ANALYSIS OF INVENTORY SYSTEMS WITH POSITIVE AND/NEGLIGIBLE SERVICE TIME

Thesis submitted to the University of Calicut for the degree of

DOCTOR OF PHILOSOPHY

under the faculty of Science

By

VINEETHA. K

Department of Statistics

University of Calicut

Kerala-673 635

India

January 2008.

CERTIFICATE

This is to certify that the thesis entitled ***ANALYSIS OF INVENTORY SYSTEMS WITH POSITIVE AND /NEGLIGIBLE SERVICE TIME*** *submitted by* ***Smt. Vineetha K.*** *to University of Calicut for the award of the degree of Doctor of Philosophy in Statistics is a record of research work carried out by her under my guidance and supervision and this thesis or part thereof has not previously formed the basis for the award of any Degree, Diploma etc., of any other University or Institution*

Calicut University Campus **Dr. N. Raju.**

Date :

DECLARATION

I hereby declare that the work reported in this thesis entitled **ANALYSIS OF INVENTORY SYSTEMS WITH POSITIVE AND /NEGLIGIBLE SERVICE TIME** is the result of investigations carried out by me under the supervision of Dr. N. Raju, Department of Statistics, University of Calicut, Calicut University P. O and it has not been submitted for the award of any Degree, Diploma etc., of any University or Institution.

Calicut University Campus **Vineetha. K.**

Date:

ACKNOWLEDGEMENTS

I express my deep felt gratitude and indebtedness to my supervisor Dr.N.Raju who encouraged me throughout and it is because of his constant help and support that I could do this work.

I acknowledge with gratitude all the help and blessings received from Professor. A.Krishnamoorthy, Head, Department of Mathematics, CUSAT. The discussions with him were a great source of inspiration to me. My sincere thanks to him for his encouragement and guidance without which the completion of this work would have been difficult.

I recall with gratitude the help from Dr.T.G.Deepak, IIST, Trivandrum and Dr.Viswanath C Narayanan, Dept. of Mathematics, Govt. Engineering College, Sree Krishnapuram. Thanks are due to Dr.K.K.Jose and other faculty members of the Department of Statistics, St.Thomas College, Pala for their encouragement and blessings.

I thank Dr.C.Chandran, Head and the faculty members Dr.M.Manoharan and Dr.K.Jayakumar of the Department of Statistics for their support during this work. I am greatly thankful to the non-teaching staff of the Department of Statistics and the staff of CDC, Project cell and Finance wing of the University for their help and support.

With great pleasure I express my affection and gratitude towards my friends in the Computer Science Department for their encouragement. My thanks are also due to Mr. Sanjai Varma, Beeta Computers, Thripoonithura for the excellent typing of this thesis.

I gratefully acknowledge DST for the financial support in the form of a research project (No.SR/WOS-A/MS-03/2005) under the Women Scientist Scheme. My sincere thanks to the University authorities for providing an excellent environment to work.

I am indebted to my father-in-law Sri.Sreekumaran Moosath and mother-in-law Smt.Subhadra Antharjanam for their constant encouragement and support. Also I express my affection to my sister-in-law Mrs.Prabha and her family for their love and care during my visits to CUSAT. I owe much to my husband Dr.K.S.Subramanian Moosath and my children Adarsh and Varadha for their support and encouragement all along.

Vineetha.k.

Contents

Chapter 1

Introduction

1.1 Inventory systems

The existence of inventory in a system generally implies the existence of an organised complex system involving inflow, accumulation and outflow of some commodities or goods or items or products. The scientific analysis of inventory systems defines the degree of interrelationship between inflow, accumulation and outflow and identifies economic control methods for operating such systems.

Inventory models play an important role in operations research and management science. It is an integral part of logistic systems common to all sectors of economy. Because of the importance of this branch of decision making in a number of practical situations in Government, military organisations, industries, hospitals *etc*.; significant development of the subject in new directions resulted.

The main factors affecting the inventory are demand, lifetime of items stored, production rate, lead time, damage rate due to external disaster, availability of space in store *etc*. If all the parameters are known beforehand, then the inventory model is called deterministic inventory model. If some or all of these parameters are not known with certainty, then it is justifiable to consider them as random variables with some probability distributions and the resulting inventory model is then called stochastic inventory model. There can be single or multi-commodity inventory systems. Inventory systems may again be classified as continuous review or periodic review. In continuous review, the system is monitored continuously over time. In periodic review systems, the system is monitored at discrete, equally spaced instants of time.

An inventory policy is a set of decision rules that dictate the 'When' and the 'How much' to order. Several policies may be used to control an inventory system; of these, the single most important policy is the (s, S) policy. The time between the order and replenishment is called lead time. If the replenishment is instantaneous, then the lead time is zero, otherwise the system is said to have positive lead time. Shortage of inventory can occur in systems with positive lead time. Then there can be back orders. In an inventory with positive lead time one of the following assumptions is made: (i) finite backlog (ii) infinite backlog (iii) no backlog.

1.2 Classical inventory

We shall refer to an inventory with negligible service time as classical inventory model. In all works reported in inventory prior to 1993 it was assumed that the time required to serve the item to the customer is negligible. If an item is available at a customer demand epoch then his waiting time is negligible. In classical inventory models, a queue of customers can be formed only when inventory level becomes zero and lead time is positive/shortage cost is not high compared to holding cost. Immediately on replenishment of items this queue will not vanish. Items are served one at a time to each waiting customer. If the number of customers waiting at the replenishment epoch turns out to outnumber the quantity replenished then only a part of the queue vanish.

The mathematical analysis of inventory problems was started by Harris [22]. He proposed the famous EOQ formula, that was popularised by Wilson. Right from the days of Harris and Wilson, the fundamental assumption in any inventory problem was that the service time is negligible.

1.3 Inventory with service time

There are various situations arising in real life where positive processing (service) time is involved before an item is delivered to the customer. Thus in an inventory with service time, a queue will be formed even when the item is available. This is not the case with inventory with negligible service time except when backlog of demand is permitted in the absence of inventory. Thus the problem in inventory with service time may appear as a problem in queue. The paper by Berman, Kim and Simshak [8] is the first attempt to introduce positive service time in inventory where it is assumed that service time is deterministic as also the inter-arrival time. Later Berman and Kim [6] extended these results to service time of random duration. Berman and Sapna [10], Parthasarathy and Vijayalakshmi [48] deal with control problems in inventory with service time. Arivarignan *et al.* [2] considers a continuous review perishable inventory control system with positive service time. One of the most recent contribution to inventory with service time is due to Schwarz *et al.* [54].

It may be noted that inventory with positive service time differs substantially in analysis from that of queue. In the latter, if a customer is available and the server is free and ready to serve, the customer starts getting service. However, in inventory with service time, the availability of the items also is needed in addition to the features required in a queue. This means that in queues we do not look at the availability of resources used for service, whereas this has to be taken into account in inventory. In a way one can regard the classical queueing problem as an inventory problem with unlimited number of items stored with the provision that holding cost is negligible!. Thus in inventory the customer is forced to wait in the absence of the item, even when server is free. This leads to the need for a separate discussion on inventory with positive service time.

1.4 Production inventory

Very little investigation on production inventory had been made in the past. This is so especially in the non-deterministic case. That is when factors such as demands, lead time etc., are not deterministic. The analysis becomes highly complex when the items are of random life times and the lead time is positive. Doshi *et al.* [17] study a production inventory system with a compound Poisson arrival of demands and a continuous review (s, S) production control policy. They derive measures of performance such as the probability distribution of the inventory level, average number of switch overs, and lost sales per unit time. Garish and Graves [18] study a continuous review (s, S) production inventory system with unit demand and fixed replenishment times. de Kok *et al.* [15] study the same problem with backorders and develop approximations for the switch over levels.

Sharafali [55] considers a production inventory system operating under the (s, S) policy where demands arrive according to Poisson process and production output is also Poisson. He assumes that the machine is subject to failure and repair time has general distribution. de Kok [14] deals with a one-product production inventory model, where the production rate can be dynamically adjusted in order to cope with random fluctuations in demand. The demand process for the product is described by a compound

Poisson process and the excess demand is lost.

Hsu and Tapiero [24] analyse a production inventory model using queueing theory techniques. Mitra [42] considers a production consumption system characterised by a finite capacity buffer that incorporates the reliability factor in to the modelling. Altiok [1] analyses a production inventory system with compound Poisson demand and phase-type distribution for the processing time. Berge *et al.* [4] consider production inventory systems in which a number of producing machines are susceptible to failure following which they must be repaired to make them operative again. William *et al.* [64] examines the impact of production level time on costs through the use of periodic review production inventory model with a non-stationary stochastic demand process.

1.5 Phase-type distribution

The idea of 'method of phases' was first proposed by A. K. Erlang. He gave a generalisation of the exponential distribution. His idea lead to what is known as Erlangian distribution. The basic idea behind Erlang's approach was the memoryless property of exponential distribution. Erlang's idea was extended by Cox who gave phase representation for all probability distribution on the positive real time which have rational Laplace-Stieltjes transform. This generalisation requires heavy use of complex analysis, due to which their numerical implementation becomes difficult. This drawback is overcome by Neuts [47] for at least a subclass of the distribution considered by Cox and has received major attention in stochastic modelling.

To get away from Poisson/exponential models Neuts [47] developed the theory of phase-type (PH) distributions and related point process. PH distributions lead themselves naturally to algorithmic implementation. They have nice closure properties and a related matrix formulation that make them attractive for use in practice.

1.5.1 Continuous PH-distribution

Let $X = (X_t : t \geq 0)$ denote a homogeneous Markov process with finite state space $\{1, 2, \dots, m+1\}$ and generator

$$Q = \begin{bmatrix} T & \eta \\ 0 & 0 \end{bmatrix}$$

where T is a square matrix of dimension m, η a column vector and 0 the zero row vector of the same dimension. The initial distribution of X shall be the row vector $\overline{\alpha} = (\alpha, \alpha_{m+1})$, with α being a row vector of dimension m. The first m states $\{1, \dots, m\}$ shall be transient, while the state $m+1$ is absorbing. Let $Z = \inf\{t \geq 0 : X_t = m+1\}$ be the random variable of the time until absorption in state $m+1$.

The distribution of Z is called phase-type distribution with representation PH (α, T). The dimension m of T is called the order of the distribution PH (α, T). The states $\{1, 2, \dots, m\}$ are called phases.
Let $\mathbf{e}$ denote the column vector of dimension m with all entries equal to one; then

$$\eta = -T\mathbf{e} \quad \text{and} \quad \alpha_{m+1} = 1 - \alpha\mathbf{e}.$$

The distribution function of Z is given by

$$F(t) = p(Z \leq t) = 1 - \alpha \exp(Tt)\mathbf{e}, \quad \text{for all } t \geq 0.$$

and density function is

$$f(t) = \alpha \exp(T \cdot t)\eta \quad \text{for all } t > 0,$$

where $\exp(T \cdot t) = \sum\limits_{n=0}^{\infty} \frac{t^n}{n!} T^n$ denotes a matrix exponential function.

1.5.2 Discrete PH-distribution

Analogous to the definition of PH distribution in continuous time, we will define discrete PH distributions in terms of Markov chains with one absorbing state. Let χ denote a Markov chain with finite state space $E = \{0, 1, \ldots m\}$ and a transition matrix

$$P = \begin{bmatrix} 1 & 0 \\ \eta & T \end{bmatrix}.$$

Let $\overline{\alpha} = (\alpha_0, \alpha)$ be the initial distribution of χ where α is of dimension m. 0 is absorbing state. All other states shall be transient.

Let $Z = \min\{n \in N_0;\ X_n = 0\}$
denote the time until absorption in state 0.
Define $p_n = P_r(Z = n)$ for all $n \in \mathbb{N}_0$. The distribution $P = (p_n : n \in \mathbb{N}_0)$ of Z is called a discrete PH distribution denoted by $\text{PH}_d(\alpha, T)$. The number m of transient states is called the order of P. A transient state is called phase. Here

$$\eta = 1 - T\mathbf{1} \text{ and } p_0 = P_r(Z = 0) = \alpha_0 = 1 - \alpha\mathbf{e}.$$

1.6 Matrix analytic method

Matrix analytic methods constitute a success story, illustrating the enrichment of a science, applied probability, by a technology, that of digital computers. Neuts has played a seminal role in these exciting developments, promoting numerical investigations as an essential part of the solution of probability models. Matrix analytic methods are popular as modelling tools because they give one the ability to construct and analyse, in a unified way and in an algorithmically tractable manner, a wide class of stochastic models.

1.6.1 Quasi-Birth and Death process (QBD)

QBD's are matrix generalisations of simple birth and death process on the non-negative integers. Consider a discrete time Markov chain $\{X_t : t \in N\}$ on the two dimensional state space $\{(n,i) : n \geq 0, 1 \leq i \leq m\}$, which we partition as $\bigcup_{n\geq 0} l(n)$, where $l(n) = \{(n,1),(n,2)\ldots(n,m)\}$ for $n \geq 0$. The first coordinate n is called the level, and the second coordinate j is called the phase of the state (n,j). The Markov chain is called a QBD if one-step transitions from a state are restricted to states in the same level or in the two adjacent levels: it is possible to move in one step from (n,j) to (n',j') only if $n' = n, n+1$ or $n-1$ (provided in the last case that $n \geq 1$).

The transition probabilities are assumed to be level-independent. That is for n or n' greater than or equal to 1, the probability $P[X_1 = (n',j)|X_0 = (n,i)]$ may depend on i, j and $n'-n$, but not on the specific values of n and n'. The infinitesimal generator $\overline{Q}$ is given by

$$\overline{Q} = \begin{bmatrix} B_0 & A_0 & & & \\ B_1 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & A_2 & A_1 & A_0 & \cdots \\ & & \vdots & \vdots & \vdots & \end{bmatrix}$$

where $B_0e + A_0e = B_1e + A_1e + A_0e = (A_0 + A_1 + A_2)e = 0$. The generator $\overline{Q}$ is assumed to be irreducible. The matrix $A = A_0 + A_1 + A_2$ is a finite generator.

Theorem 1.6.1. *The process $\overline{Q}$ is positive recurrent if and only if the minimal non-negative solution R to the matrix-quadratic equation*

$$R^2A_2 + RA_1 + A_0 = 0$$

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

$$x_0(B_0 + RB_1) = 0$$
$$x_0(I - R)^{-1}\mathbf{e} = 1$$

has a unique positive solution x_0.

If the matrix A *is irreducible, then* $s\rho(R) < 1$ *if and only if*

$$\pi A_2 e > \pi A_0 e$$

where π *is the stationary probability vector of* A.

The stationary probability vector $x = (x_0, x_1, \ldots)$ *of* $\overline{Q}$ *is given by*

$$x_i = x_0 R^i \quad \textit{for} \quad i \geq 0.$$

The equivalent equalities $RA_2e - A_0e = RB_1e - B_0e = 0$ *hold.*

1.6.2 First passage time

Let $G^v_{jj},(K,x)$ denote the conditional probability that a QBD process, starting in the state $(i+v, j)$ at time $t = 0$, reaches the state i for the first time not later than time x, after exactly k' transitions to the left, and does so by entering the state (i, j). This probability is so defined for $x \geq 0$, $v \geq 1$, $K \geq 1$, $1 \leq j \leq m$, $1 \leq j' \leq m$ and for any i such that the states (i, j'), $1 \leq j' \leq m$ are not boundary states. The probability $G^{(v)}_{jj'}(K, x)$ does not depend on i. The matrix with elements $G^{(v)}_{jj'}(K, x)$ will be denoted by $G^{(v)}(K, x)$. For ease of presentation, we also introduce the transform matrix

$$\hat{G}^{(v)}(Z, s) = \sum_{K=1}^{\infty} Z^K \int_0^{\infty} e^{-sx} dG^{(v)}(K, x),$$

for $|Z| \leq 1$ and $Re\ s \geq 0$.

1.7 Review of related works

Mathematical modelling of inventory problems started in 1915 with the work of Harris [22] in which he derives a formula popularly known as 'Harris-Wilson economic lot size formula'. A systematic analysis of the (s, S) inventory system using renewal theoretic arguments is provided in Arrow, Karlin and Scarf [3]. Veinott and Wagner [62] provide algorithms for finding the optimal (s, S) inventory policies. The application of these models in various practical situations and the related problems are dealt with in Hadley and Whitin [21], Naddor [43] and Wagner [65]. In the review article Veinott [63] provides a detailed account of the work carried out in inventory theory. Gross and Harris [19] consider the inventory systems with state dependent lead times. In [20] they developed the idea of dependence between replenishment times and the number of outstanding orders. Analysis of inventory systems with random lead times is provided in Ryshikov [51].

Sivazlian [57] considers a continuous review (s, S) inventory problem with renewal demands and zero lead times. He proves that the limiting distribution of the inventory position is uniform on $(s + 1, \ldots, S)$. Richards [49] proves the same result for the case with compound renewal demand. Richard [50] analyses a continuous review (s, S) inventory system in which the demand for items in inventory is dependent on an external environment. Sahin [52] discussed continuous review (s, S) inventory with continuous state space and constant lead times. Srinivasan [59] extended Sivazlian's result to the case of arbitrarily distributed lead times. He derived explicit expressions for probability mass function of the stock level and extracted steady state results. This was further extended by Manoharan, Krishnamoorthy and Madhusoodanan [35] to the case of non-identically distributed inter-arrival times. Sahin [53] derived the binomial moments of the transient and stationary distributions of the number of backlogs in a continuous review (s, S) inventory model with arbitrarily distributed lead times and compound renewal demand. Thangaraj and Ramanarayanan [60] deal with an inventory system with random lead time having two order levels.

Kalpakam and Arivarignan [25] studied a continuous review inventory system having an exhibiting item subject to random failure (exponentially distributed life times). They [26] extended the result to exhibiting items having Erlangian life times under renewal demands. Again in [27] they analysed a perishable inventory model having exponential lifetimes for all the items. Liu [41] analyses an (s, S) continuous review inventory system in which depletions are due to random demands and random failure. Here backlogs are allowed but the lead time is assumed to be zero. Kalpakam and Sapna [28] analyse an (s, S) perishable system with Poisson demand process and exponentially distributed lead times, allowing no backlogs. In [29] they extent it to the case of arbitrary lead time distribution.

In all the works reported in inventory prior to 1993, the service time is considered to be negligible. Berman, Kim and Shimshak [8] is the first attempt to introduce positive service times in inventory, where it was assumed that the service time is deterministic. Parthasarathy and Vijayalakshmi [48] is the first to analyse stochastic inventory models with positive service times. They obtain time dependent solution for the system state. Later Berman and Kim [5] extended this result to random service time. Berman and Sapna [9] investigate inventory control at a service facility, which uses one item of inventory for service provided. Parthasarathy and Vijayalakshmi [48] considers a transient analysis of an $(S-1, S)$ inventory model. Schwarz *et al.* [54] considers an $M|M|1$ queueing system with inventory. Their model assumes balking of customers when inventory level is zero (lost sales situation). They obtained that with infinite waiting room capacity the limiting distributions of the queue length processes are the same as in the classical $M|M|1|\infty$ system. Kazimirsky [30] considers $BMAP|G|1$ queue with service time distribution depending on number of processed items. He derives an efficient algorithm for calculating the stationary queue length distribution and Laplace-Stieltjes transform of the

Sojourn time. Sivasamy and Elangovan [58] considers a queueing system with machining input source operating under an N-policy. They developed computational algorithms for numerical evaluation of steady state probability distribution of queue length process, and first passage times through matrix analytic methods. An optimum N-policy is also obtained.

Queues with transfer of customers is studied by Qi-Min-He and Neuts [23]. They consider two $M|M|1$ parallel queues with infinite capacity and manned by one server each. They provide a number of extremely useful measures of the system that can effectively be utilised as control measures. Krishnamoorthy and Jose [33] considers an (s, S) inventory system with positive lead time and loss and retrial of customers. In another paper [34] they analyse and compare three (s, S) inventory systems with positive service time and retrial of customers. Krishnamoorthy *et al.* [31] consider an (s, S) inventory system with service time, vacation to customer and correlated lead time. Service times have PH distributions.

Krishnamoorthy and Raju [36] considers N-policy for an (s, S) inventory systems with perishable items. Later in [37] they consider N-policy for production inventory system with random life times. They obtain an optimal value of N for fixed S. The waiting time distribution for a required item is computed. Krishnamoorthy and Mohammed Ekramol Islam [32] analysed a production inventory with retrial of customers. Krishnamoorthy and Ushakumari [39] consider a D-policy for a k-out of $n : G$ system with repair. They obtain system state distribution, system reliability, expected length of time the server is continuously available, expected number of times the system is down in a cycle *etc*. They obtain an optimal D value which maximises a suitably defined cost function. In [61] they discussed $\max(N, T)$ policy for the same G system.

1.8 An outline of the present work

This thesis is organised in six chapters including this introduction. The second chapter is about control policies for inventory with service time. Control policies like N, D, T and their combinations are extensively studied in queueing systems. These can be extended to inventory systems with service time. The results thereby obtained will be of great use in management related problems. Here we are introducing N-policy for an (s, S) inventory system with positive service time. If at a service completion epoch no customer is left, then the server is switched off to be activated again at the epoch when N customers accumulate. Under specified inter-arrival and service time distributions, which are independent of each other, we obtain the necessary and sufficient condition for the system to be stable. System state distribution is shown to have product form solution. We also obtain the optimal values of the control variables s, S and N; it is seen that the cost function attains the minimum value at $s = 0$. It is also shown that the cost function is separately convex in the variables S and N. Several measures of performance of the system are evaluated. Numerical illustrations are provided.

In the third chapter we consider the effective utilisation of idle time in an (s, S) inventory with positive service time. Here we introduce some additional features in an (s, S) inventory system with positive service time. In all the models considered with positive service time, including the one we considered in chapter 2, the service starts only on the arrival of a customer; in the absence of customers (or inventory) the server remains idle. We relax this condition in our problem. If an item is available at a service completion epoch, the server process it even in the absence of customers (eg: assembling of parts); It keeps on doing this as long as unprocessed items are available in the inventory, keeping the sum-total of processed and unprocessed items (inclusive of the one being processed, if any) at most S. Thus at a customer departure epoch either there is no processed item available or one or more processed items available. As long as processed items are available, service time is negligible; otherwise a queue of customers formed. The processing of items even in the absence of customers can ensure reduction in the waiting time and hence in the associated cost. Even in the simplest of cases we will not get analytical solutions. Hence we proceed for an algorithmic analysis of the system. We establish a necessary and sufficient condition for system stability. Several performance measures are computed. An optimization problem is discussed.

In the fourth chapter we consider an inventory system with two parallel service facilities with service time and transfer of customers and/inventory. In the previous chapters we consider single server inventory systems with positive service time. Here we consider an inventory system with two parallel service facilities. When the numerical difference between the number of customers in the two queues reaches a certain fixed quantity L, a certain number $K(K < L)$ of customers are transferred from the longer to the shorter queue. In addition to the transfer of customers from the longer to the shorter queue, we transfer a certain quantity of inventory also depending on the availability. Further if one of the queues has customers, but has no inventoried items, where as the other has atleast one inventoried item to spare, then exactly one item is taken to the former and service begins thereby enhancing the efficiency of the system. Stability of the system is analysed. Several performance measures that helps in efficient design of such systems are computed.

In the previous chapters we consider exponentially distributed service time. In chapter 5 we consider arbitrarily distributed service time. Here we consider an (s, S) inventory with arbitrarily distributed service time. Several situations are considered: (i) case of zero lead time (ii) case of random lead time with different replenishment policies. In the second case we restrict to the case where the system does not admit customers when the inventory level is zero (lost sales). All these cases lead to matrices of the $M|G|1$ type. We establish product form solutions. Also the optimal values of control variables are investigated.

In the previous chapters we considered inventory systems with positive service time. In chapter 6 we consider an inventory system with negligible service time. Here we consider a production inventory system with perishable items. We introduce a control policy, D-policy for this production inventory system with items having random life times. The maximum capacity of the inventory system is S. When the inventory level is S, the production process is switched off. Thereafter the depletion in inventory is due to demand and perishability. The production starts only when the accumulated workload exceeds a threshold D (> 0). Once the production starts, it stops only when the inventory level reaches S. When the inventory level reaches zero without accumulating D, the production starts automatically so as to minimize loss of demands. Steady state probabilities and several performance measures are obtained. The case without perishability is derived as a special case.

Chapter 2

Control Policies for Inventory with Service Time

2.1 Introduction

In this chapter we consider N-policy for an (s, S) inventory system with positive service time. If at a service completion epoch no customer is left, then the server is switched off, to be activated again at the epoch when N customers accumulate. We introduce an additional control variable N to the (s, S) inventory system. Even when the server is switched off, there can be positive on hand inventory which attracts holding cost. Thus the threshold N value will depend also on the holding cost of on hand inventory during the servers inactive time, in addition to the cost incurred due to customers being made to wait. Under simple assumptions on the arrival of demands (customer arrival process) to the system and on the pattern of service, we provide an analytical expression for the optimal threshold value for N. The analytical solution to system state probability distribution is first arrived at in a heuristic fashion, and then it is given theoretical foundation.

---

* The content of this chapter is published in 'Stochastic Analysis and Applications', 24: 719–732, 2006.

The results, especially concerning N, is of great importance to production and manufacturing facilities providing service on the following count. It is not always optimal to start service immediately on arrival of the first customer to an idle server. This is so because there can be huge cost toward switching on the server. So unless the resulting busy period is long enough, the per unit time operational cost turns out to be very high compared to the cost toward customer waiting cost. Thus, it is better to start service facility only after N customers accumulate in the waiting line. This will guarantee that at least N customers are served in a busy cycle which in turn brings down the per unit time operational cost considerably. Of course, the optimal N value depends on the activation cost of the server, the waiting cost of customers, the holding cost of inventoried items and so on. The results obtained will be of great use in management-related problems.

Control policies such as N, T, D and their combinations, are extensively studied in queueing literature. It may be noted that inventory with positive service time differs substantially in analysis from that of queues. Although in the latter if a customer is available and the server is free and ready to serve, the customer starts getting service. However, in inventory with service time, the availability of the item is also needed in addition to the features required in queueing theory. This means that in queueing theory we do not look at the availability of resources used for service, whereas this has to be taken into account in the inventory. Thus, the customer is forced to wait in the absence of the item, even when server is free. This leads to the need for a separate discussion on inventory with positive service time.

Berman, Kim and Shimshak [8] introduced inventory with positive service time, where it is assumed that service time is deterministic. Later Berman and Kim [5] extended this results to random service time. The analysis is via dynamic programming and is over a finite horizon. Berman and Sapna [9] consider a finite buffer system and use Markov renewal theory for analysis. Parthasarathy and Vijayalakshmi [48] deal with control problem in inventory with service time. One of the most recent contribution to inventory with service time is due to Schwarz *et al.* [54]. Krishnamoorthy and Jose [33] compare three (s, S) inventory systems with positive service time and retrial of customers. Krishnamoorthy *et al.* [31] considers an (s, S) inventory system with service time, vacation to customer and correlated lead time. Krishnamoorthy and Raju [37] considers N-policy for an (s, S) inventory system with items having random life times.

2.2 Formation and analysis of the problem

Customers arrive to a single server counter according to a Poisson process of rate λ. Service times of customers are independent and identically distributed exponential random variables with parameter μ. Arrival and service processes are independent of each other. Also service times of customers are mutually independent. The inventory stock is governed by (s, S) policy. When the server becomes idle, it is switched off until N customers accumulate at which epoch the server is activated. If at a service completion epoch no customer is left and the inventory level drops to s, then order for replenishment is placed only on accumulation of N customers. Lead time is assumed to be zero.

Let

$N(t) =$ number of customers in the system including the one getting service (if any) at time t

$I(t) =$ Inventory level in the system at time t

$$C(t) = \begin{cases} 0 & \text{if the server is idle at time } t. \\ 1 & \text{if the server is busy at time } t. \end{cases}$$

and $X(t) = (N(t), C(t), I(t))$. Then $\{X(t);\, t \geq 0\}$ forms a continuous time Markov chain with state space

$$S = \bigcup_{i=0}^{\infty} l(i)$$

where $l(0) = \{(0,0,j)/s \leq j \leq S-1\}$.

For $1 \leq i \leq N-1$

$$l(i) = \{(i,0,j)|s \leq j \leq S-1\} \cup \{(i,1,j)|s+1 \leq j \leq S\}$$

and for $i \geq N$

$$l(1) = \{(i,1,j)|s+1 \leq j \leq S\}.$$

Arranging the states in the lexicographic order, the infinitesimal generator Q of the process $\{X(t); t \geq 0\}$ can be written as

$$Q = \begin{array}{l} \underline{0} \\ \underline{1} \\ \underline{2} \\ \vdots \\ N-2 \\ N-1 \\ N \\ N+1 \\ N+2 \\ \vdots \end{array} \begin{pmatrix} A_{00} & A_{01} & & & & & & & & \\ A_{10} & \tilde{A}_1 & \tilde{A}_0 & & & & & & & \\ & \tilde{A}_2 & \tilde{A}_1 & \tilde{A}_0 & & & & & & \\ & \ddots & \ddots & \ddots & & & & & & \\ & & & \tilde{A}_2 & \tilde{A}_1 & \tilde{A}_0 & & & & \\ & & & & \tilde{A}_2 & \tilde{A}_1 & \tilde{\tilde{A}}_0 & & & \\ & & & & & \tilde{\tilde{A}}_2 & A_1 & A_0 & & \\ & & & & & & A_2 & A_1 & A_0 & \\ & & & & & & & A_2 & A_1 & A_0 \\ & & & & & & & \ddots & \ddots & \ddots \end{pmatrix}$$

where

$$
\begin{aligned}
A_{00} &= -\lambda I_{S-s} \\
A_{01} &= [\lambda I_{S-s} 0 I_{S-s}] \\
A_{10} &= \begin{bmatrix} 0 I_{S-s} \\ \mu I_{S-s} \end{bmatrix} \\
\tilde{A}_1 &= \begin{bmatrix} -\lambda I_{S-s} & 0 \\ 0 & -(\lambda+\mu) I_{S-s} \end{bmatrix} \\
\tilde{A}_0 &= \lambda I_{2(S-s)} \\
\tilde{A}_2 &= \begin{bmatrix} 0 I_{S-s} & 0 I_{S-s} \\ 0 I_{S-s} & \mu E \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 1 \\ I_{S-s-1} & 0 \end{bmatrix} \\
\tilde{\tilde{A}}_0 &= \begin{bmatrix} \lambda E \\ \lambda I_{S-s} \end{bmatrix} \\
\tilde{\tilde{A}}_2 &= \begin{bmatrix} 0 I_{S-s} & \mu E \end{bmatrix} \\
A_1 &= -(\lambda+\mu) I_{S-s} \\
A_0 &= \lambda I_{S-s} \\
A_2 &= \mu E.
\end{aligned}
$$

2.3 Stability condition

Theorem 2.3.1. *The process* $\{X(t) | t \geq 0\}$ *will be stable if and only if* $\lambda < \mu$.

