# ANALYSIS OF SOME BATCH ARRIVAL QUEUEING SYSTEMS WITH BALKING, RENEGING, RANDOM BREAKDOWNS, FLUCTUATING MODES OF SERVICE & BERNOULLI SCHEDULLED SERVER VACATIONS

A Thesis Submitted in fulfillment of the Requirements for the Degree of
Doctor of Philosophy

Submitted by
Monita Baruah
Brunel University, United Kingdom

# ABSTRACT

The purpose of this research is to investigate and analyse some batch arrival queueing systems with Bernoulli scheduled vacation process and single server providing service. The study aims to explore and extend the work done on vacation and unreliable queues with a combination of assumptions like balking and re-service, reneging during vacations, time homogeneous random breakdowns and fluctuating modes of service. We study the steady state properties, and also transient behaviour of such queueing systems.

Due to vacations the arriving units already in the system may abandon the system without receiving any service (reneging). Customers may decide not to join the queue when the server is in either working or vacation state (balking). We study this phenomenon in the framework of two models; a single server with two types of parallel services and two stages of service. The model is further extended with re-service offered instantaneously.

Units which join the queue but leave without service upon the absence of the server; especially due to vacation is quite a natural phenomenon. We study this reneging behaviour in a queueing process with a single server in the context of Markovian and non-Markovian service time distribution. Arrivals are in batches while each customer can take the decision to renege independently. The non-Markovian model is further extended considering service time to follow a Gamma distribution and arrivals are due to Geometric distribution. The closed-form solutions are derived in all the cases.

Among other causes of service interruptions, one prime cause is breakdowns. We consider breakdowns to occur both in idle and working state of the server. In this queueing system the transient and steady state analysis are both investigated.

Applying the supplementary variable technique, we obtain the probability generating function of queue size at random epoch for the different states of the system and also derive some performance measures like probability of server's idle time, utilization factor, mean queue length and mean waiting time. The effect of the parameters on some of the main performance measures is illustrated by numerical examples to validate the analytical results obtained in the study. The *Mathematica* 10 software has been used to provide the numerical results and presentation of the effects of some performance measures through plots and graphs.

# Dedication

This thesis is dedicated in loving memory of my beloved father,

Who left me to rest in God's abode, before I could finish.

I intensely feel downhearted as I can never say to him

"I finished father"

He was the support behind me in every accomplishment of life,

I know you are still there for me, guiding and inspiring me.

No words can express your love and continuous encouragement in whatever I did, I miss you so much.

Thank you for being the best father a daughter could ever wish for

I love you with all my heart today and forever

# Acknowledgement

*Challenges are what make life interesting and overcoming them is what makes life meaningful*- Joshua J. Marim

With due humbleness and praise to almighty God for his blessings in my life to be what I am today. I thank God for giving me the strength, perseverance and patience to carry out my work successfully.

I extend my sincere thanks and gratitude to my privileged supervisor, Dr. Tillal Eldabi, for his support and overall assistance to complete my research. I appreciate his continuous encouragement and guidance and thank him for providing me his valuable time and suggesting valuable comments which helped me improve my thesis.

I am truly grateful to my privileged supervisor, Dr. Igor Smolyarenko for having faith in me and deciding to be my supervisor. I thank him for his invaluable knowledge, counseling and great effort to help me overcome the difficulties of learning and research. His insightful guidance, patience and constant encouragement made me accomplish the research in this thesis. I truly appreciate him for the interest he has taken in answering my questions whenever needed and helped me in developing both professionally and personally.

My sincere gratitude is also due to Prof Kailash C. Madan, for his generosity, extraordinary support and guidance throughout the whole process of this Ph.d Programme.

My heartiest thanks go to my family, my son Jyotishman and daughter Reedhima for their love and affection though they sacrificed a lot of my company during this period. I am ever grateful to my husband, Biraj M. Das, for his seemingly endless love, support and confidence in me which has been a great source of inspiration.

Last but not the least; I extend my gratitude to my parents, my sisters and brother for continuous support and encouragement. It is for their warm love, support and prayers that pushed me forward to achieve my goal. Thank you for the patience, understanding and support throughout my life.

# Declaration

During the course of this research and Ph.D study, the following papers have been published.

1. Monita Baruah, Kailash C. Madan and Tillal Eldabi (2012). "Balking and Re-service in a Vacation Queue with Batch Arrival and Two Types of Heterogeneous Service."Journal of Mathematics Research. Vol.4, No.4. 114-124.

2. Monita Baruah, Kailash C. Madan and Tillal Eldabi (2013)."An $M^X/(G_1,G_2)/1$ Vacation Queue with Balking and Optional Re-service." Applied Mathematical Sciences, Vol.17, N0.7, 837-856.
3. Monita Baruah, Kailash C. Madan and Tillal Eldabi (2013)."A Batch Arrival Queue with Second Optional Service and Reneging during Vacation Periods." Revista Investigacion Operacional, Vol.34, No.3, 244-258.

4. Monita Baruah, Kailash C. Madan and Tillal Eldabi (2013). "A Two Stage Batch Arrival Queue with Reneging during Vacation and Breakdown periods." American Journal of Operation Research, 3, .570-580.

5. Monita Baruah, Kailash C. Madan and Tillal Eldabi (2013). "A Batch Arrival Single Server Queue with Server Providing General Service in Two Fluctuating Modes and Reneging during Vacation and Breakdowns." Journal of Probability and Statistics, Hindawi Publishing Corporation.2014, (Article, ID.319318), 2014. Doi:/10.1155/2014/319318.

******

# CONTENTS

**Chapter 4: Analysis of a Batch Arrival Queueing System with Reneging during Server Vacations**



**Chapter 5: Transient and Steady State Solution of a Batch Arrival Vacation Queue with Time Homogeneous Breakdowns and Server Providing General Service in Two Fluctuating Modes**

**Chapter 6: Conclusions and Further Research**



******

# List of Tables

**Table Number** **Page**



******

# List of Figures

**Figure Number** **Page**



******

# Definitions and Notations

$\lambda$ : The average rate of units arriving in the queueing system.

$\mu$ : The average rate of units being served in the system.

$\rho$: A measure of traffic congestion for a single channel queuing system given by

$\rho = \frac{\lambda}{\mu}$, also known as utilization or load factor.

$P_m(t)$: Transient state probability of having *m* customers in the system at time *t*

$P_m$ : Steady state probability of having *m* customers in the system

$N(t)$: Random variable describing the total number of customers in the system, at time *t*

$N_q(t)$: Random variable describing the number of customers in the queue, at time *t*

T : Random variable representing the time it spends in the system.

$T_q$ : Random variable representing the time it spends in the queue, before receiving service.

S: Random variable representing the service time.

$V$ : Random variable representing completion of vacation time.

$R$ : Random variable representing completion of repair times.

L: The mean number of customers in the system, $E(N) = L$

$L_q$ : The mean number of customers in the queue, $E(N_q) = L_q$

$L_V$ : The mean length of the queue during server vacations.

$W$ : The mean waiting time in the system, $E(T) = W$

$W_q$ : The mean waiting time in the queue, $E(T_q) = W_q$

$P_n(x,t)$: Probability that at time *t*, the server is providing a service, there are $n$ *(≥ 0)* customers in the queue excluding the one customer in service and the elapsed service time of this customer is *x*.

$P_n(t)$: $P_n(t) = \int_0^\infty P_n(x,t)dx$, denotes the probability that at time *t*, the server is providing a service , there are $n$ (≥ 0) customers in the queue excluding the one being served, irrespective of the value *x*.

$V_n(x,t)$: Probability that at time *t*, the server is on vacation with elapsed vacation time *x* and there are $n\ (\geq 0)$ customers in the queue waiting for service.

$V_n(t)$: $V_n(t) = \int_0^\infty V_n(x,t)dx$, denotes the probability that at time *t*, the server is on vacation with $n(\geq 0)$ customers in the queue irrespective of the value *x*.

$R_n(x,t)$: Probability that at time *t*, the server is inactive and under repairs with elapsed repair time *x* and there are $n\ (\geq 1)$ customers in the queue waiting for service.

$R_n(t)$: $R_n(t) = \int_0^\infty R_n(x,t)dx$, denotes the probability that at time *t*, the server is under repairs with $n.(\geq 1)$ customers in the queue irrespective of the value *x*.

$P_{n,j}(x,t)$: Probability that at time *t*, there are $n\ (\geq 0)$ customers in the queue excluding the one receiving type j (j=1,2) service and the elapsed service time of this customer is *x*.

$P_{n,j}(t)$: $P_{n.j}(t) = \int_0^\infty P_{n,j}(x,t)dx$, denotes the probability that at time *t*, the server is providing a service, there are $n$ (≥ 0) customers in the queue excluding the one being served in type j (j=1,2), irrespective of the value *x*.

$P_n^{(j)}(x,t)$: Probability that at time $t$, there are $n\ (\geq 0)$ customers in the queue excluding the one receiving service in jth stage (j=1,2) and the elapsed service time of this customer is $x$.

$P_n^{(j)}(t)$: $P_n^{(j)}(t)=\int_0^\infty P_n^{(j)}(x,t)dx$, denotes the probability that at time $t$ , the server is providing a service, there are $n$ ($\geq$ 0) customers in the queue excluding the one being served in jth stage (j=1,2), irrespective of the value $x$.

$P_n^{(m_j)}(x,t)$: Probability that at time $t$, there are $n\ (\geq 0)$ customers in the queue excluding the one receiving service in mode j (j=1,2) service and the elapsed service time of this customer is $x$.

$P_n^{(m_j)}(t)$: $P_n^{(m_j)}(t)=\int_0^\infty P_n^{(m_j)}(x,t)dx$, denotes the probability that at time $t$, there are $n$ ($\geq$ 0) customers in the queue excluding the one being served in mode j (j=1,2), irrespective of the value $x$.

$Q(t)$: Probability that at time $t$, there are no customers in the system and the server is idle but available in the system.

******

# Chapter 1
# INTRODUCTION

## 1.1 Introduction

A flow of customers from infinite/finite population towards the service facility for receiving some kind of service forms a queue. Generally, waiting is an unpleasant experience. Waiting can be beneficial to society if both the units, one that waits and one that offers service can be managed effectively. Waiting is not only experienced by humans but also experienced in different queueing areas in our highly developed urbanized society. Queues are experienced in supermarkets to check out, vehicles waiting in traffic intersection, patients waiting in doctor's clinic for treatment, units completing work in one station waiting to access the next in a manufacturing unit of multiple work stations and so on. A few more examples where queues are common is jobs waiting to be processed in a communication system, aircraft waiting for landing (take-off) in a busy airport, merchandise waiting for shipment in a yard, calls waiting in a call center, engineering and industrial plants etc.

Queueing theory thus, is the mathematical study of waiting lines or queues. It is an important branch of Mathematics with applied probability, statistical distribution, calculus, matrix theory and complex analysis. It also falls under the area of decision science.

Queueing analysis is a mathematical model which represents the process of arrival of customers forming a queue if service is not immediately available, service process and the time taken to serve the customers. The term 'customers' used, is in the generic sense for all arriving units.

The model of a basic queueing process is as shown below:

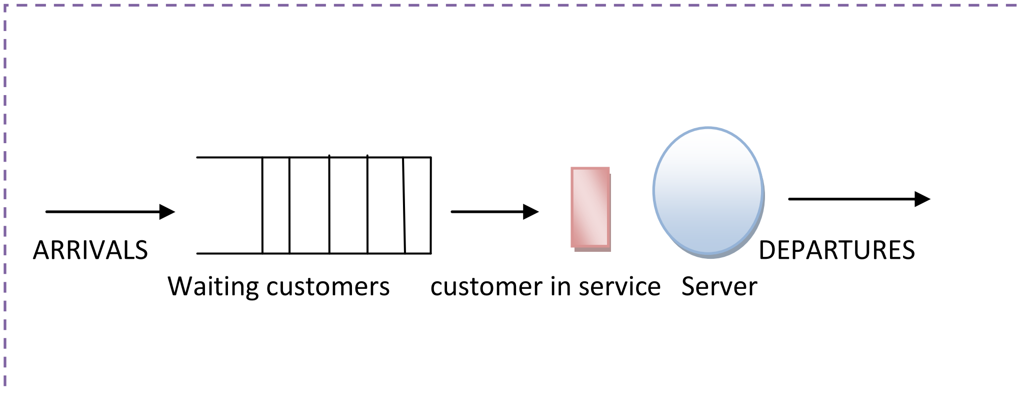


*Figure 1.1: A Classical Queueing Process*

Therefore, queueing theory is a tool for studying and analyzing different components of a queueing system and evaluating mathematical results in terms of the different performance measures. The results of queueing theory are required to obtain the characteristics of the model and to assess the effect of changes in the system. It contributes in providing vital information required for a decision maker by predicting various characteristics of the waiting line such as average waiting time in the queue, average number of customers in the queue etc.

## 1.2 Characteristics of a Queueing System

The basic features characterizing a queueing system are (i) the input (ii) the service mechanism (iii) the queue discipline (iv) service channels and (v) system capacity.

(i) **Input or Arrival Pattern:** it describes the way how units arrive and join a system. The arrivals can be one by one or in batches. The source of units may be finite or infinite. The inter-arrival time is the interval between two consecutive arrivals. In case, the arrival times are known with certainty, the queueing problems are categorized as deterministic models. However, in usual queueing situations, the process of arrivals is stochastic and it is necessary to know the probability distribution associated with the successive arrivals (inter-arrival times). The most common stochastic queueing models assume that inter-arrival times follow an Exponential distribution. The arrival pattern also describes the behaviour of the customers as some customers may wait patiently in the queue and some may be impatient if it takes a long time to receive the desired service. If an arriving customer decides not to join the queue, the customer is said to have balked. If a customer leaves the queue after joining due to impatience, it is known as reneging. In case there are two or more parallel waiting lines and a customer moves from one queue to another, the customer is said to have jockeyed. An arrival process could be stationary or non-stationary according to the probability distribution describing the arrival pattern being time-independent or dependent of time.

(ii) **Service Mechanism:** The service mechanism is concerned with service time and service facilities. It describes the way how service is rendered to an arriving unit. A unit may be served one at a time or in batches and the time required for servicing one unit or units in batches is called the service time. The service rate may depend on the number of customers in the queue. A server may work faster if he sees that the queue is building up, or conversely, he may get flustered and become less efficient. The situation in which service depends on the number of customers waiting is referred to as state-dependent service. (Gross and Harris, 1985).

(iii) **Queue Discipline**: it indicates the manner in which the units form the queue and are being offered service. The most common discipline is 'first come, first served' according to which the customers are served in order of their arrival. The following are the various queue disciplines:

- FIFO: first in, first out (or first come, first served, *FCFS*)
- LIFO: last in, first out, usually seen in a warehouse where the items that come last are taken out first.
- SIRO: service is provided in random order.
- Round Robin: Here every customer has a time slice for the service offered. If the service is not completed within this time then the customer is preempted and returns back to the queue to be served again according to FCFS discipline.

- ❖ Priority disciplines: Under this discipline, the service offered is of two types- Pre-emptive priority and Non pre-emptive priority. Under pre-emptive rule, high priority customers are given preference over low-priority customers. As such low priority customer's service is interrupted (pre-empted) to offer service to a priority customer. Once the high priority customer completes his service, the interrupted service is resumed again. Under Non pre-emptive rule, the highest priority customers go ahead in the queue but his service starts only after the completion of service of the customer in service.

(iv) **Service Channels**: A queueing system may have a single service station or a number of parallel service channels which can serve customers simultaneously. It is generally assumed that the service mechanisms of the parallel channels operate independently of each other.

A multi-channel system could be fed with a single queue or each channel may have separate queues. An arrival who finds more than one free server can choose any one of them for receiving service.

(v) **Stages of Service**: A customer may proceed through one stage or different stages for completing a service. A queueing system may have a single stage of service such as in supermarkets or a number of stages for service. In a queueing system with multiple stages, a customer joins the queue, waits for service, gets served, departs and moves to the new queue to receive the next stage of service and so on. An example of a multistage queueing system is seen in airports where passengers should proceed through different stages of service, like taking boarding tickets, immigration, security check etc.

(vi) **Capacity of the System**: A system may have maximum queue size. The capacity may be finite or infinite. A finite source restricts the number of customers in the queue to a certain limit while an infinite source does not have any limit to the size of the queue. When an arrival is not allowed to join the system due to the physical limitation of waiting room, it is called delay or loss system.

## 1.3 Queue Notation

The notation *A/B/C* denoting a queuing model was designed by Kendall (1951) where the three descriptors A, B and C denote the following:

- A: Inter-arrival distribution
- B: Service time distribution
- C: Number of service channels

The notations A and B usually represent the following distributions:

- M: Exponential( Markovian) distribution
- $E_k$: Erlang-k distribution
- G: General (arbitrary) distribution
- D: Deterministic(fixed)distribution
- GI: General Independent distribution

Kendall's notation was further extended by Lee (1966) as *A/B/C/D/E* to represent a wider range of queueing systems. Here D represents the maximum size of the queueing system and E denotes the queue discipline like *FIFO*, *LIFO* and so forth. For example:

- M/G/1: indicates a single channel queueing system having Exponential inter-arrival time distribution and arbitrary (General) service time distribution
- M/D/1/∞/FCFS: is a queueing system with Exponential inter-arrival times, Deterministic service times, a single server, no restriction on the maximum number of customers allowed in the system, and first come first served queue discipline.

The current research deals with $M^X/G/1$ queueing system that is Poisson arrivals (Exponential inter-arrival times), General Service time distribution, single server, infinite population and first come first served discipline.

The superscript 'X' refers to the customers arriving at the system in batches of variable size.

## 1.4 Performance Measures

There are many parameters which measure the effectiveness of a stochastic queueing system. Some of the parameters which are of interest for a customer arriving at a queue are mean response time and mean number of customers in the queue while some of the measures of interest to the service provider are the probability distribution of the service utilization and service cost.

The most relevant performance measures in analyzing a queueing system are

- Expected number of customers in the system denoted by $L$, is the average number of customers, both waiting and in service.
- Expected number of customers in the queue denoted by $L_q$, is the average number of customers waiting in the queue.
- Expected waiting time in the system denoted by $W$, is the average time spent by a customer in the system. (i.e. waiting time plus service time)
- Mean waiting time in the queue denoted by $W_q$, is the average waiting time spent in the queue before the commencement of service.
- The server utilization factor (or busy period) denoted by $\rho$ is the proportion of time a server actually spends with the customer. The server utilization factor is also known as traffic intensity.

The information about the performance measures of the system enables a queueing analyst to determine the appropriate measures of effectiveness of the system and design an optimal (according to some criterion) system (Gross & Harris, 1998).

## 1.5 Server Vacations

The temporary absence of the server(s) for a certain period of time in a queueing system at a service completion instant when there are no customers waiting in a queue or even when there are some is termed as server(s) vacation. Arrivals coming and waiting for service can avail the service once the server completes the vacation period. There are many situations which may lead to a server vacation for example system maintenance, machine failure, resource sharing, cyclic servers (where server is supposed to serve more than one queue) etc.

Various types of vacation policies are seen in the literature:

- Single vacation model: Here the server goes for vacation after the end of each busy period and returns back immediately after the vacation period is over even if the system is empty at that time. For example, maintenance of machines in a production process is considered as vacation.
- Multiple vacation models: here the server takes a sequence of vacations until it finds at least $k$ or more units waiting in the system. For example, a system engaged in computer and communication systems has to undertake preventive maintenance for occasional periods of time besides doing their primary job like receiving, processing and transmitting data.
- Gated vacation: When the server places a gate behind the last waiting customer and serves only those customers who are within the gate following some rules.
- Limited service discipline: Here the server takes a vacation after serving $k$ consecutive customers or after a time length $t$ or if the server is idle.
- Exhaustive service discipline: Here the server goes on serving until the system is empty, after which it takes a vacation of a random length of time.

In the current research, we assume the single vacation policy following Bernoulli schedule in all the queueing models under study.

## 1.6 Customer's Behavior

Customer's behavior has an effective role in the study of queueing systems. Customer's behavior usually relate to impatience while waiting for service. The three main forms of customer's impatience are balking, reneging and jockeying.

- Balking: the reluctance of the customer to join the queue on arrival either because the queue is too long or there is not sufficient waiting space.
- Reneging: this occurs when a customer leaves the queue after a long wait due to impatience.
- Jockeying: customers moving back and forth among several sub-queues before each of the multiple channels. This is often seen in cash counters of supermarkets with two or more parallel waiting lines and customers have a tendency to switch from one to another in order to get served faster.

The two most important forms of impatience are balking and reneging. Balking was first introduced in a queueing model by Haight (1957), where he investigated an M/M/1 queue with balking. Since then, queues with balking and reneging have gained significant importance and studied by several authors, prominent among them are Ancker et al. (1963), Haghighi et al. (1986), Zhang et al. (2005), Altman and Yechiali (2006), Choudhury and Medhi (2011), to mention a few. In the literature, it is seen that balking and reneging have mainly been treated with Markovian queueing systems with single arrivals. The author here proposes to introduce both balking and reneging in non-Markovian queueing systems. Reneging is considered only to occur during the unavailability of the server due to vacations, while balking is assumed to occur both during busy or vacation state. In most real world queueing situations, it is seen that customers seem to get discouraged for receiving a service upon the absence of the server and tend to abandon the system without receiving any service. This phenomenon is most precisely witnessed when the server is on vacation. This, in turn, leads to potential loss of customers and customer goodwill to the service provider. The current thesis deals with reneging during server vacations following Bernoulli schedule.

## 1.7 Random Breakdowns

In the real world perfectly reliable servers may not always exist. In fact, servers at times may face unpredictable breakdowns. Hence the server will not be able to continue providing the service until the system is repaired. The customer receiving service returns back to the head of the queue or might even leave the system. The service is resumed once the system is repaired.

In many queuing systems, the server is a mechanical or electronic device such as computer, networks, ATM, traffic signals etc. which might sometimes be subjected to accidental random breakdown. The repair time may be assumed to follow Deterministic, Exponential, Hyper-exponential, General distributions etc. Breakdowns can occur also when the server is not working. Also in some real situations, the server cannot be repaired in a single stage. It may take a number of stages to be repaired completely. In some cases, the server may not immediately undergo a repair process; there may be a delay in entering the repair process. This is known as delay time for the repairs to start.

Queues with random breakdowns have gained significant attention in recent years and have been studied extensively. Authors like Avi-Itzhak and Naor (1963), Federgruen and Green (1986), Aissiani and Artalejo (1998), Wang, Cao and Li (2001), Tang (2003) and in most recent years Senthil and Arumuganathan (2010), Khalaf, Madan and Lukas (2011) have studied queues with breakdowns. Most of the literature mentioned here studies breakdowns which occur only during the working state of the server and the repair process starts instantaneously while repair consists of a single phase (stage) only. However, in many real life situations the service channel may be subjected to breakdown even when it is in the idle state and repairs may also be performed in more than one stage.

The current research studies queueing systems with batch arrival wherein the service channel is subjected to random breakdown both during busy state and an idle state, the server is taken up for repairs instantaneously.

## 1.8 Fluctuating Modes of service and Fluctuating Efficiency

Extensive amount of the queueing literature deals with the queueing system where server is providing service in the same mode, i.e. at the same mean rate to all its customers. However, in the real-world, this is not always true. Service provided to the customers may be slow, normal or fast. That is, a single server providing service may not be at the same rate to all customers. This is considered as fluctuating modes of service. For instance, in the case of an online server, there is fluctuation in the speed of internet, fluctuation in service efficiency in case of human servers especially those working in banks, call centers, customer service organizations etc. This, in turn, will have an effect on the efficiency of the queueing system. Due to fluctuating modes of service the efficiency of a system also fluctuates. Fluctuating modes of service, along with server vacations, random breakdowns are close representation of realistic queueing situations.

Madan (1989) had studied the time-dependent behavior of a single channel queue with two components with fluctuating efficiency. To the best of the researcher’s knowledge, queues with fluctuating efficiency and fluctuating modes of service still has to be studied in depth. Thus this assumption with breakdowns both during the working and idle state of the server has so far not treated in the literature of queueing theory, to the best of the author’s knowledge.

## 1.9 Related Mathematical Preliminaries

We describe below, in brief, the mathematics used in the analysis of queueing theory.

### 1.9.1 Probability Generating Function

The probability generating function is a very useful tool in the analysis of queueing theory**.**

If X is discrete random variable assuming the values n = 0, 1, 2…; with probability $p_n$ and then the probability generating function is defined as

$$P(z) = E(z^n) = \sum_{n=0}^{\infty} p_n z^n \ , \ |z| \le 1$$

Thus $P(1) = \sum_{n=0}^{\infty} p_n = 1$

If $p_n$ represents the probability that there are $n$ customers in the queue then the mean number of customers in the queue could be found using the probability generating function as follows

$$\sum_{n=0}^{\infty} n p_n = \frac{d}{dz} P(z) \bigg|_{z=1}$$

In queueing theory analyses, the probability generating function is quite often useful in deriving the equations involving system state probabilities.

### 1.9.2 Laplace and Laplace Stieltjes Transform

Laplace transform serve as powerful tools in many situations. They provide an effective means of solution to many problems arising in our study (Medhi, 1994).

The Laplace transform of a probability density function $f(t)$ for $t > 0$, is defined

$$\tilde{f}(s) = \int_0^\infty \exp(-st) f(t) dt = L\{f(t)\}$$

Let $F(t)$ be a well defined function for $t$ specified for $t \geq 0$ and s be a complex number, then Laplace-Stieltjes transform which is a generalization of Laplace transform, given by

$$L(s) = F^*(s) = \int_0^\infty \exp(-st) dF(t)$$

If the real function $F(t)$ can be expressed in terms of the following integral

$F(t) = \int_0^t dF(x) = \int_0^t f(x)dx$, then

$$F^*(s) = \int_0^\infty e^{-st} dF(t) ,$$

We can relate the Laplace transform of the distribution function in terms of the density function.

$$L\{F(x)\} = \int_0^\infty e^{-sx} F(x) dx$$

$$\tilde{F}(s) = \int_0^\infty e^{-st} \left\{ \int_0^x f(t)dt \right\} dx = L\left\{ \int_0^x f(t)dt \right\} = \frac{\tilde{f}(s)}{s}$$

Such that $\tilde{f}(s) = s\tilde{F}(s)$ . Further integrating $F^*(s)$ defined above by parts, we get

$$F^*(s) = s\tilde{F}(s) - f(0) = \tilde{f}(s) - f(0)$$

### 1.9.3 Traffic Intensity ( Utilization factor)

Assuming that λ is the mean arrival rate and μ is the mean service rate, an important measure of a queueing model (M/M/c) is its traffic intensity which is given by

$\rho = \dfrac{mean\ arrival\ rate}{mean\ service\ rate} = \dfrac{\lambda}{c\mu}$, ρ gives the fraction of time the server is busy.

The necessary condition for a steady state to exist is $\rho < 1\ (\lambda < c\mu)$, which implies that arrival rate < service rate which is the stability condition for an M/M/c queuing model. A steady state solution exists since the queue size will be under control.

If $\rho > 1\ (\lambda > c\mu)$, then arrival rate > service rate and consequently the number of units in the queue increases indefinitely as time passes on and there is no steady state (assuming that customers are not restricted from entering the system).

For a $M^X/M/1$ queueing system, the utilization factor is $\rho = \dfrac{\lambda E(X)}{\mu}$ <1, where $E(X)$ is the mean of a batch of arrivals of size '*x*'. For a queueing system with a non-Markovian service time distribution, the necessary and sufficient condition for stability to exist is $\rho = \lambda E(S)$ <1, where $E(S)$ is the mean of service time. Thus, the traffic intensity $\rho$ gives the proportion of time the server is in busy state.

### 1.9.4 Transient and Steady State Behavior

If the operating characteristics of a queueing system like input, output, mean queue length etc. are dependent on time, is said to be in a transient state. This usually occurs at the early stage of the operation of the system where its behavior is still dependent on the initial conditions.

If the characteristic of the queueing system becomes independent of time, then the *steady state* condition is said to exist. However, since we are interested in the long run behavior of the system, we intend to study steady state results of the queueing system.

Let $P_n(t)$ denote the probability that there are *n* customers in the system at time *t*, then in the steady state case, we have $\lim_{t\to\infty} P_n(t) = P_n$, $n = 0,1,2......$; whenever the limit exists.

Thus $P_n$ is the limiting probability that there are n units in the system, irrespective of the number at time 0. Whenever the limit exists the system is said to reach a steady (equilibrium) state. It is independent of time and $\{P_n\}$ is said to have a steady state or stationary or equilibrium distribution. In particular, $P_0$ denotes the proportion of time that the system is empty. It follows that $\sum_{n=0}^{\infty} P_n = 1$; this is called the normalizing condition (Medhi, 2003).

While the steady state results are suited for analyzing the performance of a system on a long time scale, the analytical results of the transient behavior of a queueing system are suited for studying the dynamical behavior of systems over a finite time horizon, especially when the parameters involved are perturbed. As such, in the current research, we intend to explore a queueing system in a transient regime since for a complete description of the stochastic behavior of a queue length process, time dependent results are considered useful.

### 1.9.5 Litte‘s Law-Relation Between Expected Queue Length and Expected Waiting Time

The Little’s Law is one of the most fundamental relationships in queueing theory developed by John D.C. Little in early 1960’s. It is the most widely used formula used in queueing theory. It establishes a relationship between the average number of customers in the system, mean arrival rate and mean response time (that is, the time between entering and leaving the system after finishing a service) in the steady state.

Liitle’s law states that an average number of customers in a system (over some interval) is equal to their average arrival rate, multiplied by their average time in the system.

That is $L = \lambda W$

where $\lambda$ is the mean rate of arrival; $L$ denotes the expected number of units in the system and W is the expected waiting time of the units in the system in steady state.

Similarly “the average number of customers in the queue (over some interval) is equal to their arrival rate, multiplied by their average time spent in the queue.”

That is, $L_q = \lambda W_q$, $L_q$ and $W_q$ denoting the expected number of units in the queue and expected waiting time in the queue respectively in the steady state.

The Eilon’s proof of this theorem is mentioned in Medhi (2003).

In case of queues with bulk arrivals in batches, the expected length of the queue can be derived as $L_q = \lambda E(X) W_q$, where $E(X)$ is the mean of a batch of size 'x' of arrivals to the system.

### 1.9.6 Supplementary Variable Technique

There are different techniques used in analyzing a queueing system with fairly general assumptions such as the imbedded Markov chain, matrix-geometric method and supplementary variable technique. Cox (1955) introduced the supplementary variable technique as a general technique to study the M/G/1 queueing system.

In the supplementary variable technique, a non-Markovian process is made Markovian by the inclusion of one or more supplementary variables. Under this technique queueing models become partial differential equations with integral boundary conditions. Later Jaiswal (1968), Henderson (1972), Dshalalaw (1998), Chaudhury et al. (1999) and many other authors used this technique.

The supplementary variable technique has become more important in transient solutions of non-Markovian systems. Compared with the imbedded Markov chain technique, this method is more straightforward to obtain the steady state probabilities at an arbitrary instant and practically interesting performance measures via the supplementary variable method (Niu &Takashi, 1999). Gupta and Sikdar (2006) used this technique to develop the relation between queue length distribution when the server is busy/vacation at arbitrary and departure epoch. They justified the advantage of using this technique over other methods by that one can obtain several other results by using simple algebraic manipulations of transform equations such as mean length of the idle period. Also the supplementary variable technique has the advantage over imbedded Markov chain by that here we can study the system in continuous time instead of discrete time point (Kashyap and Chaudhury, 1988).

In the current research, we assume that server takes vacations and breakdowns may occur at random which require repair process; additional supplementary variables are being introduced like elapsed vacation time, elapsed repair time and elapsed delay times.

Several authors have used this method in their analysis for queueing system involving general distribution. ( Frey and Takashi, 1999; Madan 2000a, 2000b, 2001; Wang, Cao &Li,2001; Ke, 2003a; Niu,Shu & Takashi, 2003; Arumuganathan & Jeykumar, 2005; Kumar & Arumuganathan, 2008;Maraghi, Madan & Dowman, 2009).

## 1.10 The M/M/1 Queueing System

In such a queueing system, the arrivals occur from an infinite source in accordance with a Poisson Process with parameter λ, the inter-arrival times are independent and exponentially distributed with mean $1/\lambda$ .There is a single server and service times are independent and exponentially distributed with parameter μ (mean $1/\mu$ ).

Accordingly we have the following probabilities for arrivals and service:

Pr (arrival occurs between $t$ and $t+\Delta t$ ) = $\lambda\Delta t + 0(\Delta t)$

Pr (more than one arrival between $t$ and $t+\Delta t$) = $0(\Delta t)$

Pr (no arrival between $t$ and $t+\Delta t$) = $1-\lambda\Delta t + 0(\Delta t)$

Pr (one service completion between $t$ and $t+\Delta t$) = $\mu\Delta t + 0(\Delta t)$

Pr (more than one service completion between $t$ and $t+\Delta t$) = $0(\Delta t)$

Pr (no service completion between $t$ and $t+\Delta t$) = $1-\mu\Delta t + 0(\Delta t)$

The aim is to calculate $P_m(t)$, the probability of $m$ arrivals up to time $t$. To do so, we start with finding the probability of the state at time $(t+\Delta t)$ as follows:

$$P_m(t+\Delta t) = (1-\lambda\Delta t)(1-\mu\Delta t)P_m(t) + \lambda\Delta t(1-\mu\Delta t)P_{m-1}(t) + (1-\lambda\Delta t)\mu\Delta t P_{m+1}(t), \ m \ge 1 \tag{1.1}$$

$$P_0(t+\Delta t) = (1-\lambda\Delta t)P_0(t) + (1-\lambda\Delta t)\mu\Delta t P_1(t) \tag{1.2}$$

Simplifying the above equations and ignoring higher order terms like $(\Delta t)^2$, we get

$$P_m(t+\Delta t) = (1-\lambda\Delta t - \mu\Delta t)P_m(t) + \lambda\Delta t P_{m-1}(t) + \mu\Delta t P_{m+1}(t), \tag{1.3}$$

$$P_0(t+\Delta t) = (1-\lambda\Delta t)P_0(t) + \mu\Delta t P_1(t) \tag{1.4}$$

The corresponding differential equations are found by transposing $P_m(t)$ to the left-hand side, dividing through by Δt and taking limit as Δt→ 0, we get

$$\frac{dP_m(t)}{dt} = -(\lambda+\mu)P_m(t) + \lambda P_{m-1}(t) + \mu P_{m+1}(t) \qquad m \ge 1 \tag{1.5}$$

$$\frac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t) \tag{1.6}$$

The steady state solutions of the differential equations can be obtained by the following conditions

$$\frac{dP_m(t)}{dt} = 0 \text{ and } \lim_{t \to \infty} P_m(t) = P_m$$

Writing $\lambda/\mu = \rho$ and $\rho < 1$ for stability and subject to the condition $\sum_{m=0}^{\infty} P_m(t) = 1$ for all *t*, we get the system state probabilities to be

$$P_m = \rho^m (1-\rho) \quad m \geq 0 \tag{1.7}$$

Using the steady state (equilibrium) probabilities $P_m$, various performance measures can be calculated as follows:

a) **Probability of finding the system empty on arrival**

$P_0 = 1 - \rho$

b) **Server Utilization**

The server is busy when there is at least one customer in the system. Hence the utilization of the server is $\rho = 1 - P_0$

c) **Mean number of customers in the system**

$$L = \sum_{m=0}^{\infty} mP_m = \sum_{m=0}^{\infty} m\rho^m (1-\rho) = (1-\rho)\sum_{m=0}^{\infty} m\rho^m = \frac{\rho}{(1-\rho)}$$

d) **Mean number of customers in the queue**

$$L_q = \sum_{m=1}^{\infty} (m-1)P_m = \frac{\rho^2}{(1-\rho)}$$

e) **Mean waiting time in the system**

$W$ can be computed using Little's law;

$$W = \frac{L}{\lambda} = \frac{1}{\mu(1-\rho)}$$

f) **Mean waiting time in the queue**

Similarly using Little's law, we calculate $W_q = \frac{L_q}{\lambda} = \frac{\rho}{\mu(1-\rho)}$

## 1.11 The $M/G/1$ Queueing System

The M/G/1 queueing system assumes that there is a single server with exponential inter-arrival times with mean arrival rate λ and service times are assumed to follow a general distribution.

Let $\mu(x)dx$ be the conditional probability density of service completion during the interval (x, x+dx] given that elapsed service time is *x*, so that

$$\mu(x) = \frac{b(x)}{1-B(x)} \tag{1.8}$$

Where $b(x)$ and $B(x)$ are the density function and distribution function of service time respectively. Accordingly, we have

$$B(s) = 1 - e^{-\int_0^s \mu(x)dx} \text{ and } b(s) = \mu(s)e^{-\int_0^s \mu(x)dx} \tag{1.9}$$

For steady state, we consider the limiting probability density

$P_n(x) = \lim_{t\to\infty} P_n(x,t)$ and the limiting probability

$$P_n = \lim_{t\to\infty} P_n(t) = \lim_{t\to\infty} \int_0^\infty P_n(x,t)dx$$

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

We have the following equations

$$P_n(x+\Delta x) = P_n(x)\left[1-\left(\lambda+\mu(x)\right)\Delta x\right] + P_{n-1}(x)\lambda\Delta x, \; n \ge 1 \tag{1.10}$$

$$P_0(x+\Delta x) = P_0(x)\left[1-\left(\lambda+\mu(x)\right)\Delta x\right] \tag{1.11}$$

$$Q = Q\left(1-\lambda\Delta x\right) + \int_0^\infty \mu(x)\Delta x P_0(x)dx \tag{1.12}$$

Hence the steady state equations governing the M/G/1 system are

$$\frac{d}{dx}P_n(x) + \left(\lambda+\mu(x)\right)P_n(x) = \lambda P_{n-1}(x), \; n \ge 1 \tag{1.13}$$

$$\frac{d}{dx}P_0(x) + \left(\lambda+\mu(x)\right)P_0(x) = 0 \tag{1.14}$$

$$\lambda Q = \int_0^\infty P_0(x)\mu(x)dx \tag{1.15}$$

The above equations are to be solved with the following boundary conditions

$$P_n(0)=\int_0^\infty P_{n+1}(x)\mu(x)dx,\ n\geq 1 \tag{1.16}$$

$$P_0(0)=\int_0^\infty P_1(x)\mu(x)dx+\lambda Q \tag{1.17}$$

And the normalizing condition is $Q+\sum_{n=0}^{\infty}\int_0^\infty P_n(x)dx=1$

The generating functions are defined as

$$\begin{aligned} P_q(x,z)&=\sum_{n=0}^{\infty} z^n P_n(x) \\ P_q(z)&=\sum_{n=0}^{\infty} z^n P_n \end{aligned} \tag{1.18}$$

Multiplying equations (1.13) by $z^n$, taking summation over $n$ from 1 to $\infty$, adding to (1.14) and using the generating functions defined in (1.18) we get

$$\frac{d}{dx}P_q(x,z)+(\lambda-\lambda z+\mu(x))P_q(x,z)=0 \tag{1.19}$$

Similarly from (1.16) and (1.17) we obtain

$$zP_q(0,z)=\int_0^\infty P_q(x,z)\mu(x)dx+\lambda(z-1)Q \tag{1.20}$$

Solving (1.19) and (1.20), we get

$$P_q(x,z)=P_q(0,z)e^{-(\lambda-\lambda z)x-\int_0^x \mu(t)dt} \tag{1.21}$$

Integrating (1.21) by parts we get

$$P_q(z)=P_q(0,z)\left[\frac{1-B^*(\lambda-\lambda z)}{\lambda-\lambda z}\right] \tag{1.22}$$

where $B^*(\lambda-\lambda z)$ is the Laplace Stieltjes transform of the service time. Now multiplying equation (1.21) by $\mu(x)dx$ and then integrating over $x$, we get

$$\int_0^\infty P_q(x,z)\mu(x)dx=P_q(0,z)B^*(\lambda-\lambda z) \tag{1.23}$$

Using (1.23) in equation (1.20) we get

$$P_q(0,z)=\frac{\lambda(z-1)Q}{z-B^*(\lambda-\lambda z)} \tag{1.24}$$

Now substituting $P_q(0,z)$ in (1.22) and using the normalizing condition to obtain $Q$, we have

$$P_q(z) = \frac{\left[1 - B^*(\lambda - \lambda z)\right]\left[1 - \lambda E(S)\right]}{B^*(\lambda - \lambda z) - z} \tag{1.25}$$

where $E(S)$ is the mean service time. Equation (1.25) gives the probability generating function of the number of customers in the queue. Using the relation $L_q = \frac{d}{dz} P_q(z)\Big|_{z=1}$ and equation (1.25), various performance measures can be derived.

**a) Probability of finding the system empty on arrival**

$Q = 1 - \rho = 1 - \lambda E(S)$

**b) Server Utilization**

$\rho = 1 - Q = \lambda E(S)$

**c) Mean number of customers in the queue**

Let us write the equation (1.18) in the form $P_q(z) = \frac{N(z)}{D(z)}$ where $N(z)$ and $D(z)$ are the numerator and denominator on the R.H.S of (1.18). Thus we have

$$L_q = \frac{d}{dz} P_q(z)\Big|_{z=1} = \frac{D(1)N'(1) - N(1)D'(1)}{\left(D'(1)\right)^2}$$

This is of 0/0 form since $N(1) = D(1) = 0$. So using L'Hopital's Rule twice we get

$$\begin{aligned} L_q &= \lim_{z\to 1} \frac{d}{dz} P_q(z) = \lim_{z\to 1} \frac{D'(z)N''(z) - N'(z)D''(z)}{2\left(D'(z)\right)^2} \\ &= \frac{D'(1)N''(1) - N'(1)D''(1)}{2\left(D'(1)\right)^2} \end{aligned} \tag{1.26}$$

Using Equation (1.26) we get

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

**d) Mean number in the system**

Applying $L = \rho + L_q$, we obtain $L = \rho + \frac{\lambda^2 E(S^2)}{2(1-\rho)}$

**e) Average Waiting time in the queue:** Using the relation $W_q = \frac{L_q}{\lambda}$**,** we obtain

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

**f) Average Waiting time in the system**

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

Also the probability generating function of the number of customers in the system at random epoch $P(z)$ can be obtained using equation (1.25) and the equation given by Kashyap and Chaudhury (1988), $P(z) = zP_q(z) + Q$

The M/G/1 queueing system has been extensively studied by many authors due to its wide applicability. Various aspects of M/G/1 queueing system has been studied by Levy and Yechiali (1975), Scholl and Kleinrock (1983), Madan (1994), Li and Zhu (1996), Choudhury (2005, 2006) and Taha (2007) among many others.

## 1.12 $M^X/G/1$ Queueing System

The arrivals in a queueing system may be in groups or batches. The size of the group may be regarded as a random variable given by a probability distribution. The queueing system described in the previous section becomes a special case of the model discussed here with a batch size being equal to 1.

The $M^X/G/1$ queueing system represents a single server queueing system where the units arrive in groups according to a compound Poisson Process. The service times of the individual customers are considered to be generally distributed. The $M^X/G/1$ queueing system considers similar assumptions underlying M/G/1 queues. Further, let $\lambda c_i dt$ be the first order probability that a batch of size i arrives in a short interval of time $(t, t+dt]$ where $0 \le c_i \le 1$ and $\sum_{i=1}^{\infty} c_i = 1$ and $\lambda > 0$ is the mean rate of arrival in batches. Accordingly, the following equations govern the system at steady state

$$\frac{d}{dx}P_n(x)+(\lambda+\mu(x))P_n(x)=\lambda\sum_{i=1}^{n-1}c_iP_{n-i}(x) \qquad n\geq 1 \tag{1.27}$$

$$\frac{d}{dx}P_0(x)+(\lambda+\mu(x))P_0(x)=0 \tag{1.28}$$

$$\lambda Q=\int_0^\infty P_0(x)\mu(x)dx \tag{1.29}$$

$$P_n(0)=\int_0^\infty P_{n+1}(x)\mu(x)dx+\lambda c_{n+1}Q, \qquad n\geq 0 \tag{1.30}$$

The probability generating function for the batch size is

$$C(z)=\sum_{i=1}^\infty z^i c_i \tag{1.31}$$

From (1.27)-(1.31) we have

$$\frac{d}{dx}P_q(x,z)+(\lambda-\lambda C(z))P_q(x,z)=0 \tag{1.32}$$

$$zP_q(0,z)=\int_0^\infty P(x,z)\mu(x)dx+\lambda(C(z)-1)Q \tag{1.33}$$

Solving these equations, we get the probability generating for the number of customers in the queue at random epoch

$$P_q(z)=\frac{\left[1-G^*(\lambda-\lambda C(z))\right](1-\lambda E(I)E(S))}{G^*(\lambda-\lambda C(z))-z} \tag{1.34}$$

where $G^*(\lambda-\lambda C(z))=\int_0^\infty e^{-(\lambda-\lambda C(z))x}dG(x)$ is the Laplace Stieltjes transform of the service time and $E(I)$ is the mean of batch size of arriving customers. Similar to M/G/1 queues we can also find the probability generating function of the number of customers in the system. The various performance measures can be obtained using equation (1.34) and Little's Law, knowing that $E(I(I-1)$ is the second factorial moment for the batch size of arriving customers, as below:

**a) Probability of finding the system empty on arrival**

$Q=1-\rho=1-\lambda E(I)E(S)$

**b) Server Utilization**

The server is busy whenever there is at least one customer in the system, i.e.

$$\rho = \lambda E(I)E(S)$$

**c) Mean number in queue**

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

**d) Mean number in the system**

$$L = \rho + \frac{\lambda E(S)E(I(I-1) + (\lambda E(I))^2 E(S^2)}{2(1-\rho)}$$

**e) Mean Waiting time in the queue**

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

**f) Mean Waiting time in the system**

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

This model is studied extensively in various forms by many authors like Baba (1986), Lee and Srinivasan (1989), Choudhury (2000), Ke (2001, 2007b), Madan & Al-Rawwash (2005), among several authors. The current research aims at generalizing and extending the results of the above model by addition of the assumptions like balking, reneging, re-service, server vacations, fluctuating modes of providing service and random breakdowns.

