Investigations on Stochastic Storage Systems with Positive Service Time

Thesis submitted to

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

in partial fulfillment of the requirements

for the award of the degree of

DOCTOR OF PHILOSOPHY

under the Faculty of Science by

Manikandan R

Department of Mathematics
Cochin University of Science and Technology
Kochi - 682 022, Kerala, India

October 2013

Investigations on Stochastic Storage Systems with Positive Service Time

Ph.D. thesis in the field of Stochastic Modelling & Analysis

Author:
Manikandan R
Department of Mathematics
Cochin University of Science and Technology
Kochi - 682 022, Kerala, India
Email: rangaswamy.mani@gmail.com

Supervisor:
Dr. A. Krishnamoorthy
Emeritus Professor
Department of Mathematics
Cochin University of Science and Technology
Kochi - 682 022, Kerala, India.
Email: achyuthacusat@gmail.com

October 2013

To Amma, my inspiration

&

In memory of Achan

Dr. A. Krishnamoorthy
Emeritus Professor
Department of Mathematics
Cochin University of Science and Technology
Kochi - 682 022, India.

24th October 2013

Certificate

Certified that the work presented in this thesis entitled “Investigations on Stochastic Storage Systems with Positive Service Time ”is based on the authentic record of research carried out by Shri. Manikandan R under my guidance in the Department of Mathematics, Cochin University of Science and Technology, Kochi-682 022 and has not been included in any other thesis submitted for the award of any degree.

A. Krishnamoorthy
(Supervising Guide)

Phone : +91 484 2577518 +91 484 2862468 Email: achyuthacusat@gmail.com

Declaration

I hereby declare that the work presented in this thesis entitled "Investigations on Stochastic Storage Systems with Positive Service Time "is based on the original research work carried out by me under the supervision and guidance of Dr. A. Krishnamoorthy, Emeritus Professor, Department of Mathematics, Cochin University of Science and Technology, Kochi–682 022 and has not been included in any other thesis submitted previously for the award of any degree.

Manikandan R

Kochi– 682 022
24^{th} October 2013

Acknowledgments

This thesis is the fulfillment of my desire of research that formed in early studies of undergraduate mathematics. In the longest path of the desire, I am indebted to a lot of individuals for the realization of the thesis.

I would like to express my overwhelming gratitude to my research guide, Dr. A. Krishnamoorthy, for his expertise shown in guiding my work and the willingness to share his knowledge and experience. He has given immense freedom for us in developing ideas and he is always willing to hear and acknowledge sincere efforts. His immense knowledge and critical but valuable remarks led me to do a good research. I especially thank him for his prompt reading and careful critique of my thesis. Throughout my life I will benefit from the experience and knowledge I gained working with Dr. A. Krishnamoorthy.

I express my sincere thanks to Prof. P. G. Romeo, Head, Department of Mathematics and all the former Heads of the Department - Prof. R. S. Chakravarti, Prof. A. Vijayakumar - for permitting me to use the research facilities in the Department. I would like to thank my doctoral committee member Prof. M. N. Narayanan Namboodiri for his support and advices. I deeply obliged to Dr. B. Lakshmi for her support and inspiration during my research work. I gratefully acknowledge the help from Prof. M. Jadhavedan. I thank other faculty members of the department Mrs. Meena, Mr. Shyam Sunder Iyar and office staff of the Department of Mathematics for their help of various kinds. My gratitude also goes to the authorities of CUSAT for the facilities they provided. I thank University Grants Commission, Government of India for financially supporting me during this endeavor.

The period of my research at CUSAT was made enjoyable in large part due to the many friends who have became a part of my life. It is my pleasure to acknowledge the advice and love received from my senior researchers in my research area Dr. T.G. Deepak, Dr. S. Babu, Dr. Varghese Jacob, Dr. C. Sreenivasan, Dr. Viswanath C. Narayanan, Dr. Pramod P.K., Dr. Ajayakumar C. B., Mrs. Deepthi C.P. and Mrs. Resmi Varghese for their interest in my research work and they were always ready to share their bright ideas with me. I extend my heartfelt gratefulness to my friend Ms. Dhanya Shajin for her valuable help in my research work.

I express my sincere thanks to Dr. Sajeev S. Nair, Government Engineering College, Thrissur for spending his valuable time for the discussions with me and also for permitting to include our joint work in this thesis. I gratefully acknowledge the support and help received from Mr. Anoop N. Nair, NIT Calicut, during my research period. I find immensely enjoyable working with my dear friends and teachers- Dr. Seema Varghese, Mr. Pravas K, Mr. Manjunath A.S., Mr. Jayaprasad P. N., Mr. Jaison Jacob, Mr. Didimos K. V., Mr. Tonny K.B., Mr. Gireesan K.K., Mrs. Anu Varghese, Mrs. Pamy Sebastian, Mrs. Raji George, Mrs. Chithra M.R., Mrs. Seethu Varghese, Savitha teacher and Jaya teacher. The discussions, evening tea at ICH & TCH, conferences and parties have all bonded us as life time friends. I feel wordless to appreciate the selfless support, care and ever encouraging words of my dear friend Tibin Thomas. I express my feeling of gratitude to my friends Ambily, Jayan, Raveesh, Aneesh, Jinto, Neeraj Narayanan, Vikas, Lindo Ouseph, Ullas Kalappura, Sujith Raman (Uppsala University, Sweden), Vimalettan (ICTP-UNESCO-IAEA, Italy), Maneshettan., Satheesh Kumar (AIMS) and Prem Kumar.

I would like to thank all of my colleagues at the department of Mathematics especially Kiran Kumar V. B, Tijo James, Balesh K., Sajeesh V.K.

and all other friends in CUSAT for their great support during my Ph.D period. I express my sincere thanks to my B.Sc, M.Sc and M. Phil friends - Taju Joseph, Cynto Paul, Santhosh, Binoy Joseph, Prasanth, Diana, Abdul Rof, Amalettan and Shajeebettan for their love and support.

I record my deep and utmost gratitude to my mother, her love and care during the entire period of my research work. I would like to thank my brothers Chandrakumar, Radhakrishnan and Balakrishnan for the love and care.

Finally I thank all my well wishers and friends who have supported me in this venture. Above all, I bend my head in humble gratitude before the Omnipotent God Almighty for kindly allowing me to finish my work in time.

Manikandan R

Contents

Preface

Research on queueing systems with attached inventory has captured much attention of researchers over the last two decades. Inventory models are studied in detail in Churchman, Acoff and Arnoff [14], Hadley and Whitin [20], Naddor [46], and in Sahin [59] and in a number of research papers. In the first three, a large number of deterministic models are discussed whereas in the book by Sahin, stochastic models are highlighted. We call these models and problems as *Classical type*, since in all these the amount of time required to serve the item is negligible.

In contrast most of the real life situations need positive amount of time to serve the inventory. Such cases are referred to as *inventory with positive service time*. It may appear that there is no difference between a *queue* and an *inventory with positive service time*. However, this is not the reality. In a queue we do not speak about the resources for service – if the customers are available and server is ready to serve then the service starts. Nevertheless, this is not the case in inventory with positive service time. Server may be available to serve and there may be customers waiting to get service. However, inventory may not be available on stock. Thus a queue of customers builds up. Even in the case when lead time is zero, the above problem can very well arise. Needless to say that in the case of positive lead time the server may remain idle even when customers are waiting for want of items in the inventory.

In this thesis the queueing-inventory models considered are analyzed as continuous time Markov chains in which we use the tools such as matrix analytic methods. We obtain the steady-state distributions of various queueing-inventory models in product form under the assumption that no customer joins the system when the inventory level is zero. This is despite the strong correlation between the number of customers joining the system and the inventory level during lead time. The resulting quasi-birth-and-death (QBD) processes are solved explicitly by matrix geometric methods.

Matrix analytic methods introduced by M.F. Neuts in the second half of the 1970's, establish a success story, illustrating the enrichment of science and applied probability. Since then, matrix analytic methods have become an indispensable tool in stochastic modeling and have found applications in the analysis and design of manufacturing systems, telecommunications networks, risk/insurance models, reliability models and inventory and supply chain systems. The power and popularity of matrix analytic methods come from their flexibility in stochastic modeling, capacity for analytic exploration, natural algorithmic thinking and tractability in numerical computation.

Part of the work presented in this thesis has been published/communicated to journals

1. A survey on inventory models with positive service time, Krishnamoorthy, A., Lakshmy, B. and **Manikandan, R.** OPSEARCH (Springer), 48 (2), 153–169, 2011.

2. A revisit to queueing-inventory system with positive service time, Krishnamoorthy, A., **Manikandan, R.** and Lakshmy, B. Annals of Operations Research (Springer), DOI 10.1007/s10479-013-1437-x, 2013.

3. Production Inventory with Positive Service Time and Loss, Krishnamoorthy, A., **Manikandan, R.** and Lakshmy, B. Proceedings of the Eighteenth Ramanujan Symposium On Recent Trends in Dynamical Systems and Mathematical Modelling, University of Madras, Chennai, 25-27, September, 2013, Volume 18, 57-64, 2014.

4. On a two stage supply chain inventory with positive service time and loss, Krishnamoorthy, A. and **Manikandan, R.** (Under review).

5. Analysis of a multi-server queueing-inventory system, Krishnamoorthy, A., **Manikandan, R.** and Dhanya Shajin (Under review).

6. Classical queueing-inventory system with working vacation and vacation interruptions, Krishnamoorthy, A., **Manikandan, R.** and Sajeev, S. Nair (Under review).

7. Retrial of unsatisfied customers in a queueing-inventory system, Krishnamoorthy, A., **Manikandan, R.** and Sajeev, S. Nair (Under review).

Part of the work include in the thesis has been presented in various international conferences

1. (s, S) Inventory with Service Time in Tandem Queues, Krishnamoorthy, A., **Manikandan, R.,** and Lakshmy, B. (presented as invited talk), International Conference on Mathematical and Computational Models (ICMCM-2011), VEL TECH, Technical University, India.

2. (s, S) Inventory System with Retrial of Unsatisfied Demands in the Final Stage of a Tandem Queue, Krishnamoorthy, A., and **Manikandan, R.** (presented as plenary talk), International Conference on Stochastic Modeling and Simulation (ICSMS-2011), PSG College of Technology, India.

3. On an inventory with service time, Krishnamoorthy, A., **Manikandan, R.,** and Lakshmy, B. (presented as invited talk in the series of sessions on Stochastic Modelling), International Conference on Frontiers of Statistics and its Applications (ICFSA-2012), Pondicherry University, India.

4. Production inventory with positive service time and loss, Krishnamoorthy, A. and **Manikandan, R.** (presented as invited talk), XVIII Ramanujan Symposium on Recent Trends in Dynamical Systems and Mathematical Modelling, 2013, Ramanujan Institute of Advanced Study in Mathematics, University of Madras, Chennai, India.

5. Multi server queueing-inventory system with positive service time, Krishnamoorthy, A., **Manikandan, R.** and Dhanya Shajin (presented as invited talk), International Conference on Applied Mathematical Models (ICAMM 2014), PSG College of Technology, India.

Chapter 1

Introduction to queueing-inventory system

Inventory management is one of the most important tasks in commercial world. Inventory can be found everywhere and is an obedient companion of many human activities. Books in a bookstore, food in a refrigerator, goods in a supermarket, cars to be sold, and spare parts to be used, are all inventory of some kind. Inventory takes up space and ties up with cash/resource, which might be scarce or can be used somewhere else. In the case of business faces inventory problems in its most basic activities. Inventory is held by the selling party to meet the demand made by the buying party. The complexity of inventory problems varies significantly, depending on the situation. Consequently, inventory management becomes an issue of interest. Some of the inventory problems that arise in complex business processes

---

Part of this chapter appeared in the following paper.
A. Krishnamoorthy, B. Lakshmy and R. Manikandan : A survey on inventory models with positive service time. OPSEARCH, 48 (2), 153–169, 2011.

require sophisticated mathematical tools and advanced computing power to get a reasonably good solution. Inventory models usually consist of a demand process, goods in a warehouse, and a replenishment process of ordered goods. Thus the fundamental questions of inventory models can be described as follows: (1) when should an order be placed? and (2) how much should be ordered? Thus in inventory management, finding the optimal policy is the most important issue. There are two basic trade-offs in an inventory problem. One is the trade-off between setup costs and inventory holding costs. By placing orders frequently, the size of each order can be made relatively small. Therefore, the holding costs can be reduced. However, the total setup costs will go up. Conversely, less frequent orders will save on setup costs but incur higher holding costs. The other trade-off is between holding costs and stock out costs. Holding more inventory reduces the likelihood of stock outs, and vice versa. These trade-offs give rise to an optimization problem of finding the optimal ordering policy that minimizes the overall cost.

While dealing with inventory systems, there are many factors that should be taken into consideration when solving an inventory problem. Among them, the most important notions are listed below (for more details see Dirk Beyer *et al.* [18]).

Cost function:

One of the most important prerequisites for solving an inventory problem is an appropriate cost function. A typical cost function incorporates the following four types of costs.

- *Variable procurement cost.* This is the cost of buying items. The total purchase cost is usually expressed as cost per unit multiplied by the quantity procured. Sometimes a quantity discount applies if a large number of units are purchased at a time.

- *Fixed ordering cost.* The fixed ordering cost is associated with ordering a batch of items. The ordering cost does not depend on the number of items in the batch. It includes cost of setting up the machine, costs of issuing the purchase order, transportation cost, receiving cost, etc.

- *Holding cost.* The holding cost is associated with keeping items in inventory for a period of time. This cost is typically charged as a percentage of dollar value per unit time. It usually consists of the cost of capital, the cost of storage, the costs of obsolescence and deterioration, the costs of breakage and spoilage, etc.

- *Stock out cost.* Stock out cost reflects the economic consequences of unsatisfied demands. In cases when unsatisfied demands are backlogged, there are costs for handling back orders as well as costs associated with loss of customer goodwill on account of negative effects of backlogs on future customer demands. If all unsatisfied demands are lost, i.e., there is no backlogging, then the stock out cost will also include the cost of the foregone profit.

Demand:

Over time, demand may be constant or variable. Demand may be known in advance or may be random. Its randomness may depend on some exogenous factors such as the state of the economy, weather condition, etc. Another important factor, often ignored in the inventory literature review, is that demand can also be influenced directly or indirectly by the decision makers choice. For example, a sales promotion decision can have a positive effect on demand.

Lead time:

The lead time is defined as the amount of time required to deliver an order after the order is placed. The lead time can be constant (including zero) or random.

Review time:

There are two types of review methods. One is called continuous review, where the inventory levels are known at all times. The other is called periodic review, where inventory levels are known only at discrete points in time.

Various replenishment policies:

- (s, Q) *policy*: This policy requires two parameters for definition. The first parameter s is called the reorder level. A new order is placed as soon as the inventory falls below this level. The other parameter is the order quantity Q $(= S - s)$. Therefore, in this policy, a fixed quantity Q is ordered as soon as the actual inventory falls to the reorder level.

- (s, S) *policy*: This policy is similar to the (s, Q) policy with a difference of one parameter. Instead of a fixed quantity Q a variable quantity is ordered so that the sum of on-hand inventory and the ordered quantity become equal to the pre-defined maximum inventory level S.

- *Random replenishment quantity*: At the time of replenishment a random number of items is purchased according to a probability distribution. This random quantity belongs to the set $\{1, 2, \ldots, k\}$ such that the on-hand plus number of items purchased does not exceed a pre-specified number S.

In this thesis a few queueing-inventory models are analyzed as continuous time Markov chains. In some cases we use tools such as Matrix geometric method for detailed investigation of the problem. Algorithmically tractable tools like these help us to model and analyze the structures so obtained in a very general setup. The resulting quasi-birth-death processes are solved algorithmically by Matrix geometric method.

Phase type distribution (continuous time):

The exponential distribution is widely used in queueing models because of the exceptional mathematical tractability that flows from the memoryless property of this distribution. However, in applications this assumption is highly restrictive. This lead us to explore ways in which we can model more general distributions while maintaining some of the tractability of the exponential distribution. Thus, M. F. Neuts developed the theory of phase type (PH) distributions and related point processes. A PH distribution is obtained as the distribution of the time until absorption in a finite state space Markov chain with an absorbing state.

Consider a Markov chain $\{X(t) : t \geq 0\}$ with finite state space $\{1, 2, \ldots, m+1\}$ where state $m+1$ is absorbing, and the infinitesimal generator matrix

$$\boldsymbol{\mathcal{W}} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ m \\ m{+}1 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & \ldots & m & m+1 \end{array} \\ \begin{pmatrix} \mathcal{T}_{11} & \mathcal{T}_{12} & \ldots & \mathcal{T}_{1m} & \mathcal{T}_{1m+1} \\ \mathcal{T}_{21} & \mathcal{T}_{22} & \ldots & \mathcal{T}_{2m} & \mathcal{T}_{2m+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{T}_{m1} & \mathcal{T}_{m2} & \ldots & \mathcal{T}_{mm} & \mathcal{T}_{mm+1} \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix} \end{array} = \begin{pmatrix} \mathcal{T}_{m\times m} & \boldsymbol{\mathcal{T}}^0 \\ \mathbf{0} & 0 \end{pmatrix}$$

where the elements of the matrices $\mathcal{T}$ and $\boldsymbol{\mathcal{T}}^0$ satisfy $\mathcal{T}_{ii} < 0$ for $1 \leq i \leq m$, $\mathcal{T}_{ij} \geq 0$ for $i \neq j$; $\mathcal{T}_i^0 \geq 0$ and $\mathcal{T}_i^0 > 0$ for at least one i, $1 \leq i \leq m$ and $\mathcal{T}\boldsymbol{e} + \boldsymbol{\mathcal{T}}^0 = \mathbf{0}$. Note that the states $1, 2, \ldots, m$ are transient whereas state $m+1$ is absorbing.

The initial distribution of $\{X(t) : t \geq 0\}$ is given by $(\boldsymbol{\alpha}, \alpha_{m+1})$ with the property that $\boldsymbol{\alpha}\mathbf{e} + \alpha_{m+1} = 1$. Here the states $1, 2, \ldots, m, m+1$ are called phases.

Let $Z = \inf\{t \geq 0 : X(t) = m+1\}$ be the time until absorption in state $m+1$. Then the distribution of Z is called PH distribution with representation $(\boldsymbol{\alpha}, \mathcal{T})$. The dimension m is called the order of the distribution.

(i) The distribution function of Z is given by

$$F(t) = 1 - \boldsymbol{\alpha}\ exp(\mathcal{T}.t)\mathbf{e} \quad \text{for every } t \geq 0.$$

It has a jump of magnitude α_{m+1} at $t = 0$ and its density function is given by

$$f(t) = \boldsymbol{\alpha}\ exp(\mathcal{T}.t)\ \boldsymbol{\mathcal{T}}^0 \quad \text{for every } t > 0$$

where the function $exp(\mathcal{T}.t) = \sum_{i=0}^{\infty} \frac{t^i}{i!}\mathcal{T}^i$, the matrix exponential function and

(ii) the Laplace-Stieltjes transform of $F(.)$ is given by

$$\phi(s) = \alpha_{m+1} + \boldsymbol{\alpha}(sI - \mathcal{T})^{-1}\ \boldsymbol{\mathcal{T}}^0 \quad \text{for } Re(s) \geq 0.$$

Theorem 1.0.1 (see, *Latouche and Ramaswami* [44]). *Consider a* PH *distribution* $(\boldsymbol{\alpha}, \mathcal{T})$. *Absorption into state* $m + 1$ *occurs with probability 1 from any phase* i *in* $\{1, 2, \ldots, m\}$ *if and only if the matrix* $\mathcal{T}$ *is non singular.*

More over, $(-\mathcal{T}^{-1})_{i,j}$ *is the expected total time spent in phase* j *during the time until absorption, given that the initial phase is* i.

1.1 Quasi-birth-death processes

Consider a Markov Chain with state space $\Omega = \bigcup_{n\geq 0} \{(n, i) : 1 \leq i \leq m\}$. Here the first component n is called level of the chain and the second component i is called a phase of the n^{th} level. This Markov Chain is called a Quasi-birth-death (QBD) process if the one step transitions from a state is restricted to the same level or to the two adjacent levels. In other words,

$$(i-1, j') \rightleftharpoons (i, j) \rightleftharpoons (i+1, j'') \quad \text{for } i \geq 1$$

If the transition rates are level independent, the resulting QBD process is called level independent quasi-birth-death process (LIQBD); else it is called a level dependent quasi-birth-death process (LDQBD).

Arranging the elements of Ω in lexicographic order, the infinitesimal generator of a LIQBD process has the block tridiagonal matrix form in which three diagonal blocks repeat after some initial levels. We write such a matrix with modification depending on boundary states as

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} B_1 & A_0 & & & \\ B_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \end{bmatrix}, \tag{1.1}$$

where the sub matrices A_0, A_1, A_2 are square and have the same dimension; matrix B_1 is also square and need not have the same size as A_1. Also, $B_1\mathsf{e} + A_0\mathsf{e} = B_2\mathsf{e} + A_1\mathsf{e} + A_0\mathsf{e} = (A_0 + A_1 + A_2)\mathsf{e} = \mathbf{0}$.

1.2 Matrix geometric method

Marcel F. Neuts pioneered matrix-geometric methods in the study of queueing models in the 1970s. Since then, matrix-geometric methods have become an indispensable tool in stochastic modeling and have found applications in the analysis and design of manufacturing systems, telecommunications networks, risk/insurance models, reliability models, and inventory and supply chain systems. The power and popularity of matrix-geometric methods come from their flexibility in stochastic modeling, capacity for geometric exploration, natural algorithmic thinking, and tractability in numerical computation.

Theorem 1.2.1 (see Theorem 3.1.1. of *Neuts* [47])**.** *The process $\boldsymbol{\mathcal{H}}$ in (1.1) is positive recurrent if and only if the minimal non-negative solution*

R to the matrix-quadratic equation

$$R^2A_2 + RA_1 + A_0 = O \tag{1.2}$$

has all its eigenvalues inside the unit disk and the finite system of equations

$$\begin{aligned} \boldsymbol{x}_0\,(B_1 + RB_2) &= \boldsymbol{0} \\ \boldsymbol{x}_0(I-R)^{-1}\ \boldsymbol{e} &= 1 \end{aligned} \tag{1.3}$$

has a unique positive solution $\boldsymbol{x}_0$.

If the matrix $A = A_0 + A_1 + A_2$ is irreducible, then $sp(R) < 1$ if and only if

$$\boldsymbol{\pi} A_2\ \boldsymbol{e} > \boldsymbol{\pi} A_0\ \boldsymbol{e} \tag{1.4}$$

where $\boldsymbol{\pi}$ is the stationary probability vector of A.

The stationary probability vector $\boldsymbol{x} = (\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots)$ of $\boldsymbol{\mathcal{H}}$ is given by

$$\boldsymbol{x}_i = \boldsymbol{x}_0 R^i \quad for \ \ i \geq 1. \tag{1.5}$$

Once R, the rate matrix is obtained, the vector $\boldsymbol{x}$ can be computed. We can use an iterative procedure or logarithmic reduction algorithm (see *Latouche and Ramaswami* [45]) or the cyclic reduction algorithm (see *Bini and Meini* [5]) for computing R.

1.3 Computation of R matrix

In some cases R can be easily computed. This is especially so when the matrix A_0 has nice structure. When this feature is absent we have to be satisfied with algorithmic approach. There are several algorithms for computing rate matrix R. Here we list two of them.

1.3.1 Iterative algorithm

From (1.2), we can evaluate R in a recursive procedure as follows.

Step 0: $R(0) = \text{O}$.

Step 1:

$$R(n+1) = A_0(-A_1)^{-1} + R^2(n)A_2(-A_1)^{-1}, \quad n = 0, 1, \ldots$$

Continue **Step 1** until $R(n+1)$ is close to $R(n)$.

That is, $||R(n+1) - R(n)||_\infty < \epsilon$.

1.3.2 Logarithmic reduction algorithm

Logarithmic reduction algorithm is developed by *Latouche and Ramaswami* [45] which has extremely fast quadratic convergence. This algorithm is considered to be the most efficient one. We will list only the main steps involved in the logarithmic reduction algorithm. For full details on the logarithmic reduction algorithm refer *Latouche and Ramaswami* [45].

Step 0: $H \leftarrow (-A_1)^{-1}A_0$, $L \leftarrow (-A_1)^{-1}A_2$, $G = L$, and $T = H$.

Step 1:

$$U = HL + LH$$

$$M = H^2$$

$$H \leftarrow (I - U)^{-1}M$$

$$M \leftarrow L^2$$

$$L \leftarrow (I - U)^{-1} M$$

$$G \leftarrow G + TL$$

$$T \leftarrow TH$$

Continue **Step 1:** until $||\mathbf{e} - G\mathbf{e}||_\infty < \epsilon$.

Step 2: $R = -A_0(A_1 + A_0 G)^{-1}$.

1.4 Inventory with positive service time-a review

Research on queueing systems with attached inventory has captured much attention of researchers over the last two decades. Inventory models are studied in detail in Churchman, Acoff and Arnoff [14], Hadley and Whitin [20], Naddor [46], and in Sahin [59]. In the first three, a large number of deterministic model are discussed whereas in the book by Sahin, stochastic models are highlighted. We call these models and problems as *Classical type*, since in all these the amount of time required to serve the item is negligible.

In contrast most of the real life situations need positive amount of time to serve the inventory. Such cases are referred to as *inventory with positive service time*. It may appear that there is no difference between a *queue* and an *inventory with positive service time*. However this is not the case. In a queue we do not speak about the resources for service – if the customers are available and server is ready to serve then the service starts. Nevertheless, this is not the case in inventory with positive service time. Server may be ready to serve and there may be customers waiting to get service. However, inventory may not be available on stock. Thus a queue of customers builds up. Even in the case when lead time is zero, the above problem can very well arise. Needless to say that in the case of positive lead time the server may remain idle even when customers are waiting for want of items in inventory.

The notion of *inventory with positive service time* was introduced by Sigman and Simchi-Levi [65] with Poisson arrival of demands, arbitrarily distributed service time and exponentially distributed replenishment lead time. Among other results they proved that the resulting queueing-inventory system is stable if and only if the service rate is higher than the customer arrival rate. This was followed by large number of research works reported. A brief survey of inventory with positive service time is given in Krishnamoorthy *et al.* [39].

In what follows, we have classified the papers according to two criteria. In the first we include problems involving product form solutions and the second classification is based on queueing-inventory models that use algorithmic approach in the absence of product form solution.

1.4.1 Product form solutions in queueing-inventory models

Control policies like N, D, T and their combinations are extensively studied in queuing systems. Krishnamoorthy *et al.* [29] consider an (s, S) inventory system, where customers require a random amount (positive) of service time. With all underlying distributions independent exponentials they analyze the classical N-policy for inventory with positive service time. Lead time for replenishment of orders is assumed to be zero. Using matrix geometric method and a bit of heuristics the authors obtain the joint distribution of the system state in product form.

The paper by Schwarz *et al.* [62] requires special mention since under exponentially distributed service time and lead time and Poisson input of customers, the authors come up with product form solution for the system state distribution under the assumption that customers do not join when the inventory level is zero. This is despite the strong correlation between the number of customers joining the system during the lead time and the number of items in the inventory over that period. Their work is subsumed in Krishnamoorthy and Viswanath [42] wherein the authors have reduced the Schwarz *et al.* [62] model to a production inventory system with a single-batch bulk production of the quantum of inventory required.

Schwarz and Daduna [63] investigate an $M/M/1$ queueing system with unlimited capacity for customers where service is in the form of providing inventoried items. Customers can join even when the inventory level is zero. They derive the main performance measures from queueing and inventory perspective and study their interconnection. Wherever a performance measure does not have a closed form, the authors develop approximations. Schwarz *et al.* [64] consider queueing networks with attached inventory. At each service station an order for replenishment is placed when the inventory level at that station drops to its reorder level. They consider rerouting of customers served out from a particular station, when the immediately following station has zero inventory. Thus no customer is lost to the system. The authors derive joint stationary distribution of queue length and inventory level in explicit product form.

Saffari *et al.* [57] consider an $M/M/1$ queue with inventoried items for service. The control policy followed is (s, Q) and lead time is mixed exponential distribution. When inventory is out of stock, fresh arrivals are lost to the system. This leads to a product form solution for the system state probability.

In a very recent paper Saffari *et al.* [58] analyze an inventory model with positive service time and arbitrarily distributed lead time. They assume that no customer joins the system when the inventory level is zero. A product form solution for system state is obtained here as well. Another recent contribution of interest to inventory with positive service time involving a random environment is by Ruslan and Daduna [55] where again they establish a stochastic decomposition of the system. They prove a necessary and sufficient condition for a product form steady state distribution of the joint queueing-environment process to exist. A still more recent paper Ruslan and Daduna [56] investigate inventory with positive service time in a random environment embedded in a Markov chain. They provide a counter example to show that the steady state distribution of an $M/G/1/\infty$ system with (s, S) policy and lost sales, need not have a product form. Nevertheless, in general loss systems in a random environment have a product form steady state distribution.

Can we always get a product form solution when the lead time is zero and the probability distributions involved are all exponential? The answer is, surprisingly "NO". Krishnamoorthy *et al.* [30] considered an (s, S) inventory system with service time in which it is assumed that when the server is idle he continues to process the items. In case a processed item is available at a customer arrival epoch, then it is instantaneously served resulting in negligible service time. However, in the absence of processed item at the epoch of arrival of a customer, he has to wait until the item is processed. Of course he has to wait until all ahead of him, if any, are served. Unlike in Krishnamoorthy *et al.* [29], here the authors are not able to produce closed form solution. Instead they obtain a matrix geometric solution. Unlike its predecessor, in the present case optimal s is not zero. Whereas Krishnamoorthy *et al.* [30] failed to get closed form solution for the model where the purpose was to increase server idle time utilization and decrease waiting time of customers, Deepak *et al.* [16] (see also Krishnamoorthy *et al.* [34]) consider another variation of Krishnamoorthy *et al.* [29] where a customer demands a processed item or an unprocessed one with probability p and $1-p$, respectively, at the time when the customers enter for service. If unprocessed item is demanded, then service time is negligible whereas if processed item is needed then there is a positive service time involved which they assume to be exponential. Customers arrive according to a Poisson process. Lead time is assumed to be zero as in the last two problems discussed. Surprisingly here the authors succeeded in producing closed form solution for the system state probability, which further turned out to be in product form. Since the main objective of this thesis is to obtain product form solution for inventory with positive service time, we mention below those contributions that provide mainly algorithmic solution, without going into the details of the content of these papers. These are not referred in our main work. Hence we do not go into the details of such papers. Instead these are classified on the basis of the category they belong to, such as vacation, retrial, production, multi-server and so on. Nevertheless, chapters 5 and 6 of this thesis provide algorithmic approach to the system under study; also part of chapter 4 on multi-server queueing-inventory models adopts algorithmic approach.

1.4.2 Queueing-inventory systems involving algorithmic approach

Single server, Markovian queueing-inventory models

The contributions are:
Arivarignan *et al.* [2], Berman [6], Berman and Kim [7], Berman and Sapna [8], Berman and Sapna ([9], [10]), Berman and Kim [11], Deepak *et al.* ([16], [17]), Jayaraman *et al.* [21], Cui and Wang [15], Kalpakam and Shanthi [23], Krishnamoorthy and Islam ([25], [26]), Krishnamoorthy *et al.* [27], Krishnamoorthy and Jose [28], Krishnamoorthy *et al.* ([29], [30], [31]), Krishnamoorthy and Jose ([32], [33]), Krishnamoorthy *et al.* [34], Krishnamoorthy and Jose ([35]), Krishnamoorthy and Anbazhagan [36], Krishnamoorthy *et al.* ([37], [40], [41]), Krishnamoorthy and Viswanath ([38], [42]), Lalitha [43], Ning Zhao and Zhanotong Lian [48], Padmavathi *et al.* [49], Paul Manuel *et al.* ([50], [51]), Perumal and Arivarignan [53], Ruslan and Daduna [55], Saffari, *et al.* ([57], [58]), Sajeev S. Nair [60], Schwarz *et al.* ([62], [64]) Schwarz and Daduna [63], Sivakumar and Arivarignan ([69], [70], [72]), Sivakumar [71], Sivakumar ([66], [72], [68]), Viswanath *et al.* [74], Vineetha [75] and Yadavalli ([76], [77]).

Single server, non-Markovian queueing-inventory models

There are very few contributions beginning to this category.
Ruslan and Daduna [56], Sigman and Simchi-Levi [65], Saffari, *et al.* [58].
Fourth chapter of this thesis examines a two server and then $c(\geq 3)$ server queueing-inventory system respectively.

Multi-server queueing-inventory models

Literature on this also is pretty scarce:
Anoop N. Nair *et al.*[1], Yadavalli *et al.* ([78], [79], [80]).

Queueing-inventory model with retrial of unsatisfied customers

Though literature on retrial queues is vast, that on queueing-inventory finds very few contributions. Chapter 6 of this discuses an inventory problem with retrial of customers. Here is the list of the limitted contribution: Cui and Wang [15], Padmavathi *et al.* [49], Sivakumar ([66], [68]), Krishnamoorthy and Jose ([32], [33], [35]), Krishnamoorthy *et al.* [40], Sivakumar and Arrivagnan *et al.* [72].

Production inventory models

Production inventory could be viewed as a supply chain with two echelons. Here as well not much contributions could be found: Krishnamoorthy and Islam [25], Krishnamoorthy *et al.* ([27], [41]), Krishnamoorthy and Jose [35], Krishnamoorthy and Viswanath ([38].

Queueing-inventory models with server vacation

Jayaraman *et al.* [21], Krishnamoorthy and Viswanath [38], Sivakumar [68], Padmavathi *et al.* [49] and Viswanath *et al.* [74].

SUMMARY OF THE THESIS

In this thesis a few queueing-inventory models are analyzed as continuous time Markov chains. We obtain steady-state distributions of a few queueing-inventory models in product form under the assumption that no customer joins the system when the inventory level is zero. This is despite the strong correlation between the number of customers joining the system and the inventory level during lead time. The resulting quasi-birth-and-death (QBD) processes are solved explicitly by Matrix Geometric Methods. The inventory literature so far available assume that a customer, at the end of his service, is provided one unit of item from the inventory. However, in practice this need not hold. For example, assuming vacant job positions as inventory and job aspirants as customers, we notice that a candidate (customer) need not be offered the job at the end of the interview. It is as well the case that, a candidate rejects the offer of the position after interview. This is the motivation behind the work reported in this thesis. Further an item produced in a production process need not be of the required quality. Such items are rejected.

Now we turn to the content of the thesis. This thesis entitled "*Investigations on Stochastic Storage Systems with Positive Service Time*" is divided into 6 chapters including the introductory chapter.

In chapter 1 a detailed review of inventory models involving positive service time is given. These include classical and retrial cases. Also contributions to production inventory with service time are indicated towards the end.