Proof. Since the process $\{X(t),\, t \geq 0\}$ is level independent QBD, it will be stable if and only if $\pi A_0 e < \pi A_2 e$. (See Neuts [44]), where π is the steady state distribution of generator matrix

$$A = A_0 + A_1 + A_2 = -\mu I_{S-s} + \mu E.$$

It can be seen that $\pi = [\frac{1}{S-s}, \frac{1}{S-s}, \ldots, \frac{1}{S-s}]$ and, $\pi A_0 e = \lambda$ and $\pi A_2 e = \mu$ and the theorem follows. $\square$

2.4 The steady state probability vector of Q

Here, we calculate the steady-state probability vector of Q under the stability condition. Let the steady-state probability vector X of Q be partitioned according to the levels as

$$X = (x(0), x(1), x(2), \ldots, x(N-1), x(N), \ldots)$$

where the sub-vectors $x(i)$, $1 \le i \le N-1$, contain $2(S-s)$ elements and all other sub-vectors contain $(S-s)$ elements. The sub-vectors satisfy the equations

$$x(0)A_{00} + x(1)A_{10} = 0 \tag{2.1}$$

$$x(0)A_{01} + x(1)\tilde{A}_1 + x(2)\tilde{A}_2 = 0 \tag{2.2}$$

$$x(i-1)\tilde{A}_0 + x(i)\tilde{A}_1 + x(i+1)\tilde{A}_2 = 0; \quad 2 \le i \le N-2 \tag{2.3}$$

$$x(N-2)\tilde{A}_0 + x(N-1)\tilde{A}_1 + x(N)\tilde{\tilde{A}}_2 = 0 \tag{2.4}$$

$$x(N-1)\tilde{\tilde{A}}_0 + x(N)A_1 + x(N+1)A_2 = 0 \tag{2.5}$$

$$x(i-1)A_0 + x(i)A_1 + x(i+1)A_2 = 0; \quad i \ge N+1 \tag{2.6}$$

Again, partition the sub-vectors $x(i)$, $1 \le i \le N-1$, as

$$x(i) = (x(i,0), x(i,1))$$

where the sub-vectors $x(i,j)$, $j = 0, 1$, contain $S-s$ elements each. Equations (2.1)–(2.6) give rise to the following equations. (In what is the follow

E stands for the matrix $\begin{bmatrix} 0 & 1 \\ I_{S-s-1} & 0 \end{bmatrix}$)

$$\lambda x(0) - \lambda x(1,0) = 0 \tag{2.7}$$

$$\lambda x(i,0) - \lambda x(i+1,0) = 0, \quad 1 \leq i \leq N-2 \tag{2.8}$$

$$-\lambda x(0) + \mu x(1,1) = 0 \tag{2.9}$$

$$-(\lambda + \mu)x(1,1) + \mu x(2,1)E = 0 \tag{2.10}$$

$$\lambda x(i-1,1) - (\lambda+\mu)x(i,1) + \mu x(i+1,1)E = 0, \; 2 \leq i \leq N-2 \tag{2.11}$$

$$\lambda x(N-2,1) - (\lambda+\mu)x(N-1,1) + \mu x(N)E = 0 \tag{2.12}$$

$$\lambda x(N-1,0)E + \lambda x(N-1,1) - (\lambda+\mu)x(N) + \mu x(N+1)E = 0 \tag{2.13}$$

$$\lambda x(i-1) - (\lambda+\mu)x(i) + \mu x(i+1)E = 0, \; i \geq N+1 \tag{2.14}$$

Now equations (2.7) and (2.8) give

$$x(i,0) = x(0) \quad \text{for} \quad 1 \leq i \leq N-1. \tag{2.15}$$

Taking $x(0) = \eta(1, 1, \cdots 1)$, equation (2.9) gives

$$x(1,1) = \frac{\lambda}{\mu}\eta(1, 1 \cdots 1) \tag{2.16}$$

Now, from equation (2.10) we get

$$x(2,1) = (1 + \frac{\lambda}{\mu})(\frac{\lambda}{\mu})\eta(1, 1 \cdots 1)E^{-1}$$

which gives

$$x(2,1) = (\frac{\lambda}{\mu} + (\frac{\lambda}{\mu})^2)\eta(1, 1 \cdots 1). \tag{2.17}$$

Equation (2.11) gives

$$x(i,1) = ((\frac{\lambda}{\mu}) + (\frac{\lambda}{\mu})^2 + \cdots + (\frac{\lambda}{\mu})^i)\eta(1, \cdots 1) \text{ for } 3 \leq i \leq N-1 \tag{2.18}$$

and equation (2.12) gives

$$x(N) = (\frac{\lambda}{\mu} + (\frac{\lambda}{\mu})^2 + \cdots + (\frac{\lambda}{\mu})^N)\eta(1, \cdots 1) \tag{2.19}$$

then equation (2.13) gives

$$x(N+1) = (\frac{\lambda}{\mu})((\frac{\lambda}{\mu}) + (\frac{\lambda}{\mu})^2 + \cdots + (\frac{\lambda}{\mu})^N)\eta(1, 1 \cdots 1).$$

That is,

$$x(N+1) = \frac{\lambda}{\mu}x(N). \tag{2.20}$$

Now, from equation (2.14) we get

$$x(N+i) = (\frac{\lambda}{\mu})^i x(N) \quad \text{for} \quad i \geq 2. \tag{2.21}$$

Now, to find η, we use normalizing condition,

$$x(0)e + \sum_{i=1}^{N-1} x(i)e + \sum_{i=N}^{\infty} x(i)e = 1. \tag{2.22}$$

Under the assumption that $\frac{\lambda}{\mu} < 1$, (2.22) give

$$\frac{N(S-s)}{1-\rho}\eta = 1.$$

hence

$$\eta = \frac{1-\rho}{N(S-s)}. \tag{2.23}$$

Thus we get the steady-state probability vector explicitly as:

Theorem 2.4.1. *The steady-state probability vector* X *of* Q *partitioned as* $X = (x(0), x(1), x(2), \ldots)$ *where each* $x(i)$*,* $1 \le i \le N-1$*, again partitioned as* $x(i) = (x(i,0), x(i,1))$*,* $1 \le i \le N-1$*, is given by*

$$\begin{aligned}
x(0) &= \eta(1, 1 \cdots) \\
x(i,0) &= \eta(1, 1, \cdots 1), \quad 1 \le i \le N-1 \\
x(i,1) &= \left[\sum_{j=1}^{i} (\tfrac{\lambda}{\mu})^j\right] \eta(1, 1 \cdots 1), \quad 1 \le i \le N-1 \\
x(N) &= \left[\sum_{j=1}^{N} (\tfrac{\lambda}{\mu})^j\right] \eta(1, 1 \cdots 1) \\
x(N+i) &= (\frac{\lambda}{\mu})^i x(N) \quad \textit{for} \quad i \ge 1
\end{aligned}$$

where $\eta = \frac{1-\rho}{N(S-s)}$ *and* $\rho = \frac{\lambda}{\mu}$*.*

Remark 2.4.2. *If we put* $N = 1$ *in theorem 2.4.1 we get steady-state probabilities for the system with service of customers whenever there is at least one waiting:*

Probability of i *customers and* j *items in inventory*

$$= (\frac{\lambda}{\mu})^i (1 - \frac{\lambda}{\mu}) \frac{1}{S-s} \quad \textit{for} \quad i \ge 0.$$

Note that the first two factors constitute the system state probability in the classical $M|M|1$ *queue and the third one corresponds to inventory level probability in a system with negligible service time, zero lead time and infinite shortage cost. Thus we have a product form solution for the system state probability. This is the case even when we adopt the* N*-policy for service.*

2.5 System performance measures

1. Average number of customers in the system

$$\mu_c = \sum_{i=1}^{\infty} ix(i)e = \frac{N-1}{2} - \frac{\rho}{N}\sum_{i=1}^{N-1} i\rho^i$$
$$+ \frac{\rho}{1-\rho}(1-\rho^N) + (\frac{\rho}{1-\rho})^2\frac{(1-\rho^N)}{N}.$$

2. Average inventory size

$$\mu_{inv} = \sum_{j=s}^{S-1} jx(0,j) + \sum_{j=s}^{S-1} j\{\sum_{i=1}^{N-1} x(i,0,j)\} + \sum_{j=s+1}^{S} j$$
$$\{\sum_{i=1}^{N-1} x(i,1,j)\} \sum_{j=s+1}^{S} j\{\sum_{i=N}^{\infty} x(i,j)\} = \frac{s+S-1}{2} + \rho.$$

3. Probability that the server is idle $P_{\text{idle}} = x(0)e + \sum_{i=1}^{N-1} x(i,0)e = 1-\rho$.

4. Expected number of customers in the system when the server is inactive

$$\mu_c^1 = \frac{1}{1-\rho}\sum_{i=1}^{N-1} ix(i,0)e = \frac{N-1}{2}.$$

5. Expected inventory in the system when the server is inactive

$$\mu_{inv}^1 = \frac{1}{1-\rho}\{\sum_{j=s}^{S-1} jx(0,j) + \sum_{j=s}^{S-1} j\{\sum_{i=1}^{N-1} x(i,0,j)\} = \frac{s+S-1}{2}.$$

6. Expected number of customers when the server is active

$$\mu_c'' = \frac{1}{\rho}\Big\{\sum_{i=1}^{N-1} ix(i,1)e + \sum_{i=N}^{\infty} ix(i)e\Big\}$$
$$= \frac{N-1}{2} - \frac{1}{N}\Big\{\sum_{i=1}^{N-1} i\rho^i\Big\} + \frac{1-\rho^N}{1-\rho} + \frac{\rho}{(1-\rho)^2}\frac{(1-\rho^N)}{N}.$$

7. Average replenishment rate

$$\mu_R = \Big\{ \sum_{i=2}^{N-1} x(i,1,s+1) + \sum_{i=N}^{\infty} x(i,s+1) \Big\} \mu + x(N-1,0,s)\lambda$$
$$= \frac{\lambda}{S-s}.$$

8. Expected length of a busy cycle

$$E_{\text{busy}} = \frac{N}{\lambda(1-\rho)}.$$

Next, we use these performance measures to construct a cost function for the model described above.

2.5.1 Cost function

We define a cost function as follows:

$$F(s,Q,N) = h_1\mu_{inv} + h_2\mu_c' + c_1 P_{\text{idle}} + (K + c_2 Q)\mu_R + \frac{A}{E_{\text{busy}}} + h_3\mu_{inv}'$$

where h_1 is the holding cost per unit time per item, h_2 is the holding cost per customer per unit time when the server is inactive, c_1 is the loss per unit time due to the server kept inactive and h_3 is the holding cost of inventoried items where the server is inactive. K is the fixed cost per order, c_2 is the variable procurement cost per item, $Q = S - s$ and A is the initial set up cost or activation cost of the server. Substituting in the expression for the system parameters we get

$$F(s,Q,N) = h_1\big(\frac{s+S-1}{2} + \rho\big) + h_2\big(\frac{N-1}{2}\big) + h_3\big(\frac{s+S-1}{2}\big)$$
$$+ c_1(1-\rho) + (K + c_2 Q)\frac{\lambda}{S-s} + \frac{A\lambda(1-\rho)}{N}.$$

Because the above function is linear in s, we notice that when shortage cost is infinity and lead time is zero, the optimal s value is given by $s^* = 0$. Furthermore, we notice that there is no direct relationship between the values of S and N except that on the average $\frac{S-1}{2}$ units of the item is held in the inventory during the inactivation time of the server. Hence, F can be regarded as separable in the variables S and N. Then, we notice that the optimal value of S is given by $\sqrt{\frac{2K\lambda}{h_1+h_3}}$. We also see that the optimal value of N is given by $N^* = \sqrt{\frac{2A\lambda(1-\rho)}{h_2}}$. Thus the expected minimum cost of the system is

$$2\sqrt{\frac{K\lambda(h_1+h_3)}{2}} + (1-\rho)c_1 + \lambda c_2 + \sqrt{\frac{5}{2}A\lambda(1-\rho)h_2}. \qquad (2.24)$$

The optimal order quantity agrees with the classified formula for EOQ. It is also seen that the server need be activated on accumulation of sufficiently large number of customers when the activation cost is very high. The expression for N^* also tells that if the holding cost of inventory and/or the waiting cost of customers (incurred to the system) when the server is not active, are/is high then relatively smaller values of N turns out to be optimal.

2.5.2 Numerical illustration

Table 2.1: $S = 25$, $\lambda = 5.0$, $\mu = 6.0$, $h_1 = 20.0$, $h_2 = 2.5$, $h_3 = 20.0$, $K = 100.0$, $c_2 = 50.0$, $c_1 = 2.0$, $A = 500$.

N	2	4	6	8	10
F	874.91	793.25	781.02	786.16	798.25

In Table 2.1, the average per unit time cost is indicated against various values of N for given input parameters. The cost decreases with increasing

Table 2.2: $\lambda = 5.0$, $\mu = 6.0$, $h_1 = 20.0$, $h_2 = 2.5$, $h_3 = 20.0$, $K = 100.0$, $c_2 = 50.0$, $c_1 = 2.0$, $N = 5$, $A = 500$.

S	15	16	17	18	19	20	21	22
F	750.33	743.66	741.76	742.83	745.89	750.33	755.78	762

Table 2.3: $S = 20$, $\lambda = 5.0$, $h_1 = 20.0$, $h_2 = 2.5$, $h_3 = 20.0$, $K = 100.0$, $c_2 = 50.0$, $c_1 = 2.0$, $N = 5$.

μ	5.1	5.5	6.0	6.5	7.0	7.5
F	791.95	726.31	750.33	777.06	801.47	823.17

value of N, attains a minimum and then goes up. We notice from Table 2.2 that the above observation is true for F as a function of S. Interestingly such an observation could be made about F when viewed as a function of μ as well in Table 2.3.

2.6 Conclusion

In this chapter, an (s, S) inventory system, where customers require a random amount (positive) of service, is discussed. In addition to the control variables s and S, we have brought in another control variable, namely, N. Each time the server becomes idle it is switched off and to be activated again on accumulation of N customers. It is proved that the optimal value s^* of s is $s^* = 0$ if no shortage is permitted (shortage cost is infinity and lead time is zero). It is also established that N and S have global optimal values. The expression for the minimum cost is given by equation (2.24).

The optimal order quantity agrees with the classical formula for EOQ. It is also seen that the server need be activated on accumulation of sufficiently large number of customers when the activation cost is very high (see the expression for the optimal value of N). The expression for N^* also tells that if the holding cost of inventory and /the waiting cost of customers (incurred to the system) when the server is not active, are/is high then relatively smaller values of N turns out to be optimal. Thus, the results obtained are quite useful for optimal design of an inventory system where service time is not negligible.

Chapter 3

Effective Utilization of Idle Time in an (s, S) Inventory with Positive Service Time

3.1 Introduction

In the previous chapter we were looking exclusively at the N-policy of activation of server for cost minimization. In this chapter we have a different motive, namely to improve the utilization of the server and also reduce the waiting time of customers to the maximum extent possible. It turns out that some customers have negligible service time with zero waiting time in the queue and the others will have to wait. This is achieved through processing (servicing) inventoried items while server does not have any customer to serve. The waiting time of a customer at the service counter is caused by the time required to process the inventoried item. The server can do this even in the absence of customers and this is exactly what we assume in this

---

* The content of this chapter is published in 'Journal of Applied Mathematics and Stochastic Analysis' Article ID 69068, 1–13, 2006.Also presented in the 'Joint Statistical Meeting and International Conference on Staistics, Probability and Related areas' during January 2-5, 2007 at Cochin University of Science and Technology

chapter. Thus we note that, unlike in the previous chapter, the optimal reorder level here is NOT zero but a positive quantity even in the absence of lead time and shortage cost.

However this seemingly simple "effective utilization" principle results in analytical intractability of the model! Nevertheless it is interesting to note that Deepak et al. [16] produces analytical result to the following simple variance of model discussed here and that in chapter 2: there are two streams of customers, one stream asking for unprocessed items (and hence with zero waiting and negligible service time) whereas the other asks for processed items resulting in positive service time and sometimes waiting time in queue.

Berman *et al.* [8] introduced positive service time in inventory models in a purely deterministic set up. Berman and Kim [5] deal with stochastic inventory models with positive service time wherein the average cost is minimised using dynamic programming technique. Berman and Sapna [9] consider an (s, S) inventory problem with positive service time. They consider a finite capacity waiting station model and identify a Markov renewal process to study the system behaviour. Since the state space is finite the Markov renewal equation has a unique solution.

3.2 Description of the model and its mathematical formulation

We have a single commodity inventory system operating under the (s, S) policy, with positive service time. Items are serviced (processed — eg. assembling) even in the absence of a demand. The processed items are stacked separately. The server keeps processing the items. The total of processed and unprocessed items can not exceed S. Also when the total reaches s an order for replenishment is placed and the order materialisation takes place instantly (ie, lead time is zero).

Demands for the item arrive according to a Poisson process of rate λ. Service (processing) times are exponentially distributed with parameter μ. A queue of customers is formed in the absence of processed items at demand epochs. There is no bound assumed to the queue. Thus the queue of customers can be arbitrarily large. We investigate its stability. The investigation of the maximum number of processed items that may be stacked is also important. This is so since it is much more expensive to stack processed items than unprocessed items. Also note that unlike in classical inventory with zero lead time and inventory with positive service time and zero lead time, we expect a positive reorder level (s) as optimal since this will reduce waiting time of customers and the consequent cost, thus bringing down the otherwise avoidable holding cost of customers.

Denote by $N(t)$, the number of customers in the system at time t; by $I(t)$ the total inventory (processed plus unprocessed) at t; and by $C(t)$ the number of processed items. Thus $\{(N(t), I(t), C(t)), t \in R_+\}$ is a three dimensional Markov chain on

$$\{(i, j, k)|i = 0, S \geq j \geq s{+}1, j \geq k \geq 0\} \cup \{(i, j, 0)|i > 0, S \geq j \geq s{+}1\}.$$

$C(t)$ can be positive only when number of customers waiting in the system is zero. We investigate the optimal (s, S) values and also the maximum number of processed items that could be stored so that the average system running cost is minimum.

3.3 Analysis of the system

Let $\underline{0} = \Big((0, s+1, 0), (0, s+1, 1), \ldots (0, s+1, s+1), (0, s+2, 0) \ldots (0, s+2, s+2), \ldots (0, S, 0), (0, S, 1) \ldots (0, S, S)\Big)$ and $\underline{i} = \Big((i, s+1, 0), (i, s+2, 0), \ldots (i, S, 0)\Big)$, for $i = 1, 2, \ldots$.

Then the resulting process is a level independent Quasi birth-death process (LIQBD) with the infinitesimal generator

$$Q = \begin{array}{c} \\ \underline{0} \\ \underline{1} \\ \underline{2} \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} \underline{0} & \underline{1} & \underline{2} \end{array} \\ \begin{pmatrix} A_{00} & A_{01} & 0 & & & \cdots \\ A_{10} & A_1 & A_0 & 0 & & \cdots \\ 0 & A_2 & A_1 & A_0 & & \cdots \\ & \ddots & \ddots & \ddots & & \\ 0 & & A_2 & A_1 & A_0 & \cdots \\ & & & & & \end{pmatrix} \end{array}$$

where

$$A_{00} = \begin{bmatrix} B_{s+2} & 0 & \cdots & C'_{s+1} \\ C_{s+2} & B_{s+3} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & C_S & B_{S+1} \end{bmatrix}$$

$$A_{01} = \begin{bmatrix} D_1 \\ D_2 \\ \vdots \\ D_{S-s} \end{bmatrix}$$

$$A_{10} = \begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \mu & 0 & \cdots & 0 \\ \mu & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \mu & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 \\ \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \mu & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

is a matrix of order

$$(S-s) \times \left((S-s)s + \frac{(S-s)(S-s+1)}{2} + S - s\right)$$

$$A_0 = \lambda I_{S-s},$$

$$A_1 = -(\lambda + \mu) I_{S-s},$$

$$A_2 = \begin{bmatrix} 0 & \cdots & \mu \\ \mu & 0 \cdots & 0 \\ 0 & \mu \cdots & 0 \\ 0 & \cdots \mu & 0 \end{bmatrix}_{(S-s)\times(S-s)},$$

$$B_j = \begin{bmatrix} -(\lambda+\mu) & \mu & 0 & \cdots & 0 \\ 0 & -(\lambda+\mu) & \mu & 0 \cdots & 0 \\ \vdots & & & & \\ 0 & & & -(\lambda+\mu) & \mu \\ 0 & & & & -\lambda \end{bmatrix}_{j \times j}$$

$$C_j = \begin{bmatrix} 0 \\ \\ \lambda I_j \end{bmatrix}_{(j+1)\times j} \qquad C'_{s+1} = \begin{bmatrix} 0 & 0 \\ \lambda I_{s+1} & 0 \end{bmatrix}_{(s+2)\times(S+1)}$$

$$D_1 = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \\ \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{(s+2)\times(S-s).}$$

$$D_2 = \begin{bmatrix} 0 & \lambda & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}_{(s+3)\times(S-s)}$$

$$D_{S-s} = \begin{bmatrix} 0 & \cdots & \cdots & \lambda \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{(S+1)\times(S-s).}$$

3.3.1 Calculation of steady state probabilities

Let $\pi = (\pi_0, \pi_1, \pi_2, \cdots)$ be the stationary probability vector associated with Q. Note that π_i is the probability vector associated with level i. Then $\pi Q = 0$ and $\pi \mathbf{e} = 1$ where $\mathbf{e}$ is the column vector of 1's of appropriate dimension. The π_i are given by

$$\pi_i = \pi_1 R^{i-1} \quad \text{for} \quad i \geq 2 \tag{3.1}$$

where R is the minimal non-negative solution of the matrix equation $A_0 + RA_1 + R^2 A_2 = 0$. π_0 and π_1 are calculated from the equations

$$\pi_0 A_{00} + \pi_1 A_{10} = 0 \tag{3.2}$$

$$\pi_0 A_{01} + \pi_1 A_1 + \pi_2 A_2 = 0. \tag{3.3}$$

Using (3.1), (3.3) can be rewritten as

$$\pi_0 A_{01} + \pi_1 (A_1 + RA_2) = 0. \tag{3.4}$$

From (3.2), we have

$$\pi_0 = \pi_1 (-A_{10}) A_{00}^{-1}. \tag{3.5}$$

Using (3.5) in (3.4) we get

$$\pi_1 (-A_{10}) A_{00}^{-1} A_{01} + \pi_1 (A_1 + RA_2) = 0$$

$$\text{i.e.,} \quad \pi_1 ((-A_{10}) A_{00}^{-1} A_{01} + A_1 + RA_2) = 0 \tag{3.6}$$

Thus π_0 and $\pi_i (i \geq 2)$ are expressed in terms of π_1, and π_1 can be obtained by solving (3.6) subject to the condition that

$$\pi_1[(-A_{10})A_{00}^{-1}\mathbf{e} + \mathbf{e} + R(I - R)^{-1}\mathbf{e}] = 1. \tag{3.7}$$

Let $T = -A_{10}A_{00}^{-1}A_{01} + A_1 + RA_2$. Then (3.6) reads as $\pi_1 T = 0$. Write $T = T_1 - T_2$ where $T_1 = T_U + T_D$ and $T_2 = -T_L$. Then by block Gauss-Seidel method, the recursive scheme of equations is given by $\pi_1^{(l+1)} T_1 = \pi_1^{(l)} T_2$.

3.3.2 Stability criterion

At first glance, it may appear that the stability condition can be weaker than $\lambda < \mu$. Such suspicion arises out of the fact that some customers have zero waiting time. However, it turns out that a few of the items were processed by the server in the absence of the customers and hence such customers who encounter the system with processed items do not have to wait. In any case, service was given and thus a certain amount of time was already spent on processed items towards service. Hence here also we can expect that $\lambda < \mu$, which is true and its proof is given below.

Theorem 3.3.1. *The system is stable if and only if $\lambda < \mu$.*

This follows from the fact that if the rate of drift from level i to $i-1$ is greater than that to level $i+1$ (the two immediate neighbours of i and the process is LIQBD which is skip free to either direction) then $\pi A_2 \mathbf{e} > \pi A_0 \mathbf{e}$ (see Neuts [45]), where π is the stationary probability vector associated with $A = A_0 + A_1 + A_2$. This on simplification gives $\lambda < \mu$.

3.3.3 First passage time analysis.

Here we obtain expression for the first passage time probability from a level i to the level $i-1$ for $i \geq 1$. This provides the mean number of customers served during the transition of the number of customers from i to $i-1$. Also it provides the mean time for the above transition. These measures are helpful in the design of service facilities.

Let $G_{jj}(k,x)$ be the conditional probability that the Markov process, starting in the state $(i,j,0)$ (for $i > 1$), at time $t = 0$, reaches the level $i-1$ for the first time at or prior to x, after exactly k transitions to the left (*ie.*, after exactly k service completions) and does so by entering the state $(i-1, j', 0)$ for $s+1 \leq j' \leq S$. The matrix with elements $G_{jj'}(k,x)$ is denoted by $G(k,x)$. Let $G^*(z,\theta) = \sum_{k=1}^{\infty} z^k \int_0^{\infty} e^{-\theta x} dG(k,x)$.

Then, for $0 < z < 1$, $\theta > 0$, the Matrix $G^*(z,\theta)$ is the minimal non-negative solution to the equation,

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

$$\lim_{s \to 0, z \to 1} G^*(z,\theta) = G = (G_{jj'})$$

where

$$G_{jj'} = Pr\Big\{\tau < \infty, (N(\tau), M(\tau), M_1(\tau)) = (i-1, j', 0)\Big| \\ (N(0), M(0), M_1(0)) = (i,j,0)\Big\}$$

and τ is the first passage time from the level i to the level $i-1$

Let M_{ij} be the mean first passage time from the level $i (i > 1)$ to the level $i-1$, given that it started in the state $(i,j,0)$ and let $\tilde{M}_1$ be a row vector of dimension $S-s$ with elements M_{1j}. Let M_{2j} be the mean number of service completions during this first passage time and let $\tilde{M}_2$ be a row vector of dimension $S-s$ with elements M_{2j}.

Then

$$\tilde{M}_1 = -\frac{\partial}{\partial\theta}G^*(z,\theta)e|_{\theta=0,z=1} = -(A_1 + A_0(I+G))^{-1}e$$
$$\tilde{M}_2 = -(A_1 + A_0(I+G))A_2e\Big[\frac{\partial}{\partial z}G^*(z,\theta)e|_{\theta=0,z=1}\Big]$$

Similar to $G^*(z,\theta)$, $\tilde{M}_1$ and $\tilde{M}_2$, we define matrices $G^{*(1,0)}(z,\theta)$, $\tilde{M}_1^{(1,0)}$ and $\tilde{M}_2^{(1,0)}$ for the first passage time from the level 1 to the level 0 and $G^{*(0,0)}$ (z,θ), $\tilde{M}_1^{(0,0)}$ and $\tilde{M}_2^{(0,0)}$ for the first passage time from the level 0 to the level 0.
Then

$$G^{*(1,0)}(z,\theta) = z(\theta I - A_1)^{-1}A_{10} + (\theta I - A_1)^{-1}A_0G^*(z,\theta)G^{*(1,0)}(z,\theta)$$

and

$$G^{*(0,0)}(z,\theta) = (\theta I - A_{00})^{-1}A_{01}G^{*(1,0)}(z,\theta).$$

Hence

$$\tilde{M}_1^{(1,0)} = -\frac{\partial}{\partial\theta}G^{*(1,0)}(z,\theta)e|_{\theta=0,z=1} = -(A_1 + A_0G)^{-1}(A_0\tilde{M}_1 + e)$$
$$\tilde{M}_2 = \frac{\partial}{\partial z}G^{*(1,0)}(z,\theta)e|_{\theta=0,z=1} = -(A_1 + A_0G)^{-1}(A_0\tilde{M}_2 + A_{10}e)$$
$$\tilde{M}_1^{(0,0)} = -\frac{\partial}{\partial\theta}G^{*(0,0)}(z,\theta)e|_{\theta=0,z=1} = -A_{00}^{-1}(A_{01}\tilde{M}_1^{(1,0)} + e)$$
$$\tilde{M}_2^{(0,0)} = \frac{\partial}{\partial z}G^{*(0,0)}(z,\theta)e|_{\theta=0,z=1} = -A_{00}^{-1}A_{01}\tilde{M}_2^{(1,0)}$$

3.3.4 Performance measures

(a) Average queue size

$$\begin{aligned}
&= \sum_{i=1}^{\infty} i\pi_i \mathbf{e} \\
&= \pi_1 \mathbf{e} + \sum_{i=2}^{\infty} i\pi_1 R^{i-1} \mathbf{e} \\
&= \pi_1 (I + 2R + 3R^2 + \cdots)\mathbf{e} \\
&= \pi_1 (I - R)^{-2} \mathbf{e}
\end{aligned}$$

(b) If we partition π_i by states in level i as

$$\pi_i = (\pi_{i,s+1,0}, \ldots, \pi_{i,s+1,s+1}, \ldots \pi_{iSS}),$$

then the average inventory size (processed + unprocessed) is

$$= \sum_{i=0}^{\infty} \sum_{j=s+1}^{S} j\pi_{ij0} + \sum_{j=s+1}^{S} \sum_{k=1}^{j} j\pi_{0,j,k} = \sum_{i=0}^{\infty} \sum_{j=s+1}^{S} \sum_{k=0}^{j} j \times \pi_{ijk}.$$

(c) Average number of processed items held $= \sum_{j=s+1}^{S} \sum_{k=1}^{j} k\pi_{0jk}$.

