1

1.0

1.1

1.25

1.4

1.6

1.8

2.0

2.2

2.5

2.8

3.2

3.6

4.0

4.5

5.0

5.6

6.3

7.1

8.0

9.0

10.0

11.2

12.5

14.0

16.0

18.0

20.0

22.5

25.0

28.0

31.5

36.0

40.0

45.0

50.0

56.0

63.0

71.0

80.0

90.0

100.0

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service / Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

Si il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

**LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE**

National Library
of Canada

Bibliothèque nationale
du Canada

CANADIAN THESES
ON MICROFICHE

THÈSES CANADIENNES
SUR MICROFICHE

0-315-22192-5

NAME OF AUTHOR/NOM DE L'AUTEUR

ERIC DING-HIM HO

TITLE OF THESIS/TITRE DE LA THÈSE

ANALYSIS OF BUFFER LENGTH DISTRIBUTION
WITH RANDOM SERVER INTERRUPTIONS

UNIVERSITY/UNIVERSITÉ

CARLETON UNIVERSITY

DEGREE FOR WHICH THESIS WAS PRESENTED/

GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE

MASTER OF ENGINEERING

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE GRADE

84

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE

DR. C. M. WOODSIDE

Permission is hereby granted to the NATIONAL LIBRARY OF
CANADA to microfilm this thesis and to lend or sell copies
of the film.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-
QUE NATIONALE DU CANADA de microfilmer cette thèse et
de prêter ou de vendre des exemplaires du film.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.

L'auteur se réserve les autres droits de publication; ni la
thèse ni de longs extraits de celle-ci ne doivent être imprimés
ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/DATE

JAN. 30, 84

SIGNED/SIGNÉ

PERMANENT ADDRESS/RÉSIDENCE FIXE

N/A

TEMPORARY ADDRESS:

2750 CAROUSEL CRES. APT 414

GLoucester, ONTARIO, K1T 1Z5

ANALYSIS OF BUFFER LENGTH DISTRIBUTION
WITH RANDOM SERVER INTERRUPTIONS

By

© ERIC DING-HIM HO, B.Eng.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for
the degree of
Master of Engineering

Department of Systems and Computer Engineering
Faculty of Engineering
Carleton University
Ottawa, Ontario
December, 1983.

LICENCE TO CARLETON UNIVERSITY

In the interests of facilitating research by others at this institution and elsewhere, I hereby grant a licence to:

CARLETON UNIVERSITY

to make copies of my thesis:

ANALYSIS OF BUFFER LENGTH DISTRIBUTION
WITH SHARED SERVER INTERRUPTIONS

or substantial parts thereof, the copyright which is invested in me, provided that the licence is subject to the following conditions:

Only singled copies shall be made or authorized to be made at any one time, and only in response to a written request from the library of any university or similar institution on its own behalf or on behalf of one of its users.

This licence shall continue for the full term of the copyright, or for so long as may be legally permitted.

The Universal Copyright Notice shall appear on the title page of all copies made under the authority of this licence.

This licence does not permit the sale of authorized copies at a profit, but does permit the collection by the institution or institutions concerned of charges covering actual costs.

All copies under the authority of this licence shall bear a statement to the effect that the copy in question "is being made available in this form by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner."

The foregoing shall in no way preclude the granting by the author of a licence to the National Library of Canada to reproduce the thesis and to lend or to sell copies of the same. For this purpose it shall also be permissible for Carleton University to submit the abovementioned thesis to the National Library of Canada.

Wanda Stelmach
Signature of Witness

Jan 31, 84 M. Eng
Date Degree

[Signature]
Signature of Student

Systems & Computer Engineering
Department of Student

(ERIC DING-HIM HE)

The undersigned, recommends to the Faculty of
Graduate Studies and Research acceptance of the thesis

"Analysis Of Buffer Length Distribution With
Random Server Interruptions"

submitted by ERIC DING-HIM HO in partial fulfillment
of the requirements for the degree of
Master of Engineering.

C. M. Lindsay

Thesis Supervisor

B. Pagnier

Chairman, Department of Systems
and Computer Engineering

Carleton University
Faculty of Engineering
December 16, 1983.

Abstract

A discrete buffer system with Poisson arrival process, synchronous packet transmissions and random server interruptions as a model for a Rayleigh fading channel is studied. The objective of this study is to carry out an analysis of the Markov interrupted server by an M/G/1 queue. The modelling of the M/G/1 queue offers analytic solutions which can be easily evaluable by closed form expressions. In essence, the generating function of the queue length distribution for the described system is obtained and the inversion of this generating function is performed by a FFT technique. The analysis extends to the tail of the buffer distribution in a simple geometric form, through which the finite buffer distribution and the corresponding overflow probability can be approximated in a similar form for simple computation. Two approximations are used and results show that they are very effective in this particular analysis.

Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor, Professor C.M. Woodside for his unselfish contribution of his precious time and knowledge. The numerous comments and suggestions made by Professor Woodside have helped me a great deal during the course of this thesis.

Thanks are especially due to Diane Dodds for her help in many ways.

Lastly, I would also like to thank my wife, Nancy for her encouragement and patient support throughout the writing of this thesis.

TABLE OF CONTENTS

	Page
Abstract	i
Acknowledgements	ii
Lists Of Figures	vi
List Of Tables	viii
 Chapter 1 Introduction	
1.0 Overview	1-1
1.1 Previous Work	1-3
1.2 Present Work	1-5
1.3 Objective Of The Thesis	1-7
1.4 Thesis Organization	1-9
 Chapter 2 Summary Of Buffer Occupancy Analysis	
2.0 Introduction	2-1
2.1 The M/G/1 Queue As A Buffer Model	2-2
2.2 Notation	2-5
2.3 Infinite Buffer M/G/1 Queueing System	2-7
2.3.1 Background	2-7
2.3.2 Pollaczek-Khinchin (P-K) Transform Equation For Buffer Distribution	2-8
2.3.3 A Variation Of The P-K Transform Equation	2-10
2.4 Finite Buffer M/G/1 Queueing System	2-11
2.4.1 Background	2-11
2.4.2 Finite Buffer Queue Length Distribution and Overflow Probability	2-12
2.5 Inverting The P-K Transform Equation For The Buffer Distribution	2-14
 Chapter 3 Approximation Techniques To Large Buffer Occupancy	
3.0 Aim	3-1
3.1 Heavy Traffic Approximation On Buffer Occupancy	3-1
3.1.1 Theory Of Heavy Traffic	3-1
3.1.2 The Queue Length Distribution In Heavy Traffic	3-3

3.2	Laurent Expansion Approximation for Buffer Occupancy	3-5
3.2.1	Motivation	3-5
3.2.2	Approximating The Buffer Occupancy By Laurent Expansion.	3-6
3.2.3	Finding The Dominant Pole	3-8
3.2.4	Source Of Errors	3-9
3.2.5	Summary	3-10
3.3	Numerical Examples	3-11
3.3.1	Heavy Traffic Queue Length Distribution Of $M/E_2/1$	3-12
3.3.2	Laurent Expansion For The Queue Length Distribution Of The $M/E_2/1$ System	3-12
3.3.3	Evaluation Of The Approximation Methods For The $M/E_2/1$ System	3-13
3.4	Summary	3-16

Chapter 4 Queue Length Distribution of An Interrupted Server Model

4.0	Introduction	4-1
4.1	The System Model	4-3
4.1.1	System Operation	4-3
4.1.2	Channel Model For The Fading Process	4-5
4.1.3	Buffer Description	4-8
4.2	Equivalent Service Time Distribution Development	4-8
4.3	M/G/1 Analysis Of The Buffer Distribution	4-14
4.3.1	Infinite Buffer Distribution	4-14
4.3.2	Finite Buffer Distribution and Overflow Probability	4-15
4.4	Numerical Results	4-16
4.4.1	Infinite Buffer Distribution Of The Fading Channel Model	4-17
4.4.2	Independent Service Interruptions : A Special Case	4-21
4.4.3	Finite Buffer Distribution and Overflow Probability	4-23
4.4.4	Summary	4-27

Chapter 5 Asymptotic Solutions For The Buffer
Distribution and Overflow/Probability

5.0	Background	5-1
5.1	Heavy Traffic Buffer Distribution Approximation	5-2
5.2	Laurent Expansion Buffer Distribution Approximation	5-5
5.3	Asymptotic Solution For Buffer Overflow Probability	5-9

Chapter 6 Summary And Conclusion 6-1

Appendix A	ALGM_FB : An algorithm to compute the finite buffer distribution from the infinite buffer distribution.	A-1
------------	--	-----

Appendix B	FFT2C : The IMSL routine for the FFT inversion of the probability generating function.	B-1
------------	--	-----

Appendix C	FFT_INVERT : A program written in FORTRAN for the inversion of the probability generating function.	C-1
------------	--	-----

Appendix D	Derivatives of the service time distributions.	D-1
------------	---	-----

Appendix E	Determine the location of the dominant pole.	E-1
------------	---	-----

Appendix F	FIND_ROOT : A program to find the dominant pole of the P-K transform equation.	F-1
------------	--	-----

References		R-1
------------	--	-----

List Of Figures

	Page
2.1 Comparisons Between Infinite Buffer Distribution $q(n)$ and Finite Buffer Distribution $p(n)$ with Buffer Size, B	2-15
3.1 Comparisons Among $\tilde{q}_{HT}(n)$, $\tilde{q}_{LE}(n)$ and $q(n)$ Of The $M/E_2/1$ Queueing System	3-15
4.1 The System Model	4-4
4.2 Markov Representation Of The Fading Process	4-6
4.3 Transition Of System States For a Transmission Sequence	4-11
4.4 Infinite Buffer Length For The Interrupted Channel, $f_{rms}=1$ Hz.	4-18
4.5 Infinite Buffer Length For The Interrupted Channel, $f_{rms}=10$ Hz.	4-19
4.6 Comparisons Between M/G/1 Results and The Result From [5] For The Infinite Buffer Distribution	4-20
4.7 Finite Buffer Distribution Versus Buffer Size	4-24
4.8 Buffer Overflow Probability Versus Buffer Size	4-28
5.1 Infinite Buffer Distribution : Comparisons Between The Actual Distribution $q(n)$ and The Heavy-Traffic Approximation $\tilde{q}_{HT}(n)$, $f_{rms}=1$ Hz.	5-4

	Page
5.2 Laurent-Expansion Approximation Of The Infinite Buffer Length Distribution $\tilde{q}_{T,E}(n)$ and The Actual Distribution $q(n)$	5-8
5.3 Comparisons Of The Overflow Probabilities Of The Fading Channel, $\rho=0.9$	5-14
5.4 Comparisons Of The Overflow Probabilities Of The Fading Channel Model, $\rho=0.7$	5-15
E.1 Displays Of Some Popular M/G/1 Queues' Service Time Distributions As a Function Of z	E-4

List Of Tables

	Page
3.1 Heavy-Traffic Approximation Of The Infinite Buffer Length Distribution	3-14
3.2 Laurent-Expansion Approximation Of The Infinite Buffer Length Distribution	3-14
4.1 Values Of The Transition Probabilities Of The Fading Process, when $b=20$, $T=0.51$ and $\tau=0.01$	4-7
4.2 Comparisons Of The Infinite Queue Length Distribution Between This Analysis and Heines' (denoted as $q(n)$). This is the Special case where $\alpha + \beta = 1$	4-22
4.3a Values Of Buffer Overflow Probabilities For The Fading Channel Model, $f_{rms}=1$ Hz.	4-25
4.3b Values Of The Buffer Overflow Probabilities For The Fading Channel Model, $f_{rms}=10$ Hz.	4-26
5.1 Values Of The Mean Queue Length and The Heavy-Traffic Approximation For The Fading Channel Model, $f_{rms}=1$ Hz.	5-4
5.2 Values Of z_0 , b_1 , and The Laurent-Expansion Approximation For The Fading Channel Model	5-7
5.3 Comparisons Of The Overflow Probabilities Of The Fading Channel Model, $f_{rms}=1$ Hz.	5-12
5.4 Comparisons Of The Overflow Probabilities Of The Fading Channel Model, $f_{rms}=10$ Hz.	5-13

Chapter 1

Introduction

1.0 Overview

In many communication systems, random server interruption is a reasonable model for processes found in practical situations. Examples include: (1) breakdown of the service channel in a simple communication system, (2) preemption of data transmission by the higher priority voice packets in an integrated digital voice-data system, and (3) fading of signal strength in an intra-urban mobile radio communication system. For data communications such random interruptions prevent continuous service of data transmission and in effect, may degrade the system performance. Furthermore, in the case where data buffering is finite, they may result in severe problems of data buffer overflow if the interruption periods are long.

In physical realizations of data communication systems buffers have finite size. Yet, quite often, the effect of buffer overflow has been ignored by system designers who make the assumption that buffers are infinite in size. This practice stemmed from the fact that, from the designer's stand-point, infinite buffer analysis is less complex to analyze than the finite case. Nonetheless, with the added complexity of server interruptions, one would question the

accuracy of the infinite buffer assumption in developing design curves for data communication systems. Undoubtedly, buffer size is a problem to be considered in many engineering problems where the probability of buffer overflow and the allowable queueing delay are important factors. Examples include data compression, computer storage allocation, and statistical multiplexing of data in communication. Thus, for a specific system, given the buffer input traffic statistics and output characteristics, one may be required to find an optimal buffer size. Optimization is based on minimum cost and maximum performance, and constrained by the desire to have acceptably low traffic loss. In order to compute the traffic loss, it is necessary first to find the probability distribution of the buffer occupancy. This can be accomplished by the specification and analysis of an analytic queueing model. Although the designs of most existing queueing models were motivated by attempts to describe real situations in idealized models, the use of the queueing model has provided worthwhile results. In particular, the M/G/1 queueing model has been utilized in many applications. It serves as a basic model for analyzing communication systems with a single service resource, a Markovian arrival process and an arbitrary service distribution. Furthermore, it provides the basic framework and mathematical tools for dealing with a large class of system design issues.

1.1 Previous Work

In the past, considerable effort has been spent in analyzing buffer behavior in a variety of communication systems that allow service interruptions. In particular, a discrete time waiting system, where data units (messages, packets, bytes) are removed at equally spaced epochs, has been studied by many researchers [1-5].

Hsu [1] pioneered the study of a discrete buffered system with infinite buffer size, a fixed message length, a Poisson arrival process, and a single synchronous output channel which is subjected to interruptions controlled by a geometric process. In another words, the behavior of the output channel during each unit service interval is dictated by the outcome of a biased coin toss. Hsu obtained the generating function of the equilibrium buffer occupancy by associating a random variable with the geometric interruption process that controls the gating of the output channel. His approach did not include the geometric service-time distribution as part of the derivation. However, the solution for the buffer occupancy can be obtained using the Pollaczek-Khinchin transform equation (see Section 2.3.2).

Georganas [2] extended Hsu's work to multiple output channels. His approach used the classic imbedded Markov chain model for an M/G/1 system with bulk geometric service. The interrupted service was treated as such that the service

time was provided in multiples of a fixed time increment and the length of the service period was distributed geometrically. Thus, with the service time developed, the buffer occupancy was solved using the continuous-time M/G/1 method.

In a later work Heines [3] pointed out that the analyses of both [1] and [2] did not include the waiting that a customer must do until a service opportunity arrives. He then derived a modified version of the well known Pollaczek-Khinchin transform equation to obtain the solution for the buffer occupancy which does account for this waiting. When compared to the results from [1] and [2], Heines' results revealed that the average queue length and the variance of queue length are larger by a factor which depends on the arrival rate. The difference was expected since it corresponds to a difference of half a slot time in the average waiting time.

Kekre and Saxena [4] analyzed a queueing model for the finite buffer behavior with Poisson arrivals and random server interruptions controlled by a Bernoulli sequence of independent random variable. This behavior is typical of integrated digital voice-data systems. Results are obtained through matrix manipulations of the stationary state probabilities on buffer occupancy. The relationships between overflow probabilities, buffer size, and expected queueing delay for various traffic intensities were obtained.

Cavers and Woodside [5] analyzed the buffer requirements for data communications over fading radio links. The fading is represented by a Rayleigh fading process. Such a fading process gives server interruptions that have dependent sequences, rather than the independent interruptions given in [1]-[3]. Finite buffers were determined, and the probabilities of buffer overflow and data loss were included in the results. Their analysis encompassed more general arrival processes (Poisson and Bernoulli). In addition, they developed an analytical solution for the queue length distribution instead of a purely numerical approach.

The importance of buffer behavior has been realized in many practical applications of server interruptions. In most cases, the relationships among buffer occupancy, overflow probability, buffer size and expected queueing delay due to buffering for various traffic intensities are obtained.

1.2 Present Work

This thesis seeks to extend Heines' M/G/1 analysis technique [3] to more complex server interruption processes. In the analyses carried out in [1], [2] and [3], the server state is random and independent from one time slot to the next. This implies the probability that the server will be interrupted in the next time slot is fixed. This imposes a purely geometric distribution on the service time. In

addition, the previous studies assumed infinite buffer and made no attempt in obtaining the buffer occupancy from its generating function. This is unacceptable for the analysis of the more realistic finite buffer behavior, because the average buffer length is not sufficient for determining the buffer size needed to maintain a certain specified probability of buffer overflow.