Chapter 2 discusses a single server queueing-inventory system, with the item given with probability γ to a customer at his service completion epoch. Two control policies, (s, Q) and (s, S) are discussed. In both cases we obtain the joint distribution of the number of customers and the number of items in the inventory as the product of their marginals under the assumption that customers do not join when inventory level is zero. Optimization problems associated with both models are investigated and the optimal pairs (s, S) and (s, Q) and the corresponding expected minimum costs are obtained. Further we investigate numerically an expression for per unit time cost as a function of γ. This function exhibits convexity property. A comparison with Schwarz *et al.* [62] is provided. The case of arbitrarily distributed service time is briefly indicated. First emptiness time distribution is computed for the $M/M/1/1$ queueing-inventory system.

In Chapter 3 we discuss a production inventory system with the item produced being admitted (added to the inventory) with probability δ at the end of a production epoch as well as an item from the inventory is supplied to the customer with probability γ at the end of a service. The control policy followed is of the (s, S) type. We obtain joint distribution of the number of customers and the number of items in the inventory as the product of their marginals under the assumption that customers do not join when inventory level is zero. Performance measures that impact the system, are obtained. In particular optimal pairs (s, S) are obtained through numerical procedures for values of (γ, δ) on the set $\{0.1, 0.2, \ldots, 1\} \times \{0.1, 0.2, \ldots, 1\}$. Here also we compute the first emptiness time distribution for the $M/M/1/1$ queueing-inventory system with production.

In Chapter 4 we attempt to derive the steady-state distribution of the $M/M/c$ queueing-inventory system with positive service time. First we analyze the case of $c = 2$ servers which are assumed to be homogeneous and that the service time follows exponential distribution. The inventory replenishment follows the (s, Q) policy. We obtain a product form solution of the steady-state distribution under the assumption that customers do not join the system when the inventory level is zero. An optimization problem is also investigated to get the optimal pair (s, Q) and the corresponding expected minimum cost. As in the case of $M/M/c$ retrial queue with $c \geq 3$, we conjuncture that $M/M/c$, for $c \geq 3$ but c less than s, queueing-inventory problems do not have analytical solution. So we proceed to analyze by using algorithmic approach. All servers are assumed to be homogeneous and that the service time follows exponential distribution. Here also the inventory replenishment follows (s, Q) policy. We derive an explicit expression for the stability condition of the system. We discuss the conditional distribution of the inventory level, conditioned on the number of customers in the system and conditional distribution of the number of customers conditioned on the inventory level. Also we compute the distribution of two consecutive s to s transitions of the inventory level (that is the first return time to s). Since closed form solutions is not possible. We employ algorithmic method to compute the stationary distribution. We also obtain several system performance measures.

Chapter 5 is on queueing-inventory system under (s, Q) policy with working vacations and vacation interruptions. The notion of working vacation is introduced by Jihong Li and Naishuo Tian [22] in classical queueing theory. During working vacation also the server provides service, but at a lower rate. Further, the server can come back from the vacation mode to the normal working mode once some indices of the system, such as the number of customers achieve a certain value and there is at least one item in the inventory. More precisely, the server may come back from the vacation without completing the vacation period. This is called vacation interruption (see [22]). We assume that if there are customers in the system after a service completion during a working vacation, the server will comeback to the normal working mode provided the vacation completion is realized during the service; else the server stays in the working vacation mode. With the system having infinite capacity, we derive condition for stability of the system. Despite the corresponding queueing system (without inventory) having analytic solution, we are not able to arrive at closed form expression for system state for the queueing-inventory problem under discussion. Hence algorithmic approach is adopted. Several performance measures are evaluated. An optimization problem is also discussed.

In the 6^{th} chapter, we consider an $M/M/1/1$ queueing-inventory system. Here arrivals taking place when server busy, proceed to an orbit of infinite capacity. From the orbit the head of the queue alone retires to access the server. Failed attempts to access an idle server with positive inventory results in the retrial customer returning to orbit. The inter-retrial times are independent identically distributed exponential random variables with parameter θ, irrespective the number of customers in the orbit. We compute the condition for stability and then employ algorithmic approach for the computation of the system state probability. We also compute the expected waiting time of a customer in the orbit, distribution of the time until the first customer goes to orbit and the probability of no customer going to orbit in a given interval of time. An optimization problem is also numerically investigated. In the last section of the chapter we briefly analyze a tandem queueing-network with just two stations. The second station has characteristics indicated above in this paragraph. Station 1 is $M/M/1/\infty$ queueing-inventory system whose output proceeds to station 2, provided there is at least one item in the inventory. This combined system is analyzed. It is argued that the combined system can be decomposed into two sub systems.

Chapter 2

A revisit to queueing-inventory system with positive service time

2.1 Introduction

A close look at the literature on inventory with positive service time indicates that one unit of the inventory is provided to the customer at his departure epoch. However, this need not hold in several real life situations. For example consider a candidate who appears for an interview against a position. At the end of the interview he/she may not be offered the position. In some cases the candidate may decline the offer of the job. In this

---

Some results of this chapter are included in the following paper.
A. Krishnamoorthy, R. Manikandan and B. Lakshmy : A revisit to queueing-inventory system with positive service time. Annals of Operations Research, DOI 10.1007/s10479-013-1437-x.

case the job is taken as an inventory and the candidate as customer. In this chapter we analyze such type of situations under Poisson demand process, exponentially distributed service and lead time. We further impose the condition that no customer joins the system when the on-hand inventory is zero (those who are already present stay back in the system until served).

Two models based on the ordering policy are specifically considered: (i) The replenishment order which is placed when the inventory level goes down to s (which is called reorder level), for a fixed number Q of the item. This is referred to as (s, Q) policy. (ii) The replenishment order is to bring back the inventory to the maximum level S as and when replenishment takes place this is referred to as (s, S) policy where s is again the reorder level. Both these positions are the same when lead time is zero.

Mathematical formulation of the (s, Q) policy is given in Section 2.2.1. Stability condition of the queueing-inventory system under the (s, Q) policy is provided in Section 2.3. Further, the system state distribution is derived in that section. Several performance measures are also indicated there. In Section 2.4 mathematical description of the queueing-inventory under (s, S) policy is provided. Here again the stability condition is derived and performance measures are computed. We also establish the stochastic decomposition property of the system as done for the system under the (s, Q) policy in the previous section. In the next section three optimization problems are investigated: for given γ, (a) the optimal pair (s, Q) and the corresponding minimum cost, (b) the optimal pair (s, S) and the corresponding minimum cost and (c) the expected unit time cost of the system as a function of γ. In all the three cases we obtain through numerical experiments the *global optima*. A brief sketch of arbitrarily distributed service time with Poisson arrival of demands and exponentially distributed lead time is provided in Section 2.6. First emptiness time distribution for the $M/M/1/1$ queueing-inventory system is computed in Section 2.7.

2.2 Description of the model

We consider an $M/M/1$ queueing-inventory system with positive service time. Arrival process is assumed to be Poisson with rate λ. Each customer requires a single homogeneous item, having random duration of service time which follows exponential distribution with parameter μ. However, it is not essential that inventory is provided to the customer at the end of his service. More specifically, the item is served with probability γ at the end of a service and with probability $1-\gamma$ the item is not delivered to the customer. When $\gamma = 1$ our model reduces to Schwarz *et al.* [62]. A very crucial assumption of this model is that customers do not join the system when the inventory level is zero. This leads us to the product form solution for the models under study. We consider the two distinct replenishment policies: (i) (s, Q) and (ii) (s, S) described in the previous section.

2.2.1 Model 1: (s, Q) policy

In this model when the on-hand inventory reaches a pre-specified value $s \geq 0$, a replenishment order is placed for $Q(< \infty)$ units with $Q > s$. We fix $S = s + Q$ as the maximum number of items that could be held in the system at any given time. The lead time follows exponential distribution with parameter β. Then $\{\mathcal{X}(t)|t \geq 0\} = \{(\mathcal{N}(t), \mathcal{I}(t))|t \geq 0\}$ is a CTMC with state space

$$\Omega_1 = \bigcup_{i=0}^{\infty} \mathcal{L}(i)$$

where $\mathcal{L}(i)$ is called the i^{th} level (number of customers in the system is $i(\geq 0)$). In the i^{th} level the number of items in inventory can be anything from 0 to S. Accordingly we write $\mathcal{L}(i) = \{(i,0),\ldots,(i,s+Q)\}$. In these, the second coordinate is referred to as the phase of the system. Now we describe the transitions in the Markov chain $\{\mathcal{X}(t)|t \geq 0\}$:

(a) Transitions due to arrival of customers :

$(i,j) \to (i+1,j)$: the rate is λ, for $i \geq 0; 1 \leq j \leq S$.

(b) Transitions due to service completion consequent to which an inventoried item is served to the outgoing customer:

$(i,j) \to (i-1,j-1)$: the rate is $\gamma\mu$, for $i \geq 1; 1 \leq j \leq S$.

(c) Transitions due to service completion for which inventory is not served:

$(i,j) \to (i-1,j)$: the rate is $(1-\gamma)\mu$, for $i \geq 1; 1 \leq j \leq S$.

(d) Transitions due to replenishment's:

$(i,j) \to (i,Q+j)$: the rate is β, for $i \geq 0; 0 \leq j \leq s$.

All other transition pairs have rate zero. The infinitesimal generator $\boldsymbol{\mathcal{W}}$ of the CTMC $\{\mathcal{X}(t)|t \geq 0\}$ is

$$\mathcal{W}=\begin{bmatrix} B & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \dots \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

where B contains transition rates within $\mathcal{L}(0)$; A_0 represents the transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i+1), i \geq 0$; A_1 represents the transitions within $\mathcal{L}(i)$ for $i \geq 1$, and A_2 represents transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i-1), i \geq 1$. All these matrices are square matrices of dimension $S+1$.

2.3 Analysis of the system

In this section we perform the steady-state analysis of the (s, Q) queueing-inventory model under study by first establishing the stability condition of the system. Define A=$A_0 + A_1 + A_2$. This is the infinitesimal generator of the finite state space CTMC corresponding to the inventory level $\{0, 1, \dots, S\}$. Let $\boldsymbol{\varphi}$ denote the steady-state probability vector of A. That is,

$$\boldsymbol{\varphi} A = 0, \boldsymbol{\varphi}\boldsymbol{e} = 1. \tag{2.1}$$

Write

$$\boldsymbol{\varphi} = (\varphi_0, \varphi_1, \dots, \varphi_S)$$

where φ_k is the probability that inventory level is k, $0 \leq k \leq S$. Then using relations in (2.1) we get the components of the vector $\boldsymbol{\varphi}$ explicitly as

$$\varphi_k = \begin{cases} \left[1 + Q\frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^s\right]^{-1}, \ k = 0. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{k-1}\varphi_0, \ k = 1, 2, \cdots, s. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s}\varphi_0, \ k = s+1, s+2, \cdots, Q. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{k-Q-1}\left(\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s-(k-Q-1)} - 1\right)\varphi_0, \ k = Q+1, Q+2, \cdots, S. \end{cases}$$

Since the Markov chain is an LIQBD, it is stable if and only if the left drift rate exceeds the right drift rate. That is,

$$\boldsymbol{\varphi} A_0 \mathbf{e} < \boldsymbol{\varphi} A_2 \mathbf{e}. \tag{2.2}$$

We have the following lemma:

Lemma 2.3.1. The stability condition of the (s, Q) queueing-inventory model is given by $\lambda < \mu$.

Proof. From the well known result in Neuts [47] on the positive recurrence of A, we have $\boldsymbol{\varphi} A_0 \mathbf{e} < \boldsymbol{\varphi} A_2 \mathbf{e}$. With a bit of computation, this simplifies to the result $\lambda < \mu$. □

For future reference we define ρ as

$$\rho = \frac{\lambda}{\mu}. \tag{2.3}$$

2.3.1 Steady-state analysis

For computing the steady-state probability vector of the process $\{\mathcal{X}(t)|t \geq 0\}$, we first consider an inventory system with negligible service time and no backlog of demands. The rest of the assumptions such as those on the arrival process and lead time are the same as given earlier. Designate the Markov chain so obtained as $\{\widetilde{X}(t)\} = \{\mathcal{I}(t)|t \geq 0\}$. Here $\widetilde{X}(t) = \mathcal{I}(t)$ is the inventory level at time t. Its infinitesimal generator $\widetilde{\boldsymbol{\mathcal{W}}}$ is given by,

$$\widetilde{\boldsymbol{\mathcal{W}}} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ s \\ \\ \vdots \\ Q \\ \vdots \\ S \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & \dots & s\dots & Q & \dots & & S \end{array} \\ \left(\begin{array}{cccccccc} -\beta & & & & \beta & & & \\ \gamma\lambda & -(\gamma\lambda+\beta) & & & & & & \\ & \ddots & \ddots & & & & \ddots & \\ & & & -(\gamma\lambda+\beta) & & & & \beta \\ & & & \gamma\lambda & -\gamma\lambda & & & \\ & & & & & & & \\ & & & & \ddots & \ddots & & \\ & & & & & \gamma\lambda & -\gamma\lambda & \\ & & & & & & \gamma\lambda & -\gamma\lambda \end{array} \right) \end{array}$$

Let $\boldsymbol{\pi}$=$(\pi_0, \pi_1, \ldots, \pi_S)$ be the steady-state probability vector of the process $\{\widetilde{X}(t)\}$=$\{\mathcal{I}(t)|t \geq 0\}$. Then $\boldsymbol{\pi}$ satisfies the relations

$$\boldsymbol{\pi}\widetilde{\boldsymbol{\mathcal{W}}} = 0, \boldsymbol{\pi}\mathbf{e} = 1 \tag{2.4}$$

That is, at arbitrary epochs the inventory level distribution π_j is given by

$$\pi_j = \begin{cases} \left[1 + Q\frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^s\right]^{-1}, \; j = 0. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{j-1}\pi_0, \; j = 1, 2, \cdots, s. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s}\pi_0, \; j = s+1, s+2, \cdots, Q. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{j-Q-1}\left(\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s-(j-Q-1)} - 1\right)\pi_0, \; j = Q+1, Q+2, \cdots, S. \end{cases} \tag{2.5}$$

Using the components of the probability vector $\boldsymbol{\pi}$, we shall find the steady-state probability vector of the original system. For this, let $\boldsymbol{x}$ be the steady-state probability vector of the original system. Then the steady-state vector must satisfy the set of equations

$$\boldsymbol{x}\boldsymbol{\mathcal{W}} = 0, \boldsymbol{x}\mathbf{e} = 1. \tag{2.6}$$

Let us partition $\boldsymbol{x}$ by levels as

$$\boldsymbol{x} = (\boldsymbol{x_0}, \boldsymbol{x_1}, \boldsymbol{x_2}, \dots) \tag{2.7}$$

where the subvectors of $\boldsymbol{x}$; are further partitioned as,

$$\boldsymbol{x_i} = (x_i(0), x_i(1), x_i(2), x_i(3), ..., x_i(S)), i \geq 0. \tag{2.8}$$

Then the above system of equations reduces to

$$\boldsymbol{x_0}B + \boldsymbol{x_1}A_2 = 0 \tag{2.9}$$

$$\boldsymbol{x_i}A_0 + \boldsymbol{x_{i+1}}A_1 + \boldsymbol{x_{i+2}}A_2 = 0, i \geq 0 \tag{2.10}$$

Assume that

$$\boldsymbol{x_0} = \xi\boldsymbol{\pi} \tag{2.11}$$

and

$$\boldsymbol{x_i} = \xi \left(\frac{\lambda}{\mu}\right)^i \boldsymbol{\pi}, i \geq 1 \tag{2.12}$$

where ξ is a constant to be determined. We verify that the equations (2.9) and (2.10) are satisfied by (2.11) and (2.12). For (2.9), we have

$$\boldsymbol{x_0}B + \boldsymbol{x_1}A_2 = \xi\boldsymbol{\pi}\left(B + \frac{\lambda}{\mu}A_2\right) \tag{2.13}$$

and from relation (2.10), we have,

$$\boldsymbol{x_i}A_0 + \boldsymbol{x_{i+1}}A_1 + \boldsymbol{x_{i+2}}A_2 = \xi\left(\frac{\lambda}{\mu}\right)^{i+1}\boldsymbol{\pi}\left(B + \frac{\lambda}{\mu}A_2\right) \tag{2.14}$$

Now from the matrices B, A_2 and $\widetilde{\boldsymbol{\mathcal{W}}}$, it follows that

$$B + \frac{\lambda}{\mu}A_2 = \widetilde{\boldsymbol{\mathcal{W}}} \tag{2.15}$$

Also from (2.4) we have $\boldsymbol{\pi}\widetilde{\boldsymbol{\mathcal{W}}} = 0$. Hence the right hand side of the equation (2.13) and (2.14) are zero. Hence if we take the vector $\boldsymbol{x}$ as given by (2.6), it follows that (2.9) and (2.10) are satisfied. Now applying the normalizing condition $\boldsymbol{xe}$=1, we get

$$\xi\left[1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 + \cdots\right] = 1$$

Hence under the condition that $\lambda < \mu$, we have

$$\xi = 1 - \frac{\lambda}{\mu}. \tag{2.16}$$

Thus we arrive at our main theorem:

Theorem 2.3.1. *Under the necessary and sufficient condition* $\lambda < \mu$ *for stability, the components of the steady-state probability vector of the process* $\{\mathcal{X}(t) | t \geq 0\}$ *with generator matrix* $\boldsymbol{\mathcal{W}}$ *is given by* (2.11), (2.12) *and* (2.16). *That is,* $\boldsymbol{x_0} = (1-\rho)\boldsymbol{\pi},\ \boldsymbol{x_i} = (1-\rho)\rho^i\boldsymbol{\pi},$ *for* $i \geq 1$ *where* ρ *is as defined in* (2.3) *and the finite probability vector* $\boldsymbol{\pi}$ *is as given in* (2.5).

The consequence of Theorem 2.3.1 is that the two dimensional system can be decomposed into two distinct one dimensional objects (namely number of customers and number of inventory items in the system).

Remark 2.3.1. : From Theorem 2.3.1 we see that the system state distribution, under the stability condition, is the product of marginal distributions of the number of customers in an $M/M/1$ system and the number of items in the inventory.

2.3.2 Performance measures

- Mean number of customers in the system, $L_s = \frac{\lambda}{\mu-\lambda}$.

- Mean number of customers in the queue, $L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}$.

- Mean inventory level in the system, $I_m = \sum_{j=1}^{Q+s} j\pi_j$.

- Depletion rate of inventory, $D_{inv} = \gamma\lambda(1-\pi_0)$.

  Note that the quantity on the right hand side above is smaller than the corresponding quantity given in Schwarz *et al.* [62].

- Mean number of replenishments per unit time , $R_r = \beta\left(\sum_{j=0}^{s}\pi_j\right)$.

- Mean number of departures per unit time, $D_m = \frac{\mu^2}{\mu-\lambda}(1-\pi_0)$.

- Expected loss rate of customers, $E_{loss} = \lambda\pi_0$.

- Define the length of cycle as the time duration between two consecutive epochs at which order for replenishments are placed. So we get, Expected loss rate of customers when the inventory level is zero per cycle, $E^c_{loss} = \frac{E_{loss}}{R_r}$.

- Mean number of customers arriving per unit time, $\lambda_A = \lambda(1 - \pi_0)$.
- Mean sojourn time of the customers in the system, $W_s = \frac{L_s}{\lambda_A}$.
- Mean waiting time of a customer in the queue, $W_q = \frac{L_q}{\lambda_A}$.
- Mean number of customers waiting in the system when inventory is available, $\widetilde{W} = L_s(1 - \pi_0)$.
- Mean number of customers waiting in the system during the stock out period, $\widetilde{\widetilde{W}} = L_s\pi_0$.

2.4 Model 2: (s, S) policy

We consider a queueing-inventory system with positive service time as described at the beginning of Section 2.2. However, the inventory replenishment policy is of (s, S) type. This policy differs from the (s, Q) policy in that instead of a fixed quantity Q, a variable quantity at the time of replenishment, is purchased so that the sum of on-hand inventory and the purchased quantity equal to a predefined maximum inventory level S. This policy is also referred to as order upto S. We keep the same arrival and service processes as in Section 2.2. The lead time is also exponentially distributed with parameter β. Then the CTMC $\{Y(t)|t \geq 0\} = \{(\mathcal{N}(t), \mathcal{I}(t))|t \geq 0\}$ with state space,

$$\Omega_2 = \bigcup_{i=0}^{\infty} \mathcal{L}(i)$$

where $\mathcal{L}(i)$ is the collection of states defined as $\mathcal{L}(i) = \{(i, 0), \ldots, (i, S)\}$ as defined in Model 1. The transitions corresponding to the Markov chain $\{Y(t)\}$ are same as in Section 2.2, but the transitions corresponding to the inventory replenishment is different in that the rate of transition from (i, j) to (i, S) is β, for $i \geq 0$ and for j such that $0 \leq j \leq s$ and zero for other combinations. The infinitesimal generator $\boldsymbol{\mathcal{H}}$ of the CTMC $\{Y(t)|t \geq 0\}$ is

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \bar{B} & \bar{A}_0 & & & & \\ \bar{A}_2 & \bar{A}_1 & \bar{A}_0 & & & \\ & \bar{A}_2 & \bar{A}_1 & \bar{A}_0 & \dots & \\ & & \ddots & \ddots & \ddots & \end{bmatrix}$$

where $\bar{B}$ contains rates of transitions within $\mathcal{L}(0)$; $\bar{A}_0$ represents the transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i+1), i \geq 0$; $\bar{A}_1$ represents the rate of transitions within $\mathcal{L}(i)$ $i \geq 1$ and $\bar{A}_2$ represents the transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i-1), i \geq 1$. All entries in $\boldsymbol{\mathcal{H}}$ are square matrices of dimension $S + 1$.

2.4.1 System stability and computation of steady-state probability vector:

The Markov chain under consideration is a LIQBD process. For this chain to be stable it is necessary and sufficient that

$$\boldsymbol{\psi}\bar{A}_0\mathbf{e} < \boldsymbol{\psi}\bar{A}_2\mathbf{e} \tag{2.17}$$

where $\boldsymbol{\psi}$ is the unique non negative vector satisfying,

$$\boldsymbol{\psi}\bar{A} = 0, \boldsymbol{\psi}\mathbf{e} = 1 \tag{2.18}$$

and $\bar{A} = \bar{A}_0 + \bar{A}_1 + \bar{A}_2$, is the infinitesimal generator of the finite state CTMC on the set $\{0, 1, \dots, S\}$. Write $\boldsymbol{\psi}$ as $(\psi_0, \psi_1, \dots, \psi_S)$. Then by relation (2.18), we get the components of the probability vector $\boldsymbol{\psi}$ explicitly as,

$$\psi_k = \begin{cases} \left[\left(1+\frac{\beta}{\gamma\mu}Q\right)\left(\frac{\beta+\gamma\mu}{\gamma\mu}\right)^s\right]^{-1}, k=0. \\ \frac{\beta}{\gamma\mu}\left(\frac{\beta+\gamma\mu}{\gamma\mu}\right)^{k-1}\psi_0, k=1,2,...,s. \\ \frac{\beta}{\gamma\mu}\left(\frac{\beta+\gamma\mu}{\gamma\mu}\right)^{s}\psi_0, k=s+1,s+2,...,S-1,S. \end{cases}$$

From the relation (2.17) we have

Lemma 2.4.1. The stability condition of the queueing-inventory system under study is given by $\lambda < \mu$

Proof. : On the same lines as that of Lemma 2.3.1. □

For computing the steady-state probability vector of the process $\{Y(t)|t \geq 0\}$, we first consider an inventory system with negligible service time and no backlog of demands. Designate this CTMC by $\{\widetilde{Y}(t)|t \geq 0\} = \{\mathcal{I}(t)|t \geq 0\}$. Its infinitesimal generator $\widetilde{\boldsymbol{\mathcal{H}}}$ is a matrix of order $S+1$. Let $\widetilde{\boldsymbol{\pi}}=(\widetilde{\pi}_0, \widetilde{\pi}_1, \ldots, \widetilde{\pi}_S)$ be the steady-state probability vector of the $\widetilde{Y}(t)$ process. Then $\widetilde{\boldsymbol{\mathcal{H}}}$ satisfies the equations

$$\widetilde{\boldsymbol{\pi}}\widetilde{\boldsymbol{\mathcal{H}}} = 0, \widetilde{\boldsymbol{\pi}}\mathbf{e} = 1 \tag{2.19}$$

Its components $\widetilde{\pi}_j$ are computed as:

$$\widetilde{\pi}_j = \begin{cases} \left[\left(1+\frac{\beta}{\gamma\lambda}Q\right)\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^s\right]^{-1}, j=0. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{j-1}\widetilde{\pi}_0, j=1,2,...,s. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s}\widetilde{\pi}_0, j=s+1,s+2,...,S-1,S. \end{cases} \tag{2.20}$$

Now using the vector $\widetilde{\boldsymbol{\pi}}$, we shall find the steady-state probability vector of the original system by using the same technique as in Section 2.3.1. Thus we arrive at:

Theorem 2.4.1. *Under the necessary and sufficient condition $\lambda < \mu$ for stability, the components of the steady-state probability vector of the process $\{Y(t)|t \geq 0\}$ with generator matrix $\boldsymbol{\mathcal{H}}$ is $\boldsymbol{y_0} = (1-\rho)\widetilde{\boldsymbol{\pi}}$ and $\boldsymbol{y_i} = (1-\rho)\rho^i\widetilde{\boldsymbol{\pi}},\ i \geq 1$ where ρ is defined as in (2.3) and the finite probability vector $\widetilde{\boldsymbol{\pi}}$ in the component form is given by (2.20).*

2.4.2 Performance measures:

- Mean number of customers in the system, $L_s = \frac{\lambda}{\mu-\lambda}$.
- Mean number of customers in the queue, $L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}$.
- Mean inventory level in the system, $I_m = \sum\limits_{j=1}^{S} j\widetilde{\pi}_j$.
- Depletion rate of inventory, $D_{inv} = \gamma\lambda(1-\widetilde{\pi}_0)$.

  Note that the quantity on the right hand side is smaller than the corresponding quantity given in Schwarz *et al.* [62]
- Mean number of replenishment's per unit time, $R_r = \beta(s+1)\widetilde{\pi}_S$.
- Mean number of departures per unit time, $D_m = \frac{\mu^2}{\mu-\lambda}\left(1-\widetilde{\pi}_0\right)$.
- Expected loss rate of customers, $E_{loss} = \lambda\widetilde{\pi}_0$.
- Define the length of cycle as the time duration between two consecutive epochs at which order for replenishment are placed. So we get expected loss rate of customers when the inventory level is zero per cycle as, $E^c_{loss} = \frac{E_{loss}}{R_r}$.
- Mean number of customers arriving per unit time, $\lambda_A = \lambda(1-\widetilde{\pi}_0)$.

- Mean sojourn time of the customers in the system, $W_s = \frac{L_s}{\lambda_A}$.
- Mean waiting time of the customers in the queue, $W_q = \frac{L_q}{\lambda_A}$.
- Mean number of customers waiting in the system when inventory is available, $\widetilde{W} = L_s(1 - \widetilde{\pi}_0)$.
- Mean number of customers waiting in the system during the stock out period, $\widetilde{\widetilde{W}} = L_s\widetilde{\pi}_0$.

2.5 Optimization problem

We look for the optimal pair of control variables in the two models discussed above. Now for computing the minimal costs of (s, Q) and (s, S) models we introduce two cost functions: $\mathcal{F}_1(s, Q)$ and $\mathcal{F}_2(s, S)$ defined by

$$\mathcal{F}_1(s, Q) = h_1.I_m + c_1.E_{loss} + c_2.\widetilde{\widetilde{\mathcal{W}}} + (K + Q.c_3)\,.R_r$$

and

$$\mathcal{F}_2(s, S) = h_1.I_m + c_1.E_{loss} + c_2.\widetilde{\widetilde{\mathcal{W}}} + K.R_r + \left(\frac{\beta}{\lambda + \beta + \mu}\sum_{i=1}^{s}\pi_i.\,(S - i) + \frac{\beta S}{\lambda + \beta}.\pi_0\right).c_3$$

where K is fixed cost for placing an order, c_1 is cost incurred due to loss per customer, c_2 is waiting cost per unit time per customer during the stock out period, c_3 is variable procurement cost per item, and h is unit holding cost of inventory for one unit of time. We assign the following values to the parameters: $\lambda = 2, \mu = 3, \beta = 1, K = \$500, c_1 = \$25, c_2 = \$50, c_3 = \$25, h_1 = \$2$. We obtain the following two Tables (2.1 & 2.2) which provide the optimal pairs (s, Q) and (s, S) and also the corresponding minimum cost (in Dollars). Here γ is varied from 0.1 to 1, each time increasing it by

0.1 unit. The optimal pair (s, Q) and the corresponding cost (minimum) are given in Table 2.1. Table 2.2 contains optimal pairs (s, S) and the corresponding costs (minimum) when γ is varied from 0.1 to 1.

Table 2.1: Optimal (s, Q) pair and minimum cost

γ	0.1	0.2	0.3	0.4	0.5
Optimal (s, Q) pair & minimum cost	(1,30) 109.10	(1,29) 104.02	(1,29) 100.28	(1,28) 97.45	(1,28) 95.26
γ	0.6	0.7	0.8	0.9	1
Optimal (s, Q) pair & minimum cost	(1,28) 93.56	(1,27) 92.23	(1,27) 91.17	(1,27) 90.33	(1,27) 89.67

Table 2.2: Optimal (s, S) pair and minimum cost

γ	0.1	0.2	0.3	0.4	0.5
Optimal (s, S) pair & minimum cost	(1,13) 30.08	(1,13) 57.46	(1,14) 82.77	(1,14) 106.46	(1,14) 128.88
γ	0.6	0.7	0.8	0.9	1
Optimal (s, S) pair & minimum cost	(1,14) 150.22	(1,14) 170.62	(1,14) 190.18	(1,14) 208.99	(1,14) 227.12

2.5.1 Comparison with Schwarz *et al.* [62]

First we provide an analytical comparison (Tables 2.3 and 2.4) followed by numerical comparison (Tables 2.5 and 2.6) of our model with that of Schwarz *et al.* [62] based on a few performance measures. It may be noted that the expressions for various performance measures in column 2 and 3 in Tables 2.3 and 2.4 are in agreement when $\gamma = 1$. For numerical comparison we take the following values of the performance measures: $\lambda = 2$, $\mu = 3$, $\beta = 1$, $s = 1$ and $S = 3$. Table 2.5 and 2.6 indicate our model is superior to that of Schwarz *et al.* [62] in terms of performance measures: E_{loss} is much less, so also W_s and W_q. Of course, holding cost in our model is higher. So also are the mean inventory and mean number of arrivals per unit time.

Table 2.3: Analytical comparison with Schwarz *et al.* [62] for (s, Q) Model

Performance measures	Schwarz *et al.* [62] Model	Our Model
I_m	$\frac{Q}{Q+\frac{\lambda}{\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s}\left(\frac{Q+1}{2}+s-\frac{\lambda}{\beta}\left(1-\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)\right)$	$\sum_{j=1}^{Q+s} j\pi_j$
λ_A	$\frac{\lambda Q}{Q+\frac{\lambda}{\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s}$	$\lambda(1-\pi_0)$
E_{loss}	$\frac{\lambda}{\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s \frac{\lambda}{Q+\frac{\lambda}{\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s}$	$\lambda\pi_0$
W_s	$\frac{1}{\mu-\lambda}\left(1+\frac{\lambda}{Q\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)$	$\frac{1}{(\mu-\lambda)(1-\pi_0)}$
W_q	$\frac{\lambda}{\mu(\mu-\lambda)}\left(1+\frac{\lambda}{Q\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)$	$\frac{\lambda}{\mu(\mu-\lambda)(1-\pi_0)}$

Table 2.4: Analytical comparison with Schwarz *et al.* [62] for (s, S) Model

Performance measures	Schwarz *et al.* [62] Model	Our Model
I_m	$\frac{1}{S-s+\frac{\lambda}{\beta}}\left(\frac{\lambda}{\beta}\left(s-\frac{\lambda}{\beta}\left(1-\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)\right)+\frac{(S+1)S-(s+1)s}{2}\right)$	$\sum_{j=1}^{Q+s} j\widetilde{\pi}_j$
λ_A	$\lambda-\frac{\lambda^2}{(S-s)\beta+\lambda}\left(\frac{\lambda}{\lambda+\beta}\right)^s$	$\lambda\left(1-\widetilde{\pi}_0\right)$
E_{loss}	$\frac{\lambda^2}{(S-s)\beta+\lambda}\left(\frac{\lambda}{\lambda+\beta}\right)^s$	$\lambda\widetilde{\pi}_0$
W_s	$\frac{1}{(\mu-\lambda)}\left(1+\frac{\lambda}{\left(S-\left(s-\frac{\lambda}{\beta}\left(1-\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)\right)\right)\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)$	$\frac{1}{(\mu-\lambda)(1-\widetilde{\pi}_0)}$
W_q	$\frac{\lambda}{\mu(\mu-\lambda)}\left(1+\frac{\lambda}{\left(S-\left(s-\frac{\lambda}{\beta}\left(1-\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)\right)\right)\beta}\left(\frac{\lambda}{\lambda+\beta}\right)^s\right)$	$\frac{\lambda}{\mu(\mu-\lambda)(1-\widetilde{\pi}_0)}$

Table 2.5: (s, Q) Model

Performance measures	Schwarz *et al.* [62] Model ($with\ \gamma = 1$)	Our Model ($with\ \gamma = 0.5$)
I_m	1.1	1.6
λ_A	1.2	1.6
E_{loss}	0.8	0.4
W_s	1.6667	0.25
W_q	1.1111	0.83333

In both models it is difficult to prove analytically the convexity in γ of the cost function is because of the high non-linearity of the function. Nevertheless, all numerical experiments we have performed indicate that this cost function is either monotone decreasing in γ (for moderate values of fixed cost) or first decreases in γ, attains a minimum and then starts going up (for relatively small values of fixed cost) as in Figure 2.1; that is, in the latter case the cost function is strictly convex in γ and hence there exists a global minimum cost. This means that there is a unique probability (γ value) for providing an inventoried item to the customer, at the end of his service, that would ensure minimum cost. If fixed cost is made to tend to zero, the optimal γ value could be seen to be drifting to the left in the $(0, 1]$ interval.