(d) Average Service rate $= \lambda \sum_{j=s+1}^{S} \sum_{k=1}^{j} \pi_{0jk} + \mu \sum_{i=1}^{\infty} \sum_{j=s+1}^{S} \pi_{ij0}$.

(e) Probability of a customer getting serviced instantaneously (ie., his service time is negligible, ie., waiting time is zero) $= \sum_{j=s+1}^{S} \sum_{k=1}^{j} \pi_{0jk}$.

(f) Probability that a customer will have to wait for service

$$= \sum_{i=0}^{\infty} \sum_{j=s+1}^{S} \pi_{ij0}.$$

(g) Average replenishment rate $= \lambda \sum_{k=1}^{s+1} \pi_{0,s+1,k} + \mu \sum_{i=1}^{\infty} \pi_{i,s+1,0}$.

(h) π_{0jj} stands for the probability that there is no customer (first subscript) in the system, j items are in the inventory (middle subscript) and all these are in processed state (last subscript). Only when no customer is present, there can be processed inventory. Thus the probability that all these are processed is given by $\sum_{j=s+1}^{S} \pi_{0jj}$.

(i) Probability that there is no processed item in the inventory

$$= \sum_{i=0}^{\infty} \sum_{j=s+1}^{S} \pi_{ij0}.$$

(j) Probability that the inventory size is maximum (ie., S)

$$= \sum_{i=0}^{\infty} \pi_{iS0} + \sum_{k=1}^{S} \pi_{0Sk}.$$

(k) Average waiting time of an arbitrary customer in the system

$$= \frac{1}{\mu} \left[\sum_{j=s+1}^{S} \pi_{0j0} + \sum_{i=1}^{\infty} \sum_{j=s+1}^{S} i\pi_{ij0} \right].$$

(l) Probability that the inventory level moves from S back to S without any intervening arrivals having to wait

$$\sum_{i_1=1}^{S} \sum_{i_2=1}^{S-1} \cdots \sum_{i_{S-s}=1}^{s+1} \pi_{0,S,i_1} \pi_{0,S-1,i_2} \cdots \pi_{0,s+1,i_{S-s}}.$$

(m) Probability that during an S to S transition, all customers have to wait is given by

$$\sum_{i_0 \geq 1} \sum_{i_1 \geq i_0 - 1} \cdots \sum_{i_{S-(s+2)} \geq i_{S-(s+3)} - 1} \sum_{i_{S-(s+1)} \geq i_{S-(s+2)} - 1} \pi_{i_0,S,0} \pi_{i_1,S-1,0} \cdots \pi_{i_{S-(s+1)},s+1,0}$$

The reasoning is as follows:
There are i_0 (≥ 1) customers when a replenishment takes place. At the first departure epoch there are i_1 customers left $(i_1 \geq 0)$; if this is zero an arrival should take place before the next service commences, else the next service commences immediately; the $S-(s+1)$th service in that cycle, with $s+1$ items in the inventory, proceeds and leaves behind one or more customers at its departure and the next service starts. Also the replenishment takes place.

3.4 Control Problem

We notice a glaring departure from classical inventory and inventory with service time as in chapter 2 on the one side and the problem under discussion here on the otherside. In the former, the optimal reorder level is zero (whenever lead time is zero) whereas in the latter this is not always true.

Here a trade off between the waiting cost of customers and the holding cost of finished products has to be obtained. If the former is very high compared to the latter, always a positive reorder level (that too pretty high) is called for, whereas when the holding cost of finished products is very high in comparison with the cost towards the waiting of customers, the reorder level may tend towards zero. The numerical illustrations are suggestive of these observations.

Now we show numerically that it is optimal to place replenishment order before the inventory level drops to zero whenever the waiting cost of customers is very high compared to the holding cost of processed item. Since analytical expressions are not available for the system state probabilities, it is difficult to give a formal proof for the above statement. Nevertheless, it can intutively shown to hold. For when the customer waiting cost is very high in comparision with the holding cost of processed items, a heavy expenditure is incurred for the former in the absence of processed items. If processed item is available, the demand is immdiately met with the result that the waiting cost of customer is completely avoided. This is brought out through tables 3.4, 3.5 and 3.6. Note that the optimal values of s in these cases are 9, 3 and 1, respectively. In these tables, the service rates are given the values 2.5, 3.0, and 3.5 respectively; and all other parameters are kept fixed. In arriving at the results given in Tables 3.7, 3.8 and 3.9 we constructed a profit function

$$\hat{F}(k) = \pi_{0Sk}(k \cdot h_1' + (S-k)h_2' - \lambda h_2).$$

This is to numerically establish that for a given inventory level, say S, there is a corresponding optimal value for the number of processed items to be kept in the inventory.

Let the costs associated with the system operation be as follows. Fixed ordering cost $= K$, procurement cost $= c$ per unit item, holding cost of customers $= h_2$ per unit/time, holding cost of processed items $= h_1'$ per unit/time and holding cost of unprocessed items $= h_2'$ per unit/time.

We analyse the following cost function:

$$\begin{aligned} F(s,S) = {} & (K + c(S-s))\left(\lambda \sum_{k=1}^{s+1} \pi_{0,s+1,k} + \mu \sum_{i=1}^{\infty} \pi_{i,s+1,0}\right) \\ & + h_1' \sum_{j=s+1}^{S} \sum_{k=1}^{j} k\pi_{0jk} \\ & + h_2' \left(\sum_{i=0}^{\infty} \sum_{j=s+1}^{S} \sum_{k=0}^{j} j\pi_{ijk} - \sum_{j=s+1}^{S} \sum_{k=1}^{j} k\pi_{0jk} \right) \\ & + h_2 \pi_1 (I-R)^{-2} \mathbf{e}. \end{aligned}$$

The objective is to minimize this cost. Since analytical expressions are not available for the system state probabilities, we resort to numerical procedure. Tables 3.1, 3.2 and 3.3, respectively, show the effect of the maximum inventory level on the system running cost when service rates are 2.1, 2.5 and 3 respectively with λ fixed at 2. The values of all other parameters are also kept fixed. Tables 3.4, 3.5 and 3.6 indicate the effect of the replenishment level on the system running cost and other system parameters, with service rates varying as 2.5, 3.0 and 3.5 respectively. In all these tables, we notice that the cost first shows a decreasing trend, reaches a minimum and then climbs up.

Effect of S on the expected system cost. Here we take $\lambda = 2.0$, $s = 10$, $K = 500$, $C = 100$, $h_1' = 50$, $h_2' = 10$, $h_2 = 50.0$.
Table 3.4 to 3.6 show effect of replenishment level s on the expected system cost $\lambda = 2.0$, $S = 20$, $K = 50.0$, $c = 20.0$, $h_1' = 15.0$, $h_2' = 10.0$, $h_2 = 200.0$.

Tables 3.7 to 3.9 provide the profit gained by stacking processed items in the inventory. In these tables, we fix $\lambda = 2.0$, $S = 20$, $K = 50.0$, $c = 20.0$, $h_1' = 15.0$, $h_2' = 10.0$, $h_2 = 200.0$, the values of s and μ alone are varied.

Table 3.1: $\mu = 2.1$

S	12	16	20	24	25	26	27	28	30
Average queue size	11.203	10.784	10.495	10.264	10.212	10.162	10.113	10.067	9.978
Average inventory held	11.207	13.169	15.130	17.091	17.582	18.072	18.563	19.053	20.034
Average number of processed items	2.756	3.138	3.435	3.694	3.754	3.814	3.872	3.930	4.041
P(inventory contains only processed items)	0.0464	0.0465	0.0465	0.0466	0.0466	0.0466	0.0466	0.0466	0.0466
Average waiting time of a customer in system	5.602	5.393	5.248	5.132	5.106	5.081	5.057	5.034	4.989
Expected cost	1437.68	1122.88	1071.40	1058.70	1058.04	1058.01	1058.51	1059.47	1062.47

Table 3.2: $\mu = 2.5$

S	12	14	16	18	19	20	21	22	25
Average queue size	0.317	0.288	0.266	0.249	0.242	0.235	0.229	0.223	0.208
Average inventory held	11.354	12.356	13.356	14.355	14.854	15.354	15.853	16.353	17.850
Average number of processed items	7.541	8.022	8.429	8.796	8.968	9.136	9.298	9.457	9.913
P(inventory contains only processed items)	0.1975	0.1977	0.1979	0.1980	0.1981	0.1981	0.1982	0.1982	0.1983
Average waiting time of a customer in system	0.1587	0.1439	0.1332	0.1247	0.1210	0.1176	0.1145	0.1116	0.1039
Expected	944.04	746.21	694.26	677.72	675.25	675.06	676.56	679.32	692.58

Table 3.3: $\mu = 3.0$

S	12	14	16	18	19	20	21	22	25
Average queue size	0.020	0.017	0.015	0.014	0.013	0.012	0.012	0.011	0.0096
Average inventory held	11.474	12.476	13.478	13.978	14.478	14.979	15.479	16.480	17.981
Average number of processed items	9.333	9.889	10.374	10.602	10.822	11.036	11.246	11.653	12.241
P(inventory contains only processed items)	0.3326	0.3327	0.3328	0.3328	0.3328	0.3329	0.3329	0.3329	0.3330
Average waiting time of customer in system	0.010	0.008	0.007	0.007	0.0066	0.0063	0.006	0.0054	0.0048
Expected cost	879.46	713.31	677.01	672.11	671.50	673.73	677.92	690.23	714.49

Table 3.4: $\mu = 2.5$

S	0	3	5	8	9	10	11	12	15
Average queue size	1.640	0.920	0.623	0.346	0.285	0.235	0.193	0.160	0.091
Average inventory held	9.971	11.566	12.671	14.298	14.828	15.353	15.875	16.394	17.934
Average number of processed items	2.123	3.892	5.290	7.562	8.345	9.135	9.929	10.724	13.087
P(inventory contains only processed items)	0.1869	0.1926	0.195	0.1972	0.1977	0.1981	0.1984	0.1987	0.1992
Average waiting time of customer in system	0.820	0.460	0.311	0.173	0.142	0.117	0.097	0.080	0.045
Expected cost	454.52	337.31	297.22	271.94	269.93	270.31	272.73	276.89	298.95

Table 3.5: $\mu = 3.0$

S	0	1	2	3	4	5	6	10	15
Average queue size	0.442	0.309	0.216	0.150	0.104	0.072	0.050	0.012	0.002
Average inventory held	10.128	10.691	11.253	11.808	12.355	12.893	13.423	15.479	17.995
Average no. of processed items	3.472	4.173	4.915	5.683	6.466	7.259	8.056	11.246	15.138
P(inventory contains only processed items)	0.317	0.3219	0.325	0.328	0.329	0.331	0.331	0.333	0.333
Average waiting time of a customer in system	0.221	0.155	0.108	0.075	0.052	0.036	0.025	0.006	0.001
Expected cost	213.2	196.2	187.18	183.84	184.5	187.95	193.35	225.53	278.89

Table 3.6: $\mu = 3.5$

S	0	1	2	3	4	5	7	10	15
Average queue size	0.183	0.111	0.067	0.04	0.024	0.014	0.005	0.001	0.0001
Average inventory held	10.250	10.825	11.38	11.918	12.45	12.965	13.986	15.497	17.999
Average no. of processed items	4.452	5.2	5.965	6.738	7.513	8.288	9.833	12.133	15.88
P(inventory contains only processed items)	0.414	0.420	0.423	0.425	0.427	0.427	0.428	0.428	0.429
Average waiting time of customer in the system	0.092	0.055	0.033	0.020	0.012	0.007	0.003	0.0006	0.00005
Expected cost	163.81	158.96	159.71	163.87	170.09	177.55	194.48	222.32	274.7

Table 3.7: $\mu = 2.5$, $s = 9$

k	1	4	7	8	9	10	11	15	19
$\hat{F}(k)$	0.32	0.574	0.807	0.844	3.825	2.056	1.104	0.091	0.007

Table 3.8: $\mu = 3.0$, $s = 3$

k	1	2	3	4	5	8	11	15	19
$\hat{F}(k)$	0.231	0.272	3.821	2.231	1.301	0.257	0.05	0.006	0.0006

Table 3.9: $\mu = 3.5$, $s = 1$

k	1	2	3	4	7	11	15	19
$\hat{F}(k)$	3.539	2.194	1.359	0.842	0.199	0.029	0.004	0.0006

3.5 Conclusion

In this chapter we have considered an effective way to improve the server idle-time utilization. Eventhough the lead time is assumed to be zero, we notice that when the holding cost of customers is very high, reorder level remains positive. There are some customers requiring negligible service time (as in classical inventory models) and some others requiring positive service time. The need for stacking processed items is also brought out in this chapter.

Chapter 4

Inventory with Service Time and Transfer of Customers and/Inventory

4.1 Introduction

In the previous chapters we considered single server inventory systems with positive service time. We note that the problem in inventory with service time may appear as a problem in queue. This chapter clubs inventory with positive service time of customers waiting in two different queues with their transfer from the longer to shorter queue to improve performance of the combined system. In other words an inventory system with two parallel service facilities is considered. A certain number of customers are transferred from longer to shorter queue whenever their difference reaches a prescribed quantity. Along with this customer transfer, a certain quantity of inventory is also transferred, depending on availability. Further, if one of the queues has customers, but has no inventoried items whereas the other has at least one inventoried item to spare, then exactly one item is taken to the former

[*] The content of this paper will appear in 'Annals of Operations Research' (2008).To be presented at the'Sixth International Workshop in Matrix Analytic Method' at Beijing, June 2008.

and service begins, thereby enhancing the efficiency of the combined system and what is more is that the combined system may be stable even when one of the queues is unstable. Stability of the system is analysed. Several performance measures that help in efficient design of such systems, are computed. Some numerical results are provided.

Parallel queues providing identical service, where customers jockey for reaching the server with minimal waiting, is a common sight. Here the system itself does not provide a means to reduce the waiting time of customers. One among the queue may build up whereas the other may remain pretty small. Inorder to bring down the waiting time of customers a transfer of customers from the longer to the shorter queue by the system rather than by the customers themselves is desirable. As far as our information goes, the first work to deal with customer transfer by the system (from longer to shorter queue) in parallel queues was considered by Qi-Ming He and Neuts [23]. Despite the fact that one of the queues in a two-parallel queue system with customer transfer is unstable, the combined system can be stable. Their model runs as follows: there are two infinite capacity waiting lines manned by one server each. Arrival process to the first and second lines form independent Poisson processes of rates λ_1 and λ_2 respectively. Service times are exponentially distributed with parameters μ_1 and μ_2 and are independent of each other. Upto this stage these are two independent systems. The dependence comes in through the following: when the numerical difference between the number of customers in the two channels reaches L, a certain number K $(K < L)$ of customers are transferred from the longer to the shorter queue. Despite this the two waiting lines can grow to any large size and thus the matrix analytic method fails. However, if we now view the combined system described by a two dimensional vector, the first coordinate of which is defined as the sum of the number of customers in the two queues and the second coordinate as their difference, then the first one can be taken as the level and the second one, the phase. Since the second coordinate can assume only a finite number of values, the matrix analytic method can be employed to the two dimensional Markov chain thereby obtained. Still the system turns out to be a level dependent Quasi birth-death process (LDQBD). In order to reduce it to a level independent QBD a number of transformations are made and the resulting continuous time Markov chain is analysed. Qi-Ming He and Neuts [23] provide a number of extremely useful measures of the queueing system with transfer of customers that can effectively be utilized as control measures. In particular they demonstrate that under $\lambda_1 + \lambda_2 < \mu_1 + \mu_2$, the system is collectively stable even when one component is unstable. This means that the system performs better with transfer of customers. In the present set up, such a simple stability condition cannot be obtained due to the complexity of the model. It may be noted that if we drop the assumption of transferring one unit of inventory to the queue where customer(s) are waiting in the absence of inventory, then the same stability condition as obtained in [23] results.

Inventory with positive service time is considered by Berman *et al.* [8], Berman and Kim [6], Berman and Sapna [10], Parthasarathy and Vijayalakshmi [48]. Arivarignan *et al.* [2] consider a continuous review perishable inventory control system with positive service time. One of the most recent contributions to inventory with service time is due to Schwarz *et al.* [54]. Basically their model assumes Poisson arrival of demands, exponentially distributed service time and balking of customers when inventory level is zero (lost sales situation). They derive joint stationary distributions of queue length and inventory level in explicit product form under continuous review of inventory level and different inventory management policies. With lead time assumed to be exponentially distributed, they evaluate various performance measures. The key result they obtained is that with infinite waiting room capacity the limiting distributions of the queue length processes are the same as in the classical $M|M|1|\infty$ system. They could arrive at analytical solutions because of the assumption that no customer is permitted to join the system when inventory level is zero.

4.2 Mathematical formulation and Analysis

In this chapter we extend the Qi-Ming He-Neuts model to an inventory with service time as described below. There are two service counters manned by one server each. Customers arrive to these according to two independent Poisson processes of rates λ_1 and λ_2 respectively, demanding inventoried item—exactly one item will be served to a customer. Server at counter i serves according to exponentially distributed time with parameter μ_i, $i = 1, 2$. At both counters the same (s, S)-inventory policy is adopted for simplifying the discussion. The maximum that can be held in the inventory is S; the reorder level is s. Thus always $s < S$. However, here we make a stronger assumption that $S > 2s + 1$, the reason for which becomes evident while we proceed further. (s can be interpreted as the safety stock which could be used to serve customers during lead time). At each order placement epoch the order quantity is $Q = S - s$. Here we assume that the lead time at counter i is exponentially distributed with parameter δ_i, $i = 1, 2$.

We follow the policy of transfer of customers as in Qi-Ming He and Neuts [23]. In addition to transfer of K customers from the longer to the shorter queue, whenever their difference reaches L, we transfer certain quantity of inventory also. For example suppose queue 1 is longer than queue 2 by L units. Then K customers are chosen from the last among the waiting and passed on to the queue 2. Simultaneously we transfer a certain number of inventoried items from queue 1 to queue 2 which is given by $Inv_{1\to 2} = \min\{[I_1(t) - I_2(t)]^+, K'\}$ where K' is a positive integer less than Q provided $I_2(t) > 0$. ($[x]^+$ is the largest integer part (≥ 0) in x). If $I_2(t) = 0$, then the quantity transferred is $I_1(t) - 1$. Several clarifications are needed at this stage: (i) Note that the reorder level is s. However it can happen that $I_1(t) > s$ and consequent to the transfer the inventory level drops to level s or below that, triggering immediate order placement. Then how much to order? We restrict the order quantity to $S - s = Q$, irrespective of whether the inventory level at reorder epoch turns out to be any one among s, $s-1,\ldots,0$. (ii) An equally uncomfortable situation arises when customers are received along with a certain amount of inventoried items. For example suppose K customers are transferred from queue 1 to queue 2. Suppose that replenishment order is outstanding in queue 2 at this epoch. Further assume that as a consequence of the transfer of some inventory to queue 2, the inventory level in that queue has exceeded s, the reorder level. Subsequently the replenishment may take place when the inventory level remains above s. This will result in the inventory position in queue 2 going above S (the maximum that can be held there). (iii) Further it may appear now as if there is no pending replenishment order in queue 2, unless we introduce an entity to describe that status. To overcome all these unpleasant situations, we assume that, if an order for replenishment is pending at the queue to which customers and inventory are transferred, then that particular order is cancelled in case the inventory level goes above s consequent to the transfer. (iv) In addition if at a service completion epoch if one of the queues still has customers waiting but service could not be given for want of inventory, then depending on the availability of items in the other queue, exactly one unit of inventory is transferred to the one without inventory currently, provided it does not interrupt the service of the customer, if any, in service at the queue from which the item is transferred. (One can very well consider transfer of more than one unit of inventory). Without this assumption the structure of the infinitesimal generator of the process would have been quite handy as we will realize in the section to follow. Nevertheless these inventory transfers will ensure smoother running of the service system, by reducing waiting time thereby enhancing the efficiency of the combined system.

We suitably define the state variables so as to enable us to make the associated Markov process a QBD for which the matrix analytic method can be effectively used. With that observation we consider the process

$$\Big\{(q(t) = q_1(t) + q_2(t), J(t) = q_1(t) - q_2(t), I_1(t), I_2(t)); t \geq 0\Big\}$$

which is a QBD process with $q(t) \geq 0$ and $-(L-1) \leq J(t) \leq L-1$.

Generally, $0 \leq I_i(t) \leq S$ for $i = 1, 2$. Within these limits, however, the values assumed by $I_1(t)$ and $I_2(t)$ very much depend on the values of $(q(t), J(t))$. The state space of the above Markov chain can be divided into levels according to the values of $q(t)$. The states in each level are given as follows.

i) For $q(t) = 0$, the only possible value for $J(t)$ is zero so that the level 0, denoted by $l(0)$ is given by

$$\begin{aligned} l(0) = \{&(0,0,0,0)\ldots(0,0,0,S)(0,0,1,0)\ldots(0,0,1,S)\ldots \\ &(0,0,S,0)\ldots(0,0,S,S)\} \end{aligned}$$

ii) When $q(t) = n$; $1 \leq n \leq L-1$, the possible values for $J(t)$ are $n, n-2, n-4, \ldots -(n-4), -(n-2)$ so that for $1 \leq n \leq L-1$, the level n, denoted by $l(n)$ is given by

$$\begin{aligned} l(n) = &\Big\{(n,n,0,0)\ (n,n,1,0)\ (n,n,1,1)\ldots\ (n,n,1,S) \\ &\ldots\ \ldots\ (n,n,S,0)\ \ldots\ \ldots\ (n,n,S,S)\Big\} \\ &\cup\Big\{(n,n-2,0,0)\ (n,n-2,0,1)\ (n,n-2,1,0)\ \ldots \\ &(n,n-2,1,S)(n,n-2,2,1)\ \ldots(n,n-2,2,S)\ \ldots \\ &(n,n-2,S,1)\ldots(n,n-2,S,S)\Big\} \\ &\ldots\ \ldots\ \ldots\ \ldots \\ &\cup\Big\{(n,-n,0,0)\ (n,-n,0,1)\ \ldots\ (n.-n,0,S)\ (n,-n,1,1)\ \ldots \\ &(n,-n,1,S)\ldots\ (n,-n,S,1)\ldots(n,-n,S,S)\Big\}. \end{aligned}$$

i.e., $l(n)$ has $(n-1)(S^2-3) - 2(S(S-1)-1)$ states.

iii) For $n \geq L$ there are two cases:
1) If $n \geq L$ and $n - L$ is odd, then the possible values of $J(t)$ are $L-1, L-3 \cdots -(L-3), -(L-1)$.

So in this case

$$
\begin{aligned}
l(n) = \Big\{ & (n, L-1, 0, 0)\ (n, L-1, 0, 1)\ (n, L-1, 1, 0)\ \ldots \\
& (n, L-1, 1, S), (n, L-1, 2, 1)\ \ldots\ (n, L-1, 2, S) \ldots\ldots \\
& (n, L-1, S, 1) \ldots\ldots (n, L-1, S, S) \Big\} \cup\ \ldots\ldots\ldots\ldots \\
\cup \Big\{ & (n, -(L-1), 0, 0)\ (n, -(L-1), 0, 1) \\
& (n, -(L-1), 1, 0)\ \ldots (n, -(L-1), 1, S), \\
& (n, -(L-1), 2, 1)\ \ldots (n, -(L-1), 2, S)\ \ldots \\
& (n, -(L-1), S, 1)\ \ldots (n, -(L-1), S, S) \Big\}.
\end{aligned}
$$

i.e., $l(n)$ has $L(S^2 + 3)$ states.

2) If $n \geq L$ and $n - L$ is even, $J(t)$ assumes one of the values in the set $\{L-2, L-4, \ldots, -(L-4), -(L-2)\}$. Moreover, in this case the values assumed by $I_1(t)$ and $I_2(t)$ are the same as in the previous case so that here we arrange the states in the same order as in case 1. Hence in this case $l(n)$ has $(L-1)(S^2-3)$ states.

Now we introduce a few frequently occurring notations and abbreviations used in the sequel.

I_j : Identity matrix of order j.

0_j : Column vector of order j with all entries zeros.

e : column vector of 1's of appropriate order.

δ_{ij} : Kronecker delta.

$\otimes$: Kronecker product.

$$H_0 = [1, 0]_{1\times 2}; \quad H_1 = \begin{bmatrix} I_2 \\ 0 \end{bmatrix}_{(S+1)\times 2}; \quad H_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}_{2\times 2};$$

$$I' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_{S-1} & 0 \end{bmatrix}_{(S+1)\times(S+1)}; \quad \hat{I} = \begin{bmatrix} 0 & 0 \\ I_{S-1} & 0 \end{bmatrix}_{S\times S};$$

$$\tilde{I} = \begin{bmatrix} 0 & 0 \\ I_S & 0 \end{bmatrix}_{(S+1)\times(S+1)};$$

$$I'_j = \begin{bmatrix} 0 & 0 \\ 0 & I_{S-j+1} \end{bmatrix}_{(S+1)\times(S+1)} \text{ for } 1 \le j \le S;$$

$$I''_j = \begin{bmatrix} I_j & 0 \\ 0 & 0 \end{bmatrix}_{S\times S} \text{ for } 1 \le j \le S-1;$$

$$I'''_j = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I_{S-j+1} & 0 \end{bmatrix}_{S\times S} \text{ for } 3 \le j \le S;$$

$$\tilde{I}'_j = \begin{bmatrix} 0 & 0 \\ 0 & I_{S-j} \end{bmatrix}_{S\times S} \text{ for } 1 \le j \le S-1;$$

$$\tilde{I}''_j = \begin{bmatrix} I_j & 0 \\ 0 & 0 \end{bmatrix}_{(S+1)\times(S+1)} \text{ for } 1 \le j \le S;$$

$$\hat{I}'_j = \begin{bmatrix} 0 & 0 \\ 0 & I_{S-j+1} \end{bmatrix}_{(S+1)\times S} \text{ for } 1 \le j \le S \text{ with } \hat{I}'_1 = \begin{bmatrix} 0 \\ I_S \end{bmatrix}_{(S+1)\times S};$$

$$\pi_L(I-R)^{-1})_{n(S^2+3)+iS+j} = U(n,i,j);$$
$$\left(\pi_n\right)_{S(S+1)+n(S^2+3)+iS+j} = V(n,i,j) \text{ for } 1 \le n \le L-1$$

where $(\pi_n)_i$ represents the i^{th} component of the row vector π_n.
The matrices B_{ij}, B'_{ij}, B''_{ij} and B'''_{ij} are of orders $(S+1)\times(S+1)$, $(S+1)\times S$, $S\times S$ and $S\times(S+1)$ respectively, with the property that the $(i,j)^{\text{th}}$ entry of all these matrices is equal to 1 and all other entries are zero.

Now the infinitesimal generator matrix of the Markov process $\{(q(t),$

$J(t), I_1(t), I_2(t)), t \geq 0\}$ is given as:

$$Q_1 = \begin{bmatrix} A_{01} & A_{00} & & & & & \\ A_{12} & A_{11} & A_{10} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & A_{L-1,2} & A_{L-1,1} & A_{L-1,0} & \dots & \\ & & & A_{L,2} & A_{L,1} & A_{L,0} & \dots \\ & & & & A_{L+1,2} & A_{L+1,1} & A_{L+1,0} \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}.$$

Each block matrix appearing in Q_1 is explicitly expressed as follows.

$$A_{01} = \begin{bmatrix} I_{s+1} \otimes F_0 & 0 & \delta_1 I_{(s+1)(S+1)} \\ 0 & 0 & I_{S-s} \otimes F_1 \end{bmatrix}_{(S+1)^2 \times (S+1)^2}$$

where

$$F_0 = \begin{bmatrix} -\psi I_{s+1} & 0 & \delta_2 I_{s+1} \\ 0 & 0 & (-\psi + \delta_2) I_{S-s} \end{bmatrix}_{(S+1) \times (S+1)}$$

and

$$F_1 = \begin{bmatrix} (-\psi + \delta_1) I_{s+1} & 0 & \delta_2 I_{s+1} \\ 0 & 0 & -(\lambda_1 + \lambda_2) I_{S-s} \end{bmatrix}_{(S+1) \times (S+1)}$$

with $\psi = \lambda_1 + \lambda_2 + \delta_1 + \delta_2$.

Note that in the following the blocks in the matrices are indexed by $J(t) = q_1(t) - q_2(t)$ which has the values $\Big\{n, n-2, n-4, \dots -(n-2), -n\Big\}$ for $0 \leq l(n) \leq L-1$; $\Big\{L-2,\ L-4, \dots -(L-4), -(L-2)\Big\}$ and $\Big\{L-1, L-3, \dots -(L-3), -(L-1)\Big\}$ for $l(n) \geq L$.
Then

$$A_{00} = (A_{00}^{(1)} \quad A_{00}^{(2)}) \quad \text{where} \quad A_{00}^{(1)} = \begin{bmatrix} A_{00}^{(1,1)} & A_{00}^{(1,2)} \\ 0 & A_{00}^{(1,3)} \end{bmatrix}_{(S+1)^2 \times (S(S+1)+1)}.$$

In $A_{00}^{(2)}$, the $(1,1)^{\text{th}}$ block entry is $A_{00}^{(2,0)}$; $(2,1)^{\text{th}}$ entry is $A_{00}^{(2,1)}$; $(i,i)^{\text{th}}$ entry is $A_{00}^{(2,2)}$ for $2 \le i \le S+1$; $(i,i-1)^{\text{th}}$ entry is $A_{00}^{(2,3)}$ for $3 \le i \le S+1$ and remaining entries are zero.

The matrices appearing in each block are explicitly expressed as
$A_{00}^{(1,1)} = \begin{bmatrix} \lambda_1 \\ 0_S \end{bmatrix}$,
$A_{00}^{(1,2)} = (\lambda_1 \tilde{I}, 0)_{(S+1)\times S(S+1)}$;
$A_{00}^{(1,3)} = \lambda_1 I_{S(S+1)}$, $A_{00}^{(2,0)} = \lambda_2 I_{S+1}$, $A_{00}^{(2,1)} = \lambda_2 B_{12}$, $A_{00}^{(2,2)} = \lambda_2 \hat{I}_1'$ and $A_{00}^{(2,3)} = \lambda_2 B_{11}'$.