In this work the server exhibits sequentially dependent interruptions. Thus, in terms of a simple two state Markov chain, the probability that the server is interrupted on the next state depends on the current state. We also include the finite buffer behavior as part of the analysis. As a consequence, the generality of results has been extended.

Our analysis is based on the system model described in [5]. However, the results from this present work generalize those of [5] in two ways: (1) the buffer probabilities found here is determined for all points in time, as opposed to those found in [5] which were only computed at the time slot boundaries; (2) the finite buffer probabilities derived from the infinite buffer probabilities are exact in this analysis, while approximations are used in [5].

In addition to the M/G/1 queueing analysis, this work also explores two approximation techniques which are meant to give asymptotes for the tail distribution of large buffer occupancy. The two techniques are the Heavy Traffic

approximation and the Laurent Expansion approximation. The Heavy Traffic approximation for buffer occupancy is easy to apply, but it has limitations. On the other hand, the Laurent Expansion technique gives better results.

1.3 Objectives Of This Thesis

The objectives of this thesis are :

(1) to develop an equivalent service time distribution for the generalized random interrupt service discipline as described in [5] and use it for the M/G/1 queueing system analysis;

(2) to effectively compute the buffer occupancy from the Z-transform domain, using a modified version of the Pollaczek- Khinchin (P-K) transform equation and a fast Fourier transform (FFT) technique for the inversion;

(3) to investigate the effectiveness of the Heavy-Traffic approximation [6] for the queue length distribution under high traffic intensity, and effectiveness of the Laurent-Expansion to approximate the tail of the queue length distribution;

(4) to develop a simple expression for the buffer overflow probability that does not rely on the full knowledge of the infinite buffer distribution.

To fulfill (1), an analytic expression for the packet transmission time over the fading channel is derived. It is

found on the basis of a first order two-state markov chain which describes the channel state at the time slot boundary.

For objective (2) we employ the idea of Heines [3] and modify the P-K transform equation such that it suits the description of the packet transmission mechanism. Then we use the FFT technique introduced by Cavers [7] as a fast and efficient way to invert the transform equation. This technique has been used effectively in [5], however, in the actual application, several parameters associated with the technique require careful attention.

To study (3) we employ, as test case, a model with a Markovian arrival process, an Erlang-2 service time distribution, and a single server with infinite queue capacity; i.e. an M/G/1 queueing system. Numerical results for this case are given. The approximated results are then compared to the actual results.

To accomplish (4), we take the approximated infinite buffer distribution as obtained from the Laurent Expansion approximation and use the M/G/1/B queue length distribution (as suggested in [10]) to arrive at an expression for the buffer overflow probability which is a function of the buffer size, B and the Laurent-Expansion coefficients.

In summary, the following results are developed in this thesis :

- (1) the analysis of the queue length distribution of a Markov interrupted server model using an M/G/1 approximation;

- (2) analytical results for the tail distribution of the M/G/1 queue length distribution;
- (3) approximations for the buffer overflow probability yielding analytical results in terms of the buffer size.

At this point, it is noteworthy to point out that the terms buffer occupancy and queue length distribution bear the same meaning. They are used interchangeably in various places of this thesis, depending on the context of the description.

1.4 Thesis Organization

Chapter 2 gives a brief introduction to M/G/1 buffer occupancy analysis; both the infinite and finite buffer case are examined. The original P-K transform equation is presented as an efficient way to compute the queue length distribution for the M/G/1 system. The modified version of the P-K transform equation as suggested by Heines [3] is also given. The chapter ends with the discussion on an efficient method for the inversion of the P-K transform equation.

Chapter 3 introduces the two approximation techniques which compute the tail distribution of large buffer occupancy. Each technique is introduced and discussed. At the end of the chapter some numerical results are presented. Comparisons between the approximated results and the actual results are presented graphically.

Chapter 4 begins by presenting the model for packet transmission operating over fading channel as a case study of the random server interruption. The system model is presented. In particular, the derivation of the fading channel service time distribution is given in detail. This chapter ends by giving the numerical results for the probabilities of the buffer occupancies for the infinite and the finite buffer cases. Overflow buffer probability is also included. In addition, comparisons between results from this analysis and those of [3] and [5] are given.

Chapter 5 gives the asymptotic solutions for the tail distribution of large buffer occupancy. Both approximation techniques are used. The principal aim of this chapter is to show that the tail of the large buffer occupancy exhibits a geometric distribution.

Chapter 6 presents a summary and conclusion of the thesis. Suggestions for further research work are also given.

Chapter 2

Summary Of Buffer Occupancy Analysis

2.0 Introduction

This chapter summarizes some of the standard concepts, notation and results which have been used in the analysis of buffer occupancy. Buffer occupancy is analyzed using queueing models. Any system in which arrivals request service from a finite capacity resource may be modelled and classified as a queueing system. The term "finite capacity" refers to the fact that the resource can satisfy the demands at a finite rate only. The modelling of a communication system by a generally solvable queueing system requires a thorough understanding of the theory of queues.

Assumptions which are useful for obtaining solutions, but are often not satisfied in practice, are

- (1) the system is modelled by a stationary stochastic process,
- (2) service requests are stochastically independent,
- (3) the system is in stochastic equilibrium,
- (4) the system is ergodic.

In fact, some of these concepts are difficult to achieve. Stationarity of queues is rare, independence (arrival or service process) does not come about frequently, and ergodicity is very difficult to establish with measurements over a finite period of time. Nonetheless, the analytic queueing model approach has undoubtedly given many worthwhile results in the field of communications.

To specify a queueing system it is necessary to identify the process that describes the arrival pattern as well as the structure and discipline of the service facility. Through these parameters one can proceed to obtain useful statistics such as mean queue length, mean time spent in the system, or the utilization of the resource, with the aid of many well established formulae.

2.1 The M/G/1 Queue As A Buffer Model

In many data communication systems, the service of the resources are generally driven by non-Markovian stochastic processes, while the process of the service requests is usually considered as generated from a Poisson source. In addition the service times are assumed independently and identically distributed (iid). Under this circumstance, the M/G/1 queueing system is commonly used as a buffer model for most communication systems. In short, the M/G/1 queueing system is a single server system with Poisson arrivals and arbitrary service-time distribution. Poisson arrival process

is generally a good assumption. In fact it is empirically and analytically established that it provides quite an acceptable mathematical abstraction of message traffic in many data and computer communications [14]. The Poisson arrival process is based on the assumption that arrivals are from a sufficiently large population and are sufficiently random that the assumption of an exponential distribution for the interarrival times can reasonably be made. On the other hand, the arbitrary service-time distribution accounts for the fact that the service time is in practice non-exponentially distributed, or it is not necessarily of the memoryless type. Thus, general service is a good improvement over the Markovian service. Despite the non-Markovian service process, the M/G/1 system offers analytic solutions which can be evaluated by closed form expression.

The M/G/1 queue does have its weaknesses: (1) we cannot obtain explicit distributions for the number of customers in the system or the time in system as we can do for the M/M/1 system. Rather, we require to find the expressions for the transforms (probability generating functions) of these distributions; (2) the generating function of the service time distribution may be very difficult to obtain and perhaps impossible to obtain.

Besides the M/G/1 queueing system, there exists other models such as GI/G/1, $M^k/G/1$ and PH/PH/1, which have been used in many situations where the M/G/1 model is not

appropriate or insufficient. Nonetheless, all these models require more complicated solution techniques.

In the rest of this chapter we summarize the "imbedded Markov chain" method for analyzing the buffer behavior of a M/G/1 queueing system as given in the standard texts such as [9] and [10]. We begin the analysis by first studying the queue length distribution of the system with infinite storage space (infinite buffer). Then, we modify the infinite buffer solution to arrive at the solution for the finite buffer case, the finite buffer occupancy and the overflow probability are given. We also give a summary of the extended version of the P-K transform equation developed by Heines [3]. Heines' equation applies to a discrete M/G/1 system, which takes into the consideration that a service request may not receive immediate service if it arrives to a idle system.

To enhance the computation of the infinite queue length distribution from its probability generating function, various techniques in the inversion of the transform are given in Section 2.5. In particular, the FFT inversion technique which is very efficient in computing the probability density function from a probability generating function is outlined.

2.2 Notation

This section presents the notation, symbols, and some of the standard distributions and formulae frequently used in M/G/1 queueing system. Four categories are given below.

(I) Random Variables :

- (i) $\bar{q}$ -- number of customers left behind by a departure.
- (ii) $\bar{s}$ -- the time a customer spends in system.
- (iii) $\bar{v}$ -- number of arrivals during a service epoch.
- (iv) $\bar{w}$ -- the time a customer spends in queue.
- (v) $\bar{x}$ -- the time a customer spends in service.

(II) Symbols :

- (i) λ -- arrival rate of customers.
- (ii) μ -- service rate of customers.
- (iii) ρ -- traffic intensity or server utilization.
- (iv) $\bar{t}$ -- average arrival time.
- (v) σ_a -- variance of the arrival time.
- (vi) σ_b -- variance of the service distribution.
- (vii) T -- throughput of the system.
- (viii) $\bar{q}$ -- mean number of customers in the system.
- (ix) $\bar{Nq}$ -- mean number of customers in the queue.
- (x) $\bar{x}$ -- average service time.
- (xi) $\bar{x}^k$ -- k-th moment of $b(x)$.
- (xii) $B^{*k}(0) = (-1)^k \bar{x}^k$
- (xiii) $B^{*1}(s) = d/ds B^*(s)$
- (xiv) $B^{*n}(s) = d^2/ds^2 B^*(s)$
- (xv) $V^{(k)}(z) = d^k/dz^k V(z)$

(III) ~~D~~istributions :

- (i) $A(t)$ -- the interarrival time distribution of the arrival process.
- (ii) $b(x)$ -- service time probability density function.
- (iii) $B(x)$ -- service time probability distribution function.
- (iv) $B^*(s)$ -- the Laplace transform of $b(x)$.
- (v) $q(n)$ -- the distribution of the number of customers in the system.
- (vi) $Q(z)$ -- the probability generating function (PGF) of $q(n)$.
- (vii) $V(z)$ -- PGF of $\text{prob}[v=k]$.
- (viii) $w(y)$ -- probability density function (pdf) of the waiting time.
- (ix) $W(y)$ -- probability distribution function (PDF) of the waiting time.

(IV) Formulae :

- (i) $\rho = \lambda \bar{x}$
- (ii) $\bar{s} = \bar{w} + \bar{x}$
- (iii) $\bar{N}_q = \bar{q} - \rho$
- (iv) $\bar{q} = \rho + \frac{\lambda^2 \bar{x}^2}{2(1 - \rho)}$
- (v) $\bar{N}_q = \lambda \bar{w}$ (Little's result)
- (vi) $B^*(0) = 1$
- (vii) $Q(1) = 1$

2.3 Infinite Buffer M/G/1 Queueing System

2.3.1 Background

The infinite buffer continuous-time M/G/1 system is the general case of the M/G/1 system upon which most derivations are based. For this system, service of customers commence immediately upon their arrivals or removals from the queue. Despite the non-Markovian nature of the server, there exists several well established techniques for handling it. Among these techniques, the method of "imbedded Markov chain" due to Kendall [8] is quite popular. It involves the simplification of the description of state from a two-dimensional description into a one-dimensional description. The fundamental idea behind this method is that the behavior of the non-Markovian process at some imbedded points is completely described as a Markov process. Generally, for the non-Markovian service process the imbedded point is the departure instant, the point in time right after a customer has completed service. The solutions at these imbedded points also provide the solutions for all points in time. By far, the most well known result for this method is the Pollaczek-Khinchin (P-K) transform equation, which readily yields the moments of the distribution of the buffer occupancy. In the next section to follow, the P-K transform equation is outlined, and the results for the infinite buffer M/G/1 system are given.

2.3.2 Pollaczek-Khinchin (P-K) Transform Equation For Buffer Occupancy

This section highlights the derivation of the P-K transform equation as presented by Kleinrock [9]. We begin with the definition of the generating function (Z-transform) for the random variable $\tilde{q}$ as :

$$Q(z) \triangleq \sum_{n=0}^{\infty} P[q=n] z^n$$

Other intermediate relationships are

$$V(z) = B^*(\lambda - \lambda z) \quad (2.1)$$

$$E[v] = \rho \quad (2.2)$$

$$E[q] = \rho + \frac{E[\tilde{v}^2] - E[\tilde{v}]}{2(1 - \rho)} \quad (2.3)$$

Following the derivation in [9], we arrive at the P-K transform equation that gives the Z-transform for the buffer occupancy of the infinite buffer M/G/1 queueing system in terms of the service time distribution and the utilization factor of the server, ρ ,

$$Q(z) = \frac{B^*(\lambda - \lambda z)(1 - \rho)(1 - z)}{B^*(\lambda - \lambda z) - z} \quad (2.4)$$

where $B^*(\lambda - \lambda z)$ is the Laplace Transform of the service-time distribution evaluated at the critical point $s = \lambda - \lambda z$.

As it turns out, the P-K transform equation gives the probability distribution of the buffer occupancy $P[q=n]$ in the Z-transform domain rather than the distribution itself.

In practice, it may be difficult to invert the Z-transform. Nonetheless, the usual properties of the transforms and their derivatives provide us extremely simple ways of computing the moments of the distribution; for example, the mean buffer size can be computed using the formula

$$\bar{q} = Q^{(1)}(1) \quad (2.5)$$

It is interesting to point out that the denominator of the P-K transform equation must always contain the factor $(1-z)$ since $B^*(0)=1$. Thus we are required to use L'Hopital's rule when we attempt to set $z=1$ in (2.5). Alternatively, we can make use of the second moment of the service-time distribution $\bar{x}^2$ to get

$$\bar{q} = \rho + \lambda^2 \bar{x}^2 / 2(1-\rho) \quad (2.6)$$

where $\bar{x}^2 = B^{*''}(0)$

Equation 2.6 is commonly referred to as the Pollaczek-Khinchin (P-K) mean-value formula. It is extremely useful in the design of communication systems where average queueing delay is an important criterion. From both (2.4) and (2.6) it is obvious that the service-time distribution is the major component of the buffer analysis. Thus in modelling the communication system under consideration, we must first obtain an accurate model for the server, which might be a communication channel, a telephone switch, or a data multiplexer. Once the service-time distribution is determined, results on the behavior of the buffer model can be obtained.

2.3.3 A Variation Of The P-K Transform Equation

In [3], Heines successfully used the M/G/1 approach to analyze a buffer in a discrete-time communication system with Poisson arrivals, periodic opportunities for service and random blocking of service. He has shown the flexibility of the P-K transform equation by using a variation of the equation to account for the different service epoch lengths for customers arrive to an idle system.

Without getting into further details of Heines' derivation, we simply present his version of the probability generating function of the buffer occupancy distribution as :

$$Q(z) = \frac{f_0(V(z) - zW(z))}{V(z) - z} \quad (2.7)$$

where $V(z)$ and $W(z)$ are the generating functions for the random variables v and w respectively, and f_0 equals $\text{Prob}(q=0)$. The definitions for the random variables v and w are as follows :

v = number of arrivals during the service epoch of a customer which arrives to a non-empty queue,

w = number of arrivals during the service epoch of a customer which arrives to an empty queue.

The service epoch of a customer, according to Heines' definition, begins when the customer reaches the head of the queue, and ends when service is completed. Since each service takes a fixed time slot, the service epoch includes the waiting for the slot boundary before starting the service, if a customer arrives to an empty queue.

An equation to find f_0 is developed by noting that $Q(1)=1$. Since $Q(z)$ yields indeterminate results at $z=1$, f_0 is found after the use of the L'Hopital's rule. Thus,

$$f_0 = \frac{1 - V^{(1)}(1)}{1 - V^{(1)}(1) + W^{(1)}(1)} \quad (2.8)$$

The average queue length may also be found from $\bar{q} = Q^{(1)}(1)$. Because the result is indeterminate, the actual result is obtained from

$$\bar{q} = \frac{N^{(2)}(1) - D^{(2)}(1)}{2D^{(1)}(1)} \quad (2.9a)$$

where $N(z) = f_0(V(z) - zW(z))$ and $D(z) = V(z) - z$.

Hence,

$$\bar{q} = \frac{f_0(W^{(2)}(1) + 2W^{(1)}(1) - V^{(1)}(1)) + V^{(2)}(1)}{2(1 - V^{(1)}(1))} \quad (2.9b)$$

2.4 Finite Buffer M/G/1 Queueing System

2.4.1 Background

A finite buffer M/G/1 queueing system is a special case of the M/G/1 system with finite storage capacity. In queueing theory notation it is referred to as the M/G/1/B system in which there may be at most B customers simultaneously in the system, including the customer being served. As in the case of a store-and-forward packet switching communication system with finite buffer space, it is obvious that the limited buffer will overflow and packets

will be lost from the buffer. Under this circumstance, it is one of the system designer's interests to be able to minimize such loss while maintaining minimal buffer space, such that the overall performance/cost ratio of the system will not be degraded.

2.4.2 Finite Buffer Queue Length Distribution and Overflow Probability

For buffers with finite capacity one of the main analytical problems is the determination of the buffer overflow probability. However, we must first compute the finite buffer distribution in order to get the overflow probability. In principal, the P-K transform equation does not hold for the finite case since the expected number of arrivals during a service period must be conditioned on the system size. Moreover, the M/G/1/B queueing system can be analyzed only when an analytic expression for the service time distribution is known. Otherwise, no analytic or approximation results are known to exist.