Table 2.6: (s,S) Model

Performance measures	Schwarz *et al.* [62] Model (*with* $\gamma = 1$)	Our Model (*with* $\gamma = 0.5$)
I_m	1.4167	1.8333
λ_A	1.3333	1.6667
E_{loss}	0.66667	0.33333
W_s	1.5	1.2
W_q	1	0.8

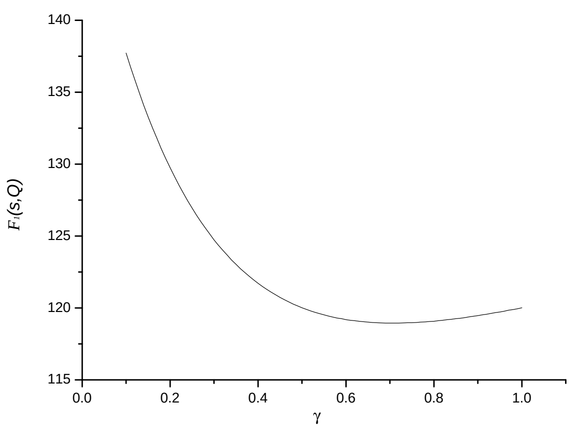

Figure 2.1: γ verses $\mathcal{F}_1(s,Q)$

2.6 $M/G/1$ type queueing-inventory system for (s,Q) policy

So far we have analyzed queueing-inventory process CTMCs'. Next we consider the case of arbitrarily distributed service time, designated as $G(.)$. Thus we have an M/G/1-type queueing-inventory system with positive service time. We assume $\int\limits_0^\infty [1-G(t)]dt$ to be finite. Denote by $t_1, t_2, \ldots$ the successive departure epochs of the first, second,... customers and let $N(t_i^+)$ denote the number of customers left behind by the i^{th} departure and $I(t_i^+)$

denote the on-hand inventory at that epoch, $i = 1, 2, 3, \ldots$. Then the embedded stochastic process $\left\{Z(t_i) = \left(N(t_i^+), I(t_i^+)\right);\ i = 1, 2, \ldots\right\}$ with state space $\Omega_3 = \{(i,j) | i \geq 0;\ 0 \leq j \leq Q + s - 1\}$ is a Markov chain. The one-step transition probability matrix of this Markov chain is

$$\boldsymbol{\mathcal{P}} = \begin{bmatrix} \widetilde{B_0} & \widetilde{B_1} & \widetilde{B_2} & \widetilde{B_3} & \cdots \\ B_0 & B_1 & B_2 & B_3 & \cdots \\ & B_0 & B_1 & B_2 & \cdots \\ & & B_0 & B_1 & \cdots \\ & & & B_0 & \cdots \\ & & & & \ddots \end{bmatrix}.$$

The $(i,j)^{th}$ (in terms of levels) entry of $\boldsymbol{\mathcal{P}}$ describes the probability of transition from i customers to j customers during a service time with different possibilities for the inventory level. These are described below:

(1) Transitions with no arrival during a service time:

$(0,j) \to (0,j-1)$: the probability is $\begin{cases} \gamma \int\limits_0^\infty e^{-(\lambda+\beta)u} dG(u), & for\ 1 \leq j \leq s. \\ \gamma \int\limits_0^\infty e^{-\lambda u} dG(u), & for\ s+1 \leq j \leq S. \end{cases}$

In the following transitions, the inventory level j is greater than zero but less than or equal to s;

a. $(0,j) \to (0,j)$: the probability is $(1-\gamma) \int\limits_0^\infty e^{-(\lambda+\beta)u} dG(u)$.

b. $(0,j) \to (0,j+Q-1)$: the probability is $\gamma \int\limits_0^\infty e^{-\lambda u}(1 - e^{-\beta u}) dG(u)$.

c. $(0,j) \to (0,j+Q)$: the probability is $(1-\gamma) \int\limits_0^\infty e^{-\lambda u}(1 - e^{-\beta u}) dG(u)$.

d. $(0,0) \to (0,Q-1)$: the probability is

$$\gamma \int\limits_{t=0}^{\infty} \int\limits_{v=u}^{t} \int\limits_{u=0}^{v} \beta e^{-\beta u} \lambda (1 - e^{-\lambda(v-u)}) du dv dG(t-v).$$

e. $(0,0) \to (0, j+Q)$: the probability is

$$(1-\gamma) \int\limits_{t=0}^{\infty} \int\limits_{v=u}^{t} \int\limits_{u=0}^{v} \beta e^{-\beta u} \lambda (1 - e^{-\lambda(v-u)}) du dv dG(t-v).$$

(2) Transitions with k arrivals during a service time:

a. $(0,j) \to (k, j-1)$: the probability is $\gamma \int\limits_{0}^{\infty} \frac{e^{-\lambda u}(\lambda u)^k}{k!} dG(u)$, for $j \geq s+1$.

b. $(0,j) \to (k,j)$: the probability is $(1-\gamma) \int\limits_{0}^{\infty} \frac{e^{-\lambda u}(\lambda u)^k}{k!} dG(u)$, for $j \geq s+1$.

(3) Transitions for $i \geq 1$; when the inventory level at the beginning of a service is greater than or equal to $s+1$, inventory level depleting by one or staying at the present position:

$(i,j) \to (i-1, j(\text{ or } j-1))$: the probability is $\begin{cases} (1-\gamma) \int\limits_{0}^{\infty} e^{-\lambda u} dG(u). \\ \gamma \int\limits_{0}^{\infty} e^{-\lambda u} dG(u). \end{cases}$

(4) Transitions for $i \geq 1$; when the inventory level at beginning of service greater than zero but less than or equal to s, inventory level depleting by one or staying at the present position:

$(i,j) \to (i-1, j(\text{ or } j-1))$: the probability is $\begin{cases} (1-\gamma) \int\limits_{0}^{\infty} e^{-(\lambda+\beta)u} dG(u). \\ \gamma \int\limits_{0}^{\infty} e^{-(\lambda+\beta)u} dG(u). \end{cases}$

(5) Transitions with replenishment of inventory, with positive inventory level at the beginning, $1 \leq j \leq s$:

$(i,j) \to (i-1, j-1+Q)$: the probability is $\gamma \int\limits_{0}^{\infty} e^{-\lambda u}(1 - e^{-\beta u}) dG(u)$.

$(i, j) \to (i-1, j+Q)$: the probability is $(1-\gamma)\int_0^\infty e^{-\lambda u}(1-e^{-\beta u})dG(u)$.

(6) Transitions for $i > 0, k \geq 0$, when there is no inventory at the beginning:

a. $(i, 0) \to (i+k-1, Q-1)$: the probability is

$$\gamma \int_0^\infty \int_u^\infty \beta e^{-\beta u} \frac{e^{-\lambda(t-u)}(\lambda(t-u))^k}{k!} dG(t-u)du.$$

b. $(i, 0) \to (i+k-1, Q)$: the probability is

$$(1-\gamma) \int_0^\infty \int_u^\infty \beta e^{-\beta u} \frac{e^{-\lambda(t-u)}(\lambda(t-u))^k}{k!} dG(t-u)du.$$

(7) Transitions for $i > 0, k \geq 0$; inventory level j is such that $1 \leq j \leq s$ and no replenishment during a service:

a. $(i, j) \to (i+k-1, j-1)$: the probability is $\gamma \int_0^\infty \frac{e^{-(\lambda+\beta)t}(\lambda t)^k}{k!} dG(t)$.

b. $(i, j) \to (i+k-1, j)$: the probability is $(1-\gamma) \int_0^\infty \frac{e^{-(\lambda+\beta)t}(\lambda t)^k}{k!} dG(t)$.

(8) Transitions for $i > 0, k \geq 0$; replenishment occurs during service, that is, $1 \leq j \leq s$ at the beginning:

a. $(i, j) \to (i+k-1, j+Q-1)$: the probability is $\gamma \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^k}{k!}(1-e^{-\beta t})dG(t)$.

b. $(i, j) \to (i+k-1, j+Q)$: the probability is $(1-\gamma) \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^k}{k!}(1-e^{-\beta t})dG(t)$.

All other transition pairs have probability zero. The above transitions probabilities could be made use of to compute the inventory level probabilities at departure epochs (we have to sum over the number of arrivals during a service time).

The distribution of the number of customers at departure epochs and at arbitrary epochs have the same form as in an $M/G/1$ queue. However, the number of items in the inventory can never be S at departure epochs. This is also the case in the $M/M/1$ set up.

2.7 Emptiness time distribution for $M/M/1/1$ queueing-inventory system

We now compute the distribution of the time till the inventory becomes empty (zero). We consider the inventory level, starting from S, until the next epoch when all items in the inventory becomes zero. Let χ denote the random variable "time until the items in the inventory becomes zero"starting with S items. We consider the CTMC $\{(\mathcal{I}(t), \mathcal{C}(t))|t \geq 0\}$. The state space of the CTMC $\{(\mathcal{I}(t), \mathcal{C}(t))|t \geq 0\}$ is

$$\{(\ell, m)\,/1 \leq \ell \leq S, m = 0, 1\} \cup \{\Delta\},$$

where $\{\Delta\}$ $(= (0,0))$ is the absorbing state which represents the state that the inventory level becomes zero, starting from the state $\{(1,1)\}$. Clearly, $\Im$ is a finite state space Markov chain. The possible transitions and the corresponding rates are given in Table 2.7.

Thus the infinitesimal generator $\boldsymbol{\mathcal{Q}}$ of the Markov chain $\{(\mathcal{I}(t), \mathcal{C}(t))|t \geq 0\}$ is of the form $\boldsymbol{\mathcal{Q}} = \begin{bmatrix} \boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{G}^0} \\ \mathbf{0} & 0 \end{bmatrix}$ with initial probability vector $\boldsymbol{\alpha} = (0, 0, \ldots, 1, 0)$ where 1 is the in the $(2S)^{th}$ position; $\boldsymbol{\mathcal{G}}$ is of order $2S+1$; $\boldsymbol{\mathcal{G}^0}$ is a $2S+1$ component column vector such that $\boldsymbol{\mathcal{G}}\mathbf{e}+\boldsymbol{\mathcal{G}^0}$=0. This time duration follows PH distribution with representation $(\boldsymbol{\alpha}, \boldsymbol{\mathcal{G}})$. Therefore the expected time

Table 2.7: The transitions in the CTMC $\{(\mathcal{I}(t), \mathcal{C}(t))|t \geq 0\}$ and corresponding rates

Form	To	Rate	
$(\ell, 0)$	$(\ell, 1)$	λ	$\ell = 1, 2, \ldots, S.$
$(\ell, 1)$	$(\ell + 1, 1)$	λ	$\ell = 1, 2, \ldots, S.$
$(\ell, 1)$	$(\ell - 1, 1)$	μ	$\ell = 2, 3, \ldots, S.$
(ℓ, m)	$(\ell + Q, m)$	β	$\ell = 2, 3, \ldots, s.;\ m = 0, 1.$
$(\ell, 1)$	$(\ell, 1)$	$-(\lambda + \beta + \mu)$	$\ell = 1, 2, \ldots, s.$
$(\ell, 0)$	$(\ell, 1)$	$-(\lambda + \mu)$	$\ell = 1, 2, \ldots, s.$
$(1, 1)$	Δ	μ	

until the inventory become zero is,

$$E(\chi) = -\boldsymbol{\alpha} \left(\boldsymbol{\mathcal{G}^{-1}}\right) \mathbf{e}.$$

Chapter 3

On a two stage supply chain inventory with positive service time and loss

3.1 Introduction

In the previous chapter we have considered the case of no inventory provided at the end of a service to the departing customer. In the present chapter we extend this concept to production inventory with positive service time. Thus we assume that the item produced is accepted with some probability and rejected with complementary probability. Similarly we assume that at the end of a service, a customer is provided/accepts the inventory with a pre-assigned probability and with complementary probability he has to go empty hand/declines the item. We impose the condition that no customer

Some results of this chapter are included in the following paper.
A. Krishnamoorthy and R. Manikandan : On a two stage supply chain inventory with positive service time and loss (Under review).

joins when the on-hand inventory is zero (those who are already present, stay back in the system until served). Thus this chapter generalizes the work reported in Krishnamoorthy and Vishwanath [42].

We arrange the presentation in this chapter as indicated below: Section 3.2 provides the mathematical formulation of the problem under study. The analysis of the system is carried out in section 3.3. In particular, we derive the long run stability of the system. Then, under this condition we show that the system state can be decomposed: that is to say, we get the system state distribution as the product of marginal distribution of the components. Next we compute system performance measures that have significant impact. Further, in order to construct an appropriate cost function, we compute the expected length of a production cycle in section 3.4. A few results on up and down crossings of level s on a production cycle are also discussed in that section. Having achieved that we construct a cost function. Then we look for the optimal pair (s, S) values that would result in cost minimization for different pairs of values of γ and δ. This is reported in section 3.5. Finally we discuss the first emptiness time distribution for the $M/M/1/1$ queueing-inventory system with production.

3.2 Description of the model

We consider an (s, S) production inventory system with a single server. Demands by customers for the item occur according to a Poisson process of rate λ. Processing of the customer request requires a random amount of time, which is exponentially distributed with parameter μ. However, as assumed in the previous chapter it is not essential that the item from inventory is provided to the customer at the end of a service. More precisely, an item from inventory is provided to a customer with probability γ at the end of his service and with probability $1-\gamma$ the customer leaves the system empty handed. When the inventory level depletes to s, the production process is immediately switched on. Each production is of 1 unit and the production process is kept in the on mode until inventory level becomes S. To produce an item it takes an amount of time which is exponentially distributed with parameter β. A produced item is not necessarily added to the inventory due to manufacturing defect: with probability δ it is accepted and with probability $1-\delta$ the item is rejected. We assume that no customer is allowed to join the queue when the inventory level is zero; such demands are considered as lost. It is assumed that the amount of time for the item produced to reach the retail shop is negligible. Thus the system is a CTMC $\{\mathcal{X}(t); t \geq 0\} = \{(\mathcal{N}(t), \mathcal{I}(t), \mathcal{P}(t)); t \geq 0\}$. The production process is in on mode if $0 \leq \mathcal{I}(t) \leq s$ and it is in off mode if $\mathcal{I}(t) = S$; but when the inventory level lies between $s+1$ and $S-1$, $\mathcal{P}(t)$ is either 0 or 1 according as the production is in off or in on mode, respectively. Thus to describe the status of the process we need write $\mathcal{P}(t) = 0$ or 1 only when $\mathcal{I}(t)$ takes values $s+1, \ldots, S-1$. Thus the state space of the CTMC is $\mathbf{\Omega}= \bigcup_{i=0}^{\infty} \mathcal{L}(i)$, where $\mathcal{L}(i)$, called level i of the CTMC, is given by, $\{(i,j); 0 \leq j \leq s\} \bigcup \{(i,j,k); s+1 \leq j \leq S-1, k = 0,1\} \bigcup \{(i,S)\}, \forall\ i \geq 0$. The number of states (called phases in that level) within i^{th} level is $2S-s$. The infinitesimal generator of this CTMC is

$$\mathcal{W} = \begin{bmatrix} B_1 & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \ldots \\ & & \ddots & \ddots & \ddots \end{bmatrix}.$$

The block matrices appearing on the right side above are explained below:

$$B_1 = \begin{pmatrix}
-\delta\beta & \delta\beta & & & & & & & & \\
& -(\lambda+\delta\beta) & \delta\beta & & & & & & & \\
& & \ddots & \ddots & & & & & & \\
& & -(\lambda+\delta\beta) & & \delta\beta & & & & & \\
& & & -(\lambda+\delta\beta) & V_1 & & & & & \\
& & & & U & V_2 & & & & \\
& & & & & U & V_2 & & & \\
& & & & & & \ddots & \ddots & & \\
& & & & & & & U & V_2 & \\
& & & & & & & & U & V_3 \\
& & & & & & & & & -\lambda
\end{pmatrix},$$

with $U = \begin{bmatrix} -\lambda & 0 \\ 0 & -(\lambda+\delta\beta) \end{bmatrix}, V_1 = \begin{bmatrix} 0 & \delta\beta \end{bmatrix}, V_2 = \begin{bmatrix} 0 & 0 \\ 0 & \delta\beta \end{bmatrix}, V_3 = \begin{bmatrix} 0 \\ \delta\beta \end{bmatrix}$;
entries of B_1 corresponding to transition rates within level 0.
$A_0 = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \lambda I_{(2S-s)-1} \end{bmatrix}$, $A_1 = B_1 - \frac{\mu}{\lambda} A_0$

$$\text{and } A_2 = \begin{pmatrix}
0 & \cdots & & & & & & & \cdots & 0 \\
& \gamma\mu & (1-\gamma)\mu & & & & & & & \\
& & \ddots & \ddots & & & & & & \\
& & & \gamma\mu & (1-\gamma)\mu & & & & & \\
& & & F_1 & F_2 & & & & & \\
& & & & F_3 & F_2 & & & & \\
& & & & & F_3 & F_2 & & & \\
& & & & & & \ddots & \ddots & & \\
& & & & & & & F_3 & F_2 & 0 \\
& & & & & & & & F_4 & (1-\gamma)\mu
\end{pmatrix},$$

with $F_1 = \begin{bmatrix} \gamma\mu \\ \gamma\mu \end{bmatrix}, F_2 = \begin{bmatrix} (1-\gamma)\mu & 0 \\ 0 & (1-\gamma)\mu \end{bmatrix}, F_3 = \begin{bmatrix} \gamma\mu & 0 \\ 0 & \gamma\mu \end{bmatrix}$

and $F_4 = \begin{bmatrix} 0 & \gamma\mu \end{bmatrix}$.

3.3 Analysis of the system

In this section we perform the steady-state analysis of the (s, S) production inventory model under study by first establishing the stability condition of the system. Define A=$A_0 + A_1 + A_2$. This is the infinitesimal generator of the finite state CTMC corresponding to the inventory level $\{0, \ldots, s\} \cup \{(j,k); s+1 \le j \le S-1, k = 0,1\} \cup \{S\}$. Let $\boldsymbol{\varphi}$ denote the steady-state probability vector of A. That is $\boldsymbol{\varphi}$ satisfies

$$\boldsymbol{\varphi} A = 0, \ \boldsymbol{\varphi}\boldsymbol{e} = 1. \tag{3.1}$$

Using the above relations, we get the components of the probability vector $\boldsymbol{\varphi}$ explicitly as:

$$\varphi\left(s-i\right) = \varphi\left(S\right)\frac{\gamma\mu}{\delta\beta-\gamma\mu}\left(1-\left(\frac{\gamma\mu}{\delta\beta}\right)^{S-s}\right)\left(\frac{\gamma\mu}{\delta\beta}\right)^{i}, 0 \le i \le s,$$

$$\varphi\left(i,0\right) = \varphi\left(S\right), s+1 \le i \le S-1,$$

$$\varphi\left(i,1\right) = \varphi\left(S\right)\frac{\gamma\mu}{\delta\beta-\gamma\mu}\left(1-\left(\frac{\gamma\mu}{\delta\beta}\right)^{S-i}\right), s+1 \le i \le S-1.$$

and the unknown probability

$$\varphi(S) = \frac{\left(\frac{\gamma\mu}{\delta\beta}-1\right)^{2}}{\left(\frac{\gamma\mu}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\mu}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\mu}{\delta\beta}-1\right)}.$$

Since the Markov chain under study is an LIQBD process, it is stable if and only if the left drift rate exceeds the right drift rate. That is,

$$\boldsymbol{\varphi} A_0\mathbf{e} < \boldsymbol{\varphi} A_2\mathbf{e}. \tag{3.2}$$

We have the following lemma:

Lemma 3.3.1. The stability condition of the (s, S) production inventory model is given by $\lambda < \mu$.

Proof. From the well known result in Neuts [47] on the positive recurrence of A, we have $\boldsymbol{\varphi}\ A_0\mathbf{e} < \boldsymbol{\varphi} A_2\mathbf{e}$. With a bit of computation, this simplifies to the result $\lambda < \mu$. For future reference we define ρ as

$$\rho = \frac{\lambda}{\mu}. \tag{3.3}$$

□

3.3.1 Steady-state analysis

For computing the steady-state probability vector of the process $\{\mathcal{X}(t); t \geq 0\}$, we first consider a production inventory system with negligible service time where no backlog of customers is allowed (that is when inventory level is zero, no demand joins the system). The rest of the assumptions such as those on the arrival process and lead time are the same as given earlier. Designate the Markov chain so obtained as $\{\widetilde{\mathcal{X}}(t); t \geq 0\}$=$\{(\mathcal{I}(t), \mathcal{P}(t))\,; t \geq 0\}$. Its infinitesimal generator $\widetilde{\boldsymbol{\mathcal{W}}}$ is given by,

$$\widetilde{\boldsymbol{\mathcal{W}}} = \begin{pmatrix} -\delta\beta & \delta\beta & & & & & & & & & \\ \gamma\lambda & -(\gamma\lambda+\delta\beta) & \delta\beta & & & & & & & & \\ & \ddots & \ddots & & \ddots & & & & & & \\ & & \gamma\lambda & -(\gamma\lambda+\delta\beta) & \delta\beta & & & & & & \\ & & & \gamma\lambda & -(\gamma\lambda+\delta\beta) & V_1 & & & & & \\ & & & & \hat{F}_1 & U & V_2 & & & & \\ & & & & & \hat{F}_3 & U & V_2 & & & \\ & & & & & & \ddots & \ddots & \ddots & & \\ & & & & & & & \hat{F}_3 & U & V_2 & \\ & & & & & & & & \hat{F}_3 & U & V_3 \\ & & & & & & & & & \hat{F}_4 & -\gamma\lambda \end{pmatrix},$$

where $\hat{F}_1 = \frac{\lambda}{\mu}F_1$, $\hat{F}_3 = \frac{\lambda}{\mu}F_3$, $\hat{F}_4 = \frac{\lambda}{\mu}F_4$, and all other sub matrices are as defined previously for matrix .

Let $\boldsymbol{\pi}$=$(\pi(0), \pi(1), \ldots, \pi(s), \pi(s+1, 1), \ldots, \pi(S-1, 1), \pi(s+1, 0), \ldots, \pi(S-$

$1,0),\pi(S))$ be the steady-state probability vector of the process $\widetilde{\mathcal{X}}(t)$=$\{\mathcal{I}(t);t\geq 0\}$. Then $\boldsymbol{\pi}$ satisfies the relations

$$\boldsymbol{\pi}\widetilde{\boldsymbol{\mathcal{W}}}=0,\ \ \boldsymbol{\pi}\mathbf{e}=1 \tag{3.4}$$

That is, at arbitrary epochs the components of the inventory level probability distribution $\boldsymbol{\pi}$ is given by:

$$\pi\left(s-i\right)=\pi\left(S\right)\frac{\gamma\lambda}{\delta\beta-\gamma\lambda}\left(1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-s}\right)\left(\frac{\gamma\lambda}{\delta\beta}\right)^{i},0\leq i\leq s,$$
$$\pi\left(i,0\right)=\pi\left(S\right),s+1\leq i\leq S-1,$$
$$\pi\left(i,1\right)=\pi\left(S\right)\frac{\gamma\lambda}{\delta\beta-\gamma\lambda}\left(1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-i}\right),s+1\leq i\leq S-1.$$

and the unknown probability

$$\pi(S)=\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^{2}}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}.$$

Using the components of the probability vector $\boldsymbol{\pi}$, we shall find the steady-state probability vector of the CTMC $\{\mathcal{X}(t);t\geq 0\}$. For this, let $\boldsymbol{x}$ be the steady-state probability vector of the original system. Then the steady-state vector must satisfy the set of equations

$$\boldsymbol{x}\boldsymbol{\mathcal{W}}=0,\ \ \boldsymbol{x}\mathbf{e}=1. \tag{3.5}$$

partition $\boldsymbol{x}$ by levels as

$$\boldsymbol{x}=(\boldsymbol{x_0},\boldsymbol{x_1},\boldsymbol{x_2},\ldots) \tag{3.6}$$

where the subvectors of $\boldsymbol{x}$ are further partitioned as, $\boldsymbol{x_i}=(x_i(0),x_i(1),\ldots,x_i(s),$ $x_i(s+1,1),\ldots x_i(S-1,1),x_i(s+1,0),\ldots x_i(S-1,0),x_i(S)),i\geq 0$. Then the above system of equations reduces to

$$\boldsymbol{x_0}B_1+\boldsymbol{x_1}A_2=0 \tag{3.7}$$

$$\boldsymbol{x_i}A_0 + \boldsymbol{x_{i+1}}A_1 + \boldsymbol{x_{i+2}}A_2 = 0, i \geq 0 \tag{3.8}$$

Assume that

$$\boldsymbol{x_0} = \xi\boldsymbol{\pi} \tag{3.9}$$

$$\boldsymbol{x_i} = \xi\left(\frac{\lambda}{\mu}\right)^i \boldsymbol{\pi}, i \geq 1 \tag{3.10}$$

where ξ is a constant to be determined. We verify that the equations (3.7) and (3.8) are satisfied by (3.9) and (3.10). For (3.7), we have

$$\boldsymbol{x_0}B_1 + \boldsymbol{x_1}A_2 = \xi\boldsymbol{\pi}\left(B_1 + \frac{\lambda}{\mu}A_2\right) \tag{3.11}$$

and from relation (3.8), we have,

$$\boldsymbol{x_i}A_0 + \boldsymbol{x_{i+1}}A_1 + \boldsymbol{x_{i+2}}A_2 = \xi\left(\frac{\lambda}{\mu}\right)^{i+1} \boldsymbol{\pi}\left(B_1 + \frac{\lambda}{\mu}A_2\right) \tag{3.12}$$

Now from the matrices B_1, A_2 and $\widetilde{\boldsymbol{\mathcal{W}}}$, it follows that

$$B_1 + \frac{\lambda}{\mu}A_2 = \widetilde{\boldsymbol{\mathcal{W}}} \tag{3.13}$$

Also from (3.4) we have $\boldsymbol{\pi}\widetilde{\boldsymbol{\mathcal{W}}} = 0$. Hence the right hand side of the equation (3.11) and (3.12) are zero. Hence if we take the vector $\boldsymbol{x}$ as given by (3.6), it follows that (3.7) and (3.8) are satisfied. Now applying the normalizing condition $\boldsymbol{x}\mathbf{e}$=1, we get

$$\xi\left[1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \left(\frac{\lambda}{\mu}\right)^3 + \cdots\right] = 1$$

Hence under the condition that $\lambda < \mu$, we have

$$\xi = 1 - \frac{\lambda}{\mu}. \tag{3.14}$$

Thus we arrive at

Theorem 3.3.1. *Under the necessary and sufficient condition $\lambda < \mu$ for stability, the components of the steady-state probability vector of the process $\{\mathcal{X}(t); t \geq 0\}$ with generator matrix $\boldsymbol{\mathcal{W}}$, is given by (3.9), (3.10) and (3.14). That is, $\boldsymbol{x_0} = (1-\rho)\boldsymbol{\pi}, \boldsymbol{x_i} = (1-\rho)\rho^i\boldsymbol{\pi}, i \geq 1$ where ρ is as defined in (3.3) and $\boldsymbol{\pi}$ is the inventory level probability vector.*

The consequence of the above Theorem 3.3.1 is that the joint distribution of the two dimensional system can be decomposed into probabilities of two distinct one dimensional objects namely, number of customers and the number of inventoried items in the system. Thus for example, when production is on, denoting by $P(z)$ and $Q(z)$ the probability generating functions of the number of customers in the system and the number of items in the inventory respectively, then the joint generating function (the generating function of the system state) , can be written as the product of the marginal generating functions. This is the case when the production is off as well (that is the inventory level is dropping from S, but is above s).

3.3.2 Performance measures

We enumerate below the long run system performance characteristics that are useful in formulating an optimization problem.

- Mean number of customers in the system, $L_s = \frac{\lambda}{\mu-\lambda}$.
- Mean number of customers waiting in the system during the stock out period,
  $W_s = L_s\pi(0)$
  $= \frac{\lambda}{\mu-\lambda}\left(\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+1}\left(\frac{1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-s}}{\left(1-\frac{\gamma\lambda}{\delta\beta}\right)}\right)\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^2}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right)\right).$

- Mean number of customers waiting in the system when inventory is available,
$\widetilde{W_s} = L_s\left(1-\pi(0)\right)$
$= \frac{\lambda}{\mu-\lambda}\left(1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+1}\left(\frac{1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-s}}{\left(1-\frac{\gamma\lambda}{\delta\beta}\right)}\right)\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^2}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right)\right).$

- Mean number of items in the inventory,
$E_{inv} = \sum\limits_{i=0}^{s} i\pi(i) + \sum\limits_{i=s+1}^{S-1} i(\pi(i,0)+\pi(i,1))$
$= \frac{2-(S-s)(S+s+3)}{2\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^2}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right).$

- Mean rate at which the production process is *switched on*,
$E_{on} = \gamma\mu\left(\sum\limits_{i=1}^{\infty}\xi\left(\frac{\lambda}{\mu}\right)^{i}\pi\left(s+1,0\right)\right)$
$= \gamma\lambda\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^2}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right).$

- Expected rate at which items are added to the inventory,
$E_{rp} = \delta\beta\left(\sum\limits_{i=0}^{s}\pi(i) + \sum\limits_{i=s+1}^{S-1} +\pi(i,1)\right)$
$= \delta\beta\left(1-(S-s)\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^2}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right)\right).$

- Expected *loss* rate of the *manufactured item* due to rejection,
$M_{loss} = (1-\delta)\beta\left(\sum\limits_{i=0}^{s}\pi(i) + \sum\limits_{i=s+1}^{S-1}\pi(i,1)\right)$
$= (1-\delta)\beta\left(1-(S-s)\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^2}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right)\right).$

- Expected *loss* rate of *customers* (customers not joining the system for want of inventory),

$$C_{loss} = \lambda\pi(0)$$
$$= \lambda\left(\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+1}\left(\frac{1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-s}}{\left(1-\frac{\gamma\lambda}{\delta\beta}\right)}\right)\left(\frac{\left(\frac{\gamma\lambda}{\delta\beta}-1\right)^{2}}{\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S+2}-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{s+2}-(S-s)\left(\frac{\gamma\lambda}{\delta\beta}-1\right)}\right)\right).$$

3.4 Analysis of the production cycle time

The production process is switched on at a service completion epoch t_0, which started with $s+1$ items in the inventory with one item from inventory supplied to the customer and the production process being in off mode. The production process, once turned on, is turned off only at an epoch t_1 at which the inventory level in the system reaches S. A production cycle starts with the switching on of the production process as inventory level drops progressively to s from S and terminates with the inventory level reaching S. We analyze the length $t_1 - t_0$ of the production cycle as the time until absorption in a CTMC $\mathbf{\Psi}= \{(\mathcal{N}(t), \mathcal{I}(t)); t \geq 0\}$, the variation of $\mathcal{N}(t)$ is from 0 to ∞ and $\mathcal{I}(t)$ varies from 0 to $S-1$. The state space of $\mathbf{\Psi}$ is given by $\bigcup\limits_{i=0}^{\infty}\left\{\widetilde{i}\right\}\bigcup\{\Delta_1\}$, where each level $\left\{\widetilde{i}\right\}$ is given by $\left\{\widetilde{i}\right\} = \{(i,j); 0 \leq j \leq S-1\}$ and Δ_1 denotes the single absorbing state, which represents switching off of the production process(that is, inventory level reaches S). Except for the absorbing state Δ_1, transitions between states in $\mathbf{\Psi}$ are the same as those in $\mathbf{\Omega}$. The infinitesimal generator $\boldsymbol{\mathcal{Q}_c}$ of the process $\mathbf{\Psi}$ has the form $\boldsymbol{\mathcal{Q}_c}=\begin{bmatrix} \boldsymbol{\mathcal{H}} & -\boldsymbol{\mathcal{H}}\mathbf{e} \\ \mathbf{0} & 0 \end{bmatrix}$, where $\boldsymbol{\mathcal{H}}$ is given by

$$\boldsymbol{\mathcal{H}}=\begin{bmatrix} \widehat{B_1} & \widehat{A}_0 & & & \\ \widehat{A}_2 & \widehat{A}_1 & \widehat{A}_0 & & \\ & \widehat{A}_2 & \widehat{A}_1 & \widehat{A}_0 & \dots \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

$$\text{with } \widehat{B_1} = \begin{pmatrix} -\delta\beta & \delta\beta & & & \\ & -(\lambda+\delta\beta) & \delta\beta & & \\ & & \ddots & \ddots & \\ & & & -(\lambda+\delta\beta) & \delta\beta \\ & & & & -(\lambda+\delta\beta) \end{pmatrix},$$

$$\widehat{A}_1 = \begin{pmatrix} -\delta\beta & \delta\beta & & & \\ & -(\lambda+\mu+\delta\beta) & \delta\beta & & \\ & & \ddots & \ddots & \\ & & & -(\lambda+\mu+\delta\beta) & \delta\beta \\ & & & & -(\lambda+\mu+\delta\beta) \end{pmatrix},$$

$$\widehat{A}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \gamma\mu & (1-\gamma)\mu & & & \\ & \ddots & \ddots & & \\ & & \gamma\mu & (1-\gamma)\mu & \\ & & & \gamma\mu & (1-\gamma)\mu \end{pmatrix}, \widehat{A}_0 = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \lambda I_{S-1} \end{pmatrix}$$

Define the row vector $\boldsymbol{\eta^\tau}$=$(\eta_0^\tau, \eta_1^\tau, \eta_2^\tau, \dots)$, where each $\boldsymbol{\eta_i}$ is a column vector with S entries, such that $\eta_i(j)$ is the expected time until absorption of the process $\boldsymbol{\Psi}$, from state (i, j). Also define the probability vector $\boldsymbol{\sigma = (\sigma_0, \sigma_1, \sigma_2, \dots)}$, where each $\boldsymbol{\sigma_i}$ is a row vector of dimension $S \times 1$ such that $\sigma_i(j)$ is the probability that the production process is switched on with i customers and j inventory in the system. Clearly $\sigma_i(j) = 0$ if $j \neq s$ and $\sigma_i(s)$ can be found using the steady-state probability vector $\boldsymbol{x}$ of the process $\{\mathcal{X}(t); t \geq 0\}$ as follows:

$$\sigma_i(s) = \frac{\xi \left(\frac{\lambda}{\mu}\right)^{i+1} \pi\left(S\right)\gamma\mu}{P_{on}} = \frac{\xi \left(\frac{\lambda}{\mu}\right)^{i+1} \pi\left(S\right)\gamma\mu}{\pi\left(S\right)\lambda} = \xi \left(\frac{\lambda}{\mu}\right)^{i} \gamma, \ for\ all\ i \geq 0.$$

Thus the expected length of the production cycle,

$$E_{cycle} = \sum_{i=0}^{\infty} \sigma_i(s)\eta_i(s) = \sum_{i=1}^{\infty} \xi \left(\frac{\lambda}{\mu}\right)^{i} \gamma\eta_i(s) \tag{3.15}$$

Now for computing the vector $\boldsymbol{\eta},$ a simple probabilistic argument shows that the vector $\boldsymbol{\eta}$ satisfies the infinite system of equations given by $\boldsymbol{\mathcal{H}\eta = -}\mathbf{e}$, which implies

$$\widehat{B_1}\boldsymbol{\eta_0} + \widehat{A}_0\boldsymbol{\eta_1} \boldsymbol{= -}\mathbf{e}\boldsymbol{,} \tag{3.16}$$

$$\widehat{A}_2\boldsymbol{\eta_{i-1}} + \widehat{A}_1\boldsymbol{\eta_i} + \widehat{A}_0\boldsymbol{\eta_{i+1}} \boldsymbol{= -}\mathbf{e}, i \geq 1. \tag{3.17}$$