Now, $A_{12} = \begin{bmatrix} A_{12}^{(1)} \\ A_{12}^{(2)} \end{bmatrix}$ where $A_{12}^{(1)} = \begin{bmatrix} 0 & 0 \\ \mu_1 I_{S(S+1)} & 0 \end{bmatrix}_{(S(S+1)+1)\times(S+1)^2}$;

$A_{12}^{(2)} = \begin{bmatrix} A_{12}^{(2,1)} & 0 \\ 0 & I_S \otimes A_{12}^{(2,2)} \end{bmatrix}$ where $A_{12}^{(2,1)} = \mu_2 \tilde{I}$
and $A_{12}^{(2,2)} = \begin{bmatrix} \mu_2 I_S & 0 \end{bmatrix}_{S\times(S+1)}$.

For $2 \le n \le L-1$,

$$A_{n2} = \begin{array}{c} \\ n \\ n-2 \\ \vdots \\ \vdots \\ \vdots \\ -(n-2) \\ -n \end{array} \begin{pmatrix} n-1 & n-3 & \cdots & \cdots & -(n-3) & -(n-1) \\ A_2^{(1)} & & & & & \\ A_2^{(2)} & A_2^{(3)} & & & & \\ & A_2^{(4)} & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & A_2^{(3)} & \\ & & & & A_2^{(4)} & A_2^{(5)} \\ & & & & & A_2^{(6)} \end{pmatrix}$$

where $A_2^{(1)}$ is a square matrix of order $S(S+1)+1$ with the only non-zero blocks $A_2^{(1,2)}$ as the $(2,2)^{\text{th}}$ (block) entry, $A_2^{(1,1)}$ as the $(2,1)^{\text{th}}$ entry and the other lower diagonal entries are equal to $\mu_1 I_{S+1}$.

$$A_2^{(2)} = \begin{bmatrix} A_2^{(2,1)} & 0 & 0 & \\ 0 & A_{12}^{(2,1)} & 0 & \\ 0 & 0 & I_{S-1}\otimes & A_{12}^{(2,2)} \end{bmatrix}_{(S^2+3)\times(S(S+1)+1)} ;$$

$A_2^{(3)}$ is a square matrix of order (S^2+3) with the lower diagonal (block) entries $A_2^{(3,4)}$ except the first two which are $A_2^{(3,1)}$ and $A_2^{(3,3)}$ respectively, and the only non-zero diagonal (block) entry $A_2^{(3,2)}$ as the second one;
$A_2^{(4)}$ is a square matrix of order (S^2+3) with the diagonal (block) entries $A_2^{(4,4)}$ except the first two which are $A_2^{(4,1)}$ and $A_{12}^{(2,1)}$ respectively, and the lower diagonal entries $A_2^{(4,5)}$ except the first two which are zero and $A_2^{(4,3)}$ respectively;

$$A_2^{(5)} = \begin{bmatrix} 0 & 0 \\ \mu_1 I_{S^2+1} & 0 \end{bmatrix}_{(S^2+3)\times(S(S+1)+1)}$$

and $A_2^{(6)}$ is a square matrix of order $S(S+1)+1$ with the diagonal entries $A_2^{(4,4)}$ except the first one which is $A_{12}^{(2,1)}$ and the lower diagonal entries $A_2^{(4,5)}$ except the first one which is $A_2^{(4,3)}$.

The matrices appearing in each of the blocks $A_2^{(i)}$, $i = 1, 2, \ldots, 6$ are as given below.

$$A_2^{(1,1)} = \begin{bmatrix} \mu_1 \\ 0_S \end{bmatrix}; \qquad A_2^{(1,2)} = \mu_1 \tilde{I}; \qquad A_2^{(2,1)} = (0, \mu_2)^T;$$
$$A_2^{(3,1)} = \mu_1 H_1; \qquad A_2^{(3,2)} = \mu_1 I'; \qquad A_2^{(3,3)} = \begin{bmatrix} 0 & \mu_1 I_S \end{bmatrix}_{S\times(S+1)};$$
$$A_2^{(3,4)} = \mu_1 I_S; \qquad A_2^{(4,1)} = \mu_2 H_2; \qquad A_2^{(4,3)} = \mu_2 B_{12}''';$$
$$A_2^{(4,4)} = \mu_2 \hat{I} \quad \text{and} \quad A_2^{(4,5)} = \mu_2 B_{11}''.$$

For $1 \leq n \leq L-2$,

$$A_{n0} = \begin{array}{c} \\ n \\ n-2 \\ \vdots \\ \vdots \\ -(n-2) \\ -n \end{array} \begin{array}{c} \begin{array}{cccccc} n+1 & n-1 & \cdots & \cdots & -(n-1) & -(n+1) \end{array} \\ \begin{pmatrix} A_0^{(1)} & A_0^{(2)} & & & & \\ & A_0^{(3)} & A_0^{(4)} & & & \\ & \ddots & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & A_0^{(3)} & A_0^{(4)} & \\ & & & & A_0^{(5)} & A_0^{(6)} \end{pmatrix} \end{array}$$

where

$$A_0^{(1)} = \lambda_1 I_{S(S+1)+1}; \qquad A_0^{(2)} = \begin{bmatrix} A_0^{(2,1)} & 0 \\ 0 & \tilde{A}_{00}^{(2)} \end{bmatrix}$$

where $A_0^{(2,1)} = \lambda_2 H_0$ and $\tilde{A}_{00}^{(2)}$ is obtained from $A_{00}^{(2)}$ by eliminating the last row (block); $A_0^{(3)} = \lambda_1 I_{S^2+3}$;

$$A_0^{(4)} = \lambda_2 I_{S^2+3}; \qquad A_0^{(5)} = \begin{bmatrix} A_0^{(5,1)} & A_0^{(5,2)} & 0 \\ 0 & A_0^{(5,3)} & 0 \\ 0 & 0 & \lambda_1 I_{S(S-1)} \end{bmatrix}$$

and $A_0^{(6)} = \lambda_2 I_{S(S+1)+1}$.

The matrices appearing in $A_0^{(5)}$ are expressed as

$$A_0^{(5,1)} = \lambda_1 H_1; \qquad A_0^{(5,2)} = \lambda_1 I' \text{ and } A_0^{(5,3)} = \begin{bmatrix} 0 & \lambda_1 I_S \end{bmatrix}_{S \times (S+1)}.$$

For $1 \leq n \leq L-1$,

$$A_{n1} = \text{diag}(A_1^{(1)}, e^T \otimes A_1^{(2)}, A_1^{(3)})$$

where e is of order $n-1$ and if $A_1^{(i)}(j,k)$ represents the entry (block) in $A_1^{(i)}$ corresponding to the pair of indices (j,k) where $0 \leq j,k \leq S$, then the non-zero (block) entries in $A_1^{(i)}$ are given by $A_1^{(1)}(0,0) = A_1^{(1,1)}$; $A_1^{(1)}(0,1) = A_1^{(1,2)}$; $A_1^{(1)}(i,i) = A_1^{(1,4)}$ for $1 \leq i \leq s$; $A_1^{(1)}(i,i) = A_1^{(1,6)}$ for $s+1 \leq i \leq S$;

$A_1^{(1)}(0, S-s) = A_1^{(1,3)}$; and $A_1^{(1)}(i, S-s+i) = A_1^{(1,5)}$ for $1 \leq i \leq s$.
$A_1^{(2)}(0,0) = A_1^{(2,1)}$; $A_1^{(2)}(1,1) = A_1^{(2,5)}$; $A_1^{(2)}(i,i) = A_1^{(2,8)}$ for $2 \leq i \leq s$; $A_1^{(2)}(i,i) = A_1^{(2,9)}$ for $s+1 \leq i \leq S$; $A_1^{(2)}(0,1) = A_1^{(2,2)}$; $A_1^{(2)}(0, S-s-1) = A_1^{(2,3)}$; $A_1^{(2)}(1, S-s) = A_1^{(2,6)}$; $A_1^{(2)}(2, S-s+1) = A_1^{(2,7)}$; and $A_1^{(2)}(i, S-s+i) = A_1^{(2,10)}$ for $2 \leq i \leq s$
$A_1^{(3)}(0,0) = A_1^{(3,1)}$; $A_1^{(3)}(i,i) = A_1^{(3,4)}$ for $1 \leq i \leq s$; $A_1^{(3)}(i,i) = A_1^{(3,6)}$ for $s+1 \leq i \leq S$; $A_1^{(3)}(0, S-s-1) = A_1^{(3,2)}$; $A_1^{(3)}(0, S-s) = A_1^{(3,3)}$ and $A_1^{(3)}(i, S-s+i) = A_1^{(3,5)}$ for $1 \leq i \leq s$.

The matrices appearing in each of the blocks $A_1^{(i)}$; $i = 1,2,3$ are given as $A_1^{(1,1)} = -\psi$; $A_1^{(1,2)} = [0, \delta_2, 0]_{1\times(S+1)}$ where δ_2 lies at the $(S-s)^{\text{th}}$ position;

$A_1^{(1,3)} = [\delta_1, 0]_{1\times(S+1)}$; $A_1^{(1,4)} = F_0 - \mu_1 I_{S+1}$; $A_1^{(1,5)} = \delta_1 I_{S+1}$;
$A_1^{(1,6)} = A_1^{(1,4)} + A_1^{(1,5)}$; $A_1^{(2,1)} = \text{diag}(-\psi, -\psi-\mu_2)$;
$A_1^{(2,2)} = (0 \;\; \delta_2 I_2 \;\; 0)_{2\times(S+1)}$
where δ_2 in the first row lies at the $(S-s)^{\text{th}}$ position;

$$A_1^{(2,3)} = \begin{bmatrix} \delta_1 & 0 \\ 0 & 0 \end{bmatrix}_{2\times S}; \quad A_1^{(2,4)} = \begin{bmatrix} 0 & 0 \\ \delta_1 & 0 \end{bmatrix}_{2\times S};$$

$A_1^{(2,5)} = F_0 - \mu_2 I_1' - \mu_1 I_{S+1}$; $A_1^{(2,6)} = \delta_1 B_{11}'$; $A_1^{(2,7)} = \delta_1 \hat{I}_1'$; $A_1^{(2,8)}$ is obtained from $A_1^{(2,5)}$ by eliminating the first row and column;

$$A_1^{(2,9)} = A_1^{(2,10)} + A_1^{(2,8)}; A_1^{(2,10)} = \delta_1 I_S;$$

$A_1^{(3,1)} = F_0 - \mu_2 I_1'$; $A_1^{(3,2)} = A_1^{(2,6)}$; $A_1^{(3,3)} = A_1^{(2,7)}$; $A_1^{(3,4)}$ is obtained from $A_1^{(3,1)}$ by eliminating first row and column; $A_1^{(3,5)} = \delta_1 I_S$ and $A_1^{(3,6)} =$

$A_1^{(3,4)} + A_1^{(3,5)}$.

$$A_{L-1,0} = \begin{array}{c} \\ L-1 \\ L-3 \\ \vdots \\ \vdots \\ \vdots \\ -(L-3) \\ -(L-1) \end{array} \begin{array}{c} \begin{array}{ccccccc} L-2 & L-4 & \cdots & L-2K & \cdots & -(L-4) & -(L-2) \end{array} \\ \begin{pmatrix} A_0^{(2)} & & & \lambda_1 D & & & \\ \lambda_1 I_{S^2+3} & \lambda_2 I_{S^2+3} & & & & & \\ \ddots & \ddots & & & & & \\ & \ddots & \ddots & & & & \\ & & \ddots & \ddots & & & \\ & & & & & \lambda_1 I_{S^2+3} & \lambda_2 I_{S^2+3} \\ & & \lambda_2 D' & & & & A_0^{(5)} \end{pmatrix} \end{array}$$

Note that in the definition of $A_{L-1,0}$, the block matrix corresponding to the pair of indices $(L-1, L-2K)$ is $\lambda_1 D$ and that corresponding to $(-(L-1), -(L-2K))$ is $\lambda_2 D'$ (because when $J(t) = \pm(L-1)$, an arrival may cause a transfer) where

$$D = \begin{bmatrix} H_0 & & & & & & & & \\ & I_{S+1} & & & & & & & \\ & \sum_{i=1}^{2} B_{i,i+1} & \hat{I'}_2 & & & & & & \\ & \sum_{i=1}^{2} B_{i,i+2} & B'_{33} & \hat{I'}_3 & & & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & & & & \\ & \sum_{i=1}^{2} B_{i,i+K'} & B'_{3,K'+1} & \cdots & B'_{K'+1,K'+1} & \hat{I'}_{K'+1} & & & \\ & & \sum_{i=1}^{3} B'_{i,i+K'-1} & & \cdots & B'_{K'+2,K'+2} & \hat{I'}_{K'+2} & & \\ & & \ddots & & & & & \ddots & \\ & & & \sum_{i=1}^{S-K'+1} B'_{i,i+K'-1} & B'_{S-K'+2,S} & \cdots & & B'_{SS} & \hat{I'}_S \end{bmatrix}$$

and

$$D' = \begin{bmatrix} H_1 & B_{32} & B'_{41} & \cdots & B'_{K'+1,1} & \sum_{i=2}^{S-K'+1} B_{i+K',i-1} & & & & \\ & B'''_{12} & B''_{21} & \cdots & \cdots & B''_{K',1} & \sum_{i=2}^{S-K'+1} B''_{i+K'-1,i-1} & & & \\ & & I''_2 & B''_{32} & \cdots & \cdots & B''_{K'+1,2} & & & \\ & & & \ddots & & & & & & \\ & & & & \ddots & & & & & \\ & & & & I''_{S-K'} & B''_{S-K'+1,S-K'} & \cdots & B''_{S,S-K'} & & \\ & & & & & & & & I''_{S-1} & B''_{S,S-1} \\ & & & & & & & & & I_S \end{bmatrix}.$$

Now for the level L of Q_1, we have

$$A_{L2} = \begin{array}{c} \\ L-2 \\ L-4 \\ \vdots \\ -(L-4) \\ -(L-2) \end{array} \begin{array}{c} \begin{array}{ccccc} L-1 & L-3 & \cdots & -(L-3) & -(L-1) \end{array} \\ \begin{pmatrix} A_2^{(2)} & A_2^{(3)} & & & \\ & A_2^{(4)} & \ddots & & \\ & & \ddots & & \\ & & & A_2^{(3)} & \\ & & & A_2^{(4)} & A_2^{(5)} \end{pmatrix} \end{array};$$

$A_{L1} = I_{L-1} \otimes A_1^{(2)}$ and

$$A_{L0} = \begin{array}{c} \\ L-2 \\ L-4 \\ \vdots \\ -(L-4) \\ -(L-2) \end{array} \begin{array}{c} \begin{array}{cccccc} L-1 & L-3 & \cdots & \cdots & -(L-3) & -(L-1) \end{array} \\ \begin{pmatrix} A_0^{(3)} & A_0^{(4)} & & & & \\ & A_0^{(3)} & A_0^{(4)} & & & \\ & & \ddots & \ddots & & \\ & & & & & \\ & & & & A_0^{(3)} & A_0^{(4)} \end{pmatrix} \end{array}.$$

For the level $L+1$, we have

$$A_{L+1,2} = \begin{array}{c} \\ L-1 \\ L-3 \\ \vdots \\ \vdots \\ -(L-3) \\ -(L-1) \end{array} \begin{array}{c} \begin{array}{ccccccc} L-2 & L-4 & \cdots & L-2K & \cdots & -(L-4) & -(L-2) \end{array} \\ \begin{pmatrix} A_2^{(3)} & & & \mu_2 E & & & \\ A_2^{(4)} & A_2^{(3)} & & & & & \\ & & \ddots & \ddots & & & \\ & & & & \ddots & \ddots & \\ & & & & & A_2^{(4)} & A_2^{(3)} \\ & & \mu_1 E' & & & & A_2^{(4)} \end{pmatrix} \end{array}.$$

Here note that the block matrix corresponding to the pair of indices $(L-1, L-2K)$ is $\mu_2 E$ and that corresponding to $(-(L-1), -(L-2K))$ is $\mu_1 E'$ (because when $J(t) = \pm(L-1)$, a service completion results in a transfer) where

$$E = \begin{bmatrix} H_2 & & & & & & \\ & \tilde{I} & & & & & \\ & \sum_{i=1}^{2} B'''_{i,i+1} & I'''_3 & & & & \\ & \sum_{i=1}^{2} B'''_{i,i+2} & B''_{33} & I'''_4 & & & \\ & \cdots & \cdots & & & & \\ & \sum_{i=1}^{2} B'''_{i,i+K'} & B''_{3,K'+1} & \cdots & B''_{K'+1,K'+1} & I'''_{K'+2} & \\ & & \sum_{i=1}^{3} B''_{i,i+K'-1} & B''_{4,K'+2} & \cdots & \cdots & I'''_{K'+3} \\ & & & \ddots & \ddots & & \\ & & & & \sum_{i=1}^{S-K'+1} B''_{i,i+K'-1} & B''_{S-K'+2,S} & \cdots B''_{SS} \end{bmatrix};$$

and $E' = \begin{bmatrix} 0 \\ \tilde{D}' \end{bmatrix}$ where the matrix $\tilde{D}'$ is obtained from D' by eliminating the last row (block).

$$A_{L+1,1} = I_L \otimes A_1^{(2)} \text{ and}$$

$$A_{L+1,0} = \begin{matrix} & L-2 & L-4 & \cdots & L-2K & \cdots & -(L-2) \\ L-1 & \lambda_2 I_{S^2+3} & & & \lambda_1 D'' & & \\ L-3 & \lambda_1 I_{S^2+3} & \lambda_2 I_{S^2+3} & & & & \\ \vdots & & & \ddots & \ddots & & \\ -(L-3) & & & & & & \lambda_2 I_{S^2+3} \\ -(L-1) & & & \lambda_2 D''' & & & \lambda_1 I_{S^2+3} \end{matrix}$$

where

$$D'' = \begin{bmatrix} I_2 & & & & & & & \\ I_{S+1} & & & & & & & \\ B'''_{13} & \tilde{I}'_1 & & & & & & \\ B'''_{14} & B''_{23} & \tilde{I}'_2 & & & & & \\ \cdots & \cdots & \cdots & & & & & \\ \cdots & \cdots & \cdots & & & & & \\ B'''_{1,K'+2} & B''_{2,K'+1} & B''_{3,K'+1} & \cdots & B''_{K',K'+1} & \tilde{I}'_{K'} & & \\ & \sum_{i=1}^{2} B''_{i,K'+i} & B''_{3,K'+2} & \cdots & \cdots & B''_{K'+1,K'+2} & \tilde{I}'_{K'+1} & \\ & & \ddots & \sum_{i=1}^{S-K'} B''_{i,K'+i} & B''_{S-K'+1,S} & \cdots & B''_{S-1,S} & \tilde{I}'_{S-1} \end{bmatrix}$$

and

$$D''' = \begin{bmatrix} I_2 & & & & & & & & \\ & \tilde{I}_2'' & B'_{31} & B'_{41} & \cdots & B'_{K'+1,1} & \sum\limits_{i=2}^{S-K'+1} B'_{K'+i,i-1} & & \\ & & I_2'' & B''_{32} & \cdots & \cdots & B''_{K'+1,2} & \sum\limits_{i=3}^{S-K'+1} B''_{K'+i,i-1} & \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \\ & & & & I''_{S-K'} & B''_{S-K'+1,S-K'} & \cdots & B''_{S-1,S-K'} & B''_{S,S-K'} \\ & & & & & I''_{S-K'+1} & \cdots & & B''_{S,S-K'+1} \\ & & & & \ddots & \cdots & \cdots & \cdots & \\ & & & & & \ddots & \cdots & \cdots & \\ & & & & & & & I''_{S-1} & B''_{S,S-1} \\ & & & & & & & & I_S \end{bmatrix}.$$

For the level $L+2$,

$$A_{L+2,2} = \begin{array}{c} \\ L-2 \\ L-4 \\ \vdots \\ \vdots \\ -(L-4) \\ -(L-2) \end{array} \begin{array}{c} \begin{array}{cccccc} L-1 & L-3 & \cdots & \cdots & -(L-3) & -(L-1) \end{array} \\ \begin{pmatrix} A_2^{(4)} & A_2^{(3)} & & & & \\ & A_2^{(4)} & A_2^{(3)} & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \ddots \\ & & & & A_2^{(4)} & A_2^{(3)} \end{pmatrix} \end{array}.$$

and for $n \geq 0$, $i = 0, 1$, we have $A_{L+2n,i} = A_{Li}$ and $A_{L+2n+1,i} = A_{L+1,i}$. Also for $n \geq 1$, $A_{L+2n,2} = A_{L+2,2}$ and for $n \geq 0$, $A_{L+2n+1,2} = A_{L+1,2}$.

Thus all the matrices (blocks) arising in the infinitesimal generator Q_1 have been explicitly expressed. Now, the QBD process $\{(q(t), J(t), I_1(t), I_2(t)) : t \geq 0\}$ is level dependent. Analysis of a level dependent QBD process is in general difficult. However, as in [23] we are fortunate to generate a level independent QBD process by reorganising the state space. The details are as follows:
For $n \geq 0$, the level $L+2n$ has $(L-1)(S^2+3)$ states and the level $L+2n+1$ has $L(S^2+3)$ states. Regroup the $(2L-1)(S^2+3)$ states in the levels $L+2n$ and $L+2n+1$ and call the new level as $L+n$.
After regrouping the states, the resulting QBD process becomes level independent. Corresponding to the new partition of the state space, we introduce

a new Markov process $\{(X(t), J(t), I_1(t), I_2(t)) : t \geq 0\}$ where

$$X(t) = \begin{cases} q(t) & \text{if } q(t) \leq L-1 \\ L + \left\lfloor \frac{q(t)-L}{2} \right\rfloor & \text{if } q(t) \geq L \end{cases}$$

where '$\lfloor x \rfloor$' is the largest integer less than or equal to x. Clearly, the Markov process $\{(X(t), J(t), I_1(t), I_2(t)) : t \geq 0\}$ is irreducible and level independent.

The infinitesimal generator Q_2 of $\{(X(t), J(t), I_1(t), I_2(t)) : t \geq 0\}$ is given by

$$Q_2 = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ L-1 \\ L \\ L+1 \\ L+2 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & 1 & 2 & \cdots & L & L+1 & L+2 & \cdots \end{array} \\ \begin{pmatrix} A_{01} & A_{00} & & & & & & \\ A_{12} & A_{11} & A_{10} & & & & & \\ \ddots & \ddots & \ddots & & & & & \\ & A_{L-1,2} & A_{L-1,1} & A^*_{L-1,0} & & & & \\ & & A^*_{L2} & A_1 & A_0 & & & \\ & & & A_2 & A_1 & A_0 & & \\ & & & & A_2 & A_1 & A_0 & \\ & & & & \ddots & \ddots & \ddots & \end{pmatrix} \end{array}$$

where $A^*_{L-1,0} = [A_{L-1,0}, 0]$, $A^*_{L2} = \begin{bmatrix} A_{L2} \\ 0 \end{bmatrix}$, $A_1 = \begin{bmatrix} A_{L1} & A_{L0} \\ A_{L+1,2} & A_{L+1,1} \end{bmatrix}$,
$A_0 = \begin{bmatrix} 0 & 0 \\ A_{L+1,0} & 0 \end{bmatrix}$ and $A_2 = \begin{bmatrix} 0 & A_{L+2,2} \\ 0 & 0 \end{bmatrix}$.

Note that $X(t)$ and $q(t)$ are related by

$$\begin{aligned} & q(t) = 2X(t) - L \text{ if } X(t) \geq L, \quad J(t) - L \text{ is even;} \\ \text{and} \quad & q(t) = 2X(t) - L + 1 \text{ if } X(t) \geq L, \quad J(t) - L \text{ is odd.} \end{aligned}$$

4.2.1 Stability condition and Stationary distribution

We first investigate the condition for the system to be stable. Starting from the fact that the system is stable if and only if $\tilde{\pi} A_0 e < \tilde{\pi} A_2 e$, where $\tilde{\pi}$ is the stationary probability vector corresponding to the generator matrix $A = A_0 + A_1 + A_2$, we get

$$(\lambda_1 + \lambda_2)\tilde{\pi}_L e < \sum_{i=1}^{L-1} \tilde{\pi}_{L-1,i} a$$

where $\tilde{\pi} = (\tilde{\pi}_{L-1}, \tilde{\pi}_L)$ with $\tilde{\pi}_{L-1}$ being the vector of probabilities associated with the first $(L-1)$ blocks, and $\tilde{\pi}_L$ that associated with the last L blocks and

$$a = (0, \mu_2, \mu_1, \mu_1 + \mu_2, \dots, \mu_1 + \mu_2)^T$$

with $S^2 + 3$ entries.

When replenishment rate is infinity, this reduces to $\lambda_1 + \lambda_2 < \mu_1 + \mu_2$ as in [23]. This is on the expected lines since when replenishment rate is infinity, servers will remain idle only in the absence of customers.

Note that if the first three elements in a were also $\mu_1 + \mu_2$ then the stability condition would have reduced to $\lambda_1 + \lambda_2 < \mu_1 + \mu_2$ which is the case of infinite replenishment rate. The numerical results obtained are supportive of this state of affairs.

Since $\{(X(t), J(t), I_1(t), I_2(t)) : t \geq 0\}$ is a level independent QBD process, its stationary distribution (if exists) has a matrix geometric solution. We refer to Neuts [45] and Latouche and Ramaswami [40] for details about the matrix geometric solution of QBD processes.

Denote by π_n, the stationary probability vector of the above level independent QBD process corresponding to the level n. Then

$$\pi Q_2 = 0 \text{ and } \pi e = 1 \tag{4.1}$$

where $\pi = (\pi_0, \pi_1, \pi_2, \cdots, \cdots)$. Because of the special structure of Q_2, π can be expressed as

$$\pi_n = \pi_L R^{n-L} : n \geq L+1, \tag{4.2}$$

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

$$R^2 A_2 + R A_1 + A_0 = 0 \tag{4.3}$$

and the vectors $\pi_0, \pi_1, \ldots, \pi_{L-1}$ are obtained by solving

$$\begin{array}{rl} \text{(i)} & \pi_0 A_{01} + \pi_1 A_{12} = 0 \\ \text{(ii)} & \pi_{i-1} A_{i-1,0} + \pi_i A_{i1} + \pi_{i+1} A_{i+1,2} = 0;\ 1 \leq i \leq L-2 \\ \text{(iii)} & \pi_{L-2} A_{L-2,0} + \pi_{L-1} A_{L-1,1} + \pi_L A^*_{L_2} = 0 \\ \text{(iv)} & \pi_{L-1} A^*_{L-1,0} + \pi_L (A_1 + R A_2) = 0 \end{array} \tag{4.4}$$

subject to the normalizing condition

$$\sum_{i=0}^{L-1} \pi_i e + \pi_L (I - R)^{-1} e = 1. \tag{4.5}$$

By the appropriate manipulation of the above system of equations, we get

$$\begin{aligned} \pi_i &= \pi_{i+1} A_{i+1,2} (-A'_i)^{-1} \quad \text{if} \quad 0 \leq i \leq L-2 \\ \pi_{L-1} &= \pi_L A^*_{L2} (-A'_{L-1})^{-1} \\ \text{and} \qquad \pi_L &= \pi_{L+1} A_2 (-A'_L)^{-1} \end{aligned} \tag{4.6}$$

where

$$A'_i = \begin{cases} A_{01} & \text{if } i = 0 \\ A_{i1} + A_{i2} (-A'_{i-1})^{-1} A_{i-1,0} & \text{if } 1 \leq i \leq L-1 \\ A_1 + A^*_{L2} (-A'_{L-1})^{-1} A^*_{L-1,0} & \text{if } i = L. \end{cases} \tag{4.7}$$

Now, to calculate the non-boundary state probability vectors we proceed as follows.

By applying block Gaussian elimination, the partitioned subvector $[\pi_L, \pi_{L+1} \ldots]$ corresponding to non-boundary states satisfies the relation

$$[\pi_L, \pi_{L+1} \ldots] \begin{bmatrix} A'_L & A_0 & \ldots & & \\ A_2 & A_1 & A_0 & \cdots & \\ 0 & A_2 & A_1 & A_0 & \cdots \\ \cdots & & \cdots & & \\ \cdots & & \cdots & & \end{bmatrix} = 0^T. \tag{4.8}$$

Let

$$\delta = \sum_{i=L}^{\infty} \pi_i e \quad \text{and} \quad x_i = \delta^{-1} \pi_{L+i} \quad \text{if } i \geq 0. \tag{4.9}$$

From (4.8), So

$$\pi_L A'_L + \pi_{L+1} A_2 = 0 \quad \text{and} \quad \pi_{L+i} = \pi_{L+i-1} R : \quad i \geq 1. \tag{4.10}$$

or, sequentially $x_i = x_0 R^i$ for $i \geq 1$. Since $\sum_{i=0}^{\infty} x_i e = 1$, we get $x_0(I - R)^{-1} e = 1$. Now from (4.2), $\pi_i = \pi_L R^{i-L}$ for $i \geq L$.
Hence

$$\pi_i = \delta x_0 R^{i-L} \text{ for } i \geq L \text{ by using (4.9).} \tag{4.11}$$

From (4.6),

$$\begin{aligned} \pi_0 &= \pi_1 A_{12} (-A'_0)^{-1} \\ &= \pi_2 A_{22} (-A'_1)^{-1} A_{12} (-A'_0)^{-1} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &= \pi_{L-1} A_{L-1,2} (-A'_{L-2})^{-1} A_{L-2,2} (-A'_{L-3})^{-1} \ldots A_{12} (-A'_0)^{-1} \\ &= \pi_L A^*_{L2} (-A'_{L-1})^{-1} A_{L-1,2} (-A'_{L-2})^{-1} \ldots A_{12} (-A'_0)^{-1}. \end{aligned}$$

That is,

$$\pi_0 = \delta x_0 A^*_{L2} (-A'_{L-1})^{-1} \prod_{j=1}^{L-1} A_{L-j,2} (-A'_{L-j-1})^{-1}.$$

In general,

$$\pi_i = \delta x_0 A^*_{L2} (-A'_{L-1})^{-1} \prod_{j=1}^{L-i-1} A_{L-j,2} (-A'_{L-j-1})^{-1}, \; 0 \leq i \leq L-1.$$

Thus we have,

$$\pi_i = \delta x_0 A^*_{L2}(-A'_{L-1})^{-1} \prod_{j=1}^{L-i-1} A_{L-j,2}(-A'_{L-j-1})^{-1}$$
$$\text{if } \ 0 \le i \le L-1$$
$$\text{and} \quad \pi_i = \delta x_0 R^{i-L} \ \text{ if } \ i \ge L, \tag{4.12}$$

where x_0 is the unique solution of

$$x_0(A'_L + RA_2) = 0; \quad x_0(I-R)^{-1}e = 1. \tag{4.13}$$

Now $\pi e = 1$ implies $\sum_{i=0}^{\infty} \pi_i e = 1$.
i.e.,

$$\delta\big[x_0 A^*_{L2}(-A'_{L-1})^{-1} \sum_{i=0}^{L-1} \prod_{j=1}^{L-i-1} A_{L-j,2}(-A'_{L-j-1})^{-1}e + x_0(I-R)^{-1}e\big] = 1$$

so that

$$\delta = \big[1 + x_0 A^*_{L2}(-A'_{L-1})^{-1} \sum_{i=0}^{L-1} \prod_{j=1}^{L-i-1} A_{L-j,2}(-A'_{L-j-1})^{-1}e\big]^{-1} \tag{4.14}$$

since $x_0(I-R)^{-1}e = 1$.