A limited number of methods can be used to obtain the finite buffer (queue length) distribution of an M/G/1/B system. Perhaps a straight forward way, which is exact but not necessarily efficient, is to get the result directly from the steady-state probabilities of the queue, since there are only a finite number of them. It thus requires to solve B

unknowns from B equations. Unfortunately, as B becomes large this method may become quite tedious. Another simple way is to analyze the buffer as if it is infinite and truncate the statistics at the buffer limit, B . This method gives close approximation only if the excluded probabilities are small. For a large finite buffer a better way is to modify the infinite buffer length distribution taking into the account that the effective arrival rate must equal the effective departure rate. Hence, the probability at the buffer size limit is adjusted according to this equality. Results from this method are exact. The procedure for this method is given next.

For the finite buffer (queue length) distribution we define the queue length steady state probabilities as $p(n)$ for $n=0,1,2,\dots,B$, whereas $q(n)$ for $n=0,1,2,\dots,\infty$ is for the infinite buffer case. In [10], it is shown that $p(n)$ is expressed in terms of $q(n)$:

$$p(n) = \begin{cases} K [1 - p(B)] q(n) & n < B \\ 1 - \mu / [\lambda + Kq(0)\mu] & n = B \end{cases} \quad (2.10)$$

where $K = 1 / \sum_{i=0}^{B-1} q(i)$

Since $\rho = \lambda/\mu = 1 - q(0)$ Equation 2.10 becomes :

$$p(n) = \frac{K}{1 + q(0)[K-1]} q(n) \quad n < B$$

$$p(B) = \frac{q(0)[K-1]}{1 + q(0)[K-1]} \quad (2.11)$$

Equation 2.11 is valid for all M/G/1/B queueing system. However it is required to first evaluate $q(n)$, using the P-K transform equation or other means. Figure 2.1 depicts the difference between the infinite buffer distribution $q(n)$ and the finite buffer distribution $p(n)$ with buffer size B . In summary, the procedures for obtaining the M/G/1/B buffer distribution are outlined in the algorithm ALGM_FB given in Appendix A.

The buffer overflow probability $P_{ovfl}(B)$ is defined as the fraction of packets (customers) lost due to arrival at a full buffer with buffer size B . As far as it is concerned, we can simply express it in terms of the buffer size. Since the buffer has a finite size of B and by the virtue of the Markovian arrival process, we can evaluate the overflow probability as :

$$P_{ovfl}(B) = p(B) \quad (2.12)$$

2.5 Inverting The P-K Transform Equation For The Buffer Distribution

In Section 2.3.2 the P-K transform equation for the queue length distribution is presented. As mentioned before, the inversion of this transform function may sometimes be difficult, or even tedious, since the transform equation is generally a transcendental function. Mathematically, the inverse of the Z-transform may be given as follows :

$$q(n) = \frac{1}{2\pi j} \oint_c Q(z) z^{-n-1} dz \quad (2.13)$$

Figure 2.1 Comparison Between Infinite Buffer Distribution, $q(n)$ and Finite Buffer Distribution, $p(n)$ with Buffer Size B .

Equation 2.13 represents a contour integral in the Z-plane over any closed path in the region of convergence that encompasses the origin in the Z-plane. This inverse Z-transform equation can be practically evaluated in several ways, which include :

- (1) Direct evaluation of (2.13) using Residue theorem.
- (2) Power series expansion of $Q(z)$.
- (3) Partial fraction expansion of $Q(z)$.
- (4) Fast Fourier Transform (FFT) inversion.

Methods (1) and (2) are not easy for this kind of application since they involve derivatives which may be quite difficult to handle. If dealing with simple expressions, method (3) is often favourable. Its principal virtue is that the result is quasi-analytic. For that method we take the Z-transform and write it as a summation of the form :

$$Q(z) = \sum_i \frac{c_i}{1 - d_i z} \quad (2.14)$$

Recognizing that each factor $c_i/(1-d_i z)$ has inverse Z-transform $c_i d_i^n$, we obtain

$$q(n) = \sum_i c_i d_i^n \quad (2.15)$$

Despite its simplicity, the use of method (3) is firmly restricted to rational polynomial generating functions. The transcendental functions often encountered in queueing analysis cannot be handled easily.

Method (4) perhaps is the most efficient way in dealing with the rather complex function of $Q(z)$, provided that the

generating function is defined in the negative powers of z and corresponds to a valid probability density function (pdf). This approach inverts a probability generating function (PGF) by means of a fast Fourier transform (FFT) routine to produce a finite set of samples from the corresponding discrete pdf. The equation for the approximate pdf is

$$q(k) \approx \frac{R^k}{N} \sum_{m=0}^{N-1} Q(R e^{j2\pi m/N}) e^{j2\pi km/N} \quad (2.16)$$

in which N is the number of distinct points, and R is the radius of integration such that $z = R e^{j\theta}$.

For a valid probability generating function, the FFT inversion offers a very fast way to obtain finite pieces (of length N samples) of the approximate pdf. In [7] Cavers outlined this approach and concluded that it is economical, widely applicable, and requires much less effort than other means.

Throughout this thesis, the FFT method is used as the standard tool to invert the probability generating function of the buffer length distribution. As mentioned before, the FFT inversion requires that the PGF is defined in the negative powers of z . If positive powers of z are used, we can still apply this method by letting $z = 1/R e^{j\theta}$. The FFT inversion for the P-K transform equation is carried out in a program written in Fortran-77, with the aid of the "FFT2C" library routine (see Appendix B) from the IMSL (International Mathematical and Statistical Libraries) package.

To perform an inversion, we are required to pay attention to the two input parameters N and R . First, if N is too small, aliasing error may be produced. Aliasing is a replication and summation of the true pdf to produce the periodic approximation. It results from the cyclic nature of the discrete Fourier transform, and can be reduced by increasing N . But, if N is too large, round off error may produce when the probability distribution is small. Through experiments, it is found that a 256 point inversion is sufficient to give consistent results. Second, the choice of R might produce some variation in the distribution due to round-off effect. If R is too small, this generally yields negative pdf. If R is too large, the probability $q(n)$ for large n becomes quite large; consequently, $\sum_{n=0}^{\infty} q(n) \gg 1$ which is incorrect for the probability distribution. It is recommended in [7] that the choice of R should be about 10% larger than the modulus of the maximum pole of the inverting function. A simple guideline to the choice of R is $R = 1.1 \times q(n+1)/q(n)$ for n is small ($5 < n < 20$). The probabilities $q(n)$ and $q(n+1)$ are obtained from an inversion using some initial guess of R . For the inversion of the P-K transform equation $R=1.05$ is often a good initial guess.

In appendix C, a program listing called FFT_INVERT is given to serve as an illustration of the inversion process. In addition, three sample runs are also given in the appendix. These sample runs are results for Chapter 4.

Chapter 3

Approximation Techniques To Large Buffer Occupancy

3.0 Aim

The principal aim of this chapter is to describe two methods for approximating the tail of the M/G/1 queue length distribution. In the case where buffer has finite size B and B is fairly large, the buffer overflow probability is dominated by the tail of the infinite buffer length distribution $q(n)$. Thus, it is desired to approximate the tail of the infinite buffer length distribution as $\tilde{q}(n)$ into a simple analytic form such that we can obtain the overflow probability in a similar analytic form. Both methods adopted here tend to give asymptotic values. The Heavy-Traffic (HT) approximation gives asymptotic values on long queues caused by heavy traffic intensity to the system; whereas, the Laurent-Expansion (LE) approximation focuses on the dominant term that governs the tail of the queue length distribution, regardless of the traffic intensity.

3.1 Heavy Traffic Approximation On Buffer Occupancy

3.1.1 Theory Of Heavy Traffic

One of the most obvious features of many queueing systems is the phenomenon of stability. Stability or

instability of stationary queues depends on the rate of customer-arrival to the queue with respect to the service rate. When the arrival rate reaches a critical value which is determined by the traffic intensity, the queue will be "just" stable; the queue length and the waiting time will be on the average be large. Under this circumstance it is referred to as "heavy-traffic".

In heavy traffic, the central result presented by Kingman [6] states that the waiting time distribution for a G/G/1 system is approximately negative exponential, with mean given by equation (3.2) :

$$W(y) \approx 1 - \exp(-y / \bar{w}) \quad (3.1)$$

$$\bar{w} = \frac{\sigma_a^2 + \sigma_b^2}{2(1 - \rho) \bar{t}} \quad (3.2)$$

Furthermore, Kingman remarked in [6] that the basic requirement for this approximation to be valid is that the denominator of (3.2) must be small compared with the square root of the numerator. This implies that ρ must be close to 1. For the M/G/1 queue, we can use the known exact mean wait $\bar{w} = \lambda \bar{x}^2 / 2(1 - \rho)$.

The application of the concept of the heavy traffic is extremely useful; in many cases, it is often possible to obtain approximate results even when no exact solution can be

found. In fact, the heavy traffic approximation is applicable to many more complex models which present formidable problems of analysis.

3.1.2 The Queue Length Distribution In Heavy Traffic

The interesting feature of the behavior of the queue in heavy traffic is the exponentially distributed waiting time. This provides us the motivation to find the queue length distribution in heavy traffic, using the relationships between the queue length distribution and the waiting time distribution. From (3.1) we can obtain the pdf of the waiting time :

$$w(y) = \frac{d}{dy} W(y) \simeq \frac{1}{\bar{w}} \exp(-y/\bar{w}) \quad (3.3)$$

Taking the Laplace Transform, we get

$$W^*(s) \simeq \frac{1/\bar{w}}{s + 1/\bar{w}} \quad (3.4)$$

Next, we define $\tilde{S}$ as the mean total time in the system such that

$$\tilde{S} = \bar{w} + \bar{x} \quad (3.5)$$

Hence, in the Laplace transform domain,

$$S^*(s) = W^*(s) B^*(s) \quad (3.6)$$

However, for M/G/1 queueing system, the queue length distribution is related to the mean total time by

$$\begin{aligned} Q(z) &= S^*(\lambda - \lambda z) \\ &= W^*(\lambda - \lambda z) B^*(\lambda - \lambda z) \end{aligned} \quad (3.7)$$

Equating the P-K transform equation (see Equation 2.4) with (3.7) we obtain, after simplification :

$$W^*(\lambda - \lambda z) = \frac{(1-\rho)(1-z)}{B^*(\lambda - \lambda z) - z} \quad (3.8)$$

By re-arranging (3.8), we get

$$W^*(\lambda - \lambda z)B^*(\lambda - \lambda z) = zW^*(\lambda - \lambda z) + (1-z)(1-\rho) \quad (3.9)$$

Consequently, we can write the queue length distribution as :

$$Q(z) = z W^*(\lambda - \lambda z) + (1-z)(1-\rho) \quad (3.10)$$

As ρ approaches 1, i.e. under heavy traffic, we can approximate $Q(z)$ as

$$\tilde{Q}_{HT}(z) = z W^*(\lambda - \lambda z) \quad (3.11)$$

From (3.4) with s substituted by $(\lambda - \lambda z)$ we have:

$$W^*(\lambda - \lambda z) \approx \frac{1/\bar{w}}{(\lambda - \lambda z) + 1/\bar{w}} \quad (3.12)$$

where, for M/G/1 system, $\bar{w} = \frac{\lambda x^2}{2(1-\rho)}$. Then, we have

$$\tilde{Q}_{HT}(z) \approx \frac{1/(1 + \lambda \bar{w})}{1 - [\lambda \bar{w}/(1 + \lambda \bar{w})] z} z \quad (3.13)$$

and after inversion we get :

$$\tilde{q}_{HT}(n) \approx \frac{1}{1 + \lambda \bar{w}} \left(\frac{\lambda \bar{w}}{1 + \lambda \bar{w}} \right)^{n-1} \quad (3.14)$$

Since $\bar{N}_q = \lambda \bar{w}$, (3.14) can be easily written as:

$$\tilde{q}_{HT}(n) = \frac{1}{\bar{N}_q} \left(\frac{\bar{N}_q}{1 + \bar{N}_q} \right)^n \quad (3.15)$$

Thus, from (3.15), we can obtain the tail of the queue length distribution using the the mean queue length, $\bar{N}_q$. As we can see (3.15) takes on a simple geometric form. Furthermore, either (2.5) or (2.6) gives the formula to

compute the mean number of customers in the system $\bar{q}$. Since $\bar{N}_q = \bar{q} - \rho$, the approximated tail distribution of the infinite queue length distribution can be obtained analytically.

3.2 Laurent-Expansion Approximation For Buffer Occupancy

3.2.1 Motivationⁱ

From the analysis of many communication systems, there exists an interesting feature which can be regarded as a general statement in many situations: the queue length distributions frequently exhibit geometric tails. In [11], Chu discovered empirically that the buffer occupancy probabilities $p(n)$ and $p(n+1)$ for his M/D/c/B model are related by a constant k such that $p(n+1)/p(n) = k$, and $k < 1$ is a function of the traffic intensity. Takahashi [12] also showed that for a given system with interarrival and service time distributions of the phase type (PH/PH/c) the queue length distribution had an asymptotically geometric tail. In [13], Geihs and Kobayashi gave the upper and lower bounds on the buffer occupancy distribution. Both bounds take the simple form Z_0^n , where Z_0 is the root of the characteristic equation determined by the input and output statistics. Thus, the buffer occupancy is bounded by the expression $AZ_0^n \leq q(n) \leq BZ_0^n$, where A and B are some constant values. In essence, all examples show that the queue length distribution is likely to have a simple geometric form $q(n) = cd^n$ for large

n, in which c and d are coefficients less than one. Such a simple analytic form makes it simple for the evaluation of the buffer overflow probability, since it is dominated by the tail of the infinite queue length distribution.

3.2.2 Approximating The Buffer Occupancy By Laurent-Expansion

The geometric tail infinite queue length distribution suggests that a simple pole near $z=1$ of $Q(z)$ may be dominant over other singularities of $Q(z)$. We denoted this dominant pole as z_0 . $Q(z)$ is analytic at $z=z_0$ if it is defined and differentiable at every point in some neighborhood of z_0 . From the P-K transform equation, we deduce that the service time distribution $B^*(\lambda - \lambda z)$ is also analytic at $z=z_0$. Since the Laurent series allows the expansion of $Q(z)$ about the point to which the function is singular, the series representation will be valid in an annulus between two concentric circles C_1 and C_2 where $Q(z)$ is analytic. The Laurent series expansion of $Q(z)$ about z_0 is therefore :

$$Q(z) = \sum_{k=0}^{\infty} a_k (z-z_0)^k + \sum_{k=1}^{\infty} b_k (z-z_0)^{-k} \quad (3.16)$$

where a_k and b_k are coefficients which can be evaluated as :

$$a_k = \frac{1}{2\pi i} \oint_C \frac{Q(z)}{(z-z_0)^{k+1}} dz$$

$$b_k = \frac{1}{2\pi i} \oint_C (z-z_0)^{k-1} Q(z) dz$$

The Laurent series converges for $0 < z - z_0 < R$, where R is the distance from z_0 to the nearest singularity of $Q(z)$. Since z_0 is assumed to be a simple pole of the first order, then b_k equals 0 for all $k > 1$, and $Q(z)$ becomes

$$Q(z) \approx \sum_{k=0}^{\infty} a_k (z - z_0)^k + b_1 (z - z_0)^{-1}$$

Thus, as z nears to z_0 , we get

$$\tilde{Q}_{LE}(z) \approx b_1 (z - z_0)^{-1}$$

Consequently, the approximated queue length distribution can be expressed as

$$\tilde{Q}_{LE}(n) \approx cd^n = -(b_1/z_0) z_0^{-n} \quad (3.17)$$

where $c = -b_1/z_0$ and $d = 1/z_0$. By using the properties of the Laurent expansion the computation of the coefficients c and d is a relatively simple task. But, first it is necessary to find the dominant pole z_0 and check that it is, in general a first order pole. With z_0 found, the coefficient b_1 can be determined using the Residue theorem which gives b_1 as :

$$\begin{aligned} b_1 &= \operatorname{Res}_{z \rightarrow z_0} Q(z) \quad \text{(residue of } Q(z) \text{ at } z=z_0) \\ &= \lim_{z \rightarrow z_0} [(z - z_0) Q(z)] \\ &= \frac{N(z_0)}{D^{(1)}(z_0)} \end{aligned} \quad (3.18)$$

$$\text{where } N(z) = B^*(\lambda - \lambda z) (1 - \rho) (1 - z) \quad (3.19)$$

$$D(z) = B^*(\lambda - \lambda z) - z \quad (3.20)$$

This determines the geometric approximation to the tail of the infinite queue length distribution, provided that z_0 can be found. In the next section we discuss the finding of the dominant pole z_0 .