## 1.13 Research Problem

In a classical queueing system it is assumed that servers are always available, this is practically unrealistic. However, in reality, the servers may become unavailable for some time due to a number of reasons. Service may be disrupted due to interruptions like vacations and breakdowns in the system. Almost all practical systems in the real world that can be modeled as queues with vacation, as a server after rendering service for some time may opt to take vacations. Some of them are call centers with multi-task employees, computer and telecommunication sectors, manufacturing industries, etc.

Another reason for the server being unavailable or interrupted is due to the unpredictable breakdown of the system. Such kind of server interruptions result in the unavailability of the server for a period of time until repaired. A queueing system may face breakdowns in a mechanical and electronic service station at any instant while providing service to the customers due to many reasons. Hence the server will not be able to continue providing the service until the system is repaired. In some situations, the repair process may take more than one stage of repair. This is mainly seen in systems where the server is mechanical or electronic. All these features led to the motivation to study through generalizing some of the classical models studied earlier. One of the possible queue behaviors in a queueing system is when represented by arrivals in batches rather than one by one. Some examples of arrivals in batches to a system are customers in elevators, supermarkets, banks, restaurants, etc. Extensive studies on vacation models with batch arrivals were conducted by many researchers. Some of the prominent works were done by authors like Baba (1986), Choudhury and Borthakur (2000), Hur and Ahn (2005), Sikdar and Gupta (2008), Ke and Chang (2009b), who studied queues with batch arrivals under different vacation policies.

One common feature that is observed is that customers receiving service sometimes may need to repeat the service for various reasons. For example, re-service in observed in health and medical clinics as after a consultation a patient may need to re-consult the doctor immediately for any queries. Rework in industrial operations is also an example of a queue with feedback or re-service. Transmission of protocol data units is sometimes repeated in computer communication systems due to an error. Most of the existing literature on repeated service deals with the customer joining the tail of the queue or forming another queue for re-services. However, the current research considers that the customers have the option of taking a re-service immediately as soon as a service is completed, which makes it different from previous studies.

Also, customers tend to leave the queue upon the absence of the server for some time. Not only this kind of abandonment is seen with human queueing situations but also in a communication system like hotlines where at times the caller is kept waiting as the service provider is busy in other work. In such cases, the caller may hang up without the service. The traditional queueing theory considers that service is provided in a single mode, usually ‘normal. It is observed that in the real world, the service offered by the server can oscillate from one mode to another as for instance, slow, normal or fast. Queueing systems whose service rates fluctuate over time are very common but are still not well understood analytically. Very few studies on fluctuating service rates are seen in the queueing literature.

Most of the studies in the queuing literature are concentrated with single assumptions like queues with vacations or queues with breakdowns etc. Further arrivals are commonly considered as single instead of batches. Moreover, they have assumed that there is no waiting space in front of the server so that if an arriving customer finds an idle server, he is immediately offered service otherwise leaves the system or joins a retrial queue. This differs from the assumption in the current research of infinite waiting space.

Reneging during vacations in a queueing system has been a new endeavor attempt since the last few years, but most of the works on reneging deals with Markovian service time and single arrival. The current study treats reneging on the non-Markovian nature of service time making it different from the studies mentioned in the literature. Moreover, reneging is considered to occur only during a vacation where vacations are based on single vacation policy and Bernoulli schedule, that is after the completion of a service, the server may take a vacation with probability p or remain in the system to continue service, if any customer, with probability $(1-p),\ 0 \le p \le 1$.

A detailed review of all the related literature is discussed in Chapter Two.

The current study, therefore, is motivated by its various real life applications. Factors like server vacations and breakdowns may contribute to affect the system's efficiency adversely. In the literature, we do not find queueing models that combine such assumptions like server vacations, breakdowns, balking or re-service in a batch arrival queueing system, which is considered in the current research.

Thus in this research, the author proposes to extend a batch arrival queueing system $(M^X/G/1)$ by considering assumptions which complement many real life situations. This generalizes the traditional queueing system in various directions with balking and re-service in two kinds or two stages of heterogeneous services, reneging during Bernoulli schedule server vacations, time homogenous breakdowns in a single server with fluctuating modes of service. All these assumptions are studied when arrivals are assumed to be in batches with general service time.

## 1.14 Research Objectives

The literature on queueing theory still lacks studies in depth on models with balking and reneging during server vacations, random breakdowns and fluctuating modes of service in a batch arrival queueing system with variants of service facilities. Significant studies have been carried out on queues with breakdowns and customer's impatience by different authors but are considered in different queueing set up than those considered in the current study. In the current research, the author has considered non-Markovian queueing systems with batch arrivals where the service times follow General (arbitrary) distribution and extended this model by considering service provided in two fluctuating modes, re-service, customer's impatience, server breakdowns and Bernoulli schedule vacation policy. This aims to generalize and analyze the basic $M^X/G/1$ by the combination of these assumptions and thus extending in many directions. As a consequence, this combination models new systems, which are close to representing a real queueing situation.

The most important problem connected with the study of queueing systems include the determination of probability distribution for the system length, waiting time, busy period and an idle period. Therefore this research is conducted with the following objectives:

1. To determine the steady state behavior of batch arrival queueing systems with two types of heterogeneous service together with balking. The customer has the option of choosing any one of the two heterogeneous services.
2. To determine the steady state behavior of a two-stage batch arrival queueing system with balking and optional re-services. A service is completed in two stages, one after the other in succession.
3. To determine the steady state behavior of a batch arrival queueing system with reneging during server vacations for Markovian and non-Markovian set up.
4. To determine the time dependent and steady state behavior of a batch arrival single server queue with breakdowns and server providing general service in two fluctuating modes.

A queueing capacity or a process can be most specifically represented by its performance measures which are quantifiable and can be documented. Thus the main objective of any queueing system is to assess the performance measures. Therefore among the other objectives mentioned above, the author also attempts to determine the some of the performance measures of interest, like the probability of an idle time, traffic intensity, mean queue size etc. of the different models under study.

## 1.15 Research Methodology

The research objectives discussed in the previous section could be achieved by deriving the probability generating function for the queue size distribution at random point of time. Thus the steady state queue probability generating functions of queue size has been obtained to achieve the research objectives from 1 to 3, discussed in Section 1.14. Both time dependent and steady state probability generating functions of the queue size have been obtained to achieve the fourth objective.

Since the queueing systems considered in this study has different states the following has been obtained:

1. Probability generating function of the number of customers in the queue when the server is providing service (or re-service).
2. Probability generating function of the number of customers in the queue when the server is on vacation.
3. Probability generating function of queue size when reneging occurs during server vacations.
4. Probability generating function of the number of customers in the queue when the server is under repair.
5. Probability generating function of the number of customers in the queue irrespective of the different states of the system.

To understand the behavior of the queueing systems under study, the author has derived some performance measures such as mean queue length, mean waiting time in the queue, proportion of time the server is idle and server utilization factor. Among different methods used in analyzing a queueing system, the supplementary variable techniques have been used to find the necessary probability generating functions.

## 1.16 Thesis Outline

The overall and organizational structure of the thesis is outlined in this section. The thesis consists of six chapters and content of each chapter is briefed as below:

- Chapter 1: Introduction explores the concept of queueing theory with a description of the queueing system, description of the theories related to the study, the objective of the research and proposed methodology of the study.
- Chapter 2: Literature Review discusses the literature related to the antecedents of the interruptions like breakdowns, server vacations, customer's impatience behavior etc. in a queueing system from many different perspectives and current work.
- Chapter 3: The author considers the basic $M^X/G/1$queueing model with Bernoulli schedule server vacation and customers or units arrive in batches of variable size. Three models have been discussed. In the first model, it is assumed that the server provides two parallel services and customer has the choice of selecting any one of the two heterogeneous services. Customers may join the queue with probability $b_1$ or balk with probability $(1-b_1)$ when the system is busy. Similarly, during vacation, customers may join the system with probability $b_2$ or balk with probability $(1-b_2)$. In the second model, we analyze a system where the server is providing service in two stages, one by one in succession. Customers may decide to balk both during working state or the vacation state of the server. Further, the model is generalized by incorporating optional re-service to develop a new model discussed as model three. As soon as service (of any one type) is complete, the customers have the choice of leaving the system or join the system for re-service. A special case with Erlang-k vacation time is also studied.
- Chapter 4: Here, in this chapter, a batch arrival queuing system with Bernoulli schedule server vacations is considered. The study is based on two aspects, Markovian and non-Markovian service time, classified as Model 1 and Model 2. Customers have the decision to renege independently and reneging is considered to follow an Exponential distribution. The study is extended by considering batch arrivals to follow Geometric distribution while service time follows Gamma distribution, symbolically denoted as $Geo(x)/\gamma/1$ where reneging takes place during vacation periods.

- Chapter 5: Here the author studies the basic model $M^X/G/1$ with server vacations and breakdowns. The new assumption here is that the server is providing service in two fluctuating modes. We assume $\pi_1$ as the probability that the server is providing service in mode 1 and $\pi_2$ as service in mode 2. We also consider that the server can experience breakdowns during the time when the server is busy in working state as well as during idle state, i.e. breakdowns are assumed to be time homogeneous. Once the repair process is complete the server immediately starts providing service in mode j (j=1,2). Both the time dependent and steady state solutions of the model are investigated.
- Chapter 6: provides the summary and implications of the study. This chapter combines the overall findings and contributions of the study. In this chapter we also present and outline the future scope for research in this area.

## 1.17 Summary

For all the models investigated in our thesis, we assume that the service times and vacation times follow the general (arbitrary) distribution except the model discussed in Chapter Four where we consider vacation time to follow a Markovian distribution. Breakdowns are assumed to follow Exponential distribution, as discussed in Chapter Five. The customers arrive in batches according to a Poisson process and are served one by one in FCFS discipline. The stochastic processes involved in the system are assumed to be independent of each other. We derive the steady state solutions of all our models using the supplementary variable technique. Further a time dependent case is also investigated for one of the models considered in the thesis. In the following chapter, we provide a detailed review of the literature on queues related to the current study.

******

# Chapter 2
# Literature Review and Current Work

## 2.1 Introduction

In this chapter, the current researcher provides a historical background of the studies on theory of queues along with a review of the literature. For any research it is imperative to have an understanding of the theories and studies developed in the past. Thus the author provides a brief description of the background of the study of queues in the past. The review of the extant literature is to explore the theoretical foundation supporting all the concepts of interruptions or disruptions in a queueing system, which is described in the literature review section. Finally, a synthesis of the various queueing systems to be studied in the current research is presented.

## 2.2 A Historical Perspective

Queueing theory had its origin in 1909 when A.K.Erlang, a Danish Engineer, published his fundamental paper relating to the congestion in telephone traffic titled 'The Theory of Probabilities and Telephone Conversations' (1909). Erlang was also responsible for the notion of statistical equilibrium, or the introduction of the so-called balance-of-state equations, and for the first consideration of the optimization of a queueing system (Gross and Harris, 1985).

The use of the term 'Queueing System' first occurred in 1951 in the journal of the *Royal Statistical Society* in the article titled 'Some Problems in Queueing Theory' by D.G. Kendall. Most of the articles used the word 'Queue' instead of 'Queueing'. He also designed the queueing notation *A/B/C* in 1953 which was later universally accepted and used. 'In this period of intensive activity, new methods were introduced; solutions to problems of interest obtained and many open problems were suggested' (Gupur, 2011). Most of the basic queueing models studied during the end of 1960's had been considered to model the real world phenomenon.

The earliest model to be analyzed was the M/M/1 queue; the first solution was the time dependent solution given by Bailey (1954) who used the generating function for the differential equations.Pollaczek (1965) did considerable work on the analytical behavior of the queueing systems. Goyal and Harris (1972) was the first to work on estimating parameters on a non-Markovian system and used the transition probabilities of the imbedded Markov chain to set up the likelihood function. Since then significant research papers and books have been contributed by authors like Cox, (1965), Leonard (1976), Kelly (1979), Mitra & Mitrani (1991), to mention a few.

With the advancement of computer technology, applied probability, methodology and application contributed immensely to the growth of the subject. Studies on queueing networks was first introduced by Jackson (1957), later Whittle (1967, 1968), Kingman (1969) treated it with respect to population processes. A vigorous growth in the application of queuing theory in computer and communication systems was founded by eminent researcher Kleinrock (1975, 1976). Later authors including Coffman and Hofri (1986), Mitra et al (1991), Doshi and Yao (1995) also published papers on queues related to communication systems.

For more research on queueing models with server vacations, the author refers to Doshi (1986), Takagi (1990) and Alfa (2003). The introduction of the matrix analytic method which developed the scope of queuing systems was by Neuts (1984).Choudhury and Madan (2005) investigated a system with modified Bernoulli vacation and N policy. Gupta and Sikdar (2006) studied a Markov Arrival Process (MAP) with single or multiple vacation policies. Banik (2010) used the embedded Markov chains to obtain the performance measure of a queueing system with single working vacation. Since then queueing theory proved to be a fertile field for researchers who worked on the study of queues with the different phenomenon and derived useful results. Many popular and important books on queuing theory and its application have been published during the last 30 years by authors like Kashyap and Chaudhury (1988), Nelson (1995), Bunday (1996), Gross and Harris (1998), Medhi (2003) and Mark (2010), among others.

## 2.3 Literature Review

Queues with vacations have been extensively studied in the past three decades. It has emerged as an important area in queueing theory and we see sufficient literature in this context. These queueing systems have got wide applicability in telecommunication engineering, computer networks, production systems and other stochastic systems. The vacations may represent server's working on some supplementary job, performing maintenance, inspection etc.

An M/G/1 queueing system where the server is unavailable for a random length of time, referred as vacation, was first studied by Miller (1964). Since then; queues with vacations attracted the attention of queueing researchers and became an active research area. In the next two decades, several mathematicians developed general models which could be used in more complex situations. Some of the early work on queues has been relevant to queues with vacations due to its wide practical applications. The fundamental result of vacation models is the stochastic decomposition theorem which was discovered by Levy and Yechiali (1975). According to this theorem, the stationary queue length or stationary waiting time can be distributed as the sum of two independent components- one of them is the queue length or waiting time of the corresponding classical queueing system without vacation as and the other is the additional queue length or delay due to vacations. Fuhrmann and Cooper (1984) extended this result further by using generalized vacations. Scholl and Kleinrock (1983) studied an M/G/1 queues with rest periods and FCFS order of service. Later, stochastic decomposition theorems on different vacation policies were studied by Keilson and Servi (1987), Takine and Hasegawa (1992), Madan (2000a), Alfa (2003), to name a few.

The GI/M/I queue with exponentially distributed vacations was studied by Tian et al. (1989) who established the stochastic decomposition properties for stationary queue length and waiting time. Takagi (1992b) studied the time dependent process of M/G/1 vacation models with exhaustive service. Choudhury (2002) studied a queuing system with two different vacation times under multiple vacation policy. Madan et. al. (2003) analyzed an M/M/2 queue with single Bernoulli schedule vacation policy and discussed two models under different conditions. A few of the notable works on M/G/1 queuing system with vacations were done by authors like Ke (2003b), Artalejo and Choudhury (2004), Choudhury (2008), Li et. al. (2009).A GI/M/C queueing system with vacations in which all servers take vacations together when the system becomes empty was studied by Tian and Zhang (2003). These servers keep taking synchronous vacations until they find customers waiting at a vacation completion instant. Using embedded Markov chain modeling and the matrix geometric solution methods, obtained explicit expressions for the stationary probability distributions and the waiting time.

Among other vacation policies, Kumar and Madheswari (2005), Banik et al. (2007) and Choudhury et al. (2007) studied queues with multiple vacation policies. A vacation queueing system where a server is turned on when there are N or more customers present and of only when the system is empty, is called an N-policy. After the server is turned off it will not operate until at least N customers are present in the system. Arumuganathan and Jeykumar (2005) considered bulk queues with multiple vacations, setup times with N-policy and closedown times. The readers are referred to the comprehensive studies on queues with vacation by Doshi (1986), Takagi (1990), Medhi (1997) and Tian and Zhang (2002).

To the best of the researcher's knowledge, most of the studies on vacation queues have been concerned with M/G/1 or bulk arrival $M^X/G/1$ queues. When arrivals are considered to be in batches, it becomes a more realistic assumption of a queueing system, instead of individual arrivals. Such models are often seen in the transportation system, hospital, logistics and shipping industry, communication network systems etc. Borthakur and Choudhury (1997) analyzed a batch arrival Poisson queue with generalized vacations. It states that the server goes on a vacation of random length as soon as the system becomes empty. On return from vacation, if he finds customer(s) waiting in the queue, it starts serving the customer one by one till the queue becomes empty; otherwise, he takes another vacation and so on. The study of $M^X/G/1$ queues with vacation has drawn the attention of many researchers; notable among them are Baba (1986, 1987), Lee et al. (1992), Choudhury and Borthakur (1995), etc. The stochastic decomposition results for a class of batch arrival Poisson queues with a grand vacation process at various points of time was derived analytically by Choudhury and Borthakur (2000).

The study of server vacations in a queueing system has been extended in many directions. Ke (2001) studied the optimal control of a $M^X/G/1$ queue with two types of generally distributed random vacation, type 1(long) and type 2 (short) vacations. Madan et al. (2004) studied a single server queue with optional phase type server vacations based on exhaustive service.

Xu et al. (2007) investigated a $M^{[X]}/G/1$ queue with multiple vacations and obtained the probability generating functions of queue length, busy period, and the stationary waiting time under the first come, first served and last come, first served service disciplines. An M/G/1 queue with multiple working vacations and exhaustive service discipline was analyzed by Wu and Takagi (2006). In recent years, significant attention has been received in studying queues with working vacations. A GI/Geo/1 queue with working vacations and vacation interruption, where the server takes the original work at a lower rate than completely stopping during the vacation period was studied by Li and Tian (2007). In the vacation interruption policy, the server can come back to the normal working level once there are customers after service completion, during the vacation period and in that case the server may not accomplish a complete vacation. They obtained the steady state distributions for the number of customers in the system at arrival epochs and waiting times, using the matrix-geometric solution method. Later, Tian et al. (2007) studied a discrete time Geom/Geom/1 queue multiple working vacations. They obtained the distribution of the number of customers in the system by using quasi-birth-death chain and matrix-geometric solution method. More recently, Khalaf et al. (2011) studied single server queueing systems with extended vacations and breakdowns. They considered that after the completion of a vacation of random length, the server has the option to go for an extended vacation or remain in the system to continue providing service. In the current study, we have considered Bernoulli schedule single vacation policy in an $M^X/G/1$ queueing system.

In real life situations, customers are often discouraged by long queues. Balking and reneging are some common forms of customer impatience occurring in a queuing system. The phenomenon of customer joining a queue and leaving without completing its service is termed as reneging. Barrer (1957) appears to be the first who studied queues with impatient customers. He generalized the standard M/M/1 queue to the case of impatient customers. Haight (1957, 1960) investigated the equilibrium characteristics of a queueing system with balking where each arrival will join the queue if the length is less than or equal to a random quantity. Another early work on reneging was carried out by Ancher and Gafarian (1963) where they considered Markovian arrival and Markovian service pattern. Since then, many significant works has been done with queues based on balking and reneging. Haghighi et al. (1986) studied both balking and reneging on multi-server queues. They derived the long run probability distribution of the number of customers in the queue and also obtained the expression for the average loss of customers in the queue. Since then a significant amount of work on the impatience phenomena has been treated under various assumptions. Balking and control retrial rate was analyzed on QBD process for a multi-server queue by Kumar and Raja (2000). They discussed the necessary and sufficient condition for stability in their work. Later, an M/Ph/1 queue with wait based balking has been analyzed by Liu and Kulkarni (2006). The differential equations for phase type service time distributions have been derived and solved it explicitly with Erlang, Hyper-exponential and Exponential distributions as special cases. El-Paoumy (2008) studied a $M^X/M/2/N$ queue with balking, reneging and heterogeneous servers. Shawky and El-Paoumy studied a $H_k/M^{a,b}/c/N$ with balking, reneging and general bulk service rule. More recently, Selvaraju and Goswami (2013) studied M/M/1 queueing systems with the reneging phenomenon in single and multiple working vacations. Even though a significant amount of studies has been carried out with balking and reneging, there are various aspects of a queuing system which still needs to be addressed.

Most of the literature in balking and reneging mentioned above deals with Markovian queues and single arrivals. As such, the current researcher attempts to contribute in this area by considering a batch arrival non-Markovian service mechanism with balking and reneging taking place only during the absence of the server, that is, in vacation state. However, balking is also assumed to take place both when the server is busy. This assumption of non-Markovian service time with arrivals in batches makes it different from the earlier studies in the theory. The study of queueing models along with balking and reneging provide a close representation of a real situation.

Sometimes the service discipline involves more than one service. In many practical situations, there are customers returning to the waiting line after receiving a service, under some decision policy to receive another service. Feedback queues are in which once a customer service becomes unsuccessful is served again and again till his service becomes successful. A queue with feedback policy was initially studied by Finch (1959), Takacs (1963), Kleinrock (1975), multiple types of feedback by Boxma and Yechiali (1997). Madan (2000a) was the first to study a $M/G/1$ queue with optional second service. In recent years, authors like Madan et al. (2004)

investigated a $M^X/G/1$ queue with optional re-service, Tadj and Ke (2008), Jeyakumar and Arumuganathan (2011) has studied queues with optional re-service. Although many research works has been carried out with batch arrival queuing models, very few studies have been carried out with re-service.

Server breakdowns or service interruption due to the failure or breakdown of the server is a very common phenomena experienced in a queuing system. For instance, in a machine processing center machine breakdowns may occur due to power failure, lack of adequate preventive maintenance or use of some inferior materials. In recent years we have seen an increasing interest in queuing systems with server breakdowns. In the past, studies in queues with breakdowns have been studied by authors like Gaver (1962) and Avi-itzhak and Naor (1963), and Mitrany and Avi-Itzhak (1968). Since then studies on queues with server breakdowns have drawn the attention of many researchers. Sengupta (1990) studied a queue with service interruptions in an alternating random environment. Jayawardene and Kella (1996) examined an M/G/1 queue with altering renewal breakdown, followed by Li et. al. (1997) who investigated the reliability analysis of M/G/1 queuing system with server breakdowns. Wang (2004) studied an M/G/1 system with a second optional service and unreliable server. He obtained the transient and steady state solutions for both queueing and reliability measures using the supplementary variable technique. Later, in recent years queues with service interruptions or breakdowns have been studied by authors like Ke (2007), Choudhury, Tadj and Paul (2007), Wang et al.(2008), Maraghi, Madan and Darby-Dowman (2009), and much more.

Although numerous studies have been done with service interruptions there are some aspects which can be still addressed further. Most of the studies on queues with breakdowns consider system breakdown during the period when the server is offering service. However, a server can fail even when it is an idle state. This is an important feature which still needs to be studied in depth. Madan (2003) studied time homogeneous system breakdowns and deterministic repair times.

In the queueing literature, a majority of the models studied assume the server to be homogeneous, where the individual service rate is same for all servers. This assumption is valid only when the service process is highly mechanically or electronically controlled. However, in real life, the server providing service to the customer does not always offer service in the same mode or rate. The service provided at the service facility may be fast, normal or slow. This not only affects the system efficiency but also the queue length and the customer's waiting time. Madan (1989) studied the time-dependent behaviour of a single channel queue consisting of two components each subject to breakdowns with fluctuating efficiency. The system works at full efficiency or reduced efficiency or does not work depending on whether either components or neither of the components is in working state.

There appear to be very few research studies on queues with fluctuating modes of service in the queueing literature, even though many research works have been carried out on bulk queueing models. Such situations are encountered in industries where the service rendered does not have a constant rate. Arumuganathan and Ramaswami (2005) investigated a $M^{[X]}/G(a,b)/1$ queue with fast and slow service rates and multiple vacations. The server does the service with a faster rate or slower rate based on queue length. In the current research, the author intends to investigate batch arrival queueing models with balking and reneging, re-service, Bernoulli schedule server vacations, random breakdowns and fluctuating modes of service. These assumptions are considered in a queueing system with a single server with different service mechanism, like two kinds of parallel services, two-stages of services and two modes of services.

## 2.4 Synthesis of the Current Work

We now use the assumptions (discussed in Chapter One) to the queueing models under study by applying the supplementary variable technique and determine the performance measures of the models. In this research, the classical batch arrival queuing model is extended by using the basic assumptions of balking and reneging, breakdowns and server vacation under different service mechanism, as these factors influence the efficiency of a queueing system in real queueing situations. The interest in studying such types of models was motivated because of their theoretical structures as well as its various practical applications in the real world.

In the queuing system studied in Chapter Three, the author has introduced balking to a single server vacation queue having two types of parallel services and two stages of service classified as two different models. The arrivals are assumed to be in batches following a compound Poisson process. A new model is further developed by generalizing the model with two stages of services considering optional re-service. The customers have the option for re-service in any one of the stages. This system can be symbolically presented as $M^{[X]}/(G_1G_2)/1/Vs$ with balking and optional re-service. In all of the models, the service time, vacation time has been considered to follow a general (arbitrary) distribution.

In Chapter Four, we extend the classical batch arrival queuing model studied by considering reneging during server vacations. Bernoulli schedule server vacations are considered and reneging during such type of vacations has not been studied earlier. Both Markovian and Non-Markovian service time distribution are treated as two separate models. A particular case of this model is treated as a separate model where batches follow Geometric pattern of arrival and service time assumed as Gamma distribution.

In Chapter five, we consider breakdown in a single server batch arrival queue where the server is providing service in two fluctuating modes and both the transient and steady state solutions of the queue size distribution has been obtained.

Although numerous research works can be seen carried out with regard to queues on vacations, no work has been found with a combination of these features in the queuing systems considered in the study. All models investigated in the research deal with a non-Markovian set-up. The new results in each chapter formed new models which can contribute to the literature of queueing theory. Therefore this study contribute by providing an analytical extension of the classical batch arrival queueing model with server vacations coupled with balking, reneging, optional re-service, random breakdown, along with fluctuating modes of providing service.

## 2.5 Summary

In this chapter, the author has provided a historical background of the study of queueing theory along with a detailed review of the literature on queues, especially on non-Makovian queues related to the different assumptions of the current research. Further, we give a description of the synthesis of the current work. In the following chapter, the analysis of the queueing system with balking and optional re-service is provided.

******

# Chapter 3

# Analysis of a Non-Markovian Batch Arrival Queue with Balking and Bernoulli Schedule Server Vacations

## 3.1 Introduction

In this chapter, we analyze a single server batch arrival queueing system with balking, and Bernoulli schedule server vacation. The single server queueing system including the $M^X/G/1$ queues with batch arrivals and vacations have been studied by many authors like Baba (1986), Rosenberg and Yechiali (1993), Choudhury (2002), Choudhury and Paul (2004), among several others. More recently, most of the studies have been devoted to various vacation policies including Bernoulli schedule. All these papers have a common assumption that the system has a single server providing only one kind of service to the arriving customers. However, in real life there are some service stations that provide more than one kind of service. Just before service starts, the customer has the option to choose one among the different kinds of service provided. We can find many real life applications of this model like automobile fuel stations, banks, post offices, etc. In the past, Beja and Teller (1975) investigated such types of queueing systems. They assumed K possible types of service in which an arriving customer has the option to choose one. The expected cost per serving customer in each kind of service was compared to the expense of losing the customer. Anabosi and Madan (2003) studied a single server queue with two types of service, optional server vacations based on Bernoulli schedule and single vacation policy, where vacation time followed the Exponential distribution. This work was later generalized by Madan et al. (2005), where they have assumed general arbitrary distributions for both service and vacation times.

A queueing system can also be described when a customer's service may be viewed as scheduled in two stages. A number of papers appeared recently in the queueing literature in which two-stage queueing systems with server vacations have been studied. Choi and Kim (2003) introduced Bernoulli feedback in a two-phase queueing system with vacations. The steady state solutions of a single server queue with two-stages of heterogeneous service and Binomial schedule server vacations having exponential vacation time have been studied by Madan (2000b). He assumed that in both stages of service, customers are served individually. Choudhury & Madan (2005) studied a two-stage batch arrival queueing system with a modified Bernoulli schedule vacation with random set up time under N-policy and later with restricted admissibility policy (Madan and Choudhury, 2006). An M/G/1 queueing system with two phases of service with breakdowns has been studied by Choudhury and Tadj (2009). A batch arrival retrial queue with two stages of service and Bernoulli schedule server vacations was investigated by Kumar and Arumuganathan (2008). Most of the literature mentioned here has studied queues with Markovian service time and with individual arrivals.

The motivation to study these queueing systems with two stages and vacations comes from a wide application of these models in the real world. In computer networks and telecommunication systems, messages are processed in two stages by a single server. An example of two stage queues is an automatic car service station where the car receives periodic checking and washing, the machines may be stopped in between once a while due to overhauling after these two stages of service. This overhauling may be considered as vacation time. Also in some computer networks and telecommunication systems messages are processed in two stages by a single server (Madan & Choudhury, 2005).

Although we see significant studies on two stages of services as well as an option of services with server vacations in the queueing literature, there are some aspects which have still not been addressed. Customers at times receiving a service may demand re-service if unsatisfied by the service or for any other reason. Similarly balking, which is a form of customer's impatience, is the reluctance of joining the queue upon arrival. The factors re-service and balking is very commonly seen occurring in any real queueing situation. In queueing literature one can find most of the studies done on two kinds of heterogeneous services or two stages of services consider that all customers arriving receive the service. However, in reality, this may not always be true.

Customers may tend to leave the system due to impatience without receiving any service, resulting in potential loss of customers to the system. Balking, when the arriving unit ceases to join the queue, is an integral customer activity in the process of receiving a service.

Haghighi et al. (1986) studied the multi-server Markovian queueing system both with balking and reneging concepts. The multi channel M/M/c/N queue with balking and reneging has been discussed by Abou-el-Ata and Hariri (1992). Later, Zhang et al. (2005) analyzed an M/M/1/N queue with balking, reneging and server vacations. An $M^X$/M/2/N with balking, reneging and heterogeneous servers was analytically studied by El-Paoumy (2008). El-Sherbiny (2008) studied Markovian queues with balking concepts with an additional server for longer queues. Transient state multi server queues with impatient customers have been studied by Al-Seed et al. (2009), Ammar et al. (2012), to name a few.

In recent years, we see very few studies on queues with repetition of services or re-service. Al-Khedairi and Tadj (2007) studied a bulk service queue with a choice of service and re-service under Bernoulli schedule. Later, Tadj and Ke (2008) studied a hysteretic bulk quorum queue with a choice of service and optional re-service. More recently, a non-Markovian queue with multiple vacations and controlled optional re-service have been analyzed by Jeykumar and Arumuganathan (2011). Although there have been significant studies on queues with feedback, it makes it different from re-service in the context that for feedback queues the customer requiring re-service joins the tail of the queue and waits for its turn for repeating the service.

Although significant studies have been done by numerous researchers in bulk queueing systems with variants of vacation policies, there are still some aspects which can be treated further.

To the best of our knowledge, no studies in the literature have combined the two assumptions balking and re-service in a $M^X/G/1$ queueing system with server vacations with variants of service facility, which led to the motivation behind this study. The purpose of this chapter is thus to generalize the model by incorporating balking and then investigate the behavior of the system with re-service.

The current study is different from those available in the literature as it combines both balking and server vacations in a batch arrival queue where the single server is classified into two types of heterogeneous services and two stages of heterogeneous services considering service time to be generally distributed. We assume that arrival units may refuse to join the system (balks) when the server is busy providing service and when the server is on vacation with different probabilities. An arriving unit may decide not to join the queue either may be by estimating the duration of waiting time for a service to get completed or by witnessing the long length of the queue. Further, the $M^X/G/1$ queue with two stages of service has been extended with optional re-service.

Thus in this chapter we analyze the queueing system in terms of three separate models; the first model denoted as $M^X/\begin{pmatrix} G_1 \\ G_2 \end{pmatrix}/1/Vs$, a single server with two types of heterogeneous services. A customer has the option of selecting any one of the two types of service; the second model denoted as $M^X/G_1G_2/1/Vs$, with two stages of service, one after another in succession. A service is complete only if the services of both the stages are complete and the third model; $M^X/G_1G_2/1$ with balking and optional re-service.

We first develop the steady-state differential equations for the queueing system by treating elapsed service time and elapsed vacation time as supplementary variables. We solve these system equations and derive explicitly the probability generating functions of various server states at a random epoch. A special case with Erlang-k vacation time is also discussed.

The rest of the chapter is organized as follows: Section 3.2 studies the steady state behavior of Model 1 along with the mean of queue size. In Section 3.3, we derive the steady state solution of the probability generating function Model 2. In Section 3.4, a two stage batch arrival queue with balking and optional re-service is discussed. The mean of queue size has been obtained for each of the models. Section 3.5 discusses a special case with Erlang-k vacation time. The summary of the chapter is given in Section 3.6.

## 3.2 Model 1: Analysis of a batch arrival queue with balking and two types of heterogeneous services

### 3.2.1 Assumptions of the Mathematical Model

The following are the assumptions describing the model:

a) Let $\lambda c_i dt$ be the first order probability that a batch of size $i$ arrives in a short interval of time $(t, t+dt]$ where $0 \le c_i \le 1$ and $\sum_{i=1}^{\infty} c_i = 1$ and λ>0 is the mean rate of arrival in batches. There is a single server providing two types of parallel service to customers, one by one on a first come first served basis (FCFS). Before a service starts, a customer can choose type 1 service with probability $\xi_1$ or choose type 2 service with probability $\xi_2$, where $\xi_1 + \xi_2 = 1$. We assume that the service time random variable $S_j$, $j = 1, 2$ of type $j$ service follows a general probability law with distribution function $G_j(s)$, probability density function $g_j(s)$ and $k$ th moment $E(S_j^k)$ $(k = 1, 2, 3, ...)$. Let $\mu_j(x)$ be the conditional probability of completion of type j service during the period $(x, x+dx]$ given that the elapsed service time is $x$, so that

$$\mu_j(x) = \frac{g_j(x)}{1 - G_j(x)}; \; j = 1,2 \, , \; g_j(s) = \mu_j(s) \exp\left[ - \int_0^s \mu_j(x) dx \right]$$

Once the service of a customer is complete then the server may decide to take a vacation with probability $p$ or may continue to serve the next customer with probability $(1-p)$ or may remain idle in the system if there is no customer requiring service. Further we assume that the vacation time random variable follows general probability law with distribution function $W(v)$, the probability density function $w(v)$ and the $k^{\text{th}}$ moment $E(V^k)$, k=1,2….; Let $\phi(x)$ be the conditional probability of completion of a vacation period during the interval $(x, x+dx]$ given that the elapsed vacation time is $x$, so that

$$\phi(x)=\frac{w(x)}{1-W(x)}, \quad \text{and } w(v)=\phi(v)\exp\left[-\int_0^v \phi(x)dx\right]$$

b) Next, we assume that $(1-b_1)$, $(0 \le b_1 \le 1)$ is the probability that an arriving batch balks during the period when the server is busy (available on the system), so that $b_1$ is the probability of joining the system and $(1-b_2)$, $(0 \le b_2 \le 1)$ is the probability that an arriving batch balks during the period when server is on vacation, so that $b_2$ is the probability of joining the system during vacation period.

The various stochastic process involved in the system are assumed to be independent of each other. The probabilities $P_{n,j}(t,x), j=1,2$, for two kinds of services and $V_n(t,x)$ for vacation time and $Q(t)$ for idle time associated for this study are defined just before Chapter One.

Now, assuming that the steady state exists (the condition(s) under which it may exist will emerge later), we define the following steady state probabilities:

$$\begin{aligned}
&\lim_{t\to\infty} P_{n,j}(t,x)=P_{n,j}(x), \quad \lim_{t\to\infty} P_{n,j}(t)=\int_0^\infty P_{n,j}(t,x)dx=P_{n,j}, \\
&\lim_{t\to\infty} V_n(t,x)=V_n(x), \quad \lim_{t\to\infty} V_n(t)=\int_0^\infty V_n(t,x)dx=V_n, \quad \lim_{t\to\infty} Q(t)=Q
\end{aligned} \tag{3.1}$$

### 3.2.2 Steady State Equations Governing the System

The following set of differential-difference equations represents the queueing system in accordance with the assumptions of our mathematical model:

$$\frac{d}{dx}P_{n,1}(x)+(\lambda+\mu_1(x))P_{n,1}(x)=(1-b_1)\lambda P_{n,1}(x)+b_1\lambda\sum_{i=1}^{n}c_iP_{n-i,1}(x), \; n\ge 1 \tag{3.2}$$

$$\frac{d}{dx}P_{0,1}(x)+(\lambda+\mu_1(x))P_{0,1}(x)=(1-b_1)\lambda P_{0,1}(x) \tag{3.3}$$

$$\frac{d}{dx}P_{n,2}(x)+(\lambda+\mu_2(x))P_{n,2}(x)=(1-b_1)\lambda P_{n,2}(x)+b_1\lambda\sum_{i=1}^{n}c_iP_{n-i,2}(x), \; n\ge 1 \tag{3.4}$$

$$\frac{d}{dx}P_{0,2}(x)+(\lambda+\mu_2(x))P_{0,2}(x)=(1-b_1)\lambda P_{0,2}(x), \tag{3.5}$$

$$\frac{d}{dx}V_n(x) + \left(\lambda + \phi(x)\right)V_n(x) = \lambda(1-b_2)V_n(x) + b_2\lambda\sum_{i=1}^{n} c_i V_{n-i}(x), \ n \geq 1 \tag{3.6}$$

$$\frac{d}{dx}V_0(x) + \left(\lambda + \phi(x)\right)V_0(x) = \lambda(1-b_2)V_n(x) + b_2\lambda\sum_{i=1}^{n} c_i V_{n-i}(x), \tag{3.7}$$

$$Q = (1-\lambda)Q + \lambda(1-b_1)Q + (1-p)\left[\int_0^\infty P_{0,1}(x)\mu_1(x)dx + \int_0^\infty P_{0.2}(x)\mu_2(x)dx\right] + \int_0^\infty V_0(x)\phi(x)dx. \tag{3.8}$$

(The detailed reasoning for equation (3.2)-(3.8) given in Appendix A.)

The above equations are to be solved subject to the boundary conditions given below at $x = 0$:

$$\begin{aligned} P_{n,1}(0) = (1-p)\xi_1&\left[\int_0^\infty P_{n+1,1}(x)\mu_1(x)dx + \int_0^\infty P_{n+1,2}(x)\mu_2(x)dx\right] \\ &+ \xi_1\int_0^\infty V_{n+1}(x)\phi(x)dx + \lambda b_1\xi_1 c_{n+1}Q; \qquad n \geq 0 \end{aligned} \tag{3.9}$$

$$\begin{aligned} P_{n,2}(0) = (1-p)\xi_2&\left[\int_0^\infty P_{n+1,1}(x)\mu_1(x)dx + \int_0^\infty P_{n+1,2}(x)\mu_2(x)dx\right] \\ &+ \xi_2\int_0^\infty V_{n+1}(x)\phi(x)dx + \lambda b_1\xi_2 c_{n+1}Q; \qquad n \geq 0 \end{aligned} \tag{3.10}$$

$$V_n(0) = p\left[\int_0^\infty P_{n,1}(x)\mu_1(x)dx + \int_0^\infty P_{n,.2}(x)\mu_2(x)dx\right] \qquad n \geq 0, \tag{3.11}$$

and the normalizing condition

$$Q + \sum_{j=1}^{2}\sum_{n=0}^{\infty}\int_0^\infty P_{n,j}(x)dx + \sum_{n=0}^{\infty}\int_0^\infty V_n(x)dx = 1 \tag{3.12}$$

Explanation of the boundary conditions from (3.9) - (3.11) are given in Appendix A.

### 3.2.3 Queue Size Distribution at Random Epoch

We define the following Probability Generating Functions:

$$\left.\begin{aligned}
&P_j(x,z)=\sum_{n=0}^{\infty} z^n P_{n,j}(x),\ \ P_j\,(z)=\sum_{n=0}^{\infty} z^n P_{n,j},\ \ j=1,2 \\
&V(x,z)=\sum_{n=0}^{\infty} z^n V_n(x), \qquad V(z)=\sum_{n=0}^{\infty} z^n V_n, \\
&C(z)=\sum_{i=1}^{\infty} z^i c_i,\ \ |z|\le 1
\end{aligned}\right| \tag{3.13}$$

Now, we multiply equation (3.2) by $z^n$, take summations over $n$ from 1 to ∞, adding with (3.3) and using (3.13), we get

$$\frac{d}{dx}P_1(x,z)+\{b_1(\lambda-\lambda C(z))+\mu_1(x)\}P_1(x,z)=0 \tag{3.14}$$

Proceeding similarly for equations (3.4)-(3.6), we obtain

$$\frac{d}{dx}P_2(x,z)+\{b_1(\lambda-\lambda C(z))+\mu_2(x)\}P_2(x,z)=0 \tag{3.15}$$

$$\frac{d}{dx}V(x,z)+\{b_2(\lambda-\lambda C(z))+\phi(x)\}V(x,z)=0 \tag{3.16}$$

We now integrate equations (3.14) – (3.16) between limits 0 and $x$ and obtain

$$P_1(x,z)=P_1(0,z)e^{-b_1\lambda(1-C(z))x-\int_0^x \mu_1(t)dt}\ ; \tag{3.17}$$

$$P_2(x,z)=P_2(0,z)\ e^{-b_1\lambda(1-C(z))x-\int_0^x \mu_2(t)dt}\ ; \tag{3.18}$$

$$V(x,z)=V(0,z)e^{-b_2\lambda(1-C(z))x-\int_0^x \phi(t)dt}\ ; \qquad \lambda>0 \tag{3.19}$$

Further integrating the above equations by parts with respect to $x$ yields

$$P_1(z) = P_1(0,z)\frac{\left[1 - G_1^*\left(b_1\left(\lambda - \lambda C(z)\right)\right)\right]}{b_1\left(\lambda - \lambda C(z)\right)} \tag{3.20}$$

$$P_2(z) = P_2(0,z)\frac{\left[1 - G_2^*\left(b_1\left(\lambda - \lambda C(z)\right)\right)\right]}{b_1\left(\lambda - \lambda C(z)\right)} \tag{3.21}$$

$$V(z) = V(0,z)\frac{\left[1 - W^*\left(b_2\left(\lambda - \lambda C(z)\right)\right)\right]}{b_2\left(\lambda - \lambda C(z)\right)} \tag{3.22}$$

Next we multiply equation (3.9) with appropriate powers of *z*, take summation over all possible values of *n* and using (3.8) and (3.13), obtain

$$\begin{aligned} zP_1(0,z) = (1-p)\xi_1\left[\int_0^\infty P_1(x,z)\,\mu_1(x)\,dx + \int_0^\infty P_2(x,z)\,\mu_2(x,z)\,dx\right] \\ + \xi_1\int_0^\infty V(x,z)\phi(x)\,dx + b_1\,\xi_1\,\lambda\,(C(z)-1)Q \end{aligned} \tag{3.23}$$

We perform the similar operations on equations (3.10) & (3.11) to obtain

$$\begin{aligned} zP_2(0,z) = (1-p)\xi_2\left[\int_0^\infty P_1(x,z)\mu_1(x)dx + \int_0^\infty P_2(x,z)\mu_2(x,z)dx\right] \\ + \xi_2\int_0^\infty V(x,z)\phi(x)\,dx + b_1\,\xi_2\lambda\,(C(z)-1)Q \end{aligned} \tag{3.24}$$

$$V(0,z) = p\left[\int_0^\infty P_1(x,z)\mu_1(x)dx + \int_0^\infty P_2(x,z)\mu_2(x)dx\right] \tag{3.25}$$

Now multiplying equations (3.17) and (3.18) by $\mu_1(x), \mu_2(x)$ respectively then integrating the resulting equation over *x*, we get

$$\int_0^\infty P_1(x,z)\mu_1(x)dx = P_1(0,z)G_1^*\left(b_1\left(\lambda - \lambda C(z)\right)\right)\ , \tag{3.26}$$

$$\int_0^\infty P_2(x,z)\mu_2(x)dx = P_2(0,z)G_2^*\left(b_1\left(\lambda - \lambda C(z)\right)\right), \tag{3.27}$$

Similarly multiplying equation (3.19) by $\phi(x)$, we get

$$\int_0^\infty V(x,z)\phi(x)dx = V(0,z)W^*\left(b_2\left(\lambda - \lambda C(z)\right)\right), \tag{3.28}$$

where $G_j^*(b_1(\lambda - \lambda C(z))) = \int_0^\infty e^{-b_1(\lambda - \lambda C(z))x} dG_j(x)$ is the Laplace-Stieltjes transform of type j, (j=1,2) service and $W^*(b_2(\lambda - \lambda C(z))) = \int_0^\infty e^{-b_2(\lambda - \lambda C(z))x} dW(x)$ is the Laplace-Stieltjes transform of vacation time.