In the next section, we discuss some important descriptors which are very useful in system designing.

4.2.2 Measures of effectiveness

Let $\pi_{L+n} = (\pi_{L+n}(0), \pi_{L+n}(1))$ where $\pi_{L+n}(0)$ corresponds to the level $L+2n$ and $\pi_{L+n}(1)$ corresponds to the level $L+2n-1$ in the level dependent case. Here remember that in the present level independent QBD process, the level $L+n$ is formed by combining levels $L+2n$ and $L+2n-1$ that existed in the level dependent QBD process. Then,

1. Mean number of customers in the system,

$$
\begin{aligned}
EN &= \sum_{n=0}^{L-1} n\pi_n e + \sum_{n=0}^{\infty}(L+2n)\pi_{L+n}(0)e + \sum_{n=0}^{\infty}(L+2n+1)\pi_{L+n}(1)e \\
&= \sum_{n=0}^{L-1}(n-L)\pi_n e + L + \pi_L[2R(I-R)^{-2}e + \begin{pmatrix} 0_{(L-1)(S^2+3)} \\ e \end{pmatrix}].
\end{aligned}
$$

2. Let $\{P_n(i), n \geq 0\}$ be the distribution of the queue length of queue i in steady state $i = 1, 2$. Then

$$
\begin{aligned}
P_{n1}(1) = P(q_1(t) = n_1) = &\sum_{n_2=\max\{0,n_1-L+1\}}^{\max\{0,L-n_1-1\}} \sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{n_1+n_2,n_1-n_2,i,j} \\
&+ \sum_{n_2=\max\{0,L-n_1-1\}+1}^{n_1+L-1} \sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{L+\lfloor\frac{n_1+n_2-L}{2}\rfloor,n_1-n_2,i,j} \quad : n_1 \geq 0.
\end{aligned}
$$

Similarly

$$
\begin{aligned}
P_{n2}(2) = P(q_2(t) = n_2) = &\sum_{n_1=\max\{0,n_2-L+1\}}^{\max\{0,L-n_2-1\}} \sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{n_1+n_2,n_1-n_2,i,j} \\
&+ \sum_{n_1=\max\{0,L-n_2-1\}+1}^{n_2+L-1} \sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{L+\lfloor\frac{n_1+n_2-L}{2}\rfloor,n_1-n_2,i,j} \quad : n_2 \geq 0.
\end{aligned}
$$

Even though in the sums we have generally taken $0 \leq i,\ j \leq S$, the variations of i and j depend on the particular values assumed by (n_1, n_2).

3. The probability that the queue size is zero is π_0e.

4. The probability that server 1 is idle,

$$
\begin{aligned}
P_1^{\text{idle}} &= \pi_0 e + \sum_{n=1}^{L-1}\sum_{i=0}^{S}\sum_{j=0}^{S}\pi_{n,-n,i,j} + \sum_{n_1=1}^{L-1-n_2}\sum_{n_2=0}^{L-2}\sum_{j=0}^{S}\pi_{n_1+n_2,n_1-n_2,0,j} \\
&\quad + \sum_{n_2=L-n_1}^{L+n_1-1}\sum_{n_1=1}^{L-1}\sum_{j=0}^{S}\pi_{L+\lfloor\frac{n_1+n_2-L}{2}\rfloor,n_1-n_2,0,j} \\
&\quad + \sum_{n_2=n_1-L+1}^{n_1+L-1}\sum_{n_1=L}^{\infty}\sum_{j=0}^{S}\pi_{L+\lfloor\frac{n_1+n_2-L}{2}\rfloor,n_1-n_2,0,j} \\
&= \pi_0 e + \sum_{n=1}^{L-1}(\pi_n)_1 + \sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=0,1}(\pi_n)_{S(S+1)+l(S^2+3)+j+2} \\
&\quad + \sum_{n=1}^{L-1}\sum_{j=0}^{S}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+j+2} \\
&\quad + \sum_{n=1}^{L-1}\sum_{i,j=1}^{S}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+is+j+2} \\
&\quad + \sum_{l=0}^{2L-2}\sum_{j=0,1}(\pi_L(I-R)^{-1})_{l(S^2+3)+j+1} \\
&= \pi_0 e + \sum_{n=1}^{L-1}(\pi_n)_1 + \sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=0,1}V(l,0,j+2) \\
&\quad + \sum_{n=1}^{L-1}\sum_{i=0}^{S}\sum_{j=1-\delta_{i0}}^{S}V(n-1,i,j+2) \\
&\quad + \sum_{l=0}^{2L-2}\sum_{j=0,1}U(l,0,j+1).
\end{aligned}
$$

The probability that server 2 is idle,

$$
\begin{aligned}
P_2^{\text{idle}} &= \pi_0 e + \sum_{n=1}^{L-1}\sum_{i=0}^{S}\sum_{j=0}^{S}\pi_{n,n,i,j} + \sum_{n_2=1}^{L-n_1-1}\sum_{n_1=0}^{L-2}\sum_{i=0}^{S}\pi_{n_1+n_2,n_1-n_2,i,0} \\
&\quad + \sum_{n_1=L-n_2}^{L+n_2-1}\sum_{n_2=1}^{L-1}\sum_{i=0}^{S}\pi_{L+\lfloor\frac{n_1+n_2-L}{2}\rfloor,n_1-n_2,i,0} \\
&\quad + \sum_{n_1=n_2-L+1}^{n_2+L-1}\sum_{n_2=L}^{\infty}\sum_{i=0}^{S}\pi_{L+\lfloor\frac{n_1+n_2-L}{2}\rfloor,n_1-n_2,i,0} \\
&= \pi_0 e + \sum_{n=1}^{L-1}(\pi_n)_1 + \sum_{n=1}^{L-1}\sum_{i=0}^{S-1}(\pi_n)_{i(S+1)+2} \\
&\quad + \sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=0,2}(\pi_n)_{S(S+1)+l(S^2+3)+j+2} \\
&\quad + \sum_{n=1}^{L-1}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+2} + \sum_{n=1}^{L-1}\sum_{i=0}^{S-1}\sum_{j=1}^{S}(\pi_n)_{i(S+1)+j+2} \\
&\quad + \sum_{l=0}^{2L-2}\sum_{j=0,2}(\pi_L(I-R)^{-1})_{l(S^2+3)+j+1}. \\
&= \pi_0 e + \sum_{n=1}^{L-1}\sum_{i=0}^{S-1}(\pi_n)_{i(S+1)+2} + \sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=0,2}V(l,0,j+2) \\
&\quad + \sum_{n=1}^{L-1}V(n-1,0,2) + \sum_{n=1}^{L-1}\sum_{i=0}^{S-1}\sum_{j=1}^{S}(\pi_n)_{i(S+1)+j+2} \\
&\quad + \sum_{l=0}^{2L-2}\sum_{j=0,2}U(l,0,j+1).
\end{aligned}
$$

Note that in this system, a server can be idle due to lack of customers or inventory.

5. Let $T^C_{R,1\to2}(T^C_{R,2\to1})$ be the transfer rate of customers from queue 1 to

queue 2 (queue 2 to queue 1). *i.e.*, the number of transfers (of customers) per unit time.
Then

$$\begin{aligned}
T^C_{R,1\to 2} &= \lambda_1 \sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{L-1,L-1,i,j} \\
&\quad + \lambda_1 \sum_{n=0}^{\infty}\sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{L+n,L-1,i,j} + \mu_2 \sum_{n=0}^{\infty}\sum_{i=0}^{S}\sum_{j=1}^{S} \pi_{L+n,L-1,i,j} \\
&= \lambda_1 \sum_{i,j=0}^{S} \pi_{L-1,L-1,i,j} + \lambda_1 \sum_{j=1}^{S^2+3} (\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+j} \\
&\quad + \mu_2 \sum_{\substack{j=2\\ j\neq 3}}^{S^2+3} (\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+j} \\
&= \lambda_1 \sum_{i,j=0}^{S} \pi_{L-1,L-1,i,j} \\
&\quad + \sum_{j=1}^{S^2+3} (\lambda_1 + \mu_2(1-\delta_{1j})(1-\delta_{3j})]U(L-1,0,j)
\end{aligned}$$

and

$$\begin{aligned}
T^C_{R,2\to 1} &= \lambda_2 \sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{L-1,-(L-1),i,j} \\
&\quad + \lambda_2 \sum_{n=0}^{\infty}\sum_{i=0}^{S}\sum_{j=0}^{S} \pi_{L+n,-(L-1),i,j} \\
&\quad + \mu_1 \sum_{n=0}^{\infty}\sum_{i=1}^{S}\sum_{j=0}^{S} \pi_{L+n,-(L-1),i,j}
\end{aligned}$$

$$= \lambda_2 \sum_{i,j=0}^{S} \pi_{L-1,-(L-1),i,j} + \lambda_2 \sum_{j=1}^{S^2+3} (\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+j}$$

$$+ \mu_1 \sum_{j=3}^{S^2+3} (\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+j}$$

$$= \lambda_2 \sum_{i,j=0}^{S} \pi_{L-1,-(L-1),i,j}$$

$$+ \sum_{j=1}^{S^2+3} (\lambda_2 + \mu_1(1-\delta_{1j})(1-\delta_{2j})]U(2L-2,0,j).$$

Note that a transfer from one station to the other can take place either through an arrival to the former or through a service completion at the latter whenever $J(t) = \pm L - 1$ as the case may be.

6. Let $T^I_{R,1\to 2}(T^I_{R,2\to 1})$ be the transfer rate of inventory from queue 1 to queue 2 (queue 2 to queue 1). Unlike the case of customer transfer, inventory transfer can occur in several situations (not necessarily with a customer transfer). For example, if a service station has sufficient stock of inventory whereas the other has no stock, then an arrival to the second station triggers an inventory transfer from the former without a customer transfer. All these things have been taken into consideration for writing the expression for $T^I_{R,1\to 2}$.

$$T^I_{R,1\to 2} = \lambda_1 \sum_{j=1}^{S-1} \sum_{i=j+1}^{S} \pi_{L-1,L-1,i,j} + \lambda_1 \sum_{i=2}^{S} \pi_{L-1,L-1,i,0}$$

$$+ \lambda_1 \sum_{n=0}^{\infty} \sum_{j=1}^{S-1} \sum_{i=j+1}^{S} \pi_{L+n,L-1,i,j}$$

$$+ \mu_2 \sum_{n=0}^{\infty} \sum_{j=1}^{S} \sum_{i=j}^{S} \pi_{L+n,L-1,i,j} + \lambda_2 \sum_{i=1}^{S} \pi_{0,0,i,0}$$

$$
\begin{aligned}
&+ \lambda_2 \sum_{\substack{n_2, n_1 \neq 0 \\ 1 \leq n_1+n_2 \leq L-1 \\ |n_1-n_2|<L}} \sum_{i=2}^{S} \pi_{n_1+n_2, n_1-n_2, i, 0} \\
&+ \mu_2 \sum_{\substack{n_2, n_1 \neq 0 \\ 1 \leq n_1+n_2 \leq L-1 \\ |n_1-n_2|<L}} \sum_{i=2}^{S} \pi_{n_1+n_2, n_1-n_2, i, 1} + \mu_2 \sum_{n=1}^{L-1} \sum_{i=2}^{S} \pi_{n,-n,i,1} \\
&+ \mu_2 \sum_{n=0}^{\infty} \sum_{P=L-2}^{-(L-1)} \sum_{i=2}^{S} \pi_{L+n,p,i,1} \\
&+ \delta_1 \sum_{\substack{n_2 \neq 0, n_1 \neq 0 \\ 1 \leq n_1+n_2 \leq L-1 \\ |n_1-n_2|<L}} \sum_{i=0,1} \pi_{n_1+n_2, n_1-n_2, i, 0} \\
&+ \delta_1 \sum_{n=1}^{L-1} \pi_{n,-n,0,0} + \delta_1 \sum_{n=0}^{\infty} \sum_{P=L-2}^{-(L-1)} \sum_{i=0,1} \pi_{L+n,p,i,0} \\
= &\lambda_1 \sum_{j=1}^{S-1} \sum_{i=j+1}^{S} \pi_{L-1,L-1,i,j} + \lambda_1 \sum_{i=2}^{S} \pi_{L-1,L-1,i,0} \\
&+ \lambda_1 \sum_{i=0}^{S-2} \sum_{j=1}^{i+1} (\pi_L (I-R)^{-1})_{(L-1)(S^2+3)+(i+1)S+j+3} \\
&+ \mu_2 \sum_{i=0}^{S-2} \sum_{j=1}^{i+2} (\pi_L (I-R)^{-1})_{(L-1)(S^2+3)+(i+1)S+j+3} \\
&+ \lambda_2 \sum_{i=0}^{S-1} (\pi_0)_{(i+1)(S+1)+1} + \lambda_2 \sum_{n=1}^{L-1} \sum_{i=1}^{S-1} (\pi_n)_{i(S+1)+2} \\
&+ \mu_2 \sum_{n=1}^{L-1} \sum_{l=1}^{n-2} \sum_{i=0}^{S-2} (\pi_n)_{S(S+1)+l(S^2+3)+(i+1)S+5} \\
&+ \mu_2 \sum_{n=1}^{L-1} \sum_{i=0}^{S-1} (\pi_n)_{S(S+1)+(n-1)(S^2+3)+(i+1)S+3}
\end{aligned}
$$

$$
\begin{aligned}
&+ \mu_2 \sum_{l=0}^{2L-2} \sum_{i=0}^{S-2} (\pi_L (I-R)^{-1})_{l(S^2+3)+(i+1)S+4} \\
&+ \delta_1 \sum_{n=1}^{L-1} \sum_{l=0}^{n-2} \sum_{j=1,3} (\pi_n)_{S(S+1)+l(S^2+3)+j+1} \\
&+ \delta_1 \sum_{n=1}^{L-1} (\pi_n)_{S(S+1)+(n-1)(S^2+3)+2} \\
&+ \delta_1 \sum_{l=0}^{2L-2} \sum_{j=1,3} (\pi_L (I-R)^{-1})_{l(S^2+3)+j} \\
= \lambda_1 &\sum_{j=0}^{S-1} \sum_{i=j+1}^{S} (1-\delta_{i1}) \pi_{L-1,L-1,i,j} \\
&+ \sum_{i=0}^{S-2} \sum_{j=1}^{i+2} (\lambda_1 (1-\delta_{i+2,j}) + \mu_2] U(L-1, i+1, j+3) \\
&+ \lambda_2 \big[\sum_{i=1}^{S} (\pi_0)_{i(S+1)+1} + \sum_{n=1}^{L-1} \sum_{i=1}^{S-1} (\pi_n)_{i(S+1)+2} \big] \\
&+ \mu_2 \big[\sum_{n=1}^{L-1} \sum_{l=1}^{n-2} \sum_{i=0}^{S-2} V(l, i+1, 5) \\
&+ \sum_{n=1}^{L-1} \sum_{i=0}^{S-1} V(n-1, i+1, 3) + \sum_{l=0}^{2L-2} \sum_{i=0}^{S-2} U(l, i+1, 4) \big] \\
&+ \delta_1 \big[\sum_{n=1}^{L-1} \sum_{l=0}^{n-1} \sum_{j=1,3} (1 - \delta_{l,n-1} \times \delta_{3j}) V(l, 0, j+1) \\
&+ \sum_{l=0}^{2L-2} \sum_{j=1,3} U(l, 0, j) \big]
\end{aligned}
$$

and

$$
\begin{aligned}
T^{I}_{R,2\to 1} &= \lambda_2 \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} \pi_{L-1,-(L-1),i,j} + \lambda_2 \sum_{j=2}^{S} \pi_{L-1,-(L-1),0,j} \\
&+ \lambda_2 \sum_{n=0}^{\infty} \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} \pi_{L+n,-(L-1),i,j} \\
&+ \mu_1 \sum_{n=0}^{\infty} \sum_{i=1}^{S} \sum_{j=i}^{S} \pi_{L+n,-(L-1),i,j} \\
&+ \lambda_1 \sum_{j=1}^{S} \pi_{0,0,0,j} + \lambda_1 \sum_{n=1}^{L-1} \sum_{j=2}^{S} \pi_{n,-n,0,j} + \mu_1 \sum_{n=1}^{L-1} \sum_{j=2}^{S} \pi_{n,n,1,j} \\
&+ \mu_1 \sum_{\substack{n_2\neq 0, n_1\neq 0 \\ 1\leq n_1+n_2\leq L-1 \\ |n_1-n_2|<L}} \sum_{j=2}^{S} \pi_{n_1+n_2,n_1-n_2,1,j} \\
&+ \mu_1 \sum_{n=0}^{\infty} \sum_{P=L-2}^{-(L-1)} \sum_{j=0,1} \pi_{L+n,p,1,j} + \delta_2 \sum_{n=1}^{L-1} \pi_{n,n,0,0} \\
&+ \delta_2 \sum_{\substack{n_1\neq 0, n_2\neq 0 \\ 1\leq n_1+n_2\leq L-1 \\ |n_1-n_2|<L}} \sum_{j=0,1} \pi_{n_1+n_2,n_1-n_2,0,j} \\
&+ \delta_2 \sum_{n=0}^{\infty} \sum_{P=L-2}^{-(L-1)} \sum_{j=0,1} \pi_{L+n,p,0,j} \\
&= \lambda_2 \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} \pi_{L-1,-(L-1),i,j} + \lambda_2 \sum_{j=2}^{S} \pi_{L-1,-(L-1),0,j} \\
&+ \lambda_2 \sum_{j=5}^{S^2+3} (\pi_L (I-R)^{-1})_{(2L-2)(S^2+3)+j} \\
&- \lambda_2 \sum_{i=0}^{S-2} \sum_{j=1}^{i+2} (\pi_L (I-R)^{-1})_{(i+1)S+j+3}
\end{aligned}
$$

$$
\begin{aligned}
&+ \mu_1 \sum_{j=5}^{S^2+3} (\pi_L (I-R)^{-1})_{(2L-2)(S^2+3)+j} \\
&- \mu_1 \sum_{i=0}^{S-2} \sum_{j=1}^{i+1} (\pi_L (I-R)^{-1})_{(i+1)S+j+3} \\
&+ \lambda_1 \sum_{j=1}^{S} (\pi_0)_{j+1} + \lambda_1 \sum_{n=1}^{L-1} \sum_{j=3}^{S+1} (\pi_n)_{S(S+1)+(n-1)(S^2+3)+j+1} \\
&+ \mu_1 \sum_{n=1}^{L-1} \sum_{j=3}^{S+1} (\pi_n)_{j+1} + \mu_1 \sum_{n=1}^{L-1} \sum_{l=0}^{n-2} \sum_{j=5}^{S+3} (\pi_n)_{S(S+1)+l(S^2+3)+j+1} \\
&+ \mu_1 \sum_{l=0}^{2L-2} \sum_{j=5}^{S+3} (\pi_L (I-R)^{-1})_{l(S^2+3)+j} + \delta_2 \sum_{n=1}^{L-1} (\pi_n)_1 \\
&+ \delta_2 \sum_{n=1}^{L-1} \sum_{l=0}^{n-2} \sum_{j=1,2} (\pi_n)_{S(S+1)+l(S^2+3)+j+1} \\
&+ \delta_2 \sum_{l=0}^{2L-2} \sum_{j=1,2} (\pi_L (I-R)^{-1})_{l(S^2+3)+j} \\
= \lambda_1 &\Big[\sum_{j=1}^{S} (\pi_0)_{j+1} + \sum_{n=1}^{L-1} \sum_{j=3}^{S+1} V(n-1,0,j+1) \Big] \\
&+ \lambda_2 \Big[\sum_{i=0}^{S-1} \sum_{j=i+1}^{S} (1-\delta_{ij}) \pi_{L-1,-(L-1),i,j} + \sum_{j=5}^{S^2+3} U(2L-2,0,j) \\
&- \sum_{i-0}^{S-2} \sum_{j=1}^{i+2} U(0,i+1,j+3) \Big] \\
&+ \mu_1 \Big[\sum_{j=5}^{S^2+3} U(2L-2,0,j) - \sum_{i=0}^{S-2} \sum_{j=1}^{i+1} U(0,i+1,j+3) \\
&+ \sum_{n=1}^{L-1} \sum_{j=3}^{S+1} (\pi_n)_{j+1} + \sum_{n=1}^{L-1} \sum_{l=0}^{n-2} \sum_{S+3}^{j=5} V(l,0,j+1)
\end{aligned}
$$

$$+\sum_{l=0}^{2L-2}\sum_{j=5}^{S+3}U(l,0,j)\Big]$$

$$+\,\delta_2\Big[\sum_{n=1}^{L-1}(\pi_n)_1+\sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=1,2}V(l,0,j+1)$$

$$+\sum_{l=0}^{2L-2}\sum_{j=1,2}V(l,0,j)\Big]$$

7. Total replenishment order rate, i.e., the rate at which orders for inventory are placed at both stations,

$$\begin{aligned}R_0^I &= \mu_2\sum_{n=1}^{L-1}\pi_n f_n+\mu_2\sum_{n=0}^{\infty}\pi_{L+n}f+\mu_1\sum_{n=1}^{L-1}\pi_n f_n'+\mu_1\sum_{n=0}^{\infty}\pi_{L+n}f'\\ &=\mu_2\Big[\sum_{n=1}^{L-1}\pi_n f_n+\pi_L(I-R)^{-1}f\Big]\\ &\quad+\mu_1\Big[\sum_{n=1}^{L-1}\pi_n f_n'+\pi_L(I-R)^{-1}f'\Big].\end{aligned}$$

Here f_n, f, f_n' and f' are column vectors having dimensions of corresponding π_i's and are defined by

$$(f_n)_{S(S+1)+l(S^2+3)+iS+s+5}=1$$
$$\text{for } l=0,1,2,\dots,n-2;\ i=0,1,\dots,S-1$$

and all other entries are zero;

$$(f)_{l(S^2+3)+iS+s+4}=1\ \text{ for }\ l=0,1,2,\dots,2L-2;\, i=0,1,\dots,S-1$$

and remaining entries are zero.

Similarly, the only non-zero entries occurring in f_n' and f' are given by

$$(f_n')_{s(S+1)+j+1}=1\quad\text{for } j=1,\dots,S;$$

$$(f_n')_{S(S+1)+l(S^2+3)+sS+j+4}=1\text{ for }\ l=0,1,\dots,n-2;\, j=1,2,\dots,S.$$

and

$$(f')_{l(S^2+3)+sS+j+3} = 1 \quad \text{for } l = 0, 1, \ldots, 2L-2;\ j = 1, 2, \ldots, S.$$

Note that a replenishment order will be placed only at a service completion epoch due to the fact that the inventory level will be reduced to s only through a service completion. Another important measure associated with the present model is the mean inventory size included in a transfer from one station to the other. Clearly, size of the inventory varies from transfer to transfer. All the possible cases are taken for the computation of such a measure. Only the final expression will be exhibited here.

8. Average size of inventory included in a transfer from queue 1 to queue 2,

$$\begin{aligned}
EI_{1\to 2} = &\sum_{j=0}^{s-1}\sum_{i=j+1}^{s} \pi_{L-1,L-1,i,j} \min\{i-j-\delta_{j0}, K'\}\frac{\lambda_1}{\Delta-\mu_2} \\
&+ \sum_{j=0}^{s}\sum_{i=s+1}^{S} \pi_{L-1,L-1,i,j} \min\{i-j-\delta_{j0}, K'\}\frac{\lambda_1}{\Delta-\mu_2-\delta_1} \\
&+ \sum_{j=s+1}^{S-1}\sum_{i=j+1}^{S} \pi_{L-1,L-1,i,j} \min\{i-j, K'\}\frac{\lambda_1}{\Delta-\mu_2-\delta_1-\delta_2} \\
&+ \sum_{n=0}^{\infty}\sum_{j=0}^{s-1}\sum_{i=j+1}^{s} \pi_{L+n,L-1,i,j}\frac{\lambda_1}{\Delta} \min\{i-j-\delta_{j0}, K'\} \\
&+ \sum_{n=0}^{\infty}\sum_{j=0}^{s}\sum_{i=s+1}^{S} \pi_{L+n,L-1,i,j}\frac{\lambda_1}{\Delta-\delta_1} \min\{i-j-\delta_{j0}, K'\} \\
&+ \sum_{n=0}^{\infty}\sum_{j=s+1}^{S-1}\sum_{i=j+1}^{S} \pi_{L+n,L-1,i,j}\frac{\lambda_1}{\Delta-\delta_1-\delta_2} \min\{i-j, K'\} \\
&+ \sum_{n=0}^{\infty}\sum_{j=1}^{s-1}\sum_{i=j}^{s} \pi_{L+n,L-1,i,j}\frac{\mu_2}{\Delta} \min\{i-j+1, K'\}
\end{aligned}$$

$$+\sum_{n=0}^{\infty}\sum_{j=1}^{s}\sum_{i=s+1}^{S}\pi_{L+n,L-1,i,j}\frac{\mu_2}{\Delta-\delta_1}\min\{i-j+1,K'\}$$

$$+\sum_{n=0}^{\infty}\pi_{L+n,L-1,s,s}\frac{\mu_2}{\Delta}$$

$$+\sum_{n=0}^{\infty}\sum_{j=s+1}^{S}\sum_{i=j}^{S}\pi_{L+n,L-1,i,j}\frac{\mu_2}{\Delta-\delta_1-\delta_2}\min\{i-j+1,K'\}$$

$$+\sum_{i=1}^{s}\pi_{0,0,i,0}\frac{\lambda_2}{\Delta-\mu_1-\mu_2}+\sum_{i=s+1}^{S}\pi_{0,0,i,0}\frac{\lambda_2}{\Delta-\delta_1-\mu_1-\mu_2}$$

$$+\sum_{\substack{n_1,n_2\\1\leq n_1+n_2\leq L-1\\|n_1-n_2|<L}}\sum_{i=2}^{s}\pi_{n_1+n_2,n_1-n_2,i,0}\frac{\lambda_2}{\Delta-\mu_2}$$

$$+\sum_{\substack{n_1,n_2\\1\leq n_1+n_2\leq L-1\\|n_1-n_2|<L}}\sum_{i=s+1}^{S}\pi_{n_1+n_2,n_1-n_2,i,0}\frac{\lambda_2}{\Delta-\delta_1-\mu_2}$$

$$+\sum_{\substack{n_2,n_1\neq 0\\1\leq n_1+n_2\leq L-1\\|n_1-n_2|<L}}\sum_{i=2}^{s}\pi_{n_1+n_2,n_1-n_2,i,1}\frac{\mu_2}{\Delta}$$

$$+\sum_{\substack{n_2,n_1\neq 0\\1\leq n_1+n_2\leq L-1\\|n_1-n_2|<L}}\sum_{i=s+1}^{S}\pi_{n_1+n_2,n_1-n_2,i,1}\frac{\mu_2}{\Delta-\delta_1}$$

$$+\sum_{n=1}^{L-1}\sum_{i=2}^{s}\pi_{n,-n,i,1}\frac{\mu_2}{\Delta-\mu_1}+\sum_{n=1}^{L-1}\sum_{i=s+1}^{S}\pi_{n,-n,i,1}\frac{\mu_2}{\Delta-\mu_1-\delta_1}$$

$$+\sum_{n=0}^{\infty}\sum_{p=L-2}^{-(L-1)}\sum_{i=2}^{s}\pi_{L+n,p,i,1}\frac{\mu_2}{\Delta}$$

$$+\sum_{n=0}^{\infty}\sum_{p=L-2}^{-(L-1)}\sum_{i=s+1}^{S}\pi_{L+n,p,i,1}\frac{\mu_2}{\Delta-\delta_1}$$

$$+\sum_{\substack{n_2\neq 0,n_1\neq 0\\ 1\leq n_1+n_2\leq L-1\\ |n_1-n_2|<L}}\sum_{i=0,1}\pi_{n_1+n_2,n_1-n_2,i,0}\frac{\delta_1}{\Delta-\mu_2}$$

$$+\sum_{n=1}^{L-1}\pi_{n,-n,0,0}\frac{\delta_1}{\Delta-\mu_1-\mu_2}$$

$$+\sum_{n=0}^{\infty}\sum_{p=L-2}^{-(L-1)}\sum_{i=0,1}\pi_{L+n,p,i,0}\frac{\delta_1}{\Delta-\mu_2}$$

$$=\sum_{j=0}^{s-1}\sum_{i=j+1}^{s}\pi_{L-1,L-1,i,j}\frac{\lambda_1}{\Delta-\mu_2}\min\{i-j-\delta_{j0},K'\}$$

$$+\sum_{j=0}^{s}\sum_{i=s+1}^{S}\pi_{L-1,L-1,i,j}\frac{\lambda_1}{\Delta-\mu_2-\delta_1}\min\{i-j-\delta_{j0},K'\}$$

$$+\sum_{j=s+1}^{S-1}\sum_{i=j+1}^{S}\pi_{L-1,L-1,i,j}\frac{\lambda_1}{\Delta-\mu_2-\delta_1-\delta_2}\min\{i-j,K'\}$$