3.2.3 Finding The Dominant Pole

To find the dominant pole z_0 , we can use a root finding technique and apply it to the denominator function $D(z)$. The dominant pole z_0 must be real, otherwise negative probability distribution may result. Thus, the basic problem is to find the real value $z=z_0$ nearest to $z=1$ such that $B^*(\lambda - \lambda z)=z$. From [13] we deduce, for an M/G/1 queue, that Equation 3.20 has exactly one real root that satisfies the condition of $B^*(\lambda - \lambda z_0)=z_0$. Since root finding is a very common task, a large number of such techniques have been devised. Techniques such as the Binary search method, Newton-Raphson method, and Secant method are quite popular and easy to use. Although any root finding method will usually converge to a root of $D(z)$, the question arises: how do we know that the found root is the dominant pole we are looking for? To answer this question, we show (in Appendix E) that in the interval $1 < z < 1+h$, there is only one value of z such that $D(z)=0$. In Appendix E, the value of h is determined and is equal to the reciprocal of the mean queue length $\bar{N}_q$. This mean queue length $\bar{N}_q$ can be found using (2.5) or (2.6). Thus, if a root of $D(z)$ is found and it is between $z=1$ and $z=1+1/\bar{N}_q$, this root is the dominant root of $D(z)$ (dominant pole of $Q(z)$).

Among various root finding methods, we use the Newton-Raphson method because of its speed. The iteration formula of Newton-Raphson method is:

$$z_{i+1} = z_i - D(z_i)/D^{(1)}(z_i)$$

where z_i is the estimate of the root of $D(z)$ at the i^{th} iteration. The iteration terminates when $|z_{i+1} - z_i| < \epsilon$, where ϵ is a small number. For $i=1$, we provide an estimate to the root as $z_1 = 1 + 1/\sqrt{Nq}$. Since $D(z)$ is always a convex function inside the interval $1 < z < 1 + 1/\sqrt{Nq}$ (shown in Appendix E), the search always converges to z_0 .

3.2.4 Source of Error

The accuracy of the approximated infinite buffer tail distribution by the Laurent-Expansion depends upon the accuracy in locating the dominant pole z_0 of the P-K transform equation, since the coefficients c and d are both functions of z_0 . The major source of error is the round-off error in the finding of z_0 . This error becomes quite serious if the traffic intensity ρ approaches 1. Under this circumstance, the value h quite often has the magnitude of 10^{-3} . If the round-off factor is of similar order, the approximated tail distribution may be off by as much as a factor of 2. In the Newton Raphson search routine, the round-off error can be reduced or avoided by decreasing the magnitude of ϵ . Thus, there exists a trade-off between accuracy and speed of the search. In general, with ϵ equals 10^{-6} we can achieve a high degree of accuracy.

3.2.5 Summary

To summarize the procedures in obtaining the Laurent-Expansion approximation for the infinite queue length distribution, we give algorithm ALGM_LE with the following steps :

- (1) Break down the generating function of the queue length distribution into two analytic functions, namely, the numerator function $N(z)$ and the denominator function $D(z)$.
- (2) Find $\bar{q}$ by using either (2.5) or (2.6), hence, $\bar{Nq}$ can be computed from $\bar{q} - \rho$.
- (3) Use the Newton-Raphson method (or other methods) to find the root z_0 nearest to 1 of the function $D(z)$. The initial estimate of z_0 is $z=1+1/\bar{Nq}$. The dominant root should lie between $z=1$ and $z=1+1/\bar{Nq}$.
- (4) Apply the Residue theorem to compute the coefficient b_1 in the Laurent series.

$$b_1 = \frac{N(z_0)}{D'(1)(z_0)}$$

- (5) Find c and d such that we can express $\tilde{q}_{LE}(n)=cd^n$.

$$c = \frac{-b_1}{z_0}$$

$$d = \frac{1}{z_0}$$

3.3 Numerical Example On The Approximations

In this section we set out to check the effectiveness of the two suggested approximation methods when applied to an M/G/1 system. First, we apply the Heavy-Traffic approximation (Section 4.4.1) which is followed by the Laurent-Expansion approximation (Section 4.4.2). Then we give the evaluation of both approximations in terms of efficiency, difficulty, and perhaps limitations.

We choose the service time distribution being Erlang-2. In M/E₂/1 queue the service time distribution is :

$$b(x) = 2\mu (2\mu x) \exp(-2\mu x)$$

Thus,

$$B^*(s) = [2\mu/(s+2\mu)]^2$$

Using the P-K transform equation for the queue length distribution, we get, in the simplest form :

$$Q(z) = \frac{(1-\rho)}{1 - (\rho + \rho^2/4)z + (\rho z/2)^2} \quad (3.21)$$

where $\rho = \lambda/\mu$

Thus, we have :

$$N(z) = (1-\rho) \quad (3.22)$$

$$D(z) = 1 - (\rho + \rho^2/4)z + (\rho z/2)^2 \quad (3.23)$$

Upon arriving to this point we can proceed to apply the approximation methods. Although the P-K transform equation of the Erlang-2 queue is rather simple, it serves the purpose of illustrating the concept of the approximations.

3.3.1 Heavy Traffic Queue Length Distribution Of M/E₂/1

From (3.15) the queue length distribution under heavy traffic is a function of the product $\lambda \bar{w}$. However for M/G/1 system, the mean waiting time is given as :

$$\bar{w} = \frac{\lambda \bar{x}^2}{2(1-\rho)}$$

For the Erlang service distribution $\bar{x}^2 = 3/2 \bar{x}^2$. After using the relationship that the product $\lambda \bar{x}$ equals the utilization factor ρ we can arrive at the expression for the approximated queue length distribution in terms of ρ .

Thus,

$$\tilde{q}_{HT}(n) \approx \frac{4(1-\rho)}{3\rho^2} \left(\frac{3\rho^2}{3\rho^2 - 4\rho + 4} \right)^n \quad (3.24)$$

Table 3.1 tabulates the approximated infinite queue length distributions for different values of ρ and Figure 3.1 depicts the results graphically.

3.3.2 Laurent Expansion For The Queue Length Distribution Of The M/E₂/1 system

From (3.22) and (3.23) the P-K transform equation of the queue length distribution is expressed in terms of a numerator function and a denominator function, $N(z)$ and $D(z)$ respectively. From (3.23) the derivative of $D(z)$ is :

$$D^{(1)}(z) = -(\rho + \rho^2/4) + \rho^2 z/2 \quad (3.25)$$

The Laurent-Expansion approximation requires the finding of the root closest to 1 of (3.23) and the coefficient b_1 . By following ALGM_LE, we obtain the dominant-root for each of the four different traffic intensity. As we can see in each case, z_0 is less than $1+1/\sqrt{Nq}$ (see $1/\sqrt{Nq}$ from Table 3.1). The expressions for the approximated infinite queue length distributions are given in Table 3.2. and these distributions are presented graphically in Figure 3.1.

3.3.3 Evaluation Of The Approximation Methods For The $M/E_2/1$ Queue

Figure 3.1 presents the curves for the infinite queue length distributions. The curves for the approximations (dotted lines) are compared with the exact values of the distributions (solid lines) obtained by numerical inversions. For the Heavy-Traffic approximation, results shows that it works very well when ρ is very close to 1. The effectiveness of this method, however, is restricted to a very narrow range of traffic intensities. When the traffic intensity falls below 90%, the approximated tail distribution is decaying much faster than the exact tail distribution. Thus, the difference between the approximation and the actual distribution increases with increasing queue length. In terms of difficulty the Heavy-Traffic approximation does not pose any. Its merit is the ease in computing the mean queue length using well established formulae for the $M/G/1$ queues.

ρ	$\bar{N}_q = \frac{3\rho^2}{4(1-\rho)}$	$\tilde{q}_{HT}(n) = \frac{1}{\bar{N}_q} \left(\frac{\bar{N}_q}{1 + \bar{N}_q} \right)^n$
0.99	73.51	(0.0136) (0.9866) ⁿ
0.95	13.53	(0.0739) (0.9321) ⁿ
0.90	6.075	(0.1646) (0.8587) ⁿ
0.80	2.400	(0.4167) (0.7059) ⁿ

Table 3.1 Heavy-Traffic Approximation of The Infinite Buffer Length Distribution

ρ	z_0	b_1	$\tilde{q}_{LE}(n) = -\frac{b_1}{z_0} \left(\frac{1}{z_0} \right)^n$
0.99	1.0135	-0.0135	(0.0134)(0.9867) ⁿ
0.95	1.0706	-0.0722	(0.0674)(0.9341) ⁿ
0.90	1.1498	-0.1570	(0.1366)(0.8696) ⁿ
0.80	1.3417	-0.3768	(0.2809)(0.7453) ⁿ

Table 3.2 Laurent-Expansion Approximation of The Infinite Buffer Length Distribution

Figure 3.1 Comparison Among $\tilde{q}_{HT}(n)$, $\tilde{q}_{LE}(n)$ and $q(n)$ Of The $M/E_2/1$ Queueing System

The Laurent Expansion method offers tight approximation to the exact distribution. For large n ($n > 10$), the difference between the actual distribution and the approximation is less than .005%, regardless of the traffic intensity. It is therefore obvious that, at least for this case, the tail of the queue length distribution can be approximated by a single term in the Laurent series expansion of the PGF of the queue length distribution. This single term is the dominant among all other terms which have little effect on the tail distribution. For the $M/E_2/1$ queue this approximation is not difficult to handle since the P-K transform equation is rather simple.

3.4 Summary

In this chapter, two approximation methods are given to approximate the tail of the infinite buffer length distribution. The Heavy-Traffic (HT) approximation is easy to apply and requires less computations. It is a good approximation when the traffic intensity to the system is always high. The Laurent-Expansion (LE) approximation, though requires more computations, gives very close approximation, regardless of the intensity of the traffic.

Of the two approximations, there exists a correlation between them. In LE approximation, the location of the dominant root z_0 lies between $z=1$ and $z=1+1/\sqrt{Nq}$. In heavy traffic, we observe that z_0 is very close to $1+1/\sqrt{Nq}$. Under

this circumstance, the coefficient d is approximately $\bar{N}_q/(1+\bar{N}_q)$, which is the same factor found in the HT approximation.

Chapter 4

Queue Length Distribution Of An Interrupted Server Model

4.0 Introduction

This chapter discusses the analysis of random service interruptions using a first-order Markov on-off model which includes correlation between successive interruption states. The interrupted server model studied here represents a mobile radio communication system whose channel characteristic suffers random fluctuation in signal strength known as fading. Fading is a phenomenon produced by partial reinforcement or cancellation among the signals which reach the receiver by different paths. When fading occurs, the channel is interrupted. Consequently, it offers no guarantee that the data, in the form of packets, are delivered without errors. Data may be lost and retransmissions of packets are required. Thus it is common for the system designer to adapt some means of error detection-retransmission mechanism as a countermeasure to the fading environment.

In [5], Cavers and Woodside validated the Markov on-off model for the fading channel. The interrupted channel is not M/G/1 but it can be modelled as such with an equivalent service time defined as the time from the packet ready to be sent until sent successful. The distribution of the

equivalent service time is given in Section 4.2. This chapter, however, deals equally with any servers with Markovian interruptions.

Occasionally, a prolonged fading may cause a large backlog of data. In physical realization the buffer is finite, and due to the random process of arrivals and retransmissions, the buffer will experience overflow. In this analysis we use an M/G/1/B approach to analyse the transmission of fixed length packets over the fading channel. Our initial objective is to formulate the equivalent service time distribution of the random dependent interrupted channel into an analytical expression so that the calculation of the probability distribution of the buffer length can be carried out using the classic imbedded Markov chain method. The same system configuration was studied by Cavers and Woodside [5] using a different approach. In [5], the steady-state balance equations of the buffer states are determined. Here, the equivalent service time distribution of the channel is obtained to be used in the M/G/1 queueing analysis. Moreover, the former work only found buffer probabilities at the discrete time slot boundaries; while this analysis determines probabilities for all instants. Another objective of our approach is to offer, as a close approximation, a fast and relatively simple analysis for the tail of the probability distribution of the buffer occupancy.

In the next section, the system model in terms of the system operation, channel model and the buffer is given.

4.1 The System Model

4.1.1 System Operation

Figure 4.1 depicts the system model in block diagram fashion. The overall system behavior described in [5] is summarized as follows: packets of N bits arrive at the transmitter and are queued in a buffer area to await transmission. The arrivals at the buffer are assumed to be generated from a Poisson process with an average rate of λ packets per transmission interval. Hence, the arrival satisfies Markov property. At every $T=N/R$ seconds, a packet is removed from the head of the queue and transmitted at R bits per second over the fading channel to the receiver. Since packets from the transmitter buffer are transmitted at fixed intervals for a fixed duration, it is said that the packets are synchronously transmitted. If the queue is empty at the time a transmission is to take place, an empty packet is sent. Because the channel suffers random interruptions, the service rate is no longer deterministic, rather, it is controlled by the channel fading process. At the receiver side, it is assumed that each packet contains sufficient redundancy to allow reliable error detection, so that if a packet is error-free it is accepted, otherwise it is discarded and retransmission is requested. However, in order to ignore the round trip delay we assume a packet numbering-continuous transmission system with a selective rejection mechanism is used.

Figure 4.1 The System Model
(Reproduced from [5])

4.1.2 Channel Model For the Fading Process

Much of the notation used below was developed in [5] and is summarized here. For the channel it is assumed that it exhibits slow, nonselective Rayleigh fading. Thus the received signal undergoes random fluctuation in strength, with a Rayleigh probability density function (pdf) of amplitude which is considered constant for the duration of a single packet. This assumption is essential since it enables us to analyze the channel state at the time slot boundaries. The instantaneous throughput, or service rate, of the transmission system is measured in packets per transmission interval and equals the probability of correct reception. The average throughput T of the channel is given as $\exp(-t/b)$, where b is the mean instantaneous signal-to-noise ratio (SNR) for a single pulse and t is the SNR threshold which determines the correctness of the received signal.

Following the development in [5], the channel state at the beginning of a packet transmission interval can be represented by a two-state Markov chain with time step τ , as given in Figure 4.2. The "Up" and "Down" (U and D, respectively) states in the two-state Markov chain is determined by the SNR threshold. The U-state represents the channel is uninterrupted, and that the packet is received error-free. On the other hand, the D-state indicates the channel is interrupted, i.e. the SNR is below the threshold, and the transmission fails. For this particular Markov model

Figure 4.2 Markov Representation Of The Fading Process

the interrupt process is sequentially dependent in such a way that the probability of the next state is conditioned on the current state. The up-to-down and down-to-up transitions have probabilities α and β respectively, selected as follows by matching the steady state probability T of the above-threshold state and also the threshold crossing frequency :

$$\alpha = 2\tau f_{rms} \sqrt{\pi t/b} \quad (4.1)$$

$$\beta = \alpha \frac{\exp(-t/b)}{1 - \exp(-t/b)} \quad (4.2)$$

where f_{rms} denoted the R.M.S. bandwidth of the equivalent lowpass fading process. A few selected values of the channel parameters used in [5] are adopted here and the corresponding values of α and β are given in Table 4.1. The channel parameters are : $b=20$ db, $T=0.51$, $\tau=0.01$.

f_{rms}	$\alpha = 2\tau f_{rms} \sqrt{\pi t/b}$	$\beta = \alpha \frac{e^{(-t/b)}}{1 - e^{(-t/b)}}$
1 Hz	0.02910	0.03028
10 Hz	0.2910	0.3028

Table 4.1 Values of the Transition probabilities
Of The Fading Process, when $b=20$, $T=0.51$
and $\tau=0.01$

4.1.3 Buffer Description

The buffer queue length at the start of each transmission interval is measured in packets. For a finite buffer system with size B it can hold at most B packets. Hence, the queue length executes a random walk with step $+1$ and -1 between 0 and B , depending on the channel state, down or up respectively. As described in Section 4.1.1, if the buffer is empty a null packet is transmitted with duration τ . Therefore, any packet arriving during an interval in which the buffer is empty must await the following interval for transmission. If the buffer is full at the start of an interval, any packets arriving during the interval are lost, since the space for the packet being transmitted is not released until transmission is complete.

4.2 Equivalent Service Time Distribution Development

In this section we set out to find the equivalent service time distribution according to the fading channel behavior. The service time distribution is a major component of the $M/G/1$ analysis in buffer occupancy. As mentioned in Section 4.0 the interrupted channel is not $M/G/1$, but can be modelled as such with an equivalent service time distribution for packet transmission times. The equivalent service epoch of a packet begins when the packet reaches the head of the queue, and ends when it is accepted at the receiver side. If the queue is empty at the time a transmission is to take

place, an empty packet is sent. Thus, a packet must wait for service even if it arrives to find the system empty. Because of this special kind of service, the equivalent service time distribution can be decomposed into three components: (1) the transmission time-- one time slot; (2) the waiting time while server (channel) is interrupted--an integral number of full time slots; (3) the waiting time from the arrival instant until first slot boundary--fraction of a slot, if the packet arrives to find the system idle.

In order to account for the third component, we are required to look at two distinct kinds of arrivals for the cases which packets arrive to find the system busy or idle. For case 1, we define the V-packets as arrivals which find the system busy. When a V-packet reaches the head of the queue, it must find the status of the channel being "Up". This situation holds because the previous packet must have a successful transmission, otherwise, no new packet transmission can take place. For case 2, we similarly define the W-packets as arrivals that find the system idle. When a W-packet is scheduled to transmit (after waiting for the completion of the empty packet), the channel state can either be "Up" or "Down". Such possibilities occur because the signal strength of the channel may be below the threshold (interrupted) while an empty packet is in transmission. Because of this the service time distribution for this case is different from that of case 1.

To analyze the equivalent service time distributions for the two cases, we begin by assuming that the arrival processes of both V-packets and W-packets are Poisson distributed. We then make the following definitions :

v = number of arrivals during the service epoch of a V-packet.

w = number of arrivals during the service epoch of a W-packet (waiting for slot boundary is included).

u = number of arrivals during the waiting for the slot boundary.

y = number of arrivals during the transmission of a W-packet.

