ISSN 0005-1179, Automation and Remote Control, 2018, Vol. 79, No. 9, pp. 1593–1608. © Pleiades Publishing, Ltd., 2018.
Original Russian Text © D. Shajin, B. Benny, R.V. Razumchik, A. Krishnamoorthy, 2018, published in Avtomatika i Telemekhanika, 2018, No. 9, pp. 59–78.

Discrete Product Inventory Control System with Positive Service Time and Two Operation Modes

D. Shajin[*,a], **B. Benny**[*,b], **R.V. Razumchik**[**,***,c], **and A. Krishnamoorthy**[*,d]

[*] *Department of Mathematics, Cochin University of Science and Technology, Cochin, India*
[**] *Institute of Informatics Problems, FRC CSC RAS, Moscow, Russia*
[***] *Peoples' Friendship University of Russia (RUDN University), Moscow, Russia*
e-mail: [a]*dhanya.shajin@gmail.com,* [b]*binithabennya@gmail.com,*
[c]*rrazumchik@ipiran.ru,* [d]*achyuthacusat@gmail.com*

Received September 28, 2016

Abstract—Consideration is given to the Markov inventory control system of a discrete product of maximum volume S under the strategies (s, Q) and (s, S) with a positive service time. Upon arrival, the order is queued if the inventory level is positive or, otherwise, leaves the system unserviced. One server handles the queued orders one-by-one in the sequence of their arrival. If the inventory level exceeds s, then the service time has the exponential distribution of intensity μ; otherwise, of intensity $\alpha\mu$, $0 < \alpha \leqslant 1$. The product in the inventory is consumed only at the instant when the service (of the order) ends. Inventory deficit is not allowed. When the inventory is empty new orders are not admitted into the system, and the service process of the queued orders (if any) is stopped. The lead time is assumed to be exponentially distributed. Analytical relations are established for the basic stationary performance characteristics of the system.

Keywords: inventory control system, positive service time, multiplicative representation

DOI: 10.1134/S0005117918090059

1. INTRODUCTION

The present paper is devoted to the study of the basic stationary characteristics of a new stochastic model of control of inventory of a discrete product with the so-called positive service time. The questions posed and resolved in the paper refer mostly to the mathematical theory of inventory control. Active research is being carried out today in this field as is witnessed by the regularly appearing publications in the leading scientific periodicals.

In the classical inventory control systems, the arriving orders are either executed (serviced) immediately or, if the system admits product deficit, are queued and executed immediately upon arrival of the current inventory. If some time is required to fulfill the order, then a queue is formed from the arriving orders. Then, the inventory is not used in a trice at the instant of order arrival, but (instantly) at the end of service (of the order). Understandably, during the time of order service new orders that can arrive to the system are accepted and queued. If at the instant when the service commences the current order the inventory turns out to be equal to zero, then the process of service is aborted because the serviced order needs a product unit, and during its service time the inventory may not be replenished. The systems of inventory control where such rules of order servicing is possible are known as the inventory control systems with positive service time.

Among the systems with positive service time, those admitting multiplicative representation for joint stationary distribution of the inventory volume and the general number of queued and in-service orders are of greatest interest. This allows one to cardinally simplify the calculation of the control performance indices. As it turned out, existence of the multiplicative representation depends on the accepted rules for reception of the arriving and servicing of the available orders during the no-inventory intervals in the system.

If the inventory replenishment does not occur instantaneously, it may happen that the inventory has been used and there are orders waiting for execution. Thus, the server idles and the time of order execution increases. It seems that the inventory control systems with positive service time were first considered in [1, 2] where analysis was carried out under the assumption of Poisson arriving flow and arbitrarily distributed times of servicing. The $M/M/1$ system with the exponentially distributed time between the order placement and the instant of inventory replenishment and the three replenishment strategies (s, Q), (s, S) and a random[1] one is considered in [3]. It turns out that if all orders arriving in default of inventory are rejected by the system, then in all three cases the joint stationary distribution of the inventory volume and number of orders in queue and on the server is represented in the multiplicative form. This result was generalized in [4] to the case of arbitrarily distributed time before replenishment. A variant of the queuing network where each unlimited-capacity system has a finite product inventory is considered in [5]. The order for inventory replenishment is placed when the volume of inventory falls down to a predefined level. The time to the inventory replenishment is random and may be dependent on the system load. If in the system, the inventory volume becomes zero, and then it does not receive new order and redirects them to any of the other systems with the positive product inventory, if any, otherwise, it is lost. In such case the joint stationary distribution is also representable in the multiplicative form.

Much attention was paid to the inventory control systems with positive service time during the two last decades. The reader can familiarise himself with a review of works up to 2011 in [6]. Among the latest publications on this subject matter we mention [7–11]. A joint stationary distribution of the volume of inventory and the numbers of queued orders and those on the server (not in the multiplicative form) was obtained in [7] for the exponential finite-capacity system for the case where the servicing intensity is played anew for each new order according to the discrete (two-point) distribution. A finite-capacity system with the general product lifetime, standby, and possible cancellation of an order (see also [9]) was considered in [8]. For a simplest order arrival flow and on condition that the new orders are lost if the product inventory is zero, the joint stationary distribution has a multiplicative form. Such representation is also possible if the time before replenishment order is other than the exponential time, in the system there is an additional flow of repeated orders (see [10]) also in the case of the replenishment strategy (r, Q) (see [11]).