Substituting relations (3.26)-(3.28) in equations (3.23) and (3.25) and denoting $b_1\left(\lambda - \lambda C(z)\right) = m$, $b_2\left(\lambda - \lambda C(z)\right) = n$, we get

$$\begin{aligned} zP_1(0,z) &= (1-p)\xi_1\left[P_1(0,z)G_1^*(m) + P_2(0,z)G_2^*(m)\right] \\ &\quad + \xi_1 V(0,z)W^*(n) + b_1\lambda\xi_1\left(C(z)-1\right)Q \end{aligned} \tag{3.29}$$

$$\begin{aligned} zP_2(0,z) &= (1-p)\xi_2\left[P_1(0,z)G_1^*(m) + P_2(0,z)G_2^*(m)\right] \\ &\quad + \xi_1 V(0,z)W^*(n) + b_1\lambda\xi_2\left(C(z)-1\right)Q \end{aligned} \tag{3.30}$$

$$V(0,z) = p\left[P_1(0,z)G_1^*(m) + P_2(0,z)G_2^*(m)\right] \tag{3.31}$$

Utilizing (3.26)-(3.28) in equations (3.23) and (3.25) we have

$$\begin{aligned} \Rightarrow z\,P_1(0,z) &= (1-p)\,\xi_1\left[P_1(0,z)G_1^*(m) + P_2(0,z)G_2^*(m)\right] \\ &\quad + \xi_1\, p\left[\begin{pmatrix} P_1(0,z)\,G_1^*(m) \\ +P_2(0,z)\,G_2^*(m) \end{pmatrix}\right] W^*(n) \;+ \lambda\, b_1\, \xi_1\left(C(z)-1\right)Q \end{aligned} \tag{3.32}$$

Similarly

$$\begin{aligned} zP_2(0,z) &= (1-p)\xi_2\left[P_1(0,z)G_1^*(m) + P_2(0,z)G_2^*(m)\right] \\ &\quad + \xi_2\, p\left[P_1(0,z)\,G_1^*(m) + P_2(0,z)\,G_2^*(m)\right]W^*(n) + \lambda b_1\,\xi_2\,(C(z)-1)\,Q \end{aligned} \tag{3.33}$$

Solving equations (3.32) and (3.33), we get the following

$$P_1(0,z)=\frac{z\lambda b_1\xi_1\left(C(z)-1\right)Q}{z^2-z\left\{(1-p)+pW^*(n)\right\}\left[\xi_1G_1^*(m)+\xi_2G_2^*(m)\right]}$$

Rewriting it as, $P_1(0,z)=\dfrac{\lambda b_1\xi_1\left(C(z)-1\right)Q}{D(z)}$ (3.34)

Similarly,

$$P_2(0,z)=\frac{\lambda b_1\xi_2(C(z)-1)Q}{D(z)}, \tag{3.35}$$

$$V(0,z)=\frac{p\lambda b_1\left(C(z)-1\right)\left[\xi_1G_1^*(m)+\xi_2G_2^*(m)\right]Q}{D(z)}, \tag{3.36}$$

where, $D(z)=\left\{(1-p)+pW^*(n)\right\}\left[\xi_1G_1^*(m)+\xi_2G_2^*(m)\right]-z$ (3.37)

Substituting the values in equations (3.34)-(3.36) in (3.20)-(3.22), we obtain

$$P_1(z)=\frac{\xi_1[1-G_1^*(m)\ ]Q}{D(z)} \tag{3.38}$$

$$P_2(z)=\frac{\xi_2[1-G_2^*(m)\ ]Q}{D(z)} \tag{3.39}$$

$$V(z)=\frac{p\dfrac{b_1}{b_2}\left[1-W^*(n)\right]\left[\xi_1G_1^*(m)+\xi_2G_2^*(m)\right]Q}{D(z)} \tag{3.40}$$

Let us now define $P_q(z)$ as the probability generating function of the queue size irrespective of the type of service the server is providing, such that adding equations (3.38)-(3.40), we get

$$P_q(z)=P_1(z)+P_2(z)+V(z)=\frac{N(z)}{D(z)} \tag{3.41}$$

The unknown probability $Q$ is determined by using the relation of the normalizing condition

$$Q+P_1(1)+P_2(1)+V(1)=1$$

Now using L'Hopital's Rule as the equation (3.41) is indeterminate of the *zero/zero* form, at $z=1$, obtain

$$P_1(1)=\lim_{z\to 1}P_1(z)=\frac{\xi_1 b_1\lambda E(I)E(S_1)Q}{1-\lambda E(I)\left[b_2 pE(V)+b_1\left\{\xi_1 E(S_1)+\xi_2 E(S_2)\right\}\right]}\quad , \tag{3.42}$$

$$P_2(1)=\underset{z\to 1}{Lim}\,P_2(z)=\frac{\xi_2 b_1\lambda E(I)E(S_2)}{1-\lambda E(I)\left[b_2 pE(V)+b_1\left\{\xi_1 E(S_1)+\xi_2 E(S_2)\right\}\right]}\quad , \tag{3.43}$$

$$V(1)=\underset{z\to 1}{Lim}\,V(z)=\frac{pb_1\lambda E(I)E(V)}{1-\lambda E(I)\left[b_2 pE(V)+b_1\left\{\xi_1 E(S_1)+\xi_2 E(S_2)\right\}\right]}\quad , \tag{3.44}$$

where $E(I)$ is the mean size of batch of arriving customers, $E(S_1)$, $E(S_2)$ are the mean service times of type 1 and type 2 services respectively, $E(V)$ is the mean of vacation time and $G_j^*(0)=1(j=1,2),\ V^*(0)=1$

The R. H. S of the results (3.42)-(3.44) respectively give the steady state probability that the server is busy providing type 1 service, type 2 service and probability of the server on vacation state. Now adding equations (3.42)-(3.44) we get

$$P_q(1)=\frac{b_1\lambda E(I)\left[\left\{\xi_1 E(S_1)+\xi_2 E(S_2)\right\}+pE(V)\right]Q}{1-\lambda E((I)\left[b_2 pE(V)+b_1\left\{\xi_1 E(S_1)+\xi_2 E(S_2)\right\}\right]}$$

The above equation gives the steady state probability that the server is busy, irrespective of whether it is providing type 1 or type 2 services and on vacation.

Let us simplify the normalizing condition $Q+P_Q(1)=1$ to get Q and thus have

$$Q=1-\frac{b_1\lambda E(I)\left[\xi_1 E(S_1)+\xi_2 E(S_2)+pE(V)\right]}{1+p(b_1-b_2)\lambda E(I)E(V)} \tag{3.45}$$

And the utilization factor $\rho=1-Q$ is given by

$$\rho=\frac{b_1\lambda E(I)\left[\xi_1 E(S_1)+\xi_2 E(S_2)\right]}{1+p(b_1-b_2)\lambda E(I)E(V)}<1$$

$$P_q(z) = \frac{Q\begin{bmatrix} \{\xi_1(G_1^*(m) + \xi_2 G_2^*(m) - 1\} \\ + p\dfrac{b_1}{b_2}(W^*(n) - 1)\{\xi_1 G_1^*(m) + \xi_2 G_2^*(m)\} \end{bmatrix}}{z - \{(1-p) + pW^*(n)\}\left[\xi_1 G_1^*(m) + \xi_2 G_2^*(m)\right]} \tag{3.46}$$

Equation (3.46) denotes the probability generating function of the queue size irrespective of the state of the system for a $M^X / \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} / 1 / Vs$ queue with balking during busy and vacation periods.

### 3.2.4 The Average Queue Size

Let $Lq$ denote the mean queue size at random epoch. Then $L_q = \lim_{z \to 1} \frac{d}{dz} P_q(z)$ . Since this formula is also of *zero/zero* form, we use L' Hopital's Rule twice and obtain

$$L_q = \lim_{z \to 1} \frac{D'(z)N''(z) - N'(z)D''(z)}{2(D'(z))^2} \tag{3.47}$$

The primes and double primes in (3.47) denote first and second derivatives at $z$ =1 respectively.

The detail of obtaining equation (3.47) is explained in appendix. Therefore we have

$$N'(1) = Q\left[b_1 \lambda E(I)\{\xi_1 E(S_1) + \xi_2 E(S_2)\} + b_1 p \lambda E(I) E(V)\right]$$

$$N''(1) = Q\begin{bmatrix} [b_1 \lambda E(I(I-1))\{\xi_1 E(S_1) + \xi_2 E(S_2) + pE(V)\} \\ + (b_1 \lambda E(I))^2 \{\xi_1 E(S_1^2) + \xi_2 E(S_2^2) + 2pE(V)[\xi_1 E(S_1) + \xi_2 E(S_2)]\} + pb_1 b_2 (\lambda E(I))^2 E(V^2) \end{bmatrix}$$

$$D'(1) = 1 - \lambda E(I)\left[b_1 \xi_1 E(S_1) + b_1 \xi_2 E(S_2) + pb_2 E(V)\right]$$

$$\begin{aligned} D''(1) = &-(b_1 \lambda E(I))^2 \left[\xi_1 E(S_1^2) + \xi_2 E(S_2^2)\right] - p(b_2 \lambda E(I))^2 E(V^2) \\ &- \lambda E(I(I-1))\left[b_1 \xi_1 E(S_1) + b_1 \xi_2 E(S_2) + pb_2 E(V)\right] \\ &- 2pb_1 b_2 (\lambda E(I))^2 E(V)\left[\xi_1 E(S_1) + \xi_2 E(S_2)\right] \end{aligned}$$

Here $E(S_1^2), E(S_2^2), E(V^2)$ are the second moments of service times of type 1, type2 and vacation time respectively. $E(I(I-1))$ is the second factorial moment of the batch of arriving customers,

$G_1^*(0)=1,\ G_2^*(0)=1$, $V^*(0)=1$ and $Q$ has been obtained in (3.45). Substituting the above expressions and Q derived in (3.45) in equation (3.47), the mean of queue size can be derived.

## 3.3 Model 2: Steady state analysis of a batch arrival queue with balking and two stages of services

In this model, we consider a queueing system where a single server provides two stages of heterogeneous services, first stage (FS) service followed by second stage (SS) service. A service is complete when both the stages of services are completed. The service discipline is assumed to be on a first come first served basis. We assume that the service time $S_j$ ( j =1, 2) of the j th stage service follows a general probability distribution with distribution function $G_j(s_j)$ and $g_j(s_j)$ being the probability density function. Let $\mu_j(x)$ be the conditional probability at stage j (j=1, 2) service during the period (*x*, *x*+*dx*] given elapsed time is *x* such that $\mu_j(x)=\frac{g_j(x)}{1-G_j(x)};\ j=1,2.$ As soon as the SS service gets completed, the server may go for vacation for a random length of time V (vacation period) with probability $p$ or may continue to serve the next unit with probability $(1-p)$ The vacation time is assumed to follow a general probability law as outlined in the first model in Section 3.2.1.

The probabilities, $P^{(1)}(t,x)$, $P^{(2)}(t,x)$, $V_n(t,x)$ and $Q(t)$ denoting the first stage, second stage service times, vacation time and idle time are defined just before Chapter One. Our model can be denoted in notation as the $M^X/G_1G_2/1/Vs/balking$ queue.

Thus the time required by an unit to complete a service cycle, in other words, the server to be ready to provide service to the next customer, is given by

$$S = S_1+S_2+V \quad \text{with probability } p,$$

$$= S_1+S_2 \quad \text{with probability } (1-p),$$

so that the Laplace- Steiltjes Transform of the total service time is given by

$$G^*(\theta)=E(e^{-\theta S})=E(e^{-\theta S_1})E(e^{-\theta S_2})=G_1^*(\theta)G_2^*(\theta)$$

Then LST for S is

$$G^*(\theta) = (1-p)G_1^*(\theta)G_2^*(\theta) + pG_1^*(\theta)G_2^*(\theta)W^*(\theta)\,, \tag{3.48}$$

and the first two moments of G are given by

$$E(S) = -\frac{d}{d\theta}G^*(\theta)\bigg|_{\theta=0} = E(S_1) + E(S_2) + pE(V)$$

$$E(S^2) = (-1)^2 \frac{d}{d\theta}G^*(\theta)\bigg|_{\theta=0} = E(S_1^2) + E(S_2^2) + 2\big[E(S_1)E(S_2) + pE(V)\{E(S_1) + E(S_2)\}\big]$$

### 3.3.1 Queue Size Distribution at a Random Epoch

Now, assuming that the steady state exists (the condition(s) under which it may exist will emerge later), we define the following steady state probabilities:

$$\begin{gathered}\lim_{t\to\infty} P_n^{(j)}(t,x) = P_n^{(j)}(x), \quad \lim_{t\to\infty} P_n^{(j)}(t) = \int_0^\infty P_n^{(j)}(t,x)x = P_n^{(j)}; \quad j = 1,2 \\ \lim_{t\to\infty} V_n(t,x) = V_n(x), \quad \lim_{t\to\infty} V_n(t) = \int_0^\infty V_n(t,x)dx = V_n\,, \quad \lim_{t\to\infty} Q(t) = Q\end{gathered} \tag{3.49}$$

A diagrammatic representation of the queueing model is as shown below:

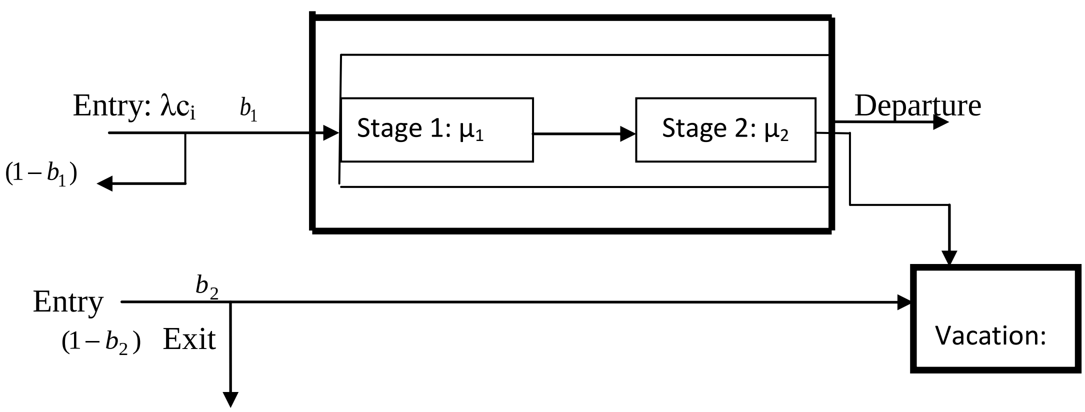


*Figure 3.1: A two-stage queuing process with balking during service and vacations.*

The steady state behavior of the queueing process at a random epoch can be performed with following set of differential-difference equations:

$$\frac{d}{dx}P_n^{(1)}(x)+\{\lambda+\mu_1(x)\}P_n^{(1)}(x)=\lambda(1-b_1)P_n^{(1)}(x)+b_1\lambda\sum_{i=1}^{n}c_iP_{n-i}^{(1)}(x) \quad ; \quad n\geq 1, \tag{3.50}$$

$$\frac{d}{dx}P_0^{(1)}(x)+\left(\lambda+\mu_1(x)\right)P_0^{(1)}(x)=\lambda(1-b_1)P_0^{(1)}(x), \tag{3.51}$$

$$\frac{d}{dx}P_n^{(2)}(x)+\{\lambda+\mu_2(x)\}P_n^{(2)}(x)=\lambda(1-b_1)P_n^{(2)}(x)+b_1\lambda\sum_{i=1}^{n}c_iP_{n-i}^{(2)}(x) \quad ; n\geq 1, \tag{3.52}$$

$$\frac{d}{dx}P_0^{(2)}(x)+\left(\lambda+\mu_2(x)\right)P_0^{(2)}(x)=\lambda(1-b_1)P_0^{(2)}(x) \tag{3.53}$$

$$\frac{d}{dx}V_n(x)+\{\lambda+\phi(x)\}V_n(x)=\lambda(1-b_2)+b_2\lambda\sum_{i=1}^{n}c_iV_{n-i}(x) \quad ; \qquad n\geq 1, \tag{3.54}$$

$$\frac{d}{dx}V_0(x)+\{\lambda+\phi(x)\}V_0(x)=\lambda(1-b_2)V_0(x). \tag{3.55}$$

The probability of idle state is defined as

$$\lambda Q=\lambda(1-b_1)Q+(1-p)\int_0^\infty P_0^{(2)}(x)\mu_2(x)dx+\int_0^\infty V_0(x)\phi(x)dx \tag{3.56}$$

The differential equations now have to be solved subject to the following boundary conditions

$$P_n^{(1)}(0)=\lambda b_1c_{n+1}Q+(1-p)\int_0^\infty P_{n+1}^{(2)}(x)\mu_2(x)dx+\int_0^\infty V_{n+1}(x)\phi(x)dx, \quad n\geq 0, \tag{3.57}$$

$$P_n^{(2)}(0)=\int_0^\infty P_n^{(1)}(x)\mu_1(x)dx \qquad n\geq 0, \tag{3.58}$$

$$V_n(0)=p\int_0^\infty P_n^{(2)}(x)\mu_2(x)dx \qquad n\geq 0, \tag{3.59}$$

The probability generating functions are as defined in equation (3.13).

Applying the standard procedure for solving the boundary conditions, and proceeding in the usual manner as discussed in Section 3.2, we obtain after simplification

$$P^{(1)}(z)=\frac{Q(1-G_1^*(m))}{D(z)}, \tag{3.60}$$

$$P^{(2)}(z)=\frac{QG_1^*(m)(1-G_2^*(m))}{D(z)}, \tag{3.61}$$

$$V(z)=\frac{Qp\frac{b_1}{b_2}\left[1-W^*(n)\right]G_1^*(m)G_2^*(m)}{D(z)},$$

(3.62)

And $D(z)=z-\{(1-p)+pW^*(n)\}\left[G_1^*(m)G_2^*(m)\right]$ , where $m=b_1(\lambda-\lambda C(z)),\ n=b_2(\lambda-\lambda C(z))$ ,

$G_j^*(m)=\int_0^\infty e^{-mx}dG_j(x);\ j=1,2$ , and $W^*(n)=\int_0^\infty e^{-nx}dW(x)$ are the L-S transform of jth stage service time (j=1, 2 ) and vacation time respectively.

Let $P_q(z)$ be the probability generating function of the queue size at random epoch. Then

$$P_q(z)=P^{(1)}(z)+P^{(2)}(z)+V(z)=\frac{N(z)}{D(z)}=\frac{Q\left[1-G_1^*(m)G_2^*(m)\right]+Qp\frac{b_1}{b_2}\left[1-W^*(n)\right]G_1^*(m)G_2^*(m)}{\{1-p+pW^*(n)\}G_1^*(m)G_2^*(m)-z} \tag{3.63}$$

Therefore, the probability generating function of the total service time at random epoch, using (3.48), is given by

$$P_q(z)==\frac{Q\left[1-\left\{(1-p\frac{b_1}{b_2})G_1^*(m)G_2^*(m)+p\frac{b_1}{b_2}W^*(n)G_1^*(m)G_2^*(m)\right\}\right]}{\{1-p+pW^*(n)\}G_1^*(m)G_2^*(m)-z} \tag{3.64}$$

Utilizing (3.48), the PGF of the steady state of the system can be expressed as

$$P_s(z)=Q+zP_q(z)=\frac{Q(1-z)[G^*(m)-1]}{G^*(m)-z} \tag{3.65}$$

We may note that equation (3.65) gives the well known Pollaczek-Khinchine formula (*Medhi, 1994*).

Since the R. H. S of (3.65) is indeterminate of the *zero/zero* form, at $z=1$, applying L' Hopital's rule we obtain

$$P_q(1)=\lim_{z\to 1}P_q(z)=\frac{Qb_1\lambda E(I)\left[E(S_1)+E(S_2)+pE(V)\right]}{1-\lambda E(I)\left[b_1\{E(S_1)+E(S_2)\}+pb_2E(V)\right]} \tag{3.66}$$

Using the normalizing condition: $Q+P_q(1)=1$, to find $Q$

We obtain

$$Q=\frac{1-\lambda E(I)\left[b_1\left\{E(S_1)+E(S_2)\right\}+pb_2E(V)\right]}{1+p(b_1-b_2)\lambda E(I)E(V)} \tag{3.67}$$

Further

$Q=1-\rho$ gives

$$\rho=\frac{b_1\lambda E(I)\left[\left\{E(S_1)+E(S_2)\right\}+pE(V)\right]}{1+p(b_1-b_2)\lambda E(I)E(V)}<1 \tag{3.68}$$

This is the stability condition under which steady state exists.

Consequently the system state probabilities that can be obtained from equations (3.60)-(3.62) as

*Prob*[ the server is busy with FS service] = $P^{(1)}(1)=b_1\lambda E(I)E(S_1)$

*Prob*[the server is busy with SS service] = $P^{(2)}(1)=b_1\lambda E(I)E(S_2)$

*Prob*[the server is on vacation] = $V(1)=pb_2\lambda E(I)E(V)$

### 3.3.2 Mean Queue Size and Mean Waiting Time

Let $L_q$ denote the mean queue size at random epoch, then

$$L_q=\left.\frac{dP_q(z)}{dz}\right]_{z=1} \tag{3.69}$$

Using L' Hopital's Rule twice as the R. H. S of $P_q(z)$ is indeterminate of the 0/0 form, the following expressions are derived from (3.64) differentiating twice. Thus we get

$$N^{/}(1)=\lambda E(I)b_1\left\{E(S_1)+E(S_2)+pE(V)\right\}Q\ ,$$

$$D'(1)=1-\lambda E(I)\left[b_1\left\{E(S_1)+E(S_2)\right\}+pb_2E(V)\right],$$

$$\begin{aligned}N''(1)=&\ b_1\lambda E(I(I-1))\left[E(S_1)+E(S_2)+pE(V)\right]Q\\&+\left(b_1\lambda E(I)\right)^2\left[E(S_1^2)+E(S_2^2)+2E(S_1)E(S_2)+2pE(V)\left\{E(S_1)+E(S_2)\right\}\right]Q\\&+pb_1b_2\left(\lambda E(I)\right)^2E(V^2)Q\end{aligned},$$

$$\begin{aligned}D^{//}(1)=&-\lambda E(I(I-1))\left[b_1\left\{E(S_1)+E(S_2)\right\}+pb_2E(V)\right]\\&-\left(b_1\lambda E(I)\right)^2\left[E(S_1^2)+E(S_2^2)++2E(S_1)E(S_2)\right]\\&-p\left(b_2\lambda E(I)\right)^2E(V^2)-2pb_1b_2\left(\lambda E(I)\right)^2E(V)\left[E(S_1)+E(S_2)\right],\end{aligned}$$

Substituting the above results in equation (3.47) and $Q$ from (3.67), the mean size of the queue can thus be obtained. Because of the complex structure of the equations we cannot write the mean queue size in explicit form.

We can further obtain $L = L_q + \rho$ and $W_q = \frac{L_q}{\lambda E(I)}$ by using the above relations in each of the models, which gives average size of the queue system and average waiting time in the queue.

## 3.4 Model 3: An $M^X / G_1 G_2 / 1 / Vs$ queue with balking and optional re-service

There are instances in real life when a customer maybe unsatisfied with the service offered and demand re-service or repeat the service. Let us investigate the second model with two stages of service and we assume that a customer desiring re-service is instantaneously taken up for re-service. As soon as the service of any stage is completed, the customer has the option to leave the service or opt for re-service. Let us consider that the probability of repeating *jth* stage *(j=1,2)* service is $r_j$ and leaving the system without re-service is $(1-r_j), j=1,2.$ The rest of the assumptions are similar to those defined in Model 2 in Section3.3. We now derive the steady state results for the probability generating function of queue size at a random point of time.

### 3.4.1 Steady State Queue Size Distribution at a Random Epoch

Let $S = S_1 + S_2$ is the total or overall service time received in both stages. We denote

$S_1 = S_1^{(1)} + S_1^{(2)}$ as the first stage service time with re-service .

Again since $S_1$ and $S_2$ are independent random variables for first and second stage service time and $G_j^*(m) = E(e^{-mx}) = \int_o^\infty e^{-mx} dG_j(x); j=1,2$ , then the L-S transform of total service time for both stages will be

$G_T(m) = E(e^{-m(S_1+S_2)}) = E(e^{-mS_1})E(e^{-mS_2}) = G_1^*(m)G_2^*(m)$, $m$ is as defined earlier.

We write $S_1 = \begin{cases} S_1 \text{ , } without\ re-service\ (1-r_1) \\ S_1^{(1)} + S_1^{(2)}, with\ re-service\ r_1 \end{cases}$

Then

$$G_{T1}(m) = (1-r_1)E\left(e^{-m(S_1)}\right) + r_1 E\left(e^{-m(S_1^{(1)}+S_1^{(2)})}\right)$$
$$= (1-r_1)G_1^*(m) + r_1\left[G_1^*(m)\right]^2$$

Similarly, the total service for second stage service with a re-service in the second stage is $G_{T_2}(m) = (1-r_2)G_2^*(m) + r_2\left[G_2^*(m)\right]^2$ and thus the total service time for both stages with or without re-services is $G_T(m) = E(e^{-mS}) = E\left(e^{-m(S_1+S_2)}\right) = E(e^{-mS_1})E\left(e^{-mS_2}\right) = G_{T_1}(m)G_{T_2}(m)$

$$G_T{}^*(m) = (1-r_1)(1-r_2)G_1^*(m)G_2^*(m) + r_1(1-r_2)\left[G_1^*(m)\right]^2 G_2^*(m) + r_2(1-r_1)G_1^*(m)\left[G_2^*(m)\right]^2 + r_1r_2\left[G_1^*(m)\right]^2\left[G_2^*(m)\right]^2 \tag{3.70}$$

Now equations (3.60) and (3.61) combine as

$$P(z) = \frac{Q\left(1-G_1(m)G_2(m)\right)}{D(z)} = \frac{Q\left[1-\left\{\begin{array}{l}(1-r_1)(1-r_2)G_1^*(m)G_2^*(m) + r_1(1-r_2)\left[G_1^*(m)\right]^2 G_2^*(m)\\ + r_2(1-r_1)G_1^*(m)\left[G_2^*(m)\right]^2 + r_1r_2\left[G_1^*(m)\right]^2\left[G_2^*(m)\right]^2\end{array}\right\}\right]}{D(z)}$$

Or simply can be written as

$$P(z) = \frac{Q\left(1-G_T{}^*(m)\right)}{D(z)} \tag{3.71}$$

Similarly equation (3.62) modifies to

$$V(z) = \frac{p\dfrac{b_1}{b_2}\left[1-W^*(n)\right]G_1^*(m)G_2^*(m)Q}{D(z)} = \frac{p\dfrac{b_1}{b_2}\left[1-W^*(n)\right]G_T^*(m)Q}{D(z)} \tag{3.72}$$

Therefore the probability generating function of the queue size at random point derived from equation (3.71) and (3.722) is obtained as

$$P_q(z) = \frac{Q\left[1-G_T^*(m)\left\{1-p\dfrac{b_1}{b_2}\left(1-W^*(n)\right)\right\}\right]}{D(z)}\ , \tag{3.73}$$

where $D(z) = \left\{1-p+pW^*(n)\right\}G_T(m) - z$, the probability of empty state Q is

$$Q=\frac{1-\lambda E(I)\left[b_1\{(1+r_1)E(S_1)+(1+r_2)E(S_2)\}+pb_2E(V)\right]}{1+p(b_1-b_2)\lambda E(I)E(V)} \quad , \tag{3.74}$$

and $G_T^*(m)$ is obtained in equation (3.70). Q has been similarly derived using the normalization condition.

Thus the above relation in equation (3.73) represents the Pollaczek-Khinchine formula for such type of models and we consider it as a classical generalization of P-K formula for $M^X/G/1/Vs$ queue with optional re-service and two stages of services.

Remark: In order to check the consistency of our results, if we consider that there is no vacations, no re-services, i.e. $p=0, r_1=r_2=0$, then equation (3.73) reduces to

$$P_q(z)=\frac{Q(1-G_T^{\;*}(m))}{G_T^*(m)-z}$$

which is the *P-k* formula for a classical $M^X/G/1$queue.

### 3.4.2 Mean queue size and mean waiting time

As discussed in Section 3.3.1, we further utilize equation (3.47) to derive the expected queue length. The equations are obtained as

$$N'(1)=\lambda E(I)b_1\{(1+r_1)E(S_1)+(1+r_2)E(S_2)+pE(V)\}Q$$

$$N''(1)=Q\begin{bmatrix} b_1\lambda E(I(I-1))\left[\{(1+r_1)E(S_1)+(1+r_2)E(S_2)\}+pE(V)\right] \\ +\left(b_1\lambda E(I)\right)^2\begin{bmatrix}(1+r_1)E(S_1^2)+(1+r_2)E(S_2^2) \\ +2\{r_1(E(S_1))^2+r_2(E(S_2))^2\} \\ +2(1+r_2)(1+r_1)E(S_1)E(S_2) \\ +2pE(V)\{(1+r_1)E(S_1)+(1+r_2)E(S_2)\}\end{bmatrix} \\ +pb_1b_2\left(\lambda E(I)\right)^2E(V^2)\end{bmatrix}$$

$$D'(1)=1-\lambda E(I)\left[b_1\{(1+r_1)E(S_1)+(1+r_2)E(S_2)\}+pb_2E(V)\right]$$

$$\begin{aligned}D''(1)=&-\lambda E(I(I-1))\left[b_1\{(1+r_1)E(S_1)+(1+r_2)E(S_2)\}+pb_2E(V)\right]\\&-\left(b_1\lambda E(I)\right)^2\begin{bmatrix}(1+r_1)E(S_1^2)+(1+r_2)E(S_2^2)+2\{r_1(E(S_1))^2+r_2(E(S_2))^2\} \\ +2(1+r_1)(1+r_2)E(S_1)E(S_2)\end{bmatrix}\\&-p\left(b_2\lambda E(I)\right)^2E(V^2)-2pb_1b_2\left(\lambda E(I)\right)^2E(V)\left[(1+r_1)E(S_1)+(1+r_2)E(S_2)\right]\end{aligned}$$

and $Q$ has been obtained in (3.74).

Similarly the mean queue size of the system can be obtained using the relation $L = \frac{d}{dz} P_s(z)$ at $z = 1$, where $L$ denotes the expected queue size for the system and $P_s(z)$, the probability generating function of the system at a random epoch. Alternatively we can find $L = L_q + \rho$. Further we can obtain the average waiting time in the queue and system by the relation: $W_q = \frac{L_q}{\lambda E(I)}$ and $W = \frac{L}{\lambda E(I)}$

## 3.5 Special Case: Erlang-*k* vacation time

Specifying the vacation time random variable, we can discuss a particular case with this queueing system. Let us consider that the vacation time random variable has an Erlang-*k* distribution, with probability density function,

$w(v) = \frac{(\phi k)^k v^{k-1} e^{-\phi v}}{(k-1)!}$; $\phi > 0$, $k \geq 1$, then

$W^*(n) = \left[ \frac{\phi k}{\phi k + n} \right]^k$; $E(V) = \frac{1}{\phi}$, where $n$ is as defined earlier.

Substituting for $W^*(n)$ in equation (3.46) we have

$$P_q(z) = \frac{Q\left[1 - \left(\xi_1 G_1^*(m) + \xi_2 G_2^*(m)\right) + p\frac{b_1}{b_2}\left[1 - \left(\frac{\phi k}{\phi k + n}\right)^k\right]\left(\xi_1 G_1^*(m) + \xi_2 G_2^*(m)\right)\right]}{\left\{(1-p) + p\left[\left(\frac{\phi k}{\phi k + n}\right)^k\right]\right\}\left\{\xi_1 G_1^*(m) + \xi_2 G_2^*(m)\right\} - z} \tag{3.75}$$

Equation (3.75) gives the probability generating function of a $M^X / \begin{pmatrix} G_1 \\ G_2 \end{pmatrix} / V_s / 1$ queue having $E_k$ vacation time distribution with balking.

Further from equation (3.64) we get

$$P_q(z)=\frac{Q\left[1-G_1^*(m)G_2^*(m)+p\frac{b_1}{b_2}\left[1-\left(\frac{\phi k}{\phi k+n}\right)^k\right]G_1^*(m)G_2^*(m)\right]}{\left\{(1-p)+p\left(\left[\frac{\phi k}{\phi k+n}\right]^k\right)\right\}G_1^*(m)G_2^*(m)-z} \tag{3.76}$$

This gives the probability generating function of queue size of a $M^X/G_1G_2/1/Vs$ queue with $E_k$ vacation time and balking.

Further considering the system with no balking, i.e. $b_1=b_2=1$ then

$$P_q(z)=\frac{Q\left[p\left\{1-\left(\frac{\phi k}{\phi k+\lambda-\lambda C(z)}\right)^k\right\}G_1^*(\lambda-\lambda C(z))G_2^*(\lambda-\lambda C(z))+1-G_1^*(\lambda-\lambda C(z))G_2^*(\lambda-\lambda C(z))\right]}{\left\{(1+p\left(\left(\frac{\phi k}{\phi k+\lambda-\lambda C(z)}\right)^k-1\right)\right\}G_1^*(\lambda-\lambda C(z))G_2^*(\lambda-\lambda C(z))-z} \tag{3.77}$$

The above equation (3.77) gives the probability generating function of queue size at a stationary point of time for a $M^X/G_1G_2/1/V_s$ queue with Erlang-*k* vacation time, studied by Choudhury and Madan (2004).

The probability generating function of queue size at a random point of time with optional re-service with Erlang-k vacation time can be derived as

$$P_q(z)=\frac{Q\left[1-G_T{}^*(m)\left\{1-p\frac{b_1}{b_2}\left(1-\left(\frac{\phi k}{\phi k+n}\right)^k\right)\right\}\right]}{\left[1-p+p\left(\frac{\phi k}{\phi k+n}\right)^k\right]G_T^*(m)}, \tag{3.78}$$

and
$$\begin{aligned}G_T{}^*(m)&=(1-r_1)(1-r_2)G_1^*(m)G_2^*(m)+r_1(1-r_2)\left[G_1^*(m)\right]^2G_2^*(m)\\&+r_2(1-r_1)G_1^*(m)\left[G_2^*(m)\right]^2+r_1r_2\left[G_1^*(m)\right]^2\left[G_2^*(m)\right]^2\end{aligned}$$

If $k$ =1 in (3.77) and arrivals are single instead of batches, then

$$P_q(z) = \frac{Q\left[(\phi+\lambda-\lambda z) - G_1^*(\lambda-\lambda z)G_2^*(\lambda-\lambda z)\{\phi+\lambda-\lambda z\}\right]}{\{\phi+(\lambda-\lambda z)(1-p)\}G_1^*(\lambda-\lambda z)G_2^*(\lambda-\lambda z) - z(\phi+\lambda-\lambda z)} \tag{3.79}$$

This gives the probability generating function of queue size at a stationary point for a $M/G_1G_2/1/Vs$ with Exponential vacation time.

## 3.6 Key Results

The probability generating function of queue size at a stationary point of time for a classical $M^X/G/1$ system is $P_q(z) = \frac{Q\left[1-G^*(\lambda-\lambda C(z))\right]}{G^*(\lambda-\lambda C(z))-z}$.

The generalization of this model with balking, Bernouli schedule server vacations and optional re-service is a re-definition of the generating function of the service time leading to the following results:

1. The probability generating function of queue size at a stationary point of time for an $M^X/\begin{pmatrix}G_1\\G_2\end{pmatrix}/1$ queue with balking and Bernoulli schedule server vacations is

$$P_q(z) = \frac{Q\left[1-\left\{1-p\frac{b_1}{b_2}+p\frac{b_1}{b_2}W^*\left(b_2(\lambda-\lambda C(z))\right)\right\}\{\xi_1 G_1^*\left(b_1(\lambda-\lambda C(z))\right)+\xi_2 G_2^*\left(b_1(\lambda-\lambda C(z))\right)\}\right]}{\left\{1-p+pW^*\left(b_2(\lambda-\lambda C(z))\right)\right\}\left\{\xi_1 G_1^*\left(b_1(\lambda-\lambda C(z))\right)+\xi_2\, G_2^*\left(b_1(\lambda-\lambda C(z))\right)\right\}-z}$$

2. The probability generating function of queue size at a stationary point of time for an $M^X/G_1G_2/1$ queue with balking and Bernoulli schedule server vacations is

$$P_q(z) = \frac{Q\left[1-\left\{1-p\frac{b_1}{b_2}+p\frac{b_1}{b_2}W^*\left(b_2(\lambda-\lambda C(z))\right)\right\}G_1^*\left(b_1(\lambda-\lambda C(z))\right)G_2^*\left(b_1(\lambda-\lambda C(z))\right)\right]}{\left\{1-p+pW^*\left(b_2(\lambda-\lambda C(z))\right)\right\}G_1^*\left(b_1(\lambda-\lambda C(z))\right)G_2^*\left(b_1(\lambda-\lambda C(z))\right)-z}$$

3. The probability generating function of the queue size at a random point of time for an $M^X/G_1G_2/1$ vacation queue with balking and optional re-service is

$$P_q(z) = \frac{Q\left[1-G_T^*\left(b_1\left(\lambda-\lambda C(z)\right)\right)\left\{1-p\frac{b_1}{b_2}\left(1-W^*\left(b_2\left(\lambda-\lambda C(z)\right)\right)\right)\right\}\right]}{\left\{1-p+pW^*\left(b_2\left(\lambda-\lambda C(z)\right)\right)\right\}G_T^*\left(b_1\left(\lambda-\lambda C(z)\right)\right)-z}$$

where $G_T^*\left(b_1\left(\lambda-\lambda C(z)\right)\right)$ is the generalized service time along with optional re-service of either stage of services.

(*Reference may be made to page 42, Sec 3.2.1, for the definitions of the various parameters*)

## 3.7 Summary

This chapter studies a non-Markovian queueing system with general service time, general vacation time and explores a queueing system with a single server offering two types of services and two stages of services as two different models. Further a third model has been discussed with the addition of optional re-service. We have derived the steady state queue size distribution at random point of time and these probability distribution allows to measures of effectiveness, mean number of customers in the system and mean number of customers in the queue. A special case is also analyzed with an Erlang-k vacation time.

******

# Chapter 4

# Analysis of a Batch Arrival Queueing System with Reneging During Server Vacations

## 4.1 Introduction

The prospect of waiting in a queueing system induces impatience on customers. The longer customers wait, the more dissatisfied they are supposed to be. If a customer leaves after joining a queue due to impatience, in queueing parlance, it is termed as reneging. The pioneering work on reneging was by Barrer (1957) who studied deterministic reneging with single server Markovian arrival and service rates. Many authors have treated the reneging phenomenon with various assumptions and we see significant contribution by numerous researchers in this area. Boxma and de Waal (1994) studied the M/M/c queues with impatient customers. Altman and Borokov (1997) investigated the stability issue in a retrial queue with impatient customers. In recent years, an M/G/1 queue with deterministic reneging was studied by Bae et al. (2001), Zhang et al. (2005) considered an M/M/1/N framework with Markovian reneging where they had derived the steady state solution along with a cost model. More recently, study of reneging within vacations in queueing systems has attracted the attention of many researchers. Altman and Yechiali (2006) studied customers' impatience in vacations systems of M/M/1, M/G/1 and M/M/c queues. In recent years, customers' impatience was studied by Dimou et al.(2011), Selvaraju and Goswami (2013), Kumar et al.(2014), investigated queueing systems with impatient customers, mentioning a few. A detailed review about the literature is discussed in Section 2.3 of Chapter Two.

Most of the literature mentioned here concentrated on studying reneging in queueing systems with single arrivals and Markovian service times. To the best of our knowledge, though reneging during the absence of the server has been studied earlier it has not been treated in the context of a batch arrival queuing system along with general service time distribution.

Further as far as the existing literature is concerned, reneging has not been treated for Bernoulli schedule server vacations. All this had led to the motivation of doing the current study.

In this chapter, customers in batches are assumed to arrive in a queuing system, join the queue but renege when the server is unavailable due to vacations. This is a common phenomenon witnessed in many real life situations especially with human customers, where the customer leaves the queue after joining without receiving any service, due to impatience upon the unavailability of the server. This can also be seen in communication networks, when there are several jobs waiting to be processed but tends to leave unprocessed if the network or signal of the server is not available. Thus reneging is one important feature which affects the efficiency of a queuing system due to the loss of customers.

Thus in this chapter, the author attempt to study a batch arrival queueing system with two different service time distributions, Markovian and non-Markovian distribution as two separate models. Reneging is considered to occur only during server vacations. Initially, we formulate a model where service time is assumed to follow Exponential distribution and then follow to generalize the first model considering a non-Markovian service time distribution. The partial generating functions for the distribution of queue sizes for both vacation state and busy state are obtained. Solving the differential equations arising, the closed-form steady states solution of the queue size at random point of time are derived obtained in both the models. Performance measures like expected value of queue length and variance of queue size are also obtained. As a special case, the $Geom/\gamma/1$ queue with reneging during server vacations is discussed with a numerical illustration.

The rest of the chapter is organized as follows. Section 4.2 discusses the steady state behavior of the queueing system with Exponential service time with customers reneging independently during server vacation. The purpose of this section is to first formulate the model in a Markovian set up in order to ease out the complex structure of the equations due to reneging, involved in the second model, as they are inherently easier to handle than the non-Markov case. In Section 4.3, we generalize Model 1 with an arbitrary (non-Markovian) service time distribution and analyze the stationary queue size distribution. Section 4.4 illustrates the model with some numerical examples to demonstrate the effects of the parameters in the various performance measures.

## 4.2 Analysis of Model 1: Markovian Case-Exponential Service time

### 4.2.1 Assumptions Underlying the Mathematical Model:

a) Arrivals are assumed to be in batches following a compound Poisson process as discussed in Chapter Three. The queueing system has a single server providing service to one customer at a time on a first come first served basis. Service time is assumed to follow an Exponential distribution so that $\mu\Delta t$ be the probability of a service completion during a small interval (t, $t+\Delta t$].

b) It is assumed that once a service is complete, the server can go on vacation for a random length of time with probability $p$ or remain in the system with probability $(1-p)$, $0 \le p \le 1$. The durations of vacation is assumed to follow an Exponential distribution with $\phi > 0$ and hence mean of vacation time is $\frac{1}{\phi}$.

c) Further, customers in the queue may renege (leave the queue) due to impatience only during the absence of the server, when the server takes a vacation. Reneging is assumed to follow an Exponential distribution with parameter $\gamma$. Thus $f(t) = \gamma e^{-\gamma t}\,\Delta t$, $\gamma > 0$. Thus $\gamma\Delta t$ is the probability that a customer can renege during a short interval of time ($t$, $t+\Delta t$] during vacation. Each customer makes a decision to renege independently; the departure rate from a queue of length $n$ is $n\gamma$.

d) We assume that all stochastic processes involved in the study are independent of each other.

The probabilities of the different states of the system $P_n(t), V_n(t)$ are defined in the section just before Chapter One.

Here in this section, after solving the balance equations we solve a differential equation for $V(z)$ and $P(z)$, the generating function of the queue size when the server is on vacation and in busy state respectively. This enables us to obtain the proportion of time the server is under vacation state or in busy state of the queueing system. Let us now carry out the analytic analysis of the model and obtain the queue size at a stationary point of time.

### 4.2.2 Steady State Solution of the Queue size at Random Epoch

The set of differential equations describing the model are

$$\frac{d}{dt}P_n(t) = -(\lambda+\mu)P_n(t) + \lambda\sum_{i=1}^{n} c_i P_{n-i}(t) + (1-p)\mu P_{n+1}(t) + \phi V_n(t) \tag{4.1}$$

$$\frac{d}{dt}P_0(t) = -\lambda P_0(t) + (1-p)\mu P_1(t) + \phi V_0(t) \tag{4.2}$$

$$\frac{d}{dt}V_n(t) = -(\lambda+\phi+n\gamma)V_n(t) + \lambda\sum_{i=1}^{n} c_i V_{n-i}(t) + (n+1)\gamma V_{n+1}(t) + p\mu P_{n+1}(t)\,,\ n\geq 1 \tag{4.3}$$

$$\frac{d}{dt}V_0(t) = -(\lambda+\phi)V_0(t) + \gamma V_1(t) + p\mu P_1(t) \tag{4.4}$$

The transition- rate diagram depicting the state of the system is shown below, where N denotes the total number of customers in the system.

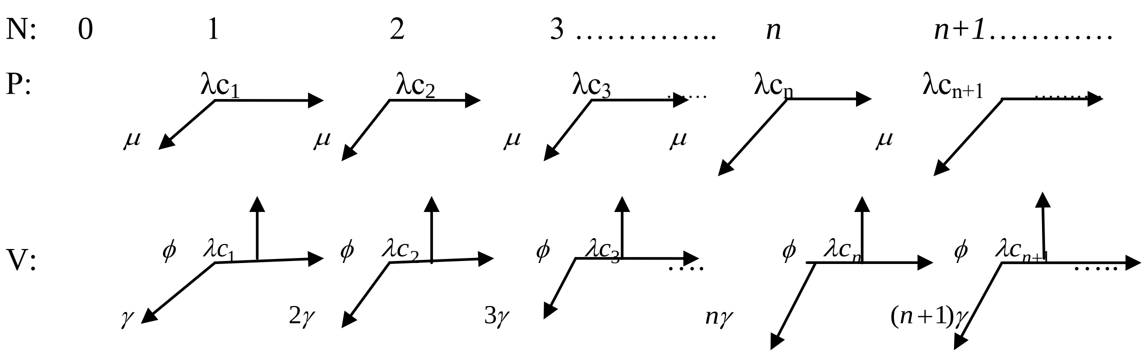


*Figure 4.1 Transition-rate diagram*

The steady state probabilities are defined as

$$\lim_{t\to\infty} P_n(t) = P_n,\quad \lim_{t\to\infty} V_n(t) = V_n$$

The stability condition for an ordinary queuing system is when the individual arrival rate is strictly less than the departure rate. However when reneging is allowed, there is no fixed departure rate because it changes with the length or size of the queue. If the queue length is large, departure rate will tend to increase which in turn will stabilize the queue. Therefore a queueing system with reneging always finds equilibrium because an instantaneously larger queue would mean a high departure rate. Thus in this case the necessary and sufficient condition for limits to exist is that the corresponding equations always have a solution.