Hence, $w = u + y$. The service time distributions for case 1 and case 2 are denoted as $b_v(x)$ and $b_w(x)$ respectively. To get these two distributions, the evolutions of system states may be represented pictorially by Figure 4.3 which displays the possible state transitions until a packet is successfully transmitted.

For case 1 (see Figure 4.3a), the probability distribution of the service time, $b_v(x) = \text{prob}[\text{transmission of a V-packet takes } j \text{ slots}]$, expressed in terms of the Dirac delta functions is :

$$b_v(x) = (1-\alpha) \delta_{x-1} + \sum_{j=0}^{\infty} \alpha \beta (1-\beta)^j \delta_{x-(2+j)} \quad (4.3)$$

where δ_{x-j} represents the service time takes on j time slots.

Taking the Laplace transform of (4.3), we get

$$BV^*(s) = \frac{1-\alpha-\beta}{1-\beta} e^{-s} + \frac{\alpha\beta}{1-\beta} \frac{e^{-s}}{1-e^{-s}(1-\beta)} \quad (4.4)$$

C = Transmission
Completed

Figure 4.3 Transition of System States For a Transmission Sequence

For case 2, since $w = u + y$, the probability distribution of the service time consists of two parts, $u(x)$ and $y(x)$. $u(x)$ is the waiting time for the slot boundary and $y(x)$ is the probability that the transmission time of a W-packet takes j time slots. For $u(x)$, if we assume that the probability of more than one arrival during an empty slot is small enough to be ignored, $u(x)$ can be considered uniformly distributed between 0 and 1. Hence,

$$u(x) = U(t) - U(t-1) \quad (4.5)$$

where $U(t)$ is the Unit step function.

For $y(x)$, the channel state at the start of a transmission is either U or D, the transmission sequences including the two possibilities give

$$y(x) = \gamma b_v(x) + (1-\gamma) \left[\sum_{j=0}^{\infty} \beta(1-\beta)^j \delta_{x-(1+j)} \right] \quad (4.6)$$

Variable γ is the probability that the channel is "U" at the start of transmission of a W-packet.

Again taking Laplace transforms, we get

$$U^*(s) = \frac{1 - e^{-s}}{s} \quad (4.7a)$$

$$Y^*(s) = \gamma BV^*(s) + (1-\gamma) \frac{e^{-s}}{1 - (1-\beta)e^{-s}} \quad (4.7b)$$

The equivalent service time distribution for case 2 is therefore

$$b_w(x) = u(x) \oplus y(x) \quad (4.8)$$

In the Laplace transform domain, it is

$$BW^*(s) = U^*(s) Y^*(s) \quad (4.9)$$

$$= \left(\frac{1 - e^{-s}}{s} \right) \left(\gamma BV^*(s) + (1-\gamma) \frac{\beta e^{-s}}{1 - (1-\beta)e^{-s}} \right)$$

To get τ we define two probability functions :

$\text{Pr}(A)$ = probability that the channel has been up (uninterrupted) for I time slots prior to the arrival of a W-packet,

$\text{Pr}(B)$ = probability that there are no arrivals during the I time slots.

From the two-state on-off Markov chain,

$$\text{Pr}(A) = \frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} e^{-(\alpha + \beta)I}$$

and because the arrivals are Poisson distributed, we get

$$\text{Pr}(B) = e^{-\lambda I} (e^{\lambda} - 1)$$

$$\text{Hence, } \tau = \sum_{I=1}^{\infty} \text{Pr}(A) \text{Pr}(B)$$

$$= \sum_{I=1}^{\infty} \left(\frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} e^{-(\alpha + \beta)I} \right) \left(e^{-\lambda I} (e^{\lambda} - 1) \right)$$

Therefore,

$$\tau = (e^{\lambda} - 1) \left(\frac{\beta}{\alpha + \beta} \frac{e^{-\lambda}}{1 - e^{-\lambda}} + \frac{\alpha}{\alpha + \beta} \frac{e^{-(\alpha + \beta + \lambda)}}{1 - e^{-(\alpha + \beta + \lambda)}} \right) \quad (4.10)$$

As we can see, τ depends on the arrival rate as well as the transition rates of the fading process.

In summary the equivalent service time distribution of the fading channel is made up of two different time distributions, $b_v(x)$ and $b_w(x)$. With the two distributions we can proceed to find the infinite buffer distribution using the modified version of the P-K transform equation (see Section 2.3.3). The expressions for the derivatives of the service time distributions are listed in Appendix D. In the next section to come, we look at the queue length distribution, using the equivalent service time distributions developed here.

4.3 M/G/1 Analysis Of The Buffer Distribution

4.3.1 Infinite Buffer Distribution

In Chapter 2, we presented the method of the "imbedded Markov chain" as the analytical tool to study the behavior of the buffer at customer departure times (end of packet transmissions). This method yields an analytic expression for the infinite buffer distribution for the M/G/1 queueing system using the P-K transform equation. Because of the special kind of service appeared in this case, Heines' version of the P-K transform equation is required since it accounts for the fact a packet which arrives to find the system ~~idle~~ will not receive immediate service.

By the virtue of the Poisson arrivals we can relate $V(z)$ and $W(z)$ of Equation 2.7 (Section 2.3.3) to the Laplace transforms of the service time distributions at the critical point $s = \lambda - \lambda z$. (See Equation 2.1 for reference)

Hence,

$$V(z) = BV^*(\lambda - \lambda z) \quad (4.11)$$

$$W(z) = BW^*(\lambda - \lambda z) \quad (4.12)$$

and f_0 is given by the equation as follows :

$$\begin{aligned} f_0 &= \frac{1 - V^{(1)}(1)}{1 - V^{(1)}(1) + W^{(1)}(1)} \\ &= \frac{1 + \lambda BV^{*'}(0)}{1 + \lambda BV^{*'}(0) - \lambda BW^{*'}(0)} \end{aligned} \quad (4.13)$$

Consequently, the infinite queue length distribution can be written in terms of the two different service distributions as

$$Q(z) = f_0 \frac{BV^*(\lambda - \lambda z) - zBW^*(\lambda - \lambda z)}{BV^*(\lambda - \lambda z) - z} \quad (4.14)$$

in which $BV^*(s)$ and $BW^*(s)$ are given as (4.4) and (4.9) respectively.

Equation 4.14 is a generalization of Heines' equation for the PGF of the infinite queue length distribution. It includes both dependent and independent server interruptions. For the special case $\alpha + \beta = 1$, Equation 4.14 is equivalent to Heines' equation, since $BW^*(s)$ equals $U^*(s)BV^*(s)$ regardless of the value of γ . In Section 4.4.2, comparisons of the results obtained from (4.14) and Heines' equation are given.

4.3.2 Finite Buffer Distribution and Overflow Probability

We now consider a finite buffer which has room for only B packets. Packets will be lost if they arrive to a full buffer. The finite buffer distribution and the overflow probability can be obtained using the M/G/1/B method. In Appendix A an algorithm to compute the finite buffer distribution in terms of the infinite buffer distribution $q(n)$ is outlined. Thus, by inverting (4.14) we can obtain the finite buffer distribution $p(n)$ following the given algorithm. The overflow probability can be obtained by using the equation $P_{\text{ovfl}}(B) = p(B)$, as given in (2.12).

4.4 Numerical Results

This section presents numerical results for the random interrupt server as represented by the fading channel model. Results are divided into three groups. In group 1, (Section 4.4.1) the transform equation of (4.14) for the infinite buffer distribution $q(n)$ is inverted numerically by the FFT inversion technique [7] for a few selected values of channel parameters. These results are compared with similar results obtained from [5]. The comparisons serve the purpose of validating the analytic approach taken in this study. Thus, instead of using the simulation approach for validation, we use the known results from [5] which have been proven to be valid. Group 2 (Section 4.4.2) is for the special case where the server behavior exhibits independent random interruptions. To accomplish that, we adjust the channel fade rate so that the condition $\beta = 1 - \alpha$ is satisfied. Results obtained from this special set of parameter are compared with those obtained from [3]. Group 3 (Section 4.4.3) considers of finite buffer probabilities and overflow probabilities for different buffer sizes. The finite buffer probabilities are presented graphically and the overflow probability are computed.

For simplicity, we opt to keep several channel parameters unchanged throughout the numerical examples. Hence, we suppose packets of 1024 bit blocks are to be transmitted at 100 Kps over a channel with mean signal-to

noise ratio $b=20$. Transmission time is therefore $\tau=0.01$, and from [5] the throughput $T=0.51$ packets per transmission interval is used for the analysis. The fade rate and the traffic intensity are the two parameters that altered in order to obtain different results. For the fade rates, we select $f_{rms}=1$ Hz and $f_{rms}=10$ Hz as representation figures for slow and fast fades, respectively. The difference in the fade rate is expected to generate some degree of variation in the results.

4.4.1 Infinite Buffer Distribution Of The Fading Channel Model

Results obtained here are based on two sets of channel parameters which are distinguished by the fade rate f_{rms} of the channel. The values of these parameters can be found in Table 4.1. For slow fade ($f_{rms}=1$ Hz) and fast fade ($f_{rms}=10$ Hz), Figures 4.4 and 4.5 show the probability distributions of the buffer occupancies for three different channel utilization factors. Quite obviously, the infinite buffer distributions exhibit geometric tails. By comparing the infinite buffer distribution $q(n)$ with those of [5] (we denoted them as $\hat{q}(n)$) we observe, from Figure 4.6, that the tails of the infinite buffer distributions found here have the same slopes to the corresponding tail distributions found in [5]. We also observe there is a slight difference in the distributions for long fade. We obtain higher probabilities in the tail (i.e. $q(n) > \hat{q}(n)$) because the probabilities

Figure 4.4 Infinite Buffer Length For The Interrupted Channel ($b=20$, $T=0.51$, $\tau=0.01$, $f_{rms}=1$ Hz)

Figure 4.5 Infinite Buffer Length For The Interrupted Channel ($b=20$, $T=0.51$, $\tau=0.01$, $f_{rms}=10$ Hz)

Figure 4.6 Comparison Between M/G/1 Results and The ~~Results From [5]~~ For The Infinite Buffer Distribution

considered here are taken for all points in time during a service interval, whereas those of [5] are considered at the slot boundaries only. This difference also explains that the probabilities of empty queues, $q(0)$, found by this analysis are lower than the corresponding ones found in [5], since in this analysis any arrivals during the service epoch of an empty packet are accounted for once they enter the buffer. Thus, the results obtained here give true probabilities of the buffer occupancies, despite using the equivalent service time distribution to model the channel.

4.4.2 Independent Service Interruptions : A Special Case.

This section gives the results for the special case where the interruption process of the server is independent from one time slot to the next. By referring to Figure 4.2, we can achieve such independence by letting $\beta = 1 - \alpha$ such that $\text{prob}(\text{the server is available}) = \beta$, regardless of the status of the current state of the server. For simplicity, we opt to keep the following channel parameters unchanged: $b=20$, $T=0.51$ and $\tau=0.01$. It is therefore required that the fade rate f_{rms} equals 17 Hz in order to make $\alpha + \beta = 1$. Under this condition α and β are found to be 0.49 and 0.51 respectively. Hence, the probability that the server is available for the next time slot is 0.51.

For comparison, we take the generating function of the queue length distribution from [3] and perform the FFT

(a) $\rho = 0.7$

n	q(n)	$\check{q}(n)$
0	.254561E+00	.252292E+00
1	.254630E+00	.254600E+00
2	.178250E+00	.178935E+00
3	.114795E+00	.115370E+00
4	.727525E-01	.731333E-01
5	.459974E-01	.462398E-01
6	.290734E-01	.292266E-01
7	.163756E-01	.184726E-01
8	.116144E-01	.116756E-01
9	.734082E-02	.737951E-02
10	.463974E-02	.466420E-02
20	.472050E-04	.474540E-04
30	.430268E-06	.482827E-06
40	.468634E-08	.491602E-08

(b) $\rho = 0.9$

n	q(n)	$\check{q}(n)$
0	.813339E-01	.801927E-01
1	.109005E+00	.108592E+00
2	.103836E+00	.103904E+00
3	.917350E-01	.919076E-01
4	.799560E-01	.801246E-01
5	.695591E-01	.697079E-01
6	.605016E-01	.606312E-01
7	.526226E-01	.527353E-01
8	.457695E-01	.458676E-01
9	.398070E-01	.398942E-01
10	.346246E-01	.346988E-01
20	.857885E-02	.859722E-02
30	.212556E-02	.213011E-02
40	.523644E-03	.527771E-03
50	.130485E-03	.130764E-03
60	.323302E-04	.323991E-04
70	.301044E-05	.802744E-05
80	.198476E-05	.198894E-05

Table 4.2 Comparison Of The Infinite Queue Length Distribution Between This Analysis and Heines' (denoted as $\check{q}(n)$). This is the special case where $\alpha = \rho - 1$

numerical inversion. The infinite queue length distribution of [3] is denoted as $\tilde{q}(n)$. Table 4.2 compares the results from the two different approaches. We can see that the infinite queue length distributions ($q(n)$ and $\tilde{q}(n)$), in general agree with each other. The probabilities obtained from this study are within 0.05% of those obtained in [3].

4.4.3 Finite Buffer Distribution and Overflow Probability

In this section we compute the finite buffer distribution $p(n)$ and the overflow probability of the fading channel model using the algorithm given in ALGM_FB. To follow the algorithm it is required first to compute the infinite buffer distribution by the FFT inversion. In Section 4.4.1 we already have the infinite buffer distributions for the following channel parameters :

- (1) $f_{rms} = 1$ Hz, $\rho = 0.9$
- (2) $f_{rms} = 1$ Hz, $\rho = 0.7$
- (3) $f_{rms} = 10$ Hz, $\rho = 0.9$
- (4) $f_{rms} = 10$ Hz, $\rho = 0.7$

The finite buffer distributions $p(n)$ for (2) and (3) above are found for a buffer size of 40. In Figure 4.7.a the probability $q(n)$ for $n > 40$ is large, therefore $p(40)$ is greater than $q(40)$. Contrarily, in Figure 4.7.b, $p(40)$ is less than $q(40)$ since $q(n)$ for $n > 40$ is small and that K_{rms} approximately 1. Tables 4.3 tabulates, for each buffer size B , the values of K and $q(0)(K-1)$ which are necessary for step

Figure 4.7 Finite Buffer Distribution Versus buffer Size

(1) $\rho = 0.9$ $q(0) = 4.4939 \times 10^{-2}$

B	$K = \frac{1}{\sum_{i=0}^{B-1} q(i)}$	$q(0)[K-1]$	$P_{\text{ovfl}}(B) = \frac{q(0)[K-1]}{1+q(0)[K-1]}$
10	5.4390	0.1995	0.1664
20	3.6324	0.1183	0.1058
30	2.8058	8.1152×10^{-2}	7.5061×10^{-2}
40	2.3337	5.9936×10^{-2}	5.6547×10^{-2}
50	2.0305	4.6308×10^{-2}	4.4258×10^{-2}

(2) $\rho = 0.7$ $q(0) = 0.1430$

B	$K = \frac{1}{\sum_{i=0}^{B-1} q(i)}$	$q(0)[K-1]$	$P_{\text{ovfl}}(B) = \frac{q(0)[K-1]}{1+q(0)[K-1]}$
10	2.1519	1.6472×10^{-1}	1.4143×10^{-1}
20	1.5721	8.1806×10^{-2}	7.5620×10^{-2}
30	1.3287	4.7003×10^{-2}	4.4893×10^{-2}
40	1.2022	2.8910×10^{-2}	2.8098×10^{-2}
50	1.1291	1.6459×10^{-2}	1.8124×10^{-2}

Table 4.3a Values of Buffer Overflow Probabilities
For The Fading Channel Model, $f_{\text{rms}} = 1$ Hz

(3) $\rho = 0.9$

$$q(0) = 6.9589 \times 10^{-2}$$

B	$K = \frac{1}{\sum_{i=0}^{B-1} q(i)}$	$q(0)[K-1]$	$P_{\text{ovfl}}(B) = \frac{q(0)[K-1]}{1+q(0)[K-1]}$
10	1.6404	4.4566×10^{-2}	4.2665×10^{-2}
20	1.1763	1.2271×10^{-2}	1.2122×10^{-2}
30	1.0611	4.2533×10^{-3}	4.2353×10^{-3}
40	1.0227	1.5799×10^{-3}	1.5775×10^{-3}
50	1.0087	6.0366×10^{-4}	6.0329×10^{-4}

(4) $\rho = 0.7$

$$q(0) = 0.2223$$

B	$K = \frac{1}{\sum_{i=0}^{B-1} q(i)}$	$q(0)[K-1]$	$P_{\text{ovfl}}(B) = \frac{q(0)[K-1]}{1+q(0)[K-1]}$
10	1.0471	1.0473×10^{-2}	1.0364×10^{-2}
20	1.0020	4.4541×10^{-4}	4.4521×10^{-4}
30	1.0001	1.8894×10^{-5}	1.8893×10^{-5}
40	1.0000037	8.2236×10^{-7}	8.2236×10^{-7}
50	1.0000002	3.5796×10^{-8}	3.5796×10^{-8}

Table 4.3b Values of Buffer Overflow Probabilities
For The Fading Channel Model, $f_{\text{rms}} = 10$ Hz

and step 3 of the algorithm ALGM_FB given in Appendix A. The corresponding overflow probabilities are computed and the curves for the overflow probability versus buffer size are given in Figure 4.8. From Figure 4.8 the overflow probabilities also exhibit geometric tails.