For future reference we define

$$\boldsymbol{P_0} = \widehat{B_1}\boldsymbol{\eta_0} + \widehat{A}_0\boldsymbol{\eta_1}, \tag{3.18}$$

and

$$\boldsymbol{P_i} = \widehat{A}_2\boldsymbol{\eta_{i-1}} + \widehat{A}_1\boldsymbol{\eta_i} + \widehat{A}_0\boldsymbol{\eta_{i+1}}, i \geq 1. \tag{3.19}$$

For solving the above infinite system of equations, we use the same technique as that was employed in the case of finding the steady-state vector; that is by seeking the help of the expected cycle time of the production process $\widetilde{E}_{cycle}$ in a production inventory system with negligible service time and no backlog of demands. For computing $\widetilde{E}_{cycle}$, we define a CTMC $\widetilde{\boldsymbol{\Psi}}$=$\{\mathcal{I}(t); t \geq 0\}$ with an absorbing state Δ_2, that represents the switching off of the production process. Here a production cycle, $\mathcal{I}(t)$ denotes the inventory level at time t. The state space of a CTMC, $\{\mathcal{I}(t); t \geq 0\}$ is given by $\{0, 1, 2, ..., S-1\} \bigcup \{\Delta_2\}$ and its infinitesimal generator is given by,

$\boldsymbol{\mathcal{G}}= \begin{bmatrix} \boldsymbol{\mathcal{D}} & -\boldsymbol{\mathcal{D}}\mathbf{e} \\ \mathbf{0} & 0 \end{bmatrix}$ with

$$\boldsymbol{\mathcal{D}} = \begin{pmatrix} -\delta\beta & \delta\beta & & & \\ \gamma\lambda & -(\gamma\lambda+\delta\beta) & \delta\beta & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma\lambda & -(\gamma\lambda+\delta\beta) & \delta\beta \\ & & & \gamma\lambda & -(\gamma\lambda+\delta\beta) \end{pmatrix}.$$

Now $\widetilde{E}_{cycle}$ is the $(s+1)^{th}$ entry of the column vector $\boldsymbol{-\mathcal{D}^{-1}}\mathbf{e}$.
Let $\boldsymbol{-\mathcal{D}^{-1}}\mathbf{e}$=$(\Gamma_0, \Gamma_1, \ldots, \Gamma_{S-1})$. Then the relation $\boldsymbol{\mathcal{D}}\left(-\boldsymbol{\mathcal{D}^{-1}}\mathbf{e}\right) = -\mathbf{e}$ gives us the following equations

$$-\delta\beta\ \Gamma_0 + \delta\beta\ \Gamma_1 = -1,$$
$$\gamma\lambda\ \Gamma_{i-1} - (\gamma\lambda+\delta\beta)\ \Gamma_i + \delta\beta\ \Gamma_{i+1} = -1,\ 1 \le i \le S-2,$$
$$\gamma\lambda\ \Gamma_{S-2} - (\gamma\lambda+\delta\beta)\ \Gamma_{S-1} = -1.$$

Some algebraic manipulation of the above equations results in the following equations:
$\Gamma_i - \Gamma_{i+1} = \frac{1}{\delta\beta}\sum\limits_{j=0}^{i}\left(\frac{\gamma\lambda}{\delta\beta}\right)^j, \Gamma_{S-1} = \frac{1}{\delta\beta}\sum\limits_{j=0}^{S-1}\left(\frac{\gamma\lambda}{\delta\beta}\right)^j$ and by solving these equations we get, $\Gamma_s = \frac{1}{\delta\beta}\left((S-s)\sum\limits_{j=0}^{s}\left(\frac{\gamma\lambda}{\delta\beta}\right)^j + \sum\limits_{j=s+1}^{S-1}(S-j)\left(\frac{\gamma\lambda}{\delta\beta}\right)^j\right)$
Now by using the relations (3.18) and (3.19), $\sum\limits_{i=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^i \boldsymbol{P_i}$ implies

$$\left(\widehat{B_1} + \frac{\lambda}{\mu}\widehat{A}_2\right)\boldsymbol{\eta_0} + \sum_{i=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^i\left(\widehat{A}_0 + \left(\frac{\lambda}{\mu}\right)\widehat{A}_1 + \left(\frac{\lambda}{\mu}\right)^2\widehat{A}_2\right)\boldsymbol{\eta_{i+1}} = -\sum_{i=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^i\mathbf{e} \tag{3.20}$$

Here we get the following identities:

$$\left(\widehat{B_1} + \frac{\lambda}{\mu}\widehat{A}_2\right) = \boldsymbol{\mathcal{D}},$$

$$\widehat{A}_0 + \left(\frac{\lambda}{\mu}\right)\widehat{A}_1 = \left(\frac{\lambda}{\mu}\right)\widehat{B_1},$$

$$\widehat{A}_0 + \left(\frac{\lambda}{\mu}\right)\widehat{A}_1 + \left(\frac{\lambda}{\mu}\right)^2\widehat{A}_2 = \left(\frac{\lambda}{\mu}\right)\widehat{B_1} + \left(\frac{\lambda}{\mu}\right)^2\widehat{A}_2 = \left(\frac{\lambda}{\mu}\right)\boldsymbol{\mathcal{D}}.$$

These identities are applied in to the equation (3.20) to get

$$\boldsymbol{\mathcal{D}}\sum_{i=0}^{\infty}\left(\frac{\lambda}{\mu}\right)^i\boldsymbol{\eta_i} = -\frac{1}{\xi}\mathbf{e}.$$

That is,

$$\sum_{i=0}^{\infty}\xi\left(\frac{\lambda}{\mu}\right)^i\boldsymbol{\eta_i} = \boldsymbol{-(\mathcal{D}^{-1})}\mathbf{e}. \tag{3.21}$$

From the equations (3.15) and (3.21), it follows that the expected duration of a production run, E_{cycle} is the same as $\widetilde{E}_{cycle}$, the expected length of a production cycle in a production inventory system with negligible service time. Thus the expected cycle time of the production process E_{cycle} is given by

$$E_{cycle} = \frac{1}{\delta\beta}\left((S-s)\sum_{j=0}^{s}\left(\frac{\gamma\lambda}{\delta\beta}\right)^j + \sum_{j=s+1}^{S-1}(S-j)\left(\frac{\gamma\lambda}{\delta\beta}\right)^j\right).$$

We record this in the following

Lemma 3.4.1. The expected length of a production cycle is given by $E_{cycle} = \frac{1}{\delta\beta}\left((S-s)\sum\limits_{j=0}^{s}\left(\frac{\gamma\lambda}{\delta\beta}\right)^j + \sum\limits_{j=s+1}^{S-1}(S-j)\left(\frac{\gamma\lambda}{\delta\beta}\right)^j\right)$
$= \frac{1}{\gamma\lambda}\left(\frac{1}{\pi(S)} - (S-s)\right).$

Corollary 1. The expected number of production up-crossings of level s is given by $\overline{E} = \left[x_0(s)\frac{\delta\beta}{\lambda+\delta\beta} + \frac{\delta\beta}{\lambda+\mu+\delta\beta}\sum\limits_{i=1}^{\infty}x_i(s)\right].E_{cycle}$
$= (1-(S-s)\pi(S))\left(\frac{\delta\beta}{\delta\beta-\gamma\lambda}\right)\left(1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-s}\right)\left(\frac{1-\rho}{\lambda+\delta\beta} + \frac{\rho}{\lambda+\mu+\delta\beta}\right).$

Corollary 2. The expected number of production down crossings of level s is given by $\underline{E} = (1-(S-s)\pi(S))\left(\frac{\gamma\lambda}{(\delta\beta-\gamma\lambda)(\lambda+\gamma\mu+\delta\beta)}\right)\left(1-\left(\frac{\gamma\lambda}{\delta\beta}\right)^{S-s}\right)$.

Some of the above down and/ up-crossings of s may not go below/above s. The expected number of such crossings are given in the following corollaries

Corollary 3. The expected number of production down crossings that goes below s in a production cycle, $P_{down} = \underline{E} * $ *Probability of a service completion before addition of an inventoried item.* That is,

$$P_{down} = \underline{E}.\left(\sum_{i=1}^{\infty}\xi\left(\frac{\lambda}{\mu}\right)^i\left(\frac{\gamma\mu}{\delta\beta+\gamma\mu}\right)+\xi\int\limits_{t=0}^{\infty}\int\limits_{v=0}^{t}\lambda e^{-\lambda v}\gamma\mu e^{-\mu(t-v)}\delta e^{-\beta t}dvdt\right)$$

$$= \underline{E}.\left(\sum_{i=1}^{\infty}\xi\left(\frac{\lambda}{\mu}\right)^i\left(\frac{\gamma\mu}{\delta\beta+\gamma\mu}\right)+\frac{\xi\delta\lambda\gamma\mu}{(\lambda+\beta)(\mu+\beta)}\right).$$

Corollary 4. The expected number of production up-crossings that go above s in a production cycle, $P_{up} = \overline{E} * $ *Probability of a unit produced before a service completion.* That is,

$$P_{up} = \overline{E}.\left(\begin{array}{l}\sum\limits_{i=1}^{\infty}\xi\left(\frac{\lambda}{\mu}\right)^i\left(\frac{\delta\beta}{\delta\beta+\gamma\mu}\right)+\left(\frac{\delta\beta}{\delta\beta+\lambda}\right)\xi \\ +\xi\int\limits_{t=0}^{\infty}\int\limits_{v=0}^{t}\lambda e^{-\lambda v}e^{-\mu(t-v)}\delta(1-e^{-\beta t}-e^{-\beta v})dvdt\end{array}\right)$$

$$= \overline{E}.\left(\sum_{i=1}^{\infty}\xi\left(\frac{\lambda}{\mu}\right)^i\left(\frac{\delta\beta}{\delta\beta+\gamma\mu}\right)+\left(\frac{\delta\beta}{\delta\beta+\lambda}\right)\xi+\frac{\xi\delta}{(\lambda+\beta)}\left[\frac{\beta(\mu+\beta)+\lambda\mu}{\mu(\mu+\beta)}\right]\right).$$

Having obtained the expected length of a production cycle we turn to compute the optimal pair (s, S) values and the corresponding minimum costs. Lemma 3.4.1 provides us the rate at which the production process is switched on in unit time.

3.5 Computing optimal (s, S) pairs and the minimum cost

We look for the optimal values of s (the level, reaching at which the production process is switched on) and the maximum inventory level S of the production inventory model under discussion. Now for checking the optimality of s and S, the following cost function is constructed. Define $\mathcal{F}(s, S)$ as the expected cost per unit time in the long run. Then

$$\mathcal{F}(s, S) = K.E_{on} + h_{inv}.E_{inv} + c_1.C_{loss} + c_2.M_{loss} + c_3.E_{rp} + c_4.W_s + c_5.\widetilde{W_s}$$

where K is the fixed cost for starting a production run, h_{inv} is the cost per unit time per inventory towards holding, c_1 is the cost incurred due to loss per customer when the inventory is out of stock, c_2 is the cost incurred due to rejection per unit manufactured item, c_3 is the cost of production per unit time, c_4 is the waiting cost per unit time per customer during the stock out period and c_5 is the waiting cost per unit time per customer when inventory is available. Though we are not able to compute explicitly the optimal values of s and S, due to the highly complex form of the cost function, we arrive at these using numerical techniques.

For the following input values $\lambda = 2, \mu = 3, \beta = 2.5, K = \$5000, h_{inv} = \$20, c_1 = \$400, c_2 = \$100, c_3 = \$200, c_4 = \$300, c_5 = \$100$ and varying δ and γ we arrive at Table 3.1. δ and γ are given values from 0.1 to 1 at 0.1 spacing. Note that the case of $\gamma = \delta = 1$ is what is discussed in Krishnamoorthy and Vishwanath [42]. The pair of values given in each cell of Table 3.1 indicates the optimal (s, S) pair and the value at the bottom of each cell corresponds to the minimum cost (in Dollars). As γ and δ are varied we get distinct optimal pairs of (s, S) and the corresponding minimum cost. We observe that the minimum cost is a decreasing function of δ, or at first decreasing and then starts growing with δ. This can be attributed to the fact that for fixed γ, and for δ increasing, initially the loss of manufactured items get reduced; but subsequently from a point on, the holding cost factor dominates the gain from acceptance of produced item. The optimal (s, S) pair first decreases with δ increasing, comes to a minimum and then starts rising up. Same is the trend shown by the minimum cost values. The explanation for this trend is that with γ increasing, customers are provided the item at the end of their service with increasing probability, so shortage is bound to occur with higher probability. To some extend, increasing δ value can cope with this, since produced items are accepted with higher probability. Nevertheless, increase in δ results in increase in the holding cost. For the given input parameters the "best" among the optimal pair is $(1, 11)$ and the minimum cost is \$461.02 which correspond to $\delta = 1$ and $\gamma = 0.1$.

Now by using the same input values of Table 3.1 and with $s = 5$ and $S = 11$ we provide a comparison of the performance measures for a few (γ, δ) pair values in Table 3.2. For example we observe from Table 3.2 that the production cycle length and loss rate of customers are largest for the (γ, δ) pair values $(1, 0.5)$ and least for $(0.5, 1)$ among the three pairs of values indicated in that table. Similarly expected inventory held is least for (γ, δ) pair value $(1, 0.5)$ and the highest for $(0.5, 1)$.

3.5.1 Emptiness time distribution for $M/M/1/1$ production inventory system

We now compute the distribution of the time till the items in the inventory becomes empty (zero) starting from the epoch at which the production is switched on reaching level S. Let χ represent this random variable. Since

Table 3.1: Optimal **(s, S)** values and minimum cost

δ \ γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	(3,11) 605.4	(1,26) 958.33	(1,12) 1189.3	(1,9) 1309.1	(1,8) 1381.7	(1,7) 1430.3	(1,7) 1465.1	(1,7) 1491.2	(1,6) 1511.6	(1,6) 1527.9
0.2	(1,10) 515.24	(2,13) 649.96	(6,20) 793.76	(1,27) 983.33	(1,15) 1120	(1,13) 1214.3	(1,13) 1282.5	(1,10) 1334.1	(1,9) 1374.4	(1,9) 1406.7
0.3	(1,10) 490.34	(1,12) 610.1	(2,14) 689.76	(4,18) 765.15	(7,25) 804.83	(1,23) 1008.3	(1,16) 1105.2	(1,19) 1180.1	(1,13) 1239.3	((1,12) 1287
0.5	(1,10) 472.89	(1,13) 584.32	(1,15) 664.66	(1,15) 722.9	(1,16) 763.58	2,18) 795.24	(4,21) 838.8	(6,26) 908.47	(1,29) 987.12	(1,24) 1058.3
0.6	(1,10) 468.89	(1,13) 578.74	(1,16) 660.13	(1,16) 721.23	(1,16) 766.65	(1,17) 797.98	(2,18) 821.01	(3,20) 849.05	(5,24) 896.26	(4,29) 959.93
0.7	(1,11) 466.11	(1,14) 574.69	(1,16) 656.36	(1,17) 720.28	(1,17) 769.82	(1,17) 806.51	(1,18) 831.65	(2,18) 849.35	(2,20) 867.81	(4,23) 899.16
0.9	(1,11) 462.32	(1,14) 569.49	(1,16) 651.76	(1,18) 732.47	(1,18) 773.71	(1,19) 818.36	(1,19) 853.53	(1,19) 879.47	(1,19) 896.85	(1,19) 907.9
1	**(1,11)** **461.02**	(1,14) 567.74	(1,16) 650.26	(1,18) 717.95	(1,19) 774.64	(1,20) 822.1	(1,20) 860.79	(1,20) 891.35	(1,20) 913.86	(1,20) 928.76

it is impossible to compute the distribution of χ for the case where the system capacity is unlimited, we specialize to the case of $M/M/1/1$ production inventory with positive service time. This will enable us to deal with a finite state space CTMC with $3S-s$ elements and having state space

$\Im = \{(0,0,1)\,,(1,0,1)\,,\ldots,(s,0,1)\,,(s,1,1)\,,(s+1,0,0)\,,(s+1,0,1)\,,\ldots,$
$(S-1,0,0)\,,(S-1,0,1),(S-1,1,0),(S-1,1,1),(S,0,0),\,(S,1,0)\}\,,$

The state $((0,0,1))$ is regarded as absorbing, state which represents the state of the inventory level becomes zero from the state $\{(1,1,1)\}$. The possible transitions and the corresponding rates are given in Table 3.3.

Thus the infinitesimal generator $\boldsymbol{\mathcal{Q}}$ of the Markov chain $\{(\mathcal{I}(t),\mathcal{C}(t),\mathcal{P}(t))|t\geq 0\}$ is of the form $\boldsymbol{\mathcal{Q}} = \begin{bmatrix} \boldsymbol{\mathcal{T}} & \boldsymbol{\mathcal{T}^0} \\ \mathbf{0} & 0 \end{bmatrix}$ with initial probability vector $\boldsymbol{\alpha} = (0,0,\ldots,1,0)$ where 1 is in the $(3S-s-2)^{th}$ position; $\boldsymbol{\mathcal{T}}$ is of order $3S-s-1$; $\boldsymbol{\mathcal{T}^0}$ is a $3S-s$ component column vector such that $\boldsymbol{\mathcal{T}}\mathbf{e}+\boldsymbol{\mathcal{T}^0}$=0. Let χ represent the random variable "time till the items in the inventory becomes zero". This time duration follows PH distribution with representation $(\boldsymbol{\alpha},\boldsymbol{\mathcal{T}})$. There-

Table 3.2: Effect of γ and δ on various performance measures

Performance measures	$\gamma = 1$ and $\delta = 0.5$	$\gamma = 0.5$ and $\delta = 1$	$\gamma = \delta = 1$
L_s	0.00085731	0.10005	0.038268
W_s	0.75643	0.0013604	0.07402
$\widetilde{W_s}$	1.2436	1.9986	1.926
E_{inv}	1.5852	7.8376	5.9064
E_{rp}	1.2436	0.99932	1.926
E_{cycle}	580.22	3.9955	10.066
C_{loss}	0.75643	0.0013604	0.07402

fore the expected time until the inventory become zero is,

$$E(\chi) = -\boldsymbol{\alpha}\left(\boldsymbol{\mathcal{T}^{-1}}\right)\mathbf{e}.$$

Table 3.3: The transitions in the CTMC $\{(\mathcal{I}(t), \mathcal{C}(t), \mathcal{P}(t)) | t \geq 0\}$ and corresponding rates

Form	To	Rate	
$(\ell, 0, 1)$	$(\ell, 1, 1)$	λ	$\ell = 1, 2, \ldots, s, s+1, \ldots, S-1.$
$(\ell, 0, 1)$	$(\ell+1, 0, 1)$	β	$\ell = 1, 2, \ldots, s, s+1, \ldots, S-1.$
$(\ell, 0, 1)$	$(\ell, 0, 1)$	$-(\lambda+\beta)$	$\ell = 1, 2, \ldots, s, s+1, \ldots, S-1.$
$(\ell, 1, 1)$	$(\ell-1, 0, 1)$	μ	$\ell = 2, 3, \ldots, s.$
$(\ell, 1, 1)$	$(\ell+1, 1, 1)$	β	$\ell = 1, 2, \ldots, s.$
$(\ell, 1, 1)$	$(\ell, 1, 1)$	$-(\beta+\mu)$	$\ell = 1, 2, \ldots, s.$
$(\ell, 0, 0)$	$(\ell, 1, 0)$	λ	$\ell = s+1, s+2, \ldots, S-1.$
$(\ell, 0, 0)$	$(\ell, 0, 0)$	$-\lambda$	$\ell = s+1, s+2, \ldots, S-1.$
$(\ell, 1, 0)$	$(\ell-1, 0, 0)$	μ	$\ell = s+1, s+2, \ldots, S-1.$
$(\ell, 1, 0)$	$(\ell, 1, 0)$	$-\mu$	$\ell = s+1, s+2, \ldots, S-1.$
$(\ell, 1, 1)$	$(\ell-1, 1, 1)$	μ	$\ell = s+1, s+2, \ldots, S-1.$
$(\ell, 1, 1)$	$(\ell+1, 1, 1)$	β	$\ell = s+1, s+2, \ldots, S-1.$
$(\ell, 1, 1)$	$(\ell, 1, 1)$	$-(\mu+\beta)$	$\ell = s+1, s+2, \ldots, S-1.$
$(S, 0, 0)$	$(S, 1, 0)$	λ	
$(S, 0, 0)$	$(S, 0, 0)$	$-\lambda$	
$(S, 1, 0)$	$(S-1, 1, 0)$	μ	
$(S, 1, 0)$	$(S, 1, 0)$	$-\mu$	
$(1, 1, 1)$	$\{*\}$	μ	

Chapter 4

Multi-server queueing-inventory system

4.1 Introduction

In chapters 2 and 3 we discussed single server queues with inventory as service item. Either bulk replenishment policy (chapter 2) or replenishment through production (chapter 3) was adopted and the optimal values of decision variables computed.

In this chapter we attempt to derive the steady-state distribution of the $M/M/c$ queueing-inventory system with positive service time. First we analyze the case of $c = 2$ servers which are assumed to be homogeneous and that the service time follows exponential distribution. The inventory replenishment follows the (s, Q) policy. We obtain a product form solution of the steady-state distribution under the assumption that customers do not

Some results of this chapter are included in the following paper.
A. Krishnamoorthy, R. Manikandan and Dhanya Shajin : Analysis of a multi-server queueing-inventory system (Under review).

join the system when the inventory level is zero. An optimization problem is also investigated to get the optimal pair (s, Q) and the corresponding expected minimum cost is obtained. As in the case of $M/M/c$ retrial queue with $c \geq 3$, we conjuncture that $M/M/c$, for $c \geq 3$, queueing-inventory problems do not have analytical solution. So we proceed to analyze those cases by using algorithmic approach. Assume that $c < s$. All servers are assumed to be homogeneous and that the service time follows exponential distribution. Here also the inventory replenishment follows (s, Q) policy. We derive an explicit expression for the stability condition of the system. We discuss the conditional distribution of the inventory level, conditioned on the number of customers in the system and conditional distribution of the number of customers conditioned on the inventory level. Also we compute the distribution of two consecutive s to s transitions of the inventory level (that is the first return time to s). Closed form solution for the system state distribution cannot be arrived so the steady-state distribution of this system is difficult to obtain as a product form. So by using algorithmic method we compute the stationary probability distribution. We also obtain several system performance measures.

This chapter organized as follows. In Section 4.2 the $M/M/2$ queueing-inventory problem is mathematically formulated. The product form solution of the steady-state probability distribution, including some important performance measures are obtained in Section 4.3. Further we provide the optimal pair (s, Q) values and the minimal cost for different values of γ. Section 4.5 discuss the $M/M/c$ with c (greater than or equal to 3 but less than s) queueing-inventory problems by using algorithmic approach. Section 4.6 gives some conditional probability distributions and few performance measures. Section 4.7 analyzes the distribution of the inventory cycle time. In Section 4.8 provides the optimal c and the corresponding minimal cost for different values of γ. Further we look for the optimal pair (s, Q) values that would result in cost minimization for different pairs of values of γ and c.

4.2 Mathematical modelling of the $M/M/2$ queueing-inventory problem

First we consider an $M/M/2$ queueing-inventory system with positive service time. Arrival process is assumed to be Poisson with rate λ. Each customer requires a single item having random duration of service which follows exponential distribution with parameter μ. However, it is not essential that inventory is provided to the customer at the end his service. More precisely, the item is served with probability γ at the end of a service and is not provided with probability $1-\gamma$. A crucial assumption of this model, as done in the previous two chapters, is that customers do not join the system when the inventory level is zero. When the number of customers is at least two and not less than two items are in inventory, the service rate is 2μ. When the on-hand inventory reaches a pre-specified value $s > 0$, a replenishment order is placed for $Q(< \infty)$ units with $Q > s$. We fix $S = Q + s$ as the maximum number of items that could be held in the system at any given time. The lead time follows exponential distribution with parameter β. Then $\{\mathcal{X}(t)|t \geq 0\} = \{(\mathcal{N}(t), \mathcal{I}(t))|t \geq 0\}$ is a CTMC with state space $\Omega_1 = \bigcup_{i=0}^{\infty} \mathcal{L}(i)$, where $\mathcal{L}(i)$ is called the i^{th} level (number of customers in the system is $i(\geq 0)$). In each of the level the number of items in the inventory can be anything from 0 to S. Accordingly we write $\mathcal{L}(i) = \{(i, 0), \ldots, (i, Q + s)\}$. In these, the second coordinate is referred to as the phase of the system. The infinitesimal generator $\boldsymbol{\mathcal{W}_1}$ of this CTMC $\{\mathcal{X}(t)|t \geq 0\}$ is

$$\mathcal{W}_1 = \begin{bmatrix} B_{00} & A_0 & & & & \\ B_{20} & B_{10} & A_0 & & & \\ & A_2 & A_1 & A_0 & & \\ & & A_2 & A_1 & A_0 & \dots \\ & & & \ddots & \ddots & \ddots \end{bmatrix},$$

where

$$[B_{00}]_{kl} = \begin{cases} -\beta, & \text{for } l = k = 0. \\ -(\lambda+\beta), & \text{for } l = k;\ k = 1, 2, ..., s. \\ -\lambda, & \text{for } l = k;\ k = s+1, s+2, ..., S. \\ \beta, & \text{for } l = k+Q;\ k = 0, 1, ..., s. \\ 0, & \text{otherwise.} \end{cases}$$

$$[B_{20}]_{kl} = \begin{cases} \gamma\mu, & \text{for } l = k-1;\ k = 1, 2, ..., S. \\ (1-\gamma)\mu, & \text{for } l = k;\ k = 1, 2, ..., S. \\ 0, & \text{otherwise.} \end{cases}$$

$$[B_{10}]_{kl} = \begin{cases} -\beta, & \text{for } l = k = 0. \\ -(\lambda+\beta+\mu), & \text{for } l = k;\ k = 1, 2, ..., s. \\ -(\lambda+\mu), & \text{for } l = k;\ k = s+1, s+2, ..., S. \\ \beta, & \text{for } l = k+Q;\ k = 0, 1, ..., s. \\ 0, & \text{otherwise.} \end{cases}$$

$$[A_0]_{kl} = \begin{cases} \lambda, & \text{for } l = k;\ k = 1, 2, ..., S. \\ 0, & \text{otherwise.} \end{cases}$$

$$[A_1]_{kl} = \begin{cases} -\beta, & \text{for } l = k = 0. \\ -(\lambda+\beta+\mu), & \text{for } l = k = 1. \\ -(\lambda+\beta+2\mu), & \text{for } l = k;\ k = 2,3,...,s. \\ -(\lambda+2\mu), & \text{for } l = k;\ k = s+1, s+2, ..., S. \\ \beta, & \text{for } l = k+Q;\ k = 0,1,...,s. \\ 0, & \text{otherwise.} \end{cases}$$

$$[A_2]_{kl} = \begin{cases} \gamma\mu, & \text{for } l = k-1;\ k = 1. \\ (1-\gamma)\mu, & \text{for } l = k = 1. \\ 2\gamma\mu, & \text{for } l = k-1;\ k = 2,3,...,S. \\ 2(1-\gamma)\mu, & \text{for } l = k;\ k = 2,3,...,S. \\ 0, & \text{otherwise.} \end{cases}$$

Note that all entries (block matrices) of $\boldsymbol{\mathcal{W}_1}$ are of the same order, namely $S+1$.

4.2.1 Analysis of the system

In this section we perform the steady-state analysis of the queueing-inventory model under study by first establishing the stability condition of the queueing-inventory system. Define $A = A_0 + A_1 + A_2$. This is the infinitesimal generator matrix of the finite state CTMC corresponding to the inventory level $\{0, 1, 2, ..., S\}$ for any level i (≥ 1). Let ζ denote the steady-state probability vector of A. That is,

$$\boldsymbol{\zeta} A = 0,\ \boldsymbol{\zeta}\mathbf{e} = 1. \tag{4.1}$$

Write

$$\boldsymbol{\zeta} = (\zeta_0, \zeta_1, ..., \zeta_s, ..., \zeta_Q, ..., \zeta_S)$$

and

$$A = \begin{bmatrix} -\beta & & & & \beta & & \\ \gamma\mu & -(\beta+\gamma\mu) & & & & & \\ & 2\gamma\mu & -(\beta+2\gamma\mu) & & & \ddots & \\ & \ddots & \ddots & & & & \\ & & 2\gamma\mu & -(\beta+2\gamma\mu) & & & \beta \\ & & & 2\gamma\mu & -2\gamma\mu & & \\ & & & & \ddots & \ddots & \\ & & & & & 2\gamma\mu & -2\gamma\mu \end{bmatrix},$$

Then using (4.1) we get the components of the vector $\boldsymbol{\zeta}$ explicitly as

$$\zeta_0 = \left\{ \begin{array}{l} 1 + \frac{\beta}{\gamma\mu}\left[1 + \left(\frac{\beta+\gamma\mu}{\gamma\mu}\right)\sum_{i=0}^{s}\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{i-2} + (Q-s-2)\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{s-1}\right] \\ +\frac{\beta}{2\gamma\mu}\left(\frac{\beta+\gamma\mu}{\gamma\mu}\right)\left[\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{s-1} - \left(\frac{\gamma\mu}{\beta+\gamma\mu}\right) + \sum_{i=0}^{s}\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{i-2}\left(\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{s-i+1} - 1\right)\right] \end{array} \right\}^{-1}$$

$$\zeta_i = \begin{cases} \frac{\beta}{\gamma\mu}\zeta_0, & \text{for } i = 1. \\ \frac{\beta}{\gamma\mu}(\frac{\beta+\gamma\mu}{2\gamma\mu})\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{i-2}\zeta_0, & \text{for } i = 2,3,...,s+1. \\ \zeta_{i+1}, & \text{for } i = s+1, s+2, ..., Q-1. \\ \frac{\beta}{2\gamma\mu}\left[(\frac{\beta+\gamma\mu}{\gamma\mu})\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{s-1} - 1\right]\zeta_0, & \text{for } i = Q+1. \end{cases}$$

and $\zeta_{Q+i} = \frac{\beta}{2\gamma\mu}\left(\frac{\beta+\gamma\mu}{\gamma\mu}\right)\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{i-2}\left[\left(\frac{\beta+2\gamma\mu}{2\gamma\mu}\right)^{s-(i-1)} - 1\right]\zeta_0,\ i = 2,3,...,s.$

Since the Markov Chain $\{\mathcal{X}(t)|t \geq 0\}$ is an LIQBD, it is stable if and only if the left drift rate exceeds the right drift rate. That is,

$$\zeta A_0\mathbf{e} < \zeta A_2\mathbf{e}.$$

Thus, we have the following lemma for stability of the system under study.

Lemma 4.2.1. The stability condition of the $M/M/2$ queueing-inventory system under consideration is given by $\lambda < \mu \left[2 - \frac{\beta\zeta_0}{\gamma\mu(1-\zeta_0)}\right]$.