Since we are considering the steady state of the system, $\lim_{t\to\infty} \frac{d}{dt} P_n(t) = 0,\ \lim_{t\to\infty} \frac{d}{dt} V_n(t) = 0,\ \ n \ge 0$

Thus, the differential-difference equations (4.1)-(4.4) under steady state, reduces to

$$(\lambda + \mu)P_n = \lambda \sum_{i=1}^{n} c_i P_{n-i} + (1-p)\mu P_{n+1} + \phi V_n \quad , n \ge 1, \tag{4.5}$$

$$\lambda P_0 = (1-p)\mu P_1 + \phi V_0 \ , \tag{4.6}$$

$$(\lambda + \phi + n\gamma)V_n = \lambda \sum_{i=1}^{n} c_i V_{n-i} + (n+1)\gamma V_{n+1} + p\mu P_{n+1} \ , n \ge 1, \tag{4.7}$$

$$(\lambda + \phi)V_0 = \gamma V_1 + p\mu P_1 \tag{4.8}$$

Defining the probability generating functions as

$$P(z) = \sum_{n=0}^{\infty} P_n z^n;\ \ V(z) = \sum_{n=0}^{\infty} V_n z^n;\ \ C(z) = \sum_{i=1}^{\infty} c_i z^i; \quad |z| \le 1 \tag{4.9}$$

Multiplying (4.5) by $z^n$, adding with (4.6) and using the probability generating functions defined in (4.9), we get

$$\left[\lambda - \lambda C(z) + \mu - (1-p)\frac{\mu}{z}\right] P(z) = \left[\mu - (1-p)\frac{\mu}{z}\right] P_0 + \phi V(z) \tag{4.10}$$

$$P(z) = \frac{p\mu P_0 - \mu(1-z)P_0 + \phi z V(z)}{z\lambda(1-C(z)) - \mu(1-z) + p\mu} \tag{4.11}$$

It follows from the definition that the series which defines (4.11) must converge everywhere in the unit circle $|z| \le 1$. Thus in this region zeroes of both numerator and denominator on the right-hand side of (4.11) must coincide. Denoting a root of the denominator in the interval (0, 1) as $z_0$, we have from equation (4.11),

$$p\mu P_0 - \mu(1-z_0)P_0 + \phi z_0 V(z_0) = 0, \qquad z_0\lambda(1-C(z_0)) - \mu(1-z_0) + p\mu = 0 \tag{4.12}$$

We now investigate the existence of any such real root, $z_0$ in the interval [0, 1] and shall further verify that this root is the only unique solution in the unit circle, $|z| \le 1$, such that the denominator is zero.

Let us denote $g(z) = z\lambda(1 - C(z)) - \mu(1 - z) + p\mu$ (4.13)

Again,

$g(0) = \mu(p - 1) < 0$

$g(1) = p\mu > 0$

Since $g$ is continuous and changes sign on the interval [0, 1] implies that there exists a real root $z_0 \in (0,1)$ such that $g(z_0) = 0$. To verify that $z_0$ is the unique solution of $g(z)$, let us apply Rouche's theorem.

Let us take $\xi(z) = (\lambda + \mu)z - \mu + p\mu$ and $\eta(z) = -\lambda z C(z)$

Now on the unit circle (i.e., on the circle where $|z|=1$), since $\lambda, \mu > 0$, $0 \le p \le 1$, and

$|C(z)| \le \sum_{i=}^{\infty} c_i |z^i| \le 1$, one has

$|\eta(z)| = |z\lambda C(z)| = \lambda |zC(z)| \le \lambda$, and $|\xi(z)| = |z(\lambda + \mu) - (1 - p)\mu| = \left|\lambda + \mu\left(1 - \frac{1-p}{z}\right)\right| \ge \lambda$,

because $|1/z|=1$ on the unit circle, and $1 > (1 - p)$, as such

$$|\xi(z)| \ge |\eta(z)| \tag{4.14}$$

So by Rouche's theorem, since $\xi(z)$ has only one zero inside the unit circle, it can be seen from (4.14), the function $g(z) = \xi(z) + \eta(z)$ has the same number of zeros inside the unit circle. Therefore $g(z)$ has only one zero in this region. Hence, there has to be just one solution of the function $g(z)$ in the interval [0, 1]. This confirms the existence of an unique root in (0, 1).

(For Rouche's theorem, reference may be made to any standard book on complex-variable theory)

Again, multiplying equation (4.7) by $z^n$ and summing over $n$ from 1 to $\infty$, using (4.8) and (4.9), we have

$$\gamma(1 - z)V'(z) - (\lambda - \lambda C(z) + \phi)V(z) = p\frac{\mu}{z}P_0 - p\frac{\mu}{z}P(z), \tag{4.15}$$

where, $\frac{d}{dz}V(z) = V'(z)$

Using (4.11) in equation (4.15) we get

$$\gamma(1 - z)V'(z) - (\lambda - \lambda C(z) + \phi)V(z) = \frac{p\mu\lambda(1 - C(z))P_0 - p\mu\phi V(z)}{z\lambda(1 - C(z)) - \mu(1 - z) + p\mu}$$

$$\gamma(1-z)V'(z)-\left[\lambda\big(1-C(z)\big)+\phi-\frac{p\mu\phi}{z\lambda\big((1-C(z)\big)-\mu(1-z)+\pi\mu}\right]V(z) = \frac{p\mu\lambda\big(1-C(z)\big)P_0}{z\lambda(1-C(z))-\mu(1-z)+p\mu} \tag{4.16}$$

Denoting, $g(z)=z\lambda\big(1-C(z)\big)-\mu(1-z)+p\mu,\ \ \frac{\lambda}{\gamma}=a.,\ \frac{\phi}{\gamma}=b$ , we re-write (4.16) as

$$V'(z)-\left[\frac{a(1-C(z)}{1-z}+\frac{b}{1-z}-\frac{p\mu b}{(1-z)g(z)}\right]V(z)=\frac{ap\mu\big(1-C(z)\big)P_0}{(1-z)g(z)} \tag{4.17}$$

Equation (4.17) is valid everywhere except at an isolated set of singular points. Therefore the equation needs to be solved separately in the interval $(0,z_0)$ and $(z_0,1]$. However, in order to calculate the queue characteristics, we are interested for the probability generating function derivatives at $z=1$, so we can restrict our solution to the interval $(z_0,1]$.

In order to solve the first order linear differential equation (4.17), an integrating factor $I$ be found as

$$I=e^{-\int\left(\frac{a(1-C(z)}{1-z}+\frac{b}{1-z}-\frac{p\mu b}{(1-z)g(z)}\right)dz}=e^{a\int_z^1\frac{1-C(t)}{1-t}dt}\,e^{b\int_z^1\left(1-\frac{p\mu}{g(y)}\right)\frac{dy}{1-y}}=e^{aJ(z)+b\delta(z)}\quad , \tag{4.18}$$

where we denote $\int_z^1\frac{1-C(t)}{1-t}=J(z),\int_z^1\left(1-\frac{p\mu}{g(y)}\right)\frac{dy}{1-y}=\delta(z)$, z > $z_0$ in the interval [0,1].

We observe that both functions $J(z)$ and $\delta(z)$ are well defined in the interval $(z_0,1)$. Further since $g(1)=p\mu$, so the combination $\left[\frac{1}{1-y}-\frac{p\mu}{g(y)(1-y)}\right]$ is finite as $y\to 1$, thus the integral $\delta(z)$ is convergent as $y\to 1$.

Multiplying (4.14) by the integrating factor $I$ in (4.15), after some algebra, we have

$$\frac{d}{dz}\Big[e^{aJ(z)+b\delta(z)}V(z)\Big]=\frac{ap\mu e^{aJ(z)+b\delta(z)}\big(1-C(z)\big)P_0}{(1-z)g(z)} \tag{4.19}$$

Integrating equation (4.15) between limits $z$ and 1, we get

$$V(1) - e^{aJ(z)+b\delta(z)} V(z) = ap\mu P_0 \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))}{(1-t)g(t)} dt$$

$$V(1) = e^{aJ(z)+b\delta(z)} V(z) + ap\mu P_0 \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))}{(1-t)g(t)} dt \tag{4.20}$$

Re-writing the above equation, we get

$$V(z) = e^{-aJ(z)-b\delta(z)} \left[ V(1) - ap\mu P_0 \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))dt}{(1-t)g(t)} \right] \tag{4.21}$$

Since $V(z)$ is a probability generating function, so must be finite everywhere in the unit circle, hence the limit below should exist and be finite. Therefore,

$$V(z_0) = \lim_{z \to z_0} V(z) = \underset{z \to z_0}{Lim}\, e^{-aJ(z)-b\delta(z)} \left[ V(1) - ap\mu P_0 \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))dt}{(1-t)g(t)} \right] \tag{4.22}$$

Since $\underset{z \to z_0}{Lim}\, e^{-aJ(z)-\delta J(z)} = \infty$ and $V(z)$ being a power series in $z$ with the coefficients as probabilities, we must have

$$\underset{z \to z_0}{Lim} \left[ V(1) - ap\mu P_0 \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))dt}{(1-t)g(t)} \right] = 0\ . \tag{4.23}$$

The limit of equation (4.23) must be finite for any value of $z$ in the interval (0, 1), but our consideration is only for all $z > z_0$ .So, the limit of the term inside the square bracket is zero, for $z \to z_0$.

From equation (4.23), we have

$$V(z_0) = \underset{z \to z_0}{Lim} \frac{\left[ V(1) - ap\mu P_0 \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))}{(1-t)g(t)} dt \right]}{e^{aJ(z)+b\delta(z)}} \ ,$$ reduces to 0/0 form.

Therefore applying L' Hopital's rule and after simplification, we get

$$V(z_0)=\lim_{z\to z_0}\frac{\left[ap\mu P_0 e^{aJ(z)+b\delta(z)}\dfrac{1-C(z)}{(1-z)g(z)}\right]}{-e^{aJ(z)+b\delta(z)}\left[\dfrac{a(1-C(z))}{1-z}+\dfrac{b}{1-z}\left(1-\dfrac{p\mu}{g(z)}\right)\right]}=\frac{\lambda}{\phi}P_0(1-C(z_0)) \tag{4.24}$$

Thus it is seen that the $V(z_0)$ in equation (4.22) is finite. Hence

$$V(1)=ap\mu F(z_0)P_0\ , \tag{4.25}$$

where $F(z_0)=\lim_{z\to z_0}\int_z^1 e^{aJ(t)+b\delta(t)}\dfrac{(1-C(t))}{(1-t)g(t)}dt$ (4.26)

Equation (4.25) gives the probability that the server is on vacation irrespective of the number of customers in the queue.

Again, for $z\to z_0$, we shall examine that the integral in (4.26) exists.

We first investigate the function $J(t)=\int_t^1\dfrac{1-C(y)}{1-y}dy$

Since $\dfrac{1-C(y)}{1-y}$ is regular near $y=1$, so it can be expanded as

$$\frac{1-C(y)}{1-y}\approx\frac{1-\left[C(1)+C'(1)(y-1)+O\left((y-1)^2\right)\right]}{1-y}=\frac{C'(1)(1-y)+O\left((1-y)^2\right)}{1-y}=C'(1)+O(1-y)$$

Neglecting further higher terms $O(1-y)$, we have $\dfrac{1-C(y)}{1-y}\approx C'(1)$

Thus, $J(t)=C'(1)(1-t)$. Again, let us investigate $\delta(t)=\int_t^1\left(1-\dfrac{p\mu}{g(y)}\right)\dfrac{dy}{1-y}$

Expanding $g(z_0)\approx g(z_0)+g'(z_0)(y-z_0)+0\left((y-z_0)^2\right)=g'(z_0)(y-z_0)$, (ignoring higher powers), we analyze the integral over the region near $z_0$ and have

$$\delta(t)\approx\int_t\left(1-\frac{p\mu}{g'(z_0)(y-z_0)}\right)\frac{dy}{1-z_0}=-\frac{p\mu}{g'(z_0)(1-z_0)}\ln(y-z_0)\Big|_t\approx\ln(t-z_0)^B$$ , where

$$B=\frac{p\mu}{g'(z_0)(1-z_0)}$$

Thus $\lim_{z\to z_0}F(z)=\lim_{z\to z_0}\int_z^1 e^{aJ(t)+b\delta(t)}\dfrac{1-C(t)}{(1-t)g(t)}dt$ , exists and is finite.

Substituting $V(1)$ in equation (4.21), we get

$$V(z) = e^{-aJ(z)-b\delta(z)} ap\mu\left[F(z_0) - \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))dt}{(1-t)g(t)}\right]P_0 \tag{4.27}$$

Therefore once $P_0$ is calculated, $P(z)$ and $V(z)$ can be completely determined from equations (4.11) and (4.27).

In order to determine $P_0$, the probability of empty state, we use the normalizing condition:

$$P(1) + V(1) = 1$$

$$\therefore \frac{p\mu P_0 + \phi V(1)}{p\mu} + V(1) = 1$$

$$P_0 + \left(\frac{\phi}{p\mu} + 1\right)V(1) = 1 \tag{4.28}$$

Now substituting $V(1)$ from (4.25) in (4.28), we get

$$P_0 + \left(\frac{\phi}{p\mu} + 1\right)ap\mu F(z_0)P_0 = 1$$

$$P_0 = \frac{1}{1 + \left(\frac{\phi}{\mu} + p\right)a\mu F(z_0)} \tag{4.29}$$

Equation (4.29) gives the probability of empty state and

$\rho = 1 - \dfrac{1}{1 + \left(\frac{\phi}{\mu} + p\right)a\mu F(z_0)} < 1$ is the utilization factor.

Thus the probability generating function of a queue size at random point of tine for a $M^X/M/1$ queueing system with reneging during server vacations can be derived in closed-form, using the relation $P_q(z) = P(z) + V(z)$.

### 4.2.3 The Mean Queue Length

The main purpose of any study of a queueing system is to assess the performance of the system through the derivation of the performance measures, as these measures of the system specifically represent how a system is working in terms of its efficiency. The most important performance measures are the mean of the queue size and mean of waiting time which actually identifies customer satisfaction and how efficient the system is. Thus we intend to derive the mean length of queue size as discussed below.

Let us take $L_V$ = mean length of the queue during vacation state.

$$L_v = V'(1) = \sum_{n=1}^{\infty} nV_n$$

From equation (4.14), using L' Hopital's Rule, we have

$$L_v = V'(1) = \lim_{z\to 1} V'(z)$$

$$L_v = \left[\frac{a(1-C(z))}{1-z} + \frac{b}{1-z}\left(1-\frac{p\mu}{g(z)}\right)\right]V(z) + \frac{ap\mu(1-C(z))P_0}{(1-z)g(z)}$$

Rewriting the above equation as

$$L_v = \left[\frac{a(1-C(z))}{1-z} + \frac{b\{z\lambda(1-C(z))-\mu(1-z)\}}{(1-z)g(z)}\right]V(z) + \frac{ap\mu(1-C(z))P_0}{(1-z)g(z)} \tag{4.30}$$

Since equation (4.27) is indeterminate of 0/0 form at $z$ =1, using L'Hopital's rule and differentiating twice, we get

$$V'(1) = \left[\frac{-aC'(1)}{-1} + \frac{bg'(1)}{-g(1)}\right]V(1) + \frac{ap\mu(-C'(1))P_0}{-g(1)}$$

$$= \left[aC'(1) - \frac{b(-\lambda C'(1)+\mu)}{p\mu}\right]V(1) + \frac{ap\mu C'(1)P_0}{p\mu}$$

$$\therefore L_v = [p\mu aE(I) - b(\mu - \lambda E(I))]aF(z_0)P_0 + aE(I)P_0, \tag{4.31}$$

Let $L_k$ = mean queue size during the period of availability of the server, so that

$$L_k = \frac{d}{dz}P(z)\bigg|_{z=1}$$

Now from equation (4.11), we have

$$L_k = \frac{(\mu P_0 + \phi V(1) + \phi V'(1))p\mu - (p\mu P_0 + \phi V(1))(-\lambda E(I) + \mu)}{(p\mu)^2}$$

$$L_k = \frac{\phi p\mu V'(1) + \phi p\mu V(1) + p\mu\lambda E(I)P_0 + \phi V(1)\lambda E(I) - \phi\mu V(1)}{p^2\mu^2}$$

After simplifying, we have

$$L_k = \frac{\phi[p\mu a E(I) - b(\mu - \lambda E(I))]aF(z_0) + \lambda E(I)P_0 + \phi aF(z_0)\lambda E(I)}{p\mu} \qquad (4.32)$$

Therefore, the mean queue size irrespective of whether the server is on vacation or working in the system is given by $L_q = L_v + L_k$ can be calculated from equations (4.31) and (4.32).

The mean of waiting time in the queue during vacation can be similarly obtained by using Little's law as $W_q = \frac{L_q}{\lambda E(I)}$

## 4.3 Analysis of Model 2: Non- Markovian Case- General Service Time

### 4.3.1 Assumptions Underlying the Mathematical Model:

To formulate the non-Markovian model, we assume that the service time $S$ follows a general probability distribution with distribution function $G(s)$, $g(s)$ being the probability density function and $E(S^n)$ is the $n$ th moment of the service time. Let $\mu(x)$ be the conditional probability of service time during the period $(x, x+dt]$ given elapsed time is $x$ such that $\mu(x) = \frac{g(x)}{1-G(x)}$ and

$$g(s) = \mu(s)\exp\left[-\int_0^s \mu(x)dx\right].$$

In addition, the assumptions of server vacation and reneging during vacations are as defined in Section 4.2.1. The different probabilities for service time, $P_n(x,t)$, vacation time $V_n(t)$ and idle state $Q(t)$ are as defined before the Chapter One.

### 4.3.2 Queue Size Distribution at a Random Epoch

We now focus on the computation of the equilibrium state of the system.

Thus assuming that the steady state exists, we define the following steady state probabilities:

$$\begin{aligned}&\lim_{t\to\infty} P_n(t,x)=P_n(x),\quad \lim_{t\to\infty} P_n(t)=\int_0^\infty P_n(t,x)dx=P_n\\&\lim_{t\to\infty} V_n(t)=V_n,\qquad \lim_{t\to\infty} Q(t)=Q\end{aligned} \tag{4.33}$$

Based on the assumptions discussed in Section 4.3.1 and 4.2.1, the differential equations (4.5) and (4.6) changes to

$$\frac{d}{dx}P_n(x)+\{\lambda+\mu(x)\}P_n(x)=\lambda\sum_{i=1}^{n}c_iP_{n-i}(x);\quad n\geq 1 \tag{4.34}$$

$$\frac{d}{dx}P_0(x)+\{\lambda+\mu(x)\}P_0(x)=0 \tag{4.35}$$

In addition, we introduce the following differential equations to account for reneging

$$\{\lambda+\phi+n\gamma\}V_n=\lambda\sum_{i=1}^{n}c_iV_{n-i}+(n+1)\gamma V_{n+1}+p\int_0^\infty P_n(x)\mu(x)dx,\ n\geq 1 \tag{4.36}$$

$$(\lambda+\phi)V_0=\gamma V_1+p\int_0^\infty P_0(x)\mu(x)dx \tag{4.37}$$

$$\lambda Q=(1-p)\int_0^\infty P_0(x)\mu(x)dx+\phi V_0 \tag{4.38}$$

The above differential equations now have to be solved subject to the following boundary conditions:

$$P_n(0)=(1-p)\int_0^\infty P_{n+1}(x)\mu(x)dx+\phi V_{n+1}+\lambda c_{n+1}Q\ ,\quad n\geq 0 \tag{4.39}$$

In the usual process of re-structuring differential equations, (4.34) and (4.35) yields

$$\frac{d}{dx}P(x,z)+\{(\lambda-\lambda C(z))+\mu(x)\}P(x,z)=0 \tag{4.40}$$

Similarly from (4.32) and (4.33), we get

$$(\lambda-\lambda C(z)+\phi))V(z)-\gamma(1-z)\frac{d}{dz}V(z)=p\int_0^\infty P(x,z)\mu(x)dx \tag{4.41}$$

Further integrating equations (4.40) over limits 0 to *x* gives us the following

$$P(x,z) = P(0,z)\exp[-(\lambda - \lambda C(z))x - \int_0^x \mu(t)dt\,] \tag{4.42}$$

Next multiplying the boundary condition (4.39) by $z^{n+1}$ and taking summation over all possible values of *n* and using the probability generating functions, we have after simplification

$$zP(0,z) = (\lambda C(z) - \lambda)Q + (1-p)\int_0^\infty P(x,z)\mu(x)dx + \phi V(z) \tag{4.43}$$

Again integrating equation (4.42) with respect to *x*, gives us

$$P(z) = P(0,z)\left[\frac{1 - G^*(\lambda - \lambda C(z))}{\lambda - \lambda C(z)}\right] \tag{4.44}$$

where $G^*(\lambda - \lambda C(z))$ is the Laplace-Stieltjes transform of service time as defined before.

Multiplying both sides of (4.42) by $\mu(x)$ and integrating over 0 to ∞, we get

$$\int_0^\infty P(x,z)\mu(x)dx = P(0,z)G^*(\lambda - \lambda C(z)) \tag{4.45}$$

From equation (4.43), using (4.45) and after simplification, we have

$$P(0,z) = \frac{(\lambda C(z) - \lambda)Q + \phi V(z)}{z - (1-p)G^*(\lambda - \lambda C(z))} \tag{4.46}$$

It follows from the definition that $P(0,z)$ being a probability function, it must converge everywhere in the unit circle $|z| \le 1$. Thus in this region the zeroes of both the numerator and denominator in the equation (4.42) must coincide. Let us denote the denominator,

$$f(z) = z - (1-p)G^*(\lambda - \lambda C(z)) \tag{4.47}$$

Since $f(0) = -(1-p)G^*(\lambda) < 0$ and $f(1) = p > 0$, there exists one real root in the unit circle, $|z| \le 1$. Further, the existence of unique real root in the interval (0, 1) can be verified in analogy with the same theory related to Model 1, except that the service time is defined by the Laplace-Stieltjes transform. (The proof of it is provided in Appendix B. 14). Similarly, we consider $z = z_0$ as the root of the denominator in the interval (0, 1).

Now since the right-hand side of equation (4.46) is indeterminate of the form zero/zero at $z = z_0$, using L' Hopital's rule, we have

$$P(0,z_0) = \lim_{z\to z_0} P(0,z) = \lim_{z\to z_0} \frac{\left(\lambda C(z)-\lambda\right)Q+\phi V(z)}{z-(1-p)G^*\left(\lambda-\lambda C(z)\right)}$$
$$= \frac{\lambda C'(z_0)Q+\phi V'(z_0)}{1+(1-p)(\lambda C'(z_0)G^{*'}(\lambda-\lambda C(z_0))} \tag{4.48}$$

Now from equation (4.41), using (4.46) and after some algebra, we have

$$V'(z)-\left[\frac{a\left(1-C(z)\right)}{(1-z)}+\frac{b}{(1-z)}\right]V(z) = \frac{-pP(0,z)G^*\left(\lambda-\lambda C(z)\right)}{\gamma(1-z)} \tag{4.49}$$

Substituting $P(0,z)$ from (4.46) in equation (4.48), we get after simplifying,

$$V'(z)-\left[\frac{a\left(1-C(z)\right)}{(1-z)}+\frac{b}{(1-z)}\left(1-\frac{pG^*(\lambda-\lambda C(z))}{f(z)}\right)\right]V(z) = \frac{pa(1-C(z))QG^*\left(\lambda-\lambda C(z)\right)}{(1-z)f(z)} \tag{4.50}$$

In order to solve the above equation, we take the integrating factor *I* to be

$$I = e^{\int_z^1 \frac{a\left(1-C(t)\right)}{1-t}+\frac{b}{1-t}\left(1-\frac{pG^*(\lambda-\lambda C(t))}{f(t)}\right)dt} = e^{aJ(z)+b\delta(z)}, \tag{4.51}$$

by taking $\int_z^1 \frac{(1-C(t))}{1-t}dt = J(z)$, $\int_z^1\left(1-\frac{pG^*(\lambda-\lambda C(t))}{f(t)}\right)\frac{dt}{1-t} = \delta(z)$, z >$z_0$ in the interval [0, 1]

Multiplying (4.49) with the integrating factor *I* defined in (4.46) and integrating between limits z and 1, we have

$$V(1)-e^{aJ(z)+b\delta(z)}V(z) = paQ\int_z^1 e^{aJ(t)+b\delta(t)}\,\frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)}dt$$

Or

$$V(z) = e^{-aJ(z)-b\delta(z)}\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\,\frac{\left(1-C(t)\right)G^*(\lambda-\lambda C(t))}{(1-t)f(t)}\,dt\right] \tag{4.52}$$

Now

$$V(z_0) = \lim_{z\to z_0} e^{-aJ(z)-b\delta(z)}\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\,\frac{\left(1-C(t)\right)G^*(\lambda-\lambda C(t))dt}{(1-t)f(t)}\right] \tag{4.53}$$

$\lim_{z\to z_0} e^{-aJ(z)-b\delta(z)} = \infty$ and since $V(z)$ is a power series in *z* with the coefficients as probabilities, so equation (4.53) must be finite for any value of *z* in the interval (0, 1). Then

$$\lim_{z\to z_0}\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))dt}{(1-t)f(t)}\right]$$ must converge and be equal to zero. Thus

$$V(1)-apQK(z_0)=0\,, \tag{4.54}$$

where $$K(z_0)=\int_{z_0}^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))dt}{(1-t)f(t)}\,, \tag{4.55}$$

provided (4.55) exists.

The existence of the limit in (4.55) can be similarly shown as in Section 4.2.2. Since

$\lim_{z\to z_0}\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)}dt$ , is finite, we examine the convergence of the expression inside the RHS of equation (4.52).

Then

$$V(z_0)=\lim_{z\to z_0}\frac{\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)}dt\right]}{e^{aJ(z)+b\delta(z)}}$$

The above equation is indeterminate of zero/zero form at $z = z_0$. Therefore applying L'Hopital's rule, we have

$$V(z_0)=\lim_{z\to z_0}\frac{apQe^{aJ(z)+b\delta(z)}\frac{(1-C(z))G^*(\lambda-\lambda C(z))}{(1-z)f(z)}}{-e^{aJ(z)+b\delta(z)}\left[\frac{a(1-C(z))}{1-z}+\frac{b}{1-z}\left(1-\frac{pG^*(\lambda-\lambda C(z))}{f(z)}\right)\right]}$$

$$=\lim_{z\to z_0}\frac{apQ(1-C(z))G^*(\lambda-\lambda C(z))}{-\left[a(1-C(z))f(z)+b\left(z-G^*(\lambda-\lambda C(z))\right)\right]}$$

$$\therefore V(z_0)=\frac{apQ(1-C(z_0))G^*(\lambda-\lambda C(z_0))}{b\left[G^*(\lambda-\lambda C(z_0))-z_0\right]}=\frac{\lambda Q(1-C(z_0))}{\phi} \tag{4.56}$$

Equation (4.55) is finite and therefore the argument (4.54) is true.

Thus, $V(1)=apQK(z_0)\,,$ (4.57)

where $K(z_0)$ is defined in (4.55). So the probability of the system in vacation state is derived using (4.57) in (4.52) as

$$V(z) = e^{-aJ(z)-b\delta(z)} apQ\left[K(z_0) - \int_z^1 e^{aJ(t)+b\delta(t)} \frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)} dt\right] \quad (4.58)$$

Also using (4.46) in (4.44), we get the probability generating function of the system in working state as

$$P(z) = \frac{(\lambda C(z)-\lambda)Q + \phi V(z)}{z-(1-p)G^*(\lambda-\lambda C(z))}\left[\frac{1-G^*(\lambda-\lambda C(z))}{\lambda-\lambda C(z)}\right] \quad (4.59)$$

In order to determine $Q$, the probability of empty state, we use the normalizing condition;

$$P_q(1) = P(1) + V(1) + Q = 1 \quad (4.60)$$

Since $P(z)$ defined in equation (4.59) is indeterminate of the 0/0 form at z =1, we apply L' Hopital's rule and obtain

$$P(1) = \frac{\phi V(1) G^{*/}(1)}{p} = \frac{\phi V(1)E(S)}{p} \quad (4.61)$$

where $E(S) = -G^{*/}(1)$ is the mean of service time.

Again utilizing equations (4.57) and (4.61) in the normalization condition, we obtain

$$\frac{\phi V(1)E(S)}{p} + apQK(z_0) + Q = 1$$

$$\frac{\phi apQK(z_0)E(S)}{p} + apQK(z_0) + Q = 1$$

$$aQK(z_0)\phi E(S) + apK(z_0)Q + Q = 1$$

$$Q = \frac{1}{aK(z_0)[p+\phi E(S)]+1} \quad (4.62)$$

Equation (4.62) gives the probability of the state when the server is idle.

Therefore the probability generating function for the number of customers in the queue at stationary point of time is derived by using the relation $P_q(z) = P(z) + V(z)$.

### 4.3.3 The Mean Queue Length

In order to facilitate analysis in this section we show some performance measures of the system like mean and variance of the queue size.

Under similar notations defined earlier, we derive the mean queue length when the server is on vacation and server is available. Thus we have

$$L_v = V'(1) = \sum_{n=1}^{\infty} nV_n$$

We use equation (4.49) and since the equation is of 0/0 form, applying L' Hopital's Rule, have

$$V'(1) = \left[ aE(I) - \frac{b(1-\lambda E(I)E(S))}{p} \right] V(1) + \frac{\lambda E(I))Q}{\gamma}$$

$$= \left[ aE(I) - \frac{b(1-\lambda E(I)E(S))}{p} \right] apQK(z_0) + aE(I)Q$$

$$L_v = \left[apE(I) - b(1-\lambda E(I)E(S))\right] aQK(z_0) + aE(I)Q \tag{4.63}$$

and $E(I)$ is the expected size of the batch of '$i$' arrivals.

Now $P(z) = \frac{N(z)}{D(z)}$ in (4.59) is of the form 0/0 at $z$ =1, so using L' Hopital's rule twice,

We get

$$L_k = \lim_{z\to 1} \frac{N''(z)D'(z) - N'(z)D''(z)}{2(D'(z))^2} \tag{4.64}$$

The primes and double primes denote the first and second derivative of equation (4.59) at z =1. Therefore, the steady state mean queue size irrespective of whether the server is on vacation or available in the system is given by $L_q = L_v + L_k$ can be calculated from equations (4.63) and (4.64) and due to the complex and long expressions of the equations, the results are not shown explicitly.

The variance of the queue size can be obtained by the relation

$$P_q''(1) + P_q'(1) - \left[P_q'(1)\right]^2 \tag{4.65}$$

Now, $P_q''(1) = P''(1) + V''(1)$ (4.66)

Therefore,

$$P''(1)=\frac{d^2}{dz^2}P(z)\Bigg|_{z=1}$$

$$=\quad\frac{2\lambda E(I)E(S)\left[\lambda E(I)Q+\phi V'(1)\right]+\phi V(1)\left[\lambda E(I(I-1))E(S)-(\lambda E(I))^2E(S^2)\right]}{2\lambda E(I)\left[1-(1-p)\lambda E(I)E(S)\right]+p\lambda E\left(I(I-1)\right)},\qquad(4.67)$$

where $E(I(I-1))$ is the second factorial moment of batch of arriving customers and $E(S^2)$ is the second moment of service time.

$$V''(1)=\frac{\begin{bmatrix}apE(I(I-1))+aE(I)\left(1-(1-p)\lambda E(I)E(S)\right)\\+b\left\{\lambda E(I(I-1))+\left(\lambda E(I)\right)^2E(S^2)\right\}\end{bmatrix}V(1)+2\left[apE(I)-b\left\{1-\lambda E(I)E(S)\right\}\right]V'(1)+apQ\left[E(I(I-1))+2E(I)\lambda E(I)E(S)\right]}{2\left[1-(1-p)\lambda E(I)E(S)\right]}\qquad(4.68)$$

Adding (4.67) and (4.68), we get

$$P_q''(1)=\frac{2\lambda E(I)E(S)\left[\lambda E(I)Q+\phi V'(1)\right]+\lambda E(I(I-1))\phi V(1)+\left(\lambda E(I)\right)^2E(S^2)}{2\lambda E(I)\left[1-(1-p)\lambda E(I)E(S)\right]+p\lambda E(I(I-1))}$$

$$+\frac{\begin{bmatrix}a\pi E(I(I-1))+aE(I)\left(1-(1-\pi)\lambda E(I)E(S)\right)\\+b\left\{\lambda E(I(I-1))+\left(\lambda E(I)\right)^2E(S^2)\right\}\end{bmatrix}V(1)+2\left[apE(I)-b\left\{1-\lambda E(I)E(S)\right\}\right]V'(1)+apQ\left[E(I(I-1))+2E(I)\lambda E(I)E(S)\right]}{2\left[1-(1-p)\lambda E(I)E(S)\right]}\qquad(4.69)$$

Utilizing the above expressions in equation (4.65), the variance of the queue size can be derived explicitly.

**Remark**: Checking the Markovian limit:

In order to examine the consistency of the above results, let us compare it with Model 1 of Markovian service time, then

$$E(S)=\frac{1}{\mu},\ K(z_0)=\underset{z\to z_0}{Lim}\,\mu\int_z^1 e^{aJ(t)+b\delta(t)}\frac{\left(1-C(t)\right)}{(1-t)f(t)}dt\ ,\quad f(z)=z-\frac{(1-p)\mu}{\mu+\lambda-\lambda C(z)}$$

then the probability of Q in equation (4.61) reduces to

$Q=\dfrac{1}{aK(z_0)\left[p\mu+\phi\right]+1}$ , which is analogous to the probability of empty state in a $M^X/M/1$ queue with reneging during vacations.

Similarly, $L_v = \left[apE(I) - b\left(1 - \frac{\lambda E(I)}{\mu}\right)\right] a\mu QK(z_0) + aE(I)Q$, is the mean queue length of the system when the server is on vacation.

Thus, $L_v = [ap\mu E(I) - b(\mu - \lambda E(I))]aQK(z_0) + aE(I)Q$, this tallies with the result obtained in the Markovaian case. This confirms the consistency of our results obtained for Model 2.

### 4.3.4 Special Case: A $Geo/\gamma/1$ Queue with Batch Arrivals and Reneging during Server Vacations

Let us investigate this model where batch arrivals are assumed to follow a Geometric distribution, mostly considered as an appropriate distribution due to its convenience for modeling purposes. The lack of memory feature in the distribution makes it more for use in stochastic modeling.

Thus we have that a batch of size 'í' arrives with probability

$$c_i = (1-v)^{i-1} v, \;\; i = 1,2,3....; \; u = 1-v$$

The probability generating function is $C(z) = \sum_{i=1}^{\infty} c_i z^i = \frac{vz}{1 - z(1-v)}$, and $C(1) = 1$

$$C'(z) = \frac{v}{[1-(1-v)z]^2} \text{ and } C'(1) = E(I) = \frac{1}{v} \tag{4.70}$$

$$C''(1) = E(I(I-1)) = \frac{2(1-v)}{v^2}$$

Let us assume that the service time $S$ follows a gamma distribution with $k$=2. Then

$$g(s) = \frac{\mu^2 e^{-\mu x} x}{\Gamma(2)} = \mu^2 e^{-\mu x} x \tag{4.71}$$

It is convenient in analysing this model assuming the service time following a 2- parameter Gamma distribution as a service could be considered as sum of the time-interval between 2 consecutive services. Variants of this distribution are widely used in the field of traffic congestion.

We shall evaluate the different probability functions $P(z)$ and $V(z)$ by considering the above assumptions. Thus let us first obtain the different functions w.r.t the above assumptions

$$G^*(\lambda-\lambda C(z))=\left[\frac{\mu}{\lambda-\lambda C(z)+\mu}\right]^2 \tag{4.72}$$

$$E(S)=-\left[\frac{d}{dz}G^*(\lambda-\lambda C(z))\right]_{z=1}=\frac{2\lambda}{v\mu},\ E(S^2)=\frac{6\lambda^2}{v^2\mu^2}+\frac{4\lambda q}{v^2\mu^2} \tag{4.73}$$

Now, from equation (4.57), we have

$$V(z)=e^{-aJ(z)-b\delta(z)}\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\,\frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)}dt\right], \tag{4.74}$$

where, $J(z)=\int_z^1\frac{1-C(t)}{1-t}dt$, and $f(z)=z-(1-p)\frac{\mu^2}{(\lambda-\lambda C(z)+\mu)^2}$

We evaluate the integral $J(z)=\int_z^1\frac{1-C(t)}{1-t}dt=\int_z^1\frac{dt}{1-(1-v)t}=\ln\left(\frac{1-(1-v)z}{v}\right),$ (4.75)

and $\delta(z)=\int_z^1\left(1-\frac{p\mu^2}{(\lambda-\lambda C(t)+\mu)^2 f(t)}\right)\frac{dt}{1-t}\quad=\int_z^1\frac{dt}{1-t}-\int_z^1\left[\frac{p\mu^2}{t(\lambda-\lambda C(t)+\mu)^2-(1-p)\mu^2}\right]\frac{dt}{1-t}$

Using the method of partial fractions, we have

$$\delta(z)=-\int_z^1\frac{dt}{t\left[\frac{\lambda}{\mu}(1-t)+(1-ut)\right]^2-(1-p)(1-ut)^2} \tag{4.76}$$

Thus, factorizing the denominator in (4.76) and considering $z_0$, $z_1$ and $z_2$ as the roots of the cubic equation in the denominator of (4.76), we have

$$\delta(z)=-\int_z^1\left[\frac{dt}{(t-z_0)(t-z_1)(t-z_2)}\right]=-\ln\left[\left(\frac{1-z_0}{z-z_0}\right)^{m_1}\left(\frac{1-z_1}{z-z_1}\right)^{m_2}\left(\frac{1-z_2}{z-z_2}\right)^{m_3}\right],$$

where $z>z_0$ and denoting

$$m_1=\frac{1}{(z_0-z_1)(z_0-z_2)},\ m_2=\frac{1}{(z_1-z_0)(z_1-z_2)},m_3=\frac{1}{(z_2-z_0)(z_2-z_1)}$$

$$e^{aJ(z)+b\delta(z)}=e^{a\ln\left(\frac{1-qz}{p}\right)-b\ln\left[\left(\frac{1-z_0}{z-z_0}\right)^{m_1}\left(\frac{1-z_1}{z-z_1}\right)^{m_2}\left(\frac{1-z_2}{z-z_2}\right)^{m_2}\right]} \tag{4.77}$$

Now, $\underset{z\to z_0}{Lim}V(z)=V(z_0)$

$$=\underset{z\to z_0}{Lim}\,e^{-aJ(z)-b\delta(z)}\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))dt}{(1-t)f(t)}\right].$$

$$\underset{z\to z_0}{Lim}\,e^{-aJ(z)-\delta J(z)}=\infty$$

Since $V(z)$ is a power series in $z$ with the coefficients as probabilities, so $V(z)$ must be finite for any value of $z$ in the interval (0, 1).

Therefore,

$$\underset{z\to z_0}{Lim}\left[V(1)-apQ\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))dt}{(1-t)f(t)}\right] \tag{4.78}$$

must be finite and equal to zero. So we have

$$\begin{aligned}&\underset{z\to z_0}{Lim}\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)}dt\\&=\underset{z\to z_0}{Lim}\int_z^1 e^{aJ(t)+b\delta(t)}\frac{\mu^2(1-ut)dt}{\left[t(\lambda-\lambda t+\mu(1-ut))^2-(1-p)\mu^2(1-ut)^2\right]}=K(z_0)\end{aligned} \tag{4.79}$$

Now $K(z_0)$ can be simplified further using (4.77) as

$$\underset{z\to z_0}{Lim}\int_z^1\left(\frac{1-ut}{v}\right)^a\left[\left(\frac{t-z_0}{1-z_0}\right)^{m_1}\left(\frac{t-z_1}{1-z_1}\right)^{m_2}\left(\frac{t-z_2}{1-z_2}\right)^{m_3}\right]^b\frac{p(1-ut)dt}{\left[t(\rho(1-t)+(1-ut))^2-(1-p)(1-ut)^2\right]}$$

$$=\frac{p}{(1-z_0)^{d_1}(1-z_1)^{d_2}(1-z_2)^{d_3}v^a}\underset{z\to z_0}{Lim}\int_z^1\frac{(t-z_0)^{d_1}(t-z_1)^{d_2}(t-z_2)^{d_3}(1-ut)^{a+1}dt}{(t-z_0)(t-z_1)(t-z_2)} \tag{4.80}$$

$$=A\int_{z_0}^1(t-z_0)^{d_1-1}(t-z_1)^{d_2-1}(t-z_2)^{d_3-1}(1-ut)^{a+1},$$

where $d_1=m_1b,\ d_2=m_2b,\ d_3=m_3b, A=\dfrac{p}{(1-z_0)^{d_1}(1-z_1)^{d_2}(1-z_2)^{d_3}v^a}$

It turns out that that the integral (4.78) is finite at $z = z_0$ , we have

$[V(1) - apQK(z_0)] = 0$ , so that $V(1) = apQK(z_0)$ .

The normalizing condition can be used to derive the probability of empty state Q.

Further, equation (4.47) reduces to

$$V'(z) = \left[\frac{a(1-C(z))}{(1-z)} + \frac{b}{(1-z)}\left(1 - \frac{p\mu^2}{[z(\lambda - \lambda C(z) + \mu)^2 - (1-p)\mu^2]}\right)\right]V(z) + \frac{pa(1-C(z))Q\mu^2}{(1-z)[z(\lambda - \lambda C(z) + \mu)^2 - (1-p)\mu^2]} \tag{4.81}$$

Since equation (4.81) is indeterminate of the 0/0 form at $z$ =1, we apply L' Hopitals's rule to evaluate the expected size of the queue during the time of server vacation. This is given by

$$L_v = V'(1) = \left[\frac{a}{v} - \frac{b}{p}\left(1 - \frac{2\lambda}{\mu v}\right)\right]V(1) + \frac{aQ}{v} \tag{4.82}$$

Again, $P(z) = \dfrac{(\lambda C(z) - \lambda)Q + \phi V(z)}{z - (1-p)G^*(\lambda - \lambda C(z))}\left[\dfrac{1 - G^*(\lambda - \lambda C(z))}{\lambda - \lambda C(z)}\right]$

Substituting the new assumptions we have

$$P(z) = \frac{(\lambda C(z) - \lambda)Q + \phi V(z)}{z(\lambda - \lambda C(z) + \mu)^2 - (1-p)\mu^2}\left[\frac{(\lambda - \lambda C(z) + \mu)^2 - \mu^2}{\lambda - \lambda C(z)}\right] \tag{4.83}$$

$$L_k = P'(1) = \frac{d}{dz}P(z)\bigg|_{z=1}$$

We see that equation (4.82) is of the form 0/0 at $z$ =1, so using L'Hopital's rule and differentiating twice we get,

$$L_k = \frac{2\lambda E(I)Q}{p\mu} + \frac{2\phi V'(1)}{p\mu} - \frac{\phi V(1)}{p\mu}\left[\frac{\lambda E(I)}{\mu} - \frac{2}{p}\right] - \frac{2\lambda E(I)\phi V(1)}{p^2\mu} \tag{4.84}$$

$$L_k = \frac{2\lambda Q}{pv\mu} + \frac{2\phi V'(1)}{p\mu} - \frac{\phi V(1)}{p\mu}\left[\frac{\lambda}{\mu v} - \frac{2}{p}\right] - \frac{2\lambda\phi V(1)}{vp^2\mu}$$

Therefore, the mean queue size irrespective of whether the server is on vacation or available in the system is given by $L_q = L_v + L_k$ can be calculated using (4.82) and (4.84), where $V(1) = apQK(z_0)$

And $Q$, the probability of empty state is given by

$$Q = \frac{1}{aK(z_0)\left[p + \frac{2\lambda\phi}{v\mu}\right] + 1} , \tag{4.85}$$

where $K(z_0)$ is derived in equation (4.80).

The variance of the queue size can be obtained by using the relation

$$P_q''(1) + P_q'(1) - \left[P_q'(1)\right]^2 \text{, where } P_q''(1) = P''(1) + V''(1) \tag{4.86}$$

$$P''(1) = \frac{d^2}{dz^2}P(z)\bigg|_{z=1} = \frac{2\left[\frac{\lambda Q}{v} + \phi V'(1)\right] - \phi V(1)\left[\frac{u}{v} - \frac{\lambda}{\mu v}\right]}{\left(\mu - \frac{2\lambda}{v} + \frac{p\mu u}{v}\right)} \tag{4.87}$$

And $V''(1) = \frac{d}{dz}V'(z)\Big|_{z=1}$

$$V''(1) = \frac{\left[ap\left\{\frac{2u}{v} + \frac{1}{v}\left(1 - \frac{2\lambda}{\mu v}\right)\right\} + \frac{4b\lambda(1+u)}{\mu v}\right]V(1) + 2V'(1)\left\{\frac{ap}{v} - b\left(1 - \frac{2\lambda}{\mu v}\right)\right\}}{2\left(1 - \frac{2\lambda}{\mu v}\right)} \tag{4.88}$$

Thus using (4.87) and (4.88), we get from $P_q''(1) = P''(1) + V''(1)$

$$P_q''(1) = \frac{2\left[\frac{\lambda Q}{v} + \phi V'(1)\right] - \phi V(1)\left[\frac{u}{v} - \frac{\lambda}{\mu v}\right]}{\left(\mu - \frac{2\lambda}{v} + \frac{p\mu v}{v}\right)} + \frac{\left[ap\left\{\frac{2u}{v} + \frac{1}{v}\left(1 - \frac{2\lambda}{\mu v}\right)\right\} + \frac{4b\lambda(1+u)}{\mu v}\right]V(1) + 2V'(1)\left\{\frac{ap}{v} - b\left(1 - \frac{2\lambda}{\mu v}\right)\right\}}{2\left(1 - \frac{2\lambda}{\mu v}\right)} \tag{4.89}$$

Now substituting the expressions from equation (4.84) and (4.88), we can derive the variance of the queue size. The mean of waiting time during vacation can be derived using Little's Law as $W_v = \frac{Lv}{\lambda E(I)}$.

