_0 r_0}(1-\gamma\Delta t)(\mu_i\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha, \tag{3.6.4}$$

$$P^{(k)}_{nrr|n_0 i_0 r_0} \cong \sum_{m=1}^{\alpha-1} P^{(k-1)}_{nm0|n_0 i_0 r_0}(\gamma\Delta t)(g_{r-m}) + P^{(k-1)}_{nrr|n_0 i_0 r_0}(1-\delta_r\Delta t) \qquad \text{for } \alpha \le r < \beta, \tag{3.6.5}$$

and

$$P^{(k)}_{n\beta\beta|n_0 i_0 r_0} \cong \sum_{m=1}^{\alpha-1} P^{(k-1)}_{nm0|n_0 i_0 r_0}(\gamma\Delta t)\left(1 - \sum_{u=1}^{\beta-m-1} g_u\right) + P^{(k-1)}_{n\beta\beta|n_0 i_0 r_0}(1-\delta_\beta\Delta t)\,. \tag{3.6.6}$$

When $n = 0$ at the end of τ_k, the service of the customer who arrives in τ_0 will have been completed in τ_k, and the sojourn time of the customer who arrives in τ_0 is approximately given by $k\Delta t$.

For k = 1, 2, …, we can use (3.6.2) to (3.6.6) to compute $P^{(k)}_{nir|n_0 i_0 r_0}$ from the values of the $P^{(k-1)}_{n'i'r'|n_0 i_0 r_0}$ where $n' = n$, $n + 1$. When the characteristics of the system at the end of τ_0 are given by n_0, i_0, and r_0, the probability that the customer who arrives in τ_0 has a sojourn time falling approximately in τ_k is given by

$$S^{(k)}_{n_0,i_0,r_0} = \sum_{i=1}^{\alpha-1} P^{(k)}_{0i0|n_0 i_0 r_0} \,. \tag{3.6.7}$$

Thus the pdf of the sojourn time evaluated at $k\Delta t$ is given by

$$f_s(k\Delta t) \cong \left(\sum_{n_0=1}^{N} \sum_{(i_0,r_0)\in R_0} S^{(k)}_{n_0,i_0,r_0} P^{(0)}_{n_0 i_0 r_0} \right) \Bigg/ \left(\sum_{n_0=1}^{N} \sum_{(i_0,r_0)\in R_0} P^{(0)}_{n_0 i_0 r_0} \right) \tag{3.6.8}$$

where $R_0 = \{(i_0,r_0) : 1 \le i_0 < \alpha,\, r_0 = 0\} \text{or} \{(i_0,r_0) : \alpha \le i_0 \le \beta,\, r_0 = i_0\}$ and N is a large positive integer.

3.7 T-Cycle

In [35], T-cycle is defined as the duration between two successive repair completions, and the length of the duration is denoted as T. The T-cycle can be divided into two time intervals of lengths T_1 and T_2 respectively:

(a) The interval from the time immediately after a repair to the time when the system is sent for repair again; and

(b) The interval from the beginning of a repair to the completion of the repair.

The expected value of T when the system is in the stationary state is an important value in the determination of the average cost of maintaining the system. We may find the expected value of T via the expected values of T_1 and T_2:

$E[T] = E[T_1] + E[T_2]$.

To find $E[T_1]$ and $E[T_2]$, we may first use the methods in Sections 3.7.1 to 3.7.2 to find the distributions of T_1 and T_2.

3.7.1 Distribution of T_1

When the system is in the stationary state, the probability of the event $F_1^{(0)}$ that,

(a) the queue length at the beginning of τ_0 is n_0;

(b) the repair process is in state r_0 at the beginning of τ_0 where $\alpha \le r_0 \le \beta$; and

(c) a completion of repair occurs in τ_0;

is given approximately by

$$\sum_{r_0=\alpha}^{\beta}\sum_{j=0}^{J} P_{n_0 r_0 r_0 j}(\delta_{r_0}\Delta t)\,. \tag{3.7.1}$$

When $F_1^{(0)}$ has occurred, the queue length, service state and repair state at the end of τ_0 will be n_0, 1, and 0, respectively. Thus we may denote the probability of $F_1^{(0)}$ by $P_{n_0 10}^{(0)}$. By using a method similar to that used in Section 3.4, it can be shown that

$$P_{n_0 10|n_0 10}^{(k)} \cong P_{n_0 10|n_0 10}^{(k-1)}(1-\gamma\Delta t)\,, \tag{3.7.2}$$

$$P_{ni0|n_0 10}^{(k)} \cong \sum_{m=1}^{i-1} P_{nm0|n_0 10}^{(k-1)}(\gamma\Delta t)(g_{i-m}) + P_{ni0|n_0 10}^{(k-1)}(1-\gamma\Delta t) \qquad \text{for } 2 \le i < \alpha, \tag{3.7.3}$$

$$P_{n_0 rr|n_0 10}^{(k)} \cong \sum_{m=1}^{\alpha-1} P_{n_0 m0|n_0 10}^{(k-1)}(\gamma\Delta t)(g_{r-m}) \qquad \text{for } \alpha \le r < \beta, \tag{3.7.4}$$

and

$$P_{n_0 \beta\beta|n_0 10}^{(k)} \cong \sum_{m=1}^{\alpha-1} P_{n_0 m 0|n_0 10}^{(k-1)}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right). \tag{3.7.5}$$

Suppose at the end of τ_k, the service state is $i \ge \alpha$. The system will then be sent for repair, and the value of T_1 is given approximately by $k\Delta t$.

For k = 1, 2, …, we can use (3.7.2) to (3.7.5) to compute $P_{n_0 ir|n_0 10}^{(k)}$ from the values of the $P^{(k-1)}_{n_0 i' r' | n_0 10}$. When the event $F^{(0)}_1$ has occurred, the probability that the server will deteriorate to a state which needs a repair at the end of τ_k is given approximately by

$$U^{(k)}_{n_0 10} = \sum_{r=\alpha}^{\beta} P^{(k)}_{n_0 rr | n_0 10} . \tag{3.7.6}$$

Thus the pdf, evaluated at $k\Delta t$, of the time elapsed before the system is sent for repair again is given by

$$f_{T_1}(k\Delta t) \cong \left(\sum_{n_0=0}^{N} U^{(k)}_{n_0 10} P^{(0)}_{n_0 10} \right) \Big/ \left(\sum_{n_0=0}^{N} P^{(0)}_{n_0 10} \right) . \tag{3.7.7}$$

3.7.2 Distribution of T_2

When the system is in the stationary state, the probability of the event $F^{(0)}_2$ that,

(a) the queue length at the beginning of τ_0 is n_0;

(b) the service process is in state i_0 at the beginning of τ_0 where $1 \le i_0 < \alpha$; and

(c) a shock with magnitude x occurs in τ_0 and deteriorates the server to state i^* where $i^* = r_0$ and $\alpha \le r_0 \le \beta$;

is given approximately by

$$P^{(0)}_{n_0 r_0 r_0} \cong \begin{cases} \displaystyle\sum_{i_0=1}^{\alpha-1} \sum_{j=0}^{J} P_{n_0 i_0 0 j} (\gamma \Delta t)(g_{r_0 - i_0}) & \text{for } \alpha \le r_0 < \beta \\ \displaystyle\sum_{i_0=1}^{\alpha-1} \sum_{j=0}^{J} P_{n_0 i_0 0 j} (\gamma \Delta t) \left(1 - \sum_{u=1}^{r_0 - i_0 - 1} g_u \right) & \text{for } r_0 = \beta \end{cases} . \tag{3.7.8}$$

We note that n_0, r_0, r_0 appearing in the left term of (3.7.8) denote, respectively, the queue length, service state, and repair state at the end of τ_0. These characteristics at the end of τ_0 are the consequences of the occurrence of the event $F^{(0)}_2$. By using a method similar to that used in Section 3.4, it can be shown that

$$P^{(k)}_{n_0 10|n_0 r_0 r_0} \cong P^{(k-1)}_{n_0 r_0 r_0|n_0 r_0 r_0}(\delta_{r_0}\Delta t), \qquad \text{for } \alpha \le r_0 \le \beta, \qquad (3.7.9)$$

and

$$P^{(k)}_{n_0 r_0 r_0|n_0 r_0 r_0} \cong P^{(k-1)}_{n_0 r_0 r_0|n_0 r_0 r_0}(1-\delta_{r_0}\Delta t) \qquad \text{for } \alpha \le r_0 \le \beta. \qquad (3.7.10)$$

Suppose at the end of τ_k, the service state is $i = 1$. Then the repair process is completed, and the value of T_2 is given approximately by $k\Delta t$.

For $k = 1, 2, \ldots$, we can use (3.7.9) and (3.7.10) to compute $P^{(k)}_{n_0 ir|n_0 r_0 r_0}$ from the values of the $P^{(k-1)}_{n_0 i'r'|n_0 r_0 r_0}$. When the event $F^{(0)}_2$ has occurred, the probability that the repair process is completed at the end of τ_k is given approximately by

$$V^{(k)}_{n_0 r_0 r_0} = P^{(k)}_{n_0 10|n_0 r_0 r_0}. \qquad (3.7.11)$$

Thus the pdf, evaluated at $k\Delta t$, of the time elapsed before the repair is completed is given by

$$f_{T_2}(k\Delta t) \cong \left(\sum_{n_0=0}^{N}\sum_{r_0=\alpha}^{\beta} V^{(k)}_{n_0 r_0 r_0} P^{(0)}_{n_0 r_0 r_0}\right) \Big/ \left(\sum_{n_0=0}^{N}\sum_{r_0=\alpha}^{\beta} P^{(0)}_{n_0 r_0 r_0}\right). \qquad (3.7.12)$$

3.8 Numerical Examples

In this section, the deteriorating M/M/1 queue of which the repair time is exponentially distributed is first considered. Let $\beta = 10$, $\mu_i = 8 - 0.7(i - 1)$ for $1 \le i \le \beta$, $\lambda = 4$, and $\delta_r = 8 - 0.7(r - 1)$ for $\alpha \le r \le \beta$, $\gamma = 0.2$, and $g_i = (1 - p)p^i$ where $p = 0.5$. By using the proposed numerical method, the results for the stationary queue length distribution, mean queue length, mean sojourn time and expected T-cycle length are found. The results can also be computed by the matrix-geometric method (see [35]). Simulation is also carried out to verify the results obtained. Some of the results obtained are shown in Tables 3.8.1 and 3.8.2.

Table 3.8.1

Comparison of stationary queue length distribution obtained by the proposed numerical method, those computed using matrix-geometric approach, and simulation procedure

Maintenance level, $\alpha = 4$

[$\Delta t = 10^{-9}$ for queue length distribution, $\Delta t = 10^{-3}$ for mean sojourn time and expected T-cycle length, $\lambda_j = \lambda, J = 2,\ N = 500$].

Queue Length, n	P(Queue Length = n)		
	Numerical method	Matrix-geometric approach	Simulation
0	0.425728	0.425728	0.424873
1	0.232254	0.232254	0.232926
2	0.130903	0.130903	0.130475
3	0.076396	0.076396	0.075914
4	0.046195	0.046195	0.046029
5	0.028905	0.028905	0.029022
6	0.018659	0.018659	0.018647
7	0.012377	0.012377	0.012432
8	0.008397	0.008397	0.008508
9	0.005801	0.005801	0.006088
10	0.004065	0.004065	0.004193
…	…	…	…
50	1.33E-08	1.33E-08	0
Mean Queue Length	1.551949	1.551949	1.563290
Mean Sojourn Time	0.388057	0.387987	0.386802
Expected T-Cycle Length	7.667340	7.667340	7.677104

Table 3.8.2

Comparison of stationary queue length distribution obtained by the proposed numerical method, those computed using matrix-geometric approach, and simulation procedure

Maintenance level, $\alpha = 9$

[$\Delta t = 10^{-9}$ for queue length distribution, $\Delta t = 10^{-3}$ for mean sojourn time and expected T-cycle length, $\lambda_j = \lambda, J = 2,\ N = 500$].

Queue Length, n	P(Queue Length = n)		
	Numerical method	Matrix-geometric approach	Simulation
0	0.364463	0.364463	0.364076
1	0.210426	0.210426	0.210650
2	0.127473	0.127473	0.127984
3	0.081149	0.081149	0.081064
4	0.054163	0.054163	0.053839
5	0.037704	0.037704	0.037416
6	0.027188	0.027188	0.027079
7	0.020169	0.020169	0.020306
8	0.015300	0.015300	0.015604
9	0.011812	0.011812	0.011993
10	0.009247	0.009247	0.00958
…	…	…	…
50	1.28E-05	1.28E-05	5.66E-06
Mean Queue Length	2.400651	2.400652	2.386995
Mean Sojourn Time	0.600273	0.600163	0.601444
Expected T-Cycle Length	9.915090	9.915090	9.859619

In Tables 3.8.1 and 3.8.2, the values for Δt have been chosen so that the results based on the proposed numerical method are very close to those obtained using the matrix-geometric approach. When compared to the simulation results, it is noted that the numerical results based on the above two methods are quite close to the simulation results.

For the case of M/M/1 queue the numerical method is able to yield results which are comparable to the matrix-geometric approach in terms of accuracy.

However the proposed numerical method appears to be more versatile than the matrix-geometric approach as it can handle the following case in which the customer interarrival time has a gamma distribution which is a special case of the CAR distribution. Suppose the parameters of the gamma distribution are chosen to be (κ, θ) = (5/4, 2/15) and the other parameter settings are the same as those used earlier. The stationary queue length distribution, mean queue length, mean sojourn time and expected T-cycle length are found. The results obtained are shown in Tables 3.8.3, 3.8.4 and 3.8.5.

Table 3.8.3

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

Maintenance level, $\alpha = 2$

[$\Delta t = 0.002$ for queue length distribution, mean sojourn time and expected T-cycle length, $J = 400$, $N = 500$].