Proof. From the well known result in Neuts [47] on the positive recurrence of A, we have $\boldsymbol{\zeta} A_0 \mathbf{e} < \boldsymbol{\zeta} A_2 \mathbf{e}$ for the Markov chain to be stable. With a bit of computation, this simplifies to the result $\lambda < \mu \left[2 - \frac{\beta\zeta_0}{\gamma\mu(1-\zeta_0)}\right]$. □

For future reference we define ρ_1 as

$$\rho_1 = \frac{\lambda}{\mu \left[2 - \frac{\beta\zeta_0}{\gamma\mu(1-\zeta_0)}\right]}. \tag{4.2}$$

4.3 Computation of the steady-state probability

For computing the steady-state probability vector of the process $\{\mathcal{X}(t)|t \geq 0\}$, we first consider a queueing-inventory system with unlimited supply of inventory items (that is classical $M/M/2$ queueing system). The rest of the assumptions such as those on the arrival process and lead time are the same as given earlier. Designate the Markov chain so obtained as $\{\mathcal{N}(t)|t \geq 0\}$, where $\mathcal{N}(t)$ is the number of customers in the system at time t. Its infinitesimal generator $\boldsymbol{\mathcal{G}_1}$ is given by,

$$\boldsymbol{\mathcal{G}_1} = \begin{bmatrix} -\lambda & \lambda & & & & \\ \mu & -(\lambda+\mu) & \lambda & & & \\ & 2\mu & -(\lambda+2\mu) & \lambda & & \\ & & 2\mu & -(\lambda+2\mu) & \lambda & \dots \\ & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Let $\boldsymbol{\pi}$ be the steady-state probability vector of $\boldsymbol{\mathcal{G}_1}$. Partitioning $\boldsymbol{\pi}$ by levels we write $\boldsymbol{\pi}$ as

$$\boldsymbol{\pi} = (\boldsymbol{\pi_0}, \boldsymbol{\pi_1}, \boldsymbol{\pi_2}, \dots). \tag{4.3}$$

Then the steady-state vector must satisfy

$$\boldsymbol{\pi}\boldsymbol{\mathcal{G}_1} = 0, \ \boldsymbol{\pi}\mathrm{e} = 1. \tag{4.4}$$

From the relation (4.4) we get the vector $\boldsymbol{\pi}$ explicitly as follows

$$\pi_i = \begin{cases} \left[1 + \frac{\lambda}{\mu}\left(1 - \frac{\lambda}{2\mu}\right)^{-1}\right]^{-1} & \text{for } i = 0. \\ \frac{\lambda}{\mu}\pi_0 & \text{for } i = 1. \\ \frac{1}{2^{i-1}}\left(\frac{\lambda}{\mu}\right)^i \pi_0 & \text{for } i \geq 2. \end{cases} \tag{4.5}$$

Further we consider an inventory system with negligible service time and no backlog of demands. The assumptions such as those on the arrival process and lead time are the same as given in the description of the model. Denote this Markov chain as $\{\mathcal{I}(t) | t \geq 0\}$. Here $\mathcal{I}(t)$ is the inventory level at time t. Its infinitesimal generator $\boldsymbol{\mathcal{G}_2}$ is given by,

$$\boldsymbol{\mathcal{G}_2} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ s \\ \\ \vdots \\ Q \\ \vdots \\ S \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & \dots & s\dots & Q & \dots & & S \end{array} \\ \left(\begin{array}{cccccccc} -\beta & & & & \beta & & & \\ \gamma\lambda & -(\gamma\lambda+\beta) & & & & & & \\ & \ddots & \ddots & & & & \ddots & \\ & & \gamma\lambda & -(\gamma\lambda+\beta) & & & & \beta \\ & & & \gamma\lambda & -\gamma\lambda & & & \\ & & & & & & & \\ & & & & \ddots & \ddots & & \\ & & & & & \gamma\lambda & -\gamma\lambda & \\ & & & & & & \gamma\lambda & -\gamma\lambda \end{array}\right) \end{array}.$$

Let $\boldsymbol{\psi}$=$(\psi_0, \psi_1, \ldots, \psi_S)$ be the steady-state probability vector of the process $\{\mathcal{I}(t)|t \geq 0\}$. Then $\boldsymbol{\psi}$ satisfies the relations

$$\boldsymbol{\psi}\boldsymbol{\mathcal{G}_2} = 0, \ \boldsymbol{\psi}\mathbf{e} = 1 \tag{4.6}$$

That is, at arbitrary epochs the inventory level distribution ψ_j is given by

$$\psi_j = \begin{cases} \left[1 + Q\frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^s\right]^{-1}, \ j = 0. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{j-1}\psi_0, \ j = 1, 2, \cdots, s. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s}\psi_0, \ j = s+1, s+2, \cdots, Q. \\ \frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{j-Q-1}\left(\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^{s-(j-Q-1)} - 1\right)\psi_0, \ j = Q+1, Q+2, \cdots, S. \end{cases} \tag{4.7}$$

Using the components of the probability vector $\boldsymbol{\psi}$, we shall find the steady-state probability vector of the original system. Let $\boldsymbol{x}$ be the steady-state probability vector of the original system. Then the steady-state vector must satisfy the set of equations

$$\boldsymbol{x}\boldsymbol{\mathcal{W}_1} = 0, \ \boldsymbol{x}\mathbf{e} = 1. \tag{4.8}$$

Partition $\boldsymbol{x}$ by levels as

$$\boldsymbol{x} = (\boldsymbol{x_0}, \boldsymbol{x_1}, \boldsymbol{x_2}, \ldots). \tag{4.9}$$

where the subvectors of $\boldsymbol{x}$ are further partitioned as

$$\boldsymbol{x_i} = (x_i(0), x_i(1), x_i(2), x_i(3), ..., x_i(S)), \ i \geq 0. \tag{4.10}$$

Then by using the relation $\boldsymbol{x}\boldsymbol{\mathcal{W}_1} = 0$, we get

$$-\beta x_i(0) + \gamma\mu x_{i+1}(1) = 0, \ i \geq 0. \tag{4.11}$$

$$\lambda x_i(j) - (\lambda + 2\mu + \beta)x_{i+1}(j) + 2(1-\gamma)\mu x_{i+2}(j) + 2\gamma\mu x_{i+2}(j+1) = 0,$$
$$i \geq 1,\ 2 \leq j \leq Q-1. \tag{4.12}$$

$$\lambda x_i(j) + \beta x_{i+1}(j-Q) - (\lambda + 2\mu)x_{i+1}(j) + 2(1-\gamma)\mu x_{i+2}(j) + 2\gamma\mu x_{i+2}(j+1) = 0,$$
$$i \geq 1, Q \leq j \leq S-1. \tag{4.13}$$

$$\lambda x_i(S) + \beta x_{i+1}(s) - (\lambda + 2\mu)x_{i+1}(S) + 2(1-\gamma)\mu x_{i+2}(S) = 0,\ i \geq 1. \tag{4.14}$$

$$-(\lambda + \beta)x_0(j) + (1-\gamma)\mu x_1(j) + \gamma\mu x_1(j+1) = 0,\ 1 \leq j \leq s. \tag{4.15}$$

$$-\lambda x_0(j) + (1-\gamma)\mu x_1(j) + \gamma\mu x_1((j+1) = 0,\ s+1 \leq j \leq Q-1. \tag{4.16}$$

$$\beta x_0(j-Q) - \lambda x_0(j) + (1-\gamma)\mu x_1(j) + \gamma\mu x_1(j+1) = 0,\ Q \leq j \leq S-1. \tag{4.17}$$

$$\beta x_0(s) - \lambda x_0(S) + (1-\gamma)\mu x_1(S) = 0. \tag{4.18}$$

$$\lambda x_0(j) - (\lambda + \beta + \mu)x_1(j) + 2(1-\gamma)\mu x_2(j) + 2\gamma\mu x_2(j+1) = 0,\ 2 \leq j \leq s. \tag{4.19}$$

$$\lambda x_0(j) - (\lambda + \mu)x_1(j) + 2(1-\gamma)\mu x_2(j) + 2\gamma\mu x_2(j+1) = 0,\ s+1 \leq j \leq Q-1. \tag{4.20}$$

$$\lambda x_0(j) + \beta x_1(j-Q) - (\lambda + \mu)x_1(j) + 2(1-\gamma)\mu x_2(j) + 2\gamma\mu x_2(j+1) = 0,$$
$$Q \leq j \leq S-1. \tag{4.21}$$

$$\lambda x_0(S) + \beta x_1(s) - (\lambda + +\mu)x_1(S) + 2(1-\gamma)\mu x_2(S) = 0. \tag{4.22}$$

Now let

$$x_i(j) = \Theta_j^i \pi_i \psi_j,\ i \geq 0,\ 0 \leq j \leq S, \tag{4.23}$$

The constants Θ_j^i's are given by

$$\Theta_0^i = 1,\ i \geq 0. \tag{4.24}$$

$$\Theta_1^i = \begin{cases} \frac{1}{\gamma}, \ i = 1. \\ \frac{2}{\gamma}, \ i \geq 2. \end{cases} \tag{4.25}$$

$$\Theta_j^0 = \left(\frac{1}{\gamma}\right)^j, \ 1 \leq j \leq S-1. \tag{4.26}$$

$$\Theta_2^i = \begin{cases} \left(\frac{\beta+\gamma\lambda}{\beta+\lambda}\right)\frac{1}{\gamma^2}, i = 1, 2. \\ \left(\frac{2\beta+(1+\gamma)\lambda}{\beta+\lambda}\right)\frac{1}{\gamma^2}, i \geq 3. \end{cases} \tag{4.27}$$

$$\Theta_j^i = \begin{cases} \left(\frac{1}{\gamma(\beta+\lambda)}\right)\delta_{j-1}^i, \ 3 \leq i \leq 2(j-1), \ 3 \leq j \leq s+1. \\ \left(\frac{(\beta+\gamma\lambda)}{\gamma(\beta+\lambda)}\right)\Theta_{j-1}^{i-2}, \ i \geq 2j-1, \ 3 \leq j \leq s+1. \\ \left(\frac{1}{\gamma\lambda}\right)\delta_{j-1}^i, \ 3 \leq i \leq 2(j-1), \ s+2 \leq j \leq Q. \\ \left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)\Theta_{j-1}^{i-2}, \ i \geq 2j-1, \ s+2 \leq j \leq Q. \end{cases} \tag{4.28}$$

where $\delta_{j-1}^i = (\lambda + 2\mu + \beta)\Theta_{j-1}^{i-1} - 2\mu\Theta_{j-1}^{i-2} - (1-\gamma)\lambda\Theta_{j-1}^i$.

$$\Theta_{Q+k}^i = \begin{cases} \frac{1}{\gamma\lambda}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^s - 1\right]^{-1}\left[\xi_{Q+k-1}^i\left(\frac{\beta+\lambda}{\lambda}\right)^s - \lambda\right], \ 3 \leq i \leq 2Q, \ k = 1. \\ \frac{1}{\gamma\lambda}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^s - 1\right]^{-1}\left[\xi_{Q+k-1}^{i-2}\gamma\lambda\left(\frac{\beta+\lambda}{\lambda}\right)^s - \lambda\right], \ i \geq 2Q+1, \ k = 1. \\ \frac{1}{\gamma\lambda(\beta+\lambda)}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^{s-(k-1)} - 1\right]^{-1}\left[\xi_{Q+k-1}^i\left[\left(\frac{\beta+\lambda}{\lambda}\right)^{s-(k-2)} - 1\right] - \beta\Theta_{k-1}^{i-1}\right], \\ \qquad 3 \leq i \leq 2(Q+k-1), \ 2 \leq k \leq s. \\ \frac{1}{\gamma\lambda(\beta+\lambda)}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^{s-(k-1)} - 1\right]^{-1}\left[\gamma\lambda\left[\left(\frac{\beta+\lambda}{\lambda}\right)^{s-(k-2)} - 1\right]\Theta_{Q+k-1}^{i-2} - \beta\Theta_{k-1}^{i-1}\right], \\ \qquad i \geq 2(Q+k)-1, \ 2 \leq k \leq s. \end{cases} \tag{4.29}$$

where, $\xi_{Q+k-1}^i = (\lambda + 2\mu)\Theta_{Q+k-1}^{i-1} - 2\mu\Theta_{Q+k-1}^{i-2} - (1-\gamma)\lambda\Theta_{Q+k-1}^i$.

$$\Theta_j^1 = \begin{cases} \frac{1}{\gamma(\beta+\lambda)}\left[(\lambda+\beta)\Theta_{j-1}^0 - (1-\gamma)\lambda\Theta_{j-1}^1\right], \ 3 \leq j \leq s+1. \\ \frac{1}{\gamma}\left[\Theta_{j-1}^0 - (1-\gamma)\Theta_{j-1}^1\right], \ s+2 \leq j \leq Q. \end{cases} \tag{4.30}$$

$$\Theta_S^0 = \left[\Theta_s^0 - (1-\gamma)\Theta_S^1\right]. \tag{4.31}$$

$$\Theta_j^2 = \begin{cases} \frac{1}{\gamma(\beta+\lambda)}\left[(\lambda+\beta+\mu)\Theta_{j-1}^1 - \mu\Theta_{j-1}^0 - (1-\gamma)\lambda\Theta_{j-1}^2\right], \ 3 \leq j \leq s+1. \\ \frac{1}{\gamma\lambda}\vartheta_{j-1}, \ s+2 \leq j \leq Q. \end{cases} \tag{4.32}$$

$$\Theta_{Q+k}^2 = \begin{cases} \frac{1}{\gamma\lambda}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^s - 1\right]^{-1}\left[\vartheta_Q\left(\frac{\beta+\mu}{\mu}\right)^s - \beta\right], \ k=1. \\ \frac{1}{\gamma(\beta+\lambda)}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^{s-(k-1)} - 1\right]^{-1}\left[\vartheta_{j-1}\left[\left(\frac{\beta+\lambda}{\lambda}\right)^{s-(k-2)} - 1\right] - \beta\Theta_{k-1}^1\right], \\ 2 \leq k \leq s. \end{cases} \tag{4.33}$$

where, $\vartheta_j = (\lambda+\mu)\Theta_j^1 - \mu\Theta_j^0 - (1-\gamma)\lambda\Theta_j^2, \ s-1 \leq j \leq S.$
Now we require $\boldsymbol{xe}$=1. That is,

$$\sum_{i=0}^{\infty}\sum_{j=0}^{Q+s}\Theta_j^i\pi_i\psi_j = 1 + Q\frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^s.$$

Let $\alpha = 1 + Q\frac{\beta}{\gamma\lambda}\left(\frac{\beta+\gamma\lambda}{\gamma\lambda}\right)^s$. So dividing each sub-vector of $\boldsymbol{x}$ by α we get the steady-state probability distribution vector of the original system.

Thus we arrive at our main theorem:

Theorem 4.3.1. *Under the necessary and sufficient condition* $\rho_1 < 1$ *for stability, the components of the steady-state probability vector of the process* $\{\mathcal{X}(t)|t \geq 0\}$ *with generator matrix* $\boldsymbol{\mathcal{W}_1}$ *is* $x_i(j) = \alpha^{-1}\Theta_j^i\pi_i\psi_j, \ i \geq 0; \ 0 \leq j \leq S$ *where* ρ_1 *is as defined in* (4.2), *the probabilities* π_i *corresponds to the distribution of number of customer in the system as given in* (4.5) *and the probabilities* ψ_j *are obtained* (4.7).

The consequence of Theorem 4.3.1 is that the two dimensional system can be decomposed into two distinct one dimensional objects one of which correspond to number of customers in an $M/M/2$ queue and the other to the number of items in the inventory.

4.3.1 Performance measures

- Mean number of customers in the system,
  $L_s = \alpha^{-1}\left(\sum_{i=1}^{\infty}\sum_{j=0}^{Q+s} i\Theta_j^i \pi_i \psi_j\right).$

- Mean number of customers in the queue,
  $L_q = \alpha^{-1}\left(\sum_{i=2}^{\infty}\sum_{j=2}^{Q+s} (i-2)\Theta_j^i \pi_i \psi_j\right).$

- Mean inventory level in the system, $I_m = \alpha^{-1}\left(\sum_{i=0}^{\infty}\sum_{j=1}^{Q+s} j\Theta_j^i \pi_i \psi_j\right).$

- Mean number of busy server,

$$P_{BS} = \alpha^{-1}\left(\begin{array}{c} \left[\sum_{i=2}^{\infty}\Theta_1^i \pi_i \psi_1 + \sum_{j=2}^{Q+s}\Theta_j^1 \pi_1 \psi_j + \Theta_1^1 \pi_1 \psi_1\right] \\ +2\left[\sum_{i=3}^{\infty}\Theta_2^i \pi_i \psi_2 + \sum_{j=3}^{Q+s}\Theta_j^2 \pi_2 \psi_j + \Theta_2^2 \pi_2 \psi_2\right] \end{array}\right).$$

- Depletion rate of inventory, $D_{inv} = \gamma\lambda\alpha^{-1}\left(\sum_{i=0}^{\infty}\sum_{j=1}^{Q+s} \Theta_j^i \pi_i \psi_j\right).$

- Mean number of replenishments per time unit,
  $R_r = \beta\alpha^{-1}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{s} \Theta_j^i \pi_i \psi_j\right).$

- Mean number of departures per unit time,
  $D_m = \mu\alpha^{-1}\left(\sum_{i=1}^{\infty}\Theta_1^i \pi_i \psi_1 + \sum_{j=1}^{Q+s}\Theta_j^1 \pi_1 \psi_j\right) + 2\mu\alpha^{-1}\left(\sum_{i=2}^{\infty}\sum_{j=2}^{Q+s} \Theta_j^i \pi_i \psi_j\right).$

- Expected loss rate of customers, $E_{loss} = \lambda\alpha^{-1}\left(\sum_{i=0}^{\infty}\Theta_0^i \pi_i \psi_0\right).$

- Expected loss rate of customers when the inventory level is zero per cycle, $E^c_{loss} = \frac{E_{loss}}{R_r}$.

- Effective arrival rate,
$\lambda_A = \lambda\alpha^{-1}\left(\sum\limits_{i=0}^{\infty}\sum\limits_{j=1}^{Q+s}\Theta^i_j\pi_i\psi_j\right)$.

- Mean sojourn time of the customers in the system, $W_s = \frac{L_s}{\lambda_A}$.

- Mean waiting time of a customer in the queue, $W_q = \frac{L_q}{\lambda_A}$.

- Mean number of customers waiting in the system when inventory is available, $\widetilde{W} = \alpha^{-1}\left(\sum\limits_{i=0}^{\infty}\sum\limits_{j=1}^{Q+s} i\Theta^i_j\pi_i\psi_j\right)$.

- Mean number of customers waiting in the system during the stock out period, $\widetilde{\widetilde{W}} = \alpha^{-1}\left(\sum\limits_{i=0}^{\infty} i\Theta^i_0\pi_i\psi_0\right)$.

4.4 Optimization problem I

In this section we provide the optimal values of the inventory level s and the fixed order quantity Q of this model. Now for computing the minimal costs of $M/M/2$ queueing-inventory model we introduce the cost function $\mathcal{F}(2, s, Q)$ defined by

$$\mathcal{F}(2, s, Q) = h.I_m + c_1.E_{loss} + c_2.\widetilde{\widetilde{W}} + (K + Q.c_3).R_r + c_4.P_{BS} + c_5.(c - P_{BS})$$

where K is fixed cost for placing an order, c_1 is the cost incurred due to loss per customer, c_2 is waiting cost per unit time per customer during the stock out period, c_3 is variable procurement cost per item, c_4 is the cost incurred per busy server, c_5 is the cost incurred per idle server and h is unit holding cost of inventory unit per unit of time. We assign the following values to the parameters: $\lambda = 5, \mu = 3, \beta = 1, K = \$500, c_1 = \$100, c_2 = \$50, c_3 =$

$\$25, c_4 = \$10, c_5 = \$20, h = \$2$. Thus we obtain Table 4.1 which provide the optimal pairs (s, Q) and also the corresponding minimum cost (in Dollars). Here γ is varied from 0.1 to 1, each time increasing it by 0.1 unit. The optimal pair (s, Q) and the corresponding cost (minimum) are given in Table 4.1.

Table 4.1: Optimal (s, Q) pair and minimum cost

γ	0.1	0.2	0.3	0.4	0.5
Optimal (s, Q) pair & minimum cost	(3,15) 82.684	(3,21) 106.87	(3,27) 130.57	(3,33) 153.76	(3,39) 176.29
γ	0.6	0.7	0.8	0.9	1
Optimal (s, Q) pair & minimum cost	(3,43) 198.10	(5,46) 219.04	(5,53) 239.03	(6,53) 258.26	(6,58) 277.18

4.5 $M/M/c$ $(c \geq 3)$ queueing-inventory system

In this section we consider an $M/M/c$ $(c \geq 3)$ queueing-inventory system with positive service time. We keep the model assumptions the same as in Section 4.2. There are c servers with $3 \leq c < s$. Hence the service rate is $i\mu$, for i varying from 0 to c, depending on the availability of the inventory and customers. When the number of customers is at least c and not less than c items are in the inventory, the service rate is $c\mu$. Write $\{\mathcal{Y}(t)|t \geq 0\} = \{(\mathcal{N}(t), \mathcal{I}(t))|t \geq 0\}$. Then $\{\mathcal{Y}(t)|t \geq 0\}$ is a CTMC with state space $\Omega_2 = \bigcup_{i=0}^{\infty} \mathcal{L}(i)$, where $\mathcal{L}(i)$ is the collection of states $\mathcal{L}(i) = \{(i, 0), \ldots, (i, Q + s)\}$ as defined in Section 4.2. The infinitesimal generator

$\boldsymbol{\mathcal{W}_2}$ of the CTMC $\{\mathcal{Y}(t)|t \geq 0\}$ is

$$\boldsymbol{\mathcal{W}_2} = \begin{bmatrix} B & \bar{A}_0 & & & & & & & & \\ A_2^1 & A_1^1 & \bar{A}_0 & & & & & & & \\ & A_2^2 & A_1^2 & \bar{A}_0 & & & & & & \\ & & \ddots & \ddots & \ddots & & & & & \\ & & & A_2^{c-2} & A_1^{c-2} & \bar{A}_0 & & & & \\ & & & & A_2^{c-1} & A_1^{c-1} & \bar{A}_0 & & & \\ & & & & & \bar{A}_2 & \bar{A}_1 & \bar{A}_0 & & \\ & & & & & & \bar{A}_2 & \bar{A}_1 & \bar{A}_0 & \dots \\ & & & & & & & \ddots & \ddots & \ddots \end{bmatrix},$$

where

$$[B]_{kl} = \begin{cases} -\beta, & \text{for } l = k = 0. \\ -(\lambda+\beta), & \text{for } l = k;\ k = 1, 2, ..., s. \\ -\lambda, & \text{for } l = k;\ k = s+1, s+2, ..., S. \\ \beta, & \text{for } l = k+Q;\ k = 0, 1, ..., s. \\ 0, & \text{otherwise.} \end{cases}$$

$$[\bar{A}_0]_{kl} = \begin{cases} \lambda, & \text{for } l = k;\ k = 1, 2, ..., S. \\ 0, & \text{otherwise.} \end{cases}$$

$$[\bar{A}_1]_{kl} = \begin{cases} -\beta, & \text{for } l = k = 0. \\ -(\lambda+\beta+i\mu), & \text{for } l = k;\ k = 1, 2, ..., c. \\ -(\lambda+\beta+c\mu), & \text{for } l = k;\ k = c+1, c+2, ..., s. \\ -(\lambda+c\mu), & \text{for } l = k,\ k = s+1, s+2, ..., S. \\ \beta, & \text{for } l = k+Q;\ k = 0, 1, ..., s. \\ 0, & \text{otherwise.} \end{cases}$$

$$
[\bar{A}_2]_{kl} = \begin{cases} i\gamma\mu, & \text{for } l = k-1;\ k = c, c+1, c+2, ..., S. \\ i\gamma\mu, & \text{for } l = k-1;\ k = 1, 2, ..., c-1. \\ c(1-\gamma)\mu, & \text{for } l = k;\ k = c, c+1, c+2, ..., S. \\ i(1-\gamma)\mu, & \text{for } l = k;\ k = 1, 2, ..., c-1. \\ 0, & \text{otherwise.} \end{cases}
$$

For $m = 1, 2, ..., c-1$,

$$
[A_2^m]_{kl} = \begin{cases} m\gamma\mu, & \text{for } l = k-1;\ m \leq k;\ k = 1, 2, ..., S. \\ k\gamma\mu, & \text{for } l = k-1;\ m > k;\ k = 1, 2, ..., S. \\ m(1-\gamma)\mu, & \text{for } l = k;\ m \leq k;\ k = 1, 2, ..., S. \\ k\gamma\mu, & \text{for } l = k;\ m > k;\ k = 1, 2, ..., S. \\ 0, & \text{otherwise.} \end{cases}
$$

$$
[A_1^m]_{kl} = \begin{cases} -\beta, & \text{for } l = k = 0. \\ -(\lambda+\beta+m\mu), & \text{for } l = k;\ m \leq k;\ k = 1, 2, ..., s. \\ -(\lambda+\beta+k\mu), & \text{for } l = k;\ m > k \geq 1. \\ -(\lambda+m\mu), & \text{for } l = k;\ k = s+1, s+2, ..., S. \\ \beta, & \text{for } l = k+Q;\ k = 0, 1, ..., s. \\ 0, & \text{otherwise.} \end{cases}
$$

4.5.1 System stability and computation of steady-state probability vector

The Markov chain under consideration is a LIQBD process. For this chain to be stable it is necessary and sufficient that

$$\boldsymbol{\xi}\bar{A}_0\mathbf{e} < \boldsymbol{\xi}\bar{A}_2\mathbf{e}. \tag{4.34}$$

where $\boldsymbol{\xi}$ is the unique non negative vector satisfying,

$$\boldsymbol{\xi}\bar{A} = 0,\ \boldsymbol{\xi}\mathbf{e} = 1. \tag{4.35}$$

and $\bar{A} = \bar{A}_0 + \bar{A}_1 + \bar{A}_2$, is the infinitesimal generator of the finite state CTMC on the set $\{0, 1, \ldots, S\}$. Write $\boldsymbol{\xi}$ as $(\xi_0, \xi_1, \ldots, \xi_S)$. Then we get from (4.35), the components of the probability vector $\boldsymbol{\xi}$ explicitly as,

$$\xi_0 = \left\{ \begin{array}{l} 1 + \sum\limits_{i=1}^{c-1} \prod\limits_{k=0}^{i-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} + \prod\limits_{k=1}^{c-1} \frac{\beta+k\gamma\mu}{k\gamma\mu} \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{s+1-c} - 1 \right] \\ +Q \left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{s+1-c} \prod\limits_{k=0}^{c-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} - \frac{s\beta}{c\gamma\mu} \left[1 + \sum\limits_{i=0}^{s-2} \prod\limits_{k=0}^{i} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right] \\ + \frac{\beta^2}{c\gamma\mu} \left[1 + \sum\limits_{i=1}^{s-2} \prod\limits_{k=1}^{i} \frac{\beta+k\gamma\mu}{k\gamma\mu} \right] \end{array} \right\}^{-1},$$

$$\xi_i = \begin{cases} \prod\limits_{k=0}^{i-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) \xi_0, & \text{for } 1 \le i \le c. \\ \left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{i-c} \prod\limits_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) \xi_0, & \text{for } c+1 \le i \le s+1. \\ \xi_{i+1}, & \text{for } s+1 \le i \le Q-1. \end{cases}$$

and

$$\xi_{Q+i} = \begin{cases} \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{s+1-c} \prod\limits_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) - \frac{\beta}{c\gamma\mu} \right] \xi_0, \text{ for } i = 1. \\ \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{s+1-c} \prod\limits_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) - \frac{\beta}{c\gamma\mu} \left[1 + \sum_{j=0}^{i-2} \prod\limits_{k=0}^{j} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right] \right] \xi_0, \\ \quad \text{for } 2 \le i \le s. \end{cases}$$

From the relation (4.34) we have

Lemma 4.5.1. The stability condition of the queueing-inventory system under study is given by $\rho_2 < 1$, where $\rho_2 = \frac{\lambda(1-\xi_0)}{\mu \left[\sum\limits_{j=1}^{c-1} j\xi_j + c \sum\limits_{j=c}^{Q+s} \xi_j \right]}$.

Proof. On the same lines as that of Lemma (4.2.1). □

Next we compute the steady-state probability vector of $\boldsymbol{\mathcal{W}_2}$ under the stability condition. Let $\boldsymbol{y}$ denote the steady-state probability vector of the generator $\boldsymbol{\mathcal{W}_2}$. So $\boldsymbol{y}$ must satisfy the relations

$$\boldsymbol{y}\boldsymbol{\mathcal{W}_2} = 0, \ \boldsymbol{y}\mathbf{e} = 1. \tag{4.36}$$

Let us partition $\boldsymbol{y}$ by levels as

$$\boldsymbol{y} = (\boldsymbol{y_0}, \boldsymbol{y_1}, \boldsymbol{y_2}, \ldots). \tag{4.37}$$

where the subvectors of $\boldsymbol{y}$ are further partitioned as,

$$\boldsymbol{y_i} = (y_i(0), y_i(1), y_i(2), \ldots, y_i(S)), \ i \geq 0. \tag{4.38}$$

The steady-state probability vector $\boldsymbol{y}$ is obtained as,

$$\boldsymbol{y_{i+c-1}} = \boldsymbol{y_{c-1}} R^i, \ i \geq 1. \tag{4.39}$$

where R is the minimal non-negative solution to the matrix quadratic equation

$$R^2 \bar{A}_2 + R\bar{A}_1 + \bar{A}_0 = 0.$$

and the vectors $\boldsymbol{y_0}$, $\boldsymbol{y_1}$, …, $\boldsymbol{y_{c-1}}$ can be obtained by solving the following equations,

$$\left.\begin{array}{l} \boldsymbol{y_0} B + \boldsymbol{y_1} A_2^1 = 0. \\ \boldsymbol{y_{i-1}} \bar{A}_0 + \boldsymbol{y_i} A_1^i + \boldsymbol{y_{i+1}} A_2^{i+1} = 0, \ 1 \leq i \leq c-1. \\ \boldsymbol{y_{c-2}} \bar{A}_0 + \boldsymbol{y_{c-1}} \left(A_1^{c-1} + R\bar{A}_2\right) = 0. \end{array}\right\} \tag{4.40}$$

Now from (4.40), we get

$$\boldsymbol{y_0} = \boldsymbol{y_1} A_2^1 (-B)^{-1} = \boldsymbol{y_1} A_2^1 (-\bar{A}_0{}')^{-1}.$$

$$\boldsymbol{y_1} = -\boldsymbol{y_2} A_2^2 \left[A_1^1 + A_2^1 (-\bar{A}_0{}')^{-1} \bar{A}_0\right]^{-1} = \boldsymbol{y_2} A_2^2 (-\bar{A}_1{}')^{-1}$$

$$\boldsymbol{y_i} = \boldsymbol{y_{i+1}} A_2^{i+1}(-\bar{A}_i{}')^{-1},\ 0 \leq i \leq c-1,$$

where

$$\bar{A}_i{}' = \begin{cases} B,\ i = 0. \\ A_1^i + A_2^i(-\bar{A}'_{i-1})^{-1}\bar{A}_0,\ 1 \leq i \leq c, \end{cases}$$

subject to normalizing condition

$$\sum_{i=1}^{c-2} \boldsymbol{y_i} + \boldsymbol{y_{c-1}}(I-R)^{-1}\mathbf{e} = 1.$$

Since R cannot be computed explicitly we explore the possibility of algorithmic computation. Thus, one can use logarithmic reduction algorithm as in [45] for computing R. We list here only the main steps involved in logarithmic reduction algorithm for computation of R.

Logarithmic Reduction Algorithm for R:

Step 0: $H \leftarrow (-\bar{A}_1)^{-1}\bar{A}_0$, $L \leftarrow (-\bar{A}_1)^{-1}\bar{A}_2$, $G = L$, and $T = H$.

Step 1:

$$\begin{gathered} U = HL + LH \\ M = H^2 \\ H \leftarrow (I-U)^{-1}M \\ M \leftarrow L^2 \\ L \leftarrow (I-U)^{-1}M \\ G \leftarrow G + TL \\ T \leftarrow TH \end{gathered}$$

Continue Step 1 until $||\mathbf{e} - G\mathbf{e}||_\infty < \epsilon$.

Step 2: $R = -\bar{A}_0(\bar{A}_1 + \bar{A}_0 G)^{-1}$.

4.6 Conditional probability distributions

We could arrive at analytical expression for system state probabilities of $M/M/2$ queueing-inventory system. However for the $M/M/c$ queueing-inventory system with $c \geq 3$, the system state distribution does not seem to have closed form owing to the strong dependence between the inventory level, number of customers and the number of servers in the system. In this section we provide conditional probabilities of the number of items in the inventory, given the number of customers in the system and also that of the number of customers in the system conditioned on the number of items in the inventory.

4.6.1 Conditional probability distribution of the inventory level conditioned on the number of customers in the system

Let $\boldsymbol{\eta}= (\eta_0, \eta_1, ..., \eta_S)$ be the probability distribution of the inventory level conditioned on the number of customers in the system. Then we get explicit form for the conditional probability distribution of the inventory level conditioned on the number of customers in the system. We formulate the result in the following lemma:

Lemma 4.6.1. Assume that i is the number of customers in the system at same point of time. Conditional on this we compute the inventory level distribution. We consider two cases as follows:
(i) When $i < c$, the inventory level probability distribution is given by,

$$\eta_0 = \left\{ \begin{array}{l} 1 + \sum_{j=1}^{i-1} \prod_{k=0}^{j-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} + \prod_{k=1}^{i-1} \frac{\beta+k\gamma\mu}{k\gamma\mu} \left[\left(\frac{\beta+i\gamma\mu}{i\gamma\mu} \right)^{s+1-i} - 1 \right] \\ +Q \left(\frac{\beta+i\gamma\mu}{i\gamma\mu} \right)^{s+1-i} \prod_{k=0}^{i-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} - \frac{s\beta}{i\gamma\mu} \left[1 + \sum_{j=0}^{s-2} \prod_{k=0}^{j} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right] \\ + \frac{\beta^2}{i\gamma\mu} \left[1 + \sum_{j=1}^{s-2} \prod_{k=1}^{j} \frac{\beta+k\gamma\mu}{k\gamma\mu} \right] \end{array} \right\}^{-1},$$

$$\eta_j = \begin{cases} \prod_{k=0}^{j-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) \eta_0, & \text{for } 1 \le j \le i. \\ \left(\frac{\beta+i\gamma\mu}{i\gamma\mu} \right)^{j-i} \prod_{k=0}^{i-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) \eta_0, & \text{for } i+1 \le j \le s+1. \\ \eta_{j+1}, & \text{for } s+1 \le j \le Q-1. \end{cases}$$

and

$$\eta_{Q+j} = \begin{cases} \left[\left(\frac{\beta+i\gamma\mu}{i\gamma\mu} \right)^{s+1-i} \prod_{k=0}^{i-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) - \frac{\beta}{i\gamma\mu} \right] \eta_0, \text{ for } j = 1. \\ \left[\left(\frac{\beta+i\gamma\mu}{i\gamma\mu} \right)^{s+1-i} \prod_{k=0}^{i-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) - \frac{\beta}{i\gamma\mu} \left[1 + \sum_{i=0}^{j-2} \prod_{k=0}^{i} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right] \right] \eta_0, \\ \text{for } 2 \le j \le s. \end{cases}$$

(ii) When $i \ge c$, the inventory level probability distribution is derived by,

$$\eta_0 = \left\{ \begin{array}{l} 1 + \sum_{j=1}^{c-1} \prod_{k=0}^{j-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} + \prod_{k=1}^{c-1} \frac{\beta+k\gamma\mu}{k\gamma\mu} \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{s+1-c} - 1 \right] \\ +Q \left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{s+1-c} \prod_{k=0}^{c-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} - \frac{s\beta}{c\gamma\mu} \left[1 + \sum_{j=0}^{s-2} \prod_{k=0}^{j} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right] \\ + \frac{\beta^2}{c\gamma\mu} \left[1 + \sum_{j=1}^{s-2} \prod_{k=1}^{j} \frac{\beta+k\gamma\mu}{k\gamma\mu} \right] \end{array} \right\}^{-1},$$

$$\eta_j = \begin{cases} \prod_{k=0}^{j-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) \eta_0, & \text{for } 1 \le j \le c. \\ \left(\frac{\beta+c\gamma\mu}{c\gamma\mu} \right)^{j-c} \prod_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} \right) \eta_0, & \text{for } c+1 \le j \le s+1. \\ \eta_{j+1}, & \text{for } s+1 \le j \le Q-1. \end{cases}$$

and

$$\eta_{Q+j} = \begin{cases} \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{s+1-c} \prod\limits_{k=0}^{c-1}\left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right) - \frac{\beta}{c\gamma\mu}\right]\eta_0, \text{ for } j=1. \\ \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{s+1-c} \prod\limits_{k=0}^{c-1}\left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right) - \frac{\beta}{c\gamma\mu}\left[1+\sum\limits_{c=0}^{j-2}\prod\limits_{k=0}^{c}\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right]\right]\eta_0, \\ \text{for } 2 \le j \le s. \end{cases}$$

Proof. Let $\mathbf{\Gamma_1}$ be the infinitesimal generator of the corresponding Markov chain.

(i) Case of $i < c$.