In the same way the mean waiting time for the system can be derived as $W_q = \frac{L_q}{\lambda E(I)}$

## 4.4 Numerical Analysis

In order to bring out the qualitative aspects of the queueing system under consideration, we perform some numerical illustrations to shed further light on the effects of the various model parameters on the system's performance.

Let us investigate the effects of the expected queue size $L_V$ and the expected waiting time $W_v$ when the system is on vacation, as reneging is only during vacation period. For computational convenience, we arbitrarily choose the values of the different parameters; such that the utilization factor ρ <1 is always satisfied.

For this purpose, we use the special case outlined in section 4.3.4, where the batch arrivals follow a Geometric distribution and service time follows a 2-parameter gamma distribution. We construct the following tables.

Table 4.1: Computed values of various queue characteristics for a $Geom(x)/\gamma/1$ queue for varying values of reneging parameter $\gamma$ with $\lambda = 1.5, E(I) = 2.5, \mu = 6, p = 0.5, \phi = 0.4$

| $\gamma$ | $Q$ | $\rho$ | $L_V$ | $W_V$ |
|---|---|---|---|---|
| 0.6 | 0.96149 | 0.03851 | 6.13606 | 1.63628 |
| 0.8 | 0.97285 | 0.02715 | 4.62723 | 1.233428 |
| 1 | 0.97745 | 0.02255 | 3.75451 | 1.0012 |
| 1.5 | 0.98231 | 0.01769 | 2.47907 | 0.66109 |
| 2.0 | 0.98541 | 0.01459 | 1.86205 | 0.49655 |
| 2.5 | 0.98921 | 0.01079 | 1.49234 | 0.397957 |
| 3.0 | 0.989353 | 0.010647 | 1.24370 | 0.33165 |
| 3.5 | 0.991454 | 0.00546 | 1.06717 | 0.28457 |

To measure the effectiveness of the queueing system with the respect to the reneging behavior, we derive the main performance measures like proportion of idle time, traffic intensity, and mean length of queue both during server vacation and during service, shown in the above Table 4.1, by taking varying values of the reneging parameter $\gamma$, keeping the other parameters fixed.

The values obtained in table 4.1 are presented by a few plots demonstrated below. Figure (4.2) shows that as the reneging rate increases the expected queue length during vacation state tends to come down. Therefore, this also affects the overall length of queue size $L_q$ and it shows a decreasing trend as shown in Figure (4.2). This also indicates that a higher reneging rate would mean the server has lesser work load and hence fraction of idle state tends to increase. The values obtained are as expected. These trends are represented by some plots and graphs.

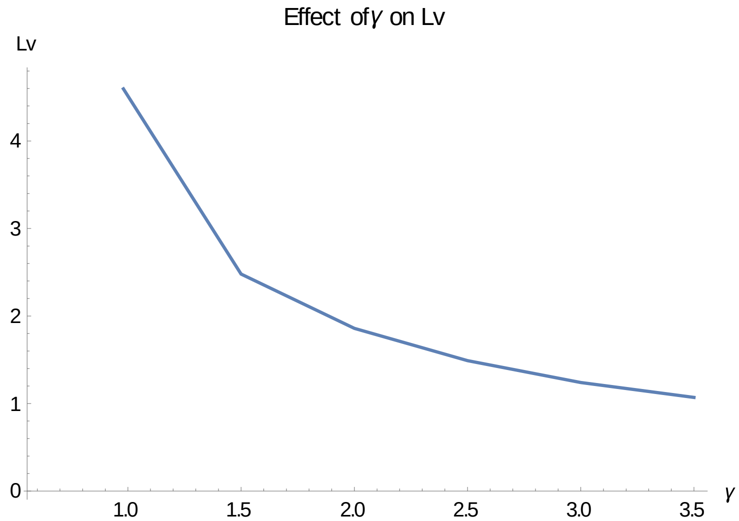


*Figure 4.2: Effect of Reneging parameter* $\gamma$ *on Lv*

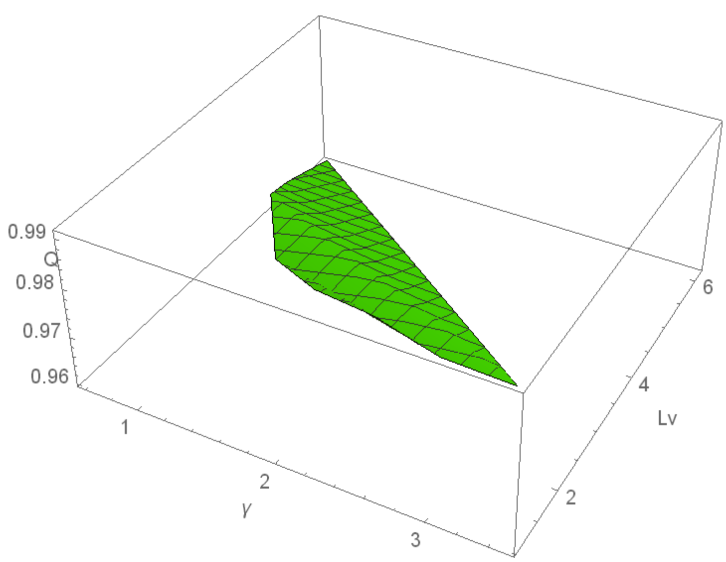


*Figure 4.3: Effect of* $\gamma$ *on* $Q$ *and* $L_v$

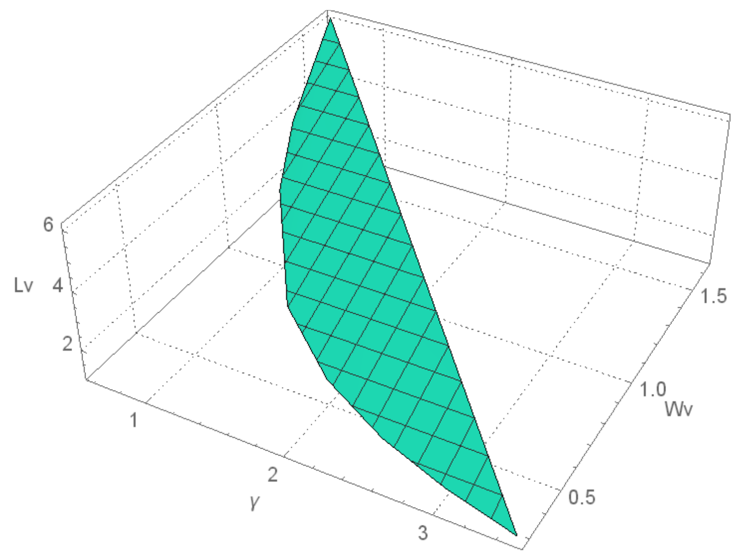

*Figure 4.4: Effect of $\gamma$ on $L_v$ and $W_v$*

It is seen that the waiting time $W_v$ during vacations will tend to decrease as the reneging parameter increases since customers leave the queue and the length of the queue decreases, which reduces the waiting time of the customers in the queue.

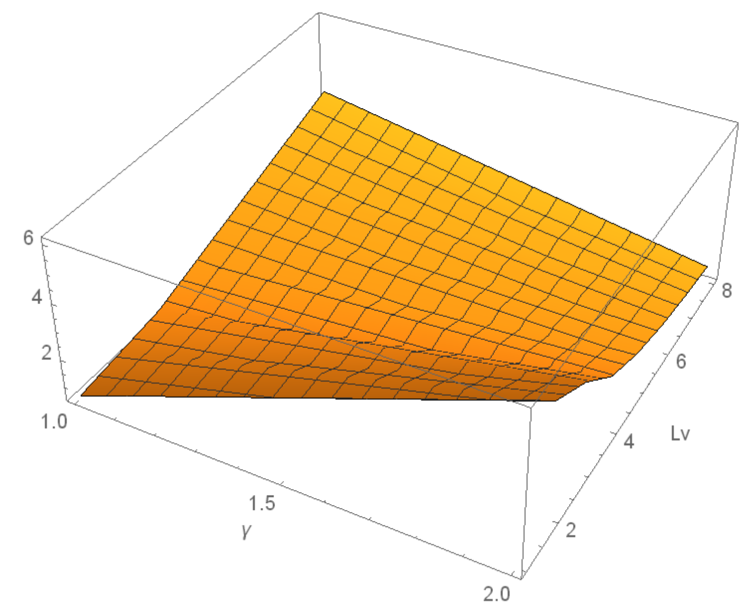

*Figure 4.5: Effect of reneging rate $\gamma$ on $L_V$*

Now, in the following Table 4.2, we fix the values of the parameters $\lambda, \mu, p, \gamma$ and $E(I)$ and vary the values of vacation parameter $\phi$, to monitor the effect of the vacation rate on the expected queue size. We also present a few plots demonstrating the behavior of queue size.

Table 4.2: Computed values of various queue characteristics for a $Geom(x)/\gamma/1$ queue for varying values of the vacation parameter $\phi$ with $\lambda = 3, \mu = 4, \gamma = 5, p = 0.5, E(I) = 2.5$

| $\phi$ | $Q$ | $\rho$ | $V(1)$ | $L_V$ | $L_k$ | $L_q$ |
|---|---|---|---|---|---|---|
| 1 | 0.959448 | 0.040552 | 0.00477 | 1.43808 | 8.5671 | 10.00518 |
| 1.5 | 0.967984 | 0.032016 | 0.00261 | 1.46021 | 9.39475 | 10.85496 |
| 2 | 0.975931 | 0.024069 | 0.001504 | 1.46946 | 10.21645 | 11.68591 |
| 2.5 | 0.98195 | 0.01805 | 0.00091 | 1.47681 | 11.02752 | 12.50433 |
| 3 | 0.98634 | 0.01366 | 0.00058 | 1.48229 | 11.8159 | 13.29819 |
| 3.5 | 0.98952 | 0.01048 | 0.00038 | 1.48631 | 12.6074 | 14.09371 |

Thus it is seen from Table 4.2 that as the parameter of vacation time increases, there is an increase in the expected queue size both during vacation and working period. It is observed that as rate of vacation time is increased, there tends to show an increase in the mean length of queue during vacation periods, which is as expected.  Since reneging takes place only during vacations and due to the increase in vacation rate the expected queue size during busy state is higher than that expected queue size in vacation state. This also increases the overall expected queue size of the system. Some plots representing the trend of the above values are displayed below.

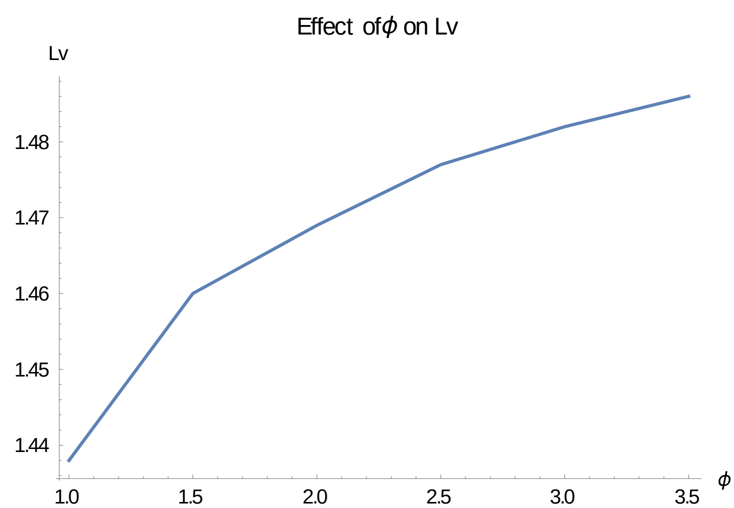


*Figure 4.6: Effect of $\phi$ on the expected queue length $L_v$*

It is seen that since the values of vacation time parameter $\phi$ are considered very small and the variation of the values of $\phi$ is also small, in comparison with the reneging parameter $\gamma = 5$, the mean length of the queue during vacation changes very little .

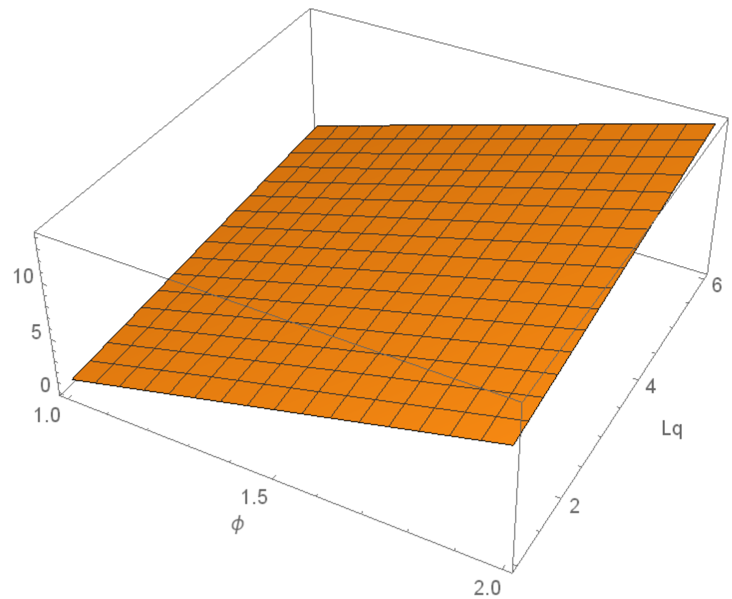


*Figure 4.7: Effect of $\phi$ on expected queue size Lq*

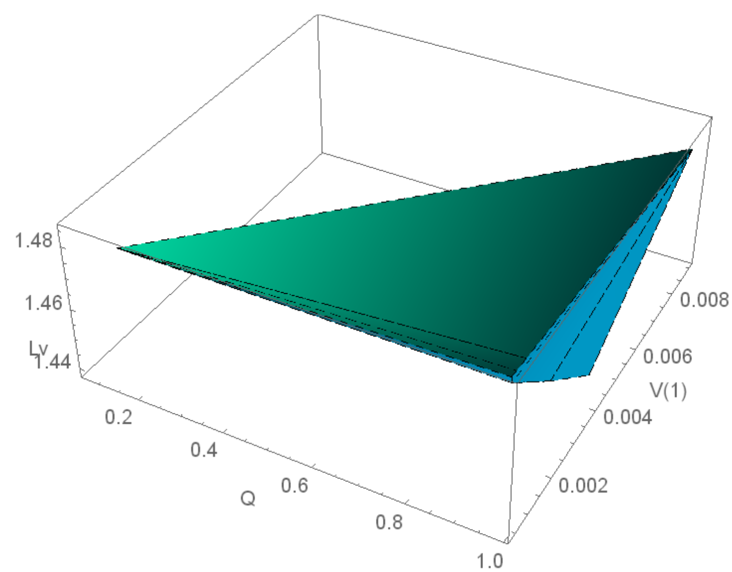


*Figure 4.8: Effect of vacation parameter on Q, V(1) and $L_v$*

Further we try to monitor the effect of the vacation probability on the various queue characteristics by fixing the values of $\lambda = 3, \mu = 4, \gamma = 5, \phi = 2$ and vary the values of $p$ from 0 to 1. When $p = 0$, the server does not take any vacation while $p = 1$ implies that the server is always in vacation state.

Table 4.3: Computed values of various queue characteristics for a $Geom(x)/\gamma/1$ queue for varying values of probability $p$ of vacation with $\lambda = 3, \mu = 4, \gamma = 5, \phi = 2, E(I) = 2.5$

| $p$ | $Q$ | $\rho$ | $V(1)$ | $L_v$ | $L_k$ | $L_q$ |
|---|---|---|---|---|---|---|
| 0.1 | 0.9998 | 0.0002 | 0.000002 | 1.49973 | 0.00002 | 1.49973 |
| 0.25 | 0.99725 | 0.00275 | 0.00009 | 1.49641 | 0.00036 | 1.49677 |
| 0.4 | 0.98811 | 0.01189 | 0.0006 | 1.4847 | 0.00151 | 1.48621 |
| 0.55 | 0.959271 | 0.040739 | 0.00278 | 1.44864 | 0.0051 | 1.45374 |
| 0.75 | 0.91152 | 0.08848 | 0.00804 | 1.39113 | 0.0107 | 1.40133 |
| 0.85 | 0.86228 | 0.13772 | 0.01402 | 1.33259 | 0.0165 | 1.34909 |
| 0.9 | 0.82277 | 0.17723 | 0.01899 | 1.28585 | 0.0211 | 1.30635 |
| 1 | 0.31025 | 0.68975 | 0.08115 | 0.67636 | 0.08115 | 0.75751 |

The figures obtained in Table4.3 are the various queue characteristics for the system for varying values of probability of taking a vacation, *p*. It is observed that as the probability of taking a vacation increase, the proportion of time the server is in idle state tends to decrease while the probability of the server in vacation state increases. It is further observed that when *p* is very small, the proportion of time the server is in vacation state is almost zero. Again, *p* =1, implies that the server is always in a vacation state, shows a much higher utilization factor $\rho$ as compared to the rest of the values.

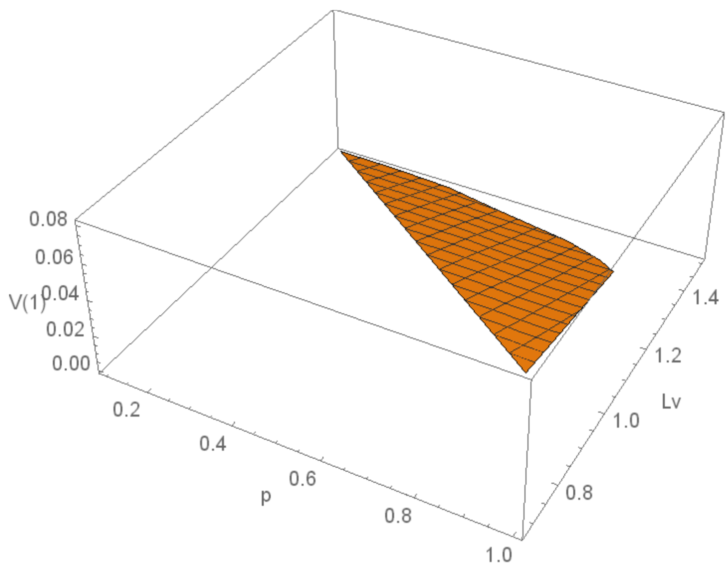


*Figure 4.9: Effect of  p on V(1) and Lv*

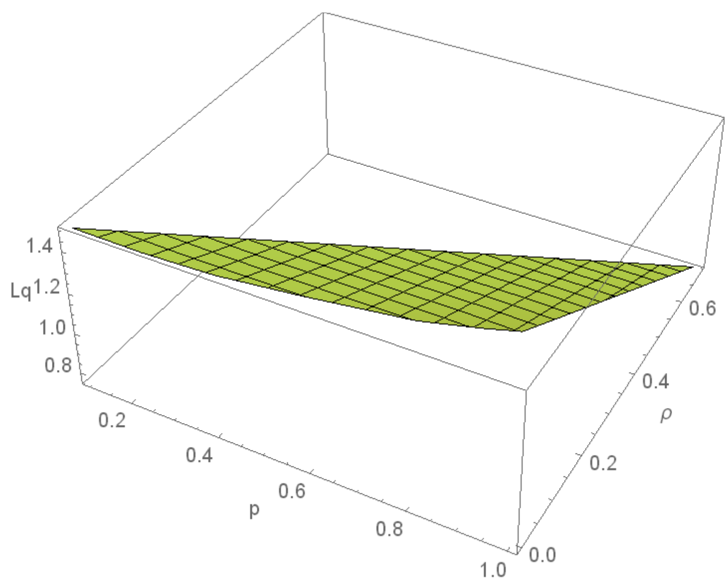


*Figure 4.10: Effect of p on expected queue length and utilization factor* $\rho$

## 4.5 Key Results

The probability generating function of the queue size at a random epoch, $P_q(z) = P(z) + V(z)$, for each of the queueing models are summarized below.

1. $M^X/M/1$ queue with reneging during server vacations:

$$P(z) = \frac{p\mu P_0 - \mu(1-z) + \phi zV(z)}{z\lambda(1-C(z)) - \mu(1-z) + p\mu} \text{ and } V(z) = e^{-aJ(z)-b\delta(z)} ap\mu\left[F(z_0) - \int_z^1 \frac{e^{aJ(t)+b\delta(t)}(1-C(t))dt}{(1-t)g(t)}\right],$$

where, $F(z_0) = \lim_{z\to z_0} \int_z^1 \frac{e^{aJ(t)+b\delta(t)}(1-C(t))dt}{(1-t)g(t)}$, $J(z) = \int_z^1 \frac{(1-C(t))dt}{1-t}$, $\delta(z) = \int_z^1 \left(1 - \frac{p\mu}{g(t)}\right)\frac{dt}{1-t}$

and $g(z) = z\lambda(1-C(z)) - \mu(1-z) + p\mu$ , $a = \frac{\lambda}{\gamma}, b = \frac{\phi}{\gamma}$.

2. $M^X/G/1$ queue with reneging during server vacations:

$$P(z) = \frac{(\lambda C(z) - \lambda)Q + \phi V(z)}{z - (1-p)G^*(\lambda - \lambda C(z))}\left[\frac{1 - G^*(\lambda - \lambda C(z))}{\lambda - \lambda C(z)}\right] \text{, and}$$

$$V(z) = e^{-aJ(z)-b\delta(z)} apQ\left[K(z_0) - \int_z^1 \frac{e^{aJ(t)+b\delta(t)}(1-C(t))G^*(\lambda - \lambda C(t))dt}{(1-t)f(t)}\right]$$

where, $K(z_0) = \int_{z_0}^1 \frac{e^{aJ(t)+b\delta(t)}(1-C(t))G^*(\lambda - \lambda C(t))dt}{(1-t)f(t)}$, with

$J(z) = \int_z^1 \frac{1-C(t)}{1-t}dt, \delta(z) = \int_z^1 \left(1 - \frac{pG^*(\lambda - \lambda C(t))}{f(t)}\right)\frac{dt}{1-t}$ and, $f(z) = z - (1-p)G^*(\lambda - \lambda C(z))$.

3. Special case: A $Geo(x)/\gamma/1$ with reneging during server vacations:

   Considering arrivals according to a Geometric distribution and service time as two parameter Gamma distribution, the different probabilities are obtained as

$$P(z) = \frac{(\lambda C(z) - \lambda)Q + \phi V(z)}{z(\lambda - \lambda C(z) + \mu)^2 - (1-p)\mu^2}\left[\frac{(\lambda - \lambda C(z) + \mu)^2 - \mu^2}{\lambda - \lambda C(z)}\right]$$

$$V(z)=e^{-aJ(z)-b\delta(z)}apQ\left[K(z_0)-\int_z^1 e^{aJ(t)+b\delta(t)}\frac{(1-C(t))G^*(\lambda-\lambda C(t))}{(1-t)f(t)}dt\right],$$

where

$$K(z_0)=\frac{p}{(1-z_0)^{d_1}(1-z_1)^{d_2}(1-z_2)^{d_3}v^a}\underset{z\to z_0}{Lim}\int_z^1\frac{(t-z_0)^{d_1}(t-z_1)^{d_2}(t-z_2)^{d_3}(1-ut)^{a+1}dt}{(t-z_0)(t-z_1)(t-z_2)}$$

and $d_1=m_1b,\ d_2=m_2b,\ d_3=m_3b, A=\dfrac{p}{(1-z_0)^{d_1}(1-z_1)^{d_2}(1-z_2)^{d_3}v^a}$

The parameters $m_1$, $m_2$ and $m_3$ are defined in Section 4.3.4, page 83.

## 4.6 Summary

In this chapter we derive the closed form solutions of probability generating function of queue size for a batch arrival queue with reneging occurring during server vacations. Both Markovian and non-Markovian cases have been treated separately. The mean and variance of queue length is further derived. A special case of $Geom(x)/\gamma/1$ is also discussed with some numerical examples illustrated and some plots to demonstrate the effect of the different assumptions like reneging and vacation on the expected length of queue size.

******

# Chapter 5

# Transient and Steady State Analysis of a Batch Arrival Vacation Queue with Time Homogeneous Breakdowns and Server Providing General Service in Two Fluctuating Modes

## 5.1 Introduction

In this chapter, the author generalizes the basic $M^X/G/1$ model by considering a queueing system with a single server where service is provided in two fluctuating modes, time homogeneous breakdowns and Bernoulli schedule server vacations. Both time dependent and steady state analysis has been done for this model.

Queueing systems whose arrival or service rates fluctuate over times are very common but still needs to be understood analytically. In real life, the service offered to each arriving unit by a server may not be at the same rate; it can be fast, regular or slow. This can be termed as fluctuating modes of service. Due to the fluctuating modes of providing service by a server, the efficiency of a queueing system may be affected. Most of the literature in queueing theory assumes that a service channel provides service in the same mode, i.e. normal, with the same mean rate of service to all customers. However, in practice, it is not always same. For instance, there are fluctuations in downloading different files by a web server, service offered by a human server in banks, post offices or supermarkets etc., may also fluctuate. The assumption of service being provided at the same rate or mode is seemed to be valid only when the server is highly mechanically or electronically controlled. This is a realistic representation of a queueing system which motivated the current author to study fluctuating modes of service in a queueing system. Wu and Takagi (2006) have studied M/G/1 queues with multiple working vacations, where they have assumed that the server works with different service times rather than completely stopping during a vacation period.

In recent years, a majority of the studies related to queues is done considering server breakdowns. We refer to some recent work done by authors like Wang, Cao and Li (2001), Wang (2004), Sherman (2006), Choudhury, Tadj and Paul (2007), Jain and Upadhyaya (2010) and Khalaf et al. (2011) who had studied queues subjected to breakdowns. Studies on queues with batch arrivals along with server vacations also gained significant importance in recent years. Numerous researchers, including Baba (1986), Lee and Srinivasan (1989), Rosenberg and Yechiali (1993), Choudhury (2002) and many more have studied queueing system with batch arrivals following different vacation policies. Presently most of the studies have been devoted to batch arrival vacation queueing system with system breakdowns. The reliability analysis of an M/G/1 queueing system with breakdowns and Bernoulli vacations was studied by Li, Shi and Chao (1997). Madan, Abu-Dayyeh and Gharaibeh (2003b) studied two models of a single bulk queueing system with random breakdowns in which they considered repair times to be exponentially distributed in the first model and deterministic in the second model. Yu et al. (2008) obtained the matrix-geometric solution of a queueing system with two heterogeneous servers.

Most of the studies mentioned assumed that breakdown occurs only during working state of the server. But in real life, a server can fail at any time, whether it is working or idle. In the present chapter we have assumed that server may be subjected to breakdowns both in working as well as an idle state. Madan (2003) studied an M/G/1 queueing system with time homogeneous breakdowns and deterministic repair times.

As discussed in the literature, to the best of the researcher's knowledge, queueing theory lacks studies on queues with assumptions like service in fluctuating modes along with time homogeneous breakdowns and general vacation time. So this study attempts to contribute in generalizing the basic queueing system to a new system, through a combination of the assumptions like fluctuating modes of services, time homogeneous breakdowns of the server and Bernoulli schedule server vacations. We can find most of the literature on queues with breakdowns devoted to steady state solutions since for a time dependent solution more advanced mathematical techniques become necessary. However, in the current research, we obtain both the time dependent as well as steady state solutions in terms of the queue size distribution as well as the probabilities on various states of the server. To monitor the behavior of the queueing system over time, time dependent solutions are useful. The service times, vacation times and repair times are assumed to follow different general (arbitrary) distribution, while the failure or breakdown time is assumed to follow an exponential distribution.

Therefore in this chapter, the author analyzes a queuing system with server vacations and time homogeneous breakdowns in the transient regime. Explicit analytical expressions are obtained for the state probabilities.

This remaining part of the chapter is organized as follows: Section 5.2 gives the assumptions under which the mathematical model is investigated. The definitions and notations are provided in Section5.3 and the equations governing the system are formulated in Section 5.4. The time dependent results are obtained in Section 5.5 while steady state results are derived in Section 5.6. In Section 5.7, we discuss some particular cases of interest. In Section 5.8, a numerical example is given to demonstrate the effect of the fluctuating modes on the different states of the system.

## 5.2 Mathematical Model and Assumptions

In addition to the assumptions of arrival pattern in batches and server vacations as defined in Section 3.2, Chapter Three, the following assumptions are further considered to describe the mathematical model:

a) There is one server providing service in two fluctuating modes. The units are served one by one on a first come first served basis. We assume that just after completion of the previous service or completion of the repair process or completion of a vacation period or just after a batch arrives at a time when the server is in idle state, the service to the next customer starts in mode j with probability $\pi_j$, j=1, 2 and $\pi_1 + \pi_2 = 1$. The service time follows general distribution $G_1$ and $G_2$ in mode 1 and 2 with rate of service $\mu_1$ and $\mu_2$ respectively. Let $G_j(x)$ and $g_j(x)$ be the distribution function and density function of the service times in mode j, j=1, 2. The conditional probability of the service time for the two modes of services is similar to as defined as in Section 4.3 of Chapter Four.

b) In addition, we assume that the system may fail or subjected to breakdown at random. The breakdowns are time–homogeneous in the sense that the server can fail even while it is idle. If a sudden breakdown occurs at a time when a customer is being provided service, then the service of this customer is interrupted and he returns back to the head of the queue. As soon as the repair is complete, this customer is taken up for service in mode *j* with probability $\pi_j, j = 1,2$. We assume that time between breakdowns occur according to an exponential distribution so that $\alpha dt$ is the first order probability of a breakdown during the interval $(t, t+dt]$. Further the repair times follows a general (arbitrary) distribution with distribution function F(x) and density function f(x). let he conditional probability for completion of repair process is $\beta(x)dx$ such that $\beta(x) = \dfrac{f(x)}{1-F(x)}$ and $f(r) = \beta(r)\exp\left(-\int_0^r \beta(x)dx\right)$

All stochastic process involved in the system are assumed to be independent of each other.

## 5.3 Equations Governing the System

The definitions of the probabilities of the different states of the system $P^{(m1)}(x,t)$, $P^{(m2)}(x,t)$, $R_n(x,t)$ and $V_n(x,t)$ are as defined earlier, before Chapter One.

In order to construct the differential-difference equations for the various state probabilities, we consider the transitions between time *t* and (t+ Δt), and follow the continuity argument (Keilson and Kooharian, 1960).

The differential-difference equations that govern the system behavior according to the assumptions made are

$$\frac{\partial}{\partial x}P_n^{(m1)}(x,t) + \frac{\partial}{\partial t}P_n^{(m1)}(x,t) + \left(\lambda + \mu_1(x) + \alpha\right)P_n^{(m1)}(x,t) = \lambda\sum_{i=1}^{n} c_i P_{n-i}^{(m1)}(x,t), \qquad n \geq 1, \tag{5.1}$$

$$\frac{\partial}{\partial x}P_0^{(m1)}(x,t) + \frac{\partial}{\partial t}P_0^{(m1)}(x,t) + \left(\lambda + \mu_1(x) + \alpha\right)P_0^{(m1)}(x,t) = 0, \tag{5.2}$$

$$\frac{\partial}{\partial x}P_n^{(m2)}(x,t) + \frac{\partial}{\partial t}P_n^{(m2)}(x,t) + (\lambda + \mu_2(x) + \alpha)P_n^{(m2)}(x,t) = \lambda\sum_{i=1}^{n} c_i P_{n-i}^{(m2)}(x,t), \qquad n \geq 1, \tag{5.3}$$

$$\frac{\partial}{\partial x}P_0^{(m2)}(x,t)+\frac{\partial}{\partial t}P_0^{(m2)}(x,t)+\left(\lambda+\mu_2(x)+\alpha\right)P_0^{(m2)}(x,t)=0, \tag{5.4}$$

$$\frac{\partial}{\partial x}R_n(x,t)+\frac{\partial}{\partial t}R_n(x,t)+\left(\lambda+\beta(x)\right)R_n(x,t)=\lambda\sum_{i=1}^{n}c_iR_{n-i}(x,t), \qquad n\geq 1, \tag{5.5}$$

$$\frac{\partial}{\partial x}R_0(x,t)+\frac{\partial}{\partial t}R_0(x,t)+\left(\lambda+\beta(x)\right)R_0(x,t)=0\ , \tag{5.6}$$

$$\frac{\partial}{\partial x}V_n(x,t)+\frac{\partial}{\partial t}V_n(x,t)+\left(\lambda+\phi(x)\right)V_n(x,t)=\lambda\sum_{i=1}^{n}c_{n-i}V_{n-i}(x,t), \qquad n\geq 1, \tag{5.7}$$

$$\frac{\partial}{\partial x}V_0(x,t)+\frac{\partial}{\partial t}V_0(x,t)+\left(\lambda+\phi(x)\right)V_0(x,t)=0\ , \tag{5.8}$$

$$\begin{aligned}\frac{\partial}{\partial t}Q(t)+(\lambda+\alpha)Q(t)=(1-p)&\left[\int_0^\infty P_0{}^{(m1)}(x,t)\mu_1(x)dx+\int_0^\infty P_0{}^{(m2)}(x,t)\mu_2(x)dx\right]\\&+\int_0^\infty R_o(x,t)\beta(x)dx+\int_0^\infty V_0(x,t)\phi(x)dx.\end{aligned} \tag{5.9}$$

We assume that initially there is no customer in the system and the server is idle, so that the initial conditions are

$$P_n^{(mj)}(0)=0;\ j=1,2\ \ R_n(0)=0\ ,\ V_n(0)=0 \text{ and } Q(0)=1 \tag{5.10}$$

The above equations can be solved subject to the following boundary conditions:

$$\begin{aligned}P_n^{(m1)}(0,t)=\pi_1(1-p)&\left[\int_0^\infty P_{n+1}^{(m1)}(x,t)\mu_1(x)dx+\int_0^\infty P_{n+1}^{(m2)}(x,t)\mu_2(x)dx\right]\\&+\pi_1\int_0^\infty R_{n+1}(x,t)\beta(x)dx+\pi_1\int_0^\infty V_{n+1}(x,t)\phi(x)dx+\pi_1\lambda c_{n+1}Q(t), \qquad n\geq 0,\end{aligned} \tag{5.11}$$

$$\begin{aligned}P_n^{(m2)}(0,t)=\pi_2(1-p)&\left[\int_0^\infty P_{n+1}^{(m1)}(x,t)\mu_1(x)+\int_0^\infty P_{n+1}^{(m2)}(x,t)\mu_2(x)dx\right]\\&+\pi_2\int_0^\infty R_{n+1}(x,t)\beta(x)dx+\pi_2\int_0^\infty V_{n+1}(x,t)\phi(x)dx+\pi_2\lambda\, c_{n+1}Q(t), \qquad n\geq 0,\end{aligned} \tag{5.12}$$

$$R_{n+1}(0,t)=\alpha\int_0^\infty P_n^{(m1)}(x,t)dx+\alpha\int_0^\infty P_n^{(m2)}(x,t)dx, \qquad n\geq 0, \tag{5.13}$$

$$R_0(0,t) = \alpha Q(t), \tag{5.14}$$

$$V_n\ (0,t) = p\left[\int_0^\infty P_n^{(m1)}(x,t)\mu_1(x)dx + \int_0^\infty P_n^{(m2)}(x,t)\mu_2(x)dx\right], \qquad n \geq 0. \tag{5.15}$$

## 5.4 Time Dependent Solution

To find the solution let us first define the probability generating functions as given below

$$P^{(mj)}(x,z,t) = \sum_{n=0}^{\infty} z^n P_n^{(mj)}(x,t);\ j=1,2;\ R(x,z,t) = \sum_{n-0}^{\infty} z^n R_n(x,t);$$
$$V(x,z,t) = \sum_{n=0}^{\infty} z^n V_n(x,t); \quad C(z) = \sum_{i=1}^{\infty} z^i c_i \tag{5.16}$$

We define the Laplace transform (L. T) of a function $f(t)$ as follows:

$$\tilde{f}(s) = L[f(t)] = \int_0^\infty e^{-st} f(t)dt$$

Also we note that Laplace transform of $\frac{d}{dt} f(t)$ is given by

$$L\left[\frac{d}{dt} f(t)\right] = s\tilde{f}(s) - f(0)$$

Taking L. T. of equations (5.1)-(5.9) using the initial conditions defined in (5.10) we have

$$\frac{\partial}{\partial x}\tilde{P}_n^{(m1)}(x,s) + (s+\lambda+\mu_1(x)+\alpha)\tilde{P}_n^{(m1)}(x,s) = \lambda\sum_{i=1}^{n} c_i \tilde{P}_{n-i}^{(m1)}(x,s), \qquad n \geq 1, \tag{5.17}$$

$$\frac{\partial}{\partial x}\tilde{P}_0^{(m1)}(x,s) + (s+\lambda+\mu_1(x)+\alpha)\tilde{P}_0^{(m1)}(x,s) = 0\ , \tag{5.18}$$

$$\frac{\partial}{\partial x}\tilde{P}_n^{(m2)}(x,s) + (s+\lambda+\mu_2(x)+\alpha)\tilde{P}_n^{(m2)}(x,s) = \lambda\sum_{i=1}^{n} c_i \tilde{P}_{n-i}^{(m2)}(x,s), \qquad n \geq 1\ , \tag{5.19}$$

$$\frac{\partial}{\partial x}\tilde{P}_0^{(m2)}(x,s) + (s+\lambda+\mu_2(x)+\alpha)\tilde{P}_0^{(m2)}(x,s) = 0\ , \tag{5.20}$$

$$\frac{\partial}{\partial x}\widetilde{R}_n(x,s)+\left(s+\lambda+\beta(x)\right)\widetilde{R}_n(x,s)=\lambda\sum_{i=1}^{n}c_i\widetilde{R}_{n-i}(x,s) \qquad n\geq 1 \tag{5.21}$$

$$\frac{\partial}{\partial x}\widetilde{R}_0(x,s)+\left(s+\lambda+\beta(x)\right)\widetilde{R}_0(x,s)=0\,, \tag{5.22}$$

$$\frac{\partial}{\partial x}\widetilde{V}_n(x,s)+\left(s+\lambda+\phi(x)\right)\widetilde{V}_n(x,s)=\lambda\sum_{i=1}^{n}c_i\widetilde{V}_{n-i}(x,s), \qquad n\geq 1, \tag{5.23}$$

$$\frac{\partial}{\partial x}\widetilde{V}_0(x,s)+\left(s+\lambda+\phi(x)\right)\widetilde{V}_0(x,s)=0, \tag{5.24}$$

$$\begin{aligned}(s+\lambda+\alpha)\widetilde{Q}(s)&=1+(1-p)\left[\int_0^\infty \widetilde{P}_0^{(m1)}(x,s)\mu_1(x)dx+\int_0^\infty \widetilde{P}_0^{(m2)}(x,s)\mu_2(x)dx\right]\\ &\quad+\int_0^\infty \widetilde{R}_0(x,s)\beta(x)dx+\int_0^\infty \widetilde{V}_0\ (x,s)\phi(x)dx\end{aligned} \tag{5.25}$$

Now multiplying equation (5.17) by $z^n$ and summing over all possible values of $n$, adding with (5.18) and using (5.16) we get

$$\frac{\partial}{\partial x}\widetilde{P}^{(m1)}(x,z,s)+\left(s+\lambda-\lambda C(z)+\mu_1(x)+\alpha\right)\widetilde{P}^{(m1)}(x,z,s)=0\,. \tag{5.26}$$

Proceeding similarly with equations (5.19), (5.21) and (5.23) we get

$$\frac{\partial}{\partial x}\widetilde{P}^{(m2)}(x,z,s)+\left(s+\lambda-\lambda C(z)+\mu_2(x)+\alpha\right)\widetilde{P}^{(m2)}(x,z,s)=0\,, \tag{5.27}$$

$$\frac{\partial}{\partial x}\widetilde{R}\ (x,z,s)+\left(s+\lambda-\lambda C(z)+\beta(x)\right)\widetilde{R}\ (x,z,s)=0\,, \tag{5.28}$$

$$\frac{\partial}{\partial x}\widetilde{V}\ (x,z,s)+\left(s+\lambda-\lambda C(z)+\phi(x)\right)\widetilde{V}\ (x,z,s)=0 \tag{5.29}$$

Now we integrate equations (5.26)-(5.29) to obtain

$$\widetilde{P}^{(m1)}(x,z,s)=\widetilde{P}^{(m1)}(0,z,s)\exp\left[-\left(s+\lambda-\lambda C(z)+\alpha\right)x-\int_0^x \mu_1(t)dt\right] \tag{5.30}$$

$$\widetilde{P}^{(m2)}(x,z,s)=\widetilde{P}^{(m2)}(0,z,s)\exp\left[-\left(s+\lambda-\lambda C(z)+\alpha\right)x-\int_0^x \mu_2(t)dt\right] \tag{5.31}$$

$$\widetilde{R}(x,z,s)=\widetilde{R}\ \ (0,z,s)\exp\left[-\left(s+\lambda-\lambda C(z)\right)x-\int_0^x \beta(t)dt\right] \tag{5.32}$$

$$\widetilde{V}(x,z,s)=\widetilde{V}(0,z,s)\exp\left[-\left(s+\lambda-\lambda C(z)\right)x-\int_0^x \phi(t)dt\right] \tag{5.33}$$

Now taking the L. T of the boundary conditions (5.11)-(5.15) we have

$$\tilde{P}_n^{(m1)}(0,s)=\pi_1(1-p)\left[\int_o^\infty \tilde{P}_{n+1}^{(m1)}(x,s)\mu_1(x)dx+\int_0^\infty \tilde{P}_{n+1}^{(m2)}(x,s)\mu_2(x)dx\right]$$
$$+\pi_1\int_0^\infty \tilde{R}_{n+1}(x,s)\beta(x)dx+\pi_1\int_0^\infty \tilde{V}_{n+1}(x,s)\phi(x)dx+\pi_1\lambda c_{n+1}\tilde{Q}(s), \qquad n\geq 0, \tag{5.34}$$

$$\tilde{P}_n^{(m2)}(0,s)=\pi_2(1-p)\left[\int_o^\infty \tilde{P}_{n+1}^{(m1)}(x,s)\mu_1(x)dx+\int_0^\infty \tilde{P}_{n+1}^{(m2)}(x,s)\mu_2(x)dx\right]$$
$$+\pi_2\int_0^\infty \tilde{R}_{n+1}(x,s)\beta(x)dx+\pi_2\int_0^\infty \tilde{V}_{n+1}(x,s)\phi(x)dx+\pi_2\lambda c_{n+1}\tilde{Q}(s), \qquad n\geq 0, \tag{5.35}$$

$$\tilde{R}_{n+1}(0,s)=\alpha\int_0^\infty \tilde{P}_n^{(m1)}(x,s)dx+\alpha\int_0^\infty \tilde{P}_n^{(m2)}(x,s)dx, \qquad n\geq 0, \tag{5.36}$$

$$\tilde{R}_0(0,s)=\alpha\tilde{Q}(s), \tag{5.37}$$

$$\tilde{V}_n(0,s)=p\left[\int_0^\infty \tilde{P}_n^{(m1)}(x,s)\mu_1(x)dx+\int_0^\infty \tilde{P}_n^{(m2)}(x,s)\mu_2(x)dx\right],\ n\geq 0. \tag{5.38}$$

For the boundary conditions (5.34) and (5.35) we multiply by $z^{n+1}$, summing over all possible values of $n$ and using the probability generating functions defined in (5.16) we get

$$z\tilde{P}^{(m1)}(0,z,s)=\pi_1+\pi_1(1-p)\left[\int_0^\infty \tilde{P}^{(m1)}(x,z,s)\mu_1(x)dx+\int_0^\infty \tilde{P}^{(m2)}(x,z,s)\mu_2(x)dx\right]$$
$$+\pi_1\int_0^\infty \tilde{R}(x,z,s)\beta(x)dx+\pi_1\int_0^\infty \tilde{V}(x,z,s)\phi(x)dx+\pi_1\big(\lambda C(z)-(s+\lambda+\alpha)\big)\tilde{Q}(s), \tag{5.39}$$

$$z\tilde{P}^{(m2)}(0,z,s)=\pi_2+\pi_2(1-p)\left[\int_0^\infty \tilde{P}^{(m1)}(x,z,s)\mu_1(x)dx+\int_0^\infty \tilde{P}^{(m2)}(x,z,s)\mu_2(x)dx\right]$$
$$+\pi_2\int_0^\infty \tilde{R}(x,z,s)\beta(x)dx+\pi_2\int_0^\infty \tilde{V}(x,z,s)\phi(x)dx+\pi_2\big(\lambda C(z)-(s+\lambda+\alpha)\big)\tilde{Q}(s) \tag{5.40}$$

In a similar way, multiplying equation (5.36) by $z^{n+1}$, adding with (5.37) and using (5.16) we get

$$\begin{aligned}\tilde{R}(0,z,s) &= \alpha z\left[\int_0^\infty \tilde{P}_n^{(m1)}(x,z,s)dx + \int_0^\infty \tilde{P}_n^{(m2)}(x,z,s)dx\right] + \alpha\tilde{Q}(s) \\ &= \alpha z\tilde{P}^{(m1)}(z,s) + \alpha z\tilde{P}^{(m2)}(z,s) + \alpha\tilde{Q}(s)\end{aligned} \tag{5.41}$$

Again multiplying equation (5.38) by $z^n$ and using (5.16) we get

$$\tilde{V}(0,z,s) = p\left[\int_0^\infty \tilde{P}^{(m1)}(x,z,s)\mu_1(x)dx + \int_0^\infty \tilde{P}^{(m2)}(x,z,s)\mu_2(x)dx\right] \tag{5.42}$$

Again integrating equations (5.30)-(5.33) by parts with respect to *x*, yields

$$\tilde{P}^{(m1)}(z,s) = \frac{\tilde{P}^{(m1)}(0,z,s)\left[1 - G_1^*(\lambda - \lambda C(z) + s + \alpha)\right]}{\lambda - \lambda C(z) + s + \alpha}, \tag{5.43}$$

$$\tilde{P}^{(m2)}(z,s) = \frac{\tilde{P}^{(m2)}(0,z,s)\left[1 - G_2^*(\lambda - \lambda C(z) + s + \alpha)\right]}{\lambda - \lambda C(z) + s + \alpha}, \tag{5.44}$$

$$\tilde{R}(z,s) = \frac{\tilde{R}(0,z,s)\left[1 - F^*(\lambda - \lambda C(z) + s)\right]}{\lambda - \lambda C(z) + s}, \tag{5.45}$$

$$\tilde{V}(z,s) = \frac{\tilde{V}(0,z,s)\left[1 - W^*(\lambda - \lambda C(z) + s)\right]}{\lambda - \lambda C(z) + s}, \tag{5.46}$$

where $G_j^*(\lambda - \lambda C(z) + s + \alpha) = \int_0^\infty e^{-(\lambda - \lambda C(z) + \alpha + s)x} dG_j(x);\ j = 1,2$ is the Laplace-Stieltjes transform of service time in mode *j*(*j*=1, 2), $F^*(\lambda - \lambda C(z) + s) = \int_0^\infty e^{-(\lambda - \lambda C(z) + s)x} dF(x)$ is the Laplace-Stieltjes transform of repair time and $W^*(\lambda - \lambda C(z) + s) = \int_0^\infty e^{-(\lambda - \lambda c(z) + s)x} dW(x)$ is the Laplace-Stieltjes transform of vacation time.