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation
0	0.189619	0.189588
1	0.162565	0.162583
2	0.123777	0.123812
3	0.095899	0.095885
4	0.075452	0.075454
5	0.060171	0.060137
6	0.048551	0.048569
7	0.039569	0.039519
8	0.032525	0.032512
9	0.026924	0.026929
10	0.022419	0.022432
…	…	…
50	3.52E-05	3.61E-05
Mean Queue Length	4.638622	4.643121
Mean Sojourn Time	0.780414	0.778980
Expected T-Cycle Length	5.372657	5.378967

Table 3.8.4

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

Maintenance level, $\alpha = 4$

[$\Delta t = 0.002$ for queue length distribution, mean sojourn time and expected T-cycle length, $J = 400$, $N = 500$].

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation
0	0.164642	0.164453
1	0.145296	0.145233
2	0.114241	0.114174
3	0.091437	0.091417
4	0.074311	0.074265
5	0.061173	0.061164
6	0.050890	0.050881
7	0.042699	0.042717
8	0.036071	0.036104
9	0.030636	0.030666
10	0.026131	0.026119
…	…	…
50	7.69E-05	7.18E-05
Mean Queue Length	5.502307	5.509168
Mean Sojourn Time	0.925434	0.924170
Expected T-Cycle Length	7.667340	7.674061

Table 3.8.5

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

Maintenance level, $\alpha = 9$

[$\Delta t = 0.002$ for queue length distribution, mean sojourn time and expected T-cycle length, $J = 400$, $N = 500$].

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation
0	0.090909	0.090995
1	0.082469	0.082612
2	0.067268	0.067400
3	0.056253	0.056379
4	0.048082	0.048200
5	0.041870	0.041962
6	0.037028	0.037029
7	0.033160	0.033190
8	0.030000	0.029998
9	0.027364	0.027395
10	0.025128	0.025168
…	…	…
50	0.003383	0.003381
Mean Queue Length	17.457620	17.378550
Mean Sojourn Time	2.932430	2.985675
Expected T-Cycle Length	9.915090	9.904495

Tables 3.8.3, 3.8.4 and 3.8.5 show that when $\Delta t = 0.002$, the results obtained using the proposed numerical method are close to the simulation results. The results based on the numerical method may be improved if the results are extrapolated based on a number of small values of Δt. For example, to improve the accuracy of P_n, we may fit a low degree polynomial function to a number of points (Δt, P_n) obtained by varying the values of Δt, and get an answer based on the polynomial for P_n when $\Delta t = 0$.

Next, the formula

$$C(\alpha) = C_H E[N_S] + \left(C_R / E[T]\right) \tag{3.8.1}$$

given in [35] is used to compute the average cost per unit time $C(\alpha)$ from the holding cost per customer per unit time C_H, the expected queue length $E[N_S]$, the fixed repair cost

$C_R = 12$, and the expected length $E[T]$ of the T-cycle. Figure 3.8.1 shows the average cost per unit time for the system at different values of the maintenance level α and holding cost C_H. Figure 3.8.2 compares the average costs when the mean arrival times given by Table 3.8.6 are used. The corresponding parameters of the gamma distributions are also shown in Table 3.8.6. In approximating the gamma distributions by the CAR distributions, we have made used of the values of Δt and J given in Table 3.8.6.

Table 3.8.6

Parameters of gamma distribution and the values of Δt and J used for obtaining CAR distribution.

Mean arrival time, $\kappa\theta$	1/6	1/5	1/4	1/2
Parameter vector, (κ, θ)	(1.25, 1/7.5)	(1.6, 0.125)	(1.25, 0.2)	(2.5, 0.2)
Length of interval, Δt	0.002	0.0022	0.003	0.005
J	400	500	500	500

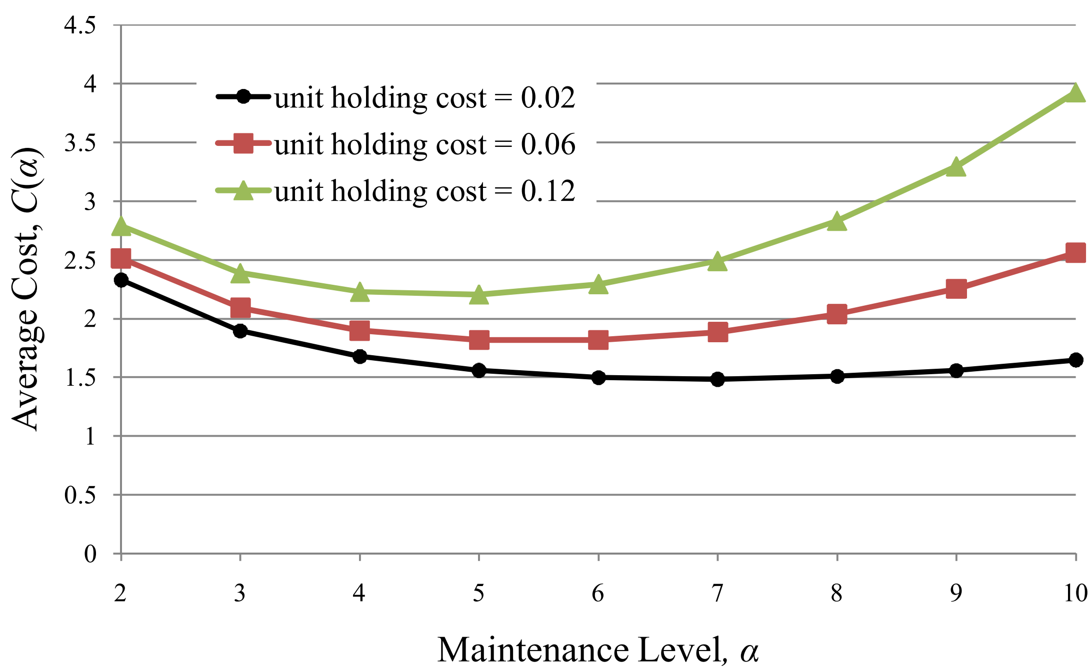

Figure 3.8.1 Average cost over maintenance level and unit holding cost $[(\kappa, \theta) = (5/4, 2/15)$, $\beta = 10$, $\mu_i = 8 - 0.7(i-1)$, $\delta_r = 8 - 0.7(r-1)$ for $\alpha \le r \le \beta$, $\gamma = 0.2$, $g_i = 0.5^{(i+1)}$ and $C_R = 12]$.

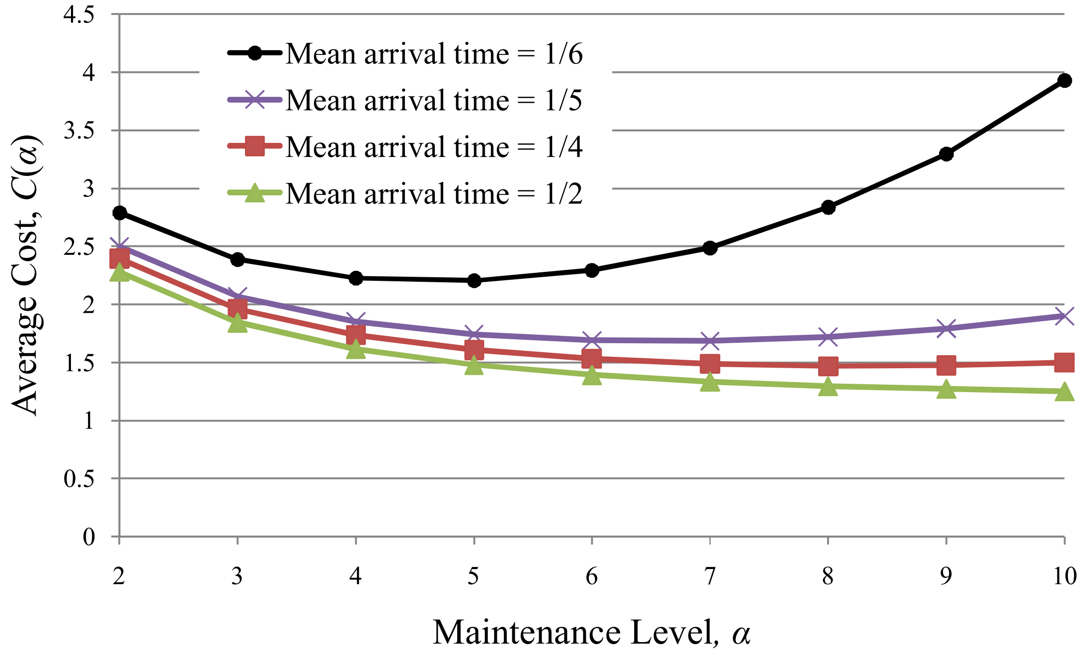

Figure 3.8.2 Average cost over maintenance level and mean of customer arrival distribution [$\beta = 10$, $\mu_i = 8 - 0.7(i-1)$, $\delta_r = 8 - 0.7(r-1)$ for $\alpha \le r \le \beta$, $\gamma = 0.2$, $g_i = 0.5^{(i+1)}$, $C_R = 12$ and $C_H = 0.12$].

Figure 3.8.1 shows that when the unit holding costs are 0.02, 0.06 and 0.12, the average cost is lowest when $\alpha = 7$, 5 and 5, respectively. Thus the optimal maintenance level tends to decrease as the unit holding cost C_H increases.

Figure 3.8.2 reveals that when the mean arrival times are 1/6, 1/5, 1/4 and 1/2, the average cost is lowest when $\alpha = 5$, 7, 8 and 10, respectively. Thus the optimal maintenance level increases as the mean of the arrival distribution increases.

Tables 3.8.7, 3.8.8 and 3.8.9 show the results obtained when the parameters of the gamma distribution are (5, 1.25), (1.675, 2) and (2.5, 1.8), respectively.

Table 3.8.7

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

Maintenance level, $\alpha = 5$

$[(\kappa, \theta) = (5, 1.25), \beta = 8, \mu_i = 1 - 0.05(i-1), \delta_r = 1 - 0.05(r-1), \gamma = 0.1, g_i = 0.5^{(i+1)}]$

[$\Delta t = 0.05$ for queue length distribution, mean sojourn time and expected T-cycle length, $J = 400$, $N = 150$].

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation
0	0.800391	0.800423
1	0.190476	0.190449
2	0.008578	0.008565
3	5.19E-04	5.25E-04
4	3.39E-05	3.48E-05
5	2.25E-06	2.49E-06
6	1.50E-07	8.10E-08
7	1.01E-08	0
8	6.74E-10	0
9	4.52E-11	0
10	3.03E-12	0
…	…	…
20	5.60E-24	0
Mean Queue Length	0.209337	0.209306
Mean Sojourn Time	1.314895	1.310193
Expected T-Cycle Length	17.275573	17.271143

Table 3.8.8

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

Maintenance level, $\alpha = 6$

[$(\kappa, \theta) = (1.675, 2)$, $\beta = 12$, $\mu_i = 1 - 0.05(i-1)$, $\delta_r = 1 - 0.05(r-1)$, $\gamma = 0.1$, $g_i = 0.5^{(i+1)}$]

[$\Delta t = 0.0375$ for queue length distribution, mean sojourn time and expected T-cycle length, $J = 400$, $N = 150$].

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation
0	0.618381	0.618486
1	0.261449	0.261437
2	0.078249	0.078256
3	0.026406	0.026352
4	0.009591	0.009579
5	0.003626	0.003616
6	0.001399	0.001389
7	5.46E-04	5.39E-04
8	2.14E-04	2.12E-04
9	8.42E-05	8.31E-05
10	3.32E-05	3.23E-05
…	…	…
20	3.07E-09	0
Mean Queue Length	0.568928	0.568496
Mean Sojourn Time	1.930162	1.916403
Expected T-Cycle Length	18.934896	18.927775

Table 3.8.9

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

Maintenance level, $\alpha = 8$

$[(\kappa, \theta) = (2.5, 1.8), \beta = 10, \mu_i = 0.7 - 0.06(i-1), \delta_r = 0.7 - 0.06(i-1), \gamma = 0.1, g_i = 0.5^{(i+1)}]$

$[\Delta t = 0.0475$ for queue length distribution, mean sojourn time and expected T-cycle length, $J = 400, N = 400]$.

Queue Length, n	P(Queue Length = n)	
	Numerical method	Simulation
0	0.436836	0.436990
1	0.269865	0.269914
2	0.120586	0.120621
3	0.067156	0.067111
4	0.040299	0.040257
5	0.024750	0.024699
6	0.015320	0.015286
7	0.009513	0.009480
8	0.005916	0.005886
9	0.003682	0.003662
10	0.002292	0.002292
…	…	…
20	2.02E-05	1.99E-05
Mean Queue Length	1.307250	1.306171
Mean Sojourn Time	5.921582	5.917252
Expected T-Cycle Length	24.314795	24.322062

3.9 Conclusion

In this chapter, the multi-state deteriorating M/M/1 queue given in [35] is studied. The customer's interarrival time distribution in the model is changed to one which has a fairly general distribution called the CAR distribution. The numerical method proposed in Chapter 2 has been successfully adapted for finding the stationary queue length distribution, stationary sojourn time distribution and expected T-cycle length. The results thus found are used to find an optimal maintenance policy such that the long run average cost is minimized. The numerical results obtained show that the optimal maintenance level increases as the unit holding cost C_H decreases or when the mean of the arrival distribution increases.

CHAPTER 4

MAINTENANCE OF A DETERIORATING QUEUE WITH NON-EXPONENTIAL SERVICE TIMES

4.1 Introduction

Consider again the multi-state M/M/1 queue studied in [35]. In Chapter 3, the distribution of the interarrival time in the above model is changed to one which has a constant asymptotic rate (CAR). In the present chapter, the distribution of the service time is instead changed to one which has a constant asymptotic rate. Then the resulting queue may be called an M/CAR/1 queue. In what follows, the deterioration of the CAR service time in the presence of a shock which occurs randomly with a random magnitude is described.