$$+\sum_{j=1}^{s-1}\sum_{i=j+1}^{s}(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3}$$

$$\frac{\lambda_1}{\Delta}\min\{i-j-\delta_{j0},K'\}$$

$$+\sum_{j=1}^{s}\sum_{i=s+1}^{S}(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3}\frac{\lambda_1}{\Delta-\delta_1}$$

$$\min\{i-j-\delta_{j0},K'\}$$

$$+\sum_{j=s+1}^{S-1}\sum_{i=j+1}^{S}(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3}\frac{\lambda_1}{\Delta-\delta_1-\delta_2}$$

$$\min\{i-j,K'\}$$

$$+\sum_{j=1}^{s-1}\sum_{i=j}^{S}(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3}$$
$$\frac{\mu_2}{\Delta}\min\{i-j+1,K'\}$$
$$+\sum_{j=1}^{s}\sum_{i=s+1}^{S}(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3}\frac{\mu_2}{\Delta-\delta_1}$$
$$\min\{i-j+1,K'\}$$
$$+\sum_{j=s+1}^{S}\sum_{i=j}^{S}(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3}\frac{\mu_2}{\Delta-\delta_1-\delta_2}$$
$$\min\{i-j+1,K'\}+(\pi_L(I-R)^{-1})_{(L-1)(S^2+3)+(s-1)S+s+3}\frac{\mu_2}{\Delta}$$
$$+\sum_{i=1}^{s}(\pi_0)_{i(S+1)+1}\frac{\lambda_2}{\Delta-\mu_1-\mu_2}$$
$$+\sum_{i=s+1}^{S}(\pi_0)_{i(S+1)+1}\frac{\lambda_2}{\Delta-\delta_1-\mu_1-\mu_2}$$
$$+\sum_{n=1}^{L-1}\sum_{i=2}^{s}(\pi_n)_{(i-1)(S+1)+2}\frac{\lambda_2}{\Delta-\mu_2}$$
$$+\sum_{n=1}^{L-1}\sum_{i=s+1}^{S}(\pi_n)_{(i-1)(S+1)+2}\frac{\lambda_2}{\Delta-\delta_1-\mu_2}$$
$$+\sum_{n=1}^{L-1}\sum_{l=1}^{n-2}\sum_{i=2}^{s}(\pi_n)_{S(S+1)+l(S^2+3)+(i-1)S+5}\frac{\mu_2}{\Delta}$$
$$+\sum_{n=1}^{L-1}\sum_{l=1}^{n-2}\sum_{i=s+1}^{S}(\pi_n)_{S(S+1)+l(S^2+3)+(i-1)S+5}\frac{\mu_2}{\Delta-\delta_1}$$
$$+\sum_{n=1}^{L-1}\sum_{i=2}^{s}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+(i-1)S+3}\frac{\mu_2}{\Delta-\mu_1}$$

$$+\sum_{n=1}^{L-1}\sum_{i=s+1}^{S}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+(i-1)S+3}\frac{\mu_2}{\Delta-\mu_1-\delta_1}$$

$$+\sum_{l=0}^{2L-2}\sum_{i=2}^{s}(\pi_L(I-R)^{-1})_{l(S^2+3)+(i-1)S+4}\frac{\mu_2}{\Delta}$$

$$+\sum_{l=0}^{2L-2}\sum_{i=s+1}^{S}(\pi_L(I-R)^{-1})_{l(S^2+3)+(i-1)S+4}\frac{\mu_2}{\Delta-\delta_1}$$

$$+\sum_{n=1}^{L-1}\sum_{l=1}^{n-2}\sum_{j=1,3}(\pi_n)_{S(S+1)+l(S^2+3)+j+1}\frac{\delta_1}{\Delta-\mu_2}$$

$$+\sum_{n=1}^{L-1}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+2}\frac{\delta_1}{\Delta-\mu_1-\mu_2}$$

$$+\sum_{l=0}^{2L-2}\sum_{j=1,3}(\pi_L(I-R)^{-1})_{l(S^2+3)+j}\frac{\delta_1}{\Delta-\mu_2}.$$

$$EI_{1\to 2}=\sum_{j=0}^{s-1}\sum_{i=j+1}^{s}\pi_{L-1,L-1,i,j}\frac{\lambda_1}{\Delta-\mu_2}\min\{i-j-\delta_{j0},K'\}$$

$$+\sum_{j=0}^{s}\sum_{i=s+1}^{S}\pi_{L-1,L-1,i,j}\frac{\lambda_1}{\Delta-\mu_2-\delta_1}\min\{i-j-\delta_{j0},K'\}$$

$$+\sum_{j=s+1}^{S-1}\sum_{i=j+1}^{S}\pi_{L-1,L-1,i,j}\frac{\lambda_1}{\Delta-\mu_2-\delta_1-\delta_2}\min\{i-j,K'\}$$

$$+\sum_{j=1}^{s-1}\sum_{i=j+1}^{s}U(L-1,i-1,j+3)\frac{\lambda_1}{\Delta}\min\{i-j,K'\}$$

$$+\sum_{j=1}^{s}\sum_{i=s+1}^{S}U(L-1,i-1,j+3)$$

$$\frac{[\lambda_1\min\{i-j,K'\}+\mu_2\min\{i-j+1,K'\}]}{\Delta-\delta_1}$$

$$
\begin{aligned}
&+ \sum_{j=s+1}^{S-1} \sum_{i=j+1}^{S} U(L-1, i-1, j+3) \frac{\lambda_1}{\Delta - \delta_1 - \delta_2} \min\{i-j, K'\} \\
&+ \sum_{j=1}^{s-1} \sum_{i=j}^{S} U(L-1, i-1, j+3) \frac{\mu_2}{\Delta} \min\{i-j+1, K'\} \\
&+ \sum_{j=s+1}^{S} \sum_{i=j}^{S} U(L-1, i-1, j+3) \frac{\mu_2}{\Delta - \delta_1 - \delta_2} \\
&\quad \min\{i-j+1, K'\} + U(L-1, s-1, s+3) \frac{\mu_2}{\Delta} \\
&+ \sum_{i=1}^{S} (\pi_0)_{i(S+1)+1} \frac{\lambda_2}{\Delta - \mu_1 - \mu_2 - \delta_1 \prod_{j=1}^{s}(1-\delta_{ij})} \\
&+ \sum_{n=1}^{L-1} \sum_{i=2}^{S} (\pi_n)_{(i-1)(S+1)+2} \frac{\lambda_2}{\Delta - \mu_2 - \delta_1 \prod_{j=2}^{s}(1-\delta_{ij})} \\
&+ \sum_{n=1}^{L-1} \sum_{l=1}^{n-2} \sum_{i=2}^{S} V(l, i-1, 5) \frac{\mu_2}{\Delta - \delta_1 \prod_{j=2}^{s}(1-\delta_{ij})} \\
&+ \sum_{n=1}^{L-1} \sum_{i=2}^{S} V(n-1, i-1, 3) \frac{\mu_2}{\Delta - \mu_1 - \delta_1 \prod_{j=2}^{s}(1-\delta_{ij})} \\
&+ \sum_{l=0}^{2L-2} \sum_{i=2}^{S} U(l, i-1, 4) \frac{\mu_2}{\Delta - \delta_1 \prod_{j=2}^{s}(1-\delta_{ij})} \\
&+ \sum_{n=1}^{L-1} \sum_{l=1}^{n-2} \sum_{j=1,3} V(l, 0, j+1) \frac{\delta_1}{\Delta - \mu_2} \\
&+ \sum_{n=1}^{L-1} V(n-1, 0, 2) \frac{\delta_1}{\Delta - \mu_1 - \mu_2} + \sum_{l=0}^{2L-2} \sum_{j=1,3} U(l, 0, j) \frac{\delta_1}{\Delta - \mu_2}.
\end{aligned}
$$

where $\Delta = \lambda_1 + \lambda_2 + \mu_1 + \mu_2 + \delta_1 + \delta_2$.

Average size of inventory included in a transfer from queue 2 to queue 1,

$$
\begin{aligned}
EI_{2\to 1} = & \sum_{i=0}^{s-1}\sum_{j=i+1}^{s} \pi_{L-1,-(L-1),i,j}\frac{\lambda_2}{\Delta-\mu_1}\min\{j-i-\delta_{i0},K'\} \\
& + \sum_{i=0}^{s}\sum_{j=s+1}^{S} \pi_{L-1,-(L-1),i,j}\frac{\lambda_2}{\Delta-\mu_1-\delta_2}\min\{j-i-\delta_{i0},K'\} \\
& + \sum_{i=s+1}^{S-1}\sum_{j=i+1}^{S} \pi_{L-1,-(L-1),i,j}\frac{\lambda_2}{\Delta-\mu_1-\delta_1-\delta_2}\min\{j-i,K'\} \\
& + \sum_{i=1}^{s-1}\sum_{j=i+1}^{s} (\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
& \frac{\lambda_2}{\Delta}\min\{j-i-\delta_{i0},K'\} \\
& + \sum_{i=0}^{s}\sum_{j=s+1}^{S} (\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
& \frac{\lambda_2}{\Delta-\delta_2}\min\{j-i-\delta_{i0},K'\} \\
& + \sum_{i=s+1}^{S-1}\sum_{j=i+1}^{S} (\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
& \frac{\lambda_2}{\Delta-\delta_1-\delta_2}\min\{j-i,K'\} \\
& + \sum_{i=1}^{s-1}\sum_{j=i}^{s}(\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
& \frac{\mu_1}{\Delta}\min\{j-i+1,K'\} \\
& + \sum_{i=1}^{s}\sum_{j=s+1}^{S} (\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
& \frac{\mu_1}{\Delta-\delta_2}\min\{j-i+1,K'\}
\end{aligned}
$$

$$+\sum_{i=s+1}^{S}\sum_{j=i}^{S}(\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3}$$
$$\frac{\mu_1}{\Delta-\delta_1-\delta_2}\min\{j-i+1,K'\}$$
$$+(\pi_L(I-R)^{-1})_{(2L-2)(S^2+3)+(s-1)S+s+3}\frac{\mu_1}{\Delta}$$
$$+\sum_{j=1}^{s}(\pi_0)_{j+1}\frac{\lambda_1}{\Delta-\mu_1-\mu_2}+\sum_{j=s+1}^{S}(\pi_0)_{j+1}\frac{\lambda_1}{\Delta-\delta_2-\mu_1-\mu_2}$$
$$+\sum_{n=1}^{L-1}\sum_{j=2}^{s}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+j+2}\frac{\lambda_1}{\Delta-\mu_1}$$
$$+\sum_{n=1}^{L-1}\sum_{j=s+1}^{S}(\pi_n)_{S(S+1)+(n-1)(S^2+3)+j+2}\frac{\lambda_1}{\Delta-\delta_2-\mu_1}$$
$$+\sum_{n=1}^{L-1}\sum_{j=2}^{s}(\pi_n)_{j+1}\frac{\mu_1}{\Delta-\mu_2}+\sum_{n=1}^{L-1}\sum_{j=s+1}^{S}(\pi_n)_{j+1}\frac{\mu_1}{\Delta-\delta_2-\mu_2}$$
$$+\sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=2}^{s}(\pi_n)_{S(S+1)+l(S^2+3)+j+4}\frac{\mu_1}{\Delta}$$
$$+\sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{i=s+1}^{S}(\pi_n)_{S(S+1)+l(S^2+3)+j+4}\frac{\mu_1}{\Delta-\delta_2}$$
$$+\sum_{l=0}^{2L-2}\sum_{j=2}^{s}(\pi_L(I-R)^{-1})_{l(S^2+3)+j+3}\frac{\mu_1}{\Delta}$$
$$+\sum_{l=0}^{2L-2}\sum_{j=s+1}^{S}(\pi_L(I-R)^{-1})_{l(S^2+3)+j+3}\frac{\mu_1}{\Delta-\delta_2}$$
$$+\sum_{n=1}^{L-1}(\pi_n)_1\frac{\delta_2}{\Delta-\mu_1-\mu_2}$$
$$+\sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=1,2}(\pi_n)_{S(S+1)+l(S^2+3)+j+1}\frac{\delta_2}{\Delta-\mu_1}$$

$$
\begin{aligned}
&+\sum_{l=0}^{2L-2}\sum_{j=1,2}(\pi_L(I-R)^{-1})_{l(S^2+3)+j}\frac{\delta_2}{\Delta-\mu_1}\\
=&\sum_{i=0}^{s-1}\sum_{j=i+1}^{s}\pi_{L-1,-(L-1),i,j}\frac{\lambda_2}{\Delta-\mu_1}\min\{j-i-\delta_{i0},K'\}\\
&+\sum_{i=0}^{s}\sum_{j=s+1}^{S}\pi_{L-1,-(L-1),i,j}\frac{\lambda_2}{\Delta-\mu_1-\delta_2}\min\{j-i-\delta_{i0},K'\}\\
&+\sum_{i=s+1}^{S-1}\sum_{j=i+1}^{S}\pi_{L-1,-(L-1),i,j}\frac{\lambda_2}{\Delta-\mu_1-\delta_1-\delta_2}\min\{j-i,K'\}\\
&+\sum_{i=1}^{s-1}\sum_{j=i+1}^{s}U(2L-2,i-1,j+3)\frac{\lambda_2}{\Delta}\min\{j-i,K'\}\\
&+\sum_{i=0}^{s}\sum_{j=s+1}^{S}U(2L-2,i-1,j+3)\frac{\lambda_2}{\Delta-\delta_2}\min\{j-i-\delta_{i0},K'\}\\
&+\sum_{i=s+1}^{S-1}\sum_{j=i+1}^{S}U(2L-2,i-1,j+3)\frac{\lambda_2}{\Delta-\delta_1-\delta_2}\min\{j-i,K'\}\\
&+\sum_{i=1}^{s-1}\sum_{j=1}^{s}U(2L-2,i-1,j+3)\frac{\mu_1}{\Delta}\min\{j-i+1,K'\}\\
&+\sum_{i=1}^{s}\sum_{j=s+1}^{S}U(2L-2,i-1,j+3)\frac{\mu_1}{\Delta-\delta_2}\min\{j-i+1,K'\}\\
&+\sum_{i=s+1}^{S}\sum_{j=i}^{S}U(2L-2,i-1,j+3)\frac{\mu_1}{\Delta-\delta_1-\delta_2}\min\{j-i+1,K'\}\\
&+U(2L-2,s-1,s+3)\frac{\mu_1}{\Delta}\\
&+\sum_{j=1}^{s}(\pi_0)_{j+1}\frac{\lambda_1}{\Delta-\mu_1-\mu_2-\delta_2\prod_{i=1}^{s}(1-\delta_{ij})}
\end{aligned}
$$

$$+\sum_{n=1}^{L-1}\sum_{j=2}^{s}V(n-1,0,j+2)\frac{\lambda_1}{\Delta-\mu_1-\delta_2\prod_{i=2}^{s}(1-\delta_{ij})}$$

$$+\sum_{n=1}^{L-1}\sum_{j=2}^{s}(\pi_n)_{j+1}\frac{\mu_1}{\Delta-\mu_2-\delta_2\prod_{i=2}^{s}(1-\delta_{ij})}$$

$$+\sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=2}^{S}V(l,0,j+4)\frac{\mu_1}{\Delta-\delta_2-\prod_{i=2}^{s}(1-\delta_{ij})}$$

$$+\sum_{l=0}^{2L-2}\sum_{j=2}^{S}U(l,0,j+3)\frac{\mu_1}{\Delta-\delta_2-\prod_{i=2}^{s}(1-\delta_{ij})}$$

$$+\sum_{n=1}^{L-1}(\pi_n)_1\frac{\delta_2}{\Delta-\mu_1-\mu_2}+\sum_{n=1}^{L-1}\sum_{l=0}^{n-2}\sum_{j=1,2}V(l,0,j+1)\frac{\delta_2}{\Delta-\mu_1}$$

$$+\sum_{l=0}^{2L-2}\sum_{j=1,2}U(l,0,j)\frac{\delta_2}{\Delta-\mu_1}$$

9. Let r_i be the rate at which jobs are completed by server i per unit time, $i=1,2$, then

$$r_1=\mu_1(1-P_1^{idle})\quad\text{and}\quad r_2=\mu_2(1-P_2^{idle}).$$

10. Since there are exactly K customers are included in each transfer, the total number of customers transferred per unit time is $T_R^C K$ where $T_R^C = T_{R,1\to 2}^C + T_{R,2\to 1}^C$. Since the total arrival rate of customers is $\lambda_1+\lambda_2$, the mean number of transfers per customer (during its sojourn time in the system), $N_{TR}^C = \frac{T_R^C K}{\lambda_1+\lambda_2}$. Replenishment realization rate, i.e., the rate at which ordered inventory reaches the stations,

$$R_R^I=\delta_1\big[\sum_{n=0}^{L-1}\pi_n f_n''+\sum_{n=0}^{\infty}\pi_{L+n}f''\big]+\delta_2\big[\sum_{n=0}^{L-1}\pi_n f_n'''+\sum_{n=0}^{\infty}\pi_{L+n}f'''\big]$$

where f_n'', f'', f_n''' and f''' are column vectors of appropriate size and the non-zero entries occurring in these vectors are given by

$$(f_0'')_i=1\quad\text{for } i=0,1,2,\ldots,(s+1)(S+1);$$

for $1 \le n \le L-1$, $(f''_n)_i = 1$ for $i = 1, 2, \ldots s(S+1)+1$;

$$(f''_n)_{S(S+1)+l(S^2+3)+i+1} = 1 \qquad \text{for } l = 0, 1, 2, \ldots, n-2;\ i = 1, 2, \ldots, sS+3$$

and

$$(f''_n)_{S(S+1)+(n-1)(S^2+3)+i+1} = 1 \quad \text{for } i = 1, \ldots, (s+1)S+1.$$
$$(f'')_{l(S^2+3)+i} = 1 \quad \text{for } l = 0, 1, \ldots, 2L-2; i = 1, 2 \ldots, sS+3;$$
$$(f'''_0)_{i(s+1)+j} = 1 \quad \text{for } i = 0, 1, 2, \ldots, S; j = 1, 2, \ldots, s+1.$$

For $1 \le n \le L-1$,
$(f'''_n)_1 = 1$
$(f'''_n)_{i(S+1)+j+1} = 1$ for $i = 0, 1, \ldots S-1$; $j = 1, 2, \ldots s+1$
$(f'''_n)_{S(S+1)+l((S^2+3)+i+1} = 1$ for $l = 0, 1, 2, \ldots n-2$; $i = 1, 2, \ldots s+3$
$(f'''_n)_{S(S+1)+l((S^2+3)+iS+j+5} = 1$ for $l = 0, 1, 2, \ldots n-2$; $i = 1, 2, \ldots S-2$; $j = 0, 1, \ldots s-1$
$(f'''_n)_{S(S+1)+(n-1)(S^2+3)+iS+j+2} = 1$ for $i = 1, 2, \ldots S$; $j = 1, 2 \ldots s$ and for $i = 0$, $j = 0, 1, 2, \ldots s$
$(f''')_{l(S^2+3)+i} = 1$ for $l = 0, 1, 2, \ldots 2L-2$; $i = 1, 2, \ldots s+3$ and
$(f''')_{l(S^2+3)+iS+j+4} = 1$ for $l = 0, 1, 2, \ldots 2L-2$; $i = 1, 2, \ldots S-2$; $j = 0, 1, 2, \ldots s-1$.

11. Since, on average there are $EI_{1\to2}$ and $EI_{2\to1}$ units of inventory that are included in a transfer from queue 1 to queue 2 and queue 2 to queue 1 respectively, the total units of inventory transferred per unit time from queue 1 to queue 2 is $T^I_{R,1\to2} \times EI_{1\to2}$ and from queue 2 to queue 1 is $T^I_{R,2\to1} \times EI_{2\to1}$. Hence the total units of inventory transferred per unit time is $ET^I_R = T^I_{R,1\to2} \times EI_{1\to2} + T^I_{R,2\to1} \times EI_{2\to1}$. Now the replenishment realization rate of inventory is R^I_R and since there are exactly $S-s$ units that are included in each replenishment, the total units of inventory reached in the system per unit time is $(S-s)R^I_R$. Hence the mean number of transfers per unit inventory is $N^I_{TR} = \frac{ET^I_R}{(S-s)R^I_R}$.

12. Even though an inventory order is placed when the inventory level drops to s, the inventory level may reach above s due to a transfer from the other station. This results in the cancellation of the order that has already been placed. On the average, how many such orders will be cancelled per unit time is a significant measure. Let $R_C^{(i)}$ be the rate at which replenishment orders are cancelled at station i $(i = 1, 2)$.
Then

$$
\begin{aligned}
R_C^{(1)} &= \lambda_2 \sum_{i=1}^{s} \sum_{j=s+1}^{S} \pi_{L-1,-(L-1),i,j} \\
&\quad + \lambda_2 \sum_{j=s+2}^{S} \pi_{L-1,-(L-1),0,j} + \lambda_2 \sum_{n=0}^{\infty} \sum_{i=1}^{s} \sum_{j=s+1}^{S} \pi_{L+n,-(L-1),i,j} \\
&\quad + \mu_1 \sum_{n=0}^{\infty} \sum_{i=2}^{s} \sum_{j=s+1}^{S} \pi_{L+n,-(L-1),i,j} + \mu_1 \sum_{n=0}^{\infty} \sum_{j=s+2}^{S} \pi_{L+n,-(L-1),1,j} \\
&= \lambda_2 \sum_{i=1}^{s} \sum_{j=s+1}^{S} \pi_{L-1,-(L-1),i,j} \\
&\quad + \lambda_2 \sum_{j=s+2}^{S} \pi_{L-1,-(L-1),0,j} \\
&\quad + \lambda_2 \sum_{i=1}^{s} \sum_{j=s+1}^{S} (\pi_L (I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
&\quad + \mu_1 \sum_{i=2}^{s} \sum_{j=s+1}^{S} (\pi_L (I-R)^{-1})_{(2L-2)(S^2+3)+(i-1)S+j+3} \\
&\quad + \mu_1 \sum_{j=s+2}^{S} (\pi_L (I-R)^{-1})_{(2L-2)(S^2+3)+j+3}.
\end{aligned}
$$

$$R_C^{(1)} = \lambda_2 \sum_{i=0}^{s} \sum_{j=s+\delta_{i0}+1}^{S} (\pi_{L-1,-(L-1),i,j}$$

$$+ \sum_{i=1}^{s} \sum_{j=s+1}^{S} (\lambda_2 + \delta_{i1}\mu_1) U(2L-2, i-1, j+3)$$

Similarly,

$$\begin{aligned}
R_C^{(2)} = {} & \lambda_1 \sum_{i=s+1}^{S} \sum_{j=1}^{s} \pi_{L-1,L-1,i,j} \\
& + \lambda_1 \sum_{i=s+2}^{S} \pi_{L-1,L-1,i,0} \\
& + \lambda_1 \sum_{i=s+1}^{S} \sum_{j=1}^{s} (\pi_L (I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3} \\
& + \mu_2 \sum_{i=s+1}^{S} \sum_{j=2}^{s} (\pi_L (I-R)^{-1})_{(L-1)(S^2+3)+(i-1)S+j+3} \\
& + \mu_2 \sum_{i=s+2}^{S} (\pi_L (I-R)^{-1})_{(L-1)(S^2+3)+(s+1)S+4} \\
= {} & \lambda_1 \sum_{j=0}^{s} \sum_{i=s+\delta_{0j}+1}^{S} (\pi_{(L-1),-(L-1),i,j} \\
& + \sum_{j=1}^{s} \sum_{i=s+1}^{S} (\lambda_1 + \delta_{2j}\mu_2) U(L-1, i-1, j+3) \\
& + \mu_2 \sum_{i=s+2}^{S} U(L-1, s+1, 4)
\end{aligned}$$

13. Probability that there are exactly n customer (batch) transfers from queue 1 to queue 2 before the first transfer from queue 2 to queue $1 = \left(\frac{T^C_{R,1\to 2}}{T^C_R}\right)^n \frac{T^C_{R,2\to 1}}{T^C_R}$; $n \geq 0$.
Hence, mean number of customer (batch) transfers from queue 1 to queue 2 before the first transfer from queue 2 to queue 1

$$= \frac{T^C_{R,1\to 2} \times T^C_{R,2\to 1}}{(T^C_R)^2} \times \frac{1}{\left(1 - \frac{T^C_{R,1\to 2}}{T^C_R}\right)^2} = \frac{T^C_{R,1\to 2} \times T^C_{R,2\to 1}}{(T^C_{R,2\to 1})^2}.$$

Similarly, mean number of customer (batch) transfers from queue 2 to queue 1 before the first transfer from queue 1 to queue $2 = \frac{T^C_{R,1\to 2} \times T^C_{R,2\to 1}}{(T^C_{R,1\to 2})^2}$.
Similar is the case for inventory batch transfer.

14. Time between two successive customer transfers conditioned on the fact that the first transfer is from queue 1 to queue 2, follows a PH distribution with parameters (β_1, T) where β_1 is a $(2L-1)$ dimensional row vector with $(2L-2K)^{\text{th}}$ entry 1 and all other entries are zero and

$$T = \begin{bmatrix} -(\lambda_1+\lambda_2+\mu_1+\mu_2) & \lambda_1+\mu_2 & \cdots \\ \lambda_2+\mu_1 & -(\lambda_1+\lambda_2+\mu_1+\mu_2) & \lambda_1+\mu_2 \\ \ddots & \ddots & \ddots \\ & \lambda_2+\mu_1 & -(\lambda_1+\lambda_2+\mu_1+\mu_2) \end{bmatrix}_{2L-1\times 2L-1}.$$

Similarly time between two successive transfers conditioned on the fact that the first transfer is from queue 2 to queue 1 follows a PH distribution with parameters (β_2, T) where β_2 is a $(2L-1)$ dimensional row vector with $2K^{\text{th}}$ entry 1 and all other entries zero.

Hence the time between successive transfers $= -\beta_i T^{-1} e$; $i = 1, 2$ as the case may be.

As a particular case, if $L = 2$, $K = 1$ then the mean time between two successive (customer) transfers

$$= \frac{2(\lambda_1+\lambda_2+\mu_1+\mu_2)}{(\lambda_1+\mu_2)^2 + (\lambda_2+\mu_1)^2}.$$

4.2.3 Discussion on numerical results

Using the results obtained in previous section, in this section, we study some design issues related to the present model and make an attempt to establish that it is very often much superior to the one with two independent (without transfer of customers or inventory) service stations. Below we discuss in brief the outcome of our numerical study.

1. We note from table 4.1 that with replenishment rates δ_1 and δ_2 increasing the condition $\lambda_1 + \lambda_2 < \mu_1 + \mu_2$ becomes more and more conspicuous. To be more specific the stability condition $\frac{\tilde{\pi} A_0 e}{\tilde{\pi} A_2 e} < 1$ approaches the condition $\rho = \frac{\lambda_1+\lambda_2}{\mu_1+\mu_2} < 1$, for larger δ_1 and δ_2 which is in agreement with the result in [23] in that their model can be considered as a system with infinite resource (inventory) with service; or equivalently a service system with infinite replenishment rate. In this case a customer waiting for service due to the absence of inventory does not arise.
2. Table 4.2 provides performance measures of the system for varying pairs (δ_1, δ_2). We notice that the probability of being idle of the servers decrease as the values of (δ_1, δ_2) increase; this also holds for the EN values. Also the transfer rate $T^I_{R,1\to 2}$ decreases with increase in value of δ_1. Note, however, that $T^I_{R,2\to 1}$ increases with δ_2 increasing. This is so because $\frac{\lambda_1}{\mu_1} > 1$.
3. Table 4.3 compares the system that we have described with one where there are two independent (with no customer/inventory transfer) service stations. On comparing values in table 4.3, we see that the combined system (present system) is far more effective than the latter. For the parameters given, the model we have described has probability of being empty given by $\pi_0 e = 0.11419$, whereas the independent service system has probability of being empty given by $P^{ID}_{\text{empty}} = 0.04409$. Similarly for our system $EN = 4.18205$ while the independent service system has mean system size $EN^{ID} = 12.30469$, which implies that

both customer and inventory transfers greatly increase the efficiency of the system, which is intuitive. Hence the interactive system we have described and modelled here performs better than the independent service system. Also, the parameters that we have chosen for the preparation of table 1 and table 2 indicate that the present system can be stable even when one among the individual systems is unstable.

Table 4.1: (δ_1, δ_2) verses "traffic intensity". ($\lambda_1 = 1.5$, $\lambda_2 = 1.1$, $\mu_1 = 1.15$, $\mu_2 = 3.1$, $S = 10$, $s = 4$, $L = 7$, $K = 3$, $K' = 3$, $\rho = 0.61176$)

(δ_1, δ_2)	(1,1)	(10,10)	(50,50)	(75,75)
$\frac{\tilde{\pi} A_0 e}{\tilde{\pi} A_2 e}$	0.62203	0.61176	0.61176	0.61176

Table 4.2: (δ_1, δ_2) verses system performance. ($\lambda_1 = 1.5$, $\lambda_2 = 1.1$, $\mu_1 = 1.15$, $\mu_2 = 3.1$, $S = 10$, $s = 4$, $L = 7$, $K = 3$, $K' = 3$)

(δ_1, δ_2)	P_1^{idle}	P_2^{idle}	$T^I_{R,1\to 2}$	$T^I_{R,2\to 1}$	$EI_{1\to 2}$	$EI_{2\to 1}$	EN
$(1, 1)$	0.08402	0.52350	1.06295	0.30899	0.40620	0.10706	11.90100
(10,10)	0.07207	0.50553	0.80340	0.34682	0.24312	0.12187	11.66534
(50,50)	0.07206	0.50552	0.75888	0.35910	0.23514	0.12914	11.66496
(75,75)	0.07206	0.50552	0.75476	0.36021	0.23504	0.12987	11.66496

Table 4.3: Comparison of the model with a system having two independent service stations. ($\lambda_1 = 1.5$, $\lambda_2 = 1.0$, $\mu_1 = 3.5$, $\mu_2 = 1.1$, $S = 10$, $s = 4$, $L = 7$, $K = 3$, $K' = 3$, $\delta_1 = \delta_2 = 1.0$)

$\pi_0 e$	EN	P^{ID}_{empty}	EN^{ID}
0.11419	4.18205	0.04409	12.30469

Chapter 5

$M/G/1$ Type Inventory Systems with/without Positive Lead Time

5.1 Introduction

In the previous chapters we considered inventory systems with exponentially distributed service times. In this chapter we consider inventory models where the distribution of the service time does not possess the memoryless property, that is we consider non-exponential distributions. Here we assume service times as independently and identically distributed with distribution function $G(\cdot)$. Arrival of demands form a Poission process. Arrival and service processes are assumed to be independent. Our modelling leads to an $M/G/1$ type service inventory system.