Now multiplying (5.30), (5.31), (5.32) and (5.33) by $\mu_1(x), \mu_2(x), \beta(x)$ and $\phi(x)$ respectively and then integrating over *x*, we get

$$\int_0^\infty \tilde{P}^{(m1)}(x,z,s)\mu_1(x)dx = \tilde{P}^{(m1)}(0,z,s)G_1^*\left(\lambda - \lambda C(z) + s + \alpha\right), \tag{5.47}$$

$$\int_0^\infty \tilde{P}^{(m2)}(x,z,s)\mu_2(x)dx = \tilde{P}^{(m2)}(0,z,s)G_2^*\left(\lambda - \lambda C(z) + s + \alpha\right), \tag{5.48}$$

$$\int_0^\infty \tilde{R}(x,z,s)\beta(x)dx = \tilde{R}(0,z,s)F^*\left(\lambda - \lambda C(z) + s\right), \tag{5.49}$$

$$\int_0^\infty \tilde{V}(x,z,s)\phi(x)dx = \tilde{V}(0,z,s)W^*\left(\lambda - \lambda C(z) + s\right) \tag{5.50}$$

Utilizing (5.47)-(5.50) in equations (5.39) and (5.40) and using (5.41)-(5.42), we get

$$z\tilde{P}^{(m1)}(0,z,s) = \pi_1 + \pi_1(1-p)\tilde{P}^{(m1)}(0,z,s)G_1^*(s+\lambda-\lambda C(z)+\alpha)$$

$$+\pi_1(1-p)\tilde{P}^{(m2)}(0,z,s)G_2^*(s+\lambda-\lambda C(z)+\alpha)+\pi_1\tilde{R}(0,z,s)F^*(s+\lambda-\lambda C(z))+\alpha\pi_1\tilde{Q}(s)$$
$$+\pi_1 p\tilde{P}^{(m1)}(0,z,s)G_1^*(s+\lambda-\lambda C(z)+\alpha)W^*(s+\lambda-\lambda C(z))$$

$$+\pi_1 p\tilde{P}^{(m2)}(0,z,s)G_2^*(s+\lambda-\lambda C(z)+\alpha)W^*(s+\lambda-\lambda C(z))+\pi_1\left(\lambda C(z)-(s+\lambda+\alpha)\right)\tilde{Q}(s)$$

Re-arranging the terms of the above equation, we have

$$\begin{aligned}
&\left[z-\pi_1\left\{(1-p)+pW^*(s+\lambda-\lambda C(z))\right\}G_1^*(s+\lambda-\lambda C(z)+\alpha)\right]\tilde{P}^{(m1)}(0,z,s)\\
&-\pi_1\left\{(1-p)+pW^*(s+\lambda-\lambda C(z))\right\}G_2^*(s+\lambda-\lambda C(z)+\alpha)\tilde{P}^{(m2)}(0,z,s)\\
&-\pi_1\tilde{R}(0,z,s)F^*(s+\lambda-\lambda C(z)) = \pi_1+\pi_1\left(\lambda C(z)-\lambda-s-\alpha\right)\tilde{Q}(s)
\end{aligned} \tag{5.51}$$

Similarly proceeding with equation (5.40) we have

$$\begin{aligned}
&-\pi_2\left\{(1-p)+pW^*\left(s+\lambda-\lambda C(z)\right)\right\}G_1^*(s+\lambda-\lambda C(z)+\alpha)\tilde{P}^{(m1)}(0,z,s)\\
&+\left[z-\pi_2\left\{(1-p)+pW^*(s+\lambda-\lambda C(z))\right\}G_2^*\left(s+\lambda-\lambda C(z)+\alpha\right)\right]\tilde{P}^{(m2)}(0,z,s)\\
&-\pi_2\tilde{R}(0,z,s)F^*(s+\lambda-\lambda C(z)) = \pi_2+\pi_2\left(\lambda C(z)-\lambda-s-\alpha\right)\tilde{Q}(s)
\end{aligned} \tag{5.52}$$

And equation (5.41) can be expressed as

$$\frac{\alpha z\left[1-G_1^*\left(s+\lambda-\lambda C(z)+\alpha\right)\right]}{s+\lambda-\lambda C(z)+\alpha}\tilde{P}^{(m1)}(0,z,s)+\frac{\alpha z\left[1-G_2^*\left(s+\lambda-\lambda C(z)+\alpha\right)\right]}{s+\lambda-\lambda C(z)+\alpha}\tilde{P}^{(m2)}(0,z,s) - \tilde{R}(0,z,s) = -\alpha\tilde{Q}(s) \tag{5.53}$$

Solving equations (5.51)-(5.53) for $\tilde{P}^{(m1)}(0,z,s), \tilde{P}^{(m2)}(0,z,s)$ and $\tilde{R}(0,z,s)$ by Cramer's rule and considering $\left(s+\lambda-\lambda C(z)\right)=\psi$ and $\left(s+\lambda-\lambda C(z)+\alpha\right)=\psi+\alpha$, we obtain

$$\tilde{P}^{(m1)}\left(0,z,s\right)=\frac{\pi_1\left[\tilde{Q}(s)\left((\psi+\alpha)-\alpha F^*(\psi\ )\right)-1\right]}{D(z,s)} \tag{5.54}$$

$$\tilde{P}^{(m2)}\left(0,z,s\right)=\frac{\pi_2\left[\tilde{Q}(s)\left((\psi+\alpha)-\alpha F^*(\psi)\right)-1\right]}{D(z,s)} \tag{5.55}$$

$$\tilde{R}(0,z,s)=\frac{\alpha\begin{bmatrix}\tilde{Q}(s)\begin{bmatrix}\left\{(1-p)+pW^*\left(\psi\right)\right\}\left\{\pi_1G_1^*(\psi+\alpha)+\pi_2G_2^*(\psi+\alpha)\right\}\\ -z\left\{\pi_1G_1^*(\psi+\alpha)+\pi_2G_2^*(\psi+\alpha)\right\}\end{bmatrix}\\ -z\left\{\dfrac{\pi_1(1-G_1^*\left(\psi+\alpha\right))}{\psi+\alpha}+\dfrac{\pi_2(1-G_2^*\left(\psi+\alpha\right))}{\psi+\alpha}\right\}\end{bmatrix}}{D(z,s)} \tag{5.56}$$

Now from (5.42) and using (5.54) and (5.55) we have

$$\tilde{V}(0,z,s)=p\left[\tilde{P}^{(m1)}(0,z,s)G_1^*(\psi+\alpha)+\tilde{P}^{(m2)}(0,z,s)G_2^*(\psi+\alpha)\right]$$

$$\tilde{V}(0,z,s)=\frac{p\left[\tilde{Q}(s)\left\{(\psi+\alpha)-\alpha F^*(\psi)\right\}-1\right]\left[\pi_1G_1^*(\psi+\alpha)+\pi_2G_2^*(\psi+\alpha)\right]}{D(z,s)} \tag{5.57}$$

And

$$D(z,s)=\left\{(1-p)+pW^*\left(\psi\right)\right\}\left\{\pi_1G_1^*\left(\psi+\alpha\right)+\pi_2G_2^*\left(\psi+\alpha\right)\right\} + \frac{z\ \alpha F^*\left(\psi\right)\left\{\pi_1(1-G_1^*(\psi+\alpha)+\pi_2\left(1-G_2^*(\psi+\alpha)\right)\right\}}{\psi+\alpha}-z \tag{5.58}$$

Further substituting the above equations (5.54)-(5.57) in equations (5.43)-(5.46) we get

$$\tilde{P}^{(m1)}(z,s)=\tilde{P}^{(m1)}(0,z,s)\left[\frac{1-G_1^*(\psi+\alpha)}{\psi+\alpha}\right] \tag{5.59}$$

$$\tilde{P}^{(m2)}(z,s)=\tilde{P}^{(m2)}(0,z,s)\left[\frac{1-G_2^*(\psi+\alpha)}{\psi+\alpha}\right] \tag{5.60}$$

$$\tilde{R}(z,s)=\tilde{R}(0,z,s)\left[\frac{1-F^*(\psi)}{\psi}\right] \tag{5.61}$$

$$\tilde{V}(z,s)=\tilde{V}(0,z,s)\left[\frac{1-W^*(\psi\ )}{\psi}\right] \tag{5.62}$$

Let $\tilde{P}_q(z,s)$ denote the probability generating function of the queue size irrespective of the state of the system. Then adding equations(5.59)-(5.62) we get

$$\tilde{P}_q(z,s)=\tilde{P}^{(m1)}(z,s)+\tilde{P}^{(m2)}(z,s)+\tilde{R}(z,s)+\tilde{V}(z,s)=\frac{N(z,s)}{D(z,s)}$$

(5.63)The probability generating functions for the different states of the system in equations (5.59)-(5.62) satisfy the normalization condition

$$\tilde{P}_q(1.s)+\tilde{Q}(s)=\frac{1}{s} \tag{5.64}$$

It follows from the definition that $\tilde{P}_q(z,s)$ being a generating function for a normalized set of probabilities, the series which defines it must converge everywhere in the unit circle $|z|\leq 1$. Thus in this region zeroes of both numerator and denominator on the right-hand side of (5.63) must coincide. The denominator on the right-hand side is continuous and differentiable in the interval [0, 1] and therefore, let us examine the existence of a single root in the interval (0,1), say, $z_0$ , for which $D(z_0,s)=0$. For convenience, we denote, $\pi_1 G_1^*(\psi+\alpha)+\pi_2 G_2^*(\psi+\alpha)=A^*(\psi+\alpha)$ .Thus, $A^*(\psi+\alpha)$ is the effective probability generating function of the service time offered in the two modes.

For z = 0, $D(z,s)$ in (5.58) becomes

$$D(0,s)=\{1-p(1-W^{*}(s+\lambda))\}A^{*}(s+\lambda+\alpha) \tag{5.65}$$

Since $0\le p\le 1$, $\pi_1+\pi_2=1$ and $W^{*}(s+\lambda), A^{*}(s+\lambda+\alpha), j=1,2,$ being Laplace -Stieltjes transform are non-negative and $\le 1$, therefore from (5.65), $D(0,s)>0$.

Again for $z=1$,

$$D(1,s)=A^{*}(s+\alpha)\{1-p(1-W^{*}(s))\}+\frac{\alpha F^{*}(s)\{1-A^{*}(s+\alpha)\}}{s+\alpha}-1 \tag{5.66}$$

Since it is known that $p\ge 0$, $p(1-W^{*}(s))\ge 0$, let us examine equation (5.66) for $p = 0$. Then (5.66) reduces to

$$D(1,s)\big|_{p=0}=A^{*}(s+\alpha)+\frac{\alpha F^{*}(s)\left[1-A^{*}(s+\alpha)\right]}{s+\alpha}-1$$

Now, $F^{*}(s)\le 1, A^{*}(s+\alpha)\le 1$ being Laplace-Stieltjes transform are non-negative for $\Re(s)\ge 0$

Let us examine this equation by contradiction

We consider that: $D(1,s)\big|_{p=0}\ge 0$

$$\frac{\alpha F^{*}(s)\left[1-A^{*}(s+\alpha)\right]}{s+\alpha}\ge 1-A^{*}(s+\alpha),\ 1-A^{*}(s+\alpha)\ge 0$$

$$\frac{\alpha F^{*}(s)}{s+\alpha}\ge 1,$$

$$F^{*}(s)\ge\frac{s+\alpha}{\alpha}$$

$$F^{*}(s)\ge\frac{s}{\alpha}+1$$

$\Rightarrow\ F^{*}(s)\ge\dfrac{s}{\alpha}+1$, is not true.

Thus $D(1,s)\big|_{p=0}<0$

Again, $D(1,s)\big|_{p=0}=D(1,s)+p(1-W^{*}(s))<0$

Since, $p(1-W^{*}(s))\ge 0$, so it implies that $D(1,s)$ must be $<0$.

$\therefore D(1,s)<0$

Hence the above argument provides evidence that there exists at least one root, $z_0$ in the interval (0 ,1). Further, let us apply Rouche's theorem to verify that $D(z,s)$ has only one root in the unit circle $|z|\le 1$.

$$D(z,s)=\left\{(1-p)+pW^*(\psi)\right\}A^*(\psi+\alpha)+\frac{z\ \alpha F^*(\psi)\left\{1-A^*(\psi+\alpha)\right\}}{\psi+\alpha}-z$$

We consider $f(z)=-z, g(z)=\left(1-p+pW^*(\psi)\right)A^*(\psi+\alpha)+\dfrac{z\alpha F^*(\psi)\left\{1-A^*(\psi+\alpha)\right\}}{\psi+\alpha}$

Now, $|f(z)|=|-z|=1$ for $|z|=1$

(5.67)

$$\begin{aligned}|g(z)|&=\left|\left(1-p(1-W^*(\psi)\right)A^*(\psi+\alpha)+\frac{z\alpha F^*(\psi)(1-A^*(\psi+\alpha)}{\psi+\alpha}\right|\\ &=\left|\left(1-p(1-W^*(s)\right)A^*(s)+\frac{\alpha F^*(s)(1-A^*(s+\alpha))}{s+\alpha}\right| \qquad (5.68)\\ &\leq\left|\left(1-p(1-W^*(s)\right)A^*(s+\alpha)\right|+\left|\frac{\alpha F^*(s)(1-A^*(s+\alpha)}{s+\alpha}\right|\end{aligned}$$

Now $F^*(s)=\int_0^\infty e^{-sx}dF(x), A^*(s+\alpha)=\int_0^\infty e^{-(s+\alpha)x}dA(s)$, are Laplace –Steiltjes transforms for $\Re(s)\geq 0$. So for $\Re(s)$ >0, $\left|F^*(\psi)\right|,\left|W^*(\psi)\right|$ is <1. Also,

$\left|A^*(s+\alpha)\right|$ <1, for $\alpha$>0, $\Re(s)$ >0, $\left|\dfrac{\alpha}{s+\alpha}\right|$ <1, for $s$, $\alpha$ >0, and $C(z)=\sum_{i=}^{n}c_i z^i$, being a power series with probabilities as coefficients, converges if and only if $|z|\leq 1$, which therefore implies that $|g(z)|$ < 1.

Thus it is seen that, $|f(z)|\geq|g(z)|$ for the contour $|z|=1$.Since $f(z)$ has one zero inside the unit circle, it implies by Rouche's theorem that $f(z)+g(z)$ also has one zero inside the unit circle $|z|=1$. This confirms the verification that there exists only one real root of the function $D(z,s)$ in the interval (0, 1).

Considering $z_0$ as the unique root of the denominator in (0, 1) on the right hand side of (5.63), the numerator vanishes for $z=z_0$.

Thus the unknown probability $\tilde{Q}(s)$ can be determined from the equation $N(z_0,s)=0$, which after some algebra is obtained as

$$\tilde{Q}(s)=\frac{(1-A^*(\psi_1+\alpha))\left\{\psi_1+\alpha z_0(1-F^*(\psi_1)\right\}}{\psi_1(\psi_1+\alpha)}\left[M(z_0,s)\right]^{-1}, \qquad (5.69)$$

where we denote $\psi_1 = \lambda - \lambda C(z_0)$ ,and

$$M(z_0,s) = \{(\psi_1+\alpha) - \alpha F^*(\psi_1)\}\left\{\frac{1-A^*(\psi_1+\alpha)}{\psi_1+\alpha} + \frac{p(1-W^*(\psi_1))}{\psi_1}\right\}$$
$$+\left[\alpha\{1-p+pW^*(\psi_1)\} - z_0\right]\frac{A^*(\psi_1+\alpha)(1-F^*(\psi_1))}{\psi_1} - \frac{pA^*(\psi_1+\alpha)(1-W^*(\psi_1))}{\psi_1}$$

## 5.5 Steady State Results

We drop the argument ($t$) to define the steady state probabilities and the corresponding probability generating functions and for that matter, the argument $s$ wherever it appears in the time dependent analysis up to this point. Now by applying the well known Tauberian property, the corresponding steady state probabilities can be obtained, which is given by

$$\lim_{s\to 0} s\,\tilde{f}(s) = \lim_{t\to\infty} f(t) \tag{5.70}$$

Thus multiplying (5.54) by $s$, taking limit as $s \to 0$ and applying (5.70) we have

$$P^{(m1)}(z) = \lim_{s\to 0} s\tilde{P}^{(m1)}(z,s)$$
$$= \frac{\lim_{s\to 0}\left[\{s\pi_1\tilde{Q}(s)((\psi+\alpha)-\alpha F^*(\psi)) - s\}\right]}{D(z)}\left[\frac{1-G_1^*(\psi+\alpha)}{\psi+\alpha}\right] \tag{5.71}$$
$$P^{(m1)}(z) = \frac{\left[\pi_1 Q((\psi+\alpha)-\alpha F^*(\psi))\right]}{D(z)}\left[\frac{1-G_1^*(\psi+\alpha))}{\psi+\alpha}\right]$$

Similarly proceeding with equations (5.55)-(5.57) we have

$$P^{(m2)}(z) = \underset{s\to 0}{Lim}\, s\tilde{P}^{(m2)}(z,s) = \frac{\left[\pi_2 Q\{(\psi+\alpha)-\alpha F^*(\psi)\}\right]}{D(z)}\left[\frac{1-G_2^*(\psi+\alpha)}{\psi+\alpha}\right] \tag{5.72}$$

$$R(z) = \frac{\alpha\left[Q\{\pi_1 G_1^*(\psi+\alpha)+\pi_2 G_2^*(\psi+\alpha)\}\left[\{1-p+pW^*(\psi)\}-z\right]\right]}{D(z)}\left[\frac{1-F^*(\psi)}{\psi}\right] \tag{5.73}$$

$$V(z)=\frac{pQ\left[\left((\psi+\alpha)-\alpha F^{*}(\psi)\right)\right]\left[\pi_1 G_1^{*}(\psi+\alpha)+\pi_2 G_2^{*}(\psi+\alpha)\right]}{D(z)}\left[\frac{1-W^{*}(\psi)}{\psi}\right], \tag{5.74}$$

where, $\psi=\lambda-\lambda C(z),$ and

$$D(z)=\left\{(1-p)+pW^{*}(\psi)\right\}\left\{\pi_1 G_1^{*}(\psi+\alpha))+\pi_2 G_2^{*}(\psi+\alpha))\right\}$$
$$+\frac{z\alpha F^{*}(\psi))}{\psi+\alpha}\left\{\pi_1\left(1-G_1^{*}(\psi+\alpha)\right)+\pi_2\left(1-G_2^{*}(\psi+\alpha)\right)\right\}-z$$

Equations (5.71)-(5.74) gives the marginal probability generating function of the server's state queue size distribution under the stability condition ρ <1.Now for the steady states we add equations (5.71)-(5.74) and get

Let $P_q(z)=P^{(m1)}(z)+P^{(m2)}(z)+R(z)+V(z)=\dfrac{N(z)}{D(z)}$ (5.75)

According to normalization condition, we have $P_q(1)+Q=1$. We see that for z =1, $P_q(z)$ is indeterminate of the 0/0 form. Using L' Hopital's rule in equation (5.75), where we differentiate both numerator and denominator with respect to *z*, and accordingly get

$$P^{(m1)}(1)=\lim_{z\to 1}P^{(m1)}(z)=\frac{Q\pi_1\left[\lambda E(I)+\alpha\lambda E(I)E(R)\right]\left[\frac{1-G_1^{*}(\alpha)}{\alpha}\right]}{D'(1)}, \tag{5.76}$$

$$P^{(m2)}(1)=\lim_{z\to 1}P^{(m2)}(z)=\frac{Q\pi_2\left[\lambda E(I)+\alpha\lambda E(I)E(R)\right]\left[\frac{1-G_2^{*}(\alpha)}{\alpha}\right]}{D'(1)}, \tag{5.77}$$

$$R(1)=\lim_{z\to 1}R(z)=\frac{\alpha Q\left[\pi_1 G_1^{*}(\alpha)+\pi_2 G_2^{*}(\alpha)\right]\left[1-p\lambda E(I)E(V)\right]E(R)}{D'(1)}, \tag{5.78}$$

$$V(1)=\lim_{z\to 1}V(z)=\frac{pQ\left[\lambda E(I)+\alpha\lambda E(I)E(R)\right]\left[\pi_1 G_1^{*}(\alpha)+\pi_2 G_2^{*}(\alpha)\right]E(V)}{D'(1)}, \tag{5.79}$$

where $C'(1)=E(I)$ is the mean batch size of arriving customers, $E(R)$ is the mean of repair time and $E(V)$ is the mean vacation time and

$$D'(1)=\lim_{z\to 1} D(z)=\left[1-p\lambda E(I)E(V)\right]\left[\pi_1 G_1^*(\alpha)+\pi_2 G_2^*(\alpha)\right]$$
$$-\lambda E(I)(1+\alpha E(R))\left[\frac{\pi_1\left(1-G_1^*(\alpha)\right)+\pi_2\left(1-G_2^*(\alpha)\right)}{\alpha}\right]$$

Now utilizing the above equations (5.76)-(5.79) in the normalizing condition, obtain Q as

$$Q=\frac{\left[\pi_1 G_1^*(\alpha)+\pi_2 G_2^*(\alpha)\right]\left[1-p\lambda E(I)E(V)\right]-\lambda E(I)\left[1+\alpha E(R)\right]\left[\dfrac{\pi_1\left(1-G_1^*(\alpha)\right)+\pi_2\left(1-G_2^*(\alpha)\right)}{\alpha}\right]}{\left[\pi_1 G_1^*(\alpha)+\pi_2 G_2^*(\alpha)\right]\left[1+\alpha E(R)\right]}, \tag{5.80}$$

and the stability condition emerges to be

$$\frac{\left[\pi_1 G_1^*(\alpha)+\pi_2 G_2^*(\alpha)\right]\left[\alpha E(R)+p\lambda E(I)E(V)\right]-\lambda E(I)\left[1+\alpha E(R)\right]\left[\dfrac{\pi_1\left(1-G_1^*(\alpha)\right)+\pi_2\left(1-G_2^((\alpha)\right)}{\alpha}\right]}{\left[\pi_1 G_1^*(\alpha)+\pi_2 G_2^*(\alpha)\right]\left[1+\alpha E(R)\right]}<1 \tag{5.81}$$

We can substitute the value of Q in equations (5.76)-(5.79) and obtain clear and explicit expression for $P^{(mi)}(1); i=1,2$, the steady state probability that the server is active and providing service to customers in two different fluctuating modes $\pi_1$ and $\pi_2$ at any random point of time, R(1) the steady state probability that the server is in failed state and is under repairs, and V(1) the steady state probability that the server goes for vacation at any random point of time.

## 5.6 Particular Cases

### 5.6.1 Exponential Service Time, Vacation Completion Time and Repair Time

Here we consider that the service times, vacation completion time and repair times to be exponentially distributed. We have

$$G_1^*(\psi+\alpha)=\frac{\mu_1}{\mu_1+\psi+\alpha}, G_2^*(\psi+\alpha))=\frac{\mu_2}{\mu_2+\psi+\alpha}, F^*(\psi)=\frac{\beta}{\beta+\psi}, W^*(\psi_2(z))=\frac{\phi}{\phi+\psi}$$

$E(R)=\dfrac{1}{\beta}\ E(V)=\dfrac{1}{\phi}$, where $\psi=\lambda-\lambda C(z)$

Thus our equations in (5.76)-(5.79) changes to

$$P^{(m1)}(z)=\frac{\frac{\pi_1 Q}{\mu_1+\psi+\alpha}\left[\frac{(\psi+\alpha)(\beta+\psi)-\alpha\beta}{\beta+\psi}\right]}{D_1} \tag{5.82}$$

$$P^{(m2)}(z)=\frac{\frac{\pi_2 Q}{\mu_2+\psi+\alpha}\left[\frac{(\psi+\alpha)(\beta+\psi)-\alpha\beta}{\beta+\psi}\right]}{D_1} \tag{5.83}$$

$$R(z)=\frac{\alpha Q\left[\frac{\pi_1\mu_1}{\mu_1+\psi+\alpha}+\frac{\pi_2\mu_2}{\mu_2+\psi+\alpha}\right]\left[\left\{\frac{(1-p)(\phi+\psi)+p\phi}{\phi+\psi}\right\}-z\right]\frac{1}{\beta+\psi}}{D_1}, \tag{5.84}$$

$$V(z)=\frac{pQ\left[\frac{\pi_1\mu_1}{\mu_1+\psi+\alpha}+\frac{\pi_2\mu_2}{\mu_2+\psi+\alpha}\right]\left[\frac{(\psi+\alpha)(\beta+\psi)-\alpha\beta}{\beta+\psi}\right]\frac{1}{\phi+\psi}}{D_1}, \tag{5.85}$$

And

$$D_1=\left\{\frac{(1-p)(\psi+\phi)+p\phi}{\psi+\phi}\right\}\left\{\frac{\pi_1\mu_1}{\mu_1+\psi+\alpha}+\frac{\pi_2\mu_2}{\mu_1+\psi+\alpha}\right\}+\frac{z\alpha\beta}{\beta+\psi}\left[\frac{\pi_1}{\mu_1+\psi+\alpha}+\frac{\pi_2}{\mu_2+\psi+\alpha}\right]-z$$

The probability that the server is idle but available in the system is given by

$$Q=\frac{\left[\frac{\pi_1\mu_1}{\mu_1+\alpha}+\frac{\pi_2\mu_2}{\mu_2+\alpha}\right]\left[1-\frac{p\lambda E(I)}{\phi}\right]-\lambda E(I)\left(1+\frac{\alpha}{\beta}\right)\left[\frac{\pi_1}{\mu_1+\alpha}+\frac{\pi_2}{\mu_2+\alpha}\right]}{\left(1+\frac{\alpha}{\beta}\right)\left(\frac{\pi_1\mu_1}{\mu_1+\alpha}+\frac{\pi_2\mu_2}{\mu_2+\alpha}\right)} \tag{5.86}$$

**Remark**: It can be noted that the results obtained are consistent with the existing literature. For example, if we assume the case where there are no breakdowns and no server vacation, so that $\alpha=0,\Rightarrow R(z)=0$ , $p=0\Rightarrow V(z)=0$ and service is provided in only one mode, i.e., $\pi_1=1,\pi_2=0$, then the above equations in (5.82)-(5.84) reduces to

$$P(z)=\frac{Q(\lambda C(z)-\lambda)}{(\mu+\lambda-\lambda C(z))D_1(z)},\quad D_1(z)=z-\frac{\mu}{(\mu+\lambda-\lambda C(z))}$$

Thus the probability generating function of the queue size is

$$P_q(z) = \frac{Q(\lambda C(z) - \lambda)}{z(\lambda - \lambda C(z) + \mu) - \mu} \quad , \tag{5.87}$$

with the probability of idle state $Q = 1 - \dfrac{\lambda E(I)}{\mu}$ .

Equation (5.87) agrees with the Pollaczeck-Khinchine equation (*Gross & Harris*, *1985*) of the classical $M^X/M/1$ queueing system.

### 5.6.2 Erlang-k Service Time

In this case, the service time has an Erlang-k distribution, then

$g(s) = \dfrac{\mu^k x^{k-1} e^{-\mu x}}{(k-1)!}$, $\mu > 0$, $k = 1,2,3....$ and

the Laplace –transform $G^*(\psi + \alpha) = \left[\dfrac{k\mu}{\psi + \alpha + k\mu}\right]^k$, then the probability generating functions (5.70)-(5.73) reduces to

$$P^{(m1)}(z) = \frac{\pi_1 Q[(\psi+\alpha) - \alpha F^*(\psi)]}{D(z)} \left[ \frac{1 - \left( \dfrac{k\mu_1}{\psi + \alpha + k\mu_1} \right)^k}{\psi + \alpha} \right] \tag{5.88}$$

$$P^{(m2)}(z) = \frac{\pi_2 Q[(\psi+\alpha) - \alpha F^*(\psi)]}{D(z)} \left[ \frac{1 - \left( \dfrac{k\mu_2}{\psi + \alpha + k\mu_2} \right)^k}{\psi + \alpha} \right] \tag{5.89}$$

$$R(z) = \frac{\alpha Q\left[ \dfrac{\pi_1 (k\mu_1)^k}{(\psi + \alpha + k\mu_1)^k} + \dfrac{\pi_2 (k\mu_2)^k}{(\psi + \alpha + k\mu_2)^k} \right] \left[ \{1 - p + pW^*(\psi)\} - z \right]}{D(z)} \left[ \frac{1 - F^*(\psi)}{\psi} \right] \tag{5.90}$$

$$V(z) = \frac{pQ[(\psi+\alpha) - \alpha F^*(\psi)] \left[ \dfrac{\pi_1 (k\mu_1)^k}{(\psi + \alpha + k\mu_1)^k} + \dfrac{\pi_2 (k\mu_2)^k}{(\psi + \alpha + k\mu_2)^k} \right]}{D(z)} \left[ \frac{1 - W^*(\psi)}{\psi} \right] \tag{5.91}$$

and

$$D(z)=\{(1-p)+pW^*(\psi)\}\left\{\pi_1\left(\frac{k\mu_1}{\psi+\alpha+k\mu_1}\right)^k+\pi_2\left(\frac{k\mu_2}{\psi+\alpha+k\mu_2}\right)^k\right\}$$
$$+\frac{z\alpha F^*(\psi)}{\psi+\alpha}\left[1-\left\{\pi_1\left(\frac{k\mu_1}{\psi+\alpha+k\mu_1}\right)^k+\pi_2\left(\frac{k\mu_2}{\psi+\alpha+k\mu_2}\right)^k\right\}\right]-z ,$$

$$Q=(1-p\lambda E(I)E(V))\left\{\pi_1\left(\frac{k\mu_1}{\alpha+k\mu_1}\right)^k+\pi_2\left(\frac{k\mu_2}{\alpha+k\mu_2}\right)^k\right\}$$
$$-\lambda E(I)\left(\frac{1}{\alpha}+E(R)\right)\left\{1-\left[\pi_1\left(\frac{k\mu_1}{\alpha+k\mu}\right)^k+\pi_2\left(\frac{k\mu_2}{\alpha+k\mu_2}\right)^k\right]\right\}$$

Thus adding equations (5.88)-(5.91), $P_q(z)=P^{(m1)}(z)+P^{(m2)}(z)+R(z)+V(z)$, gives the steady state probability generating function of the queue size at random epoch for a $M^X/E_k/1$queue with two modes of service, random breakdowns and Bernoulli schedule server vacations.

## 5.7 Numerical Analysis

In this section we illustrate some numerical results to show the effect of the different parameters on the different states of the system

Example 1: We consider the service times, repair times and vacation time to be Exponentially distributed. All the values are chosen arbitrarily so that the stability condition is not violated.

Arrivals are assumed to be one by one so that $E(I)=1,\ E(I(I-1))=0$. Let us take various values for the rates of service in mode 1 and mode 2 and fix the values of the remaining parameters of rate of arrival, breakdown, repair times, vacation time and probability of taking a vacation as $\lambda=3,\alpha=5,\beta=8,\phi=6,\ p=0.6$.

Table 5.1: Computed Values of Probabilities of Different States of the System for varying values of fluctuating modes of services with $\lambda = 3, \alpha = 5, \beta = 8, \phi = 6, p = 0.6$

| $\mu_1 = 10$ | $\mu_2 = 6$ | $Q$ | $\rho$ | $P^{(m1)}(1)$ | $P^{(m2)}(1)$ | $R(1)$ | $V(1)$ |
|---|---|---|---|---|---|---|---|
| $\pi_1 =$ 0.5 | $\pi_2 = 0.5$ | 0.0407 | 0.9593 | 0.165 | 0.2254 | 0.2694 | 0.3002 |
| = 0.6 | = 0.4 | 0.0600 | 0.9400 | 0.1944 | 0.1766 | 0.2692 | 0.2999 |
| = 0.7 | = 0.3 | 0.079 | 0.921 | 0.2223 | 0.1295 | 0.2693 | 0.3001 |
| $\mu_1 = 7$ | $\mu_2 = 12$ | | | | | | |
| $\pi_1 =$ 0.3 | $\pi_2 = 0.7$ | 0.1340 | 0.8660 | 0.1121 | 0.1843 | 0.2693 | 0.3001 |
| = 0.4 | = 0.6 | 0.1179 | 0.8821 | 0.1521 | 0.1608 | 0.2693 | 0.3001 |
| = 0.5 | = 0.5 | 0.0999 | 0.9001 | 0.1941 | 0.1369 | 0.2693 | 0.3001 |
| $\mu_1 = 9$ | $\mu_2 = 9$ | | | | | | |
| $\pi_1 =$ 0.3 | $\pi_2 = 0.7$ | 0.0976 | 0.9024 | 0.0998 | 0.2332 | 0.2691 | 0.2999 |
| = 0.5 | = 0.5 | 0.0975 | 0.9025 | 0.1665 | 0.1665 | 0.2691 | 0.2999 |
| = 0.6 | = 0.4 | 0.0971 | 0.9029 | 0.2002 | 0.1335 | 0.2692 | 0.2999 |

The above Table 5.1 shows a clear trend that probability of the system in service state depends on the increase or decrease of the two modes. Thus it is observed that fluctuating modes has an effect on the probabilities of service, which in turn, may tend to fluctuate the efficiency of the system. On the other hand, fixed values for the rates of breakdown, repair times and vacation, the probability of the different states of the system remains to be same.

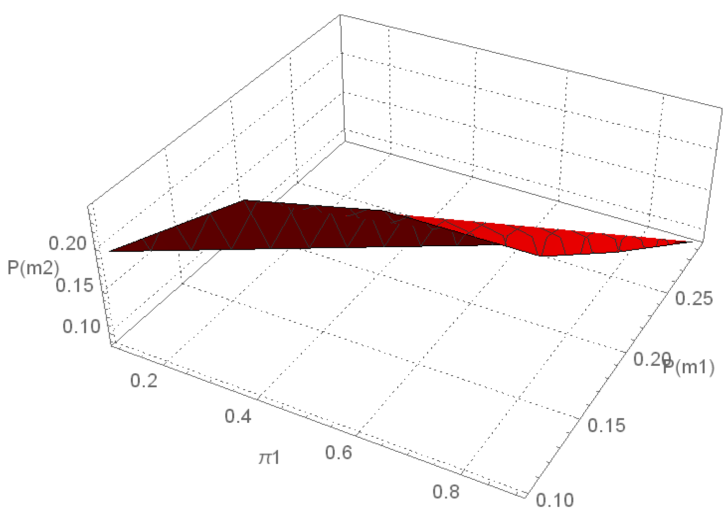


*Figure 5.1: Effect of mode $\pi_1$ on $P^{(m1)}(1)$ and $P^{(m2)}(1)$*

We next compute the different probabilities for varying values of the breakdown parameter α in the following table in order to monitor the effect of breakdown on the queueing system.

Table 5.2: Computed Values of Probabilities of Different States of the System for varying values of breakdown rate α, repair rate β and vacation probability $p$ with $\lambda = 3, \mu_1 = 10, \mu_2 = 6, \pi_1 = 0.6, \pi_2 = 0.4, \phi = 6$

| | | $Q$ | $\rho$ | $P^{(m1)}(1)$ | $P^{(m2)}(1)$ | $R(1)$ | V(1) |
|---|---|---|---|---|---|---|---|
| p=0.6,β=8 | α=3 | 0.1358 | 0.8642 | 0.1903 | 0.1829 | 0.1908 | 0.2999 |
| | α=5 | 0.06 | 0.94 | 0.1944 | 0.1766 | 0.2692 | 0.2999 |
| | α=7 | 0.005 | 0.995 | 0.1986 | 0.1733 | 0.3228 | 0.3024 |
| | | | | | | | |
| α=7, β=8 | p=0.25 | 0.0984 | 0.9016 | 0.197 | 0.1719 | 0.4083 | 0.125 |
| | p=0.4 | 0.0584 | 0.9416 | 0.1972 | 0.1721 | 0.3738 | 0.2002 |
| | p=0.6 | 0.005 | 0.995 | 0.1986 | 0.1733 | 0.3228 | 0.3024 |
| | | | | | | | |
| α=7,p=0.6 | β=8 | 0.005 | 0.995 | 0.1986 | 0.1733 | 0.3228 | 0.3024 |
| | β=10 | 0.0435 | 0.9565 | 0.1973 | 0.1721 | 0.2886 | 0.3004 |
| | β=12 | 0.0739 | 0.9261 | 0.197 | 0.1718 | 0.2579 | 0.2999 |

Example 2: We now illustrate an example considering the service time as a Erlang-$k$ distribution and that the service is active in only one mode such that $\pi_1 = 1, \pi_2 = 0$, the repair time and vacation time are assumed to follow Exponential distribution, such that

$$G^*(\alpha) = \left[\frac{k\mu}{\alpha + k\mu}\right]^k, E(V) = \frac{1}{\phi}, E(R) = \frac{1}{\beta}.$$

Let us further consider that arrivals are in batches following a Geometric distribution, such that

$$C(z) = \sum_{i=1}^{n} c_i z^i = \frac{rz}{1 - z(1-r)};\ C(1) = 1, C'(1) = E(I) = \frac{1}{r}$$

then the probabilities of the different states of the system are

$$P(1) = \frac{Q\frac{\lambda}{r}\left(1 + \frac{\alpha}{\beta}\right)}{\Delta\alpha}\left[1 - \left(\frac{k\mu}{\alpha + k\mu}\right)^2\right],$$

$$R(1) = \frac{Q\frac{\alpha}{\beta}\left[\frac{k\mu}{\alpha + k\mu}\right]^k\left[1 - \frac{p\lambda}{r\phi}\right]}{\Delta},$$

$$V(1)=\frac{Qp\frac{\lambda}{r\phi}\left(1+\frac{\alpha}{\beta}\right)\left(\frac{k\mu}{k\mu+\alpha}\right)^k}{\Delta},$$

where $\Delta=\left\{\left(\frac{k\mu}{\alpha+k\mu}\right)^k\right\}\left(1-\frac{p\lambda}{r\phi}\right)-\frac{\lambda}{r\alpha}\left(1+\frac{\alpha}{\beta}\right)\left\{1-\left(\frac{k\mu}{\alpha+k\mu}\right)^k\right\}$.

The probability of idle state Q is given by

$$Q=\frac{\left[\left(\frac{k\mu}{\alpha+k\mu}\right)^k\right]\left(1-\frac{p\lambda}{r\phi}\right)-\frac{\lambda}{r\alpha}\left(1+\frac{\alpha}{\beta}\right)\left\{1-\left(\frac{k\mu}{\alpha+k\mu}\right)^k\right\}}{\left[\left(\frac{k\mu}{\alpha+k\mu}\right)^k\left(1+\frac{\alpha}{\beta}\right)\right]}$$

For illustration purpose, we fix the values of the vacation parameter $\phi$, service rate α and $p$, probability of taking a vacation. The values of the breakdown parameter α vary from 0.5 to 2.5. All the values are chosen such that the stability condition is satisfied. The following table gives the computed values of the probabilities of the different states of the system.

Table 5.3: Computed values of various queue characteristics for an Erlang-2 queue for varying values of breakdown rate with $\lambda=1.2, \mu=4, p=0.6, \phi=6, E(I)=1.5, k=2$

| $\alpha$ | $\beta$ | $Q$ | $\rho$ | $P(1)$ | $R(1)$ | $V(1)$ |
|---|---|---|---|---|---|---|
| 0.5 | 8 | 0.30770 | 0.69023 | 0.46406 | 0.048235 | 0.2338 |
| 1 | 8 | 0.25076 | 0.74924 | 0.47814 | 0.09111 | 0.18001 |
| 1.5 | 8 | 0.19833 | 0.80167 | 0.49217 | 0.12947 | 0.17999 |
| 2 | 8 | 0.14975 | 0.85025 | 0.50625 | 0.164 | 0.18 |
| 2.5 | 8 | 0.10444 | 0.89556 | 0.52027 | 0.19522 | 0.17998 |

We observe from Table 5.3 that as the rate of breakdown increases the probability of the server in busy state tends to increase. The breakdown rates are small as compared to the fixed repair rate. The probabilities of the system in repair state are seen to be approximately close for various values of the breakdown parameter $\alpha$ due to a fixed repair rate. So having a high repair rate increases the utilization factor. Also the proportion of time in vacation state is almost same. Some plots are demonstrated below showing the behavior of the states due to the effect of breakdown parameter α.

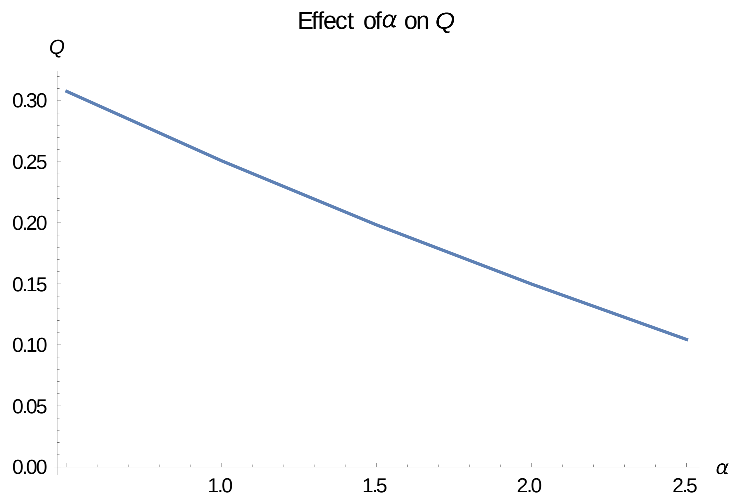


*Figure 5.2: Effect of α on Q*

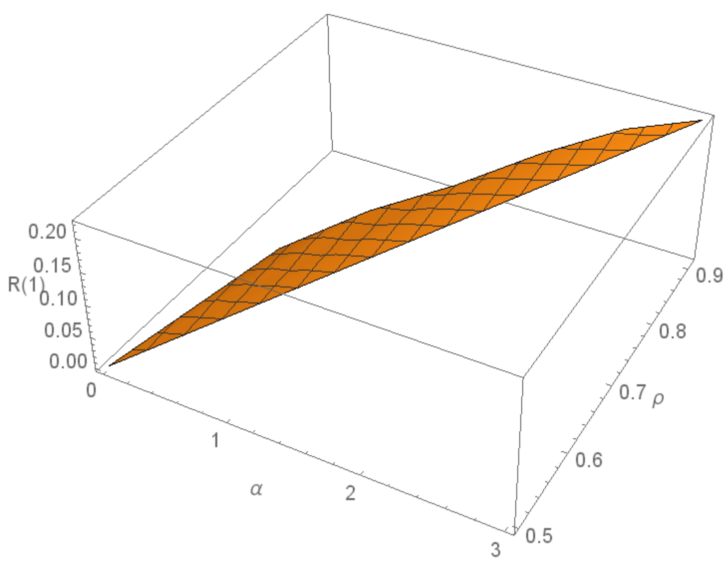


*Figure 5.3: Effect of α on R(1) and ρ*

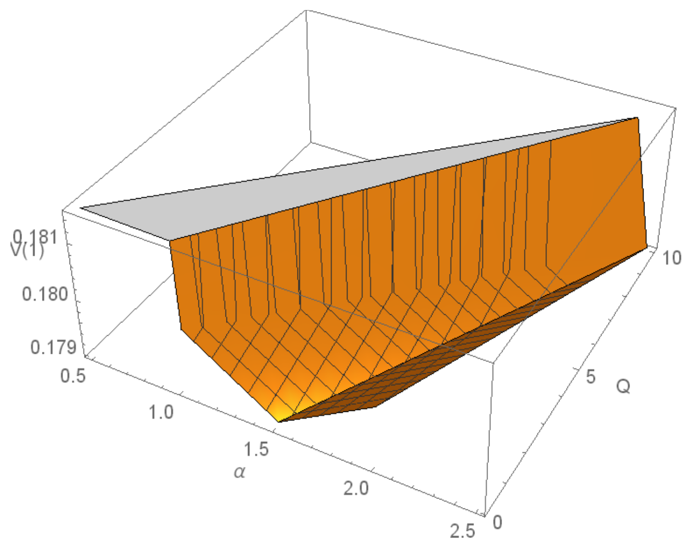


*Figure 5.4: Effect of α on Q and V(1)*

We next compute some queue characteristics for varying values of the repair parameter β and probability of taking a vacation *p* and keep the values of the rest of the parameters fixed. All the values are arbitrarily chosen such that the stability condition is not violated.