As in Chapter 2, let $g(t)$ be the probability density function (pdf) of the service time. When the service is in the initial state of "1", the rate of the service time distribution evaluated at $t = s\Delta t$ is given by

$$\mu_{1s} = \frac{g(s\Delta t)}{\int_{s\Delta t}^{\infty} g(u)du}, \qquad s \geq 1.$$

When the service time has a CAR distribution, we may assume that there is a large positive integer I such that

$$\mu_{1I} \cong \lim_{s\to\infty} \mu_{1s}.$$

This means $\mu_{1s} \cong \mu_{1I}$ for $s \geq I$.

When the server experiences a shock at time $t = s\Delta t$, the shock is assumed to have magnitude x with probability g_x. The service state will then be changed from 1 to

$i = \min(1 + x, \beta)$. When $i \geq \alpha$, the server will be sent for repair. However when $i < \alpha$, we assume that the service rate will deteriorate from μ_{1s} to $\mu_{is} = \mu_{1s} f^{(i)}$ where $f^{(1)} = 1$ and $f^{(i)} < 1$ for $2 \leq i \leq \beta$ is a constant. When the server who is in state $i < \alpha$ experiences yet another shock with magnitude x^* at time $t = s^*\Delta t$, it is assumed that the service state will change to $i^* = \min(i + x^*, \beta)$. When $i^* \geq \alpha$, the server will be sent for repair. But when $i^* < \alpha$, the service rate will deteriorate from μ_{is^*} to $\mu_{i^*s^*} = \mu_{1s^*} f^{(i^*)}$. The above description shows that the rate of the server depends on i (or i^*) which will be increased to a larger value by a shock, and also on s (or s^*) which will increase with time for s (or s^*) $< I$ and remain at I when s (or s^*) $\geq I$. We may thus use (i, s) (or (i^*, s^*)) to denote the overall state of the server.

The adopted preventive maintenance policy requires the server to be repaired when the first component i of the service state (i, s) exceeds $\alpha - 1$ where $\alpha \leq \beta$, and it is assumed that the repair time is exponentially distributed with rate δ_r where $r = i$ is defined to be the repair state. The server does not provide service to the customers during a repair.

4.2 Notations and Assumptions

The following notations are used throughout Chapter 4:

β	largest possible service state
α	maintenance level for the system, $\alpha \leq \beta$
μ_{is}	service rate in state (i, s) of the service process
δ_r	repair rate in state r of the repair process
λ	arrival rate of the arrival process
γ	shock rate

n number of customers in the system

g_x probability that the random amount of the shock is x

τ_k interval given by $((k-1)\Delta t, k\Delta t]$, $k = 0, 1, 2, \ldots$

n_k queue length of the system at the end of τ_k

(ξ_k, ω_k) state vector of the service process at the end of τ_k, $\xi_k \in \{1, 2, 3, ..., \beta\}$ and $\omega_k \in \{0, 1, 2, 3, ..., I\}$

φ_k state number of the repair process at the end of τ_k, $\varphi_k \in \{0, \alpha, \alpha+1, ..., \beta\}$

ψ_k state number of the arrival process at the end of τ_k, $\psi_k \in \{0, 1\}$

P_{nisrj} the probability that at the end of τ_k,

(a) the number of customers in the system is n (including the customer that is being served);

(b) the state vector of the service process is (i, s);

(c) the repair process is in state r; and

(d) the arrival process is in state j

Assumptions:

1. The first component i of the service state vector (i, s) is ordered. The server has the largest service rate when $i = 1$.

2. Repair is performed immediately on the system when the first component i of the service state vector (i, s) exceeds $\alpha - 1$.

3. Each successful repair brings the service state vector back to (1, 0).

4. $P_{nisrj} = \lim_{k \to \infty} P_{nisrj}^{(k)}$ exists.

4.3 Derivation of Equations for the Stationary Probabilities

Given that a customer arrives at a time in τ_0, we may define the state number ψ_k of the arrival process at the end of τ_k as

$$\psi_k = \begin{cases} 1, & \text{if } k = 0 \text{ or the next customer arrives in } \tau_k,\ k \geq 1. \\ 0, & \text{if the next customer does not arrive in } \tau_k,\ k \geq 1. \end{cases}$$

Next, define the state vector of the service process at the end of τ_k as

$(\xi_k, \omega_k) =$

- $(1, 0)$, if $k = 0$ or a repair is completed in τ_k, $k \geq 1$.
- $(i, 0)$, if
  - the server is active and having the state vector $(i, \min(k-1, I))$ at the end of τ_{k-1} for $1 \leq i < \alpha$, and the service ends in τ_k, $k \geq 1$; or
  - the server is idle and having the state vector $(i, 0)$ at the end of τ_{k-1} and the server remains idle in τ_k or a customer arrives in τ_k, $k \geq 1$.
- (i, k^*), if the server is active and having the state vector $(i, \min(k-1, I))$ at the end of τ_{k-1} for $1 \leq i < \alpha$, and no shocks or no service completions occur in τ_k, $k \geq 1$ where $k^* = \min(k, I)$.

$(i + x, k^*)$, if the server is active and having the state vector $(i, \min(k-1, I))$ at the end of τ_{k-1} for $1 \le i < \alpha$, and a shock with magnitude x occurs in τ_k and deteriorates the first component of the service state vector to $i + x < \alpha$, $k \ge 1$, and $k^* = \min(k, I)$.

$(\min((i + x), \beta), 0)$, if

- the server is active and having the state vector $(i, \min(k-1, I))$ at the end of τ_{k-1} for $1 \le i < \alpha$, and a shock with magnitude x occurs in τ_k and deteriorates the first component of the service state vector to $i + x \ge \alpha$, $k \ge 1$; or
- the server is idle and having the state vector $(i, 0)$ at the end of τ_{k-1} $1 \le i + x < \alpha$, and a shock with magnitude x occurs in τ_k and deteriorates the service state vector to $(\min((i + x), \beta), 0)$.

$(r, 0)$, if the service state vector is $(r, 0)$ at the end of τ_{k-1} for $\alpha \le r \le \beta$, and no repair completions occur in τ_k, $k \ge 1$.

The state number of the repair process at the end of τ_k is defined as

$$\varphi_k = \begin{cases} 0, & \text{if} \\ & \bullet\ k = 0;\text{ or} \\ & \bullet\ \xi_{k-1} = i < \alpha \text{ at the end of } \tau_{k-1} \text{ and, no shocks occur in } \tau_k;\text{ or} \\ & \bullet\ \xi_{k-1} = i < \alpha \text{ at the end of } \tau_{k-1} \text{ and, a shock with magnitude } x < \alpha - i \text{ occurs in } \tau_k,\ k \geq 1;\text{ or} \\ & \bullet\ \text{the repair state is } r \text{ at the end of } \tau_{k-1} \text{ for } \alpha \leq r \leq \beta \text{ and a repair completion occurs in } \tau_k,\ k \geq 2. \\ \min(i + x, \beta), & \text{if the service state vector is } (i, \min(k-1, I)) \text{ at the end of } \tau_{k-1} \text{ for } 1 \leq i < \alpha, \text{ and a shock with magnitude } x \geq \alpha - i \text{ occurs in } \tau_k,\ k \geq 1. \\ r, & \text{if the repair state is } r \text{ at the end of } \tau_{k-1} \text{ for } \alpha \leq r \leq \beta \text{ and, no repair completions occur in } \tau_k,\ k \geq 2. \end{cases}$$

Let n_k be the queue length at the end of τ_k and $\boldsymbol{h}_k = (n_k, \xi_k, \omega_k, \varphi_k, \psi_k)$. We may refer to $\boldsymbol{h}_k$ as the vector of characteristics of the queue at the end of τ_k.

Let $P_{nisrj}^{(k)}$ be the probability that at the end of τ_k, the number of customers in the system is n (including the customer that is being served), the service process is in state (i, s), the repair process is in state r and the arrival process is in state j. Assume that

$$P_{nisrj} = \lim_{k \to \infty} P_{nisrj}^{(k)}$$

exists. To find the P_{nisrj}, we first make the following observations.

Suppose at the end of τ_{k-1}, the queue length n is not empty (i.e. $n_{k-1} = n \geq 1$), the server is in state $(i, s-1)$ where $i < \alpha$, and the arrival process is in state j. In this case the server is still active and we define the repair state number to be zero. This means the vector of characteristics at the end of τ_{k-1} is given by $\boldsymbol{h}_{k-1} = (n, i, s-1, 0, j)$. With this value of $\boldsymbol{h}_{k-1}$, only one of the following events can occur in τ_k:

(a) A customer enters the system with the arrival rate λ, and at the end of τ_k, the vector of characteristics becomes $\boldsymbol{h}_k = (n+1, i, s^*, 0, 1)$;

(b) A customer leaves the system with the departure rate μ_{is^*}, and $\boldsymbol{h}_k = (n-1, i, 0, 0, 0)$;

(c) A shock with magnitude x occurs and deteriorates the service state to $i^* = \min(i+x, \beta)$, yielding $\boldsymbol{h}_k = (n, i^*, s^*, r^*, 0)$;

(d) No customers enter or leave the system, and no shocks arrive, yielding $\boldsymbol{h}_k = (n, i, s^*, 0, 0)$

where

$$s^* = \begin{cases} \min(s, I), & \text{if } 1 \le i + x < \alpha, \\ 0, & \text{if } i + x \ge \alpha, \end{cases} \qquad x \ge 0.$$

and

$$r^* = \begin{cases} 0, & \text{if } 1 < i + x < \alpha, \\ \min(i+x, \beta), & \text{if } i + x \ge \alpha, \end{cases} \qquad x \ge 1.$$

However if at the end of τ_{k-1}, the system is empty (i.e. $n_{k-1} = 0$), the state number i of the idle server is less than α and no customer arrives in τ_{k-1} with $\boldsymbol{h}_{k-1} = (0, i, 0, 0, 0)$, then one of the following events can occur in τ_k:

(e) A customer enters the system with arrival rate λ, and $\boldsymbol{h}_k = (1, i, 0, 0, 1)$;

(f) A shock with magnitude x occurs and deteriorates the service state to i^*, yielding $\boldsymbol{h}_k = (0, i^*, 0, r^*, 0)$;

(g) No customers enter the system and no shocks arrive, yielding $\boldsymbol{h}_k = (0, i, 0, 0, 0)$.

Suppose at the end of τ_{k-1}, the queue length is $n_{k-1} = n \ge 0$, the repair process is in state

$\varphi_{k-1} = r \geq \alpha$, the service state vector is $(r, 0)$, and the arrival process is in state j, yielding $\boldsymbol{h}_{k-1} = (n, r, 0, r, j)$. Then one of the following events can occur in τ_k:

(h) A customer enters the system with arrival rate λ, and $\boldsymbol{h}_k = (n + 1, r, 0, r, 1)$;

(i) A completion of repair occurs with the repair rate δ_r, bringing the service state vector back to (1, 0) and, yielding $\boldsymbol{h}_k = (n, 1, 0, 0, 0)$;

(j) No customers enter the system and no completion of repair occurs, yielding $\boldsymbol{h}_k = (n, r, 0, r, 0)$.

Figures 4.3.1 to 4.3.10 illustrate the occurrence of events (a)–(j) described above. In the figures,

1) the number inside the rectangle denotes the queue length at the end of indicated small time interval.
2) the number inside the ellipse denotes the state of the service process at the end of indicated small time interval.
3) the number inside the triangle denotes the state of the repair process at the end of indicated small time interval.
4) the number inside the circle denotes the state of the arrival process at the end of indicated small time interval.
5) the symbol ‘x’ indicates that a customer enters the system at the indicated time.
6) the symbol ‘↓’ indicates that a customer leaves the system at the indicated time.
7) the symbol ‘↡’ indicates that a repair is completed at the indicated time.
8) the symbol ‘↑’ indicates that a shock deteriorates the system at the indicated time.

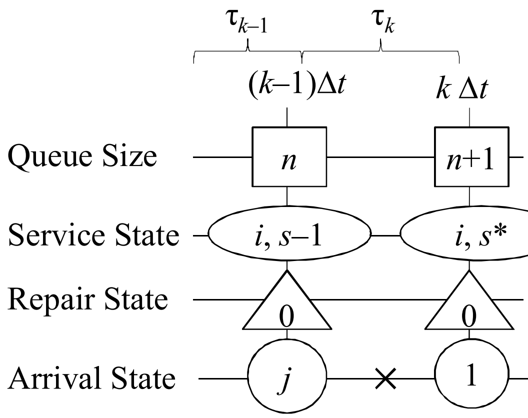

Figure 4.3.1 Transitions of queue length and states when Event (a) occurs.

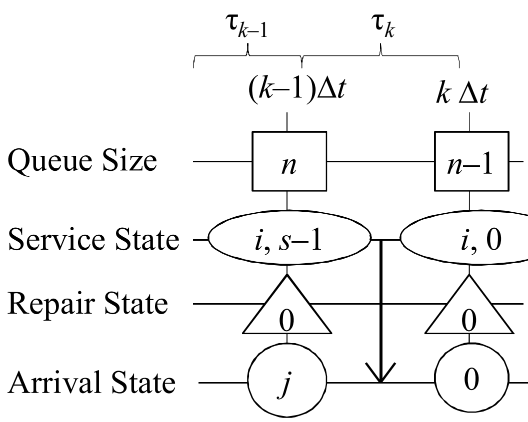

Figure 4.3.2 Transitions of queue length and states when Event (b) occurs.

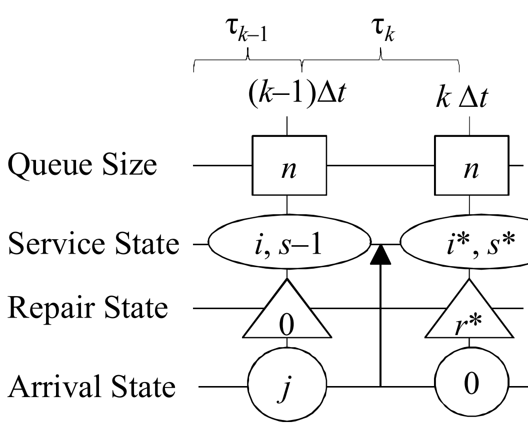

Figure 4.3.3 Transitions of queue length and states when Event (c) occurs.