Only recently investigations in integrated models appeared concerning the problem of how the classical performance measures (eg. queue length, waiting time, etc.) are influenced by the management of attached inventory and vice versa. How inventory management has to react to queueing of demands and customers which is due to incorporated service facilities. An early contribution is [56] where approximation procedures are used to find performance descriptions for models in which the interaction of queueing for service and inventory control is integrated. In a sequence of papers Berman and his co-workers [6, 10, 7, 11] investigate the behaviour of service systems with an attached inventory. All these models assume that the demands, which arrive during the time the inventory is zero, is back ordered.

In this chapter we consider two cases of $M/G/1$ type service inventory systems: without lead time and with positive lead time. In the first case we consider an (s, S) inventory system with positive service time, that is an $M/G/1$ type system. Lead time is zero. Arrival process is Poisson with rate λ and service time is distributed as $G(\cdot)$ with $\mu = \int_0^\infty [1 - G(t)]dt < \infty$. The queue discipline is FIFS and items are served one by one. We analyse the systems at departure epochs and get a product form solution to the system probability and from that we obtain the arbitrary time system state distribution. Different performance measures are derived and control policy is discussed. In the second case we consider $M/G/1$ type system with positive lead time and lost sales. Lead time is assumed to be exponentially distributed. When the inventory level is zero, no customer is allowed. However, those who are already present waiting are not turned away. Thus the queue size is given by the set of non negative integers. We discuss two models. In the first model the system follows (r, Q) policy in which when the inventory level becomes r, an order for replenishment is placed. The quantity ordered is of fixed size Q $(Q > r)$. In the second model we consider (r, s) policy in which with each replenishment the inventory level is restocked to exactly $S < \infty$ units with $r < S$ no matter how many items are still present in inventory. In both these models we get analytical solution to system state probability distribution. To be precise we obtain the long run system probability distribution in product form. Several performance measures are considered.

The lost sales situation arises in many retail establishments [13], where the intense competition allows customers to choose another brand or to go to another store. But there are other areas of applications, where lost sales models are appropriate as well. For example, these models apply to cases such as essential spare parts where one must go to the outside of the normal ordering system when a stockout occurs [12, p. 605]. The essential spare part problem is central for many repair procedures, where broken down units arrive at a repair station, queue for repair, and are repaired by replacing a failed part by a spare part from the inventory. A similar problem arises in production processes where rough material items are needed to let the production process run. Both of these problems are modelled using pure service systems, but these queueing theoretical models neglect inventory management. Schwarz et al. [54] consider $M/M/1$ queueing systems with inventory with lost sales. They analyse several single server queueing systems of $M/M/1$-type with an attached inventory. The systems under investigation differ with respect to the size of replenishment orders and the order policy. Every system under consideration has the property that no customers are allowed to join the queue as long as the inventory is empty. This corresponds to the lost sales case of inventory management. However, if inventory is at hand, customers are still admitted to enter the waiting room even if the number of customers in the system exceeds the inventory on hand. In all the models they discussed product form solutions were obtained to the steady state probabilities. That means in the long run and in equilibrium the queue length process and the inventory process behave as if they are independent. This is a rather strange observation because, these processes strongly interact independently of whether being in equilibrium or not.

As in Schwarz et al [54] we discuss the same models in $M/G/1$ type systems. The stationary distribution of $M/G/1$ queue at arbitrary epochs is the same as that at departure epochs. We get a product form solution for the system state distribution.

5.2 $M/G/1$ type inventory system without lead time

Here we consider an (s, S) inventory system with positive service time. Inter arrival distribution is exponential with parameter λ and service times are independently and identically distributed with distribution function $G(\cdot)$; it is natural and essential to assume that $\mu = \int_0^\infty (1 - G(t))dt < \infty$. Lead time is assumed to be zero.

5.2.1 Mathematical modelling and Analysis

Let $N(t)$= Number of customers at time t
$I(t)$= Number of items in inventory at time t.

Then $\{(N(t), I(t));\ t \in R^+\}$ forms a Markov chain at post departure epochs with state space

$$\{(0, s), (0, s+1), \ldots, (0, S-1)\} \cup \ \{\cup_{i\geq 1}\{(i, s+1), (i, s+2) \ldots (i, S)\}\}.$$

Let $\underline{0} = \{(0, s), (0, s+1), \ldots (0, S-1)\}$ and $\underline{i} = \{(i, s+1) \ldots (i, S)\}$ for $i \geq 1$. The transition probability matrix associated with this Markov chain is given by

$$\mathcal{P} = \begin{array}{c} \\ \underline{0} \\ \underline{1} \\ \underline{2} \\ \underline{3} \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccc} \underline{0} & \underline{1} & \underline{2} & \underline{3} & \cdots \end{array} \\ \left(\begin{array}{ccccccc} A_0 & A_1 & A_2 & A_3 & \cdots & & \\ A_0 & A_1 & A_2 & A_3 & \cdots & & \\ 0 & A_0 & A_1 & A_2 & A_3 & \cdots & \\ 0 & 0 & & A_0 & A_1 & A_2 & \cdots \\ & \ddots & \ddots & \ddots & & & \end{array} \right) \end{array}$$

Note that this matrix is of the $M/G/1$ type (Neuts [46]). In the above,

$$A_0 = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_0 \\ a_0 & 0 & \cdots & 0 & 0 \\ 0 & a_0 & \cdots & 0 & 0 \\ & & & a_0 & 0 \end{bmatrix}_{(S-s-1)\times(S-s-1)}$$

$$A_i = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_i & 0 \\ 0 & 0 & \cdots & 0 & 0 & a_i \\ a_i & 0 & \cdots & 0 & 0 & 0 \\ 0 & a_i & \cdots & 0 & 0 & 0 \\ 0 & & \ddots & & & \\ 0 & & & \ddots & & \\ 0 & 0 & \cdots & a_i & 0 & 0 \end{bmatrix}_{(S-s-1)\times(S-s-1)}$$

where $a_i = \Pr(i$ arrivals during a service time)
$= \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^i}{i!}dG(t); \; i = 0, 1, \ldots$

Theorem 5.2.1. *The system state distribution has a product form solution given by*

$$\boldsymbol{\pi}_{ij} = \lim_{n\to\infty} P(N_n = i, I_n = j),$$

where N_n is the queue length immediately after the nth service completion and I_n is the corresponding inventory level.

$$= \pi_i \frac{1}{Q}; \quad \textit{with } j = s, s+1, \ldots, S-1 \textit{ for } i = 0$$
$$\textit{and } j = s+1, \ldots S \textit{ for } i \geq 1$$

where π_i is the stationary probability of i customers in the $M/G/1$ queue and $Q = S - s$.

Proof. Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots)$ be the stationary probability vector associated with the Markov chain, where

$$\boldsymbol{\pi}_0 = (\boldsymbol{\pi}_{(0,s)}, \boldsymbol{\pi}_{(0,s+1)}, \ldots \boldsymbol{\pi}_{(0,S-1)})$$
$$\text{and} \quad \boldsymbol{\pi}_i = (\boldsymbol{\pi}_{(i,s+1)}, \boldsymbol{\pi}_{(i,s+2)}, \ldots \boldsymbol{\pi}_{(i,S)}) \quad \text{for } i \geq 1$$

As in $M/G/1$ queue, here also if $\rho = \lambda\mu < 1$ the Markov chain is ergodic. Hence the stationary probabilities are given by the unique solution of $\boldsymbol{\pi} = \boldsymbol{\pi}\mathcal{P}$

$$\begin{aligned} \text{i.e.,} \quad \boldsymbol{\pi}_0 A_0 + \boldsymbol{\pi}_1 A_0 &= \boldsymbol{\pi}_0 \\ \boldsymbol{\pi}_0 A_1 + \boldsymbol{\pi}_1 A_1 + \boldsymbol{\pi}_2 A_0 &= \boldsymbol{\pi}_1 \\ \boldsymbol{\pi}_0 A_2 + \boldsymbol{\pi}_1 A_2 + \boldsymbol{\pi}_2 A_1 + \boldsymbol{\pi}_3 A_0 &= \boldsymbol{\pi}_2 \\ &\vdots \end{aligned} \tag{5.1}$$

Substituting

$$\begin{aligned} \boldsymbol{\pi}_0 &= \big(\boldsymbol{\pi}_{(0,s)}, \boldsymbol{\pi}_{(0,s+1)}, \ldots \boldsymbol{\pi}_{(0,S-1)}\big) \\ &= \pi_0(\frac{1}{Q}, \ldots, \frac{1}{Q}) = \pi_0 \cdot \frac{1}{Q}(1, \ldots, 1) \\ &= \pi_0 \frac{1}{Q} e, \quad \text{where } e = (1, \ldots, 1) \\ \text{and } \boldsymbol{\pi}_i &= \big(\boldsymbol{\pi}_{(i,s+1)}, \boldsymbol{\pi}_{(i,s+2)}, \ldots \boldsymbol{\pi}_{(i,S)}\big) \\ &= \pi_i \frac{1}{Q} e \text{ for } i \geq 1 \end{aligned}$$

in (5.1) we get the solution which turns out to be unique due to normalizing condition.
In the above π_i; $i \geq 0$, is the stationary probabilities that there are i customers at a departure epoch (and hence at arbitrary epoch) in an $M/G/1$ queue. Thus the stationary probability distribution of the system at departure epoch is given by $\boldsymbol{\pi}_{ij} = \pi_i \frac{1}{Q}$, for $i = 0$ and $j = s, s+1, \ldots S-1$ and for $i = 1$ we have $j = s+1 \ldots, S$. □

Theorem 5.2.2. *For the $M/G/1$ queue the distribution of the number of customers at departure epoch is the same as that of the number of customers at arbitrary epoch, hence we have $\lim_{t \to \infty} Pr(N(t) = i,\ I(t) = j) = \pi_i \frac{1}{Q}$; $j = s, s+1 \ldots S-1$ for $i = 0$ and $j = s+1, s+2 \ldots S$ for $i \geq 1$.*

5.2.2 Pollaczek-Khintchine-type formula

Since the system state distribution is in product form the Pollaczek-Khintchine type formula can be derived as

$$\boldsymbol{\pi}(z,r) = \frac{(1-\rho)(1-z)K(z)U(r)}{K(z)-z}$$

where $\frac{(1-\rho)(1-z)K(z)}{K(z)-z}$ is the Pollaczek-Khintchine formula for $M/G/1/\infty$ queue and $U(r) = E(r^I) = \sum_{k=s+1}^{S} \frac{1}{Q} r^k = \frac{1}{Q} r^{s+1} \frac{1-r^{Q+1}}{1-r}$ is the generating function of the number of items in the inventory.

5.2.3 Busy-Period Analysis

In this section we determine the distribution of the busy period for $M/G/1$-type (s, S) inventory system without lead time. Here the busy period analysis starts with an arrival to an idle server having inventory level i, $(1 \leq i \leq S-1)$ (since the optimum of s is $s^* = 0$) until the server again becomes idle with an inventory level j, $(1 \leq j \leq S-1)$. Let $B(x)$ denote the cumulative distribution function of the busy period X of the $M/G/1$-type inventory system with service cumulative distribution $G(t)$. Then we condition X on the length of the first service time inaugurating the busy period when the inventory level is $i \geq 0$. Since each arrival during that service time will contribute to the busy period by having arrivals come during its service time, we can look at each arrival during the first service time of the busy period as essentially generating its own busy period.

Thus

$$\begin{aligned}
B(x) = \text{Pr}\big\{&\text{busy period generated by all arrivals during } (0,t] \\
&\text{is less than or equal to } x-t \,|\, \text{first customers} \\
&\text{service completed in } (t, t+dt) \text{ with initial} \\
&\text{inventory level } i \text{ and final level } j\big\} \\
= \int_0^x &\sum_{n=1}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} B^{*n}(x-t)\, dG(t)
\end{aligned}$$

where B^{*n} is the n-fold convolution of B with itself. Now we can evaluate the mean length of busy period.

$$E(X) = \int_0^{\infty} [1 - B(x)]\, dx.$$

With some computations this reduces to
E($X|$ inventory level at the start of the cycle $= i$)

$$= \frac{\mu\pi_{(0,i)}}{1 - \lambda\mu\pi_{(0,i)}} = \frac{\mu\pi_{(0,i)}}{1 - \rho\pi_{(0,i)}}$$

5.2.4 Performance measures

i) Average inventory size

$$\begin{aligned}
\mu_{inv} &= \sum_{j=s}^{S-1} j\boldsymbol{\pi}(0,j) + \sum_{i=1}^{\infty} \sum_{j=s+1}^{S} j\boldsymbol{\pi}(i,j) \\
&= s + \frac{Q+1}{2}.
\end{aligned}$$

ii) Average number of customers in the system

$$\mu_c = \sum_{i=1}^{\infty} i\boldsymbol{\pi}_i = \sum_{i=1}^{\infty} i\pi_i \frac{1}{Q}.$$
$$= \frac{1}{Q}\sum_{i=1}^{\infty} i\pi_i = \frac{1}{Q}(\text{ Expected number of customers in the } M/G/1 \text{ queue}).$$

iii) Expected cycle length from replenishment to replenishment is equal to expected time for Q services. There are the following four cases.

Case-1 At a replenishment epoch there are at least Q customers.
Then

$$E(\text{ cycle length}) = E(\text{ time to serve } Q \text{ customers})$$
$$= Q \cdot \mu$$

where $\mu = E$ (service time per customers).

Case-2 At a replenishment epoch there are $k(< Q)$ customers. During the service time of these k customers if at least $Q - k$ customers join then expected cycle length is $= Q \cdot \mu$.

Case-3 After each service, server has to wait for the next arrival i.e., there are Q idle periods in between Q services.
Distribution of service time of these Q customers $= [G * \exp(\lambda)]^{*Q}$
$\therefore E$ (cycle length) $= (\frac{1}{\lambda} + \mu)Q$.

Case-4 There are $j(< Q)$ idle periods in between Q services.
E (cycle length) $= \mu Q + \frac{j}{\lambda}; j = 1, 2, \ldots Q - 1$.

5.2.5 Cost function

We define a cost function as follows:

$$F(s,Q) = h\mu_{inv} + \frac{K + CQ}{\text{(expected cycle length)}}$$

where h is the holding cost of inventory per unit per unit time, K is the fixed order cost and C is the cost of inventory per unit.

$$\begin{aligned} F(s,Q) &= h(s + \frac{Q+1}{2}) + \frac{K + CQ}{\sum_{j=0}^{Q} \frac{j}{\lambda} + \mu Q} \\ &= h(s + \frac{Q+1}{2}) + \frac{K + CQ}{\mu Q + \frac{1}{\lambda}\frac{Q(Q+1)}{2}}. \end{aligned}$$

The above function is linear in s. Since $h > 0$, value s^* of s that minimizes the expected cost per cycle is given by $s^* = 0$. Now we have to find Q^* the optimum value of Q that minimizes

$$F(Q) = h(\frac{Q+1}{2}) + \frac{K + CQ}{\mu Q + \frac{1}{\lambda}\frac{Q(Q+1)}{2}}.$$

Note that $\frac{d^2F(Q)}{dQ^2} \geq 0$. Hence the function is convex with respect to Q. So there is a value of Q, say Q^* for which

$$F(Q^* + 1) \geq F(Q^*) \quad \text{and} \quad F(Q^* - 1) \geq F(Q^*).$$

5.3 $M/G/1$-Type system with positive lead time and lost sales

In this case we consider $M/G/1$ type inventory system with positive lead time. An additional feature we are considering is “lost sales”. That means when the inventory level is zero, no customer is allowed to join the queue.

However, customers join when a service is going on, independent of the number of inventoried items. Here we consider two models. In the first model the replenishment policy is (r, Q). In this model an order for replenishment is placing when the inventory level reaches the point r. Here a fixed quantity Q is ordered for each time. In the second model we consider the (r, S) policy. Here with each replenishment the inventory level is restocked to exactly $S < \infty$ units, ($r < S$), no matter how many items are still present in the inventory.

5.3.1 $M/G/1$ type system with (r, Q)-policy

We have a single server with infinite waiting room under FCFS regime and an attached inventory of capacity M (identical) items. Each customer needs exactly one item from the inventory for service, and the on-hand inventory decreases by one at the moment of service completion. If the server is ready to serve a customer which is at the head of the line and there is no item of inventory this service starts only at the time instant (and then immediately) when the next replenishment arrives at the inventory. Customers arriving during a period when the server waits for replenishment order are rejected and lost to the system (lost sales).

Customers are of stochastically identical behaviour. To the server there is a Poisson (λ) arrival stream ($\lambda > 0$). Customers request an amount of service time which is arbitrarily distributed with distribution function $G(\cdot)$. Here $\mu = \int_0^\infty [1 - G(t)]dt$. If the on-hand inventory reaches a prespecified value $r \geq 0$, a replenishment order is instantaneously triggered. The size of the replenishment order is fixed to $Q < \infty$ units, $Q > r$. We fix $M = r + Q$. The replenishment lead time is exponentially distributed with parameter v, $v > 0$. Let $X(t_i)$ denote the number of customers left behind at the i^{th} departure epoch and $Y(t_i)$ denote the on-hand inventory at time t_i, $i = 1, 2, \ldots$. Then the embedded stochastic process $Z = \{(X(t_i), Y(t_i))\}$

where t_i, $i = 1, 2, \ldots$ are the successive times of completion of service, is a Markov chain with the state space

$$E_z = \{(n, k);\ n \in N_0;\ 0 \le k \le Q + r\}.$$

The transition probability matrix of the embedded Markov chain is given by

$$\mathcal{P} = \begin{array}{c} \\ \underline{0} \\ \underline{1} \\ \underline{2} \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccc} \underline{0} & \underline{1} & \underline{2} & \cdots \end{array} \\ \left(\begin{array}{ccccc} A_0 & A_1 & A_2 & \cdots & \\ A_0 & A_1 & A_2 & \cdots & \\ \ddots & A_0 & A_1 & A_2 & \cdots \\ & \ddots & & & \end{array} \right) \end{array}$$

where

$$\underline{0} = \{(0, k);\ 0 \le k \le r + Q\}$$
$$\underline{i} = \{(i, k);\ 0 \le k \le r + Q\} \quad i = 1, 2, \ldots.$$

$$A_0 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ \vdots \\ (0,r) \\ (0,r+1) \\ \vdots \\ (0,Q+r) \end{array} \begin{array}{c} \begin{array}{ccccccccccc} (0,0) & (0,1) & \cdots & (0,r-1) & (0,r) & \cdots & (0,Q-1) & (0,Q) & \cdots & & (0,Q+r) \end{array} \\ \left(\begin{array}{ccccccccccc} 0 & 0 & \cdots & 0 & 0 & \cdots & a_0 & 0 & \cdots & & 0 \\ b_0 & 0 & \cdots & 0 & 0 & \cdots & 0 & rb_0 & \cdots & & 0 \\ & b_0 & & & & & & & & & \\ 0 & 0 & \cdots & b_0 & 0 & \cdots & 0 & 0 & \cdots & rb_0 & 0 \\ 0 & 0 & \cdots & 0 & a_0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ & & & & & & & & & & \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & a_0 & 0 \end{array} \right) \end{array}$$

$$A_i = \begin{bmatrix} 0 & 0 & \cdots & \cdots & a_i & 0 & 0 & \cdots & 0 & 0 \\ b_i & 0 & \cdots & \cdots & 0 & rb_i & 0 & \cdots & 0 & 0 \\ 0 & b_i & \cdots & \cdots & 0 & 0 & rb_i & \cdots & 0 & 0 \\ & & \cdots & b_i & 0 & 0 & 0 & \cdots & rb_i & 0 \\ & & \cdots & & a_0 & & & & & \\ & & & & & & & & & \\ & & & & & & & & a_0 & 0 \end{bmatrix}_{(Q+r+1)\times(Q+r+1)}$$

where a_i =Pr(i arrivals during a service time) as in the $M/G/1$ queue

$$= \int_0^\infty (\lambda t)^i \frac{e^{-\lambda t}}{i!} dG(t).$$

$b_i = Pr(i$ arrivals during a service time and no replenishment).

$= \int_0^\infty dG(t) e^{-\alpha t} \wedge_i (t)$ where $\wedge_i(t)$ denotes the probability of i arrivals.

$rb_i =$ Pr(i arrivals during a service time and replenishment)

$$= \int_0^\infty dG(t)[1 - e^{-\alpha t}] \wedge_i (t).$$

Theorem 5.3.1. *The embedded Markov chain Z is ergodic if and only if $\lambda < \mu$ and $\nu < \infty$. If Z is ergodic then it has a unique stationary distribution of product form:*

$$\boldsymbol{\pi}(n,k) = A^{-1} C(k) \pi_n \textit{ for } n \in \mathbb{N}_0,\ 1 \leq k \leq Q + r \tag{5.2}$$

$$\boldsymbol{\pi}(n,0) = A^{-1} \frac{\lambda}{v} \pi_n, \textit{ for } n \in \mathbb{N}_0 \tag{5.3}$$

with

$$C(k) = (\frac{\lambda + \nu}{\lambda})^{k-1}; \quad k = 1, 2, \dots r$$

$$C(k) = (\frac{\lambda + \nu}{\lambda})^{r}; \quad k = r+1, \dots Q$$

$$C(k+Q) = (\frac{\lambda + \nu}{\lambda})^{r} - (\frac{\lambda + \nu}{\lambda})^{k-1}; \quad k = 1, 2, \dots r.$$

The normalization constant is

$$A = (Q(\frac{\lambda + \nu}{\lambda})^r + \frac{\lambda}{\nu})$$

$\pi_n =$ *Pr(n no. of customers in an $M/G/1$ queue).*

Proof. Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \cdots)$ denote the stationary probability vector associated with Z; where $\boldsymbol{\pi}_i = (\boldsymbol{\pi}_{(i,0)}, \boldsymbol{\pi}_{(i,1)}, \dots)$, stationary vector associated to level i. Then $\boldsymbol{\pi}$ is given by

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathcal{P}$$

$$(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \dots) = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \dots) \begin{pmatrix} A_0 & A_1 & \cdots & \cdots \\ A_0 & A_1 & \cdots & \cdots \\ & A_0 & A_1 & \cdots \\ & \ddots & \ddots & \cdots \end{pmatrix}$$

i.e.,

$$\begin{aligned} \boldsymbol{\pi}_0 A_0 + \boldsymbol{\pi}_1 A_0 &= \boldsymbol{\pi}_0 \\ \boldsymbol{\pi}_0 A_1 + \boldsymbol{\pi}_1 A_1 + \boldsymbol{\pi}_2 A_0 &= \boldsymbol{\pi}_1 \\ \boldsymbol{\pi}_0 A_2 + \boldsymbol{\pi}_1 A_2 + \boldsymbol{\pi}_2 A_1 + \boldsymbol{\pi}_3 A_3 &= \boldsymbol{\pi}_2 \\ &\vdots \end{aligned} \tag{5.4}$$

□

Substituting the values (5.2) and (5.3) in the above system of equations we get the product form solution to the embedded Markov chain Z. For an $M/G/1$ queue the steady state solution at the departure epochs is the same as the solution at arbitrary epochs. Thus we can conclude that the above result is the steady state solution at arbitrary epoch.

5.3.2 Measures of system performance

1. Average on-hand inventory

$$\begin{aligned} \overline{I} &= \sum_{k=1}^{Q+r} k \sum_{n \geq 0} \boldsymbol{\pi}(n, k) \\ &= \sum_{k=1}^{Q+r} k \sum_{n \geq 0} A^{-1} C(k) \cdot \pi_n. \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{r} k \sum_{n\geq 0} A^{-1}(\frac{\lambda+\nu}{\lambda})^{k-1}\pi_n + \sum_{k=r+1}^{Q} k \sum_{n\geq 0} A^{-1}(\frac{\lambda+\nu}{\lambda})^{r}\pi_n \\
&\quad + \sum_{k=1}^{r}(k+Q)\sum_{n\geq 0} A^{-1}((\frac{\lambda+\nu}{\lambda})^{r} - (\frac{\lambda+\nu}{\lambda})^{k-1})\pi_n. \\
&= \sum_{n\geq 0} A^{-1}\pi_n \sum_{k=1}^{r} k(\frac{\lambda+\nu}{\lambda})^{k-1} + \sum_{n\geq 0} A^{-1}\pi_n \sum_{k=r+1}^{Q} k(\frac{\lambda+\nu}{\lambda})^{r} \\
&\quad + \sum_{n\geq 0} A^{-1}\pi_n \sum_{k=1}^{r}(k+Q)((\frac{\lambda+\nu}{\lambda})^{r} - (\frac{\lambda+\nu}{\lambda})^{k-1}).
\end{aligned}
$$

2. Mean number of customers arriving per unit time

$$
\begin{aligned}
\lambda_A &= \lambda \sum_{n\geq 0}\sum_{k=1}^{M} \boldsymbol{\pi}(n,k) \\
&= \lambda \cdot \sum_{n\geq 0}\sum_{k=1}^{M} A^{-1}C(k)\cdot \pi_n.
\end{aligned}
$$

3. Mean number of departures per unit time

$$
\lambda_D = \mu \cdot \sum_{n\geq 1}\sum_{k=1}^{M} \boldsymbol{\pi}(n,k) = \mu \sum_{n\geq 1}\sum_{k=1}^{M} A^{-1}C(k)\pi_n.
$$

4. Intensity of arrival of replenishment orders

$$
\lambda_R = \nu \sum_{n\geq 0}\sum_{i=0}^{r} \boldsymbol{\pi}(n, i+Q).
$$

5. Average number of lost sales incurred per unit of time

$$
\begin{aligned}
\overline{LS} &= \lambda \cdot \sum_{n\geq 0} \boldsymbol{\pi}(n,0) \\
&= \lambda \sum_{n\geq 0} A^{-1}\frac{\lambda}{\nu}\pi_n.
\end{aligned}
$$

6. Expected number of lost sales per cycle

$$\overline{LS_c} = \frac{\overline{LS}}{\lambda_R}.$$

7. Safety stock or the expected net inventory position just before a replenishment order arrives

$$s = \frac{\nu}{\lambda_R} \sum_{k=1}^{r} k \cdot \sum_{n \geq 0} \boldsymbol{\pi}(n, k).$$
$$= \frac{\nu}{\lambda_R} \sum_{n \geq 0} \sum_{k=1}^{r} k \cdot A^{-1} (\frac{\lambda + \nu}{\lambda})^{k-1} \pi_n.$$

8. Mean number of customers in the system $\overline{L_0}$ and the mean number of waiting customers $\overline{L}$ are the same as in the usual $M/G/1$ queueing system. Using Little's formula the mean sojourn time $\overline{W}_0$ and the mean waiting time $\overline{W}$ of customers are

$$\overline{W}_0 = \frac{\overline{L_0}}{\lambda_A}$$
$$\overline{W} = \frac{\overline{L}}{\lambda_A}.$$

5.3.3 $M/G/1$ type system with (r, S)-policy

If the inventory reaches a specified value $r > 0$, the replenishment order is instantaneously triggered. With each replenishment the inventory level is restocked to exactly $S < \infty$ units with $r < S$, no matter how many items are still present in the inventory. We set $M = S$. The replenishment lead time is exponetially distributed with parameter ν, $\nu > 0$. During the time the inventory is zero, no customer is admitted to join the queue. Let $(X(t_i))$ denote the number of customers left behind at the time of i^{th} departure (t_i) and $Y(t_i)$ denote the inventory level at time t_i, then $Z = \{(X(t_i), Y(t_i))\}$;

denote an embedded Markov chain with state space

$$E_z = \{(n,k);\ n \in N_0;\ 0 \le k \le S\}.$$

Theorem 5.3.2. *The embedded Markov chain is ergodic if $\lambda < \mu$. If $\lambda < \mu$ then Z has a unique stationary distribution of product form*

$$\boldsymbol{\pi}(n,k) = A^{-1}C(k)\pi_n \quad n \in \mathbb{N}_0 \textit{ and } 1 \le k \le S \tag{5.5}$$

$$\boldsymbol{\pi}(n,0) = A^{-1}\frac{\lambda}{\nu}\pi_n \quad \textit{for } n \in \mathbb{N}_0 \tag{5.6}$$

$$\textit{with } C(k) = (\frac{\lambda+\nu}{\lambda})^{k-1}; \quad k = 1,2,\ldots r$$

$$C(k) = (\frac{\lambda+\nu}{\lambda})^{r}; \quad k = r+1,\ldots S$$

and normalization constant $A = (S - r + \frac{\lambda}{\nu})(\frac{\lambda+\nu}{\lambda})^r$.

Proof. The transition probability matrix $\mathcal{P}$ associated with Z is given by

$$\mathcal{P} = \begin{bmatrix} A_0 & A_1 & \cdots & \cdots \\ A_0 & A_1 & \cdots & \cdots \\ & A_0 & A_1 & \cdots \\ & \ddots & \ddots & \\ & & & \end{bmatrix}$$

with

$$A_0 = \begin{array}{c} (0,0) \\ (0,1) \\ \vdots \\ (0,r) \\ (0,r+1) \\ \vdots \\ (0,S) \end{array} \begin{pmatrix} & (0,0) & (0,1) & \cdots & (0,r-1) & (0,r) & \cdots & (0,S-1) & (0,S) \\ & 0 & 0 & \cdots & 0 & 0 & \cdots & a_0 & 0 \\ & b_0 & 0 & \cdots & 0 & 0 & \cdots & rb_0 & 0 \\ & 0 & b_0 & & & & & & \\ & 0 & 0 & \cdots & b_0 & 0 & \cdots & rb_0 & 0 \\ & 0 & 0 & \cdots & 0 & a_0 & \cdots & 0 & 0 \\ & & & & & & a_0 & & \\ & 0 & 0 & 0 & 0 & \cdots & a_0 & 0 & \end{pmatrix}$$

$$A_i = \begin{bmatrix} 0 & 0 & \cdots & \cdots & a_i & 0 \\ b_i & & & & rb_i & 0 \\ & b_i & \cdots & \cdots & & \\ & & b_i & \cdots & rb_i & 0 \\ & & & a_0 & 0 & 0 \\ & & & & a_0 & \\ & & & & & a_0 0 \end{bmatrix}_{(S+1)\times(S+1)}$$

where a_i, b_i and rb_i are like in (r, Q) policy.