Analysis of the results obtained, generally speaking, suggests the following expected conclusion: the multiplicative representation of the stationary distribution can be obtained only in rare cases. A system with positive service time operating in the random medium was considered in [12]. The following rule of system–medium interaction was examined. The discrete set of states was decomposed into two disjoint subsets of blocking and reception. As long as the medium is in a state from the reception subset, the system operates as an ordinary system with positive service time unlimited inventory. When the medium is in a state from the subset of blockings, the process of servicing is stopped and the new arrivals are lost. In this case, one can establish the necessary and sufficient condition for representability of the joint stationary distribution in the multiplicative

[1] The strategy (s, S) implies that each time as the inventory volume drops to s and an order is placed increasing the inventory volume up to S; however, if the inventory volume exceeds s, the order is not placed. The strategy (s, Q) operates in a similar way, but the order is placed for $Q = S - s$ units of product. Under the random strategy, added to the inventory is a random amount of product units, but such that the total volume does not exceed S.

form. It deserves noting that if in the states from the blocking set permitted is reception of new orders or servicing of the orders of the in-system orders, the multiplicative representation is already impossible. It turns out that in the systems with positive service time "correct" definition of the blocking set leads multiplicative representation for the stationary distribution. For the systems $M/M/1$ such blocking set is given by $\{(n, 0), n \geqslant 0\}$, where n is the total number of queued orders and those in the server and the second coordinate means that the volume of inventory is zero. However, as is shown in [12], this set can be extended by adding to it the states where the arriving order is rejected if the product inventory is positive.

The present paper that was written in continuation of [13–15] considers the system $M/M/1$ with positive service time and two modes of server operation, ordinary where the volume of inventory is above the threshold value and slow where the volume of inventory is below the threshold value. It is assumed that under the zero volume of inventory the new orders are rejected and servicing of the queued orders is suspended (see, for example, [16]). In what follows, a multiplicative representation for the joint stationary distribution under the strategies (s, Q) and (s, S) will be established for this system. Since under zero inventory the orders are lost, introduction of two operational modes seems to be a possible variant of improving the control performance. By varying the rate of service under low volume of inventory enables one can reduce the part of lost orders, but at that the time of order sojourn in the system grows. Having set the aim of minimizing some cost functional, the present authors also formulate and solve numerically the problem of determining the optimal level of rate reduction.

The next section gives a detailed description of the system under the strategy (s, Q). Sections 3 and 4 are devoted to finding a joint stationary distribution under the strategies (s, Q) and (s, S) and some indices of the control performance under the strategy (s, Q). Numerical results are given in Section 5. The main results are formulated in conclusion together with the possible lines of future studies.

2. SYSTEM DESCRIPTION

Consider a system of controlling a single-product inventory with positive service time. The product volume is measured in discrete units, the maximal volume of inventory is fixed at S. The arriving Poissonian flow of orders has the parameter λ. Each arriving order is admitted to the system, that is, the order is satisfied if the current volume of inventory is positive. At that, the product stock is not consumed at accepting the order to the system. Instead, the order is queued. The system has one server handling one-by-one the queued orders according to the FCFS discipline. At the instant of completing servicing of the order, one product unit is consumed. As soon as the volume of product inventory falls down to $s > 0$, an order for product replenishment by $Q = S - s$ is placed. The times of fulfilling the orders for product replenishment are independent of the processes of order arrivals and servicing and have an exponential distribution with the parameter β. Inventory deficit is not permitted, that is, if at the instant of finishing servicing an order the inventory volume on the server is zero, then until the instant of product inventory replenishment servicing of other queued orders and acception of the arriving orders is stopped. Reordering for replenishment prior to the delivery of the preceding order is also inadmissible. The time of order servicing is exponentially distributed, but the value of the distribution parameter depends on the current volume of inventory. If the volume of product inventory does not exceed s, then the servicing intensity is equal to μ_2, otherwise, it is equal to $\mu_1 = \alpha\mu_2$, where $0 < \alpha \leqslant 1$.