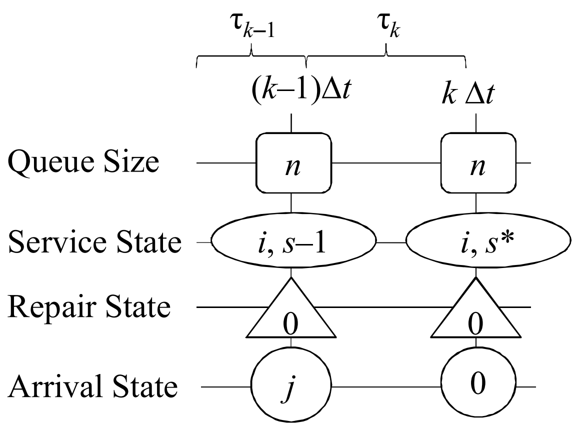

Figure 4.3.4 Transitions of queue length and states when Event (d) occurs.

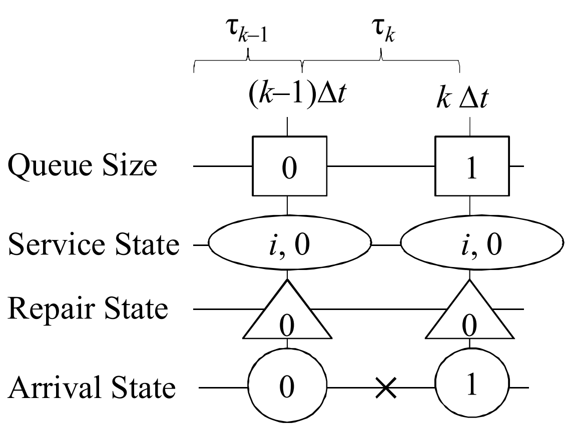

Figure 4.3.5 Transitions of queue length and states when Event (e) occurs.

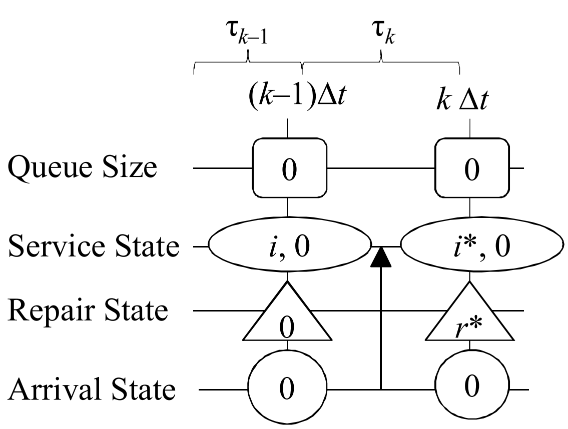

Figure 4.3.6 Transitions of queue length and states when Event (f) occurs.

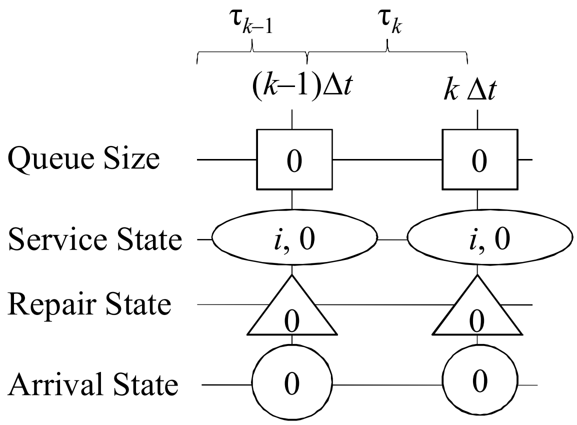

Figure 4.3.7 Transitions of queue length and states when Event (g) occurs.

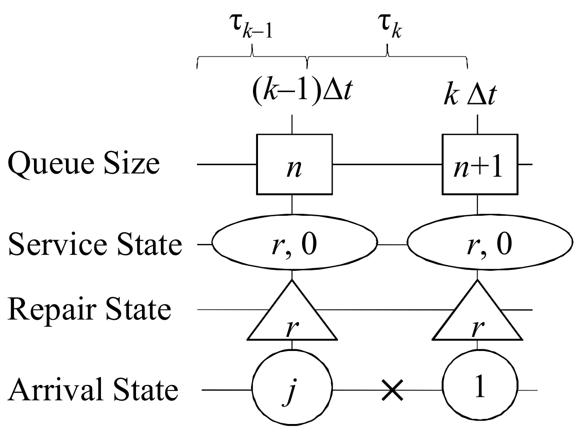

Figure 4.3.8 Transitions of queue length and states when Event (h) occurs.

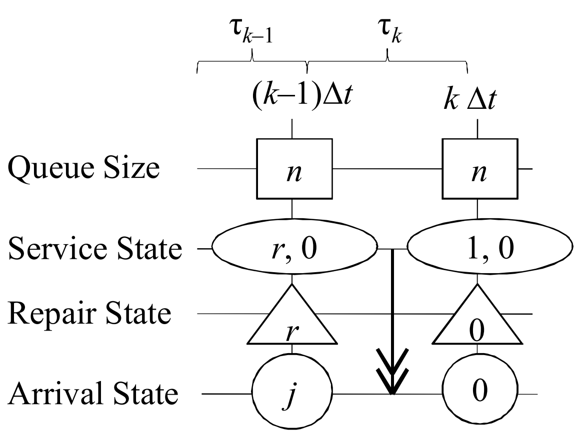

Figure 4.3.9 Transitions of queue length and states when Event (i) occurs.

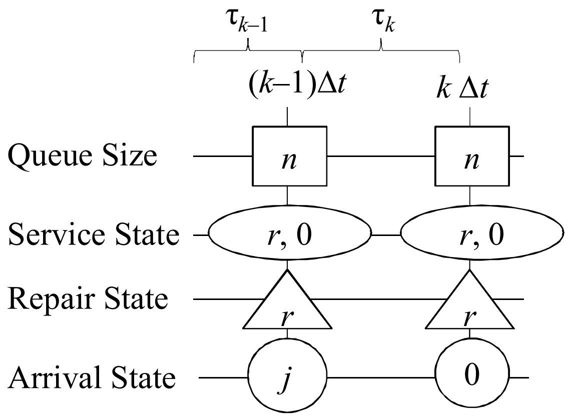

Figure 4.3.10 Transitions of queue length and states when Event (j) occurs.

By setting $n_{k-1} = 0$, $(\xi_{k-1}, \omega_{k-1}) = (1, 0)$, $\varphi_{k-1} = 0$ and $\psi_{k-1} = 0$ and letting Event (e) occur in τ_k, we get

$$P_{11001}^{(k)} \cong P_{01000}^{(k-1)}(1-\gamma\Delta t)(\lambda\Delta t)\,. \tag{4.3.1}$$

When $k\to\infty$, we get from (4.3.1),

$$P_{11001} \cong P_{01000}(1-\gamma\Delta t)(\lambda\Delta t)\,. \tag{4.3.2}$$

Similarly, with the aid of Figures 4.3.1 to 4.3.10, we can obtain the following equations.

$$\begin{aligned} P_{01000} \cong{}& P_{01000}(1-\gamma\Delta t)(1-\lambda\Delta t) + \sum_{m=\alpha}^{\beta} P_{0m0m0}(\delta_m\Delta t) + \sum_{m=0}^{I-1} P_{11m00}(1-\gamma\Delta t)(\mu_{1(m+1)}\Delta t) \\ &+ P_{11I00}(1-\gamma\Delta t)(\mu_{1I}\Delta t) + P_{11001}(1-\gamma\Delta t)(\mu_{11}\Delta t) \end{aligned}, \tag{4.3.3}$$

$$\begin{aligned} P_{0i000} \cong{}& \sum_{m=1}^{i-1} P_{0m000}(\gamma\Delta t)(g_{i-m}) + P_{0i000}(1-\gamma\Delta t)(1-\lambda\Delta t) \\ &+ \sum_{m=0}^{I-1} P_{1im00}(\mu_{i(m+1)}\Delta t)(1-\gamma\Delta t) + P_{1iI00}(\mu_{iI}\Delta t)(1-\gamma\Delta t) \\ &+ P_{1i001}(\mu_{i1}\Delta t)(1-\gamma\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha, \tag{4.3.4}$$

$$P_{0r0r0} \cong \sum_{m=1}^{\alpha-1} P_{0m000}(\gamma\Delta t)(g_{r-m}) + P_{0r0r0}(1-\delta_r\Delta t)(1-\lambda\Delta t) \qquad \text{for } \alpha \le r < \beta, \tag{4.3.5}$$

$$P_{0\beta0\beta0} \cong \sum_{m=1}^{\alpha-1} P_{0m000}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right) + P_{0\beta0\beta0}(1-\delta_\beta\Delta t)(1-\lambda\Delta t)\,. \tag{4.3.6}$$

When n = 1,

$$P_{1i001} \cong P_{0i000}(1-\gamma\Delta t)(\lambda\Delta t) \qquad \text{for } 1 \le i < \alpha, \qquad (4.3.7)$$

$$P_{11100} \cong P_{11000}(1-\gamma\Delta t)(1-\mu_{11}\Delta t)(1-\lambda\Delta t) + P_{11001}(1-\gamma\Delta t)(1-\mu_{11}\Delta t)(1-\lambda\Delta t), \qquad (4.3.8)$$

$$P_{11s00} \cong P_{11(s-1)00}(1-\gamma\Delta t)(1-\mu_{1s}\Delta t)(1-\lambda\Delta t) \qquad \text{for } 2 \le s < I, \qquad (4.3.9)$$

$$P_{11I00} \cong P_{11(I-1)00}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t)(1-\lambda\Delta t) + P_{11I00}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t)(1-\lambda\Delta t), \qquad (4.3.10)$$

$$\begin{aligned} P_{1i100} \cong & \sum_{m=1}^{i-1} P_{1m000}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{1m001}(\gamma\Delta t)(g_{i-m}) \\ & + P_{1i000}(1-\gamma\Delta t)(1-\mu_{i1}\Delta t)(1-\lambda\Delta t) \\ & + P_{1i001}(1-\gamma\Delta t)(1-\mu_{i1}\Delta t)(1-\lambda\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha, \qquad (4.3.11)$$

$$\begin{aligned} P_{1is00} \cong & \sum_{m=1}^{i-1} P_{1m(s-1)00}(\gamma\Delta t)(g_{i-m}) \\ & + P_{1i(s-1)00}(1-\gamma\Delta t)(1-\mu_{is}\Delta t)(1-\lambda\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha,\ 2 \le s < I, \qquad (4.3.12)$$

$$\begin{aligned} P_{1iI00} \cong & \sum_{m=1}^{i-1} P_{1m(I-1)00}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{1mI00}(\gamma\Delta t)(g_{i-m}) \\ & + P_{1i(I-1)00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(1-\lambda\Delta t) \\ & + P_{1iI00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(1-\lambda\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha, \qquad (4.3.13)$$

$$\begin{aligned} P_{1r0r0} \cong & \sum_{m=1}^{\alpha-1}\sum_{q=0}^{I} P_{1mq00}(\gamma\Delta t)(g_{r-m}) + \sum_{m=1}^{\alpha-1} P_{1m001}(\gamma\Delta t)(g_{r-m}) \\ & + P_{1r0r0}(1-\delta_r\Delta t)(1-\lambda\Delta t) + P_{1r0r1}(1-\delta_r\Delta t)(1-\lambda\Delta t) \end{aligned} \qquad \text{for } \alpha \le r < \beta, \qquad (4.3.14)$$

$$\begin{aligned} P_{1\beta0\beta0} \cong & \sum_{m=1}^{\alpha-1}\sum_{q=0}^{I} P_{1mq00}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right) + \sum_{m=1}^{\alpha-1} P_{1m001}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right), \\ & + P_{1\beta0\beta0}(1-\delta_\beta\Delta t)(1-\lambda\Delta t) + P_{1\beta0\beta1}(1-\delta_\beta\Delta t)(1-\lambda\Delta t) \end{aligned} \qquad (4.3.15)$$

$$P_{1r0r1} \cong P_{0r0r0}(1-\delta_r\Delta t)(\lambda\Delta t) \qquad \text{for } \alpha \le r \le \beta. \qquad (4.3.16)$$

For $n \geq 1$,

$$P_{n1000} \cong \sum_{m=r}^{\beta} P_{nm0m0}(\delta_m \Delta t) + \sum_{m=r}^{\beta} P_{nm0m1}(\delta_m \Delta t)$$
$$+ \sum_{m=0}^{I-1} P_{(n+1)1m00}(1-\gamma\Delta t)(\mu_{1(m+1)}\Delta t) + P_{(n+1)1I00}(1-\gamma\Delta t)(\mu_{1I}\Delta t), \quad (4.3.17)$$
$$+ \sum_{m=1}^{I-1} P_{(n+1)1m01}(1-\gamma\Delta t)(\mu_{1(m+1)}\Delta t) + P_{(n+1)1I01}(1-\gamma\Delta t)(\mu_{1I}\Delta t)$$

$$P_{ni000} \cong \sum_{m=0}^{I-1} P_{(n+1)im00}(1-\gamma\Delta t)(\mu_{i(m+1)}\Delta t) + P_{(n+1)iI00}(1-\gamma\Delta t)(\mu_{iI}\Delta t)$$
$$+ \sum_{m=1}^{I-1} P_{(n+1)im01}(1-\gamma\Delta t)(\mu_{i(m+1)}\Delta t) + P_{(n+1)iI01}(1-\gamma\Delta t)(\mu_{iI}\Delta t) \quad \text{for } 2 \leq i < \alpha. \quad (4.3.18)$$