The infinitesimal generator $\mathbf{\Gamma_1}$ is given by,

$$\mathbf{\Gamma_1} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \vdots \\ c \\ \vdots \\ s \\ \\ \vdots \\ Q \\ \vdots \\ S \end{array} \begin{array}{ccccccccccc} 0 & 1 & \cdots & i & \cdots & c & \cdots & s\cdots & Q & \cdots & S \\ -\beta & & & & & & & & \beta & & \\ \gamma\mu & -(\gamma\mu+\beta) & & & & & & & & & \\ & 2\gamma\mu & -(2\gamma\mu+\beta) & & & & & & & & \\ & \ddots & \ddots & & & & & & & \ddots & \\ & & i\gamma\mu & -(i\gamma\mu+\beta) & & & & & & & \\ & & & \ddots & \ddots & & & & & & \\ & & & & i\gamma\mu & -(i\gamma\mu+\beta) & & & & & \\ & & & & & \ddots & \ddots & & & & \\ & & & & & i\gamma\mu & -(i\gamma\mu+\beta) & & & & \beta \\ & & & & & & i\gamma\mu & -i\gamma\mu & & & \\ & & & & & & & & & & \\ & & & & & & & \ddots & \ddots & & \\ & & & & & & & & i\gamma\mu & -i\gamma\mu & \\ & & & & & & & & & i\gamma\mu & -i\gamma\mu \end{array}.$$

and

The inventory level distribution $\boldsymbol{\eta}$ can be obtained from the equations $\boldsymbol{\eta}\mathbf{\Gamma_1}= 0$ and $\boldsymbol{\eta}\mathbf{e}= 1$, we get

$$\eta_j = \begin{cases} \prod\limits_{k=0}^{j-1}\left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right)\eta_0, \text{ for } 1 \le j \le i. \\ \left(\frac{\beta+i\gamma\mu}{i\gamma\mu}\right)^{j-i} \prod\limits_{k=0}^{i-1}\left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right)\eta_0, \text{ for } i+1 \le j \le s+1. \\ \eta_{j+1}, \text{for} s+1 \le j \le Q-1. \end{cases}$$

$$\eta_{Q+j}=\begin{cases}\left[\left(\frac{\beta+i\gamma\mu}{i\gamma\mu}\right)^{s+1-i}\prod\limits_{k=0}^{i-1}\left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right)-\frac{\beta}{i\gamma\mu}\right]\eta_0,\ \text{for } j=1.\\ \left[\left(\frac{\beta+i\gamma\mu}{i\gamma\mu}\right)^{s+1-i}\prod\limits_{k=0}^{i-1}\left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right)-\frac{\beta}{i\gamma\mu}\left[1+\sum\limits_{i=0}^{j-2}\prod\limits_{k=0}^{i}\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right]\right]\eta_0,\\ \text{for } 2\le j\le s.\end{cases}$$

$$\text{where, }\eta_0=\left\{\begin{array}{l}1+\sum\limits_{j=1}^{i-1}\prod\limits_{k=0}^{j-1}\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}+\prod\limits_{k=1}^{i-1}\frac{\beta+k\gamma\mu}{k\gamma\mu}\left[\left(\frac{\beta+i\gamma\mu}{i\gamma\mu}\right)^{s+1-i}-1\right]\\ +Q\left(\frac{\beta+i\gamma\mu}{i\gamma\mu}\right)^{s+1-i}\prod\limits_{k=0}^{i-1}\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}-\frac{s\beta}{i\gamma\mu}\left[1+\sum\limits_{j=0}^{s-2}\prod\limits_{k=0}^{j}\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right]\\ +\frac{\beta^2}{i\gamma\mu}\left[1+\sum\limits_{j=1}^{s-2}\prod\limits_{k=1}^{j}\frac{\beta+k\gamma\mu}{k\gamma\mu}\right]\end{array}\right\}^{-1}.$$

(ii) Case of $i\ge c$.

The infinitesimal generator $\mathbf{\Gamma_2}$ is given by,

$$\mathbf{\Gamma_2}=\begin{array}{c|cccccccccccc} & 0 & 1 & \dots & c & \dots & i & \dots & s\dots & Q & \dots & S\\ \hline 0 & -\beta & & & & & & & & \beta & & \\ 1 & \gamma\mu & -(\gamma\mu+\beta) & & & & & & & & & \\ 2 & & 2\gamma\mu & -(2\gamma\mu+\beta) & & & & & & & & \\ \vdots & & \ddots & \ddots & & & & & & & \ddots & \\ c & & & c\gamma\mu & -(c\gamma\mu+\beta) & & & & & & & \\ \vdots & & & & \ddots & \ddots & & & & & & \\ i & & & & & c\gamma\mu & -(c\gamma\mu+\beta) & & & & & \\ \vdots & & & & & & \ddots & \ddots & & & & \\ s & & & & & & c\gamma\mu & -(c\gamma\mu+\beta) & & & & \beta\\ & & & & & & & c\gamma\mu & -c\gamma\mu & & & \\ \vdots & & & & & & & & & & & \\ Q & & & & & & & & \ddots & \ddots & & \\ \vdots & & & & & & & & & c\gamma\mu & -c\gamma\mu & \\ S & & & & & & & & & & c\gamma\mu & -c\gamma\mu \end{array}.$$

By solving the equations $\boldsymbol{\eta}\mathbf{\Gamma_2}=0$ and $\boldsymbol{\eta}\mathbf{e}=1$, we get

$$\eta_j = \begin{cases} \prod_{k=0}^{j-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right) \eta_0, & \text{for } 1 \le j \le c. \\ \left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{j-c} \prod_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right) \eta_0, & \text{for } c+1 \le j \le s+1. \\ \eta_{j+1}, & \text{for } s+1 \le j \le Q-1. \end{cases}$$

and

$$\eta_{Q+j} = \begin{cases} \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{s+1-c} \prod_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right) - \frac{\beta}{c\gamma\mu}\right] \eta_0, \text{ for } j = 1. \\ \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{s+1-c} \prod_{k=0}^{c-1} \left(\frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right) - \frac{\beta}{c\gamma\mu}\left[1 + \sum_{c=0}^{j-2} \prod_{k=0}^{c} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right]\right] \eta_0, \\ 2 \le j \le s. \end{cases}$$

$$\text{where, } \eta_0 = \left\{ \begin{array}{l} 1 + \sum_{j=1}^{c-1} \prod_{k=0}^{j-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} + \prod_{k=1}^{c-1} \frac{\beta+k\gamma\mu}{k\gamma\mu} \left[\left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{s+1-c} - 1\right] \\ +Q\left(\frac{\beta+c\gamma\mu}{c\gamma\mu}\right)^{s+1-c} \prod_{k=0}^{c-1} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu} - \frac{s\beta}{c\gamma\mu}\left[1 + \sum_{j=0}^{s-2} \prod_{k=0}^{j} \frac{\beta+k\gamma\mu}{(k+1)\gamma\mu}\right] \\ + \frac{\beta^2}{c\gamma\mu}\left[1 + \sum_{j=1}^{s-2} \prod_{k=1}^{j} \frac{\beta+k\gamma\mu}{k\gamma\mu}\right] \end{array} \right\}^{-1}.$$

□

4.6.2 Conditional probability distribution of the number of customers given the number of items in the inventory

Let $p_i, i \ge 0$, denote the probability that there are i customers in the system conditioned on the inventory level at j. We have three different cases:
(i) When $j = 0$,

$$p_i = \frac{\mu}{\mu+\lambda+\beta} p_{i+1}, \text{ for } i \ge 1.$$

and

$$p_0 = \frac{\mu}{\mu+\lambda+\beta} p_1, \text{ for } i = 0.$$

(ii) When $0 < j < c$,

$$p_i = \begin{cases} \frac{\lambda^i}{i!\mu^i} p_0, & \text{for } i < j. \\ \frac{\lambda^i}{j! j^{i-j} \mu^i} p_0, & \text{for } i \geq j;\ i < c. \\ \frac{\lambda^i}{j! j^{i-j} \mu^i} p_0, & \text{for } i \geq j;\ 0 < j \leq c;\ i \geq c. \end{cases}$$

(iii) When $j \geq c$,

$$p_i = \begin{cases} \frac{\lambda^i}{i!\mu^i} p_0, & \text{for } 1 \leq i < c. \\ \frac{\lambda^i}{c! c^{i-c} \mu^i} p_0, & \text{for } i \geq c;\ j \leq i. \\ \frac{\lambda^i}{c! c^{i-c} \mu^i} p_0, & \text{for } c \leq i \leq j. \end{cases}$$

4.6.3 Performance measures

- Mean number of customers in the system, $L_s = \sum_{i=1}^{\infty} \sum_{j=0}^{Q+s} i y_i(j)$.
- Mean number of customers in the queue, $L_q = \sum_{i=c+1}^{\infty} \sum_{j=0}^{Q+s} (i-c) y_i(j)$.
- Mean inventory level in the system, $I_m = \sum_{i=0}^{\infty} \sum_{j=1}^{Q+s} j y_i(j)$.
- Mean number of busy server,
  $P_{BS} = \sum_{k=1}^{c} k \left[\sum_{i=k+1}^{\infty} y_i(k) + \sum_{j=k+1}^{Q+s} y_k(j) + y_k(k) \right]$
- Mean number of idle server , $P_{IS} = \left(c - \sum_{i=0}^{\infty} y_i(0) \right)$.
- Depletion rate of inventory, $D_{inv} = \gamma\lambda \left(\sum_{i=0}^{\infty} \sum_{j=1}^{Q+s} y_i(j) \right)$.

- Mean number of replenishments per time unit, $R_r = \beta \left(\sum_{i=0}^{\infty} \sum_{j=0}^{s} y_i(j) \right)$.
- Mean number of departures per unit time,
$D_m = \sum_{k=1}^{c-1} \left[k\mu \left(\sum_{i=k}^{\infty} y_i(k) + \sum_{j=k}^{Q+s} y_k(j) \right) \right] + c\mu \left[\sum_{i=c}^{\infty} \sum_{j=c}^{Q+s} y_i(j) \right]$.
- Expected loss rate of customers, $E_{loss} = \lambda \left(\sum_{i=0}^{\infty} y_i(0) \right)$.
- Expected loss rate of customers when the inventory level is zero per cycle, $E_{loss}^{c} = \frac{E_{loss}}{R_r}$.
- Mean number of customers arriving per unit time,
$\lambda_A = \lambda \left(\sum_{i=0}^{\infty} \sum_{j=1}^{Q+s} y_i(j) \right)$.
- Mean sojourn time of the customers in the system, $W_s = \frac{L_s}{\lambda_A}$.
- Mean waiting time of a customer in the queue, $W_q = \frac{L_q}{\lambda_A}$.
- Mean number of customers waiting in the system when inventory is available, $\widetilde{W} = \sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} i y_i(j)$.
- Mean number of customers waiting in the system during the stock out period, $\widetilde{\widetilde{W}} = \sum_{i=1}^{\infty} i y_i(0)$.

4.7 Analysis of inventory cycle time

We define the inventory cycle time random, Γ_{cycle} as the time interval between two consecutive instants at which the inventory level drops to s. Thus the inventory cycle time is a random variable whose distribution depends on the number of customers at the time when inventory level dropped to s at the beginning of the cycle and the inventory level process prior to replenishment. ***We proceed with the assumption that*** $\boldsymbol{\gamma = 1}$. If the number of customers present in the system is at least $Q+c$ when the order for replenishment is placed, then we need not have to look at future arrivals to get a nice form for the cycle time distribution. In fact it is sufficient that there are at least Q customers at that epoch. However in this case the service rate during lead time may drop below $c\mu$ even when there are at least c items in the inventory. This is so since number of customers may go below c.

4.7.1 When the number of customers $\ell \geq Q + c$

When the number of customers is at least $Q + c$, future arrivals need not be considered. The service rate of the $M/M/c$ queueing-inventory system depends on the number of customers, number of servers and number of items in the inventory. Thus we consider the following cases:

Case 1. Replenishment occurs before inventory level hits $c - 1$.
We consider the state (ℓ, s) as the starting state; thus the inventory level decreases from s to a particular level $s - k$, k vary from 0 to $s - c$ due to service completion at rate $c\mu$, during the lead time. At level $s - k$, the replenishment occurs and it is absorbed to $\{\Delta_1\}$, where the absorbing state is defined as $\{\Delta_1\} = \{(\ell - k, Q + s - k)|0 \leq k \leq s - c\}$. Therefore the time until absorption to $\{\Delta_1\}$ follows Erlang distribution of order k with parameter $c\mu$, it is denoted as $E(c\mu; k)$. Now, the number of customers in the system is $\ell - k$ or larger with the corresponding inventory level $Q+s-k$, for k varying from 0 to $s - c$. Similarly, the inventory level reaches s from $Q + s - k$ with $Q - k$ service completions all of which have rate $c\mu$. This

time duration also follows Erlang distribution of order $Q-k$. Write this as $E(c\mu; Q-k)$. Thus under the condition that there are at least $Q+c$ customers at the beginning of the cycle and that the inventory level does not fall below c, the inventory cycle time, Γ_{cycle} has Erlang distribution of order Q with parameter $c\mu$. That is,

$$\begin{aligned}\Gamma_{cycle} &\sim E(c\mu; k) * E(c\mu; Q-k) \\ &\sim E(c\mu; Q).\end{aligned}$$

where the symbol "$\sim$" stands for "has distribution". The probability of replenishment taking place before inventory level drops to $c-1$, is given by $\int\limits_0^{\infty} \sum\limits_{k=0}^{s-c} \frac{e^{-\mu v}(\mu v)^k \beta e^{-\beta v}}{k!} dv$.

Case 2. Replenishment after hitting $c-1$ but not zero.
The inventory level decreases from s to k, when k varies from 1 to $c-1$. The first $s-c+1$ services are at the same rate $c\mu$. Thereafter it shows down to $(c-1)\mu$ and finally to $k\mu$, when replenishment occurs. Consequently the inventory level rises to $Q+k$. Now on the service rate stays at $c\mu$. Thus in the cycle, the distribution of the time until replenishment takes place is the convolution of generalized Erlang distribution and that of an Erlang distribution $E(Q+k-s; c\mu)$. The conditional distribution of replenishment realization after $s-k-1$ service are completed, but before $(s-k)^{th}$ is completed, can be computed as in case 1. At the same level $s-k$, the replenishment will occur and it is absorbed to $\{\Delta_2\}$, where the absorbing state is defined as $\{\Delta_2\} = \{(\ell-(s-k), Q+k) | 1 \leq k \leq c-1\}$. Thus, the time until absorption to $\{\Delta_2\}$ follows generalized Erlang distribution with parameters $c\mu, (c-1)\mu, \ldots, (k+1)\mu$ of order $s-k$ and k vary from 1 to $c-1$. It is denoted as $\mathcal{G}E\left(c\mu, (c-1)\mu, \ldots, (k+1)\mu;\ s-k\right)$. Then from

$\{\Delta_2\}$ the inventory level reaches s due to service completion with parameter $c\mu$. Thus the time duration follows Erlang distribution with parameter $c\mu$ of order $Q+k-s$, k vary from 1 to $c-1$. That is, $E\,(c\mu; Q+k-s)$. Hence the inventory cycle time, Γ_{cycle} follows generalized Erlang distribution of order Q. Therefore, Γ_{cycle} is defined as

$$\Gamma_{cycle} \sim \mathcal{G}E\left(\underbrace{c\mu,\ldots,c\mu}_{s-c+1\ times},(c-1)\mu,\ldots,(k+1)\mu;\ s-k\right)*E\,(c\mu; Q+k-s)$$

where $\mathcal{G}E(.)$ stands for generalized Erlang distribution.

Case 3. Replenishment after inventory level reaching zero.
Then the inventory level reaches 0 from the level s due to service completion with parameters $c\mu$ (repeated $s-c+1$ times). Thus the time until absorption to $\{\Delta_3\} = \{(\ell - s, Q)\}$ follows generalized Erlang distribution of order s and parameters $c\mu, (c-1)\mu, \ldots, \mu$. When the inventory level hits 0, the system becomes idle for a random duration of time which follows exponential distribution with parameter β. After replenishment, the system starts service and consequently the inventory level reaches s from Q due to service completion with parameter $c\mu$. This part has Erlang distribution with parameter $c\mu$ and order $Q-s$. Thus, Γ_{cycle} follows generalized Erlang distribution of order Q. That is,

$$\Gamma_{cycle} \sim \mathcal{G}E\left(\underbrace{c\mu,\ldots,c\mu}_{s-c+1\ times},(c-1)\mu,\ldots,\mu;\ s\right)*exp(\beta)*E(c\mu; Q-s)$$

The cases we are going to consider hereafter result in cycle time distribution that are phase type with not necessarily unique representation. However, one can sort out the problem of minimal representation. Obviously this is the one which considers that many arrivals needed to have exactly Q services in this cycle.

4.7.2 When the number of customers $\ell < Q + c$

In this case we may have to consider future arrivals as well, since number of customers available at the start of the cycle may be such that the service rate falls below $c\mu$. Thus the cycle time will have more general distribution, namely the phase type. We go about doing this. Our procedure is such that the moment we have enough customers to serve during the remaining part of the cycle, we stop looking at future arrivals. Thus consider a Markov chain on the state space

$$\begin{aligned}&\{(s,\ell)\,,(s-1,\ell-1)\,,\ldots,(0,\ell-s)\,,(s,\ell+1)\,,(s-1,\ell)\\&\ldots,(0,\ell-s+1)\,,\ldots,(s+Q,\ell)\,,\ldots,(s+Q-\ell,0)\,,\\&(s+Q,\ell+1)\,,\ldots,(s,\ell)\,,\ldots(s+Q,s+Q-\ell-1)\,,\\&\cdots,\\&(s+Q-1,\ell-1)\,,(s+Q,s+Q-\ell)\,,\ldots,(s,s-\ell-1)\}\,.\end{aligned}$$

The initial state (s,ℓ). Thus the initial probability vector will have one at the position corresponding to (s,ℓ) and the rest of the elements zero. The absorption state in this Markov chain is $(s,*)$, where $*$ belonging to $\{0,1,2,\ldots,Q+\ell-s\}$ and is a departure epoch. Let $\mathcal{T}$ be the block with transitions among transient states and $\mathcal{T}^*$ be the column vector with transition rates to the absorbing states as elements. Then the cycle time has distribution $1-\boldsymbol{\alpha}e^{\mathcal{T}t}\mathbf{e}$ where $\boldsymbol{\alpha}$ is the initial probability vector with 1 at the position indicating the inventory level s as first coordinate and the number of customers $(=\ell)$ at the beginning of the cycle as second coordinate. Note that the phase type representation obtained is not unique since the service rate strongly depends on both inventory level and number of customers in the system. The case of $\ell < s$: Here again the procedure is similar to that corresponding to $\ell \geq s$, but less than $Q + c$. The initial state is (s,ℓ). After exactly Q service completions with a replenishment within this cycle and with arrivals truncated at that epoch which ensure rate $c\mu$ for as many services as possible. The absorption state of the Markov chain generated corresponds to a departure epoch with s items in the inventory. Here again the cycle time has a PH distribution with representation which is not unique because the service rates may change depending on the number of customers in the system and the number of items in the inventory.

4.8 Optimization problem II

We look for the optimal pair of control variables in the model discussed above. Now for computing the minimal cost of (s, Q) model we introduce the cost function: $\mathcal{F}(c, s, Q)$ which is defined by,

$$\mathcal{F}(c, s, Q) = h.I_m + c_1.E_{loss} + c_2.\widetilde{\widetilde{W}} + (K + Q.c_3)\,.R_r + c_4.P_{BS} + c_5.\left(c - P_{BS}\right)$$

where $s = 40$, $S = 81$ and $K, c_1, c_2, c_3, c_4, c_5$, h are the same input parameters as described in Section 4.4. We provide optimal c and corresponding minimum cost for various γ values. From Table 4.2 we notice that the optimal value of c is 6 for various γ values, presumably become of the high holding cost.

Table 4.2: Optimal server c and minimum cost

γ	0.1	0.2	0.3	0.4	0.5
Optimal c	6	6	6	6	6
& minimum cost	148.78	166.36	183.93	201.51	219.09
γ	0.6	0.7	0.8	0.9	1
Optimal c	6	6	6	6	6
& minimum cost	236.66	254.24	271.82	289.40	306.98

Table 4.3: Optimal (s, Q) values and minimum cost

c \ γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
3	**(4,12)** **95.857**	(4,15) 113.57	(4,19) 131.86	(4,23) 151.49	(4,27) 171.55	(4,31) 191.59	(4,37) 211.50	(4,39) 230.82	(4,42) 249.87	(4,45) 268.55
4	(5,15) 126.43	(5,19) 147.25	(5,23) 166.70	(5,27) 186.06	(5,31) 205.42	(5,34) 224.69	(5,38) 243.74	(5,41) 262.55	(5,45) 281.05	((5,48) 299.25
5	(6,16) 154.11	(6,22) 177.53	(6,26) 198.43	(6,31) 218.40	(6,34) 237.91	(6,38) 257.07	(6,42) 275.93	(6,45) 294.49	(6,48) 312.75	(6,52) 330.72
6	(7,16) 177.90	(7,23) 202.28	(7,28) 223.79	(7,32) 244.06	(7,36) 263.65	(7,40) 282.79	(7,43) 301.56	(7,46) 320.00	(7,49) 338.13	(7,53) 355.97
7	(8,16) 200.30	(8,23) 224.90	(8,28) 246.58	(8,32) 266.91	(8,36) 286.47	(8,40) 305.52	(8,43) 324.18	(8,46) 342.49	(8,49) 360.51	(8,53) 378.24
8	(9,16) 222.30	(9,23) 246.98	(9,28) 268.68	(9,32) 288.97	(9,36) 308.45	(9,40) 329.01	(9,43) 345.92	(9,46) 364.12	(9,49) 382.02	(9,53) 399.65
9	(10,16) 244.29	(10,23) 268.97	(10,28) 290.65	(10,32) 310.87	(10,36) 330.25	(10,40) 349.08	(10,43) 367.49	(10,46) 385.57	(10,49) 403.36	(10,53) 420.89
10	(11,16) 266.27	(11,23) 290.94	(11,28) 312.60	(11,32) 332.76	(11,36) 352.06	(11,40) 370.78	(11,43) 389.08	(11,46) 407.05	(11,49) 424.73	(11,53) 442.17

In Table 4.3, we examine the optimal pair (s, Q) and the corresponding minimum cost for various of γ and c, keeping other parameters fixed (as in Section 4.4).

Chapter 5

Queueing-inventory system with working vacations and vacation interruptions

5.1 Introduction

In this chapter we discuss about queueing-inventory system under (s, Q) policy with working vacations and vacation interruptions. This investigation appears almost unrelated to problems discussed in the rest of the thesis. Nevertheless, if we replace the assumption of working vacation by the usual notion of idleness of the server due to the absence of customers and/ inventory then we recoop the model discussed in chapter 2. The notion of working vacation is introduced by Jihong Li and Naishuo Tian [22]. During working vacation also the server provides service, however, at a lower

Some results of this chapter are included in the following paper.
A. Krishnamoorthy, R. Manikandan and Sajeev S.Nair : Classical queueing-inventory system with working vacations and vacation interruptions (Under review).

rate. Further, the server can come back from the vacation mode to the normal working mode once some indices of the system, such as the number of customers achieve a certain value and items of the inventory are available during a working vacation. More precisely, the server may come back from the vacation without completing the vacation period. This is called vacation interruption (see [22]). We assume that if there are customers in the system at a service completion epoch during a working vacation, the server will comeback to the normal working mode; else the server stays in the working vacation mode. With the system having infinite capacity, we derive condition for stability of the system. Despite the corresponding queueing system (without inventory) having analytic solution, we are not able to arrive at even closed form expression for system state distribution for the queueing-inventory problem under discussion. Hence algorithmic approach is adopted which is given in Section 5.3. Several performance measures are evaluated in Section 5.3.3. An optimization problem is also discussed in Section 5.4.

5.2 Mathematical formulation

Consider a single server queueing-inventory system with working vacation and vacation interruptions. The server takes vacation only in the absence of customers in the system and not due to inventory level falling to zero at a service completion epoch. We assume that if there are customers in the system after a service completion during a working vacation period, the server will come back to the normal working mode. On the other hand if there are no customers in the system at the end of service in vacation mode, the server continues the vacation. This vacation duration follows exponential distribution with parameter θ.

Customers arrive to a single server counter according to a Poisson process of rate λ. They do not join the system when inventory level is zero. Service time follows exponential distribution with parameter μ_v during vacation period and μ_b during normal period. We assume that even when vacation mode is realized during a service in that mode, switching to normal mode is done starting with the next customers service only, provided there is at least one waiting on completion of the present service. The inventory replenishment is governed by the (s, Q) policy. Here s is the reorder level and $Q(= S - s)$ is the fixed order quantity. We assume $(S > 2s)$ to avoid perpetual reordering. Lead time is exponentially distributed with rate β. Then $\{\mathcal{X}(t)|t \geq 0\} = \{(\mathcal{N}(t), \mathcal{M}(t), \mathcal{I}(t))|t \geq 0\}$ is a CTMC with state space Ω is given by

$$\Omega = \bigcup_{i=0}^{\infty} \mathcal{L}(i)$$

where the state space of the CTMC is partitioned in to levels $\mathcal{L}(i)$ defined as

$$\mathcal{L}(0) = \{(0, 0, 0), (0, 0, 1), \ldots, (0, 0, Q + s)\}$$

and $\mathcal{L}(i) = \{(i, 0, 0), (i, 0, 1), \ldots, (i, 0, Q + s), (i, 1, 1), \ldots, (i, 1, Q + s)\}$, for $i \geq 1$. Now we describe the transitions in the Markov chain:

(a) Transitions due to arrival of customers:

$(i, 0, j) \rightarrow (i + 1, 0, j)$: the rate is λ, for $i \geq 0; 1 \leq j \leq Q + s$.

$(i, 1, j) \rightarrow (i + 1, 1, j)$: the rate is λ, for $i \geq 0; 1 \leq j \leq Q + s$.

(b) Transitions due to service completion during working vacation mode:

$(i, 0, j) \to (i-1, 0, j-1)$: the rate is μ_v, for $i = 1; 1 \leq j \leq Q+s$.

$(i, 0, j) \to (i-1, 1, j-1)$: the rate is μ_v, for $i \geq 2; 2 \leq j \leq Q+s$.

$(i, 0, 1) \to (i-1, 0, 0)$: the rate is μ_v, for $i \geq 2$.

(c) Transitions due to service completion during normal mode:

$(i, 1, j) \to (i-1, 0, j-1)$: the rate is μ_b, for $i = 1; 1 \leq j \leq Q+s$.

$(i, 1, j) \to (i-1, 1, j-1)$: the rate is μ_b, for $i \geq 2; 2 \leq j \leq Q+s$.

$(i, 1, 1) \to (i-1, 0, 0)$: the rate is μ_b, for $i \geq 2$.

(d) Transitions due to replenishment:

$(i, 0, j) \to (i, 0, Q+j)$: the rate is β, for $i \geq 0; 0 \leq j \leq s$.

$(i, 1, j) \to (i, 1, Q+j)$: the rate is β, for $i \geq 1; 1 \leq j \leq s$.

(e)Transitions due to vacation realization:

$(i, 0, j) \to (i, 1, j)$: the rate is θ, for $i \geq 1; 1 \leq j \leq Q+s$.

All other transition pairs have rate zero. The infinitesimal generator $\boldsymbol{\mathcal{W}}$ of this CTMC is expressed in a block partitioned form:

$$\mathcal{W} = \begin{bmatrix} C_1 & C_0 & & & & \\ C_2 & A_1 & A_0 & & & \\ & A_2 & A_1 & A_0 & & \\ & & A_2 & A_1 & A_0 & \dots \\ & & & \ddots & \ddots & \ddots \end{bmatrix},$$

where C_1 is a square matrix of dimension $S+1$ that represents transitions within $\mathcal{L}(0)$; C_0 and A_0 represent transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i+1)$ for $i \geq 0$, with dimensions $(S+1)\times(2S+1)$ and $(2S+1)\times(2S+1)$ respectively; C_2 has dimension $(2S+1)\times(S+1)$ and represents transitions from $\mathcal{L}(1)$ to $\mathcal{L}(0)$; A_1 is a square matrix of dimension $2S+1$ which represents transitions within $\mathcal{L}(i)$, $i \geq 1$, and A_2 is a square matrix of order $2S+1$ that represents transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i-1)$, $i \geq 2$.

5.3 Analysis of the system

In this section we discuss the steady-state analysis of the queueing-inventory system under study by first establishing the stability condition of the system. Define $A{=}A_0+A_1+A_2$. Let the steady-state probability vector of the generator matrix A be $\boldsymbol{\pi} = (\pi_0(0), \pi_0(1), \ldots, \pi_0(S), \pi_1(1), \pi_1(2), \ldots \pi_1(S))$. Then the relations $\boldsymbol{\pi}A = 0$ and $\boldsymbol{\pi}\mathbf{e} = 1$ gives the following equations,

$$\pi_0(Q) = \frac{\beta}{\theta+\mu_v}\pi_0(0)$$

$$\pi_0(1) = \pi_0(2) = \cdots = \pi_0(Q-1) = \pi_0(Q+1) = \pi_0(Q+2) = \cdots = \pi_0(S) = 0$$

$$\pi_1(1) = \frac{\beta}{\mu_b}\pi_0(0)$$

$$\pi_1(2) = \frac{\beta(\beta+\mu_b)}{\mu_b^2}\pi_0(0)$$

$$\pi_1(3) = \frac{\beta(\beta+\mu_b)^2}{\mu_b^3}\pi_0(0)$$

$$\vdots$$

$$\pi_1(s+1) = \frac{\beta(\beta+\mu_b)^s}{\mu_b^{s+1}}\pi_0(0)$$

$$\pi_1(s+1) = \pi_1(s+2) = \cdots = \pi_1(Q)$$

$$\pi_1(1) + \pi_1(Q+1) = \pi_1(Q)$$

$$\pi_1(2) + \pi_1(Q+2) = \pi_1(Q)$$

$$\pi_1(3) + \pi_1(Q+3) = \pi_1(Q)$$

$$\vdots$$

$$\pi_1(s) + \pi_1(S) = \pi_1(Q)$$

The LIQBD process with infinitesimal generator $\boldsymbol{\mathcal{W}}$ is stable if and only if $\boldsymbol{\pi}A_0\mathbf{e} < \boldsymbol{\pi}A_2\mathbf{e}$. That is,

$$\Longleftrightarrow \mu_b\left(\pi_1(1)+\pi_1(2)+\cdots+\pi_1(S)\right) > \lambda\left(\pi_1(1)+\pi_1(2)+\cdots+\pi_1(S)\right)+\lambda\pi_0(Q)$$

$$\Longleftrightarrow \mu_b Q\pi_1(Q) > \lambda\left(\pi_0(Q)+Q\pi_1(Q)\right)$$

$$\Longleftrightarrow \mu_b Q\frac{\beta(\beta+\mu_b)^s}{\mu_b^{s+1}}\pi_0(0) > \lambda\left(\frac{\beta}{\theta+\mu_v}\pi_0(0)+Q\frac{\beta(\beta+\mu_b)^s}{\mu_b^{s+1}}\pi_0(0)\right)$$

$$\Longleftrightarrow \lambda < \frac{\mu_b Q\frac{(\beta+\mu_b)^s}{\mu_b^{s+1}}}{\frac{1}{\theta+\mu_v}+Q\frac{(\beta+\mu_b)^s}{\mu_b^{s+1}}}$$

$$\Longleftrightarrow \lambda < \frac{\mu_b}{1+\frac{\mu_b^{s+1}}{(\theta+\mu_v)Q(\beta+\mu_b)^s}}$$

Thus we have the following result for the stability of the system:

Lemma 5.3.1. The CTMC Ω is stable if and only if $\lambda < \frac{\mu_b}{1+\frac{\mu_b^{s+1}}{(\theta+\mu_v)Q(\beta+\mu_b)^s}}$.

Proof. From the well known result in Neuts [47] on the positive recurrence of A, we have $\boldsymbol{\pi} A_0\mathbf{e} < \boldsymbol{\pi} A_2\mathbf{e}$. With a bit of computation, this simplifies to the result $\lambda < \frac{\mu_b}{1+\frac{\mu_b^{s+1}}{(\theta+\mu_v)Q(\beta+\mu_b)^s}}$. □

It may be noted that the above condition is weaker than the one corresponding to $M/M/1$ queueing-inventory systems discussed in chapters 2 and 3. Next we compute the steady-state probability vector $\boldsymbol{x}$ of the infinitesimal generator $\boldsymbol{\mathcal{W}}$ under the stability condition. The steady-state probability vector $\boldsymbol{x}$ be partitioned according to the levels as

$$\boldsymbol{x} = (\boldsymbol{x_0}, \boldsymbol{x_1}, \boldsymbol{x_2}, ...), \tag{5.1}$$

where the subvectors of $\boldsymbol{x}$ are further partitioned as

$$\boldsymbol{x_0} = (x_0(0,0), x_0(0,1), x_0(0,2), ..., x_0(0,S))\,, \tag{5.2}$$

$$\boldsymbol{x_i} = (x_i(0,0), x_i(0,1), x_i(0,2), ..., x_i(0,S), x_i(1,1), x_i(1,2), ..., x_i(1,S))\,,\ i \geq 1. \tag{5.3}$$

Suppose $\boldsymbol{x_{i+1}} = \boldsymbol{x_1}R^i,$ for $i \geq 1$. Then from $\boldsymbol{x\mathcal{W}} = 0$, we get

$$\boldsymbol{x_1}A_0 + \boldsymbol{x_2}A_1 + \boldsymbol{x_3}A_2 = 0$$

$$\Longrightarrow \boldsymbol{x_1}A_0 + \boldsymbol{x_2}RA_1 + \boldsymbol{x_1}R^2A_2 = 0$$

$$\Longrightarrow \boldsymbol{x_1}\left(A_0 + RA_1 + R^2A_2\right) = 0$$

Choose R such that $A_0 + RA_1 + R^2A_2 = 0$. Also we have

$$\boldsymbol{x_0}C_0 + \boldsymbol{x_1}C_2 = 0$$

$$\boldsymbol{x_0}C_1 + \boldsymbol{x_1}A_1 + \boldsymbol{x_2}A_2 = 0$$

$$\Longrightarrow \boldsymbol{x_0} C_1 + \boldsymbol{x_1} \left(A_1 + RA_2\right) = 0$$

$$\Longrightarrow \boldsymbol{x_1} = -\boldsymbol{x_0} C_1 \left(A_1 + RA_2\right)^{-1}$$

$$= \boldsymbol{x_0} V, \text{ where } V = -C_1 \left(A_1 + RA_2\right)^{-1}$$

Hence from the above we get $\boldsymbol{x_0} \left(C_0 + VC_2\right) = 0$. First take $\boldsymbol{x_0}$ as the steady-state vector of C_0+VC_2. Then $\boldsymbol{x_1} = \boldsymbol{x_0} V$ and $\boldsymbol{x_{i+1}} = \boldsymbol{x_1} R^i$, for $i \geq 1$. Now the steady-state probability distribution of the system is obtained by dividing each $\boldsymbol{x_i}$, with the normalizing constant

$$\left[\boldsymbol{x_0} + \boldsymbol{x_1} + \cdots\right] \mathbf{e} = \left[\boldsymbol{x_0} + \boldsymbol{x_1} \left(I - R\right)^{-1}\right] \mathbf{e}.$$

Once the matrix R is obtained, the vector $\boldsymbol{x}$ can be computed by exploiting the special structure of the coefficient matrices. One can use logarithmic reduction algorithm for computing R. We will list the main steps involved in the logarithmic reduction algorithm.

Logarithmic Reduction Algorithm for R:

Step 0: $H \leftarrow (-A_1)^{-1} A_0$, $L \leftarrow (-A_1)^{-1} A_2$, $G = L$, and $T = H$.

Step 1:

$$U = HL + LH$$
$$M = H^2$$
$$H \leftarrow (I - U)^{-1} M$$
$$M \leftarrow L^2$$
$$L \leftarrow (I - U)^{-1} M$$
$$G \leftarrow G + TL$$
$$T \leftarrow TH$$

Continue Step 1 until $||\boldsymbol{e} - G\boldsymbol{e}||_{\infty} < \epsilon$.

Step 2: $R = -A_0 (A_1 + A_0 G)^{-1}$.