Let $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1 \dots)$ be the stationary probability vector associated with Z then $\boldsymbol{\pi}$ is given by $\boldsymbol{\pi} = \boldsymbol{\pi}\mathcal{P}$.

$$\begin{aligned} \boldsymbol{\pi}_0 &= \boldsymbol{\pi}_0 A_0 + \boldsymbol{\pi}_1 A_0 \\ \boldsymbol{\pi}_1 &= \boldsymbol{\pi}_0 A_1 + \boldsymbol{\pi}_1 A_1 + \boldsymbol{\pi}_2 A_0 \\ \boldsymbol{\pi}_2 &= \boldsymbol{\pi}_0 A_2 + \boldsymbol{\pi}_1 A_2 + \boldsymbol{\pi}_2 A_1 + \boldsymbol{\pi}_3 A_0. \\ &\vdots \end{aligned}$$

□

Substituting (5.5) and (5.6) in the above system of equations we get the product form solution to the embedded Markov chain Z. Since in $M/G/1$ queue the steady state probabilities at departure epochs are the same as the continuous time probabilities, here also we get a product form solution at arbitrary epochs.

5.3.4 Performance measures

1. Average on-hand inventory

$$\overline{I} = \sum_{k=1}^{S} k \sum_{n\geq 0} \boldsymbol{\pi}(n,k)$$
$$= \sum_{k=1}^{r} k \sum_{n\geq 0} A^{-1}(\frac{\lambda+\nu}{\lambda})^{k-1}\pi_n + \sum_{k=r+1}^{S} k \sum_{n\geq 0} A^{-1}(\frac{\lambda+\nu}{\lambda})^{r}\pi_n$$

2. Mean number of customers arriving per unit time

$$\lambda_A = \lambda \cdot \sum_{n\geq 0}\sum_{k=1}^{S} \boldsymbol{\pi}(n,k)$$
$$= \lambda \cdot \sum_{n\geq 0}\sum_{k=1}^{M} A^{-1}C(k)\cdot \pi_n.$$

3. Mean number of departures per unit time

$$\lambda_D = \mu \cdot \sum_{n\geq 1}\sum_{k=1}^{S} \boldsymbol{\pi}(n,k) = \mu \sum_{n\geq 1}\sum_{k=1}^{S} A^{-1}C(k)\pi_n.$$

4. Intensity of arrival of replenishment orders

$$\lambda_R = \nu \sum_{n\geq 0}\sum_{i=0}^{r} \boldsymbol{\pi}(n,i)$$
$$= \nu \sum_{n\geq 0}\sum_{i=0}^{r} A^{-1}(\frac{\lambda+\nu}{\lambda})^{i-1}\pi_n$$
$$= \nu \cdot \sum_{n\geq 0} A^{-1}\pi_n \cdot \frac{(\frac{\lambda+\nu}{\lambda})^r - 1}{\nu/\lambda}$$
$$= \sum_{n\geq 0} A^{-1}\pi_n \lambda[(\frac{\lambda+\nu}{\lambda})^r - 1]$$

5. Average number of lost sales incurred per unit time

$$\begin{aligned}\overline{LS} &= \lambda \sum_{n\geq 0} \boldsymbol{\pi}(n,0)\\ &= \lambda \cdot \sum_{n\geq 0} A^{-1}\frac{\lambda}{\nu}\pi_n.\end{aligned}$$

6. Expected number of lost sales per cycle

$$\overline{LS_c} = \frac{\overline{LS}}{\lambda_R}.$$

7. Safety stock

$$\begin{aligned}s &= \frac{\nu}{\lambda_R}\sum_{k=1}^{r} k \sum_{n\geq 0} \boldsymbol{\pi}(n,k).\\ &= \frac{\nu}{\lambda_R}\sum_{k=1}^{r} k \cdot \sum_{n\geq 0} A^{-1}(\frac{\lambda+\nu}{\lambda})^{k-1}\pi_n.\end{aligned}$$

8. Mean number of customers in the system $\overline{L_0}$ and the mean number of waiting customers $\overline{L}$ are the same as in the usual $M/G/1/\infty$ queueing system. Using Little's formula the mean sojourn time $\overline{W}_0$ and the mean waiting time $\overline{W}$ of customers are

$$\begin{aligned}\overline{W}_0 &= \frac{\overline{L_0}}{\lambda_A}\\ \overline{W} &= \frac{\overline{L}}{\lambda_A}.\end{aligned}$$

5.4 Conclusion

In this chapter we considered an inventory with arbitrarily distributed service time. To start with we considered an inventory system without lead time. We got an $M/G/1$-type transition probability matrix and by analysis we derived a product form solution to the system state probability distribution. In the second case we considered two models with positive lead time. We discussed two replenishment policies. Here also we got product form solution. That means in the long run and in equilibrium the queue length process and the inventory process behave as if they are independent. This is rather a strange observation because these process strongly interact, independently of whether being in equilibrium or not.

Chapter 6

D-Policy for a Production Inventory System with Perishable Items

6.1 Introduction

So far we were concentrating on invenory with positive (random) service time. In this chapter we concentrate on an inventory system with negligible service time coupled with a control policy. We consider a production inventory system of items having random lifetimes. We introduce a D-policy. According to this policy the production process is activated only when the accumulated workload exceeds a threshold $D\ (> 0)$. Workload accumulates due to demands and perishability. Once the production process is switched on, it continues till the inventory level reaches a maximum $S\ (> 0)$. When the inventory level reaches zero, and the accumulated workload has not reached the threshold value D, then the production starts immediately to minimize loss of demands. Whenever the inventory level reaches S, the production is stopped. Thereafter there is depletion in the inventory due to demand and/ decay. We assume that only items that are kept in the inventory are subject to decay. The advantage of this problem is that if the production cost is high, we can find an optimal D which minimizes the overall cost or maximizes the net profit. We obtain the system state distribution, expected length of time the production is continuously on and an optimal D which maximizes the profit.

Very little investigation on production inventory has been made in the past. Kalpakam and Sapna [28] analyse a perishable (s, S) system with Poisson demands and exponentially distributed lead times. They analyse the case of arbitrarily distributed lead times in [29]. In both the cases demand occurring during stock-out periods are assumed to be lost. Sharafali [55] considers a production inventory operating under (s, S) policy where demands arrive according to Poisson process and production output is also Poisson. Altiok [1] analyses a production inventory system with compound Poisson demand and phase-type distribution for the processing time. Krishnamoorthy and Raju [36] discussed the case N-policy for a production inventory system with random lifetimes. N-policy is introduced to inventory by Krishnamoorthy and Raju. N is the number of backlogs required to start the production. They obtain some performance measures and an optimal N which minimizes a suitable cost function. Krishnamoorthy and Mohammed Ekramol Islam [32] analysed a production inventory with retrial of customers. Krishnamoorthy and Ushakumari [38] consider a D-policy for a k-out-of-n: G system with repair. They obtain system state distribution, system reliability, expected number of times the system is down in a cycle *etc*. They obtain an optimal D value which maximizes a suitably defined cost function.

6.2 Mathematical modelling and analysis

In this model no backlog is allowed and the machine does not breakdown. Demands arrive according to Poisson process of rate λ. Lifetime of items is exponentially distributed with parameter β. The production time is also exponentially distributed with parameter μ. We obtain the stationary distribution of the system by identifying a continuous time bivariate Markov chain.

Let $I(t); t \geq 0$ be the inventory level at time t.
$I(t)$ can take values in $\{0, 1, 2, \dots S\}$.
Let

$$X(t) = \begin{cases} 1 & \text{if the production process is on at time } t \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{(I(t), X(t));\ t \geq 0\}$ is a two-dimensional Markov process in the state space

$$E = \{(0,1), (1,1) \cdots (S-1,1)\} \cup \{(1,0), (2,0), \cdots (S,0)\}.$$

Let

$$\begin{aligned} G(r,a) &= Pr \text{ (the production is not activated at the } r^{\text{th}} \text{ demand and/ failure)} \\ &= Pr \text{ (the workload is less than } D \text{ at the } r^{\text{th}} \text{ demand and/ failure)} \\ &= 1 - \sum_{j=0}^{r-1} e^{-\mu D} \frac{(\mu D)^j}{j!}; \quad r = 0, 1, \dots S-1 \end{aligned}$$

where $a = \mu D$.
Also $P_{ij}(t) = Pr\Big((I(t), X(t)) = (i,j) | (I(0), X(0)) = (S,0)\Big)$ for $(i,j) \in E$. The Kolmogorov forward differential equations are given by

$$\begin{aligned} P'_{S,0}(t) =& -(\lambda + S\beta)P_{S,0}(t) + \mu P_{S-1,1}(t). \\ P'_{i,0}(t) =& -(\lambda + i\beta)P_{i,0}(t) + [\lambda + (i+1)\beta]P_{i+1,0}(t)G(S-i,a) \\ & \text{for } i = S-1, S-2, \dots, 1. \\ P'_{i,1}(t) =& -(\lambda + i\beta + \mu)P_{i,1}(t) + \mu P_{i-1,1}(t) \\ & + (\lambda + (i+1)\beta)P_{i+1,0}(1 - G(S-i,a)) \\ & + (1 - \delta_{i,S-1})(\lambda + (i+1)\beta)P_{i+1,1}(t) \\ & \text{for } i = S-1, S-2, \dots, 1 \\ P'_{0,1}(t) =& -\mu P_{0,1}(t) + (\lambda + \beta)P_{11}(t) + (\lambda + \beta)P_{1,0}(t). \end{aligned}$$

6.2.1 Limiting probabilities

Let

$$\begin{aligned} q_{i,0} &= \lim_{t\to\infty} P_{i,0}(t), \quad for \quad i = 1,2,\dots S \\ q_{i,1} &= \lim_{t\to\infty} P_{i,1}(t), \quad for \quad i = 0,1,2,\dots S-1 \\ q_{S,0} &= \lim_{t\to\infty} P_{S,0}(t). \end{aligned}$$

Then

$$\begin{aligned} q_{i,0} &= \frac{\lambda + S\beta}{\lambda + i\beta} \prod_{j=1}^{S-i} G(j,a) q_{S,0}, \quad for \;\; i = 0,1,\dots S-1 \\ q_{i,1} &= \frac{\lambda + S\beta}{\mu} \sum_{j=1}^{S-i} \prod_{k=i+1}^{S-j} \frac{\lambda + k\beta}{\mu} \prod_{r=1}^{j-1} G(r,a) q_{S,0}, \quad for \;\; i = 0,1,\dots S-1 \end{aligned}$$

where
(i) $\prod_{k=j}^{y} \frac{\lambda+k\beta}{\mu} = 1 \quad$ if $y < j$
(ii) $\prod_{r=1}^{k} G(r,a) = 1 \quad$ if $k < r$
(iii) $q_{S,0}$ is determined such that the total probability is one.

6.2.2 Performance measures

(i) Expected inventory level is given by

$$\begin{aligned} Sq_{S,0} + \sum_{i=1}^{S-1} iq_{i,0} + \sum_{i=0}^{S-1} iq_{i,1} = \Big\{ S + \sum_{i=1}^{S-1} i \frac{\lambda + S\beta}{\lambda + i\beta} \prod_{j=1}^{S-i} G(j,a) \\ + \sum_{i=0}^{S-1} i \frac{\lambda + S\beta}{\mu} \Big[\sum_{j=1}^{S-i} \prod_{k=i+1}^{S-j} \frac{\lambda + k\beta}{\mu} \prod_{r=1}^{j-1} G(r,a) \Big] \Big\} q_{S,0}. \end{aligned}$$

(ii) Now we derive the distribution of the time during which the production is continuously on. To derive the distribution consider the Markov chain on the state space

$$E = \{(0,1), (1,1), (2,1), \cdots, (S-1,1), (S,0)\},$$

with $(S,0)$ as absorbing state. Then the distribution of the duration of time the production is continuously on is of phase-type with distribution function $F(x)$ given by

$$F(x) = 1 - \underline{\alpha}\exp(Tx)\underline{e}; \quad x \geq 0,$$

where T is a non-singular matrix obtained by deleting the last row and last column of the infinitesimal generator matrix of the Markov chain on E, $\underline{e} = (1,1,\ldots,1)$ is a row vector of order S with all entries equal to one and $\underline{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{S-1})$ is the initial probability vector of order S with

$$\alpha_{S-i} = \frac{e^{-\mu D}(\mu D)^{i-1}}{(i-1)!}; \quad i = 1, 2, \ldots, S-2.$$

$$\alpha_1 = 1 - \sum_{i=1}^{S-2} \frac{e^{-\mu D}(\mu D)^{i-1}}{(i-1)!}$$

and $\alpha_0 = 0$.

The expected length of time the production is continuously on is given by $(-1)\underline{\alpha}T^{-1}\underline{e}$. Inversion of T may often be difficult. So we use a recursive method to find the expected value.

Let $T_{i,1}$ = time to reach $(i+1,1)$ from $(i,1)$ for $i = 0, 1, \ldots, S-2$ and $T_{S-1,1}$ = time to reach $(S,0)$ from $(S-1,1)$
$E[T_{i,1}] = \frac{1}{\mu} + \frac{\lambda+i\beta}{\mu}E[T_{i-1,1}]; i = 1, 2, \ldots, S-1$
$E[T_{0,1}] = \frac{1}{\mu}$; since once the process is in state $(0,1)$, it is not possible to move further to the left. Then we get

$$E[T_{i,1}] = \frac{1}{\mu} + \frac{1}{\mu}\sum_{j=1}^{i}\prod_{k=j}^{i}\frac{\lambda + k\beta}{\mu}; \quad i = 0, 1, \ldots S-1.$$

Now the expected time $E(B)$ the production is continuously on is given by

$$E(B) = E[T_{S-1,1}]e^{-\mu D} + \{E[T_{S-2,1}] + E[T_{S-2,1}]\}(\mu D)e^{-\mu D} + \cdots$$
$$+ \{E[T_{0,1}] + E[T_{1,1}] + \cdots + E[T_{S-1,1}]\}[1 - \sum_{j=0}^{S-2} \frac{(\mu D)^j e^{-\mu D}}{j!}]$$
$$= \sum_{t=1}^{S-1}[\frac{1}{\mu} + \frac{1}{\mu}\sum_{j=1}^{S-t-1}\prod_{k=j}^{S-t-1}\frac{\lambda + k\beta}{\mu}]G(t,a).$$

(iii) The average amount of time the production is continuously off is given by

$$E(I) = \sum_{i=1}^{S-1}\frac{i}{\lambda + (S-i+1)\beta}G(i,a) + \frac{S}{\lambda+\beta}[1 - \sum_{i=1}^{S-1}G(i,a)].$$

The expected cycle length (*i.e.*, the expected time for $(S,0)$ to $(S,0)$) is $E(B) + E(I)$.

(iv) The expected number of visits to state $(0,1)$ in a cycle is $\frac{q_{0,1}}{q_{S,0}}$. Therefore the expected amount of time the system is empty in a cycle is

$$\frac{1}{\mu}\frac{q_{0,1}}{q_{S,0}} = \frac{1}{\mu}[\frac{\lambda + S\beta}{\mu}\sum_{j=1}^{S}\prod_{k=1}^{S-j}\frac{\lambda + k\beta}{\mu}\prod_{r=1}^{j-1}G(r,a)]$$

6.3 Control problem

We derive the optimal value of the control variable D to get the maximum profit. Let K be the profit per unit time by way of the switched off production and C be the loss per unit time when the system is empty. Then the expected profit per cycle is

$$E(P) = KE(I) = K\{\sum_{i=1}^{S-1}\frac{i}{\lambda + (S-i+1)\beta}G(i,a) + \frac{S}{\lambda+\beta}[1 - \sum_{i=1}^{S-1}G(i,a)]\}.$$

Expected loss per cycle is

$$\begin{aligned} E(L) &= C \times \text{ Expected amount of time the system is empty} \\ &= C[\frac{\lambda + S\beta}{\mu} \sum_{j=1}^{S} \prod_{k=1}^{S-j} \frac{\lambda + k\beta}{\mu} \prod_{r=1}^{j-1} G(r,a)]. \end{aligned}$$

Expected net profit $= E(P) - E(L)$.

6.4 Special case

Here we consider D-policy for a production inventory system without perishability (*i.e.*, $\beta = 0$).
We have limiting probabilities as

$$\begin{aligned} q_{i,0} &= \prod_{j=1}^{S-i} G(j,a) q_{S,0}; \quad i = 1, 2, \ldots, S-1 \\ q_{i,1} &= \Big[\sum_{j=1}^{S-i} (\frac{\lambda}{\mu})^j \prod_{k=1}^{S-i} G(k,a)\Big] q_{S,0}; \quad i = 0, 1, \ldots, S-1 \end{aligned}$$

where $q_{S,0}$ is determined such that total probability is one.
Expected inventory level is given by

$$\Big[S + \sum_{i=1}^{S-1} i \prod_{j=1}^{S-i} G(j,a) + \sum_{i=0}^{S-1} i \{\sum_{j=1}^{S-i} (\frac{\lambda}{\mu})^j \prod_{k=1}^{S-j} G(k,a)\}\Big] q_{S,0}.$$

6.5 Numerical illustration

For the values of the parameters $S = 10$, $\lambda = 1$, $\beta = 0.05$, $\mu = 2$, expected profit per cycle, expected loss per cycle and net profit were obtained and are tabulated below.

Table 6.1: $S = 10$, $\lambda = 1$, $\beta = 0.05$, $\mu = 2$, $K = 10$, $C = 10$

D	expected profit per cycle	Expected loss per cycle	Net profit
10	36.74523	24.81855	11.92668
10.1	36.74818	24.8214	11.92678
10.2	36.75079	24.82395	11.92685
10.3	36.7531	24.82621	11.92689
10.4	36.75513	24.82823	11.9269*
10.5	36.75692	24.83002	11.92688
10.6	36.75849	24.83162	11.92687
10.7	36.75987	24.83303	11.92684
10.8	36.76109	24.83429	11.9268
10.9	36.76216	24.83541	11.92675
11	36.7631	24.83639	11.92671
11.1	36.76392	24.83727	11.92666
11.2	36.76465	24.83804	11.92661
11.3	36.76528	24.83873	11.92655
11.4	36.76583	24.83933	11.92651
11.5	36.76632	24.83986	11.92645
11.6	36.76674	24.84034	11.92641
11.7	36.76711	24.84075	11.92637
11.8	36.76744	24.84112	11.92632
11.9	36.76772	24.84144	11.92629

We get an optimal D value 10.4 with the corresponding net profit 11.9269.

6.6 Concluding remarks

In this thesis we have discussed inventory models with negligible and/ positive service time. In the second chapter we discussed N-policy for an (s, S) inventory system with positive service time. In the third chapter we discussed the problem of effective utilization of idle time of a server in an (s, S) inventory system with positive and/ negligible service time. In the fourth chapter we discussed the transfer of customers along with inventory, if available, in two parallel service facilities following (s, S) policy with positive service time. In all these chapters we consider exponentially distributed service time. But in the fifth chapter we introduce arbitrarily distributed service time. Here we consider $M|G|1$ type inventory system. In the sixth chapter we consider a production inventory with items having random lifetimes with negligible service time. Here we introduce a control policy called D-policy. All these problems are of high importance in any industrial related management problems and production process.

Bibliography

[1] T. Altiok. (R, r) production/inventory systems. *Operations Research*, 37(2):266–276, 1989.

[2] G. Arivarignan, C. Elango, and N. Arumugam. A continuous review perishable inventory control system at service facilities. In J. R. Artalejo and A. Krishnamoorthy, editors, *Advances in Stochastic Modelling*. Notable Publications, 2002.

[3] K. Arrow, S. Karlin, and H. Scarf. *Studies in Applied Probability and Management Science*. Standford University Press, Standford, 1958.

[4] M. Berge, M. J. M. Posner, and H. Zhao. Production inventory system with unreliable machines. *Operations Research*, 42(1):111–118, 1994.

[5] O. Berman and E. Kim. *Stochastic inventory management at service facilities with non-instantaneous order replenishment*. PhD thesis, Joseph L.Rotman School of management, University of Toronto, 1999. Working Paper.

[6] O. Berman and E. Kim. Stochastic models for inventory management at service facilities. *Comm. Stat. –Stochastic models*, 15:695–718, 1999.

[7] O. Berman and E. Kim. Dynamic order replenishment policy in internet- based supply chains. *Mathematical models of operations research*, 53:371–390, 2001.

[8] O. Berman, E. Kim, and D. G. Shimshack. Deterministic approximations for inventory management at service facilities. *IEEE Transactions*, 25(5):98–104, 1993.

[9] O. Berman and K. P. Sapna. *Inventory management at service facilities with positive lead time*. PhD thesis, Joseph L. Rotman School of management, University of Toronto, 1999. Working paper.

[10] O. Berman and K. P. Sapna. Inventory management at service facilities for systems arbitrarily distributed service times. *Comm. Stat.–Stochastic Models*, 16(384):343–360, 2000.

[11] O. Berman and K. P. Sapna. Optimal service rates of a service facility with perishable inventory items. *Naval Research Logistics*, 49:464–482, 2002.

[12] D. J. Buchanan and R. F. Love. A (Q, R) inventory model with lost sales and erlang distributed lead times. *Naval Research Logistics Quarterly*, 32:605–611, 1985.

[13] F. M. Ceng and S. P. Sethi. Optimality of state dependent (s, s) policies in inventory models with markov-modulated demand and lost sales. *Production and Operations Management*, 8:183–192, 1999.

[14] A. G. de Kok. Approximations for a lost sales production/inventory control model with service level constrains. *Management Sci.*, 31(6):729–737, 1985.

[15] A. G. de Kok, H. C. Tijims, and F. A. van der Duyn Schouten. Approximations for the single product production-inventory problem with compound poisson demand and service level constrains. *Adv. Appl. Prob.*, 16:378–401, 1984.

[16] T. G. Deepak, V. C. Joshua, and A. Krishnamoorthy. Queues with postponed work. *TOP*, 12(3–4), 2004.

[17] B. T. Doshi, F. A. van der Duyn Schouten, and A. J. J. Talman. A production-inventory model with a mixture of backorders and lost sales. *Mgmt. Sci.*, 24:1078–1086, 1978.

[18] B. Garish and S. C. Graves. A one-product production-inventory problem under continuous review policy. *Opns. Res.*, 28:1228–1236, 1980.

[19] D. Gross and C. M. Harris. *Fundamentals of Queueing Theory*. John Wiley and Sons, Inc., New York, third edition, 1998.

[20] D. Gross and C. M. Harris. *Fundamentals of Queueing Theory*. John Wiley and Sons (Asia), third edition, 2002.

[21] G. Hadley and Whitin. *Analysis of inventory systems*. Prentice Hall Inc., Englewood Cliffs, New Jersey, 1963.

[22] F. Harris. *Operations and Costs*. Factory Management. A. W. Shah Co., Chicago, 1915.

[23] Qi-Ming He and M. F. Neuts. Two $M/M/1$ queues with transfer of customers. *Queueing systems – Theory and Applications*, 42:377–400, 2002.

[24] L. Hsu and C. S. Tapiero. Maintenance of an unreliable $M|G|1$ queueing job shop. *Queue. Syst.*, 2:333–350, 1987.

[25] S. Kalpakam and G. Arivarignan. Analysis of an exhibit inventory system. *Stochastic Analysis and Applications*, 3(4):447–466, 1985.

[26] S. Kalpakam and G. Arivarignan. A continuous review inventory system with arbitrary inter arrival times between demands and an exhibit item subject to random failure. *Opsearch*, 22:153–168, 1985.

[27] S. Kalpakam and G. Arivarignan. On exhibiting inventory system with Erlangian life times under renewal demand. *Ann. Inst. Stat. Math.*, 41(3):601–616, 1989.

[28] S. Kalpakam and K. P. Sapna. A control policy of an inventory system with compound Poisson demand. *Stochastic Analysis and Applications*, 11(4):459–482, 1993.

[29] S. Kalpakam and K. P. Sapna. Continuous review (s, S) inventory system with random life time and positive lead times. *Operations Research Letters*, 16:115–119, 1994.

[30] A. V. Kazimirsky. Analysis of BMAP/G/1 queue with reservation of service. *Stochastic Analysis and Applications*, 24:703–718, 2006.

[31] A. Krishnamoorthy, T. G. Deepak, Viswanath C. Narayanan, and B. Krishnakumar. On an (s, S) inventory policy with service time, vacation to server and correlated lead time. *Jour. of Quality Tech. and Quality Mgmt.*, 2007.

[32] A. Krishnamoorthy and Mohammad Ekramol Islam. Production inventory with retrial of customers in an (s, S) policy. To appear in Stochastic Modelling and Application, 2004.

[33] A. Krishnamoorthy and K. P. Jose. An (s, S) inventory system with positive lead-time loss and retrial of customers. *Stochastic Modelling, Analysis and Applications*, 2005.

[34] A. Krishnamoorthy and K. P. Jose. Comparison of inventory systems with service, positive lead-time, loss and retrial of customers. *JAMSA*, 2006.

[35] A. Krishnamoorthy, T. P. Madhusoodanan, and M. Manoharan. On the (s, S) inventory policy with independent non-identically distributed interarrival demand times and random lead times. *Cahiers du C. E. R. O.*, 29(3,4):239–248, 1987.

[36] A. Krishnamoorthy and N. Raju. n-policy for a (S, s) inventory system with random life times. *The Korean Journal of Computational and Applied Mathematics*, 1997.

[37] A. Krishnamoorthy and N. Raju. n-policy for the production inventory system with random life times. *Ricerca Operativa*, 31(97):37–46, 1998.

[38] A. Krishnamoorthy and P. V. Ushakumari. Reliability of k-out-of-n system with repair and retrial of failed units. *TOP*, 7(2):293–304, 1999.

[39] A. Krishnamoorthy, P. V. Ushakumari, and B. Lakshmi. k-out-of-n system with the repair: the N-policy. *Asia Pacific Journal of Operations Research*, 19:47–61, 2002.

[40] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modelling*. ASA-SIAM: Philadelphia, 1999.

[41] L. Liu. (s, S) continuous review models for inventory with random life times. *Operations Research Letters*, 9(3):161–167, 1990.

[42] D. Mitra. Stochastic theory of a fluid model of multiple failure-susceptible producers and consumers coupled by a buffer. *Adv. Appl. Porb.*, 26:646–676, 1988.

[43] E. Naddor. *Inventory systems*. John Wiley & Sons, New York, 1966.

[44] M. F. Neuts. A versatile Markovian point process. *J. Appl. Probab.*, 16:764–779, 1979.

[45] M. F. Neuts. *Matrix-geometric solutions in stochastic models-An algorithmic approach*. John-Hopkins Univ. Press, 1981. Also published by Dover, 2002.

[46] M. F. Neuts. *Structured Stochastic Matrices of M/G/1 type and Their Applications*. Marcel Dekker, New York, 1989.

[47] M. F. Neuts. *Matrix-geometric solutions in stochastic models: An algorithmic approach*. Dover Publications, Inc., New York, second edition, 1994.

[48] P. R. Parthasarathy and V. Vijayalakshmi. Transient analysis of an $(S-1, S)$ inventory model-a numerical approach. *Intern. J. Computer Math.*, 59:177–185, 1996.

[49] F. R. Richards. Comments on the distribution of inventory positions in a continuous review (s, S) inventory systems. *Operations Research*, 23:366–371, 1975.

[50] F. R. Richards. A continuous review (s, S) inventory systems in a random environment. *Journal of Applied Probability*, 15:654–659, 1978.

[51] J. I. Ryshikov. *Lagerhattung*. Akademic Verlag, Berlin, 1973.

[52] I. Sahin. On the stationary analysis of continuous review (s, S) inventory systems with constant lead times. *Operations Research*, 27:719–729, 1979.

[53] I. Sahin. On the continuous review (s, S) inventory model under compound renewal demand and random lead times. *J. Appl. Prob.*, 20:213–219, 1983.

[54] M. Schwarz, C. Sauer, H. Daduna, R. Kulik, and R. Szekli. $M|M|1$ queueing system with inventory. *Queueing Systems: Theory and Applications*, 54(1):55–78, 2006.

[55] M. Sharafali. On a continuous review production-inventory problem. *Operations Research Letters*, 3:199–204, 1984.

[56] K. Sigman and O. Simchi-Levi. Light traffic heuristic for an $M|G|1$ queue with limited inventory. *Annals of Operations Research*, 40:371–380, 1992.

[57] B. D. Sivalzlian. A continuous review (s, S) inventory system with arbitrary interarrival distribution between unit demands. *Oper. Res.*, 22:65–71, 1974.

[58] R. Sivasamy and R. Elangovan. A queueing system with a machining input source operating under an N-policy. *Advances in Stochastic modelling*, pages 367–376, 2002.

[59] S. K. Srinivasan. General analysis of (s, S) inventory systems with random lead times and unit demands. *J. Math. Phy. Sci.*, 13:107–129, 1979.

[60] V. Thangaraj and R. Ramanarayanan. An operating policy in inventory systems with random lead times and unit demands. *Maths. Operations forsch U. Statist. Ser. Optim.*, 14:111–124, 1983.

[61] P. V. Ushakumari and A. Krishnamoorthy. k-out-of-n system with repair: the $\max(N, T)$ policy. *Performance Evaluation*, 57:221–234, 2004.

[62] A. F. Veinott and H. M. Wagner. Computing optimal (s, S) inventory policies. *Management Science*, 11:525–552, 1965.

[63] A. F. Veinott Jr. The status of mathematical inventory theory. *Management Science*, 12:745–777, 1966.

[64] William G. Vendemia, B. Eddy Patuwo, and Ming S. Hung. Evaluation of lead time in production/inventory systems with non stationary stochastic demand. *Journal of Operational Research Society*, 46:221–233, 1995.

[65] H. Wagner. *Principles of Operations Research*. Prentice Hall, New Jersey, 2 edition, 1975.