When $n = 2$,

$$P_{2i101} \cong P_{1i000}(1-\gamma\Delta t)(1-\mu_{i1}\Delta t)(\lambda\Delta t)$$
$$+ P_{1i001}(1-\gamma\Delta t)(1-\mu_{i1}\Delta t)(\lambda\Delta t) \quad \text{for } 1 \leq i < \alpha, \quad (4.3.19)$$

$$P_{2is01} \cong P_{1i(s-1)00}(1-\gamma\Delta t)(1-\mu_{is}\Delta t)(\lambda\Delta t) \quad \text{for } 1 \leq i < \alpha,\ 2 \leq s < I, \quad (4.3.20)$$

$$P_{2iI01} \cong P_{1i(I-1)00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(\lambda\Delta t)$$
$$+ P_{1iI00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(\lambda\Delta t) \quad \text{for } 1 \leq i < \alpha. \quad (4.3.21)$$

For $n \geq 2$,

$$P_{n1100} \cong P_{n1000}(1-\gamma\Delta t)(1-\mu_{11}\Delta t)(1-\lambda\Delta t), \quad (4.3.22)$$

$$P_{n1s00} \cong P_{n1(s-1)00}(1-\gamma\Delta t)(1-\mu_{1s}\Delta t)(1-\lambda\Delta t)$$
$$+ P_{n1(s-1)01}(1-\gamma\Delta t)(1-\mu_{1s}\Delta t)(1-\lambda\Delta t) \quad \text{for } 1 \leq s < I, \quad (4.3.23)$$

$$P_{n1I00} \cong P_{n1(I-1)00}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t)(1-\lambda\Delta t) + P_{n1(I-1)01}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t)(1-\lambda\Delta t)$$
$$P_{n1I00}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t)(1-\lambda\Delta t) + P_{n1I01}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t)(1-\lambda\Delta t), \quad (4.3.24)$$

$$P_{ni100} \cong \sum_{m=1}^{i-1} P_{nm000}(\gamma\Delta t)(g_{i-m}) + P_{ni000}(1-\gamma\Delta t)(1-\mu_{i1}\Delta t)(1-\lambda\Delta t) \quad \text{for } 2 \leq i < \alpha, \quad (4.3.25)$$

$$P_{nis00} \cong \sum_{m=1}^{i-1} P_{nm(s-1)00}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{nm(s-1)01}(\gamma\Delta t)(g_{i-m}) + P_{ni(s-1)00}(1-\gamma\Delta t)(1-\mu_{is}\Delta t)(1-\lambda\Delta t) + P_{ni(s-1)01}(1-\gamma\Delta t)(1-\mu_{is}\Delta t)(1-\lambda\Delta t) \quad \text{for } 2 \le i < \alpha,\ 2 \le s < I, \tag{4.3.26}$$

$$P_{niI00} \cong \sum_{m=1}^{i-1} P_{nm(I-1)00}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{nm(I-1)01}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{nmI00}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P_{nmI01}(\gamma\Delta t)(g_{i-m}) + P_{ni(I-1)00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(1-\lambda\Delta t) + P_{ni(I-1)01}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(1-\lambda\Delta t) + P_{niI00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(1-\lambda\Delta t) + P_{niI01}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(1-\lambda\Delta t) \quad \text{for } 2 \le i < \alpha, \tag{4.3.27}$$

$$P_{nr0r0} \cong \sum_{m=1}^{\alpha-1}\sum_{q=0}^{I} P_{nmq00}(\gamma\Delta t)(g_{r-m}) + \sum_{m=1}^{\alpha-1}\sum_{q=1}^{I} P_{nmq01}(\gamma\Delta t)(g_{r-m}) + P_{nr0r0}(1-\delta_r\Delta t)(1-\lambda\Delta t) + P_{nr0r1}(1-\delta_r\Delta t)(1-\lambda\Delta t) \quad \text{for } \alpha \le r < \beta, \tag{4.3.28}$$

$$P_{n\beta0\beta0} \cong \sum_{m=1}^{\alpha-1}\sum_{q=0}^{I} P_{nmq00}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right) + \sum_{m=1}^{\alpha-1}\sum_{q=1}^{I} P_{nmq01}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right), + P_{n\beta0\beta0}(1-\delta_\beta\Delta t)(1-\lambda\Delta t) + P_{n\beta0\beta1}(1-\delta_\beta\Delta t)(1-\lambda\Delta t) \tag{4.3.29}$$

$$P_{nr0r1} \cong P_{(n-1)r0r0}(1-\delta_r\Delta t)(\lambda\Delta t) + P_{(n-1)r0r1}(1-\delta_r\Delta t)(\lambda\Delta t) \quad \text{for } \alpha \le r \le \beta. \tag{4.3.30}$$

For $n \ge 3$,

$$P_{ni101} \cong P_{(n-1)i000}(1-\gamma\Delta t)(1-\mu_{i1}\Delta t)(\lambda\Delta t) \quad \text{for } 1 \le i < \alpha, \tag{4.3.31}$$

$$P_{nis01} \cong P_{(n-1)i(s-1)00}(1-\gamma\Delta t)(1-\mu_{is}\Delta t)(\lambda\Delta t) + P_{(n-1)i(s-1)01}(1-\gamma\Delta t)(1-\mu_{is}\Delta t)(\lambda\Delta t) \quad \text{for } 1 \le i < \alpha,\ 2 \le s < I, \tag{4.3.32}$$

and

$$P_{niI01} \cong P_{(n-1)i(I-1)00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(\lambda\Delta t) + P_{(n-1)i(I-1)01}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(\lambda\Delta t) + P_{(n-1)iI00}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(\lambda\Delta t) + P_{(n-1)iI01}(1-\gamma\Delta t)(1-\mu_{iI}\Delta t)(\lambda\Delta t) \quad \text{for } 1 \le i < \alpha. \tag{4.3.33}$$

4.4 Stationary Queue Length Distribution

Before solving (4.3.3) to (4.3.33) in Section 4.3 to obtain the stationary queue length distribution, we first let b_{isj}, c_{rj}, d_{isj}, e_{rj}, f_{isj}, h_i, u_r and v_{isj} be constants and introduce the following notations:

1) $P_{n****} = \{P_{nisrj} : (1 \le i < \alpha, 0 \le s \le I, r = 0, j = 0,1) \text{ or } (i = r, s = 0, \alpha \le r \le \beta, j = 0,1)\};$

2) $P_{n**0*} = \{ P_{nis0j}: 1 \le i < \alpha, 0 \le s \le I, j = 0,1\};$

3) $(P_{m****}, P_{(m+1)****}, P_{(m+2)**0*})$ denotes the set of equations of the form

$$\sum_{i=1}^{\alpha-1}\sum_{s=0}^{I}\sum_{j=0}^{1} b_{isj}P_{mis0j} + \sum_{r=\alpha}^{\beta}\sum_{j=0}^{1} c_{rj}P_{mr0rj} + \sum_{i=1}^{\alpha-1}\sum_{s=0}^{I}\sum_{j=0}^{1} d_{isj}P_{(m+1)is0j} + \sum_{r=\alpha}^{\beta}\sum_{j=0}^{1} e_{rj}P_{(m+1)r0rj} + \sum_{i=0}^{\alpha-1}\sum_{s=0}^{I}\sum_{j=0}^{1} f_{isj}P_{(m+2)is0j} \cong 0 \quad ;$$

4) $(P_{misrj} \mid P_{0i0*0}, P_{(m+1)**0*})$ denotes the equation of the form

$$P_{misrj} \cong \sum_{i=1}^{\alpha-1} h_i P_{0i000} + \sum_{r=\alpha}^{\beta} u_r P_{0r0r0} + \sum_{i=1}^{\alpha-1}\sum_{s=0}^{I}\sum_{j=0}^{1} v_{isj}P_{(m+1)is0j} \quad .$$

With the above notations, (4.3.7) to (4.3.18) in the case when $n = 1$ can be represented as

$(P_{0****}, P_{1****}, P_{2**0*})$, (4.4.1)

and (4.3.17) to (4.3.30) in the case when $n = 2$ may be represented as

$(P_{1****}, P_{2****}, P_{3**0*})$. (4.4.2)

When $n \ge 3$, (4.3.17) to (4.3.18) together with (4.3.22) to (4.3.33) can be represented as

$(P_{(n-1)****}, P_{n****}, P_{(n+1)**0*})$. (4.4.3)

It can be shown that from the set of equations given by (4.4.1), we can get

$(P_{1isrj} \mid P_{0i0*0}, P_{2**0*})$ for $(i,s,r,j) \in R_0$ (4.4.4)

where

$$R_0 = \{(i,s,r,j): 1 \le i < \alpha, 1 \le s \le I, r = 0, j = 0,1\} \cup \{(i,s,r,j): \alpha \le i \le \beta, s = 0, r = i, j = 0,1\}.$$

By substituting the expression of the P_{1isrj} given by (4.4.4) into (4.4.2), and solving for the P_{2isrj}, we get

$(P_{2isrj} \mid P_{0i0*0}, P_{3**0*})$ for $(i,s,r,j) \in R_0$. (4.4.5)

By substituting the expression of the P_{2isrj} given by (4.4.5) into (4.4.3) when n = 3 and solving for the P_{3isrj}, we get

$(P_{3isrj} \mid P_{0i0*0}, P_{4**0*})$ for $(i,s,r,j) \in R_0$. (4.4.6)

Next for $n \ge 4$, repeat the process of substituting the expression of the $P_{(n-1)isrj}$ given by

$(P_{(n-1)isrj} \mid P_{0i0*0}, P_{n**0*})$ for $(i,s,r,j) \in R_0$. (4.4.7)

into (4.4.3) and solving for the P_{nisrj} to get

$(P_{nisrj} \mid P_{0i0*0}, P_{(n+1)**0*})$ for $(i,s,r,j) \in R_0$. (4.4.8)

When $n = N$ is large enough, we may set all the $P_{(n+1)**0*}$ in (4.4.8) to be zero and obtain

$(P_{nisrj} \mid P_{0i0*0})$ for $(i,s,r,j) \in R_0$. (4.4.9)

For $n = N - 1, N - 2,\ldots,1$, we may perform the substitution of $(P_{(n+1)isrj} \mid P_{0i0*0})$ into (4.4.8) and obtain

$(P_{nisrj} \mid P_{0i0*0})$ for $(i,s,r,j) \in R_0$. (4.4.10)

When n = 1, (4.4.10) yields $(P_{1isrj} \mid P_{0i0*0})$. By using the results given by $(P_{1isrj} \mid P_{0i0*0})$ and the equations (4.3.3) to (4.3.6), we get the following system of $N_0 = \beta$ equations:

$(P_{0i0*0} \mid P_{0i0*0})$ for $1 \le i \le \beta$. (4.4.11)

An inspection of (4.4.11) reveals that among the N_0 equations, only $N_0 - 1$ of them are linearly independent. Hence, we need to include another linearly independent equation so that the resulting system of N_0 equations has a unique solution. Equating the sum of the left sides of the equations given by (4.4.10) to the sum of the right sides of (4.4.10), we get an equation of the form,

$$\sum_{n=1}^{N}\sum_{i}\sum_{s}\sum_{j}P_{nisrj}=\sum_{i}k_{i}P_{0i0r0} \quad (4.4.12)$$

where $1 \le i < \alpha$, $r = 0$ or $i = r$, $\alpha \le r \le \beta$, and the k_i are constants.

As $\sum_{n=0}^{N}\sum_{i}\sum_{s}\sum_{j}P_{nisrj}\cong 1$, we get from (4.4.12) an equation involving only P_{0i0r0}, $1 \le i < \alpha$, $r = 0$ or $i = r$, $\alpha \le r \le \beta$. This equation derived from (4.4.12), and $N_0 - 1$ equations chosen from (4.4.11), constitute a system of N_0 equations which can be solved to yield numerical answers for P_{0i0r0}, $1 \le i < \alpha$, $r = 0$ or $i = r$, $\alpha \le r \le \beta$. Then using (4.4.10), we can get numerical answers for P_{nisrj} where $n \ge 1$, $1 \le i < \alpha, 0 \le s \le I, r = 0, j = 0, 1$ or $i = r, s = 0, \alpha \le r \le \beta, j = 0, 1$. The stationary probability that the queue length is n is then given by the sum of the P_{nisrj} over all i, s, r and j,

$$P_n=\sum_{i}\sum_{s}\sum_{j}P_{nisrj}\,. \quad (4.4.13)$$

In Equation (4.4.13), the sum over the value of r is not included as the value of r depends on i as summarized below:

$$r=\begin{cases} 0 & \text{for } 1 \le i < \alpha \\ i & \text{for } \alpha \le i \le \beta \end{cases}\,.$$

4.5 Sojourn Time Distribution

Suppose the system is in the stationary state. Let $t = 0$ be a reference point in time under this condition of the system and assume that a customer arrives at $t = 0$. The distribution of the sojourn time of the arriving customer will be derived in this section.