4.4.4 Summary

By comparing the numerical results with similar results from [3] and [5] we learn how well the M/G/1 approach works, despite using an equivalent service time distribution. We also see how efficient it is to compute finite buffer distribution and the buffer overflow probability, provided that the infinite buffer distribution is known.

Figure 4.8 Buffer Overflow Probability Versus Buffer Size

Chapter 5

Asymptotic Solutions For The Buffer Distribution
and Overflow Probability

5.0 Background

In Chapter 3, we suggested two approximation methods to approximate the tail of the queue length distribution. As it turned out for the $M/E_2/1$ case, the effectiveness of the Heavy-Traffic approximation is limited by the traffic intensity. It requires that the traffic intensity very close to 100%; more specifically, the approximation is unacceptable when the utilization falls below 90%. On the other hand, the Laurent-Expansion approximation offers tight lower asymptote to the tail of the queue length distribution.

In this chapter, we apply the two approximation methods to the more complex and practical interrupted server model. The more complicated P-K transform equation for the infinite buffer length distribution of the fading channel developed in Chapter 4 is used here. Numerical examples are used to illustrate the capability of each approximation. Section 5.1 gives the results of the Heavy-Traffic approximation to the slow fading process of the interrupted server model. Section 5.2 deals with the Laurent-Expansion approximation; results from this section are carried over for the calculations of an approximation for the finite buffer overflow probabilities given in Section 5.3.

5.1 Heavy Traffic Buffer Distribution Approximation

In this section, we apply the Heavy-Traffic approximation method to evaluate the buffer tail distribution of the fading channel in heavy traffic and perhaps study the limitation of this approximation. From Equation 3.15 the Heavy-Traffic approximation for the buffer tail distribution requires the computation of the mean queue length $\bar{N}_q$ as the main parameter. In the interrupted server model with synchronous transmissions and re-transmissions mechanism the buffer space occupied by the packet in transmission is not released until transmission is complete. The mean queue length $\bar{N}_q$ therefore equals the mean number of packets in the system $\bar{q}$. Hence, the approximated Buffer distribution for the tail is

$$\tilde{q}_{HT}(n) = \frac{1}{\bar{q}} \left(\frac{\bar{q}}{1 + \bar{q}} \right)^n \quad (5.1)$$

Thus, by referring to equation (4.14) for the infinite buffer distribution of the fading channel model, we can find $\bar{q}$ from $Q^{(1)}(1)$. After using the L'Hopital rule, we have

$$\bar{q} = Q^{(1)}(1) = \frac{N^{(2)}(1) - D^{(2)}(1)}{2D^{(1)}(1)}$$

where $N(z) = f_0 [BV^*(\lambda - \lambda z) - zBW^*(\lambda - \lambda z)]$

$$N^{(1)}(z) = f_0 [BV^{*(1)}(\lambda - \lambda z) - zBW^{*(1)}(\lambda - \lambda z) - BW^*(\lambda - \lambda z)]$$

$$N^{(2)}(z) = f_0 [BV^{*(2)}(\lambda - \lambda z) - zBW^{*(2)}(\lambda - \lambda z) - 2BW^{*(1)}(\lambda - \lambda z)]$$

$$D(z) = BV^*(\lambda - \lambda z) - z$$

$$D^{(1)}(z) = BV^{*(1)}(\lambda - \lambda z) - 1$$

$$D^{(2)}(z) = BV^{*(2)}(\lambda - \lambda z)$$

The definitions of $BV^*(\lambda - \lambda z)$ and $BW^*(\lambda - \lambda z)$ are given in (4.4) and (4.9) respectively, and they are once again presented here :

$$BV^*(\lambda - \lambda z) = e^{-(\lambda - \lambda z)} \left(1 - \alpha - \frac{1 - e^{-(\lambda - \lambda z)}}{1 - (1 - \beta) e^{-(\lambda - \lambda z)}} \right)$$

$$BW^*(\lambda - \lambda z) = \frac{1 - e^{-(\lambda - \lambda z)}}{\lambda - \lambda z} \left(\gamma BV^*(\lambda - \lambda z) + \frac{(1 - \gamma) \beta e^{-(\lambda - \lambda z)}}{1 - (1 - \beta) e^{-(\lambda - \lambda z)}} \right)$$

Other formulae that are required are :

$$(i) T = \beta / \alpha + \beta = 0.51$$

$$(ii) \lambda = \rho T$$

$$(iii) \gamma = (e^\lambda - 1) \left(\frac{\beta}{\alpha + \beta} \cdot \frac{e^{-\lambda}}{1 - e^{-\lambda}} + \frac{\alpha}{\alpha + \beta} \cdot \frac{e^{-(\alpha + \beta + \lambda)}}{1 - e^{-(\alpha + \beta + \lambda)}} \right)$$

$$(iv) f_0 = \frac{1 + \lambda BV^{*'}(0)}{1 + \lambda BV^{*'}(0) - \lambda BW^{*'}(0)}$$

$$= \frac{1 - \rho}{(1 - \rho + \gamma \rho) + \lambda \left((1 - \gamma) / \beta + 1/2 \right)} \quad (5.2)$$

The derivatives of $BV^*(\lambda - \lambda z)$ and $BW^*(\lambda - \lambda z)$ are given in Appendix D.

For numerical results, we take f_{rms} equals 1 Hz (slow fade). The corresponding up-to-down and down-to-up transition probabilities, α and β respectively, are 0.02910 and 0.03028 as obtained from Table 4.1. Table 5.1 lists the values of $\bar{q}$ and the expressions of the approximated buffer tail distribution for the slow fade rate with different values of traffic intensities. The comparisons of the approximated distributions and the exact distributions are depicted in Figure 5.1.

ρ	$\bar{c}_{HT}$	$\tilde{q}_{HT}(n) = \frac{1}{\bar{q}} \left(\frac{\bar{q}}{1 + \bar{q}} \right)^n$
0.99	832.09	$(0.0012)(0.9988)^n$
0.95	151.91	$(0.0066)(0.9935)^n$
0.90	67.45	$(0.0148)(0.9854)^n$
0.70	12.77	$(0.0783)(0.9274)^n$

Table 5.1 Values of The Mean Queue Length and The Heavy-Traffic Approximation For The Fading Channel Model, $f_{rms} = 1$ Hz

Figure 5.1 Infinite Buffer Distribution :
Comparison Between The Actual Distribution $q(n)$ and The Heavy-Traffic Approximation $\tilde{q}_{HT}(n)$, $f_{rms} = 1$ Hz

As we can see from Figure 5.1, the effectiveness of the approximation depends on the traffic intensity . When ρ falls below 90%, the approximated distributions decay much faster than the actual distributions. However, this is expected since the approximation is derived under heavy traffic. For ρ very close to 1, the approximation is justified.

5.2 Laurent Expansion Buffer Tail Distribution Approximation

In Chapter 3, we showed that the tail of the infinite queue length distribution can be expressed in a simple geometric form, and the Laurent series expansion can be used to approximate the tail distribution from the P-K transform equation. To test this idea for the complicated interrupted server, we take Equation 4.14 and express it as $Q(z)=N(z)/D(z)$ in which we have :

$$N(z) = f_0 \left(BV^*(\lambda - \lambda z) - z BW^*(\lambda - \lambda z) \right)$$

$$D(z) = BV^*(\lambda - \lambda z) - z$$

The Laurent-Expansion approximation for the infinite buffer tail distribution uses the dominant pole of the P-K transform equation as part of the result. To obtain the dominant pole of $Q(z)$ (the root closest to $z=1$ of $D(z)$), we follow the guidelines of steps (2) and (3) of ALGM_LE (Section 3.2.5). The program FIND_ROOT (Appendix F) implements the Newton Raphson search method to find the roots of $D(z)$.

As a numerical example, we now consider the fading channel with $f_{rms}=1$ Hz and $\rho=0.9$. Other parameters are $\alpha=0.02910$, $\beta=0.03028$, and $\lambda=0.459$. Also, we get from Table 5.1 the mean queue length $\bar{N}_q$ equals 67.48. Hence h equals 0.0148 (1/67.48). Using $z=1.0148$ as the initial estimation for the root, program FIND_ROOT gives $z_0=1.011954$. Since it lies between 1 and 1.0148, the dominant pole of the P-K transform equation is found. To evaluate the special coefficient b_1 in the Laurent-Expansion approximation, we have

$$b_1 = N(z_0)/D'(z_0) = -0.0109509$$

Thus, for the approximated infinite buffer tail distribution, we have $\tilde{q}_{LE}(n)=cd^n$, in which $c=-b_1/z_0$ and $d=1/z_0$. Hence, for this particular example $\tilde{q}_{LE}(n)=(0.010822)(0.9882)^n$.

This example illustrates the relatively simple way in obtaining the approximation of the infinite buffer tail distribution of a fairly complex interrupted server queueing system, without requiring to invert the probability generating function that describes the queue. To justify the effectiveness of the approximation, the results of the tail distributions for two different fade rates (1 Hz and 10 Hz) and different values of ρ are given. Table 5.2 summarizes the results and Figure 5.2 depicts the approximated and the exact tail distribution. Figure 5.2 reveals that for the buffer length n larger than 10, the infinite buffer distributions obtained from the Laurent Expansion approximation method are very close, regardless of the traffic intensity.

(a) $f_{rms} = 1$ Hz

ρ	z_0	b_1	$\tilde{q}_{LE}(n) = cd^n$
0.99	1.001184	-0.001145	(0.00114)(0.9988) ⁿ
0.95	1.005915	-0.005664	(0.00563)(0.9946) ⁿ
0.90	1.011954	-0.010951	(0.01082)(0.9882) ⁿ
0.70	1.039351	-0.030992	(0.02982)(0.9621) ⁿ

$$c = -b_1/z_0, \quad d = 1/z_0$$

(b) $f_{rms} = 10$ Hz

ρ	z_0	b_1	$\tilde{q}_{LE}(n) = cd^n$
0.99	1.009329	-0.009260	(0.009266)(0.9907) ⁿ
0.95	1.048165	-0.048183	(0.045969)(0.9541) ⁿ
0.90	1.100478	-0.102733	(0.093353)(0.9087) ⁿ
0.70	1.368645	-0.382452	(0.279441)(0.7307) ⁿ

Table 5.2 Values Of z_0 , b_1 , and The Laurent-Expansion Approximation For The Fading Channel Model

Figure 5.2 Laurent-Expansion Approximation Of The Infinite Buffer Length Distribution $q_{LE}(n)$ and The Actual Distribution $q(n)$

An interesting observation arises from this approximation is that the position of the dominant pole of the P-K transform equation depends on both the traffic intensity and the fade rate of the channel. For low traffic intensity and high fade rate, the position of the dominant pole is further away from $z=1$. However, the position of the dominant pole z_0 has little effect on the effectiveness of the approximation. But, the accuracy of the approximation depends on the distance of z_0 and $z=1$, the further the distance, the better accuracy the approximation. This is true because the round-off effect in $(1-z_0)$ decreases as z_0 is further away from $z=1$.

5.3 Asymptotic Solution For Buffer Overflow Probability

We now turn our attention to the asymptotic solution for the fading channel model in which there is only a finite number of buffer spaces. In particular, we are interested in the overflow probability. Section 4.4.3 has given us this probability $P_{ovfl}(B)$ from the infinite buffer length distribution $q(n)$. Here we look for a simple method which will give us the overflow probability without directly computing the infinite buffer length distribution.

In Chapter 3 we see that the tail distribution of the infinite buffer length can be approximated by a simple geometric form with coefficients c and d such that

Since the buffer overflow probability is dominated by the tail of the buffer distribution, it is desired to find a similar geometric and analytic solution for the overflow probability.

Since $P_{\text{ovfl}}(B)$ equals $p(B)$ (refer to Equation 2.12), and from the algorithm ALGM_FB for the computation of the finite buffer distribution (Appendix A), it is required to compute the constant $K = 1 / \sum_{i=0}^{B-1} q(i)$ for a finite buffer size B such that

$$P_{\text{ovfl}}(B) = p(B) = \frac{q(0) \cdot (K-1)}{1 + q(0) \cdot (K-1)}$$

Using the Laurent-Expansion tail approximation $\tilde{q}_{LE}(n) = cd^n$, we can then compute K in the following manners :

$$\begin{aligned} K &= \left[\sum_{i=0}^{B-1} q(i) \right]^{-1} \\ &= \left[1 - \sum_{i=B}^{\infty} q(i) \right]^{-1} \\ &= \left[1 - cd^B(1+d+d^2+\dots) \right]^{-1} \\ &= \left[1 - cd^B/(1-d) \right]^{-1} \end{aligned}$$

We have simplified the computation of K in terms of three parameters, c , d , and B . Since c and d can always be found from the Laurent-Expansion approximation (Section 5.2), the computation of K becomes rather simple. Hence, after some simplifications, $P_{\text{ovfl}}(B)$ can be written as :

$$\tilde{P}_{\text{ovfl}}(B) = \frac{f_0 \cdot cd^B}{1 - d - (1-f_0) cd^B} \quad (5.3)$$

in which f_0 equals $q(0)$, f_0 can be evaluated by (5.2).

As an example, we take the channel fade rate as $f_{\text{rms}}=1$ and $\rho=0.9$, the following results are obtained from the

Laurent-Expansion approximation given in the previous section (see Table 5.2a):

$$c = 1.0822 \times 10^{-2}$$

$$d = 0.9882$$

$$f_0 = 4.4939 \times 10^{-2}$$

Then, for $B=10$, $\tilde{P}_{ovfl}(B)=0.1648$, as compared to 0.1664 found in Section 4.4.3.

If the buffer size B is large enough that $\tilde{P}_{ovfl}(B)$ is small, $O(10^{-3})$, the term $(1-f_0)cd^B$ is much less than $(1-d)$ and can be ignored. Consequently, with only a small degree of further approximation, we get

$$\tilde{P}_{ovfl}(B) = \frac{f_0}{(1-d)} c d^B \quad (5.4)$$

From the above equation, we observe that the overflow probability can be approximated by a simple geometric form. We hereafter call this the second approximation of the overflow probability, $\tilde{P}_{ovfl}(B)$, (5.3) is referred to as the first approximation, $\bar{P}_{ovfl}(B)$.

In Tables 5.3 and 5.4 and Figures 5.3 and 5.4, the first and second approximations of the buffer overflow probabilities are compared with the exact values. From both figures, the first approximation is very good since the tail approximation of the infinite buffer distribution is very close. For the second approximation, when the overflow probability is less than 0.01, the approximation is acceptable.

22

OF / DE

(a)

 $\rho = 0.7$

B	$P_{\text{ovfl}}(B)$	$\tilde{P}_{\text{ovfl}}(B)$ (First Approx.)	$\hat{P}_{\text{ovfl}}(B)$ (Second Approx.)
10	1.4143×10^{-1}	1.4111×10^{-1}	7.6453×10^{-2}
20	7.5620×10^{-2}	7.5420×10^{-2}	5.1951×10^{-2}
30	4.4893×10^{-2}	4.4765×10^{-2}	3.5302×10^{-2}
40	2.8098×10^{-2}	2.8010×10^{-2}	2.3988×10^{-2}
50	1.8124×10^{-2}	1.8061×10^{-2}	1.6300×10^{-2}

(b)

 $\rho = 0.9$

B	$P_{\text{ovfl}}(B)$	$\tilde{P}_{\text{ovfl}}(B)$ (First Approx.)	$\hat{P}_{\text{ovfl}}(B)$ (Second Approx.)
10	0.1664	0.1648	3.6601×10^{-2}
20	0.1058	0.1051	3.2505×10^{-2}
30	7.5061×10^{-2}	7.4682×10^{-2}	2.8804×10^{-2}
40	5.6547×10^{-2}	5.6322×10^{-2}	2.5636×10^{-2}
50	4.4258×10^{-2}	4.4111×10^{-2}	2.2766×10^{-2}

Table 5.3 Comparisons Of The Overflow Probabilities
Of The Fading Channel Model, $f_{\text{rms}} = 1$ Hz

(a)

 $\rho = 0.7$

B	$P_{\text{ovfl}}(B)$	$\tilde{P}_{\text{ovfl}}(B)$ (First Approx.)	$\hat{P}_{\text{ovfl}}(B)$ (Second Approx.)
10	1.0364×10^{-2}	1.0372×10^{-2}	1.0009×10^{-2}
20	4.4521×10^{-4}	4.3494×10^{-4}	4.3428×10^{-4}
30	1.8893×10^{-5}	1.8845×10^{-5}	1.8844×10^{-5}
40	8.2236×10^{-7}	8.1763×10^{-7}	8.1763×10^{-7}
50	3.5796×10^{-8}	3.5477×10^{-8}	3.5477×10^{-8}

(b)

 $\rho = 0.9$

B	$P_{\text{ovfl}}(B)$	$\tilde{P}_{\text{ovfl}}(B)$ (First Approx.)	$\hat{P}_{\text{ovfl}}(B)$ (Second Approx.)
10	4.2665×10^{-2}	4.3030×10^{-2}	2.7315×10^{-2}
20	1.2122×10^{-2}	1.2196×10^{-2}	1.0486×10^{-2}
30	4.2353×10^{-3}	4.2544×10^{-3}	4.0254×10^{-3}
40	1.5775×10^{-3}	1.5779×10^{-3}	1.5453×10^{-3}
50	6.0329×10^{-4}	5.9797×10^{-4}	5.9323×10^{-4}

Table 5.4 Comparisons Of The Overflow Probabilities
Of The Fading Channel Model, $f_{\text{rms}} = 10$ Hz

Figure 5.3 Comparisons Of The Overflow Probabilities
Of The Fading Channel Model, $\rho=0.9$

Figure 5.4 Comparisons Of The Overflow Probabilities Of The Fading Channel Model, $\rho=0.7$

Chapter 6

Summary and Conclusion

This thesis considered the buffer analysis of an interrupted server with Poisson arrivals, slotted service intervals, and Markov interruption process of the first order. The analysis was carried out using the M/G/1 approach by approximating the service time distribution of the interrupted server. The basic framework for the M/G/1 analysis was given in Chapter 2. In particular, the well known Pollaczek-Khinchin (P-K) transform equation was presented. Although the P-K transform equation yields readily the probability generating function of the queue length distribution for the M/G/1 system, the complexity of the interrupted server model required a modified version of the P-K transform equation in order to account for the special effect of the slotted service. The probability generating function of the queue length distribution was inverted by a FFT inversion technique (Section 2.5).