Table 5.4: Computed values of various queue characteristics for Erlang-2 queue for varying values of repair rate β and vacation probability $p$ with $\lambda = 0.75, \mu = 4, p = 0.6, \alpha = 3, \phi = 6, E(I) = 1.5, k = 2$

| $\beta$ | $\alpha$ | $p$ | $Q$ | $\rho$ | $P(1)$ | $R(1)$ | $V(1)$ |
|---|---|---|---|---|---|---|---|
| 3 | 3 | 0.25 | 0.14258 | 0.85742 | 0.33398 | 0.47656 | 0.04687 |
| | | 0.6 | 0.10977 | 0.89023 | 0.33404 | 0.44383 | 0.1125 |
| | | 0.8 | 0.09102 | 0.90898 | 0.33397 | 0.42499 | 0.149998 |
| 4 | 3 | 0.25 | 0.21066 | 0.78934 | 0.33397 | 0.40847 | 0.04687 |
| | | 0.6 | 0.17316 | 0.82684 | 0.33396 | 0.38034 | 0.11249 |
| | | 0.8 | 0.15174 | 0.84826 | 0.33398 | 0.36492 | 0.15000 |
| 5 | 3 | 0.25 | 0.26172 | 0.73828 | 0.33397 | 0.35742 | 0.046875 |
| | | 0.6 | 0.22071 | 0.77929 | 0.33398 | 0.33282 | 0.1125 |
| | | 0.8 | 0.19727 | 0.80273 | 0.33397 | 0.31874 | 0.14999 |
| 6 | 3 | 0.25 | 0.30144 | 0.69856 | 0.33399 | 0.31772 | 0.046877 |
| | | 0.6 | 0.25769 | 0.74231 | 0.33398 | 0.29583 | 0.11250 |
| | | 0.8 | 0.23269 | 0.76731 | 0.33396 | 0.28334 | 0.15000 |

In the above Table 5.4, the effect of breakdowns, vacation and repair on the different states of the queueing system is demonstrated. It is seen that increasing the values of the breakdown parameter decreases the probability of the system in idle state as the system will be busy under repairs, which in turn tends to increase the probability of the system in repair state. Similarly an increase in repair parameter shows a decrease in the probability of the system in repair state, which is obvious. The trends shown by the table are as expected. Some plots are illustrated below for representation of the effects of the various parameters involved.

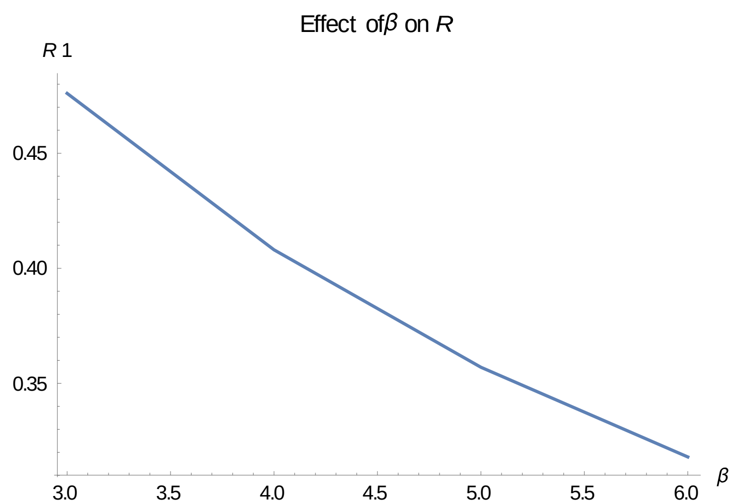


*Figure 5.5: Effect of β on Probability of Repair state.*

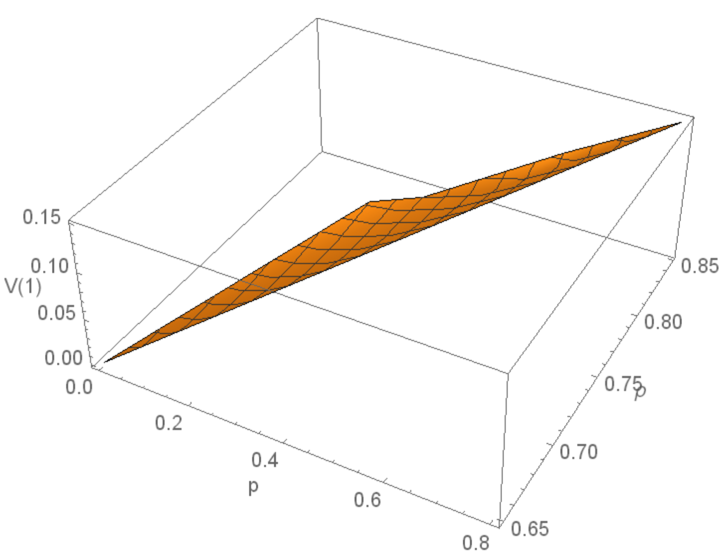


*Figure 5.6: Effect of p on ρ and proportion of time in vacation state.*

## 5.8 Key Results

### 1. Time dependent solution

The Laplace transform of the probability generating function of the queue size is

$$\tilde{P}_q(z,s) = \tilde{P}^{(m1)}(z,s) + \tilde{P}^{(m2)}(z,s) + \tilde{R}(z,s) + \tilde{V}(z,s)\text{, where}$$

$$\tilde{P}^{(mj)}(z,s) = \frac{\pi_j\left[\tilde{Q}(s)\left\{(\psi+\alpha) - \alpha F^*(\psi)\right\} - 1\right]}{D(z,s)}\left[\frac{1-G_j^*(\psi+\alpha)}{\psi+\alpha}\right],\quad j = 1,2$$

$$\tilde{R}(z,s) = \frac{\alpha\left[\tilde{Q}(s)\left\{1-p+pW^*(\psi)\right\} - z\right]A^*(\psi+\alpha) - z\left[\frac{1-A^*(\psi+\alpha)}{\psi+\alpha}\right]}{D(z,s)}\left[\frac{1-F^*(\psi)}{\psi}\right],$$

$$\tilde{V}(z,s) = \frac{p\left[\tilde{Q}(s)\left\{(\psi+\alpha) - \alpha F^*(\psi)\right\} - 1\right]A^*(\psi+\alpha)}{D(z,s)}\left[\frac{1-W^*(\psi)}{\psi}\right],$$

And the denominator, $D(z,s) = \left\{1-p+pW^*(\psi)\right\}A^*(\psi+\alpha) + \dfrac{z\alpha F^*(\psi)\left\{1-A^*(\psi+\alpha)\right\}}{\psi+\alpha} - z$, with

$$\psi = s+\lambda-\lambda C(z),\quad A^*(\psi+\alpha) = \pi_1 G_1^*(\psi+\alpha) + \pi_2 G_2^*(\psi+\alpha)$$

### 2. Steady State Solutions

The probability generating function of the queue size at a random point of time is obtained as

$$P_q(z) = P^{(m1)}(z) + P^{(m2)}(z) + R(z) + V(z)\text{, where}$$

$$P^{(mj)}(z) = \frac{\left[\pi_j Q\left\{(\psi+\alpha) - \alpha F^*(\psi)\right\}\right]}{D(z)}\left[\frac{1-A^*(\psi+\alpha)}{\psi+\alpha}\right],\quad j = 1,2$$

$$R(z) = \frac{\alpha\left[QA^*(\psi+\alpha)\left[1-p+pW^*(\psi)\right] - z\right]}{D(z)}\left[\frac{1-F^*(\psi)}{\psi}\right],$$

$$V(z) = \frac{pQ\left[(\psi+\alpha) - \alpha F^*(\psi)\right]A^*(\psi+\alpha)}{D(z)}\left[\frac{1-W^*(\psi)}{\psi}\right],\quad\text{with}$$

$$D(z) = \left\{1-p+pW^*(\psi)\right\}A^*(\psi+\alpha) + \frac{z\alpha F^*(\psi)\left[1-A^*(\psi+\alpha)\right]}{\psi+\alpha} - z\quad\text{and}$$

$$\psi = \lambda - \lambda C(z), A^*(\psi+\alpha) = \pi_1 G_1^*(\psi+\alpha) + \pi_2 G_2^*(\psi+\alpha).$$

## 5.9 Summary

In this chapter, both the transient and steady state analysis of a batch arrival queueing system with Bernoulli schedule single vacation policy with service in two fluctuating modes has been studied. Time homogeneous breakdowns are assumed to occur. Some special cases are discussed and numerical examples are illustrated to demonstrate the effect of the various assumptions on the behavior of the queueing system.

******

# Chapter 6

# Conclusions and Further Research

## 6.1 Conclusions

This thesis is concerned with investigating the classical batch arrival $M^X/G/1$ queueing system with different types of service facility along with Bernoulli schedule server vacation, where servers take a vacation with probability $p$ or remain in the system with probability $1-p$. We investigate and analyze this model by the addition of some assumptions like balking and reneging, optional re-service, breakdowns, and server vacation thus developing and extending it many directions.

In each of the chapters, there is a combination of more than one new assumption to the general basic model which develops into an advanced queueing system, thus extending it in many directions. In the current research, we aimed at deriving the steady state probability generating function of queue size, solution of some performance measures like the average length of the queue, average waiting time, utilization factor (intensity parameter) and average idle time. The supplementary variable technique has been used to solve the system of equations. The elapsed service time for heterogeneous servers, two stages of service and service in two modes, elapsed vacation time and elapsed repair times has been introduced as supplementary variables. Numerical examples have been presented to demonstrate how the various parameters of the models influence the system.

## 6.2 Contributions of the Study

The contributions of this research have been outlined chapter wise as below:

Chapter Three: A batch arrival single server queueing model with two types of heterogeneous services and two stages of service has been generalized as two separate models by combining assumptions of balking and optional re-service. Steady state probability generating functions for the queue length is obtained and expressed in equations (3.38)-(3.40) and (3.52)-(3.54) for Model 1 and Model 2 respectively. The server takes a vacation for a random period of time and resumes service even if there is none waiting in queue. The probability generating function of queue size at stationary point of time for a batch arrival queue with balking and re-service has been derived in equation (3.61).

Chapter Four: We further generalize the basic $M^X/G/1$ by introducing reneging during server vacations. The closed form expressions for the steady state probability generating functions of the queue size has been derived in both Markovian and non-Markovian service time distribution as two separate models. In case the system is inoperative for some random length of time, it increases the probability of customers' losses due to impatience.

Chapter Five: The basic $M^X/G/1$queueing system studied in Chapter Four is further extended by assuming that service offered oscillates between two fluctuating modes and time homogeneous breakdowns. Breakdowns can occur both in the working and idle state of the server. After any service completion, the server may take a single vacation of random length. During the working state, if the system breaks down, the customer whose service is interrupted returns back to join the head of the queue. Using the elapsed service time, elapsed repair time and elapsed vacation time as supplementary variables, the set of time dependent differential equations were obtained. The solutions to these equations have been obtained in equations (5.59)-(5.62). These equations have been used to obtain steady state solutions given by equations (5.70)-(5.73).

For each of the above-mentioned queueing systems, the necessary and sufficient condition for a system to be stable and some useful performance measures like utilization factor, average queue size, average waiting time in the queue and the probability of the steady state system have been obtained. Some important particular cases have also been discussed in an attempt to connect the current research to the queueing systems studied by earlier researchers. This is done by dropping the assumptions of re-service, balking, breakdowns, and vacations in our results which reduce the system to some models investigated earlier by queueing theorists. Numerical calculations have been done along with presentation of graphs for some of the queueing models. These illustrations demonstrate the effect of queue parameters especially reneging, breakdowns and vacations on different queue measures like mean queue size, mean waiting time, the proportion of server's idle time, utilization factor etc.

The models studied in the current research differ from any other models previously studied due to the combination of assumptions considered in the queueing systems and we do not find any queuing model in the literature that generalizes a queueing system with similar assumptions having diverse service mechanism.

The results obtained in the thesis can be treated as performance evaluation means for the concerned system which may be suited to much congestion situations arising in many practical applications encountered in computer and communications systems, distribution and service sectors, health care systems, production and manufacturing systems etc. The insights provided in this thesis are not only limited to theoretical aspects as a mathematical model but is essentially an approximation of a real process.

## 6.3 Future Work

On the basis of the findings of this research, the following suggestions for further studies can be carried out on the following queueing systems.

- ❖ A batch arrival queueing system with reneging during random breakdowns and server vacations following general arbitrary distributions based on multiple vacation policy. That is the server continues taking vacation unless it finds at least k ($k \geq 1$) units waiting. Queues with multiple vacation policy are more complicated than with single vacation.
- ❖ A batch arrival queue with two types of heterogeneous services with balking and reneging. Balking and reneging can be possibly considered to occur at any time. (In this case Chapter Three would be a special case by dropping the assumption of reneging)
- ❖ A batch arrival vacation queue with balking and reneging during breakdowns and vacation periods with general breakdown and vacation times, general delay times where the system does not immediately enter a repair process but has to wait for a random length of time before being repaired.

- ❖ A two- stage batch arrival queue with an optional second service with system failures and general vacation times. Balking and reneging occurring at any stage of the system. The breakdown rates can be assumed to be different for the two stages of service.
- ❖ A batch arrival vacation queue with two stages of general service where breakdown may occur in one of the stages so that one stage is working and the other is in failed state. In this way, there is fluctuating efficiency, reduced efficiency when one of the stage is working and full efficiency when both the stages are working. We can assume reneging to occur when the server is providing service at only one stage while the other stage is in failed state.

## 6.4 Potential Applications

Starting with a congestion problem in telephone traffic, the range of applications has grown not only to include telecommunications and computer science but also to commercial queueing organizations such as banks, ATMs, post offices etc, transportation service systems like vehicles waiting at toll stations, traffic signals, trucks or ships waiting to be loaded etc. Queueing is also experienced in social service systems like judicial process, hospitals, businesses such as manufacturing industries, inventory systems, etc.

We identify some potential queueing situations following the assumptions of the current research which are batch arrivals, balking and reneging, vacations and breakdowns and some real life situations where queueing theorists have implemented theoretical models.

**Batch Arrivals**: According to Marc Schleyer (2012), there are various applications in the field of logistic and information systems where it is important to know how many units or orders are waiting to be processed. Regarding logistic systems, one of the main reasons considering building a batch is that a specific number of material units are gathered together into one and then taken to a transport carrier, in order to reduce the handling effort. Another typical example of batch arrivals is when a family of four arrives at an airport for boarding a flight. Elevators in building form a common example of these types of systems. In the setting of information technologies, individual information packets are grouped together in larger entities for transmission. These information packets can be modeled as batch arrivals. Queues in many fields are highly stochastic with random arrivals and service (Wang et al. 2014). Mehri et al. (2011)

performed a case study with application of queueing models in a toll motorway. In the study, data on car arrivals and service durations of a motorway was collected and the probability distribution that well described the intervals between arrivals and service durations was determined. They applied the Pollaczek-Khitchinne formula that enabled to derive the performance measures of the M/G/c models. Later they used simulation technique to derive approximate values for equilibrium probabilities on the number of customers present in the system ( $P_n$ ).

**Impatience**:  one of the most important features as customers feel concerned while waiting for some service. When a customer encounters a queue, he often makes a quick estimate of the expected waiting time to decide whether to join the queue or not, based on the amount of time he might be willing to wait. This is mostly experienced when the system has only a single server and customers may decide to balk, that is leave the queue either  without joining or after waiting in the queue for some amount of time( renege). A typical situation is noticed in a communication network where various types of jobs like data files have to wait in queue if they are not served immediately upon arrival in the network. The jobs can become impatient due to high waiting times and uncertainty of receiving service and may leave the queue un-served. This is often referred as Optical Burst Switching (OBS) network. Bocquet (2005) studied the operation of an OBS controller which is seen as a queue with reneging (Selvaraju et al. 2013).  Queues that form to embark on ferries can be typically observed and late arrivals are considered as customers balking. An application of queueing analysis with balking and reneging behavior for passengers waiting for public transport services has been studied by Wang et al. (2014), which relates significantly to the models studied in the current research.

Bus bridging is commonly applied in developed countries in response to severe rail disruptions. An affected train brings a random number of passengers to a bridging station at one time (batch arrivals). Some passengers upon reaching a bridging station may choose to leave immediately without waiting for bridging buses. Some of the remaining passengers may gradually lose their patience and eventually leave before service starts, (balking and reneging), (Wang et al., 2014). Recently, Bai et al. (2016) have done a study on applications of queueing model for managing small projects under uncertainties. They consider projects (customers) arrivals randomly according to a Poisson Process and project processing time as an Exponential distribution because both customer (project) arrivals and completion time of projects are uncertain.

Moreover, they capture customer abandonment dynamics (reneging), where the contractors leave the project unsatisfied before completion. They assumed that each customer randomly reneges at a constant rate α, where α is a measure of customer's abandonment. For tractability, they further assumed that abandonment times are exponentially distributed and independent of service times. This study replicates our model studied in Chapter Four with arrivals of a batch size as one, connecting it to an $M/M/1$ queue with reneging. An M/M/1 queue with reneging will be a particular case of the $M^X/M/1/reneging$ model where reneging occurs during service instead of vacations.

**Vacations and Breakdowns**: It is noticed that various types of service interruptions occur during service in most systems like banks, post offices, supermarkets, doctor's chamber etc., when the server is off for a certain period of time, due to a break or some other secondary services. In many industrial systems, servers are subject to breakdown due to failure of a mechanical or electrical component. Machine breakdowns can occur randomly while the machine is operating. The repair process starts immediately. In a typical production system, it is usually assumed that when a machine (or operator) becomes idle, it remains idle till the next job arrives. However it is sometimes considered as a waste of valuable resources. A remedy is sometimes adopted wherein the machine/operator is scheduled for alternative (secondary) services. As far as the system is concerned, it perceives as if the machine/operator has taken a vacation. When the machine/operator has finished with the alternative services, it becomes available to process the jobs waiting in queue (Gupta et al. 2000). This kind of breakdown may also be observed in computer network server due to any faulty server and for which the server has to go through a repair process. The files waiting need to be served once it is repaired.

Queues with a breakdown as considered in the current research may be highly intractable to model in real life systems as queuing models are developed depending on certain assumptions; however, Taha (1981) stated that for the purpose of applying queuing models to real life situations, the practitioner may take the advantage of the possibility that certain assumptions of the models can be violated or simplified without considerable error in the system's performance measures.

**Potential application of a queueing model**: Health care systems have a complex queueing network. A recent study found that in 2001, 7.7% of 36.6 million adults who sought care in ED reported trouble in receiving emergency care, and that more than half of these cited long waiting times as a cause (Kennedy, 2004). The management of healthcare facilities such as outpatient clinics in a Hospital is also very complex and demanding.

A potential application of one of my models discussed in Chapter Four can be visualized in an outpatient clinic of any local health care centre. Patient arrivals to a medical centre can be occasionally seen to be in bulk (batches) and there may be a preferred 'server' (medical practitioner/doctor) for whom loyal customers may form a queue. We can consider the preferred doctor as a single server. However, there are times when a doctor may need to attend the inpatients in emergency department and so he might not be available for a certain period of time. This temporary unavailability of the server can be considered as a vacation of a random length of time. During this time, some of the patients being impatient due to the discomfort of the illness or any other reason may tend to see an alternate doctor and leave the queue (reneging). So in order to mathematically investigate such a phenomenon in a real life system, we can formulate the model of our queueing system: *A batch arrival queueing system with reneging during server vacations.*

In queuing situations it is necessary to estimate the probability distribution of the pattern of arrivals. Arrivals to a service system usually occur in a random, unpredictable fashion. Even if a good forecast of the total number of arrivals is available, there is often still considerable uncertainty about the precise timing of arrivals. The arrival times between successive patient arrivals can be approximated by the Exponential distribution. One more reason is that the number of customers at an instance has nothing to do with the time since the last arrival (memoryless property). The theory of compound Poisson process can be introduced to establish a stochastic model that applies to the bulk arrivals.

Moreover, it is noticed that there is variability in the service time offered to each patient. The service time durations, the amount of time taken for consultation with the physician are random variables approximated by a probability distribution. For analytically tractable results, an Exponential distribution for service times can be assumed. However, other distributions which are used to model service time of activities in a health care environment include Triangular, Gamma, Log-normal, Weibull distributions or a combination of two or three Exponential terms (Bhattacharjee et al., 2014).

Patients may be discouraged by a long waiting time due to uncomfortable conditions, poor ventilation, noises and the existing illness etc. They might be inclined to leave the queue and possibly check with an alternate doctor. There are different ways to model customer abandonment in the queueing literature. One possible approach is to allow customers to abandon (renege) if the waiting time exceeds a pre-specified time (impatience time). Unfortunately, this analysis is analytically intractable in general. In practice, due to the complexity of the model and lack of appropriate analytical methods for underlying problems, the practitioner has to often resort to simulation methods for analysis (Bai et al. 2016). However, based on the literature on some of the previous works on applications of queuing theory, it can be assumed that a patient reneges independently according to an Exponential distribution. Moreover, the time a server (doctor) is on vacation (checking patients in Emergency departments or inpatients) can be further modeled with an arbitrary distribution. Thus this phenomenon can be expressed mathematically as a $M^X/G/1$ queue with reneging during server vacations studied in Chapter Four of the current research.

For the purpose of an actual implementation of the model, a framework of different steps involved in the process is required. They are (i) preliminary analysis and data collection, (ii) performance analysis and (iii) decision making. The first step involves an understanding and observation of the various processes and services of the system, discussion/interview with management and clinical staff. It includes collection of data to determine the characteristics of a system as arrival pattern of patients, service rates involved in the process of care (Bhattacharjee et al. 2014). The second step is evaluating the performance measures based on the theoretical model. Lastly, based on the outcomes of the performance measures, health care managers can make effective decisions for optimum utilization of resources with quality health care services. As has been found by several researchers analytical queueing models are quite efficient in representing patient flows and performance of hospital systems (Green, 2006). Moreover, (e.g., outpatient systems), the only limitation is that many health care environments system may not usually reach a steady state, and hence the use of steady state results might not lead to a perfect mathematical investigation of the system.

However, for the purpose of implementing queueing models into real life systems, it is conducive to convert the real life system into a simplified representative containing model description and analysis. Then either using existing models or developing new ones, tools, and algorithms, the queueing system is analyzed and performance measures evaluated. Finally, a decision making process uses the performance measures in optimization control, design and operation process for improving the efficiency of the system.

## 6.5 Limitations

While there is no doubt about the advantages of using queueing theory to model the behavior of a queuing system in order to optimize its performance and service levels, there are also certain limitations in the theory. This is because the classical queueing theory may be too restrictive to model real life situations. Also, we rely solely on the assumptions of queueing theory to model the system and to arrive at solutions. However, it is a growing area and the use of more complex models to solve the challenges of modern day queueing theory is robust enough to provide close enough approximations. Because of the complexity of the mathematics that is involved in the theory, queue modeling is often seen as a challenging research

******

# Appendix A

## A. Explanation of Equations

### A.1.1 Equation (3.2) and (3.4)

$$P_{n,j}(x+\Delta x) = (1-\lambda\Delta x)(1-\mu_j(x)\Delta x)P_{n,j}(x) + (1-b_1)(\lambda\Delta x)P_{n,j}(x)$$
$$+ b_1\lambda\Delta x\sum_{i=1}^{n} c_i P_{n-i,j}(x) \qquad ; \quad j=1,2; \quad n \geq 1$$

Connecting the system probabilities at *x* with $(x+\Delta x)$ by considering $P_{n,j}(x+\Delta x)$ which means there are $n\,(n \geq 0)$ customers in the queue excluding one customer in type j (j=1,2) service and the elapsed service time of the customer is $(x, x+\Delta x)$. Then all possible mutually exclusive transitions that can occur during a short interval of time $(x, x+\Delta x]$ are as follows:

(i) There are *n* customers in the queue excluding one customer in type j (j=1, 2) service and elapsed service time is *x*, no arrivals and no service completion during (x, x+Δx]. The joint probability for these events is given by $(1-\lambda\Delta x)(1-\mu_j(x)\Delta x)P_{n,j}(x)$.

(ii) There are *n* customers in the queue excluding the one in service since the elapsed service time is *x* and an arriving batch balks with probability $(1-b_1)$. The joint probability for this case is $(1-b_1)\lambda\Delta x P_{n,j}(x)$.

(iii) There are *(n-i)* customers in the queue excluding the one in type j(j=1, 2) service since the elapsed service time is *x*, a batch of size *i* customers arrive and join the queue with probability $b_1$.The joint probability for this case is: $b_1\lambda\Delta x\sum_{i=1}^{n} c_i P_{n-i,j}(x)$.

### A.1.2 Equation (3.3)

$$P_{0,j}(x+\Delta x) = (1-\lambda\Delta x)(1-\mu_j(x)\Delta x)P_{0,j}(x) + (1-b_1)(1-\lambda\Delta x)\mu_j(x)P_{1.j}(x); \quad j=1,2$$

Connecting the system probabilities at *x* with $(x+\Delta x)$ by considering $P_{0,j}(x+\Delta x)$ which means there are no customers in the queue during $(x, x+\Delta x)$. Then all possible mutually exclusive transitions that can occur during a short interval of time $(x, x+\Delta x]$ are as follows:

(i) There are no customers in the queue, no arrivals and no service completion during (x, x+Δx]. The joint probability for these events is given by $(1-\lambda\Delta x)(1-\mu_j(x)\Delta x)P_{0j}(x)$.

(ii) There is 1 customer in service of type j, service is completed and the elapsed service time of this customer is $x$ and an arriving batch balks with probability $(1-b_1)$. The joint probability for this case is $(1-b_1)\lambda\Delta x P_{1,j}(x)$.

The rest of differential-difference equations (3.4)-(3.7) in the chapter follow similar reasoning, with regard to the different states of the system.

### A.1.3 Boundary Condition (3.9)

$$P_{n,1}(0)=(1-p)\xi_1\left[\int_0^\infty P_{n+1,1}(x)\mu_1(x)dx+\int_0^\infty P_{n+1,2}(x)\mu_2(x)dx\right]+\xi_1\int_0^\infty V_n(x)\phi(x)dx+\lambda b_1\xi_1 c_{n+1}Q;\ \ n\geq 0$$

$P_{n,1}(0)=$ Probability that at time 0, there are $n\,(n\geq 0)$ customers in the queue excluding the customer in the service of type1 and elapsed service time of this customer is 0 (service started in type 1). Then we have the following mutually exclusive cases:

(i) There are $(n+1)$ customers in the queue excluding the one being served in service of type 1 or type 2 given that elapsed service time is $x$, server completes service of this customer, does not go for vacation and starts serving the next customer in type 1 with probability $\xi_1$. This is given by the probability

$$(1-p)\xi_1\left(\int_0^\infty P_{n+1,1}(x)\mu_1(x)dx+\int_0^\infty P_{n+1,2}(x)\mu_2(x)dx\right)$$

(ii) There are $(n+1)$ customers in the queue and server is on vacation given that the elapsed vacation time is $x$. Vacation just completed and starts serving the next customer in type 1 with probability $\xi_1$.This case has the probability

$$\xi_1\int_0^\infty V_{n+1,1}(x)\phi(x)dx$$

(iii) There are no customers in the system, server is idle but available in the system, a batch of size $(n+1)$ customers arrive and decides to join the queue with probability $b_1$. This case has the probability $\lambda\xi_1 b_1 c_{n+1} Q$.

Equation (3.10) has the same explanation.

**A.1.4 Boundary Conditions (3.11)**

$$V_n(0) = p\left[\int_0^\infty P_{n,1}(x)\mu_1(x)dx + \int_0^\infty P_{n,.2}(x)\mu_2(x)dx\right] \quad n \geq 0$$

$V_n(0) =$ Probability that at time 0, there are $n(n \geq 0)$ customers in the queue and server is on vacation with elapsed vacation time 0. (Vacation has just started). This has the following mutually exclusive cases:

(i) There are $n(\geq 0)$ customers in the queue excluding one being served in type 1 or type 2, given that elapsed service time of this customer is *x*, service is completed and goes for vacation with probability *p*. This has the probability $p\left[\int_0^\infty P_{n,1}(x)\mu_1(x)dx + \int_0^\infty P_{n,2}(x)\mu_2(x)dx\right]$.

**A.1.5 Equation (3.56)**

$$Q = (1-\lambda\Delta x)Q + \lambda\Delta x(1-b_1)Q + (1-p)\int_0^\infty P_0^{(2)}(x)\mu_2(x)\Delta x dx + \int_0^\infty V_0(x)\phi(x)\Delta x dx$$

$Q =$ There are no customers in the system and the server is idle but available in the system. This has the following mutually exclusive cases:

(i) The system is idle but available and there are no arrivals. This has probability $(1-\lambda\Delta x)Q$

(ii) There are no customers in the system, an arriving batch balks with probability $(1-b_1)$ and the server is idle but available. This has probability $\lambda\Delta x(1-b_1)Q$.

(iii) There are no customers in the queue excluding one being served in stage 2, service of this customer is completed with elapsed service time *x*, does not take vacation, thus decides to stay in the system. This case has probability $(1-p)\int_0^\infty P_0^{(2)}(x)\mu_2(x)\Delta x dx$.

(iv) There are no customers in the queue and the server is on vacation with elapsed vacation time $x$. The vacation period is completed and the server is in idle state. This case has probability $\int_0^\infty V_0\,(x)\phi(x)\Delta x dx$.

### A.1.6 Boundary Condition (3.57)

$$P_n^{(1)}(0) = \lambda b_1 c_{n+1} Q + (1-p)\int_0^\infty P_{n+1}^{(2)}(x)\mu_2(x)dx + \int_0^\infty V_{n+1}(x)\phi(x)dx, \quad n \geq 0,$$

$P_n{}^{(1)}(0) =$ Probability that at time 0, there are $n$ customers in the queue excluding one customer in first stage service, given elapsed service time of this customer is 0 (service just started .Then we have three cases

(i) At time $t$, there are (n+1) customers in the queue excluding the customer being served in second stage given elapsed service time of this customer is $x$, second stage service completed, does not go for vacation and starts serving the next customer in first stage. This is given by $(1-P)\int_0^\infty P_{n+1}{}^{(2)}(x)\mu_2(x)dx$

(ii) There are $(n+1)$ customers in the queue and server is on vacation given that the elapsed vacation time is $x$. Vacation just completed and starts serving the next customer in first stage. This case has the probability $\int_0^\infty V_{n+1}(x)\phi(x)dx$

(iii) There are no customers in the system, server is idle but available in the system, a batch of size $(n+1)$ customers arrive and decides to join the queue with probability $b_1$. This case has the probability $\lambda b_1 c_{n+1} Q$ .

### A.1.7 Boundary Condition (3.58)

$$P_n^{(2)}(0) = \int_0^\infty P_n^{(1)}(x)\mu_1(x)dx \qquad n \geq 0,$$

$P_n^{(2)}(0) =$ At time 0, there are $n$ customers in the queue, excluding one customer in first stage service; service is completed, with elapsed service time 0 and starts serving the next customer in stage two. This case has the following probability $\int_0^\infty P_n^{(1)}(x)\mu_1(x)dx$

### A.1.8 Boundary Condition (3.59)

$$V_n(0) = p\int_0^\infty P_n^{(2)}(x)\mu_2(x)dx \qquad n \geq 0,$$

$V_n(0) =$ Probability that at time 0, there are $n$ customers in the queue, the server is on vacation with elapsed vacation time 0, this means vacation just started.

Then this state happens with the following probability. That is there are $n$ (n ≥0) customers in the queue excluding one customer in stage two, service of this customer is completed with elapsed service time $x$ and vacations just started.

### A.1.9 Equation (4.1)

$$P_n(t+\Delta t) = (1-\lambda\Delta t)(1-\mu\Delta t)P_n(t) + (1-p)(1-\lambda\Delta t)\mu\Delta t P_{n+1}(t) + (1-\mu\Delta t)\lambda\Delta t\sum_{i=1}^{n} c_i P_{n-i}(t)$$
$$+\phi\Delta t(1-\lambda\Delta t)V_n(t); \qquad n \geq 1$$

Connecting the system probabilities at time $t$ with $(t+\Delta t)$ by considering $P_n(t+\Delta t)$ which means there are $n$ customers in the queue. Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) There are $n$ customers in the queue at time $t$, there are no arrivals and no service completed during (t, t+Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)(1-\mu\Delta t)P_n(t)$ .

(ii) There are $(n+1)$ customers in the queue at time $t$, a service is completed and there are no arrivals and the server does not take vacation. The joint probability is given by $(1-p)(1-\lambda\Delta t)\mu\Delta t P_{n+1}(t)$

(iii) At time t, there are $(n-i)$ customers, there is no service completed and a batch of '$i$' customers arrives. This is given by $(1-\mu\Delta t)\lambda\Delta t\sum_{i=1}^{n} c_i P_{n-i}(t)$ .

(iv) At time t, there are $n$ customers in the queue, there are no arrivals and the server is in vacation state. This joint probability is given as $(1-\lambda\Delta t)\phi\Delta t V_n(t)$

### A.2.0 Equation (4.2)

$$P_0(t+\Delta t) = (1-\lambda\Delta t)P_0(t) + (1-p)\mu\Delta t P_1(t) + \phi\Delta V_0(t)$$

Connecting the system probabilities at time $t$ with $(t+\Delta t)$ by considering $P_0(t+\Delta t)$ which means there are no customers in the queue. Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) There are no customers in the queue at time t and there are no arrivals, the probability is given by $(1-\lambda\Delta t)P_0(t)$.

(ii) There is one customer in the queue at time *t*, the server does not go for vacation and service of the customer is completed. This joint probability of these events is given by $(1-p)\mu\Delta t P_1(t)$.

(iii) There are no customers in the queue at time *t* and server is in vacation state. The joint probability is given by $\phi\Delta t V_0(t)$.

### A.2.1 Equation (4.3)

$$V_n(t+\Delta t) = (1-\lambda\Delta t)(1-\phi\Delta t)(1-n\gamma\Delta t)V_n(t) + (1-\phi\Delta t)(1-n\gamma\Delta t)\lambda\Delta t\sum_{i=1}^{n} c_i V_{n-i}(t)$$
$$+ (1-\lambda\Delta t)(1-\phi\Delta t)(n+1)\gamma\Delta t V_{n+1}(t) + p(1-\lambda\Delta t)\mu\Delta t P_{n+1}(t)$$

Connecting the system probabilities at time $t$ with $(t+\Delta t)$ by considering $V_n(t+\Delta t)$ which means there are $n$ customers in the queue and the server is on vacation. Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) There are $n$ customers in the queue at time $t$, the server is on vacation, there are no arrivals, vacation period not completed and no reneging during (t, t+Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)(1-\phi\Delta t)(1-n\gamma\Delta t)V_n(t)$.

(ii) There are $(n-i)$ customers in the queue at time $t$, the server is on vacation, vacation period not completed and no reneging occurs, a batch of 'i' customers arrive with probability $\lambda c_i$ during (t,t+Δt]. The joint probability for these events is given by $\lambda\Delta t(1-n\gamma\Delta t)\sum_{i=1}^{n} c_i V_{n-i}(t)$.

(iii) There are $(n+1)$ customers in the queue at time *t*, the server is on vacation, no arrivals and no vacation completion. A customer reneges, during (t, t+Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)(1-\phi\Delta t)(n+1)\gamma\Delta t V_{n+1}(t)$.

(iv) There are $(n+1)$ customers in the queue at time *t*, a service is completed and there are no arrivals during (t, t+Δt].The server goes for vacation after completion of the service with probability *p*. The joint probability for these events is given by $p(1-\lambda\Delta t)\mu\Delta t P_{n+1}(t)$.

### A.2.2. Equation (4.4)

$$V_0(t+\Delta t) = (1-\lambda\Delta t)(1-\phi\Delta t)V_0(t) + (1-\lambda\Delta t)\gamma\Delta t V_1(t) + p(1-\lambda\Delta t)\mu\Delta t P_1(t)$$

Connecting the system probabilities at time *t* with $(t+\Delta t)$ by considering $V_0(t+\Delta t)$ which means there are no customers in the queue and the server is on vacation. Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) There are no customers in the queue at time *t*, the server is on vacation, and there are no arrivals, vacation period not completed during (t, t + Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)(1-\phi\Delta t)V_0(t)$.

(ii) There is 1 customer in the queue at time *t*, the server is on vacation, no arrivals and the customer reneges, during (t, t+Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)\gamma\Delta t V_1(t)$.

(iii) There is one customer in the service at time *t*, service is completed and there are no arrivals during (t, t+Δt].The server goes for vacation after completion of the service. The joint probability for these events is given by $(1-\lambda\Delta t)\mu\Delta t P_1(t)$.

The differential equation (4.35) can be similarly explained with the same reasoning as in A.2.1 and A.2.2.

### A.2.3 Equation (4.36)

$$V_n(t+\Delta t) = (1-\lambda\Delta t)(1-\phi\Delta t)(1-n\gamma\Delta t)V_n(t) + (1-\phi\Delta t)\lambda\Delta t\sum_{i=1}^{n} c_i V_{n-i}(t) + (1-\lambda\Delta t)(1-\phi\Delta t))(n+1)\gamma V_{n+1}(t)$$

$$+\, p\int_0^\infty P_n(x)\mu(x)dx, \quad n \geq 1$$

Connecting the system probabilities at time *t* with $(t+\Delta t)$ by considering $V_n(t+\Delta t)$ which means there are *n* customers in the queue and the server is on vacation. Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) There are *n* customers in the queue at time *t*, the server is on vacation, there are no arrivals, vacation period not completed and no reneging during (t, t+Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)(1-\phi\Delta t)(1-n\gamma\Delta t)V_n(t)$.

(ii) There are $(n-i)$ customers in the queue at time *t*, the server is on vacation, vacation period not completed and reneging occurs, a batch of customers arrive with probability $\lambda c_i$ during (t,t+Δt]. The joint probability for these events is given by $\lambda\Delta t(1-n\gamma\Delta t)\sum_{i=1}^{n} c_i V_{n-i}(t)$.

(iii) There are $(n+1)$ customers in the queue at time *t*, the server is on vacation, no arrivals and a customer reneges, during (t, t+Δt]. The joint probability for these events is given by $(1-\lambda\Delta t)(n+1)\gamma\Delta t V_{n+1}(t)$.

(iv) There are *n* customers in the queue at time *t* excluding one customer in service, service is completed with elapsed service time *x*. The server goes for vacation after completion of the service with probability *p*. The joint probability for these events is given by $p\int_0^\infty P_n(t)\mu(x)dx$.

Equation (4.37) can be explained with a similar reasoning as in A.2.2.

**A.2.4 Equation (4.38)**

$$Q = (1-\lambda)Q + (1-p)\int_0^\infty P(x)\mu(x)dx + \phi V_0$$

$Q =$ There are no customers in the system and the server is idle but available in the system. This has the following mutually exclusive cases:

(i) The system is idle but available and there are no arrivals. This has probability $(1-\lambda)Q$

(ii) There are no customers in the queue excluding one being served, service is completed with elapsed service time *x*, does not take vacation, thus decides to stay in the system. This case has probability $(1-p)\int_0^\infty P_0(x)\mu(x)dx$.

(iii) The server is in vacation and there are no customers in the queue. The vacation period is completed and the server is in idle state. This case has probability $\phi V_0$

The same reasoning can be considered to explain the remaining equations in Chapter Four, bearing in mind the different states of the server in each equation.

**A.2.5. Equation (5.1)**

$$P_n^{(mj)}(x+\Delta x, t+\Delta t) = (1-\lambda\Delta t)(1-\mu_j(x)\Delta x)(1-\alpha\Delta t)P_n^{(mj)}(x.t) + (1-\mu\Delta x)(1-\alpha\Delta t)\lambda\Delta t\sum_{i=1}^{n} c_i P_{n-i}^{(mj)}(x,t);\ n \geq 1,\ j = 1,2$$

Connecting the system probabilities at time *t* with $(t+\Delta t)$ by considering $P_n(t+\Delta t, x+\Delta x)$ which means there are *n* customers in the queue and the server is active in mode j, (j=1,2). Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) There are *n* customers in the queue at time *t*, there are no arrivals, no service completion, no breakdowns during the time interval $(t, t+\Delta t]$ .The joint probability is given by $(1-\lambda\Delta t)(1-\mu(x)\Delta x)(1-\alpha\Delta t)P_n^{(mj)}(x,t)$ .

(ii) At time t, there are $(n-i)$ customers in the queue, there is no service completed, no breakdowns and a batch of 'ï' customers arrives. This is given by

$$(1-\mu_j(x)\Delta x)(1-\alpha\Delta t)\lambda\Delta t\sum_{i=1}^{n} c_i P_{n-i}{}^{(mj)}(x,x)$$

Re-arranging the terms and dividing throughout by Δ and taking Δ→0, we arrive at difference-differential equation (5.1). The same analogy follows for the rest of the equations.

Equations (5.2)-(5.4) follow the same reasoning.

### A.2.6 Equation (5.5)

$$R_n(x+\Delta x,.t+\Delta t) = (1-\lambda\Delta t)(1-\beta(x)\Delta x)R_n(x.t) + (1-\beta\Delta x)\lambda\Delta t\sum_{i=1}^{n} c_i R_{n-i}(x,t);\ n \geq 1,$$

Connecting the system probabilities at time *t* with $(t+\Delta t)$ by considering $R_n(t+\Delta t, x+\Delta x)$ which means there are *n* customers in the queue and the server is under repairs due to breakdown. Then all possible mutually exclusive transitions that can occur during a short interval of time $(t, t+\Delta t]$ are as follows:

(i) At time *t*, there are *n* customers in the queue and the server is under repairs due to breakdowns. There is no arrivals and no repair completion, during the time $(t, t+\Delta t]$. This joint probability is given by $(1-\lambda\Delta t)(1-\beta(x)\Delta x)R_n(x,t)$.

(ii) At time *t*, there are $(n-i)$ customers in the queue and the server is under repairs due to breakdowns. No repairs completed and a batch of 'i' customers arrive during the time $(t, t+\Delta t]$. This is given by $(1-\beta\Delta(x))\lambda\Delta\sum_{i=}^{n} c_i R_{n-i}(x,t)$.

Equations (5.6)-(5.8) can be similarly explained.

### A.2.7. Equation (5.9)

$$Q(t+\Delta t) = (1-\lambda\Delta t)(1-\alpha\Delta t)Q(t) + (1-p)\left[\int_0^\infty P_0^{(m1)}(x,t)\mu_1(x)dx + \int_0^\infty P_o^{(m2)}(x,t)\mu_2(x)dx\right]$$
$$+\int_0^\infty R_0(x,t)\beta(x)dx + \int_0^\infty V_0(x,t)\phi(x)dx$$

$Q(t+\Delta)$ = There are no customers in the system and the server is idle but available in the system. This has the following mutually exclusive cases:

(i) At time $t$, the system is idle but available, there are no arrivals and no breakdowns. This has probability $(1-\lambda\Delta x)(1-\alpha\Delta t)Q(t)$

(ii) There are no customers in the queue excluding one being served in mode 1 or mode 2, service of this customer is completed with elapsed service time $x$, does not take vacation, thus decides to stay in the system. This case has probability $(1-p)\left[\int_0^\infty P_0^{(m1}(x,t)\mu_1(x)dx + \int_0^\infty P_0^{(m2)}(x,t)\mu_2(x)dx\right]$.

(iii) There are no customers in the queue and the server is under repairs, with elapsed repair time $x$. The repair period is completed and server is in idle state. This case has probability $\int_0^\infty R_0(x,t)\beta(x)dx$

(iv) There are no customers in the queue and the server is on vacation with elapsed vacation time $x$. The vacation period is completed and the server is in idle state. This case has probability $\int_0^\infty V_0(x,t)\phi(x)\Delta x dx$.

**A.2.8.Boundary condition (5.11)**

$$P_n^{(m1)}(0,t) = \pi_1(1-p)\left[\int_0^\infty P_{n+1}^{(m1)}(x,t)\mu_1(x)dx + \int_0^\infty P_{n+1}^{(m2)}(x,t)\mu_2(x)dx\right]$$
$$+\pi_1\int_0^\infty R_{n+1}(x,t)\beta(x)dx + \pi_1\int_0^\infty V_{n+1}(x,t)\phi(x)dx + \pi_1\lambda c_{n+1}Q(t), \qquad n \geq 0,$$

$P_n^{(m1)}(0,t) =$ Probability that at time *t*, there are *n* customers in the queue excluding one customer in service in mode 1, given elapsed service time of this customer is 0 (service just started) .Then we have three cases:

(i) At time *t*, there are (n+1) customers in the queue excluding the customer being served in mode 1 or in mode 2 given elapsed service time of this customer is *x*, service completed, does not go for vacation and starts serving the next customer in mode 1 with probability $\pi_1$ . This is given by $\pi_1(1-p)\left[\int_0^\infty P_{n+1}^{(m1)}(x,t)\mu_1(x)dx + \int_0^\infty P_{n+1}^{(m2)}(x,t)\mu_2(x)dx\right]$

(ii) There are $(n+1)$ customers in the queue at time *t* and server is under repairs given that the elapsed repair time is *x*. Repair is just completed and starts serving the next customer in mode 1 with probability $\pi_1$. This case has the probability $\pi_1\int_0^\infty R_{n+1}(x,t)\beta(x)dx$

(iii) There are $(n+1)$ customers in the queue at time *t* and server is on vacation given that the elapsed vacation time is *x*. Vacation just completed and starts serving the next customer in mode 1. This case has the probability $\pi_1\int_0^\infty V_{n+1}(x,t)\phi(x)dx$

(iv) At time *t*, there are no customers in the system, server is idle but available in the system, a batch of size $(n+1)$ customers arrive and decides to join the queue. This case has the probability $\lambda\pi_1 c_{n+1}Q(t)$ .