Let $P^{(k)}_{nisr|n_0 i_0 s_0 r_0}$ be the probability that at the end of τ_k, the service state vector is (i, s), the repair state is r and there are n customers in the queue formed by the customers who arrive before t = 0 and still remain in the system, given that at the end of τ_0, the queue length is n_0, the service state vector is (i_0, s_0) and the repair state is r_0. When the system is in the stationary state, we note the probability of the event $E^{(0)}$ that

(i) the queue length at the beginning of τ_0 is $n_0 - 1$;

(ii) the state vector of the service process is $(i_0, s_0 - 1)$ at the beginning of τ_0;

(iii) the repair process is in state r_0 at the beginning of τ_0; and

(iv) a customer arrives in τ_0;

is given approximately by

$$P_{(n_0-1)i_0(s_0-1)r_0 0}(\lambda\Delta t) + P_{(n_0-1)i_0(s_0-1)r_0 1}(\lambda\Delta t). \tag{4.5.1}$$

When $E^{(0)}$ has occurred, the queue length, service state vector and repair state at the end of τ_0 will be n_0, (i_0, s_0), and r_0, respectively. Thus we may denote the probability of $E^{(0)}$ by $P^{(0)}_{n_0 i_0 s_0 r_0}$. By using a method similar to that used in Section 4.3, it can be shown that

$$\begin{aligned} P^{(k)}_{0i00|n_0 i_0 s_0 r_0} &\cong \sum_{m=0}^{I-1} P^{(k-1)}_{1im0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(\mu_{i(m+1)}\Delta t) \\ &+ P^{(k-1)}_{1iI0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(\mu_{iI}\Delta t) \end{aligned} \qquad \text{for } 1 \le i < \alpha, \tag{4.5.2}$$

$$\begin{aligned} P^{(k)}_{n100|n_0 i_0 s_0 r_0} &\cong \sum_{m=\alpha}^{\beta} P^{(k-1)}_{nm0m|n_0 i_0 s_0 r_0}(\delta_m \Delta t) + \sum_{m=0}^{I-1} P^{(k-1)}_{(n+1)1m0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(\mu_{1(m+1)}\Delta t) \\ &+ P^{(k-1)}_{(n+1)1I0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(\mu_{1I}\Delta t) \end{aligned}, \tag{4.5.3}$$

$$P^{(k)}_{n1s0|n_0 i_0 s_0 r_0} \cong P^{(k-1)}_{n1(s-1)0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(1-\mu_{1s}\Delta t) \qquad \text{for } 1 \le s < I, \tag{4.5.4}$$

$$P^{(k)}_{n1I0|n_0 i_0 s_0 r_0} \cong P^{(k-1)}_{n1(I-1)0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t) + P^{(k-1)}_{n1I0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(1-\mu_{1I}\Delta t), \tag{4.5.5}$$

$$\begin{aligned} P^{(k)}_{ni00|n_0 i_0 s_0 r_0} &\cong \sum_{m=0}^{I-1} P^{(k-1)}_{(n+1)im0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(\mu_{i(m+1)}\Delta t) \\ &+ P^{(k-1)}_{(n+1)iI0|n_0 i_0 s_0 r_0}(1-\gamma\Delta t)(\mu_{iI}\Delta t) \end{aligned} \qquad \text{for } 2 \le i < \alpha, \tag{4.5.6}$$

$$P^{(k)}_{nis0|n_0 i_0 s_0 r_0} \cong \sum_{m=1}^{i-1} P^{(k-1)}_{nm(s-1)0|n_0 i_0 s_0 r_0} (\gamma\Delta t)(g_{i-m}) + P^{(k-1)}_{ni(s-1)0|n_0 i_0 s_0 r_0} (1-\gamma\Delta t)(1-\mu_{is}\Delta t) \quad \text{for } 2 \le i < \alpha,\ 1 \le s < I, \qquad (4.5.7)$$

$$P^{(k)}_{niI0|n_0 i_0 s_0 r_0} \cong \sum_{m=1}^{i-1} P^{(k-1)}_{nm(I-1)0|n_0 i_0 s_0 r_0} (\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P^{(k-1)}_{nmI0|n_0 i_0 s_0 r_0} (\gamma\Delta t)(g_{i-m}) + P^{(k-1)}_{ni(I-1)0|n_0 i_0 s_0 r_0} (1-\gamma\Delta t)(1-\mu_{iI}\Delta t) + P^{(k-1)}_{niI0|n_0 i_0 s_0 r_0} (1-\gamma\Delta t)(1-\mu_{iI}\Delta t) \quad \text{for } 2 \le i < \alpha, \qquad (4.5.8)$$

$$P^{(k)}_{nr0r|n_0 i_0 s_0 r_0} \cong \sum_{m=1}^{\alpha-1} \sum_{q=0}^{I} P^{(k-1)}_{nmq0|n_0 i_0 s_0 r_0} (\gamma\Delta t)(g_{r-m}) + P^{(k-1)}_{nr0r|n_0 i_0 s_0 r_0} (1-\delta_r \Delta t) \quad \text{for } \alpha \le r < \beta, \qquad (4.5.9)$$

and

$$P^{(k)}_{n\beta 0\beta|n_0 i_0 s_0 r_0} \cong \sum_{m=1}^{\alpha-1} \sum_{q=0}^{I} P^{(k-1)}_{nmq0|n_0 i_0 s_0 r_0} (\gamma\Delta t)\left(1 - \sum_{u=1}^{\beta-m-1} g_u\right) + P^{(k-1)}_{n\beta 0\beta|n_0 i_0 s_0 r_0} (1-\delta_\beta \Delta t)\,. \qquad (4.5.10)$$

When n = 0 at the end of τ_k, the service of the customer who arrives in τ_0 will have been completed in τ_k, and the sojourn time of the customer who arrives in τ_0 is given approximately by $k\Delta t$.

For k = 1, 2, …, we can use (4.5.2) to (4.5.10) to compute $P^{(k)}_{nisr|n_0 i_0 s_0 r_0}$ from the values of the $P^{(k-1)}_{n'i's'r'|n_0 i_0 s_0 r_0}$ where $n' = n$, $n + 1$. When the characteristics of the system at the end of τ_0 are given by n_0, i_0, s_0, and r_0, the probability that the customer who arrives in τ_0 has a sojourn time falling approximately in τ_k is given by

$$S^{(k)}_{n_0, i_0, s_0, r_0} = \sum_{i=1}^{\alpha-1} P^{(k)}_{0i00|n_0 i_0 s_0 r_0}\,. \qquad (4.5.11)$$

Thus the pdf of the sojourn time evaluated at $k\Delta t$ is given by

$$f_s(k\Delta t) \cong \left(\sum_{n_0=1}^{N} \sum_{(i_0, s_0, r_0)\in R_0} S^{(k)}_{n_0, i_0, s_0, r_0} P^{(0)}_{n_0 i_0 s_0 r_0} \right) \Big/ \left(\sum_{n_0=1}^{N} \sum_{(i_0, s_0, r_0)\in R_0} P^{(0)}_{n_0 i_0 s_0 r_0} \right), \qquad (4.5.12)$$

where $R_0 = \{(i_0, s_0, r_0) : 1 \le i_0 < \alpha, 1 \le s_0 \le I, r_0 = 0\} \cup \{(i_0, s_0, r_0) : \alpha \le i_0 \le \beta, s_0 = 0, r_0 = i_0\}$

and N is a large positive integer.

4.6 T-cycle

Let T, T_1 and T_2 be as defined in Section 3.7. We may use a method similar to that in Section 3.7 to find the distributions of T_1 and T_2, and find the expected value of T via the expected values of T_1 and T_2:

$$E[T] = E[T_1] + E[T_2].$$

4.6.1 Distribution of T_1

When the system is in the stationary state, the probability of the event $F_1^{(0)}$ that,

(a) the queue length at the beginning of τ_0 is n_0;

(b) the repair process is in state r_0 at the beginning of τ_0 where $\alpha \le r_0 \le \beta$; and

(c) a completion of repair occurs in τ_0;

is given approximately by

$$\sum_{r_0=\alpha}^{\beta} \left(P_{n_0 r_0 0 r_0 0}(\delta_{r_0} \Delta t) + P_{n_0 r_0 0 r_0 1}(\delta_{r_0} \Delta t) \right). \tag{4.6.1}$$

When $F_1^{(0)}$ has occurred, the queue length, service state vector and repair state at the end of τ_0 will be n_0, (1, 0), and 0, respectively. Thus we may denote the probability of $F_1^{(0)}$ by $P^{(0)}_{n_0 100}$. By using a method similar to that used in Section 4.3, it can be shown that,

when $n_0 = 0$,

$$P^{(k)}_{n_0 100 | n_0 100} \cong P^{(k-1)}_{n_0 100 | n_0 100}(1 - \gamma \Delta t), \tag{4.6.2}$$

$$P^{(k)}_{n_0 i00|n_0 100} \cong \sum_{m=1}^{i-1} P^{(k-1)}_{n_0 m00|n_0 100}(\gamma\Delta t)(g_{i-m}) + P^{(k-1)}_{n_0 i00|n_0 100}(1-\gamma\Delta t) \qquad \text{for } 2 \le i < \alpha, \qquad (4.6.3)$$

$$P^{(k)}_{n_0 r0r|n_0 100} \cong \sum_{m=1}^{\alpha-1} P^{(k-1)}_{n_0 m00|n_0 100}(\gamma\Delta t)(g_{r-m}) \qquad \text{for } \alpha \le r < \beta, \qquad (4.6.4)$$

$$P^{(k)}_{n_0 \beta 0\beta|n_0 100} \cong \sum_{m=1}^{\alpha-1} P^{(k-1)}_{n_0 m00|n_0 100}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right). \qquad (4.6.5)$$

For $n_0 \ge 1$,

$$P^{(k)}_{n_0 1s0|n_0 100} \cong P^{(k-1)}_{n_0 1(s-1)0|n_0 100}(1-\gamma\Delta t)\,, \qquad \text{for } 1 \le s < I, \qquad (4.6.6)$$

$$P^{(k)}_{n_0 1I0|n_0 100} \cong P^{(k-1)}_{n_0 1(I-1)0|n_0 100}(1-\gamma\Delta t) + P^{(k-1)}_{n_0 1I0|n_0 100}(1-\gamma\Delta t)\,, \qquad (4.6.7)$$

$$P^{(k)}_{n_0 i10|n_0 100} \cong P^{(k-1)}_{n_0 100|n_0 100}(\gamma\Delta t)(g_{i-1}) \qquad \text{for } 2 \le i < \alpha, \qquad (4.6.8)$$

$$P^{(k)}_{n_0 is0|n_0 100} \cong \sum_{m=1}^{i-1} P^{(k-1)}_{n_0 m(s-1)0|n_0 100}(\gamma\Delta t)(g_{i-m}) + P^{(k-1)}_{n_0 i(s-1)0|n_0 100}(1-\gamma\Delta t) \qquad \text{for } 2 \le i < \alpha,\ 2 \le s < I, \qquad (4.6.9)$$

$$P^{(k)}_{n_0 iI0|n_0 100} \cong \sum_{m=1}^{i-1} P^{(k-1)}_{n_0 m(I-1)0|n_0 100}(\gamma\Delta t)(g_{i-m}) + \sum_{m=1}^{i-1} P^{(k-1)}_{n_0 mI0|n_0 100}(\gamma\Delta t)(g_{i-m}) + P^{(k-1)}_{n_0 i(I-1)0|n_0 100}(1-\gamma\Delta t) + P^{(k-1)}_{n_0 iI0|n_0 100}(1-\gamma\Delta t) \qquad \text{for } 2 \le i < \alpha, \qquad (4.6.10)$$

$$P^{(k)}_{n_0 r0r|n_0 100} \cong \sum_{q=0}^{I} P^{(k-1)}_{n_0 1q0|n_0 100}(\gamma\Delta t)(g_{r-1}) + \sum_{m=2}^{\alpha-1}\sum_{q=1}^{I} P^{(k-1)}_{n_0 mq0|n_0 100}(\gamma\Delta t)(g_{r-m}) \qquad \text{for } \alpha \le r < \beta, \qquad (4.6.11)$$

and

$$P^{(k)}_{n_0 \beta 0\beta|n_0 100} \cong \sum_{q=0}^{I} P^{(k-1)}_{n_0 1q0|n_0 100}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-2} g_u\right) + \sum_{m=2}^{\alpha-1}\sum_{q=1}^{I} P^{(k-1)}_{n_0 mq0|n_0 100}(\gamma\Delta t)\left(1-\sum_{u=1}^{\beta-m-1} g_u\right). \qquad (4.6.12)$$

Suppose at the end of τ_k, the first component i of the service state vector (i, s) exceeds $\alpha-1$. The system will then be sent for repair, and the value of T_1 is given approximately by $k\Delta t$.

For $k = 1, 2, \ldots$, we can use (4.6.2) to (4.6.12) to compute $P^{(k)}_{n_0 isr|n_0 100}$ from the values of the $P^{(k-1)}_{n_0 i' s' r' | n_0 100}$. When the event $F^{(0)}_1$ has occurred, the probability that the server will deteriorate to a state which needs a repair at the end of τ_k is given approximately by

$$U^{(k)}_{n_0 100} = \sum_{r=\alpha}^{\beta} P^{(k)}_{n_0 r 0 r | n_0 100} . \tag{4.6.13}$$

Thus the pdf, evaluated at $k\Delta t$, of the time elapsed before the system is sent for repair again is given by

$$f_{T_1}(k\Delta t) \cong \left(\sum_{n_0=0}^{N} U^{(k)}_{n_0 100} P^{(0)}_{n_0 100} \right) \Bigg/ \left(\sum_{n_0=0}^{N} P^{(0)}_{n_0 100} \right) . \tag{4.6.14}$$

4.6.2 Distribution of T_2

When the system is in the stationary state, the probability of the event $F^{(0)}_2$ that,

(a) the queue length at the beginning of τ_0 is n_0;

(b) the service process is in state (i_0, s_0) at the beginning of τ_0 where $1 \le i_0 < \alpha$, $0 \le s_0 \le I$; and

(c) a shock with magnitude x occurs in τ_0 and deteriorates the first component i_0 of the service state vector to state i^* where $i^* = r_0$ and $\alpha \le r_0 \le \beta$;

is given approximately by

$$P^{(0)}_{n_0 r_0 0 r_0} \cong \begin{cases} \displaystyle\sum_{i_0=1}^{\alpha-1} \sum_{s_0=0}^{I} \left(P_{n_0 i_0 s_0 00} + P_{n_0 i_0 s_0 01} \right) (\gamma \Delta t)(g_{r_0 - i_0}) & \text{for } \alpha \le r_0 < \beta \\ \displaystyle\sum_{i_0=1}^{\alpha-1} \sum_{s_0=0}^{I} \left(P_{n_0 i_0 s_0 00} + P_{n_0 i_0 s_0 01} \right) (\gamma \Delta t) \left(1 - \sum_{u=1}^{r_0 - i_0 - 1} g_u \right) & \text{for } r_0 = \beta \end{cases} . \tag{4.6.15}$$

We note that n_0, (r_0, 0), r_0 appearing in the left term of (4.6.15) denote, respectively, the queue length, service state vector, and repair state at the end of τ_0. These characteristics at the end of τ_0 are the consequences of the occurrence of the event $F^{(0)}_2$. By using a method similar to that used in Section 4.3, it can be shown that

$$P^{(k)}_{n_0 100|n_0 r_0 0 r_0} \cong P^{(k-1)}_{n_0 r_0 0 r_0|n_0 r_0 0 r_0}(\delta_{r_0}\Delta t) \qquad \text{for } \alpha \le r_0 \le \beta, \tag{4.6.16}$$

and

$$P^{(k)}_{n_0 r_0 0 r_0|n_0 r_0 0 r_0} \cong P^{(k-1)}_{n_0 r_0 0 r_0|n_0 r_0 0 r_0}(1-\delta_{r_0}\Delta t) \qquad \text{for } \alpha \le r_0 \le \beta. \tag{4.6.17}$$

Suppose at the end of τ_k, the first component of the service state vector (i, s) is $i = 1$. Then the repair process is completed, and the value of T_2 is given approximately by $k\Delta t$.