A modified version of the P-K transform equation is developed. It is similar to the version proposed by Heines [3]. Nonetheless, the present version was derived in terms of the service time distribution. Consequently, it can handle more complex service interruption processes. The approximated service time distribution derived in this work

was justified by the close agreements between the infinite queue length distribution obtained here and similar results from [3] and [5].

This work also includes two approximation methods to estimate the tail of the infinite queue length distribution. The Heavy-Traffic approximation and the Laurent-Expansion approximation are outlined in Chapter 3. The Heavy-Traffic approximation, despite the fact that the traffic intensity must be very close to 1, is very easy to apply. It gives, in heavy traffic, asymptote to the tail of the infinite buffer distribution of any M/G/1 queues with very little computational effort, provided that the mean queue length can be found. The Laurent-Expansion approximation, however, offers very close estimate to the tail of the infinite buffer distribution, regardless of the traffic intensity. This approximation was justified by two different kinds of service characteristics. The Laurent-Expansion method can be applied to almost any kind of M/G/1 systems so long as they have a valid probability generating function.

Results also included are finite buffer probability and buffer overflow probability (Chapters 2, 4, and 5). The two probabilities do not require much extra computational effort from the infinite buffer distribution since the M/G/1 analysis simplifies the calculations. Finally, two approximations were derived for the overflow probability (Section 5.3), and they were shown to be quite effective.

These approximations are easy to apply and they give analytic results in terms of the buffer size. Thus, we may use the approximations as design tools for the consideration of finite buffer requirements.

In conclusion, this work has provided a framework to the study of buffer behavior of systems with complex service characteristics; particularly, systems with random service interruptions. Further work may be extended to M/G/c queues, or Markov interrupt processes of higher orders.

Appendix A

ALGM_FB --- An algorithm to compute the finite buffer distribution $p(n)$ from the infinite buffer distribution $q(n)$. Buffer size is B .

- (1) Obtain the infinite buffer distribution $q(n)$ using the P-K transform equation, or other methods.
- (2) From the distribution $q(n)$, compute :

$$K = \frac{1}{\sum_{i=0}^{B-1} q(i)}$$

- (3) Compute the value of $q(0)(K-1)$.

Note : $q(0)$ can be obtained from the probability generating function of the buffer distribution as $q^{(1)}(1)$.

- (4) Compute $p(n)$ using the formula :

$$p(n) = \begin{cases} \frac{K}{1 + q(0)(K-1)} q(n) & n < B \\ \frac{q(0)(K-1)}{1 + q(0)(K-1)} & n = B \end{cases}$$

Appendix B

IMSL ROUTINE NAME - FFT2C

PURPOSE - COMPUTE THE FAST FOURIER TRANSFORM OF A
COMPLEX VALUED SEQUENCE OF LENGTH EQUAL TO
A POWER TWO

USAGE - CALL FFT2C (A,M,IWK)

ARGUMENTS A - COMPLEX VECTOR OF LENGTH N, WHERE $N=2^M$.
ON INPUT A CONTAINS THE COMPLEX VALUED
SEQUENCE TO BE TRANSFORMED.
ON OUTPUT A IS REPLACED BY THE
FOURIER TRANSFORM.

M - INPUT EXPONENT TO WHICH 2 IS RAISED TO
PRODUCE THE NUMBER OF DATA POINTS, N
(I.E. $N = 2^M$).

IWK - WORK VECTOR OF LENGTH M+1.

PRECISION/HARDWARE - SINGLE AND DOUBLE/H32
- SINGLE/H36,H48,H60

REQD. IMSL ROUTINES - NONE REQUIRED

NOTATION - INFORMATION ON SPECIAL NOTATION AND
CONVENTIONS IS AVAILABLE IN THE MANUAL
INTRODUCTION OR THROUGH IMSL ROUTINE UHELP

REMARKS 1. FFT2C COMPUTES THE FOURIER TRANSFORM, X, ACCORDING
TO THE FOLLOWING FORMULA:

$$X(K+1) = \text{SUM FROM } J = 0 \text{ TO } N-1 \text{ OF} \\ A(J+1) * \text{CEXP}((0.0, (2.0 * \pi * J * K) / N)) \\ \text{FOR } K=0,1,\dots,N-1 \text{ AND } \pi=3.1415\dots$$

NOTE THAT X OVERWRITES A ON OUTPUT.

2. FFT2C CAN BE USED TO COMPUTE

$$X(K+1) = (1/N) * \text{SUM FROM } J = 0 \text{ TO } N-1 \text{ OF} \\ A(J+1) * \text{CEXP}((0.0, (-2.0 * \pi * J * K) / N)) \\ \text{FOR } K=0,1,\dots,N-1 \text{ AND } \pi=3.1415\dots$$

BY PERFORMING THE FOLLOWING STEPS;

```
DO 10 I=1,N
  A(I) = CONJG(A(I))
10 CONTINUE
CALL FFT2C (A,M,IWK)
DO 20 I=1,N
  A(I) = CONJG(A(I))/N
20 CONTINUE
```

Algorithm

FFT2C computes the fast Fourier transform (FFT) of a complex vector of length equal to a power of two.

The output vector X is defined mathematically as

$$X_{K+1} = \sum_{j=0}^{N-1} A_{j+1} e^{2\pi i j k / N} \quad \text{where } k=0,1,\dots,N-1; N=2^M; i=\text{SQRT}(-1)$$

Computationally, the above summation which defines X requires M^2 complex multiplications and additions. FFT2C utilizes a modification of the Singleton version of the Cooley-Tukey FFT algorithm which requires only $N \log_2 N$ basic sets of operations.

See reference:

Singleton, Richard C., "On computing the fast Fourier transform", Comm. ACM, 10(10)1967, 647-654.

Programming Notes

Note that the forward and inverse transforms are sometimes defined differently in the mathematical literature.

Example

This example computes the FFT of a sequence of length 8. Therefore, $M=3$ ($8=2^3$) and the input is $A(1), \dots, A(8)$.

Input:

```

INTEGER M,IWK(4)
COMPLEX A(8)
M      = 3
A(1)   = (1.0,2.0)
A(2)   = (4.0,1.0)
A(3)   = (3.0,3.0)
A(4)   = (0.0,1.0)
A(5)   = (2.0,4.0)
A(6)   = (3.0,0.0)
A(7)   = (2.0,2.0)
A(8)   = (1.0,3.0)
CALL FFT2C (A,M,IWK)
      .
      .
      .
END

```

Output:

```

A(1) = (16.0,16.0)
A(2) = (.121,1.12)
A(3) = (1.0,7.0)
A(4) = (-.707,-5.12)
A(5) = (0.0,6.0)
A(6) = (-4.12,-3.12)
A(7) = (-5.0,-5.0)
A(8) = (.707,-.879)

```

FFT2C-2

June, 1981

Appendix C

```
*****
*
*   PROGRAM NAME : FFT_INVERT
*
*   AUTHOR : ERIC HO
*
*   LANGUAGE : FORTRAN-77
*
*   DESCRIPTION :
*
*       The main function of this program is to invert a
*       Z-transform function using a FFT routine available from
*       the IMSL package. The FFT routine used is called 'FFT2C'
*       which handles complex function.
*
*       This program is specially written to perform the
*       inversion of the P-K transform equation for the infinite
*       queue length distribution. This program takes in 5 input
*       parameters, they are
*       (1) RHO
*       (2) ALPHA
*       (3) BETA
*       (4) N - the number of points of the FFT
*       (5) R - the radius of integration
*       On the output, besides the desired probability distribution,
*       some intermediate results are also given.
*
*   PROGRAM EXECUTION :
*
*       To run FFT_INVERT, the following steps are required :
*
*       ! FORTRAN FFT_INVERT OVER *G
*       ! LYNX *G OVER *F (UN=IMSLD.IMSL)
*       ! *F
*
*       These three execution steps will produce a run file in *F.
*
*****
C
C   ***** DECLARATION *****
C
C   DOUBLE COMPLEX A(512),CJ,Z,NZ,DZ,BZ,UZ,YZ
C   DIMENSION IMK(10)
C   DOUBLE PRECISION PI2,THETA,R,PI
C   DOUBLE PRECISION SUM,Q(0:511)
C   DOUBLE PRECISION RHO,ALPHA,BETA,LAMDA,A1,A2
C   DOUBLE PRECISION B1,B2,U1,V1,Y1,F0
C   DATA PI/3.1415926D0/
C
C   ***** READ IN PARAMETERS *****
C
```

```

WRITE (108,30)
10 FORMAT(' ENTER FUNCTION PARAMETERS: ')
OUTPUT ' RHO, ALPHA, BETA '
READ(105,20) RHO, ALPHA, BETA
20 FORMAT(4G)
WRITE(108,30)
30 FORMAT(' ENTER NO. POINTS, N AND RADIUS, R ')
OUTPUT, ' NOTE: N MUST BE A POWER OF 2 '
READ(105,20) N, R
WRITE (108,40)
40 FORMAT(' ENTER NO. OF CONSECUTIVE POINTS TO BE PRINTED ')
READ(105,20) M

C
C ***** TRIM N TO POWER OF TWO *****
C
CALL TRIM(N,NBITS)

C
SUM=0.00
CJ=CDSORT((-1.00,0.00))
PI2=2.000*PI
THETA=PI2/N

C
OUTPUT RHO
LAMDA=BETA/(ALPHA+BETA)*RHO

C
CALL GET GAMMA(ALPHA,BETA,LAMDA,A1,A2)
B1=(1.00-ALPHA-BETA)/(1.00-BETA)
B2=(ALPHA*BETA)/(1.00-BETA)
V1=LAMDA*(ALPHA+BETA)/BETA
U1=LAMDA*(A1+A2/BETA**2)
Y1=LAMDA/2.00
FO=(1.00-V1)/(1.00-V1+U1+Y1)
OUTPUT FO

C
DO 200 L=0,(N-1)
Z=1.00/(R*CDEXP(CJ*L*THETA))
BZ=B1*CDEXP(LAMDA*(Z-1.00))+B2/(CDEXP(LAMDA*(1.00-Z))-1.00+BETA)
UZ=A1*CDEXP(LAMDA*(Z-1.00))+A2/(CDEXP(LAMDA*(1.00-Z))-1.00+BETA)
YZ=(1.00-CDEXP(LAMDA*Z-LAMDA))/(LAMDA-(LAMDA*Z))
NZ=FO*(BZ-Z*UZ*YZ)
DZ=BZ-Z
A(L+1)=NZ/DZ
200 CONTINUE

C
C ***** INVERT THE Z-TRANSFORM. THE ARRAY A IS USED *****
C ***** BOTH TO SUPPLY THE VECTOR OF VALUES TO BE *****
C ***** INVERTED AND TO RETURN THE RESULTING TRANS- *****
C ***** FORM. THE LENGTH OF A IS TAKEN TO BE *****
C ***** 2**NBITS. *****
C
CALL FFT2C(A,NBITS,IWK)
OUTPUT 'PASSED FFT2C'

C
C ***** THE RESULT IS SCALED *****
C

```

```

      DO 300 L=0,(N-1)
        A(L+1)=A(L+1)*R**L/N
300  CONTINUE
C
C
      DO 400 L=0,(N-1)
        Q(L)=DREAL(A(L+1))
        SUM=SUM+Q(L)
400  CONTINUE
C
      OUTPUT SUM
      OUTPUT '
      OUTPUT ' BUFFER LENGTH      PROBABILITY DISTRIBUTION '
      OUTPUT '           N              Q(N) '
      OUTPUT '
      DO 500 I=0,M
        WRITE(108,1000) I,Q(I)
500  CONTINUE
C
      OUTPUT '
      DO 600 I=(M+10),(N/10*10),10
        WRITE(108,1000) I,Q(I)
600  CONTINUE
C
1000 FORMAT(8X,I3,10X,E12.6)
C
      STOP
      END
C
C
C
*****
*
*   SUBROUTINE TRIM(N,NBITS)
*
*   The returned value of N is the greatest power of
*   two not exceeding the input value of N. NBITS is set
*   to the corresponding number of bits.
*
*****
C
      SUBROUTINE TRIM(N,NBITS)
        NBITS=0
        NCOPY=N/2
        N=1
10     IF(NCOPY.EQ.0) GOTO 20
        NCOPY=NCOPY/2
        N=2*N
        NBITS=NBITS+1
        GOTO 10
      RETURN
      END

```



```

*****
*
*      SUBROUTINE GET GAMMA(ALPHA,BETA,LAMDA,A1,A2)
*
*      This subroutine computes the value of GAMMA and
*      return two values A1 and A2 which are functions of
*      GAMMA. GAMMA is the probability that the channel
*      is uninterrupted at the start of a transmission.
*
*****
C
      SUBROUTINE GET GAMMA(ALPHA,BETA,LAMDA,A1,A2)
      DOUBLE PRECISION GAMMA,K1,K2
      DOUBLE PRECISION ALPHA,BETA,LAMDA,A1,A2
C
C
      K1=(BETA/(ALPHA+BETA))*DEXP(-1.DO*LAMDA)/(1.DO-DEXP(-1.DO*LAMDA))
      K2=(ALPHA/(ALPHA+BETA))*DEXP(-1.DO*(ALPHA+BETA+LAMDA))/
      (1.DO-DEXP(-1.DO*(ALPHA+BETA+LAMDA)))
      GAMMA=(DEXP(LAMDA)-1.DO)*(K1+K2)
      A1=GAMMA*(1.DO-ALPHA-BETA)/(1.DO-BETA)
      A2=(GAMMA*ALPHA*BETA)/(1.DO-BETA)+(1.DO-GAMMA)*BETA
      OUTPUT GAMMA
      RETURN
      END

```

LEFT RUN.

ENTER FUNCTION PARAMETERS:

RHO, ALPHA, BETA

70.9, 0.02910, 0.03028

ENTER NO. POINTS, N AND RADIUS, R

NOTE: N MUST BE A POWER OF 2

7256, 1.05

ENTER NO. OF CONSECUTIVE POINTS TO BE PRINTED

720

RHO = .900000000000000000

GAMMA = .930150356458071625

FO = .449379446625369056E-01

PASSED FFT2C

SUM = .949031695644556392

BUFFER LENGTH N	PROBABILITY DISTRIBUTION Q(N)
--------------------	----------------------------------

0	.449379E-01
1	.394096E-01
2	.231486E-01
3	.145420E-01
4	.114762E-01
5	.105163E-01
6	.101795E-01
7	.100070E-01
8	.987620E-02
9	.975653E-02
10	.964056E-02
11	.952652E-02
12	.941395E-02
13	.930274E-02
14	.919285E-02
15	.908427E-02
16	.897697E-02
17	.887093E-02
18	.876615E-02
19	.866260E-02
20	.856028E-02
30	.760123E-02
40	.674964E-02
50	.599345E-02
60	.532199E-02
70	.472575E-02
80	.419632E-02
90	.372621E-02
100	.330877E-02
110	.293811E-02
120	.260898E-02
130	.231672E-02
140	.205720E-02
150	.182671E-02
160	.162196E-02
170	.143992E-02
180	.127781E-02
190	.113287E-02
200	.100211E-02
210	.881711E-03
220	.765678E-03
230	.642334E-03
240	.483207E-03
250	.186942E-03

STOP

Sample Run No. 2 : $f_{rms} = 10 \text{ Hz}$

IFFT_RUN.