5.3.1 Busy period analysis

For the system under study, we define busy period the time duration between the arrival of a customer to an empty system with positive inventory and the first epoch thereafter when the system is left with no customer immediately after a service completion. Thus it is precisely the first passage time from the state $(1, 0, j)$, for $1 \leq j \leq S$, to the state $(0, 0, \widetilde{j})$, for $0 \leq \widetilde{j} \leq S-1$. Busy cycle for the given system is the time interval between two successive departures, which leave the system empty (in terms of customers). Thus the busy cycle is the first return time to the state $(0, 0, \widetilde{j})$, for $0 \leq \widetilde{j} \leq S$ with at least one visit to any other state. Before analyzing the busy period structure we introduce the notion of fundamental period. For the QBD process under consideration, it is the first passage time from level i, where $i > 1$, to the level $i-1$. The cases $i = 1$ and $i = 0$ corresponding to the boundary states need to be discussed separately. It should be noted that due to the structure of the QBD process the distribution of the first passage time is invariant in i $(i \geq 2)$.

Let $G_{j\widetilde{j}}(k, \tau)$ denote the conditional probability that the QBD process starting in the state $(i, 0, j)$, for $1 \leq j \leq S$ and $i > 1$, at time 0, reaches the state $(i-1, 0, \widetilde{j})$, where $0 \leq \widetilde{j} \leq S-1$, for the first time, involving exactly k transitions and completing before time τ. Thus

$$G_{j\widetilde{j}}(k, \tau) = P[\tau < \infty : \chi(\tau) = \widetilde{j}/\chi(0) = j]$$

where τ is the first passage time from the level i to the level $i-1$ and χ is the QBD process under reference. Because of the structure of $\boldsymbol{\mathcal{W}}$, the probability $G_{j\widetilde{j}}(k, \tau)$ does not depend on i. The matrix with elements $G_{j\widetilde{j}}(k, \tau)$ is denoted by $G(k, \tau)$. For convenience, we write the joint transform matrix,

$$\widetilde{G}_{j\widetilde{j}}(z, \theta) = \sum_{k=1}^{\infty} z^k \int_0^{\infty} e^{-\theta\tau} \mathrm{d}G_{j\widetilde{j}}(k, \tau) \quad ; \quad |z| \leq 1, \theta > 0$$

and the matrix

$$\widetilde{G}(z,\theta) = (\widetilde{G}_{j\widetilde{j}}(z,\theta)).$$

The matrix $\widetilde{G}(z,\theta)$is the unique solution to the equation (see Neuts [47])

$$\widetilde{G}(z,\theta) = z(\theta I - A_1)^{-1}A_2 + (\theta I - A_1)^{-1}A_0\widetilde{G}^2(z,\theta). \tag{5.4}$$

Then the matrix $G = \widetilde{G}(1,0)$ takes care of the first passage times, except for the boundary states. If we know the matrix R then matrix G can be computed using the result (see [44])

$$G = -(A_1 + RA_2)^{-1}A_2. \tag{5.5}$$

We use logarithmic reduction method to compute G. For the boundary level states 1 and 0 let $G^{(1,0,j)}_{j\widetilde{j}}(k,\tau)$, for $1 \le j \le S$ and $G^{(0,0,\widetilde{j})}_{j\widetilde{j}}(k,\tau)$, for $0 \le \widetilde{j} \le S-1$, be the conditional probability discussed above for the first passage times from level 1 to level 0 and the first return time to the level 0 respectively. Then as in (5.4) we get

$$\widetilde{G}^{(1,0,j)}(z,\theta) = z(\theta I - A_1)^{-1}C_2 + (\theta I - A_1)^{-1}A_0\widetilde{G}(z,\theta)\widetilde{G}^{(1,0,j)}(z,\theta),\ 1 \le j \le S. \tag{5.6}$$

and

$$\widetilde{G}^{(0,0,\widetilde{j})}(z,\theta) = [\lambda/(\lambda+\theta), 0, \widetilde{j}]\widetilde{G}^{(1,0,j)}(z,\theta),\ 1 \le j \le S,\ 0 \le \widetilde{j} \le S-1. \tag{5.7}$$

Note that $\widetilde{G}^{(1,0,j)}(z,\theta)$, for $1 \le j \le S$ is a $(2S+1)\times(S+1)$ matrix. Thus the LST of the busy period is the first element of $\widetilde{G}^{(1,0,j)}(1,0)$. For future reference use the notations $G_{10} = \widetilde{G}^{(1,0,j)}(1,0)$, $G_{00} = \widetilde{G}^{(0,0,\widetilde{j})}(1,0)$, for $1 \le j \le S,\ 0 \le \widetilde{j} \le S-1$. Due to the positive recurrence of the QBD process, matrices G , G_{10}, and G_{00} are all stochastic. If we let $C_0 = (-A_1)^{-1}A_2$ and

$C_2 = (-A_1)^{-1}A_0$, then G is the minimal non negative solution (see [47]) to the matrix equation $G = C_0 + C_2G^2$. From equations (5.6) and (5.7) we get

$$G_{10} = -(A_1 + A_0G)^{-1}C_2 \tag{5.8}$$

and

$$G_{00} = [1, 0, \widetilde{j}]G_{10} \tag{5.9}$$

for $1 \leq \widetilde{j} \leq S - 1$ respectively. Equation (5.4) is equivalent to

$$zA_2 - (\theta I - A_1)\widetilde{G}(z, \theta) + A_0\widetilde{G}^2(z, \theta) = 0. \tag{5.10}$$

Let

$$D = -\left.\frac{\partial \widetilde{G}(z, \theta)}{\partial \theta}\right|_{z=1,\theta=0}$$

and

$$\widetilde{D} = \left.\frac{\partial \widetilde{G}(z, \theta)}{\partial z}\right|_{z=1,\theta=0}.$$

Differentiation of (5.10) with respect to θ and z followed by setting $z = 1$ and $\theta = 0$ leads to (see Neuts [47])

$$D = -A_1^{-1}G + C_2(GD + DG)$$

and

$$\widetilde{D} = C_0 + C_2(G\widetilde{D} + \widetilde{D}G).$$

With **0** as starting value for D and $\widetilde{D}$, successive substitutions in the above equations yield the values of D and $\widetilde{D}$. Applying an exactly similar reasoning to (5.6) and (5.7), we get

$$D_{10} = -(A_1 + A_0G)^{-1}(I + A_0D)G_{10},$$

and

$$D_{00} = [1/\lambda, 0, \widetilde{j}]G_{10} + [1, 0, \widetilde{j}]D_{10},\ 0 \leq \widetilde{j} \leq S - 1$$

where

$$D_{10} = -\left.\frac{\partial \widetilde{G}^{(1,0,j)}(z,\theta)}{\partial\theta}\right|_{z=1,\theta=0}, \text{ for } 1 \leq j \leq S$$

$$D_{00} = -\left.\frac{\partial \widetilde{G}^{(0,0,\widetilde{j})}(z,\theta)}{\partial\theta}\right|_{z=1,\theta=0}, \text{ for } 0 \leq \widetilde{j} \leq S-1.$$

The first element of the vector D_{10} and D_{00} are mean lengths of a busy period and a busy cycle respectively. With the notation

$$\widetilde{D}_{10} = \left.\frac{\partial \widetilde{G}^{(1,0,j)}(z,\theta)}{\partial z}\right|_{z=1,\theta=0}$$

it follows from equations (5.6) that

$$\widetilde{D}_{10} = -(A_1 + A_0 G)^{-1}(C_2 + A_0 D G_{10}).$$

The first component of the vector $\widetilde{D}_{10}$ is the mean number of service completions in a busy period.

5.3.2 Stationary waiting time distribution in the queue

In this section the LST of waiting time distribution and mean waiting time of a customer in the queue are discussed. The stationary waiting time distribution of the queueing-inventory system is in general, analytically intractable. However, we obtain the LST of the waiting time of a customer in the queue and derive an expression for its mean. First note that an arriving customer will enter into service immediately with probability $\boldsymbol{z_0} = \boldsymbol{x_0}\mathbf{e}$. With probability $1-\boldsymbol{z_0}$ the arriving customer has to wait before getting into service. Any such customer is served only in the normal mode. If the tagged customer joins as the first one, his waiting time would be equal to the service time of the customer in service. Thus in this case the mean waiting time of the customer is $\frac{1}{\mu_v}$ or $\frac{1}{\mu_b}$ depending upon the nature of the service of the customer ahead of the tagged customer. The waiting time may be viewed as the time until absorption in a Markov chain with a highly sparse structure. The state space (that includes the arriving customer in its count) of the Markov chain is given by

$$\widetilde{\Omega} = \{*\} \bigcup \{(i, j) |\ i \geq 2, 1 \leq j \leq Q + s\}. \tag{5.11}$$

The state $*$ indicate that the tagged customer is taken for service. That is, $*$ is obtained by lumping $\{(0, j) |\ 1 \leq j \leq Q + s\}$. Its generator matrix $\boldsymbol{\mathcal{H}}$ is given by

$$\boldsymbol{\mathcal{H}} = \begin{pmatrix} 0 & \boldsymbol{0} & & & \\ \boldsymbol{a} & \widetilde{A}_1 & & & \\ & A_2 & \widetilde{A}_1 & & \\ & & A_2 & \widetilde{A}_1 & \\ & & & \ddots & \ddots \end{pmatrix}, \tag{5.12}$$

where

$$\widetilde{A}_1 = A_1 + \lambda I, \quad \boldsymbol{a} = A_2\mathbf{e}. \tag{5.13}$$

The initial probability vector of $\boldsymbol{\mathcal{H}}$ is denoted by $\boldsymbol{\sigma}$ and in partitioned form it is given by

$$\boldsymbol{\sigma} = (\boldsymbol{z_0}, \boldsymbol{z_2}, \boldsymbol{z_3} \cdots),$$

where $\boldsymbol{z_0} = \boldsymbol{x_0}\mathbf{e}$ and $\boldsymbol{z_i}$ denotes the steady-state probability that an arrival finds the server busy in normal mode and the number of customers in the system including the current arrival is i.

$$\boldsymbol{z_i} = \left(\mathbf{0}_{1\times S}, \boldsymbol{x_i}(1,1), \boldsymbol{x_i}(1,2) \cdots, \boldsymbol{x_i}(1, Q+s)\right), \text{ for } i \geq 2. \tag{5.14}$$

Define $\widetilde{\mathbf{W}}(t),\ t \geq 0$ to be the probability that an arriving customer will enter into service no later than t units of time from his arrival, when the server is in normal mode. We will now derive the LST, $\widetilde{w}(\theta)$, of the stationary waiting time in the queue of an arriving customer during the normal mode of service. Using the structure of $\boldsymbol{\mathcal{H}}$, it can readily be verified that

Theorem 5.3.1. *The LST, $\widetilde{w}(\theta)$, of $W(t)$ is given by*

$$\widetilde{w}(\theta) = \boldsymbol{z_0} + \sum_{i=0}^{\infty} \boldsymbol{z_i}[(\theta I - \widetilde{A}_1)^{-1} A_2]^i (\theta I - \widetilde{A}_1)^{-1} A_2 \mathbf{e}. \tag{5.15}$$

Corollary 5. The mean waiting time μ'_W, in the queue of an arriving customer is given by

$$\mu'_W = [\boldsymbol{z_2}(I-R)^{-1} - \boldsymbol{z_2}\sum_{k=0}^{\infty} R^k P^{k+1} + \boldsymbol{z_2}(I-R)^{-2}\widetilde{P}](I-P+\widetilde{P})^{-1}(-\widetilde{A}_1)^{-1}\mathbf{e}, \tag{5.16}$$

where

$$P = (-\widetilde{A}_1)^{-1} A_2, \quad \widetilde{P} = \mathbf{e}\boldsymbol{p}, \tag{5.17}$$

and $\boldsymbol{p}$ is the invariant probability vector of P. That is,

$$\boldsymbol{p}P = \boldsymbol{p}, \quad \boldsymbol{p}\mathbf{e} = 1. \tag{5.18}$$

Note: In the computation of the mean waiting time μ'_W, we need to evaluate the infinite sum $\sum_{k=0}^{\infty} R^k P^{k+1}$. On noting that P is a stochastic matrix, we get $\boldsymbol{z_2}\sum_{k=0}^{\infty} R^k P^{k+1}\mathbf{e} = 1 - \boldsymbol{z_0}$ and hence in truncating the infinite sum we find N^* such that $|\boldsymbol{z_2}\sum_{k=0}^{N^*} R^k P^{k+1}\mathbf{e} - (1 - \boldsymbol{z_0})| < \epsilon$, where ϵ is a pre-determined sufficiently small quantity.

5.3.3 System performance measures

- Mean number of customers in the system,
$L_s = \sum_{i=1}^{\infty} \left(\sum_{j=1}^{Q+s} i x_i\,(1,j) + \sum_{j=0}^{Q+s} i x_i\,(0,j) \right).$

- Mean inventory level,
$I_m = \sum_{i=0}^{\infty} \sum_{j=1}^{Q+s} j x_i\,(0,j) + \sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} j x_i\,(1,j).$

- Mean number of replenishments per time unit,
$R_r = \beta \left(\sum_{j=0}^{s} \left(\sum_{i=0}^{\infty} x_i\,(0,j) + \sum_{i=1}^{\infty} x_i\,(1,j) \right) \right).$

- Rate of service when the server is in normal mode,
$P_n = \mu_b \left(\sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} x_i\,(1,j) \right).$

- Rate of service when the server is in vacation mode,
$P_v = \mu_v \left(\sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} x_i\,(0,j) \right).$

- Rate at which the server goes to vacation mode,
$$\Gamma = \tfrac{\mu_v}{\lambda+\mu_v+\beta} \left(\sum_{i=1}^{\infty} x_i(0,1) + \sum_{j=1}^{s} x_1(1,j) \right) + \tfrac{\mu_b}{\lambda+\mu_b+\beta} \left(\sum_{i=1}^{\infty} x_i(1,1) + \sum_{j=1}^{s} x_1(1,j) \right) + \tfrac{\mu_v}{\lambda+\mu_v} \left(\sum_{j=s+1}^{Q+s} x_1(1,j) \right) + \tfrac{\mu_b}{\lambda+\mu_b} \left(\sum_{j=s+1}^{Q+s} x_1(1,j) \right).$$

- Rate of vacation realization, $R_v = \theta \left(\sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} x_i\,(0,j) \right).$

- Expected loss rate of customers, $E_{loss} = \lambda \left(\sum_{i=0}^{\infty} x_i\,(0,0) \right).$

- Mean number of customers waiting in the system when inventory is available, $W_{inv} = \left(\sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} i x_i(0,j) + \sum_{i=1}^{\infty} \sum_{j=1}^{Q+s} i x_i(1,j) \right)$.
- Mean number of customers waiting in the system during the stock out period, $\widetilde{W_{inv}} = \left(\sum_{i=1}^{\infty} i x_i(0,0) \right)$.

5.4 Optimization problem

We look for the optimal pair of the values of the control variables. Now for computing the minimal cost and the optimal pair (s, Q) we introduce the cost function $\mathcal{F}(s, Q)$ defined by

$$\mathcal{F}(s, Q) = h.I_m + c_1.E_{loss} + c_2.\widetilde{W_{inv}} + (K + Q.c_3).R_r$$

where K is the fixed cost for placing an order, c_1 is the cost incurred due to loss per customer, c_2 is the waiting cost per unit time per customer during the stock out period, c_3 is the variable procurement cost per item and h is the unit holding cost of inventory for one unit of time. Though we are not able to compute explicitly the optimal values of s and Q, due to the complexity of the cost function, we arrive at these by using numerical procedures. Thus for the following input values of the parameters:
$\lambda = 5, \mu_v = 3, \mu_b = 10, \beta = 3, K = \$500, h = \$5, c_1 = \$100, c_2 = \$50$ and $c_3 = \$50$ we get the optimal pair (s, Q) as $(4, 15)$ and the corresponding minimum cost is \$134.9468.

Chapter 6

Retrial of unsatisfied customers in a queueing-inventory system

6.1 Introduction

In chapters 2 through 5 we assumed that customers join an infinite capacity waiting station on arrival, if the server is busy. If the server is idle and at least one item is in the inventory the arriving customer enters for service immediately. If customers upon arrival encounter an idle server with no inventory, then it does not join the system and is lost for ever. In the present chapter we consider $M/M/1/1$ queueing-inventory system with service time where, on arrival, if a customer encounters a busy server, proceeds to an orbit of infinite capacity. In the orbit a queue of customers is formed. The

Some results of this chapter are included in the following paper.
A. Krishnamoorthy, R. Manikandan and Sajeev S. Nair : Retrial of unsatisfied customers in a queueing-inventory system (Under review).

head of the queue retries to access an idle server with at least one item in the inventory, failing which it goes back to orbit and occupies the first position in the queue. The inter retrial times are exponentially distribution with parameter θ, independent of the number of customers in the orbit, provided there is at least one. With arrival of customers according to a Poisson process of rate λ_2, service time exponentially distributed with parameter μ_2 and lead time for replenishment of inventory following exponential distribution with parameter β_2, the process $\{(\mathcal{N}_2(t), \mathcal{C}(t), \mathcal{I}_2(t))|t \geq 0\}$, forms a CTMC on the state space Ω_2 given by

$$\Omega_2 = \left((\mathbb{Z}_+ \bigcup \{\mathbf{0}\}) \times \{0,1\} \times \{1,2,\ldots,S_2\}\right) \bigcup \left((\mathbb{Z}_+ \bigcup \{\mathbf{0}\}) \times \{\mathbf{0}\} \times \{\mathbf{0}\}\right).$$

Retrial of unsatisfied customers is extensively discussed in queueing literature (see Falin and Templeton [19], Artalejo and Gomez Corral [4]). However, in the context of inventory with retrial of unsatisfied customers, not much work is reported, especially those involving positive service time. The negligible service time case is discussed in Ushakumari [73] and Artalejo *et al.* [3]. Whereas the former provides analytical solution (for the case of constant retrial), the latter provides an algorithmic approach in a more general set up (linear retrial rate). Those involving positive service time also has limited literature (see for example Krishnamoorthy *et al.* [40], Cui and Wang [15] and Padmavathi *et al.* [49]). A few other references are also provided in chapter 1.

This chapter is arranged as follows. Section 6.2 deals with the mathematical formulation of the problem. In Section 6.3 the condition for stability of the system is investigated, followed by the computation of the steady-state probability vector. Performance measures are provided in Section 6.4. In particular we compute the expected waiting time of a customer in the orbit, distribution of time until the first customer goes to orbit (during a cycle that is appropriately defined) and probability of no customer going to orbit in a given interval of time. Section 6.5 discusses an optimization problem. In Section 6.6 we analyze briefly a tandem queueing-inventory network.

6.2 Mathematical formulation of the problem

With arrival constituting a Poisson process of rate λ_2, service time independent identically distributed exponential random variables with parameter μ_2, lead time for replenishment having exponential distribution with parameter β_2 and inter-retrial time of head of the queue in the orbit following exponential distribution with parameter θ, the process $\{(\mathcal{N}_2(t), \mathcal{C}(t), \mathcal{I}_2(t))|t \geq 0\}$ forms a CTMC on the state space Ω_2 described in the introduction. It is to be noted that we make a strong assumption on customers getting into the system: when inventory level is zero, no customer joins the system. The replenishment policy followed is (s_2, Q_2) (This notation is needed since towards the end of this chapter we examine a queueing-inventory network with the first station having the classical $M/M/1/\infty$ pattern, whereas the second station has retrial component attached to it). Further, as considered in all previous chapters it is assumed here also that at the end of a service a customer is provided one unit of the item with probability γ. We expected "the assumption that no customer joins when inventory is zero" would enable us to arrive at, in the least, a closed form solution of the system state distribution, if not decomposition of the system. Nevertheless it turned out to be otherwise. Thus we are forced to adopt algorithmic approach for the analysis of the system described.

The state space of the CTMC is partitioned in to levels $\mathcal{L}(i)$ defined as

$$\mathcal{L}(i) = \{(0,0,j)\,|1 \leq j \leq s_2 + Q_2\} \cup \{(i,k,j)\,/i \geq 1; k = 0,1; 0 \leq j \leq s_2 + Q_2\}.$$

The transitions in the Markov chain are listed below:
(a) Transitions due to arrival of customers :

$(i,0,j) \to (i,1,j)$: the rate is λ_2, for $i \geq 0;\ 1 \leq j \leq S_2$.

$(i,1,j) \to (i+1,1,j)$: the rate is λ_2, for $i \geq 0;\ 1 \leq j \leq S_2$.

(b) Transitions due to service completion of customers:

$(i,1,j) \to (i,0,j-1)$: the rate is $\gamma\mu_2$, for $i \geq 0;\ 1 \leq j \leq S_2$.

$(i,1,j) \to (i,0,j)$: the rate is $(1-\gamma)\,\mu_2$, for $i \geq 0;\ 1 \leq j \leq S_2$.

(c) Transitions due to replenishments:

$(i,0,j) \to (i,0,Q_2+j)$: the rate is β_2, for $i \geq 0;\ 0 \leq j \leq s_2$.

$(i,1,j) \to (i,1,Q_2+j)$: the rate is β_2, for $i \geq 0;\ 0 \leq j \leq s_2$.

(d) Transitions due to retrial of customers:

$(i,0,j) \to (i-1,1,j)$: the rate is θ, for $i \geq 1;\ 1 \leq j \leq S_2$.

All other transition pairs have rate zero. The infinitesimal generator $\boldsymbol{\mathcal{W}}$ of this CTMC is given by

$$\mathcal{W} = \begin{bmatrix} \widehat{B_0} & \widehat{B_1} & & & \\ \widehat{B_2} & \widehat{A_1} & \widehat{A_0} & & \\ & \widehat{A_2} & \widehat{A_1} & \widehat{A_0} & \ldots \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

where $\widehat{B_0}$, $\widehat{B_1}$ and $\widehat{B_2}$ contains transition rates within $\mathcal{L}(0)$, transition from $\mathcal{L}(0)$ to $\mathcal{L}(1)$ and transition from $\mathcal{L}(1)$ to $\mathcal{L}(0)$ respectively; $\widehat{A_0}$ represents the transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i+1)$, $i \geq 1$; $\widehat{A_1}$ represents the transitions within $\mathcal{L}(i)$ for $i \geq 1$, and $\widehat{A_2}$ represents transitions from $\mathcal{L}(i)$ to $\mathcal{L}(i-1)$, $i \geq 2$. All these matrices are square matrices of order $2S_2 + 1$.

6.3 System stability and computation of steady-state probability vector

The Markov chain under consideration is a LIQBD process. For this chain to be stable it is necessary and sufficient that

$$\boldsymbol{\xi}\widehat{A_0}\mathbf{e} < \boldsymbol{\xi}\widehat{A_2}\mathbf{e}. \tag{6.1}$$

where $\boldsymbol{\xi}$ is the unique non negative vector satisfying,

$$\boldsymbol{\xi}\widehat{A} = 0, \ \boldsymbol{\xi}\mathbf{e} = 1 \tag{6.2}$$

and $\widehat{A} = \widehat{A_0} + \widehat{A_1} + \widehat{A_2}$, is the infinitesimal generator of the finite state CTMC. Let $\boldsymbol{\xi} = (\xi_0(0), \xi_0(1), \ldots, \xi_0(S_2), \xi_1(1), \xi_1(2), \ldots \xi_1(S_2))$ be the steady-state vector of the generator matrix $\widehat{A}$. Then $\boldsymbol{\xi}\widehat{A} = 0$ gives the following equations

$$-\beta_2\xi_0(0) + \gamma\mu_2\xi_1(1) = 0 \tag{6.3}$$

$$-(\lambda_2 + \theta + \beta_2)\xi_0(i) + (1-\gamma)\mu_2\xi_1(i) + \gamma\mu_2\xi_1(i+1) = 0, \ 1 \leq i \leq s_2 \tag{6.4}$$

$$-(\lambda_2+\theta)\xi_0(i)+(1-\gamma)\mu_2\xi_1(i)+\gamma\mu_2\xi_1(i+1) = 0,\ s_2+1 \le i \le Q_2-1 \quad (6.5)$$

$$\beta_2\xi_0(i)-(\lambda_2+\theta)\xi_0(Q_2+i)+(1-\gamma)\mu_2\xi_1(Q_2+i)+\gamma\mu_2\xi_1(Q_2+i+1) = 0,\ 0 \le i \le s_2-1 \quad (6.6)$$

$$\beta_2\xi_0(s_2) - (\lambda_2+\theta)\xi_0(S_2) + (1-\gamma)\mu_2\xi_1(S_2) = 0 \quad (6.7)$$

$$(\lambda_2+\theta)\xi_0(i) - (\beta_2+\mu_2)\xi_1(i) = 0,\ 1 \le i \le s_2 \quad (6.8)$$

$$(\lambda_2+\theta)\xi_0(i) - \mu_2\xi_1(i) = 0,\ s_2+1 \le i \le Q_2 \quad (6.9)$$

$$\beta_2\xi_1(i) + (\lambda_2+\theta)\xi_0(Q_2+i) - \mu_2\xi_1(Q_2+i) = 0,\ 1 \le i \le s_2 \quad (6.10)$$

The LIQBD process with infinitesimal generator $\boldsymbol{\mathcal{W}}$ is stable if and only if $\boldsymbol{\xi}\widehat{A_0}\mathbf{e} < \boldsymbol{\xi}\widehat{A_2}\mathbf{e}$. That is,

$$\iff \theta\left(\xi_0(1)+\xi_1(2)+\cdots+\xi_0(S_2)\right) > \lambda_2\left(\xi_1(1)+\xi_1(2)+\cdots+\xi_1(S_2)\right).$$

$$\iff \theta\left(\xi_0(1)+\xi_1(2)+\cdots+\xi_0(S_2)\right) > \lambda_2\left(\frac{\lambda_2+\theta}{\mu_2}\right)\left(\xi_0(1)+\xi_1(2)+\cdots+\xi_0(S_2)\right)$$

$$\iff \theta > \lambda_2\left(\frac{\lambda_2+\theta}{\mu_2}\right)$$

$$\iff \frac{\lambda_2}{\mu_2} < \frac{\theta}{\lambda_2+\theta}.$$

Thus we have the following lemma for the stability of the second station:

Lemma 6.3.1. The CTMC Ω_2 is stable if and only if $\lambda_2 < \frac{\mu_2\theta}{\lambda_2+\theta}$.

Now we compute the steady-state probability vector of $\boldsymbol{\mathcal{W}}$ under the stability condition. Let $\boldsymbol{y}$ denote the steady-state probability vector of the infinitesimal generator $\boldsymbol{\mathcal{W}}$. Then the steady-state probability vector must satisfy the relations,

$$\boldsymbol{y}\boldsymbol{\mathcal{W}} = 0,\ \boldsymbol{y}\mathbf{e} = 1. \quad (6.11)$$

Let us partition $\boldsymbol{y}$ by levels as

$$\boldsymbol{y} = (\boldsymbol{y_0}, \boldsymbol{y_1}, \boldsymbol{y_2}, \dots), \tag{6.12}$$

where the subvectors of $\boldsymbol{y}$ are further partitioned as,

$$\boldsymbol{y_i} = (y_i(0,0), y_i(0,1), y_i(0,2), ..., y_i(0,S_2), y_i(1,1), y_i(1,2), ..., y_i(1,S_2)), \; i \geq 0. \tag{6.13}$$

Since the state space Ω_2 is a LIQBD process, its steady-state vector is given by

$$\boldsymbol{y_i} = \boldsymbol{y_0} R^i, \; i \geq 1. \tag{6.14}$$

(see Neuts [47]), where R is the minimal non-negative solution to the matrix quadratic equation $R^2 + R\widehat{A_1} + \widehat{A_0} = 0$. For finding the boundary vectors $\boldsymbol{y_0}$ and $\boldsymbol{y_1}$, we have from $\boldsymbol{y}\boldsymbol{\mathcal{W}}$=0,

$$\boldsymbol{y_0}\widehat{B_1} + \boldsymbol{y_1}\widehat{A_1} + \boldsymbol{y_2}\widehat{A_2} = 0$$

$$\Longleftrightarrow \boldsymbol{y_0}\widehat{B_1} + \boldsymbol{y_1}\left(\widehat{A_1} + R\widehat{A_2}\right) = 0$$

$$\Longleftrightarrow \boldsymbol{y_1} = -\boldsymbol{y_0}\widehat{B_1}\left(\widehat{A_1} + R\widehat{A_2}\right)^{-1}$$

$$\Longleftrightarrow \boldsymbol{y_1} = \boldsymbol{y_0} D, \; \text{where } D = -\widehat{B_1}\left(\widehat{A_1} + R\widehat{A_2}\right)^{-1}.$$

Further,

$$\boldsymbol{y_0}\widehat{B_0} + \boldsymbol{y_1}\widehat{B_2} = 0$$

$$\Longleftrightarrow \boldsymbol{y_0}\left(\widehat{B_0} + D\widehat{B_2}\right) = 0.$$

First we take $\boldsymbol{y_0}$ as the stady state vector of the generator matrix $\widehat{B_0}+D\widehat{B_2}$. Then $\boldsymbol{y_i}$, for $i \geq 1$, can be found using the formula $\boldsymbol{y_1} = \boldsymbol{y_0}D$ and $\boldsymbol{y_i} = \boldsymbol{y_1}R^{i-1}$, for $i \geq 2$. Finally, the steady-state probability distribution of the system under study is obtained by dividing each $\boldsymbol{y_i}$ with normalizing condition

$$\boldsymbol{y_0}\mathbf{e} + (\boldsymbol{y_1} + \boldsymbol{y_2} + \dots)\,\mathbf{e} = \boldsymbol{y_0}\left(I + D\,(I - R)^{-1}\right)\mathbf{e}$$

Once the matrix R is obtained, the steady-state probability vector $\boldsymbol{y}$ can be computed by exploiting the special structure of the coefficient matrices. We can use logarithmic reduction algorithm for computing R. We will list only the main steps involved in the logarithmic reduction algorithm for computing R.

Logarithmic Reduction Algorithm for R:

Step 0: $H \leftarrow (-\widehat{A}_1)^{-1}\widehat{A}_0$, $L \leftarrow (-\widehat{A}_1)^{-1}\widehat{A}_2$, $G = L$, and $T = H$.
Step 1:

$$\begin{gathered} U = HL + LH \\ M = H^2 \\ H \leftarrow (I - U)^{-1}M \\ M \leftarrow L^2 \\ L \leftarrow (I - U)^{-1}M \\ G \leftarrow G + TL \\ T \leftarrow TH \end{gathered}$$

Continue Step 1 until $||\mathbf{e} - G\mathbf{e}||_\infty < \epsilon$.

Step 2: $R = -\widehat{A}_0(\widehat{A}_1 + \widehat{A}_0 G)^{-1}$.

6.4 Performance measures

6.4.1 Expected waiting time of a customer in the orbit

For computing the expected waiting time in the orbit of a tagged customer who joins as r^{th} customer in the orbit, we consider the CTMC, $\boldsymbol{\Psi_1} = \left\{(\widehat{\mathcal{N}_2(t)}, \mathcal{C}(t), \mathcal{I}_2(t))|t \geq 0\right\}$ where $\widehat{\mathcal{N}_2(t)}$ denotes the rank, which is the position of the tagged customer in the orbit at the time he joins the system. The state space of the CTMC $\boldsymbol{\Psi_1}$ is given by
$\Im_1 = \{(i, 0, m),\ 1 \leq i \leq r;\ 0 \leq m \leq S_2\} \bigcup \{\Delta_1\}$, where $\{\Delta_1\}$ is an absorbing state which corresponds to the tagged customer being taken for service. The infinitesimal generator of the chain $\boldsymbol{\Psi_1}$ is given by
$\boldsymbol{\mathcal{H}_1} = \begin{bmatrix} \boldsymbol{\mathcal{G}_1} & \boldsymbol{\mathcal{G}_1}^{\mathbf{0}} \\ \mathbf{0} & 0 \end{bmatrix}$, where $\boldsymbol{\mathcal{G}_1}^{\mathbf{0}}$ is an $\{(r-1)(2S_2+1)+S_2\} \times 1$ matrix

such that $\boldsymbol{\mathcal{G}_1}^{\mathbf{0}}(i,1) = \theta$, for $1 \leq i \leq S_2$ and $\boldsymbol{\mathcal{G}_1} = \begin{bmatrix} B & 0 & 0 & ... & ... & 0 \\ \tilde{A}_2 & B & 0 & ... & ... & 0 \\ 0 & \tilde{A}_2 & B & ... & ... & 0 \\ & & \ddots & \ddots & & \\ & & & & \tilde{\tilde{A}}_2 & \tilde{B} \end{bmatrix}$,

where $B = \begin{bmatrix} B_1 & 0 & B_2 & 0 & 0 & 0 \\ 0 & B_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & B_6 & 0 & 0 & 0 \\ B_8 & 0 & 0 & B_9 & 0 & B_{10} \\ B_3 & B_{11} & 0 & 0 & B_{12} & 0 \\ 0 & B_5 & B_{13} & 0 & 0 & B_{14} \end{bmatrix}$ with $B_1 = \begin{bmatrix} -\beta_2 & 0 \\ 0 & -(\beta_2+\theta)I_s \end{bmatrix}$,

$$B_2 = \beta_2 I_{s+1},\ B_3 = \begin{bmatrix} 0 & \gamma\mu_2 \\ 0 & 0 \end{bmatrix}_{(S-2s-1)\times(s+1)},\ B_4 = -\theta I_{S-2s-1},$$

$$B_5 = \begin{bmatrix} 0 & \gamma\mu_2 \\ 0 & 0 \end{bmatrix}_{(s+1)\times(S-2s)},\ B_6 = -\theta I_{s+1},$$

$$B_8 = \begin{bmatrix} \gamma\mu_2 & (1-\gamma)\mu_2 & & & & \\ & \gamma\mu_2 & (1-\gamma)\mu_2 & & & \\ & & \ddots & \ddots & & \\ & & & \gamma\mu_2 & (1-\gamma)\mu_2 \end{bmatrix}_{s\times(s+1)}, B_9 = -\beta_2 I_s,$$

$$B_{11} = \begin{bmatrix} (1-\gamma)\mu_2 & & & \\ \gamma\mu_2 & (1-\gamma)\mu_2 & & \\ & \ddots & \ddots & \\ & & \gamma\mu_2 & (1-\gamma)\mu_2 \end{bmatrix}_{(S-2s-1)\times(S-2s-1)},$$

$$B_{13} = \begin{bmatrix} (1-\gamma)\mu_2 & & & \\ \gamma\mu_2 & (1-\gamma)\mu_2 & & \\ & \ddots & \ddots & \\ & & \gamma\mu_2 & (1-\gamma)\mu_2 \end{bmatrix}_{(s+1)\times(s+1)},$$

$B_{10} = \begin{bmatrix} 0 & \beta_2 I_s \end{bmatrix}_{s\times(s+1)}$, $B_{12} = -\mu_2 I_{S-2s-1}$, $B_{14} = -\mu_2 I_{s+1}$,
$\tilde{B}(i,j) = B(i+1, j+1)$ for $1 \le i,\ j \le 2S$;
$\tilde{\tilde{A}}(i,j) = \tilde{\tilde{A}}_2(i+1, j)$ for $1 \le i \le 2S,\ 1 \le j \le 2S+1$;

$$\tilde{A}_2 = \begin{bmatrix} 0 & 0 & 0 & F_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & F_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & F_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ with } F_1 = \begin{bmatrix} 0 \\ \theta I_s \end{bmatrix}_{(s+1)\times s},$$

$F_2 = \theta I_{Q-s}$ and $F_3 = \theta I_{(s+1)\times(s+1)}$.