For $k = 1, 2, \ldots$, we can use (4.6.16) and (4.6.17) to compute $P^{(k)}_{n_0 isr|n_0 r_0 0 r_0}$ from the values of the $P^{(k-1)}_{n_0 i's'r'|n_0 r_0 0 r_0}$. When the event $F_2^{(0)}$ has occurred, the probability that the repair process is completed at the end of τ_k is given approximately by

$$V^{(k)}_{n_0 r_0 0 r_0} = P^{(k)}_{n_0 100|n_0 r_0 0 r_0}\,. \tag{4.6.18}$$

Thus the pdf, evaluated at $k\Delta t$, of the time elapsed before the repair is completed is given by

$$f_{T_2}(k\Delta t) \cong \left(\sum_{n_0=0}^{N}\sum_{r_0=\alpha}^{\beta} V^{(k)}_{n_0 r_0 0 r_0} P^{(0)}_{n_0 r_0 0 r_0}\right) \Bigg/ \left(\sum_{n_0=0}^{N}\sum_{r_0=\alpha}^{\beta} P^{(0)}_{n_0 r_0 0 r_0}\right). \tag{4.6.19}$$

4.7 Numerical Examples

Consider again the case of a deteriorating M/M/1 queue with the same set of parameters as specified in the first example of Section 3.8: $\beta = 10$, $\mu_i = 8 - 0.7(i - 1)$ for $1 \le i \le \beta$ and $I = 3$, $\lambda = 4$, $\delta_r = 8 - 0.7(r - 1)$ for $\alpha \le r \le \beta$, $\gamma = 0.2$, and $g_i = (1 - p)p^i$ where $p = 0.5$. By using the proposed numerical method, the results for the stationary queue length distribution, mean queue length, mean sojourn time and expected T-cycle length are found. We may compare the results thus obtained with those computed by the method used in [35]. Simulation is again carried out to verify the results obtained. Some of the results obtained are shown in Tables 4.7.1 and 4.7.2.

Table 4.7.1

Comparison of stationary queue length distribution obtained by the proposed numerical method, those computed using matrix-geometric approach, and simulation procedure

Maintenance level, $\alpha = 4$

[$\Delta t = 10^{-9}$ for queue length distribution, $\Delta t = 10^{-3}$ for mean sojourn time and expected T-cycle length, $\mu_{is} = \mu_i$, $I = 3$, $N = 500$].

Queue Length, n	P(Queue Length = n)		
	Numerical method	Matrix-geometric approach	Simulation
0	0.425728	0.425728	0.427626
1	0.232254	0.232254	0.233299
2	0.130903	0.130903	0.130803
3	0.076396	0.076396	0.075977
4	0.046195	0.046195	0.045670
5	0.028905	0.028905	0.028482
6	0.018659	0.018659	0.018256
7	0.012377	0.012377	0.012109
8	0.008397	0.008397	0.008226
9	0.005801	0.005801	0.005662
10	0.004065	0.004065	0.003945
…	…	…	…
50	1.33E-08	1.33E-08	0
Mean Queue Length	1.551949	1.551949	1.533833
Mean Sojourn Time	0.388057	0.387987	0.385118
Expected T-Cycle Length	7.667336	7.667340	7.692020

Table 4.7.2

Comparison of stationary queue length distribution obtained by the proposed numerical method, those computed using matrix-geometric approach, and simulation procedure

Maintenance level, $\alpha = 9$

[$\Delta t = 10^{-9}$ for queue length distribution, $\Delta t = 10^{-3}$ for mean sojourn time and expected T-cycle length, $\mu_{is} = \mu_i$, $I = 3$, $N = 500$].

Queue Length, n	P(Queue Length = n)		
	Numerical method	Matrix-geometric approach	Simulation
0	0.364463	0.364463	0.365332
1	0.210426	0.210426	0.211206
2	0.127473	0.127473	0.127579
3	0.081149	0.081149	0.081466
4	0.054163	0.054163	0.054116
5	0.037704	0.037704	0.037418
6	0.027188	0.027188	0.026955
7	0.020169	0.020169	0.019792
8	0.015300	0.015300	0.015095
9	0.011812	0.011812	0.011594
10	0.009247	0.009247	0.009100
…	…	…	…
50	1.28E-05	1.28E-05	5.82E-06
Mean Queue Length	2.400651	2.400652	2.375517
Mean Sojourn Time	0.600259	0.600163	0.599452
Expected T-Cycle Length	9.914940	9.915090	9.909612

Tables 4.7.1 and 4.7.2 show that the results based on the proposed numerical method are very close to those obtained using the matrix-geometric approach and the simulation procedure.

Next, consider an example in which the service time has a gamma distribution with parameter (κ, θ) = (5/4, 2/25) and mean service time, $E(S_m)$ = 0.1. Suppose the other parameter settings are, β = 10, $f^{(i)} = (11-i)/10$ for $1 \le i \le \beta$, λ = 6, $\delta_r = r$ for $\alpha \le r \le \beta$, $\gamma = 0.08$, and $g_i = (1-p)p^i$ where p = 0.6. The results for the stationary queue length distribution, mean queue length, mean sojourn time and expected T-cycle length are computed using the proposed numerical method. The results obtained are shown in Table 4.7.3.

Table 4.7.3

Comparison of stationary queue length distribution computed using the proposed numerical method, and simulation procedure

[$\Delta t = 0.0015$ for queue length distribution, mean sojourn time and expected T-cycle length, $I = 400$, $N = 300$].

Maintenance Level, α	$\alpha = 4$		$\alpha = 6$	
Queue Length, n	P(Queue Length = n)		P(Queue Length = n)	
	Numerical method	Simulation	Numerical method	Simulation
0	0.346995	0.347099	0.306561	0.305688
1	0.230229	0.230297	0.210037	0.209710
2	0.142501	0.142571	0.135975	0.135877
3	0.088364	0.088324	0.089217	0.089100
4	0.055801	0.055698	0.060258	0.060191
5	0.036155	0.036009	0.042104	0.042067
6	0.024145	0.024128	0.030453	0.030420
7	0.016664	0.016676	0.022753	0.022827
8	0.011897	0.011903	0.017498	0.017626
9	0.008777	0.008754	0.013791	0.013884
10	0.006675	0.006678	0.306561	0.305688
…	…	…	…	…
50	1.34E-05	1.19E-05	2.96E-05	3.12E-05
Mean Queue Length	2.222147	2.221585	2.951975	2.969179
Mean Sojourn Time	0.374662	0.370374	0.496362	0.497001
Expected T-Cycle Length	18.682908	18.683166	22.688576	22.627509

Table 4.7.3 shows that when Δt = 0.0015, the results obtained using the proposed numerical method are close to the simulation results. The results based on the numerical method may be improved by using the extrapolation procedure described in Section 2.5.

Next use the formula in (3.8.1) to compute the average cost per unit time, $C(\alpha)$. Figure 4.7.1 shows the average cost per unit time for the system at different values of the maintenance level α and holding cost C_H. In Figure 4.7.2, the average costs are compared when the arrival rates are given by λ = 2, 4, 6 and fixed repair cost $C_R = 12$, respectively.

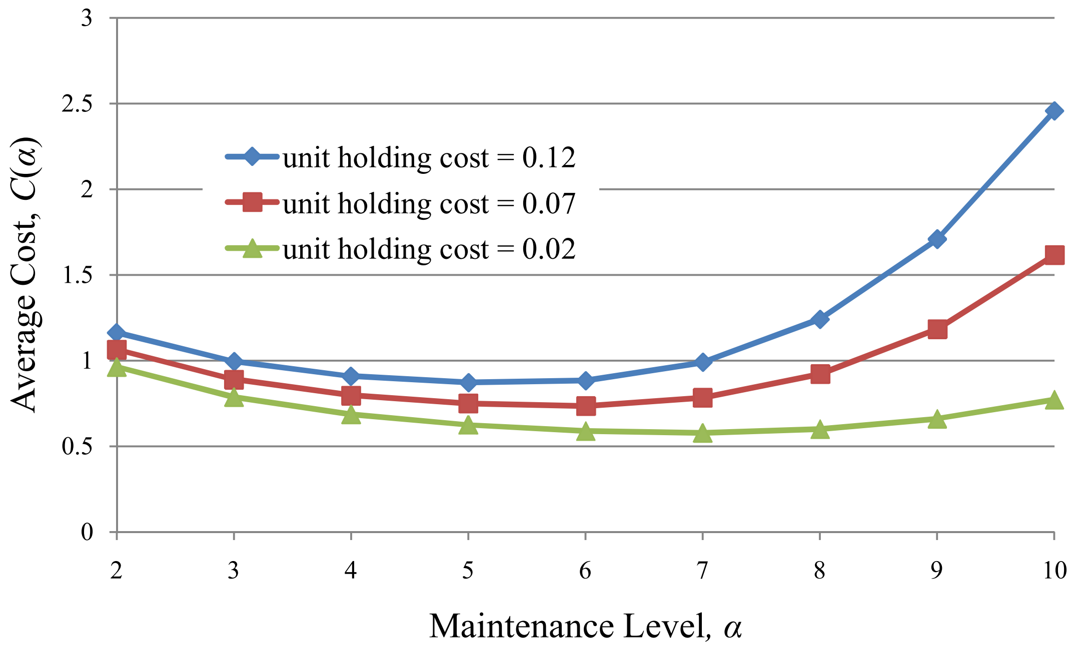

Figure 4.7.1 Average cost over maintenance level and unit holding cost [(κ, θ) = (5/4, 2/25), $\beta = 10$, $f^{(i)} = (11-i)/10$, $\lambda = 6$, $\delta_r = r$ for $\alpha \le r \le \beta$, $\gamma = 0.08$, $g_i = (1-p)p^i$ where $p = 0.6$ and $C_R = 12$].

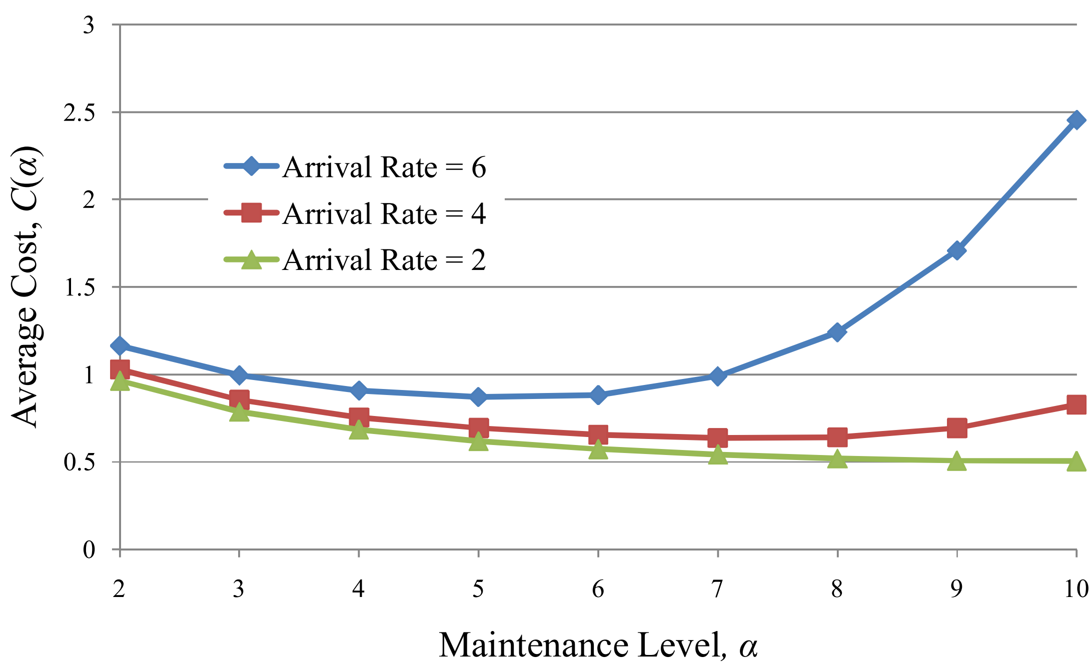

Figure 4.7.2 Average cost over maintenance level and arrival rate [(κ, θ) = (5/4, 2/25), $\beta = 10$, $f^{(i)} = (11-i)/10$, $\lambda = 6$, $\delta_r = r$ for $\alpha \le r \le \beta$, $\gamma = 0.08$, $g_i = (1-p)p^i$ where $p = 0.6$, $C_R = 12$ and $C_H = 0.12$].