ENTER FUNCTION PARAMETERS:

RHO, ALPHA, BETA

70.7, 0.2910, 0.3028

ENTER NO. POINTS, N AND RADIUS, R

NOTE: N MUST BE A POWER OF 2

7128, 0.8

ENTER NO. OF CONSECUTIVE POINTS TO BE PRINTED

710

RHO = .700000000000000000

GAMMA = .642346752158639925

FD = .222265164016270344

PASSED FFT2C

SUM = 1.00000922529364510

BUFFER LENGTH N	PROBABILITY DISTRIBUTION Q(N)
0	.222268E+00
1	.215206E+00
2	.155784E+00
3	.110669E+00
4	.799545E-01
5	.582456E-01
6	.425302E-01
7	.310709E-01
8	.227013E-01
9	.165866E-01
10	.121189E-01
20	.525448E-03
30	.227823E-04
40	.987790E-06
50	.428285E-07
60	.185695E-08
70	.805137E-10
80	.349089E-11
90	.151353E-12
100	.656128E-14
110	.284288E-15
120	.122881E-16

STOP

Sample Run No. 3 : $f_{rms} = 17 \text{ Hz}$

FFT_RUN.

ENTER FUNCTION PARAMETERS:

RHO, ALPHA, BETA

70.9, 0.49, 0.51

ENTER NO. POINTS, N AND RADIUS, R

NOTE: N MUST BE A POWER OF 2

7256, 0.95

ENTER NO. OF CONSECUTIVE POINTS TO BE PRINTED

710

RHO = .900000000000000000

GAMMA = .596447670040579476

FO = .813338755591703947E-01

PASSED FFT2C

SUM = .999999866415608218

BUFFER LENGTH N	PROBABILITY DISTRIBUTION Q(N)
--------------------	----------------------------------

0	.813339E-01
1	.109005E+00
2	.103836E+00
3	.917350E-01
4	.799560E-01
5	.695591E-01
6	.605016E-01
7	.526226E-01
8	.457695E-01
9	.398090E-01
10	.346246E-01
20	.857885E-02
30	.212556E-02
40	.526644E-03
50	.130485E-03
60	.323302E-04
70	.801044E-05
80	.198476E-05
90	.491776E-06
100	.121854E-06
110	.301950E-07
120	.748277E-08
130	.185447E-08
140	.459558E-09
150	.113786E-09
160	.280489E-10
170	.683219E-11
180	.158792E-11
190	.309206E-12
200	.109957E-13
210	-.478988E-13
220	-.512322E-13
230	-.442094E-13
240	-.379340E-13
250	-.369268E-13

STOP

Appendix D

Derivatives of the Equivalent Service Time Distribution

Notation :

$$(1) A1 = (1-\alpha-\beta)/(1-\beta)$$

$$(2) A2 = \alpha\beta/(1-\beta)$$

$$(3) A3 = \gamma A1$$

$$(4) A4 = \gamma A2 + (1-\gamma)\beta$$

$$(5) \phi = \exp(s)$$

Formulae :

$$(1) BV^*(s) = A1 \phi^{-1} + A2 (\phi - 1 + \beta)^{-1}$$

$$(2) BV^{*'}(s) = A1 \phi^{-1} + A2 \phi (\phi - 1 + \beta)^{-2}$$

$$(3) BV^{*''}(s) = A1 \phi^{-1} + A2 \phi (\phi - 1 + \beta)^{-2} [2(\phi - 1 + \beta)^{-1} - 1]$$

$$(4) BW^*(s) = U^*(s) Y^*(s) \\ = (1 - \phi^{-1})/s [A3 \phi^{-1} + A4 (\phi - 1 + \beta)^{-1}]$$

$$(5) BW^{*'}(s) = U^{*'}(s) Y^*(s) + U^*(s) Y^{*'}(s)$$

$$(6) BW^{*''}(s) = U^{*''}(s) Y^*(s) + U^{*'}(s) Y^{*'}(s) + U^*(s) Y^{*''}(s)$$

$$(7) U^*(s) = (1 - \phi^{-1})/s$$

$$(8) U^{*'}(s) = (1 - \phi^{-1}(1+s))/s^2$$

$$(9) U^{*''}(s) = (2 - \phi^{-1}(s^2 + 2s + 2))/s^3$$

$$(10) Y^*(s) = A3 \phi^{-1} + A4 \phi (\phi - 1 + \beta)^{-1}$$

$$(11) Y^{*'}(s) = -A3 \phi^{-1} - \phi A4 (\phi - 1 + \beta)^{-2}$$

$$(12) Y^{*''}(s) = A3 \phi^{-1} + A4 \phi (\phi - 1 + \beta)^{-2} (2(\phi - 1 + \beta)^{-1} - 1)$$

Intermediate Results :

$$(1) U^*(0) = 1$$

$$(2) Y^*(0) = 1$$

$$(3) BV^*(0) = 1$$

$$(4) BV^{*'}(0) = -(\alpha + \beta)/\beta$$

$$(5) BV^{**}(0) = 1 + \alpha(\beta+2)/\beta^2$$

$$(6) U^{*'}(0) = -1/2$$

$$(7) U^{**}(0) = 1/3$$

$$(8) Y^{*'}(0) = -(\gamma + (1-\gamma)/\beta)$$

$$(9) Y^{**}(0) = \gamma(1 + \alpha(\beta+2)/\beta^2) + (1-\gamma)(2/\beta - 1)/\beta$$

$$(10) BW^{*'}(0) = -(\gamma(\alpha + \beta) + (1-\gamma) + \beta/2)/\beta$$

$$(11) BW^{**}(0) = \gamma \left(\frac{(\alpha + \beta)(4-3\beta) - 1}{\beta(1-\beta)} \right) + (4/\beta - 2/3)$$

Equalities :

$$(1) B^{*(1)}(\lambda - \lambda s) = -\lambda B^{*'}(s)$$

$$(2) B^{*(2)}(\lambda - \lambda s) = \lambda^2 B^{**}(s)$$

Appendix E

In this appendix, we show that the service time distribution $B^*(\lambda - \lambda z)$ has two values of z in a closed interval $1 \leq z \leq 1+h$ such that $B^*(\lambda - \lambda z) = z$. We also derive the value of h from a theorem, which is an extension of the Theorem of the Mean but at the same time a special case of the Taylor's Theorem with Remainder [15].

The theorem states : suppose that a function f and its first derivative f' are continuous on the closed interval I with a and $a+h$ as endpoints and $f''(x)$ exists for each x between a and $a+h$. Then there exist a ξ between a and $a+h$ such that

$$\Delta f(a, h) - df(a, h) = 1/2 f''(\xi) h^2 \quad (1)$$

where $\Delta f(a, h) = f(a+h) - f(a)$,

and $df(a, h) = f'(a) h$.

In case $|f''(x)| \geq |f''(a)|$ and $f''(x)$ and $f''(a)$ have the same sign for any x interior to I , (1) can be expressed as

$$\Delta f(a, h) - df(a, h) \geq 1/2 f''(a) h^2 \quad (2)$$

Now, if we define $f(z) = B^*(\lambda - \lambda z)$ and $a=1$, we have

$$f(1) = 1$$

$$f'(1) = B^{*(1)}(\lambda - \lambda z) \Big|_{z=1} = \lambda \bar{x}$$

$$f''(1) = B^{*(2)}(\lambda - \lambda z) \Big|_{z=1} = \lambda^2 \bar{x}^2$$

Since $B^*(s)$ is the Laplace Transform of the service time distribution, $B^*(s)$ and $B^{*(1)}(s)$ must be continuous on the closed interval I and that $B^{*(2)}(s)$ is also continuous, provided h is sufficiently small. Also, since $s = \lambda - \lambda z$, $z=1$ and $z=1+h$ corresponds to $s=0$ and $s=-\lambda h$ respectively.

We can find the signs of the service time distribution and its derivatives from :

$$\begin{aligned} B^*(s) &= \int_0^{\infty} e^{-sx} b(x) dx \\ B^{*'}(s) &= - \int_0^{\infty} x e^{-sx} b(x) dx \\ B^{*''}(s) &= \int_0^{\infty} x^2 e^{-sx} b(x) dx \end{aligned}$$

Since e^{-sx} and $b(x)$ are always positive, we deduce that both $B^*(s)$ and $B^{*''}(s)$ are always positive, and $B^{*'}(s)$ is always negative. However, since we have

$$\begin{aligned} B^{*(1)}(\lambda - \lambda z) &= -\lambda B^{*'}(s) \\ \text{and} \quad B^{*(2)}(\lambda - \lambda z) &= \lambda^2 B^{*''}(s) \end{aligned}$$

the service time distribution $B^*(\lambda - \lambda z)$ and its higher derivatives are always positive.

From $B^{*''}(s)$, we can deduce that for s less than 0, $B^{*''}(s_1) > B^{*''}(s_2)$ if $|s_1| > |s_2|$. Under this circumstance, we can use (2) to determine the value of h .

To evaluate h , let us assume that for $z=1+h$, $B^*(\lambda - \lambda z)$ equals z .

Hence,

$$\begin{aligned} \Delta f(1+h) &= f(1,h) - f(1) \\ &= (1+h) - 1 \\ &= h \\ df(1,h) &= f'(1) h \end{aligned}$$

Thus from (2)

$$h - f'(1)h \geq 1/2 f''(1) h^2 \quad (3)$$

From (3), we can determine h :

$$h = 0$$

or

$$h \leq \frac{2(1-f'(1))}{f''(1)} = \frac{2(1-\lambda\bar{x})}{\lambda^2 \bar{x}^2} = \frac{1}{Nq}$$

Since $B^{*(2)}(\lambda - \lambda z)$ is always positive for z interior to I , the locus of $B^*(\lambda - \lambda z)$ is concave upward on I . In this case, the curves $g(z)=z$ and $f(z)=B^*(\lambda - \lambda z)$ can only intersect at two points. Figure E.1 shows some popular M/G/1 queues' service time distribution as a function of z . All four graphs depict the characteristic of the concave upward service time distribution on interval I . Note that if the traffic intensity very close to 1, function $g(z)=z$ becomes the tangent to $B^*(\lambda - \lambda z)$ at the point $z=1$. Thus, the concave upward service time distribution on I removes any doubts that there are more than two values of z such that $B^*(\lambda - \lambda z)=z$.

Figure E.1 Displays Of Some Popular M/G/1 Queues' Service Time Distributions As a Function Of z .

Appendix F

Program FIND_ROOT

```

*****
C
C      AUTHOR : ERIC HU
C
C      LANGUAGE : FORTRAN-77
C
C      DESCRIPTION :
C
C          THIS IS A PROGRAM TO COMPUTE THE ROOT OF A FUNCTION,
C          USING THE NEWTON RAPHSON METHOD. THE FUNCTION AND ITS
C          FIRST DERIVATIVE HAVE TO BE ENTER INTO THE PROGRAM THROUGH
C          EDITING CERTAIN LINES OF CODES.
C
C          THIS PROGRAM IS TAILORED SPECIALLY FOR THE FUNCTION D(Z)
C          OF THE P-K TRANSFORM EQUATION. THE INPUT PARAMETERS ARE :
C          (1) KHI
C          (2) ALPHA
C          (3) BETA
C          (4) ROOT
C
C          AN ESTIMATED VALUE OF THE ROOT IS REQUIRED FOR THE STARTING
C          VALUE OF THE NEWTON RAPHSON METHOD.
C          THE OUTPUT CONSISTS OF THE VALUE OF THE DESIRE ROOT AND B,
C          THE COEFFICIENT OF THE LAURENT-EXPANSION.
C
*****
C
C      ***** DECLARATION *****
C
C      COMPLEX ROOT,OLDROOT,NEURROOT,Z
C      REAL A1,A2,B1,B2,K1,K2,V1,U1,Y1
C      REAL B,RHO,ALPHA,BETA
C      COMPLEX DZ,DDZ,NZ
C      COMPLEX BZ,UZ,YZ
C      REAL LAMDA
C      REAL NORM
C
C      OUTPUT ' ENTER RHO,ALPHA,BETA : '
C      READS,RHO,ALPHA,BETA
C
C      LAMDA=BETA/(ALPHA+BETA)*RHO
C      K1=(BETA/(ALPHA+BETA))*EXP(-LAMDA)/(1-EXP(-LAMDA))
C      K2=(ALPHA/(ALPHA+BETA))*EXP(-1*(ALPHA+BETA+LAMDA))/
C      (1-EXP(-1*(ALPHA+BETA+LAMDA)))
C      GAMMA=(EXP(LAMDA)-1)*(K1+K2)
C      A1=GAMMA*(1-ALPHA-BETA)/(1-BETA)
C      A2=(GAMMA*ALPHA*BETA)/(1-BETA)+(1-GAMMA)*BETA
C      B1=(1-ALPHA-BETA)/(1-BETA)
C      B2=(ALPHA*BETA)/(1-BETA)
C      V1=LAMDA*(ALPHA+BETA)/BETA
C      U1=LAMDA*(A1+A2/BETA*B2)
C      Y1=LAMDA/2.0
C      NORM=(1-V1)/(1-V1+U1+Y1)
C      OUTPUT ' ENTER AN ESTIMATED ROOT : '
C      READS,ROOT
C      OLDROOT=ROOT
C      NEURROOT=OLDROOT
C
C      ***** ITERATION OF THE NEWTON RAPHSON METHOD *****
C
C      10 CONTINUE
C      OLDROOT=NEURROOT
C      Z=NEURROOT
C      BZ=B1*CEXP(LAMDA*(Z-1))+B2/(CEXP(LAMDA*(1-Z))-1+BETA)
C      UZ=A1*CEXP(LAMDA*(Z-1))+A2/(CEXP(LAMDA*(1-Z))-1+BETA)
C      YZ=(1-CEXP(LAMDA*(Z-LAMDA)))/(LAMDA-(LAMDA*Z))
C      NZ=NORM*(BZ-Z*UZ*YZ)
C      DZ=BZ-Z
C      DDZ=(B1*CEXP(LAMDA*(Z-1))*LAMDA+LAMDA*CEXP(LAMDA*(1-Z))*
C      B2/(CEXP(LAMDA*(1-Z))-1+BETA)*Z)-1
C
C      NEURROOT=OLDROOT-DZ/DDZ
C
C      IF (ABS((NEURROOT-OLDROOT)).GE.0.0001) GOTO 10
C      ROOT=NEURROOT
C
C      R=NZ/DDZ
C
C      OUTPUT ROOT,R
C      STOP
C      END

```

Sample Run of FIND_ROOT :

!#F

ENTER RHO,ALPHA,BETA :
?0.7,0.02910,0.03028

ENTER AN ESTIMATE ROOT :
?1.0783

ROOT = (1.039352 ,0.000000)
B = -.3086337E-01
STOP

!#F

ENTER RHO,ALPHA,BETA :
?0.9,0.02910,0.03028

ENTER AN ESTIMATE ROOT :
?1.0148

ROOT = (1.011954 ,0.000000)
B = -.1094789E-01
STOP

!#F

ENTER RHO,ALPHA,BETA :
?0.99,0.02910,0.03028

ENTER AN ESTIMATE ROOT :
?1.0012

ROOT = (1.001189 ,0.000000)
B = -.1124725E-02
STOP

References

- [1]---J. Hsu, "Buffer Behavior with Poisson Arrival and Geometric Output Process", IEEE Transaction on Communications, Dec.1974, pp.1940-1941.
- [2]---N.D. Georganas, "Buffer Behavior with Poisson Arrivals and Bulk Geometric Service", IEEE Transactions on Communications, August 1976, pp.938-940.
- [3]---T.S. Heines, "Buffer Behavior in Computer Communication Systems", IEEE Transactions on Computers, Vol.C-28, August 1979.
- [4]---H.B. Kekre and C.L. Saxena, "Finite buffer behavior in an integrated digital voice data system for Erlang Arrival", Performance of Data Communication Systems and their Applications, North-Holland Publishing Company, 1981, pp.55-65.
- [5]---J.k. Cavers and C.M. Woodside, "Buffer Size for Data Communications Over Fading Radio Link", Computer Networks, North-Holland Publishing Company, Vol.4, 1980, pp.167-174.
- [6]---J.F.C. Kingman, "The Heavy Traffic Approximation in The Theory of Queues", Proceedings of the Symposium on Congestion Theory, 1964, pp.137-159.
- [7]---J.K. Cavers, "On the fast Fourier transform inversion of probability generation functions", J. Inst. Maths. Application, Vol.22, 1978, pp.275-282.
- [8]---D.G. Kendall, "Some Problems in the Theory of Queues", Journal of the Royal Statistical Society, Ser.B-13, 1951, pp-151-185.
- [9]---L. Kleinrock, "Queueing Systems, Volume I:Theory", John Wiley & Sons, Inc., 1975, pp.180-196.
- [10]---D. Gross and C.M. Harris, "Fundamental of Queueing Theory", Wiley, 1974, pp.250-255.
- [11]---W.W. Chu, "Buffer behavior for Poisson arrivals and multiple synchronous constant output", IEEE Transactions on Computers, Vol.C-19, No.6, June 1970 pp.530-534.

- [12]---Y. Takashashi, "Asymptotics Exponentiality of the Tail of the Waiting-time Distribution in a PH/PH/c Queue", Advances in Applied Probability, 1981, pp.619-630.
- [13]---k.Geihs and H.kobayash, "Bounds on Buffer Overflow Probability in Communication Systems", Performance Evaluation 2(1980), North-Holland Publishing Company, 1982.
- [14]---E.Fuchs and P.E. Jackson, "Estimates of distribution of random variable for certain computer communications traffic models", Proc. ACM Symp. Problems on the optimization Data Communication Systems, Pine Mountain, GA, Oct.1969, pp.202-225.
- [15]---C.B. Morrey, "University Calculus with Analytic Geometry", Addison-Wesley Publishing Company, Inc., 1962. ppl44-148.

END

1 2 0 6 8 6

FIN