Now the waiting time distribution $\mathcal{W}^r$ of the tagged customer who joins as the r^{th} customer in the orbit, is the time until absorption in the CTMC $\mathbf{\Psi_1}$, and given by the column vector

$$\mathcal{W}^r = \hat{I}_{2S_2}(-\boldsymbol{\mathcal{G}_1}^{-1})\mathbf{e},$$

where $\hat{I}_{2S_2} = [0 \quad I_{2S_2}]_{(2S_2)\times\{(r-1)(2S_2+1)+S_2\}}$. Hence, the expected waiting time of a general customer is given by

$$E(\mathcal{W}_L) = \sum_{r=1}^{\infty} \hat{\pi}_r \mathcal{W}^r,$$

where $\hat{\pi}_r$ is a $1 \times 2S_2$ dimensional row vector defined by

$$\hat{\pi}_r(i) = \hat{\pi}_r(i+1), \text{ for} 1 \leq i \leq 2S_2.$$

In a similar manner, we can find the second moment of the waiting time of an orbital customer as

$$E(\mathcal{W}_L^2) = \sum_{r=1}^{\infty} \hat{\pi}_r \mathcal{W}_2^r,$$

where $\mathcal{W}_2^r = 2\hat{I}_{2S_2}(\boldsymbol{\mathcal{G}_1}^{\mathbf{-2}})\mathbf{e}$ (see Neuts [47]).

6.4.2 Distribution of the time until the first customer goes to the orbit

We now compute the distribution of the time till the first customer in a cycle goes to orbit. By a cycle we shall mean that starting with no customer in orbit, until the next epoch when all customers in the orbit are served out. We also assume that at the beginning of a cycle the inventory level is S_2 and there is no customer in service. A customer arrives and straight enters for service. During this service time, if another customer arrives, then he is the first to go to orbit. Let χ denote the random variable "time until the first customer goes to orbit in a cycle".

We consider the CTMC $\mathbf{\Psi_2} = \{(\mathcal{C}(t), \mathcal{I}_2(t)) | t \geq 0\}$, where $\mathcal{C}(t)$ and $\mathcal{I}_2(t)$ are same as defined in the beginning of this chapter. The state space of this CTMC $\mathbf{\Psi_2}$ is

$$\Im_2 = \{\mathbf{0}\} \bigcup \{(\ell, m) |\ \ell = 0, 1;\ 1 \leq m \leq S_2\} \bigcup \{\Delta_2\}$$

where $\{\Delta_2\}$ is the absorbing state which represents the state "first customer to go to orbit" from the state $\{(1,m)|\ 1 \le m \le S_2\}$. Clearly, $\Im_2$ is a finite state space Markov chain. The possible transitions and the corresponding rates are given in Table 6.1.

Table 6.1: The transitions in the CTMC $\boldsymbol{\Psi_2}$ and corresponding rates

Form	To	Rate	
$(0,0)$	$(0,0)$	$-\beta_2$	
$(0,m)$	$(0,m)$	$-(\beta_2+\lambda_2)$	$m=1,2,\ldots,s_2.$
$(0,m)$	$(0,m)$	$-\lambda_2$	$m=s_2+1,s_2+2,\ldots,S_2.$
$(1,m)$	$(1,m)$	$-(\lambda_2+\mu_2+\beta_2)$	$m=1,2,\ldots,s_2.$
$(1,m)$	$(1,m)$	$-(\lambda_2+\mu_2)$	$m=s_2+1,s_2+2,\ldots,S_2.$
$(1,m)$	$(0,m-1)$	μ_2	$m=1,2,\ldots,S_2.$
(ℓ,m)	$(\ell,m+Q_2)$	β_2	$\ell=0,1;\ m=0,1,\ldots,s_2.$
$(0,m)$	$(1,m)$	λ_2	$m=1,2,\ldots,S_2.$
$(1,m)$	$\{\Delta_2\}$	λ_2	$m=1,2,\ldots,S_2.$

Thus the infinitesimal generator $\boldsymbol{\mathcal{H}_2}$ of the Markov chain $\boldsymbol{\Psi_2}$ is of the form $\boldsymbol{\mathcal{H}_2} = \begin{bmatrix} \boldsymbol{\mathcal{G}_2} & \boldsymbol{\mathcal{G}_2}^{\mathbf{0}} \\ \mathbf{0} & 0 \end{bmatrix}$ with initial probability vector $\boldsymbol{\alpha} = (0,0,\ldots,1,0)$ where 1 is the in the ${S_2}^{th}$ position; $\boldsymbol{\mathcal{G}_2}$ is of order $2S_2{+}1$; $\boldsymbol{\mathcal{G}_2}^{\mathbf{0}}$ is a $2S_2{+}1$ component column vector such that $\boldsymbol{\mathcal{G}_2}\mathbf{e}{+}\boldsymbol{\mathcal{G}_2}^{\mathbf{0}}{=}0$. Let χ represent the random variable "time till first customer goes to orbit". This time duration follows PH distribution with representation $(\boldsymbol{\alpha},\boldsymbol{\mathcal{G}_2})$. Therefore the expected time until the first customer goes to the orbit is

$$E(\chi) = -\boldsymbol{\alpha}\left(\boldsymbol{\mathcal{G}_2^{-1}}\right)\mathbf{e}.$$

6.4.3 Probability that all customer arrivals (demands) in a time duration of length t do not go to the orbit

Consider an interval of duration t in the steady-state regime. The objective is to compute the probability that no customer arriving during this time period goes to orbit. This means that all customer arrivals in this interval either meet an idle server with positive inventory or during the stock out period. Thus customers do not join the second station when the inventory level is zero (by model assumption). Assume that there is no customer in the orbit at the beginning of this interval. Choose n_1 of the arrivals to find positive inventory and server idle. The remaining $n - n_1$ are chosen such that upon their arrival the server is found to be idle with no item in the inventory. Then the required probability P_t is given by

$$P_t = \sum_{n=1}^{\infty} \mathcal{J} \begin{pmatrix} n \\ n_1 \end{pmatrix} \sum_{n_1=1}^{n} \sum_{i=1}^{n_1} \left(\sum_{j=1}^{Q_2+s_2} y_0(0,j) \left(1 - e^{-\mu(x_i - x_{i-1})}\right) \right)^{n_1} (y_0(0,0))^{n-n_1}$$

where $\mathcal{J} = \left(\frac{n!}{t^n} \int\limits_0^{s_1} \cdots \int\limits_{x_{n-2}}^{s_{n-1}} \int\limits_{x_{n-1}}^{s_n} dx_n \cdots dx_1 \right)$.

6.4.4 Other performance measures

- Mean number of customers in the orbit,
  $L_O = \left(\sum\limits_{i=1}^{\infty} \sum\limits_{j=0}^{Q_2+s_2} i y_i(0,j) + \sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{Q_2+s_2} i y_i(1,j) \right)$.
- Mean inventory level, $E_{inv} = \sum\limits_{i=0}^{\infty} \sum\limits_{j=0}^{Q_2+s_2} j y_i(0,j) + \sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{Q_2+s_2} j y_i(1,j)$.
- Depletion rate of inventory, $D_{inv} = \gamma\mu_2 \left(\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{Q_2+s_2} y_i(1,j) \right)$.

- Mean number of replenishments per unit time,
  $R_r = \beta_2 \left(\sum_{j=0}^{s_2} \left(\sum_{i=0}^{\infty} y_i(0,j) + \sum_{i=1}^{\infty} y_i(1,j) \right) \right).$
- Expected loss rate of customers, $E_{loss} = \lambda_2 \left(\sum_{i=1}^{\infty} y_i(0,0) \right).$
- Probability that the server is busy, $P_{busy} = \sum_{i=0}^{\infty} \sum_{j=1}^{Q_2+s_2} y_i(1,j).$
- Successful rate of retrials, $E_{retrial} = \theta \left(\sum_{i=1}^{\infty} \sum_{j=1}^{Q_2+s_2} y_i(0,j) \right).$
- Mean number of departures per unit time,
  $D_m = \mu_2 \left(\sum_{i=0}^{\infty} \sum_{j=1}^{Q_2+s_2} y_i(1,j) \right).$
- Mean number of customers waiting in the orbit when inventory is available, $\widetilde{W_O} = \left(\sum_{i=1}^{\infty} \sum_{j=1}^{Q_2+s_2} i y_i(0,j) + \sum_{i=1}^{\infty} \sum_{j=1}^{Q_2+s_2} i y_i(1,j) \right).$
- Mean number of customers waiting in the orbit during the stock out period, $\widetilde{\widetilde{W_O}} = \left(\sum_{i=1}^{\infty} i y_i(0,0) \right).$

6.5 Optimization problem

In this section we provide the optimal values of the inventory level s_2 and the fixed order quantity Q_2 of the model. For checking the optimality of s_2 and Q_2, the following cost function is constructed. Define $\mathcal{F}(s_2, Q_2)$ as the expected total cost per unit time in the long run. Then

$$\mathcal{F}(s_2, Q_2) = h.E_{inv} + c_1.E_{loss} + c_2.(1 - P_{busy}) + (K + Q_2.c_3).R_r$$

where K is the fixed cost for placing an order, c_1 is the cost incurred due to loss per customer, c_2 is the waiting cost per unit time per customer during the stock out period, c_3 is the variable procurement cost per item and h is the unit holding cost of inventory for one unit of time. Table 6.2 provides the optimal pair (s_2, Q_2) and the corresponding minimum cost (in Dollars). Here γ is varied from 0.1 to 1, at an interval of 0.1. The values for the input parameters are given as follows $\lambda_2 = 2,\ \mu_2 = 5,\ \theta = 4,\ \beta_2 = 3,\ K = \$500,\ c_1 = \$25, c_2 = \$50, c_3 = \$35, h = \$3.5$. We provide a numerical comparison based on a few performance measures in Table 6.3.

Table 6.2: Optimal (s_2, Q_2) pair and minimum cost

γ	0.1	0.2	0.3	0.4	0.5
Optimal (s_2, Q_2) pair & minimum cost	(1,29) 242.353	(1,29) 241.978	(1,29) 241.585	(1,29) 241.181	(1,29) 240.767
γ	0.6	0.7	0.8	0.9	1
Optimal (s_2, Q_2) pair & minimum cost	(1,29) 240.347	(1,29) 239.922	(1,29) 239.494	(1,29) 239.062	(1,29) 238.629

For numerical comparison we assign the same input values as for Table 6.2 with $s_2 = 10$ and $S_2 = 31$. For example we observe from Table 6.3 that the mean number of replenishments and loss rate of customer is larger for $\gamma = 1$ compared to that for γ $(= 0.5)$. Further P_{busy} and E_{inv} are higher for $\gamma = 0.5$ compared to that for $\gamma = 1$. These are all on expected lines.

6.6 Tandem queueing-inventory network

Now we assume that the model discussed so far in this chapter is the second station in a tandem queueing-inventory network. The first station follows $M/M/1/\infty$ queueing-inventory. Thus arrival process to this forms a Pois-

Table 6.3: Effect of γ on various performance measures

Performance measures	with $\gamma = 0.5$	with $\gamma = 1$ (classical queueing-inventory system)
P_{busy}	0.39999911	0.39999864
E_{inv}	20.6666431	20.3333149
D_{inv}	0.14285707	0.14285702
R_r	0.03213748	0.06427498
L_O	1.09998846	1.09998834
E_{loss}	0.00000070	0.00000286

son process of rate λ_1, service times are independent identically distributed exponential random variables with parameter μ_1. We follow (s_1, Q_1) policy for inventory replenishment. The distribution for replenishment time is exponential with parameter β_1. As done in all our earlier discussions throughout this thesis, here also we make the crucial assumption that no customer joins when inventory in this station is empty. As obtained in Schwarz *et al.* [64] or in Krishnamoorthy and Viswanath [42], we have stochastic decomposition property of the system state holding for station one. That is, $P(\mathcal{N}_1 = i,\ \mathcal{I}_1 = i_1) = P(\mathcal{N}_1 = i)\,.\,P(\mathcal{I}_1 = i_1)$ where $\mathcal{N}_1$ is the number of customers and $\mathcal{I}_1$ the number of items in the inventory in station 1 in the steady state. In other words, $P(\mathcal{N}_1 = i)\,.\,P(\mathcal{I}_1 = i_1) = \left(1 - \frac{\lambda_1}{\mu_1}\right)\left(\frac{\lambda_1}{\mu_1}\right)^i . \boldsymbol{\pi}$ where $\boldsymbol{\pi} = (\pi_0, \pi_1, \ldots, \pi_{S_1})$ with $\pi_{i_1} = P(\mathcal{I}_1 = i_1),\ i_1 = 0, 1, \ldots, S_1$. Explicit expression for π_{i_1} is given by:

$$\pi_{i_1} = \begin{cases} \left[1 + Q_1 \frac{\beta_1}{\gamma\lambda_1}\left(\frac{\beta_1+\gamma\lambda_1}{\gamma\lambda_1}\right)^{s}_{1}\right]^{-1},\ i_1 = 0. \\ \frac{\beta_1}{\gamma\lambda_1}\left(\frac{\beta_1+\gamma\lambda_1}{\gamma\lambda_1}\right)^{i_1-1}\pi_0,\ i_1 = 1, 2, \cdots, s_1. \\ \frac{\beta_1}{\gamma\lambda_1}\left(\frac{\beta_1+\gamma\lambda_1}{\gamma\lambda_1}\right)^{s_1}\pi_0,\ i_1 = s_1+1, s_1+2, \cdots, Q_1. \\ \frac{\beta_1}{\gamma\lambda_1}\left(\frac{\beta_1+\gamma\lambda_1}{\gamma\lambda_1}\right)^{i_1-Q_1-1}\left(\left(\frac{\beta_1+\gamma\lambda_1}{\gamma\lambda_1}\right)^{s_1-(i_1-Q_1-1)} - 1\right)\pi_0, \\ \quad i_1 = Q_1+1, Q_1+2, \cdots, S_1. \end{cases} \tag{6.15}$$

The output of station 1 is Poisson of rate $\lambda_1(1 - \pi_0)$ by Burkes theorem (see [13]). This is fed into station 2 that was described in the earlier part of this chapter. Since no served customer is blocked in station 1 (except for want of inventory), the two stations behave almost like two independent stations, except that the output from station 1 flows to station 2. This flow to station 2 is prevented when inventory in that station is zero. Thus the effective inflow of customers to station 2 is $[\lambda_1 (1-\pi_0) \times$ probability that station 2 has inventory$]$ which we designate as λ_2. With this we can write combined system state distribution as the product of the probability distribution of the status of station 1 $\times$ probability distribution of the status of station 2. Let $\mathcal{N}_1(t)$ is the umber of customers in station 1, $\mathcal{N}_2(t)$ is the number of customers in the orbit of station 2, $\mathcal{I}_i(t)$ is the number of inventoried items in station i $(= 1, 2)$ and for station 2, $\mathcal{C}(t)$ is the status of the server at time t, that is $\mathcal{C}(t) = \begin{cases} 0, \text{ if server is idle at time t.} \\ 1, \text{ if server is busy at time t.} \end{cases}$

The combined system $\{(\mathcal{N}_1(t), \mathcal{I}_1(t), \mathcal{N}_2(t), \mathcal{I}_2(t), \mathcal{C}(t)), t \geq 0\}$ is a CTMC with state space

$$\{(n_1, i_1, n_2, i_2, k) \,| n_1, n_2 \geq 0; 0 \leq i_1 \leq s_1 + Q_1; 0 \leq i_2 \leq s_2 + Q_2; k = 0, 1\}.$$

Assume that the whole system is stable. For station 1 to be stable it is necessary and sufficient that $\lambda_1 < \mu_1$. The stability condition for station 2 is given by $\lambda_2 < \frac{\mu_2 \theta}{\lambda_2 + \theta}$. Both these conditions should hold in order for the combined system to be stable. We write $\mathcal{X}$ for $\lim_{t \to 0} \mathcal{X}(t)$. Thus under the condition that the whole system is stable, the probability distribution of the system state is given by,

$$P\{\mathcal{N}_1 = n_1, \mathcal{I}_1 = i_1; \mathcal{N}_2, \mathcal{I}_2 = i_2, \mathcal{C} = 0\} = \left(1 - \frac{\lambda_1}{\mu_1}\right)\left(\frac{\lambda_1}{\mu_1}\right)^{n_1} \pi_{i_1} y_{n_2}(0, i_2)$$

(obviously $\mathcal{C} = 0$ for $i_2 = 0$) and

$$P\{\mathcal{N}_1 = n_1, \mathcal{I}_1 = i_1; \mathcal{N}_2, \mathcal{I}_2 = i_2, \mathcal{C} = 1\} = \left(1 - \frac{\lambda_1}{\mu_1}\right)\left(\frac{\lambda_1}{\mu_1}\right)^{n_1} \pi_{i_1} y_{n_2}(1, i_2),\ i_2 > 0.$$

Concluding remarks:

This thesis was an attempt to arrive at product form solution for queueing-inventory problems. In chapters 2 and 3 we succeeded in achieving that. In chapter 4 (multi-server system), we could achieve product form solution only in the case when number of servers is restricted to 2. In remaining part of that chapter and the rest of the chapters we were forced to satisfy with algorithmic solution. Performance measures of significance were computed for all models discussed. We wish to highlight quite a few distributions that we derived in this thesis. Though these are of great significance, it is surprising that no attempt was made earlier to derive such distributions. In the context of storage systems (continuous state space) first emptiness probability is discussed. However, for the discrete state space system this distribution was not seen to be studied earlier.

We have to admit the fact that exponentially distributed service/ production time are not very common. Nevertheless, there are a few cases where it works. Despite the assumption that customers do not join when inventory level is zero in a retrial queue, we are not able to produce product form solution. We did notice in our attempt towards this end that we can have analytic solution if we proceed on the lines of the 2-server queueing-inventory problem discussed in chapter 4. However, we have not reported that in this thesis.

There are several avenues for future studies based on this thesis. Introducing vacation to server when either no customer in the system or inventory is empty is one possibility. Also the case of server breakdown/ production mechanism breakdown could be studied. In all these cases we do not expect closed form solution. Moving from exponential distribution to more complex distributions enhances applicability of the findings reported in the thesis.

Appendix A

Notations and abbreviations used in the thesis

Notations:

- $\mathcal{N}(t)$: number of customers in the system at time t.
- $\mathcal{N}_1(t)$: Number of customers in station 1.
- $\mathcal{N}_2(t)$: Number of customers in the orbit of station 2.
- $\mathcal{I}_i(t)$: Number of inventoried items in station i $(= 1, 2)$.
- $\mathcal{I}(t)$: Inventory level in the system at time t.
- $\mathcal{C}(t)$: Status of the server is idle/ busy at time t.

  That is, $\mathcal{C}(t) = \begin{cases} 0, \text{ if server is idle at time t.} \\ 1, \text{ if server is busy at time t.} \end{cases}$
- $\mathcal{P}(t)$: Status of the production process at time t.

  That is, $\mathcal{P}(t) = \begin{cases} 0, \text{ if production is off at time t.} \\ 1, \text{ if production is on at time t.} \end{cases}$

- $\mathcal{M}(t)$: Status of the server is vacation/ normal mode at time t.
  That is, $\mathcal{M}(t) = \begin{cases} 0 \text{ if server is in vacation mode at time } t. \\ 1 \text{ if server is in normal mode at time } t. \end{cases}$
- I_k : Identity matrix of order k.
- $\mathbf{e}$: Column vector of 1's with appropriate dimension.
- $\mathbf{0}$: Vector consisting of 0's with appropriate dimension.
- $\mathbb{Z}_+$: The set of positive integers.

Abbreviations:

- PH : Phase type.
- CTMC: Continuous Time Markov Chain.
- QBD: Quasi-Birth-Death.
- LST: Laplace-Stieltjes Transform.
- LIQBD: Level Independent Quasi-Birth-Death.
- LDQBD: Level Dependent Quasi-Birth-Death.

Bibliography

[1] Anoop N. Nair, M.J. Jacob and A. Krishnamoorthy (2013). The multi server $M/M/(s, S)$ queueing inventory system. Ann Oper Res, DOI 10.1007/s10479-013-1405-5.

[2] G. Arivarignan, C. Elango and N. Arumugam (2002). A continuous review perishable inventory control system at service facilities. In Advances in Stochastic Modelling. J. R. Artalejo and A. Krishnamoorthy (eds.), Notable Publications, NJ, USA, 29 - 40.

[3] J. R. Artalejo, A. Gomez-Corral, and M. F. Neuts (2001). Analysis of multiserver queues with constant retrial rate, European Journal of Operational Research, vol. 135, no. 3, 569-581.

[4] J. R. Artalejo and A. Gomez-Corral (2008). Retrial Queueing Systems: A Computational Approach, Springer, Berlin.

[5] Bini, D., and B. Meini (1995). On cyclic reduction applied to a class of Toeplitz matrices arising in queueing problems. In Computations with Markov Chains, Ed., W. J. Stewart, Kluwer Academic Publisher, 21-38.

[6] O. Berman, E. H. Kaplan, and D. G. Shimshak (1993). Deterministic approximations for inventory management at service facilities. IIE Transactions,25: 5, 98 - 104.

[7] O. Berman, E. Kim (1999). Stochastic models for inventory managements at service facilities. Commun.Statist. -Stochastic Models, 15(4), 695 - 718.

[8] O. Berman, K. P. Sapna (2000). Inventory management at service facilities for systems with arbitrary distributed service times. Commun.Statist.-Stochastic Models, 16(3-4), 343 - 360.

[9] O. Berman, K. P. Sapna (2001). Optical control of service for facilities holding inventory. Computers and Operations Research, 28, 429 - 441.

[10] O. Berman, K. P. Sapna (2002). Optimal service rates of a service facility with perishable inventory items. Naval Research Logistics, 49, 464 - 482.

[11] O. Berman, E. Kim (2004). Dynamic inventory strategies for profit maximization in a service facility with stochastic service, demand and lead time. Math.Meth.Oper.Res, 60, 497 - 521.

[12] L. Bright and P. G. Taylor (1995). Equilibrium distribution for level dependent quasi-birth-and-death processes. Commun. Statist. -Stochastic Models, 11, 497 - 525.

[13] P. J. Burke (1956). The output of a queuing system. Operations Res., 4, 699-704.

[14] C. Churchman, E. Ackoff and E. L. Arnoff (1961). Introduction to operations research. John Wiley & Sons INC. New York.

[15] J. Cui, J.Wang (2013). A queueing-inventory system with registration and orbital searching processes. Z. Zhang et al.(eds.), LISS 2012: Proceedings of 2nd International Conference on Logistics, Informatics

and Service Science, DOI 10.1007/978-3-642-32054-5_62, Springer-Verlag Berlin Heidelberg.

[16] T. G. Deepak, Viswanath C. Narayanan, V. C. Joshua (2007). On an inventory with positive service time having optional processing time. Bulletin of Kerala Mathematics Association, Vol.4, No.2, 75 - 86.

[17] T. G. Deepak, A. Krishnamoorthy, Viswanath C. Narayan and K. Vineetha (2008). Inventory with service time and transfer of customers and/inventory. Ann.Oper.Res, 160: 191 - 213.

[18] Dirk Beyer, Feng Cheng, Suresh P. Sethi and Michael Taksar (2010). Markovian Demand Inventory Models, International Series 3 in Operations Research & Management Science 108, DOI 10.1007/978-0-387-71604-6 1, Springer Science+Business Media, LLC.

[19] G. I. Falin and J. G. C. Templeton (1997). Retrial Queues. Chapman & Hall, London.

[20] G. Hadley and T. M. Whitin (1963). Analysis of inventory systems. Prentice Hall Inc., Englewood Cliffs, New Jersey.

[21] R. Jayaraman B. Sivakumar and G. Arivarignan (2013). A perishable inventory system with postponed demands and multiple server vacations. Modelling and Simulation in Engineering, doi:10.1155/2012/620960.

[22] Jihong Li and Naishuo Tian (2007). The M/M/1 queue with working vacations and vacation interruptions. Journal of Systems Science and Systems Engineering, vol. 16, no. 1, 121-127.

[23] S. Kalpakam, S. Shanthi (2001). A perishable inventory system with modified $(S-1, S)$ policy and arbitrary processing times. Computers and Operations Research, 28(5), 453 - 471.

[24] A. Krishnamoorthy and N. Raju (1998). N-Policy for (s,S) perishable inventory system with positive lead time. Korean J.Comput.& Appl.Math, Vol.5, No.1, 253 - 261.

[25] A. Krishnamoorthy, M. E. Islam (2003). Production inventory with retrial of customers in an (s,S) policy. Stochastic Modelling and Applications. Vol.6, No.2, 1 - 11.

[26] A. Krishnamoorthy and M. E. Islam (2004). (s,S) inventory system with postponed demands, Stochastic Analysis and Applications, vol. 22, no. 3, 827 - 842.

[27] A. Krishnamoorthy, Viswanath C. Narayanan,and M. E. Islam (2004). On production inventory with service time and retrial of customers. In Proceedings of the 11th international Conference on Analytical and Stochastic Modelling Techniques and Analysis, 238 - 247, Magdeburg, SCS-Publishing House.

[28] A. Krishnamoorthy and K. P. Jose (2005). An (s,S) inventory system with positive lead time, loss and retrial of customers. Stochastic Modelling and Applications, vol. 8, no. 2, 56 - 71.

[29] A. Krishnamoorthy, T. G. Deepak, V. C. Narayanan and K. Vineetha (2006 a). Control policies for inventory with service time. Stochastic Analysis and Applications, 24(4), 889 - 899.

[30] A. Krishnamoorthy, V. C. Narayanan, T. G. Deepak and K. Vineetha (2006 b). Effective utilization of server idle time in an (s,S) inventory with positive service time. Journal of Applied Mathematics and Stochastic Analysis, DOI 10.1155/JAMSA/2006/69068, 1 - 13.

[31] A. Krishnamoorthy, M. E. Islam, and Viswanath C. Narayanan (2006 c). Retrial inventory with batch markovian arrival and positive service time. Stochastic Modelling and Applications, 9(2): 38 - 53.

[32] A. Krishnamoorthy and K. P. Jose (2007 a). Comparison of inventory systems with service positive lead time, less and Retrial of customers. Journal of Applied Mathematics and Stochastic Analysis, Doi:10.1155/2007/37848.

[33] A. Krishnamoorthy and K. P. Jose (2007 b), Matrix Analytic Solution to an Inventory with Service, Reneging of Customers and Shortage, OPSEARCH, Vol.44, No.2. 147 - 159.

[34] A. Krishnamoorthy K. P. Jose and Viswanath C. Narayanan (2008). Numerical investigation of a PH/PH/1 inventory system with positive service time and shortage. Neural,Parallel and Scientific Computations, 16, 579 - 592.

[35] A. Krishnamoorthy and K. P. Jose (2008). Three production inventory systems with service, loss and retrial of customers. Information and Management Sciences, Vol.19, 3, 367 - 389.

[36] A. Krishnamoorthy and N. Anbazhagan (2008). Perishable inventory system at service facilities with N-policy.Stochastic Analysis and Applications, 26: 120 - 135.

[37] A. Krishnamoorthy, Sajeev S. Nair and Viswanath C. Narayanan (2009). An inventory model with retrial and orbital search. Bulletin of Kerala Mathematics association, Special issue, 47 - 65, Guest editor: S. R. S. Varadhan FRS.

[38] A. Krishnamoorthy and Viswanath C. Narayanan (2010). Production inventory with service time and vacation to the server. IMA Journal of Management Mathematics, Doi:10.1093/imaman/dpp025.

[39] A. Krishnamoorthy, B. Lakshmy, R. Manikandan (2011). A survey on inventory models with positive service time. OPSEARCH 48(2):153-169.

[40] A. Krishnamoorthy, Sajeev S. Nair, Viswanath C. Narayanan (2012). An inventory model with server interruptions and retrials. Oper Res Int J, 12:151–171.

[41] A. Krishnamoorthy, Sajeev S. Nair and Viswanath C. Narayanan (2013). Production inventory with service time and interruptions. International Journal of Systems Science, http://dx.doi.org/10.1080/00207721.2013.837538.

[42] A. Krishnamoorthy, Viswanath C Narayanan (2013). Stochastic Decomposition in Production Inventory with Service Time. European Journal of Operational Research, 228: 358-366.

[43] Lalitha, K.(2010): Studies on classical and retrial inventory with positive service time. Doctoral thesis, submitted to Cochin University of Science & Technology.

[44] Latouche, G., and V. Ramaswami (1999). Introduction to Matrix Analytic Methods in Stochastic Modeling. SIAM., Philadelphia, PA.

[45] Latouche, G., and V. Ramaswami (1993). A logarithmic reduction algorithm for quasi-birth-and-death processes. Journal of Applied Probability, 30, 650-674.

[46] E. Naddor (1996). Inventory systems. John Wiley & Sons, New York.

[47] Neuts, M.F (1994) Matrix-Geometric Solutions in Stochastic Models - An Algorithmic Approach, 2nd ed., Dover Publications, Inc., New York.

[48] Ning Zhao and Zhanotong Lian (2011). A queueing-inventory system with two classes of customers. Int.J.Production Economics, 129: 225-231.

[49] I. Padmavathi, B. Sivakumar, G. Arivarignan (2013). A retrial inventory system with single and modified multiple vacation for server. Ann Oper Res, DOI 10.1007/s10479-013-1417-1.

[50] Paul Manuel, B. Sivakumar and G. Arivarignan (2007). A perishable inventory system with service facilities, MAP arrivals and PH-service times. J Syst Sci Syst Eng, 16(1), 062 - 073.

[51] Paul Manual, B. Sivakumar and G. Arivarignan (2008). A Perishable inventory system with service facilities and retrial customers, Computers and Industrial Engineering, 54, 484 - 501.

[52] P. R. Parthasarathy and V. Vijayalakshmi (1996). Transient Analysis of an Inventory Model: a numerical approach. Intern. J. Computer Math., 59, 177 - 185.

[53] V. Perumal and G. Arivarignan (2002). A continuous review perishable inventory system at infinite capacity service facilities. ANJAC Journal of Science, 1: 37 - 45.

[54] Qi-Ming He (2014). Fundamentals of Matrix-Analytic Methods, Springer Science+Business Media New York.

[55] K. Ruslan and H. Daduna (2012). Loss systems in a random environment - steady state analysis. http://preprint.math.uni-hamburg.de/public/papers/prst/prst2012-04.pdf.

[56] K. Ruslan and H. Daduna (2013). Loss systems in a random environment - embedded Markov chains analysis. http://preprint.math.uni-hamburg.de/public/papers/prst/prst2013-02.pdf.

[57] M. Saffari, R. Haji, F Hassanzadeh (2011). A queueing system with inventory and mixed exponentially distributed lead times. Int. J. Adv. Manuf. Technol. 53:1231-1237.

[58] M. Saffari, S. Asmussen, R. Haji (2013). The $M/M/1$ queue with inventory, lost sale, and general lead times, Queueing Syst., DOI 10. 1007/s11134-012-9337-3.

[59] I. Sahin (1990). Regenerative inventory systems: Operating characteristics and optimization, Springer-Verlag.

[60] Sajeev S. Nair (2012). On (s, S) inventory policy with/without retrial and interruption of service/production. Ph. D. thesis, Cochin University of Science and Technology, India.

[61] Sennott, L. I., Humblet, P. A. and Tweedie, R. L. (1983). Mean drifts and the non-ergodicity of Markov chains. Operations Research 31: 783-789.

[62] M. Schwarz, C. Sauer, H. Daduna, R. Kulik, and R. Szekli (2006). $M/M/1$ queueing system with inventory. Queueing Systems: Theory and Applications, 1, 54, 55 - 78.

[63] M. Schwarz and H. Daduna (2006). Queueing systems with inventory management with random lead times and with backordering, Mathematical Methods of Operations Research, vol. 64, no. 3, 383-414.

[64] M. Schwarz, C. Wichelhaus, and H. Daduna (2007). Product Form Models for Queueing Networks with an Inventory, Stochastic Models, vol. 23, no. 4, 627-663.

[65] Sigman K, Simchi-Levi D (1992). Light traffic heuristic for an $M/G/1$ queue with limited inventory. Ann Oper Res 40:371-380.

[66] B. Sivakumar (2008). Two-commodity inventory system with retrial demand. European Journal of Operational Research, 187(1), 70–83.

[67] B. Sivakumar (2009). A perishable inventory system with retrial demands and a finite population. Journal of Computational and Applied Mathematics, 224, 29-8.

[68] B. Sivakumar (2011). An inventory system with retrial demands and multiple server vacation. Quality Technology & Quantitative Management, 8(2), 125–146.

[69] B. Sivakumar and G. Arivarignan (2005). A perishable inventory system with service facilities and negative customers. Advanced Modelling and Optimization, Vol.7,No.2, 193 - 210.

[70] B. Sivakumar and G. Arivarignan (2006).A perishable inventory system at service facilities with negative customers. Information and Management Sciences, vol.17, no.2, 1 - 18.

[71] B. Sivakumar, C. Elango, G. Arivarignan (2006). A perishable inventory system with service facilities and batch Markovian demands. International Journal of Pure and Applied Mathematics, Vol.32, No.1, 33 - 49.

[72] B. Sivakumar and G. Arivarignan (2009). A stochastic inventory system with postponed demands. Performance Evaluation 66: 47-58.

[73] Ushakumari, P. V (2006). On (s, S) inventory system with random lead time and repeated demands. Journal of Applied Mathematics and Stochastic Analysis, 2006, 1-22. Article ID 81508.

[74] Viswanath C. Narayanan, T. G. Deepak, A. Krishnamoorthy and B. Krishnakumar (2008). On an (s,S) inventory policy with service time, Vacation to server and correlated lead time. Journal of Quality Technology Quantitative Management, Vol.5, No.2, 129 - 143.

[75] Vineetha, K (2008). Analysis of inventory systems with positive and / negligible service time. Ph. D. thesis, University of Calicut, India.

[76] V. S. S. Yadavalli, B. Sivakumar, G. Arivarignan (2007). Stochastic inventory management at a service facility with a set of reorder levels. ORiON, 23(2), 137 - 149.

[77] V. S. S. Yadavalli, B. Sivakumar, G. Arivarignan (2008). Inventory system with renewal demands at service facilities. Int.J.Production Economics, 114, 252 - 264.

[78] V. S. S. Yadavalli, B. Sivakumar and G. Arivarignan (2009). An inventory system with multi server facility with negative demands. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5223937.

[79] V. S. S. Yadavalli, B. Sivakumar and G. Arivarignan and O. Adetunji (2011). A multi-server perishable inventory system with negative customer. Computers & Industrial Engineering, 61(2), 254 - 273.

[80] V. S. S. Yadavalli, B. Sivakumar and G. Arivarignan and O. Adetunji (2012). A finite source multi-server inventory system with service facility. Computers & Industrial Engineering, 63, 739 - 753.