Figure 4.7.1 shows that when the unit holding costs are 0.02, 0.07 and 0.12, the average cost is lowest when α = 7, 6 and 5, respectively. Thus the optimal maintenance level decreases as the unit holding cost C_H increases.

Figure 4.7.2 reveals that when the arrival rates are 2, 4, and 6, the average cost is lowest when α = 10, 7 and 5, respectively. Thus the optimal maintenance level increases as the mean of the arrival distribution increases.

4.8 Conclusion

In this chapter, the multi-state deteriorating single server queue given in [35] is studied again by assuming that the service time has a CAR distribution. The basic characteristics of the queue are evaluated and the optimal maintenance policy for the system is determined. Although the service time distribution used in this chapter is fairly general, the model may still be improved further. For example, we may take into account the deterioration due to usage by introducing a correlation structure between two consecutive service times.

CONCLUDING REMARKS

The thesis introduces a new methodology for finding the stationary queue length distribution in the one-server queue in which the distributions of the service time and interarrival time have respectively a fairly general distribution called the CAR distribution. It is shown that the proposed methodology can also be used to find the stationary waiting time distribution.

The proposed numerical method can be adapted to investigate a multi-state deteriorating single server queue in which the service rate deteriorates due to random shocks, and the interarrival time or service time in the queue is assumed to have a CAR distribution. Approximate results for the stationary queue length distribution, stationary sojourn time distribution and expected T-cycle length are found. More accurate results can be obtained by using a smaller value of the length Δt of the time interval, and the results can be improved by using extrapolation. It would be theoretically possible to apply the method to the multi-state deteriorating single server queue in which both the service time and interarrival time distributions have respectively a constant asymptotic rate. However we may encounter dimensionality problem as the procedure involves the solution of a large number of equations.

The proposed numerical method may also be used to study other more general queueing systems such as a system involving two or more queues, or a system of which the consecutive service times are correlated.

REFERENCES

[1] Erlang, A. K. (1909). The theory of probabilities and telephone conversations. *Nyt Tidsskrift for Matematik B*, **20**, 33–39.

[2] Kendall, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov Chain. *Annals of Mathematical Statistics*, **24**(3), 338–354.

[3] Lee, A. M. (1966). *Applied Queueing Theory*. New York: MacMillan.

[4] Ackroyd, M. H. (1980). Computing the waiting time distribution for the G/G/1 queue by signal processing methods. *IEEE Transactions on Communications*, **28**(1), 52–58.

[5] Fryer, M. J., & Winsten, C. B. (1986). An algorithm to compute the equilibrium distribution of a one-dimensional bounded random walk. *Operations Research*, **34**(3), 449–454.

[6] Grassmann, W. K., & Jain, J. L. (1989). Numerical solutions of the waiting time distribution and idle time distribution of the arithmetic GI/G/1 queue. *Operations Research*, **37**(1), 141–150.

[7] Haβlinger, G. (2000). Waiting time, busy periods and output models of a server analyzed via Wiener-Hopf factorization. *Performance Evaluation*, **40**(1–3), 3–26.

[8] Konheim, A. G. (1975). An elementary solution of the queueing system GI/G/1. *SIAM Journal on Computing*, **4**, 540–545.

[9] Ponstein, J. (1974). Theory and solution of a discrete queueing problem. *Statistica Neerlandica*, **20**(3), 139–152.

[10] Alfa, A. S., & Li, W. (2001). Matrix-geometric analysis of the discrete time GI/G/1 system. *Stochastic Models*, **17**(4), 541–554.

[11] Alfa, A. S. (2003). Combined elapsed time and matrix-analytic method for the discrete-time GI/G/1 and GI^X/G/1 systems. *Queueing Systems*, **45**(1), 5–25.

[12] Rao, B. V., & Feldman, R. M. (1999). Numerical approximations for the steady-state waiting times in a GI/G/1 queue. *Queueing Systems*, **31**(1/2), 25–45.

[13] Kim, N. K., & Chaudhry, M. L. (2008). The use of the distributional Little's law in the computational analysis of discrete-time GI/G/1 and GI/D/c queues. *Performance Evaluation*, **65**(1), 3–9.

[14] Neuts, M. F. (1981). *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Baltimore: The Johns Hopkins University Press.

[15] White, H. C., & Christie, L. S. (1958). Queuing with preemptive priorities or with breakdown. *Operations Research*, **6**(1), 79–95.

[16] Avi-Itzhak, B., & Naor, P. (1963). Some queueing problems with the service station subject to breakdown. *Operations Research*, **11**(3), 303–322.

[17] Gaver, D. P. (1962). A waiting line with interrupted service including priorities. *Journal of Royal Statistical Society Series B*, **24**(1), 73–90.

[18] Mitrany, I. L., & Avi-Itzhak, B. (1968). A many-server queue with service interruptions, *Operations Research*, **16**(3), 628–638.

[19] Thiruvengadam, K. (1963). Queueing with breakdowns. *Operations Research*, **11**(1), 62–71.

[20] Barlow, R. E., & Hunter, L. C. (1960). Optimum preventive maintenance policies. *Operations Research*, **8**(1), 90–100.

[21] Boland, P. J., & Proschan, F. (1982). Periodic replacement with increasing minimal repair costs at failure. *Operations Research*, **30**(6), 1183–1189.

[22] Nakagawa, T. (1981). A summary of periodic replacement with minimal repair at failure. *Journal of the Operations Research Society of Japan*, **24**(3), 213–227.

[23] Nakagawa, T., & Kowada, M. (1983). Analysis of a system with minimal repair and its application to replacement policy. *European Journal of Operational Research*, **12**(2), 176–182.

[24] Pham, H., & Wang, H. (1996). Imperfect maintenance. *European Journal of Operational Research*, **94**(3), 425–438.

[25] Sheu, S. H., Griffith, W. S., & Nakagawa, T. (1995). Extended optimal replacement model with random minimal repair costs. *European Journal of Operational Research*, **85**(3), 636–649.

[26] Ahmadi, R., & Newby, M. (2011). Maintenance scheduling of a manufacturing system subject to deterioration. *Reliability Engineering & System Safety*, **96**(10), 1411–1420.

[27] Chen, J., & Li, Z. (2008). An extended extreme shock maintenance model for a deteriorating system. *Reliability Engineering & System Safety*, **93**(8), 1123–1129.

[28] Hsieh, C. C., & Chiu, K. C. (2002). Optimal maintenance policy in a multistate deteriorating standby system. *European Journal of Operational Research*, **141**(3), 689–698.

[29] Hu, Q. (1995). The optimal replacement of a Markov deteriorative under stochastic shocks. *Microelectron Reliability*, **35**(1), 27–31.

[30] Lam, Y., & Zhang, Y. L. (2004). A shock model for the maintenance problem of a repairable system. *Computers & Operations Research*, **31**(11), 1807–1820.

[31] Li, W. J., & Pham, H. (2005). An Inspection-Maintenance Model for Systems with Multiple Competing Processes. *IEEE Transactions on Reliability*, **54**(2), 318–327.

[32] Sheu S. H., Lin Y. B., & Liao G. L. (2006). Optimum policies for a system with general imperfect maintenance. *Reliability Engineering & System Safety*, **91**(3), 362–369.

[33] Wang, H., & Pham, H. (1999). Some maintenance models and availability with imperfect maintenance in production systems. *Annals of Operations Research*, **91**(1), 305–18.

[34] Wortman, M. A., Klutke, G. A., & Ayhan, H. (1994). A maintenance strategy for systems subjected to deterioration governed by random shocks. *IEEE Transactions on Reliability*, **43**(3), 439–445.

[35] Yang, W. S., Lim, D. E., & Chae, K. C. (2009). Maintenance of deteriorating single server queues with random shocks. *Computers & Industrial Engineering*, **57**(4), 1404–1406.

[36] Yang, W. S., Lim, D. E., & Chae, K. C. (2011). Maintenance of multi-state production systems deteriorated by random shocks and production. *Journal of Systems Science and Systems Engineering*, **20**(1), 110–118.

[37] Cho, D. I., & Parlar, M. (1991). A survey of maintenance models for multi-unit systems. *European Journal of Operational Research*, **51**(1), 1–23.

[38] McCall, J. J. (1965). Maintenance policies for stochastically failing equipment: A survey. *Management Science*, **11**(5), 493–524.

[39] Pierskalla, W. P., & Voelker, J. A. (1976). A survey of maintenance models: the control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, **23**(3), 353–388.

[40] Sherif, Y. S., & Smith, M. L. (1981). Optimal maintenance models for systems subject to failure: a review. *Naval Research Logistics Quarterly*, **28**(1), 47–74.

[41] Valdez-Flores, C., & Feldman, R.M. (1989). A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics*, **36**(4), 419–446.

[42] Wang, H. (2002). A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, **139**(3), 469–489.

[43] Esary, J. D., Marshall, A. W., & Proschan, F. (1973). Shocks models and wear processes. *The Annals of Probability*, **1**(4), 627–649.

[44] Gut, A. (1990). Cumulative shock models. *Advances in Applied Probability*, **22**(2), 504–507.

[45] Gut, A., & Hüsler, J. (2005). Realistic variation of shock models. *Statistics & Probability Letters*, **74**(2), 187–204.

[46] Sumita, U., & Shanthikumar, J. G. (1985). A class of correlated cumulative shock models. *Advances in Applied Probability*, **17**(2), 347–366.

[47] Gut, A., & Hüsler, J. (1999). Extreme shock models. *Extremes*, **2**(3), 293–305.

[48] Ross, S. M. (1981). Generalized Poisson shock models. *The Annals of Probability*, **9**(5), 896–898.

[49] Shanthikumar, J. G., & Sumita, U. (1983). General shock models associated with correlated renewal sequences. *Journal of Applied Probability*, **20**(3), 600–614.

[50] Li, Z. H. (1984). Some distributions related to Poisson processes and their application in solving the problem of traffic jam. *Journal of Lanzhou University*, **20**, 127–136.

[51] Li, Z., & Kong, X. (2007). Life behavior of δ-shock model. *Statistics & Probability Letters*, **77**(6), 577–587.

[52] Li, Z. H., Chan, L. Y., & Yuan, Z. Q. (1999). Failure time distribution under the δ-shock model and its application to economic design of systems. *International Journal of Reliability, Quality & Safety Engineering*, **6**(3), 237–247.

[53] Liang, X. L., Lam, Yeh., & Li, Z. H. (2011). Optimal replacement policy for a general geometric process model with δ-shock. *International Journal of Systems Science*, **42**(12), 2021–2034.

[54] Tang, Y. Y., & Lam, Y. (2006). A δ-shock maintenance model for a deterioration system. *European Journal of Operational Research*, **168**(2), 541–556.

[55] Igaki, N., Sumita, U., & Kowada, M. (1995). Analysis of Markov renewal shock models. *Journal of Applied Probability*, **32**(3), 821–831.

[56] Gut, A. (2001). Mixed shock models. *Bernoulli*, **7**(3), 541–555.

[57] Wang, Y. P., & Pham, H. (2011). Dependent competing-risk degradation systems. In H. Pham (Ed.), *Safety and Risk Modeling and its Applications* (pp. 197–218). London: Springer London.

[58] Anderson, M. Q. (1981). Monotone optimal maintenance policies for stochastically failing equipment. *Naval Research Logistics*, **28**(3), 347–358.

[59] Kolesar, P. (1966). Minimum cost replacement under Markovian deterioration. *Operations Research*, **12**(9), 694–706.

[60] Wood, A. P. (1988). Optimal maintenance policies for constantly monitored systems. *Naval Research Logistics*, **35**(4), 461–71.

[61] Luss, H. (1976). Maintenance policies when deterioration can be observed by inspections. *Operations Research*, **24**(2), 359–366.

[62] Mine, H., & Kawai, H. (1975). An optimal inspection and replacement policy. *IEEE Transactions on Reliability*, **24**(5), 305–309.

[63] Ohnishi, M., Kawai, H., & Mine, H. (1986). An optimal inspection and replacement policy for a deteriorating system. *Advanced Applied Probability*, **23**(4), 973–988.

[64] Sim, S. H., & Endrenyi, J. (1988). Optimal preventive maintenance with repair. *IEEE Transactions on Reliability*, **37**(1), 92–96.

[65] Feldman, R. M. (1976). Optimal replacement with semi-Markov shock models. *Journal of Applied Probability*, **13**(1), 108–117.

[66] Gottlieb, G. (1982). Optimal replacement for shock models with general failure rate. *Operations Research*, **30**(1), 82–92.

[67] Hsieh, C. C. (2005). Replacement and standby redundancy policies in a deteriorating system with aging and random shocks. *Computers & Operations Research,* **32**(9), 2297–2308.

[68] Kao, E. P. C. (1973). Optimal replacement rules when changes of states are semi-Markovian. *Operations Research*, **21**(6), 1231–1249.

[69] Lam, C. T., & Yeh, R. H. (1994). Optimal maintenance policies for deteriorating systems under various maintenance strategies. *IEEE Transactions on Reliability*, **43**(3), 423–430.

[70] Lee, E. Y., & Lee, J. (1993). A model for a system subject to random shocks. *Journal of Applied Probability*, **30**(4), 979–984.

[71] Park, Y. I., Chae, K. C., & Lee, H. S. (2000). A random review replacement model for a system subject to compound Poisson shocks. *Stochastic Analysis and Applications*, **18**(1), 145–157.

[72] Kaufman, D. L., & Lewis, M. E. (2007). Machine maintenance with workload considerations. *Naval Research Logistics*, **54**(7), 750–766.

[73] Chakravarthy, S. R. (2012). Maintenance of a deteriorating single server system with Markovian arrivals and random shocks. *European Journal of Operational Research*, **222**(3), 508–522.

[74] Gross, D., & Harris, C. M. (1998). *Fundamentals of Queueing Theory*. New York: Wiley.