Equation (5.12) has the same reasoning.

### A.2.9. Boundary condition (5.13)

$$R_{n+1}(0,t) = \alpha \int_0^\infty P_n^{(m1)}(x,t)dx + \alpha \int_0^\infty P_n^{(m2)}(x,t)dx, \qquad n \geq 0,$$

$R_{n+1}(0,t) =$ Probability that at time $t$, there are *(n+1)* customers in the queue and the server is under repairs due to breakdowns, given elapsed service time of this customer is 0 (repairs just started). Then we have the following two cases

(i) There are $n$ customers in the queue at time $t$, excluding the customer served in mode 1 given elapsed service time is $x$ at the moment when server breaks down i.e. during the service of this customer. This case has probability $\alpha \int_0^\infty P_n^{(m1)}(x,t)dx$

(ii) There are $n$ customers in the queue at time $t$, excluding the customer served in mode2 given elapsed service time is $x$ at the moment when server breaks down i.e. during the service of this customer. This case has probability $\alpha \int_0^\infty P_n^{(m2)}(x,t)dx$

### A.3.0. Equation (5.14)

$$R_0(0,t) = \alpha Q(t),$$

$R_0(0,t) =$ Probability that at time $t$, there are no customers in the queue and at the moment the server is idle, the server breakdowns and repairs start. This is given by $\alpha Q(t)$.

Equation (5.15) has a similar reasoning with respect to server vacations explained in the previous chapters.

# Appendix B

### B.1.1 Details of obtaining $g(x)$

$$\mu(x) = \frac{g(x)}{1-G(x)}$$

$$\int_0^x \mu(u)du = -\ln(1-G(u))]_0^x$$

$$-\int_0^x \mu(u)du = \ln\left(\frac{1-G(x)}{1-G(0)}\right), \; G(0)=0$$

$$e^{-\int_0^x \mu(u)du} = 1-G(x)$$

$$-\mu(x)e^{-\int_0^x \mu(u)du} = -g(x)$$

$$g(x) = \mu(x)e^{-\int_0^x \mu(u)du}$$

### B.1.2 Details of integration by parts

$$P(x,z) = P(0,z)e^{-mx-\int_0^x \mu(t)dt}$$

$$\int_0^\infty P(x,z)dx = \int_0^\infty P(0,z)e^{-mx-\int_0^x \mu(t)dt}dx$$

Let $u = e^{-\int_0^x \mu(t)dt} \Rightarrow du = -\mu(x)e^{-\int_0^x \mu(t)dt}dx$

Also let $dv = e^{-mx} \Rightarrow v = -\frac{e^{-mx}}{m}$

Then we have $P(0,z)\int_0^\infty udv$ .

$$\int_0^\infty u\,dv = uv - \int_0^\infty v\,du \Rightarrow \left[\frac{-e^{-mx-\int_0^x \mu(t)dt}}{m}\right] - \int_0^\infty \left(\frac{e^{-mx}}{m}\right)\mu(x)e^{-\int_0^x \mu(t)dt}dx$$

Now $$= \frac{1}{m} - \frac{1}{m}\int_0^\infty e^{-mx}\mu(x)e^{-\int_0^x \mu(t)dd}dx = \frac{1}{m} - \frac{1}{m}\int_0^\infty e^{-mx}dG(x)$$

$$= \frac{1}{m}\left[1 - G^*(m)\right]$$

$$\therefore \int_0^\infty P(x,z)dx = P(o,z)\frac{\left[1-G^*(m)\right]}{m}$$

### B.1.3 Details of obtaining equation (3.75):

$$L_q = \frac{d}{dz}P_q(z)\Big|_{z=1}$$

$$P_q(z) = \frac{N(z)}{D(z)} \quad \left(N(1) = 0, D(1) = 0\right)$$

$$\frac{dP_q(z)}{dz} = \frac{D(z)N'(z) - N(z)D'(z)}{\left(D(z)\right)^2} = \frac{0}{0}$$

Applying L'Hopital's Rule for the first time

$$\frac{d\left[D(z)N'(z) - N(z)D'(z)\right]}{d\left(D(z)^2\right)} = \frac{D(z)N''(z) - N(z)D''(z)}{2D(z)D'(z)} = \frac{0}{0}$$

Applying L'Hopital's Rule twice

$$\frac{d\left[D(z)N''(z) - N(z)D''(z)\right]}{d\left(2D(z)D'(z)\right)} = \frac{D'(z)N''(z) + D(z)N'''(z) - N'(z)D''(z) - N(z)D'''(z)}{2\left[D'(z)D'(z) + D(z)D''(z)\right]}$$

Thus

$$L_q = \lim_{z\to 1}\frac{D'(z)N''(z) - N'(z)D''(z)}{2\left(D'(z)\right)^2}$$

### B.1.4. Details of verification of a single root in equation (4.47)

$$f(z) = z - (1-p)G^*(\lambda - \lambda C(z))$$

Let us take $\eta(z) = z$ and $\xi(z) = -(1-p)G^*(\lambda - \lambda C(z))$

It is seen that $\eta(z)$ has one zero inside the unit circle $|z| \le 1$

Now $|\eta(z)| = |z| = 1$ and $|\xi(z)| = \left|-(1-p)G^*(\lambda - \lambda C(z))\right| = (1-p)\left|G^*(\lambda - \lambda C(z))\right|$

Now since $(1-p)$ <1 and $\left|G^*(\lambda - \lambda C(z))\right| \le 1$, being a Laplace-Stieltjes transform. Therefore $|\xi(z)|$<1.

So by Rouche's theorem, since $|\eta(z)| > |\xi(z)|$, then $f(z) = \eta(z) + \xi(z)$ has the same number of zeros as $\eta(z)$. Hence there exists only a single root for the function $f(z)$.

### B.1.5 Details of equation (5.64)

$$\tilde{P}_q(1.s) + \tilde{Q}(s) = \frac{1}{s}$$

LHS:

$$\frac{\tilde{P}^{(m1)}(1,s) + \tilde{P}^{(m2)}(1,s) + \tilde{R}(1,s) + \tilde{V}(1,s) + Q\tilde{D}(1.s)}{D(1,s)}$$

$$\tilde{P}^{(m1)}(1,s) = \frac{\pi_1\left[\tilde{Q}\left((s+\alpha) - \alpha F^*(s)\right) - 1\right]}{D(1,s)}\left[\frac{1-G_1^*(s+\alpha)}{s+\alpha}\right]$$

$$\tilde{P}^{(m2)}(1,s) = \frac{\pi_2\left[\tilde{Q}\left((s+\alpha) - \alpha F^*(s)\right) - 1\right]}{D(1,s)}\left[\frac{1-G_2^*(s+\alpha)}{s+\alpha}\right]$$

$$\tilde{R}(1,s) = \frac{\alpha\begin{bmatrix}\tilde{Q}\begin{bmatrix}\{(1-p)+pW^*(s)\}\{A\} \\ -A\end{bmatrix} \\ -\left\{\dfrac{\pi_1(1-G_1^*(s+\alpha))}{s+\alpha} + \dfrac{\pi_2(1-G_2^*(s+\alpha))}{s+\alpha}\right\}\end{bmatrix}}{D(1,s)}\left[\frac{1-F^*(s)}{s}\right]$$

$$\tilde{V}(1,s)=\frac{p\left[\tilde{Q}\{(s+\alpha)-\alpha F^{*}(s\_\}A-A\right]}{D(1,s)}\left[\frac{1-W^{*}(s)}{s}\right]$$

$$\tilde{D}(1,s)=\{(1-p)+pW^{*}(s)\}\{A\}+\frac{\alpha F^{*}(s)\left\{\pi_1(1-G_1^{*}(s+\alpha))+\pi_2\left(1-G_2^{*}(s+\alpha)\right)\right\}}{s+\alpha}-1$$

where we denote A= $\pi_1 G_1^{*}(s+\alpha)+\pi_2 G_2^{*}(s+\alpha)$

Simplifying the numerator we have

$$[1-A]Q-\frac{\alpha F^{*}(s)Q}{s+\alpha}\left[\pi_1(1-G_1^{*}(s+\alpha))+\pi_2(1-G_2^{*}(s+\alpha))\right]\left[\frac{\pi_1(1-G_1^{*}(s+\alpha))+\pi_2\left(1-G_2^{*}(s+\alpha)\right)}{s+\alpha}\right]$$
$$+\alpha QA\left[\frac{1-F^{*}(s)}{s}\right]-\alpha QpA\left\{\frac{1-F^{*}(s)}{s}\right\}+\alpha QpW^{*}(s)A\left[\frac{1-F^{*}(s)}{s}\right]-\alpha QA\left[\frac{1-F^{*}(s)}{s}\right]$$
$$-\alpha\left[\frac{\pi_1\left(1-G_1^{*}(s+\alpha)\right)+\pi_2\left(1-G_2^{*}(s+\alpha)\right)}{s+\alpha}\right]\left[\frac{1-F^{*}(s)}{s}\right]+pQA(s+\alpha)\left[\frac{1-W^{*}(s)}{s}\right]$$
$$-pQ\alpha FA\left[\frac{1-W^{*}(s)}{s}\right]-pA\left[\frac{1-W^{*}(s)}{s}\right]+QA-pAQ$$
$$+pW^{*}(s)AQ+\frac{\alpha F^{*}(s)Q\left[\pi_1\left(1-G_1^{*}(s+\alpha)\right)+\pi_2\left(1-G_2^{*}(s+\alpha)\right)\right]}{s+\alpha}-Q$$
$$=\frac{A}{s}\left[1-p+pW^{*}(s)\right]+\frac{\alpha F^{*}(s)}{s}\left[\frac{\pi_1(1-G_1^{*}(s+\alpha))+\pi_2(1-G_2^{*}(s+\alpha))}{s+\alpha}\right]-\frac{1}{s}$$
$$=\frac{1}{s}\left[(1-p+pW^{*}(s))A+\frac{\alpha F^{*}(s)}{s+\alpha}\left(\pi_1(1-G_1^{*}(s+\alpha))+\pi_2(1-G_2^{*}(s+\alpha))\right)-1\right]=\frac{1}{s}D(1,s)$$

Thus LHS= $\dfrac{\frac{1}{s}D(1,s)}{D(1,s)}=\dfrac{1}{s}$ = RHS

### B.1.6. Some expressions and their values at z = 1

| Expression | At z =1 |
|---|---|
| $C(z)$ | 1 |
| $C'(z)$ | $E(I)$ |
| $C''(z)$ | $E(I(I-1))$ |
| $m=\lambda-\lambda C(z)+\alpha$ | $\alpha$ |
| $\frac{dm}{dz}=-\lambda C^{/}(z)$ | $-\lambda E(I)$ |
| $G^{*}\left(\lambda-\lambda C(z)\right)$ | $G^{*}(0)=1$ |
| $G^{*}\left(\lambda-\lambda C(z)+\alpha\right)$ | $G^{*}(\alpha)$ |
| $\frac{d}{dz}G^{*}\left(\lambda-\lambda C(z)\right)=\frac{dG^{*}(n)}{dn}\bullet\frac{dn}{dz}$ | $G^{*/}(0)\left(-\lambda C'(1)\right)=\left(-E(S)\right)\left(-\lambda E(I)\right)=\lambda E(I)E(S)$ |

******

## List of References


1. Aissani, A & Artalejo, J. R. (1998). 'On the single server retrial queue subject to breakdowns.' *Queuing systems*, 30 (3-4), 309-321.
2. Alfa, A. S and Rao, Srinivasa, S. S. T. (2000). 'Supplementary variable technique in stochastic models.' *Probability in the Engineering and Information Sciences*, 14, 203-218.
3. Alfa, A. S. (2003). 'Vacation models in discrete time.' *Queueing Systems*, 44(1), 5-30.
4. Al-Jararha, J. & Madan, K. (2003). 'An M/G/1 queue with second optional service with general service time distribution.' *International Journal of Information and Management Sciences*. 14(2), 47-56.
5. Al-Jararha, J. & Madan, K. C. (2003). 'An M/G/1 queue with second optional service and general service time distribution.' *International Journal of Information and Management Sciences*, 14(2), 47-56.
6. Altman, E. & Yechiali, U. (2006).'Analysis of customers' impatience in queue with server vacations.' *Queuing Systems*, 52(4), 261-279.
7. Anabosi, R. F. & Madan, K. C. (2003). 'A single server queue with two types of service, Bernoulli schedule server vacations and a single vacation policy.' *Pakistan Journal of Statistics*, 19(3), 331-342.
8. Ancker Jr., C.J. & Gafarian, A.V.,(1963). 'Some queuing problems with balking and reneging.' *Operations Research*, 11(1), 88-100.
9. Artalejo, J.R. & Choudhury, G. (2004). 'Steady state analysis of an M/G/1 Queue with repeated attempts and two-phase service.' *Quality Technology of Quantitative Management*. 1(2), 189-199.
10. Arumuganathan, R. & Jeyakumar, S. (2005). 'Steady state analysis of a bulk queue with multiple vacations set up times with N-policy and closedown times.' *Applied Mathematical Modelling*, 29, 927-986.
11. Arumuganathan, R. & Ramaswami, K. S. (2005). 'Analysis of a bulk queue with fast and slow service rates and multiple vacations.' *Asia-Pacific Journal of Operation Research,* 22(2), 239-260.
12. Atencia, I. & Moreno, P. (2006). 'Geo/G/1 retrial queue with $2^{nd}$ optional service.' *International Journal of Operation Research*, 1(4), 340-362.

13. Avi-Itzhak, B. and Naor, P. (1963). ‘Some queueing problems with the service subject to breakdowns.’ *Operation Research*, 11(3), 303-320.

14. Baba, Y. (1986). ‘On the $M^X/G/1$ queue with vacation time’, *Operation Research Letters*, 5, 93-98.

15. Baba, Y. (1987). ‘On the $M^X/G/1$ queue and without vacation time, non preemptive, last come first served discipline, *Journal of Operation Research*, Japan, 30,150-159.

16. Bae, J, Kim, S. and Lee, E.Y. (2001). ‘The virtual waiting time of the M/G/1 queue with impatient customers.’ *Queuing Systems: Theory and Application*, 38(4), 485-494.

17. Bailey, N. T.J. (1954). ‘A continuous time treatment of a simple queue using generating functions.’ *J. Roy. Statist. Society*, B, 16, 288-291.

18. Bai, J., So, K.C. and Tang, C. (2016). ‘A queuing model for managing small projects under uncertainties.’ *European Journal of Operation Research*, 253, 777-790.

19. Banik, A. D., Gupta, U. C. and Pathak, S.S. (2007). ‘On the GI/M/1/N queue with multiple working vacations: analytic analysis and computation.’ *Applied Mathematical Modelling*, 31(9), 1701-1710.

20. Banik, A.D. (2010). ‘Analysis of single working vacation in GI/M/1/N and G/M/1/∞ queueing system.’ *International Journal of Operation Research*, 7(3), 314-333.

21. Barrer, D. Y. (1957). ‘Queueing with impatient customers and ordered service.’ *Operation Research*, 5(5), 650-656.

22. Beja, A & Teller, A. (1975). ‘Relevant policies for Markovian queueing systems with many types of servers.’ *Management Sciences*, 21(9), 1049-1054.

23. Bhat, U.N. (1969). ‘Sixty years of queueing theory.’ *Management Science*, 15B, 280-294.

24. Bhat, U.N. (2008). ‘An Introduction to Queueing Theory: modeling and analysis in applications.’ New York: Springer-Verlag, GmbH.

25. Bhatacharjee, P. & Ray, P.K. (2014). ‘Patient flow modeling and performance analysis of health care delivery processes in hospitals: a review and reflection.’ *Computers and Industrial Engineering*, ELSEVIER, 78, 299-312.

26. Borthakur, A. and Choudhury, G. (1997). ‘On a batch arrival Poisson queue with generalized Vacation.’ *Sankhya-Series* B, 59, 369-383.

27. Borthakur, A. and Choudhury, G. (2000). ‘The stochastic decomposition results of batch arrival Poisson queue with a grand vacation process.’ *Sankhya-Series B*, 62, Pt.3, 448-462.

28. Boxma, O. J. & De Wall, P. R. (1994). Multiserver queues with impatient customers.' *ITC*, 14, 743-756.

29. Boxma, O. J. & Yechiali, U. (1997). 'An M/G/1 queue with multiple types of feedback and gated vacations.' *Journal of Applied Probability*, 34,773-784.

30. Boxma, O. J. & de Waal, P.R. (1999). 'Multiserver queues with impatient customers'. ITC, 14, 743-756.

31. Bunday, B. D. (1996). 'An introduction to queueing Theory.' Oxford, England, *Oxford University Press.*

32. Chaudhury, M. L & Gupta, U. C. (1999). 'Modelling and analysis of $M/G^{a,b}/1/N$ queue-a simple alternative approach.' *Queueing Systems*, 31, 95-100.

33. Choi, D. and Kim, T. (2003). 'Analysis of a two-phase queueing system with vacations and Bernoulli feedback.' *Stochastic Analysis and Application*, 21(5), 1009-1019.

34. Choudhury, G. (2000). 'An $M^X/G/1$queueing system with a setup period and a vacation period.' *Queueing Systems*, 36, 23-38.

35. Choudhury, G. (2002). 'Analysis of the $M^X$/G/1 queuing system with vacation times.' *Sankhya- Series* B, 64(1), 37-49.

36. Choudhury, G. (2003b). A batch arrival queueing system with an additional service channel.' *International Journal of Information and Management Science*, 14(2).17-30.

37. Choudhury, G. (2005). 'An M/G/1 queueing system with two phase service under D-policy.' *International Journal of Information and Management Science*, 16(4), 1-17.

38. Choudhury, G. (2006). 'An M/G/1 queue with an optional server vacation.' *International Journal of Information and Management Sciences*, 17(3), 19-30.

39. Choudhury, G. (2008). 'A single server queueing system with two phases of service and vacations.' *Quality Technology and Quantitative Management,* 5(1), 33-49.

40. Choudhury, G. and Borthakur, A. (1995). 'On a Poisson queue with random idle period.' *Contr. Appl. Math. Stat*., 1, 33-42.

41. Choudhury, G. and Borthakur, A. (2000). 'The stochastic decomposition results o batch arrival Poisson queue with a grand vacation process.' *Sankhya Series B*, 62, 448-462.

42. Choudhury, G. & Madan, K. C. (2004). 'Steady state analysis of an $M^X/G_1$, G2)/1 queue with restricted admissibility and modified Bernoulli server vacations.' *Journal of Probability and Statistical Science*.2 (2), 167-185.

43. Choudhury, G. and Paul, M. (2004). ‘A batch arrival queue with additional service channel under N-policy.’ *Applied Mathematics and Computation*, 156, 115-130.

44. Choudhury, G. & Madan, K.C. (2005). ‘A two-stage batch arrival queuing system with a random set up time under modified Bernoulli schedule vacation.’ *Revista Investigacion Operacional*, 26(3), 238-250.

45. Choudhury, G. Tadj, L. & Paul, M. (2007). ‘Steady state analysis of an $M^X/G/1$ queue with two phase service and Bernoulli vacation schedule under multiple vacation policy.’ *Applied Mathematical Modelling*, 31(6), 1079-1091.

46. Choudhury, G. and Deka, K. (2008). ‘An M/G/1 retrial queue system with two phases of service subject to server breakdown and repair.’ *Performance Evaluation*, 65, 714-724.

47. Choudhuy, G. and Tadj, L. (2009). ‘An M/G/1 queue with two phases of service subject to server breakdown and delayed repair.’ *Applied Mathematical Modelling*, 33, 2699-2709.

48. Choudhury, A & Medhi, P. (2011). ‘Balking and reneging in multiserver Markovian queuing Systems.’ *International Journal of Mathematics in Operational Research*, 3(4), 377-394.

49. Choudhury, G. and Ke, J.C. (2012). ‘A batch arrival retrial queue with general retrial times under Bernoulli vacation schedule for unreliable sever and delaying repair.’ *Applied Mathematical Modelling*, 36, 6050-6060.

50. Choudury, G. and Deka, M. (2013). ‘A batch arrival unreliable server Bernoulli vacation queue with two phases of service and delayed repair.’ *International Journal of Operation Research*, 10(3), 134-152.

51. Coffman, E. G. Jr. and Hofri, M. (1986). ‘Queueing models of secondary storage devices.’ *Queueing Systems*, 1,129-168.

52. Cox, D. (1955). ‘The analysis of non-Markovian processes by the inclusion of supplementary variables.’ *Proceedings of the Cambridge Philosophical Society*, 51, 433-441.

53. Cox, D. R. (1965). ‘Some problems of statistical analysis connected with congestion.’ *Proc. Symp. Congestion Theory*, University of North Carolina, Chapel Hills, 289-316.

54. Daley, D. J. (1965). ‘General customer impatience in the queue GI/G/1.’*Journal of Applied Probability*. 2, 186-205.

55. Dimou, S., Economou, A., Fakinos, D., (2011). 'The single server queue with geometric abandonments.' *Journal of Statistical Planning and Inference*. 141, 2863-2877.
56. Doshi, B. T. (1985). 'A note on stochastic decomposition in a GI/G/1 queue with vacations or set up times.' *Journal of Applied Probability*, 22, 419-428.
57. Doshi, B.T. (1986). 'Queueing Systems with Vacations-a survey.' *Queueing Systems*, 1(1), 29-66.
58. Doshi, B. T. (1991). 'Analysis of a two phase queueing system with general service times.' *Operation Research Letters*, 10(5), 265-275.
59. Doshi, B. T. and Yao, D. (1995). 'Editorial introduction.' *Queueing Systems*, 20, 1-5, *Special Issue on Telecommunication System.*
60. Dshalalow, J. H. (1998). 'Queues with hysteretic control by vacations and post-vacation periods.' *Queueing Systems*, 29(2-4), 231-268.
61. El-Paoumy, M.S. (2008). 'On-Poisson arrival queue: $M^X/M/2/N$ with balking, reneging and heterogeneous servers.' *Applied Mathematical Sciences*, 2(24), 1169-1175.
62. Erlang, A. K. (1909). 'The theory of probabilities and telephone conversations.' *Nyt. Tidsskrift for Mathematik*, B, 20, 33-39.
63. Finch, P. D. (1959). 'Cyclic queues with feedback.' *J. R. Statistical Society*, B 21,153-157.
64. Federgruen, A. and Green, L. (1986). 'Queueing systems with service interruptions.' *Operation Research*, 34, 752-768.
65. Fomundam, S., and Herrmann, J. W. (2007). 'A survey of queueing theory applications in healthcare.' *Technical Report*, University of Maryland.
66. Frey, A., and Takashi, Y. (1999). 'An $M^X/GI/1/N$ queue with close-down and vacation times.' *Journal of Applied Mathematics and Stochastic Analysis*, 12(1), 63-83.
67. Fry, T. C. (1928). 'Probability and its Engineering uses.' Van Nostrand, New York.
68. Fuhrmann, S.W. (1984). 'A note on the M/G/1 queue with server vacations.' *Operation Research*, 32(6), 1368-1373.
69. Fuhrmann, S. W & Cooper, R. B. (1985). 'Stochastic decomposition in M/G/1 queue with generalized vacations.' *Operation Research*, 33, 1117-1129.
70. Gaver, D.P. (1959). 'Imbedded Markov chain analysis of a waiting line process in continuous time.' *Ann. Mathematical Statistics*. 30, 698-720.

71. Gaver, D. P. (1962). ‘A waiting line with interrupted service including priorities.’ *J. Roy, Stat* B-24:73-90.
72. Glenebe, G. and Pujolle, G. (1987). ‘Introduction to queueing networks.’ John Wiley and Sons, New York.
73. Goyal, T. L. and Harris, C. M. (1972). ‘Maximum likelihood estimation for queues with state dependent service.’ *Sankhya*, A34, 65-80.
74. Gray, W. J. Wang, P. P and Scott, M. (2004). ‘A queuing model with multiple types of server breakdowns.’ *Quality Technology of Quantitative Management*, 1 (2), 245-255.
75. Green, L. (2006): ‘Patient flow: Reducing delay in health care delivery.’ *Queuing Analysis in Health Care*, 281-307, In. R.W. Hall (Ed.), USA: Springer-Verlag.
76. Gupta, S.M., and Kavusturucu, A. (2000). ‘Production systems with interruptions, arbitrary topology and finite buffers.’ *Annals of Operation Research*, 13, 145-176.
77. Gupta, U. C & Sikdar, K. (2006). ‘Computing queue length distribution in MAP/G/1/N queue under single and multiple vacations.’ *Applied Mathematics and Computation*, 174, 1498-1525.
78. Gupta, V, Harchol-Balter, M., Scheller, W. A., Yechiali, U. (2006). ‘Fundamental characteristics of queue with fluctuating load.’ *Technical Report, CMU-CS-06-117*, School of Computer Science, CM University.
79. Gupur, G. (2011). ‘Advances in queueing models’ research.’ *ACTA Analysis, Functionalis Applicata*, 13(3), 225-245.
80. Haight, F. A. (1957). ‘Queueing with balking.’ *Biometrika*, 44(3/4), 360-369.
81. Haight, F. A. (1959). Queueing with reneging.’ *Metrika 2*, 186-197.
82. Henderson, W. (1972). ‘Alternative approaches to the analysis of M/G/1 and G/M/1 queue. *J. Oper. Res. Soc.*, 15, 92-101.
83. Hur, S. and Ahn, S. (2005). ‘Batch arrival queues with vacations and server set up.’ *Applied Mathematical Modelling*, 29, 1164-1181.
84. Jackson, J. R. (1957). ‘Networks of waiting lines.’ *Operation Research*, 5, 518-521.
85. Jain, M & Upadhyaya, S. (2010). ‘Optimal repairable $M^{[X]}$/G/1 with multi-optional services and Bernoulli vacation.’ *International Journal of Operation Research*, 7(1), 109-132.
86. Jaiswal, N. K. (1968). ‘Priority queues’ New York, Academic Press.

87. Jayawardene, A. K. & Kella, O. (1996). ‘M/G/1 queue with altering renewal breakdowns.’ *Queueing Systems*, 22, 1-2.

88. Jeyakumar, S .and Arumuganathan, R. (2011). ‘A non-Markovian queue with multiple vacations and control policy on request for re-service.’ *Quality Technology of Quantitative Management*, l8 (3), 253-269.

89. Kashyap, B. R. and Choudhury, M. L. (1988). ‘An Introduction to Queuing Theory.’ Ontario, Canada: A &A publication.

90. Ke, J.C. (2001). ‘The control policy of an $M^X/G/1$ queueing system with server setup and two vacation types.’ *Mathematical Methods of Operation Research*, 54, 471-490.

91. Ke, J.C. (2003a). ‘The analysis of general input queue with N policy and exponential vacation.’ *Queueing Systems,* 45(2), 135-160.

92. Ke, J.C. (2003b). ‘The optimal control of an M/G/1 queueing system with server startup and two vacation types.’ *Applied Mathematical Modelling*, 27, 437-450.

93. Ke, J.C. (2007). ‘Batch Arrival Queues under Vacation policies with Server Breakdowns and startup/closed down times.’ *Applied Mathematical Modelling*, l31 (7), 1282-1292.

94. Ke, J.C. (2007b). ‘Operating characteristic analysis on the $M^X/G/1$ system with a variant vacation policy.’ *Applied Mathematical Modelling*, 31(7), 1321-1337.

95. Ke, J.C. & Wang, K. H. (1999). ‘Cost analysis of the M/M/R machine repair problem with balking, reneging and server breakdowns.’ *The Journal of Operational Research Society*, 50(3), 275-282.

96. Ke, J.C. and Chang, F. M. (2009b). ‘$M^{[x]}/(G_1, G_2)/1$ retrial queue under Bernoulli vacations schedule with general repeated attempts and starting failures.’ *Applied Mathematical Modelling*, 33, 3186-3196.

97. Keilson, J. and Kooharian, A. (1960). ‘On time dependent queueing process.’ *Ann. Math. Stat*., 31, 104-112.

98. Keilson, J, & Servi, L.D (1987). ‘Dynamics of the M/G/1 vacation model.’ *Operation Research*, 35, 575-582.

99. Kelly, F. P. (1979). ‘Reversibility and Stochastic Networks.’ John Wiley & Sons, New York.

100. Kendall, D. G. (1951). ‘Some problems in the theory of queues.’ *Journal of the Royal Statistical Society-Series B*, 13(2), 151-185.

101. Khalaf, R. F., Madan, K. C. & Lukas, C.A. (2011a). 'An $M^X/G/1$ queue with Bernoulli schedule vacation times, random breakdowns, general delay times and general repair times.' *Applied Mathematical Sciences*,' 5(1), 35-51.

102. Khalaf, R. F., Madan, K. C. & Lucas, C.A. (2012). 'A $M^X/G/1$ queueing system with server breakdowns, server vacations, delay time and a standby server.' *International Journal of Operation Research*, 15(1), 30-47.

103. Khedhairi A. Al and Tadj, L. (2007). 'A bulk service queue with a choice of service and re-service under Bernoulli schedule.' *Int. J. Contemp. Mathematical Science*, 2(23), 107-1120.

104. Kingman, J. F. C. (1969). 'Markov population processes.' *Journal of Applied Probability*, 6, 1-18.

105. Kleinrock, L. (1975, 1976). 'Queueing Systems', John Wiley, New York.

106. Kumar Krishna, B. & Raja, J. (2000). 'On multi server feedback retrial queues with balking and control retrial rate.' *Ann. Oper. Res*., 143, 211-232

107. Kumar, B. K., Arivudainambi, D. & Vijaykumar, A. (2002). 'An M/G/1/1 queue with unreliable server and no waiting capacity.' *International Journal of Information and Management Science*, 13(2), 35-50.

108. Kumar, B. K. & Madheswari, S. P. (2005). 'An M/M/2 queuing system with heterogeneous server and multiple vacations.' *Mathematical and Computer Modelling*, 41, 1415-1429.

109. Kumar, M. S. and Arumuganathan, R. (2008). 'On the single serve batch arrival retrial queue with general vacation time under Bernoulli vacation schedule and two phases of heterogeneous services.' *Quality Technology and Quantitative Management*. 5(2), 145-160.

110. Kumar, R. and Sharma, S. K. (2012). 'A Markovian feedback queue with retention of reneged customers and balking.' *AMO-Advanced Modeling and Optimization*, 14(3), 2012.

111. Lee, A.M. (1966). 'Applied Queueing Theory'. Macmillan London.

112. Lee, H.W, Lee, S.S., Chae, K. C., & Nadarajan, R. (1992). 'On a batch service queue with single vacation.' *Applied Mathematical Modelling*, 16, 36-42.

113. Lee, H., Lee, S.S, Park, J.O and Chae, K.C. (1994). Analysis of $M^X/G/1$ queue with N-policy and multiple vacations.' *Journal of Applied Probability*, 31, 476-496.

114. Lee, H. S., & Srinivasan, M.M. (1989). 'Control policies for the $M^X/G/1$ queueing system.' *Management Science*, 35(6), 708-721.

115. Leonard, K. (1976). 'Queueing systems.' Vol.1, Theory, John Wiley & Sons, New York.

116. Levy, Y. and Yechiali, U. (1975). 'Utilization of idle time in an M/G/1 queueing system.' *Management Science*, 22, 202-211.

117. Levy, Y. and Yechiali, U (1976). An M/M/s queue with server vacations. *INFOR Journal*, 14(2), 153-163.

118. Li, H., and Zhu, Y. (1996). 'Analysis of M/G/1 queue with delayed vacation and exhaustive service discipline.' *European Journal of Operation Research*, 92(1), 125-134.

119. Li, H. and Zhu, Y. (1997). 'M (n)/G/1/N queue with generalized vacations.' *Comput. Oper. Res.*, 24, 301-316.

120. Li, W., Shi, D., and Chao, X. (1997). 'Reliability analysis of M/G/1 queueing system with server breakdowns and vacations.' *Journal of Applied Probability*, 34,456-555.

121. Li, J. and Tian, N. (2007). 'The discrete time GI/Geo/1 queue with working vacations and vacation interruption.' *Appl. Math. Comput.* , 185(1), 1-10.

122. Li, J. H., Tian, N., Zhang, Z. G. & Luh, H. P. (2009).'Analysis of the M/G/1 with exponentially working vacations- a matrix analytic approach.' *Queueing Systems*, 61,139-166.

123. Little, J. D.C. (1961). 'A Proof of the Queueing Formula: L=AW.' *Operation Research*, 9(3), 383-387.

124. Liu, L and Kulkani, V.G. (2006). 'Explicit solutions for the steady state distribution in M/Ph/1 queue with workload dependent balking.' *Queueing Systems: Theory and Practice*, 52, 251-260.

125. Madan, K.C. (1988). 'A cyclic queuing system with three servers and optional two-way feedback.' *Microelectron, Rel.*, 28:6, 873-875.

126. Madan, K. C. (1989). 'Time–dependent behaviour of a single channel queue with fluctuating efficiency.' *Tamkang Journal of Management Science*, 10(2), 71-80.

127. Madan, K. C. (1992). 'A bulk queueing system with random failures and two phase repairs.' *Microelectron, Reliability*, 32(5), 669-677.

128. Madan, K.C. (2000a). 'An M/G/1 queue with second optional service.' *Queueing Systems*, 34(1-4), 37-46.

129. Madan, K.C. (2000b). 'On a single server queue with two –stage heterogeneous service and binomial schedule server vacations.' *The Egyptian Statistical Journal*, 44(1), 39-55.

130. Madan, K. C. (2001). 'On a single server queue with two-stage heterogeneous service and deterministic server vacations.' *International Journal of Systems Science*, 32(7), 837-844.

131. Madan, K. C., (2003). 'An M/G/1 type queue with time homogeneous breakdowns and deterministic repair times.' *Soochow Journal of Mathematics*, 29(1), 103-110.

132. Madan, K. C. & Balkizi, A. (2002). 'An M/G/1 queue with additional second stage service and optional re-service.' *International Journal of Information and Management Sciences*, 13(4), 13-31.

133. Madan, K.C and Abu-Dayyeh, W. (2002). 'Restricted admissibility of batches into an M/G/1 type bulk queue with modified Bernoulli schedule server vacations.' *ESAIM Probability Statistics*, 6,113-125.

134. Madan, K. C., Abu-Dayyeh, W. & Gharaibeh, M. (2003). 'Steady state analysis of two $M^X/M^{a,b}/1$ queue models with random breakdowns.' *International Journal of Information and Management Sciences*, 14(3), 37-51.

135. Madan, K.C., Al-Nasser, A.D & Al –Masri, A.Q.(2004). On $M^X/\begin{pmatrix} G_1 \\ G_2 \end{pmatrix}/1$ Queue with optional re-service. *Applied Mathematics and Computation*.Elsevier.152, 71-88.

136. Madan, K.C & Al-Rawwash, M. (2005). 'On the $M^X/G/1$ queue with feedback and optional server vacations based on a single vacation policy.' *Applied Mathematics and Computation*, 160,909-919.

137. Madan, K. C. and Choudhury, G. (2004). 'Steady state analysis of an $M^X/(G_1G_2)/1$ queue with restricted admissibility and modified Bernoulli schedule server vacation.' *Journal of Probability and Statistical Science* 2(2), 167-185.

138. Madan, K. C. & Choudhury, G. (2005). 'A two-stage batch arrival queueing system with a modified Bernoulli schedule vacation under N-policy.' *Mathematical and Computer Modelling*, (USA), 42, 71-85.

139. Madan, K. C. & Choudhury, G. (2005). 'A single server queue with two phases of heterogeneous service under Bernoulli schedule and general vacation times.' *International of Information and Management Sciences*, 16(2), 1-16.

140. Madan, K.C., Al-Rawi, Z.R & Al-Nasser, A.D. (2005). On $M^X/(G1,G2)/1/G(BS)/V_S$ vacation queue with two types of general heterogeneous service.' *Journal of Applied Mathematics and Decision Sciences*. (3), 123-135. http://dx.doi.org/10.1155/JAMDS.2005.123

141. Maraghi, F.A., Madan, K.C & Darby-Dowman, K.(2009). 'A batch arrival queueing system with random breakdown and Bernoulli schedule server vacation having general vacation time distribution.' *International Journal of Information and Management Sciences*. 20(1), 55-70.

142. Maraghi, F. A., Madan, K. C. & Darby-Dowman, K. (2010). 'Bernoulli Schedule Vacation Queues with Batch Arrivals and Random System Breakdowns having General Repair Times Distribution.' *International Journal of Operation Research*, (USA). 7(2), 240-256.

143. Mark, S.D. (2010). Service Science, John Wiley & Sons Inc.

144. Medhi, J. (1997). 'A single server Poisson input.' *Advances in Combinatorial Methods and Applications to Probability and Statistics*, 317-388, Birkhausser, Boston.

145. Mehri, H. and Djemel, T. ( ). 'Queuing theory in toll motorway' *Studia Informatica Universalis*, 96-130.

146. Miller, L. (1964). 'Alternating priorities in multi-class queues.' *Ph.D dissertation*, Cornell University, Ithaca, New York.

147. Mitra, D. & Mitrani, I. (1991). 'Editorial introduction.' *Queueing Systems*, 9, 1-4, Special Issue on Communication Systems.

148. Mitrani, I. L and Avi-Itzhak, B. (1968). 'A many server queue with service interruptions.' *Operation Research*, 16, 628-638.

149. Molina, E. C. (1927). 'Application of the theory of probability to telephone trunking problems.' *B .S. T. J.*, 6. 461-494.

150. Nelson, R. (1995). 'Probability, Stochastic Processes and Queueing Theory.' New York: Springer-Verlag.

151. Neuts, M. F. (1984). ‘Matrix-analytic methods in queuing theory.’ *European Journal of Operational Research*, 15(1), 2-12.

152. Niu, Z. & Takahashi, Y. (1999). ‘A finite-capacity queue with exhaustive vacation/close-down/set up times and Markovian arrival processes.’ *Queueing Systems*, 31(1-2), 1-23.

153. Niu, Z., Shu, T. and Takashi, Y. (2003). ‘A vacation queue with setup and closedown times and batch Makovian arrival processes.’ *Performance Evaluation*, 54,225-248.

154. Pollaczek, F. (1965). ‘Concerning an analytical method for the treatment of queuing problems.’ *Proc. Symp. Congestion Theory*, Univ. of North Carolina, Chapel Hill, 1-42.

155. Prabhu, N. U. (1974). ‘Stochastic control of queueing systems.’ *Nav. Res. Log. Q*, 21,411-418.

156. Rosenberg, E. and Yechiali, U. (1993). ‘The $M^X/G/1$queue with single and multiple vacations under LIFO service regime.’ *Operation Research Letters*, 14, 171-179.

157. Scholl, M. & Kleinrock, L. (1983). ‘On the M/G/1 queue with rest periods and certain service-independent queuing discipline.’ *Operation Research*. 31(4), 705-719.

158. Selvam, D. and Sivasankaran, V. (1994). ‘A two-phase queueing system with server vacations.’ *Operation Research Letters*, 15(3), 163-168.

159. Servi, L. D. (1986). ‘D/G/1 queues with vacations.’ *Operation Research*, 34(4), 619-629.

160. Sengupta, B. (1990). ‘A queue with service interruptions in an alternating random environment.’ *Operation Research,* 38(2), 308-312.

161. Senthil, M. & Arumuganathan, R. (2010). ‘An $M^{[X]}$/G/1 retrial queue with two-phase service subject to active server breakdowns and two types of repair.’ *International Journal of Operation Research*, 8(3), 261-291.

162. Shanthikumar, J.G. (1988). ‘On stochastic decomposition in the M/G/1 type queues with generalized vacations.’ *Operations Research*, 36, 566-569.

163. Shermen, N. P. (2006). ‘Analysis and control of unreliable, single-server retrial queues with infinite capacity orbit and normal queue.’ Unpublished Ph.d thesis. *Air University*.

164. Sikdar, K and Gupta, U. C. (2008). “On the batch arrival batch service queue with finite buffer under server’s vacation; $M^{\{X\}}/G^{[Y]}/1/N$ queue.’ *Computer and Mathematics with Applications*, 56, 2861-2873.

165. Simon, B. (1984). ‘Priority queues with feedback.’ *J. Assoc. Comput. Machinery,* 31(1), 134-149.

166. Shawky, A. I. & El-Paoumy, M.S. (2008). 'The truncated Hyper-Poisson queue: $H_k / M^{a,b} / c / N$ with balking, reneging and general bulk service queue.' *Yugoslav Journal of Operation Research*, 8(1), 23-36.
167. Tadj, L & Ke, J.C. (2008). A hysteretic bulk queue with a choice of a service and optional re-service.' *Qualitative Technology of Quantitative Management*.5 (2), 161-178.
168. Taha, H.A. (1981). 'Queueing theory in practice.' *INTERFACES*, 11(1), 43-49.
169. Taha, H.A. (2007). Operation research: *an Introduction*. 8th edn. New Jersey; Pearson Prentice Hall.
170. Takacs, L. (1963). 'A single–server queue with feedback.' *Bell System*, 42,505-519.
171. Takagi, H. (1990). 'Time-dependent analysis of an M/G/1 vacation models with exhaustive service.' *Queueing Systems*, 6(1), 369-390.
172. Takagi, H. (1991). 'Queueing analysis: a foundation of performance evaluation'.Vol.1, Vacations and Priority systems, Part 1, North Holland, Amsterdam.
173. Takagi, H. (1992b). 'Time dependent process of M/G/1 vacation models with exhaustive service.' *Journal of Applied Probability*, 29, 418-429.
174. Takine, T. and Hasekawa, T. (1992). 'A generalization of the decomposition property in the M/G/1 queue with server vacations.' *Operation Res. Lett*., 12, 97-99.
175. Takine, T and Sengupta, B. (1997). 'A single server queue with server interruptions.' *Queueing Systems*, 26(3-4), 285-300.
176. Tang, Y. (2003). 'Single server M/G/1 queuing system subject to breakdowns- some reliability and queuing problems.' *Microelectronics Reliability*, 149, 588-613.
177. Tian, N., Zhang, D. & Cao, C. (1989). 'The GI/M/1 queue with exponential vacation.' *Queueing Systems*, 5(4), 331-344.
178. Tian, N., Zhang, Z. G. (2002). 'The discrete- time GI/Geo/1 queue with multiple vacations.' *Queueing Systems*, 40(3), 283-294.
179. Tian, N., and Zhang, Z. G. (2003). 'Stationary distribution of GI/M/c queue with Ph type vacations.' *Queueing Systems*, 44, 183-202.
180. Wang, J., Cao, J. & Li, Q. (2001). "Reliability analysis of the retrial queue with server breakdowns and repairs.' *Queueing Systems*, 38(4), 363-380.
181. Wang, J. (2004). 'An M/G/1 queue with second optional service and server breakdown.' *Computers and Mathematics with Applications*, 47, 1713-1723.

182. Wang, J. Liu, B. and Li, J. (2008). 'Transient Analysis of an M/G/1 Queue subject to Disaster and Server Failures.' *European Journal of Operation Research*. 189, 1118-1132.

183. Wang, J & Li, J. (2010). 'Analysis of the $M^{[X]}/G/1$ queue with second multi-optional service and unreliable server.' *Acta Mathematicae Applicatae Sinca, English Series*, 26(3), 353-368.

184. Wang, Y., Guo, J., Ceder, A. Currie, C., Dong, W., and Yuan, H. (2014). 'Waiting for public transport: queuing analysis with balking and reneging behaviors of impatient customers.' *Transportation Research*, B, 63, 53-76.

185. White, H. & Christie, L. S. (1958). 'Queueing with Preemptive Priorities or with Breakdowns.' *Journal of Operation Research Society of America*, 6, 79-95.

186. Whittle, P. (1967). 'Nonlinear migration processes'. *Bull. Int. Stat. Inst*., 42, 642-646.

187. Whittle, P. (1968). 'Equilibrium distributions for an open migration process.' *J. Appl. Prob.*, 5, 567-571.

188. Wu, D.A. and Takagi, H. (2006). 'M/G/1 queue with multiple working vacations.' *Performance Evaluation*, 63: 654-681.

189. Xu, Q., Bao, S. Ma, Z. & Tian, N. (2007). '$M^{[X]}/G/1$Queue with multiple vacations.' *Stochastic Analysis and Application*, 25, 127-140.

190. Yechiali, U. & Naor, P. (1971). 'Queuing problems with heterogeneous arrivals and service.' *INFORM*, 19(3), 722-734.

191. Yue, J., Yue, D., and Tian, R. (2008). 'Matrix-geometric solution of a repairable queueing system with two heterogeneous servers.' Acta. Mathematicae Applicate Sinca. 31, 894-900.

192. Zhang, Z. G. & Tian, N. (2001). 'Discrete time Geo/G/1 queue with multiple adaptive vacations.' *Queueing Systems*, 38,419-429.

193. Zhang, Z. G and Tian, N. (2004). 'An analysis of queueing System with multi task server.' *European Journal of Operation Research*. 156, 375-389.

194. Zhang, Y., Yue, D. & Yue, W. (2005). 'Analysis of an M/M/1/N queue with balking, reneging and server vacations.' *International Symposium on OR and its Applications*.

******