Under the assumptions made about the arrival flow, the process of servicing, and inventory replenishment, operation of the system under consideration can be described by the Markov process $\{\Omega(t) = (N(t), I(t)),\ t \geqslant 0\}$, where $N(t)$ is the number of orders on the queue and server at the instant t, and $I(t)$ is the volume of inventory at the instant t. The set of states of the process

$\Omega(t)$ is $\mathcal{X} = \{(n,i); n \geqslant 0, 0 \leqslant i \leqslant S\}$. The infinitesimal transition matrix $\mathcal{Q}$ of the process $\Omega(t)$ is a block tridiagonal matrix

$$\mathcal{Q} = \begin{pmatrix} A_{00} & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{1}$$

where the square matrices A_{00}, A_0, A_1, and A_2 of size $(S+1)$ are given by

$$(A_{00})_{ij} = \begin{cases} \beta, & j = i + Q,\ 1 \leqslant i \leqslant s+1 \\ -\beta, & j = i,\ i = 1 \\ -(\lambda+\beta), & j = i,\ 2 \leqslant i \leqslant s+1 \\ -\lambda, & j = i,\ s+2 \leqslant i \leqslant S+1 \\ 0, & \text{otherwise,} \end{cases}$$

$$(A_0)_{ij} = \begin{cases} \lambda, & j = i,\ 2 \leqslant i \leqslant S+1 \\ 0, & \text{otherwise,} \end{cases}$$

$$(A_1)_{ij} = \begin{cases} \beta, & j = i + Q,\ 1 \leqslant i \leqslant s+1 \\ -\beta, & j = i,\ i = 1 \\ -(\lambda+\beta+\mu_1), & j = i,\ 2 \leqslant i \leqslant s+1 \\ -(\lambda+\mu_2), & j = i,\ s+2 \leqslant i \leqslant S+1 \\ 0, & \text{otherwise,} \end{cases}$$

$$(A_2)_{ij} = \begin{cases} \mu_1, & j = i-1,\ 2 \leqslant i \leqslant s+1 \\ \mu_2, & j = i-1,\ s+2 \leqslant i \leqslant S+1 \\ 0, & \text{unless otherwise stated.} \end{cases}$$

It follows from the structure of the matrix $\mathcal{Q}$ that the process $\Omega(t)$ is a quasi birth-and-death process (QBD).

3. JOINT STATIONARY DISTRIBUTION OF THE INVENTORY VOLUME AND THE NUMBERS OF QUEUED AND IN-SERVER ORDERS

Denote by $p_{n,i}$, $n \geqslant 0$, $0 \leqslant i \leqslant S$, the stationary probability that the process $\Omega(t)$ is in the state (n,i), that is, queued, and on the server there are n orders, and the volume of inventory is i, that is,

$$p_{n,i} = \lim_{t\to\infty} \mathbf{P}\{N(t) = n, I(t) = i\},\ n \geqslant 0,\ 0 \leqslant i \leqslant S.$$

The stationary probabilities $p_{n,i}$ exist under a certain constraint which will be discussed in what follows (see the inequality (5)). Introduce a vector $\vec{p} = (\vec{p}_0, \vec{p}_1, \dots)$ with the coordinates $\vec{p}_n = (p_{n,0}, p_{n,1}, \dots, p_{n,S})$, $n \geqslant 0$. As is known from the general theory of QBD (see [17]), the stationary distribution $\vec{p}$ exists if and only if

$$\vec{\pi} A_0 \vec{1} < \vec{\pi} A_2 \vec{1}, \tag{2}$$

where $\vec{\pi} = (\pi_0, \pi_1, \dots, \pi_S)$ is a unique positive solution of the equation system

$$\vec{\pi}(A_0 + A_1 + A_2) = \vec{0}, \tag{3}$$
$$\vec{\pi}\vec{1} = 1. \tag{4}$$

If the matrix $A = A_0 + A_1 + A_2$ is irreducible, then the vector $\vec{\pi}$ satisfying Eqs. (3) and (4) can be selected. In the case under consideration, the matrix A is given by

$$A = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ s \\ s+1 \\ \vdots \\ Q \\ Q+1 \\ \vdots \\ S \end{array} \begin{array}{c} \begin{array}{ccccccccccc} 0 & 1 & \dots & s & s+1 & \dots & Q & Q+1 & \dots & S \end{array} \\ \begin{pmatrix} -\beta & & & & & & \beta & & & \\ \mu_1 & -(\beta+\mu_1) & & & & & & \beta & & \\ & \ddots & \ddots & & & & & & \ddots & \\ & & \mu_1 & -(\beta+\mu_1) & & & & & & \beta \\ & & & \mu_2 & -\mu_2 & & & & & \\ & & & & \ddots & \ddots & & & & \\ & & & & & \mu_2 & -\mu_2 & & & \\ & & & & & & \mu_2 & -\mu_2 & & \\ & & & & & & & \ddots & \ddots & \\ & & & & & & & & \mu_2 & -\mu_2 \end{pmatrix} \end{array}$$

and is the infinitesimal matrix of some Markov process $\{\tilde{I}(t), t \geqslant 0\}$ with finite set of states. As can be seen from the structure of the matrix A, the process $\{\tilde{I}(t), t \geqslant 0\}$ coincides in substance with the process $\{I(t), t \geqslant 0\}$ of variation of the volume of inventory in the considered system with the exception of the fact that there are no transitions related with product consumption (with the intensity λ). The set of its states is $\{0, 1, \dots, S\}$, and all states communicate with each other. Therefore, the process $\{\tilde{I}(t),\ t \geqslant 0\}$ is irreducible, and the matrix A is indecomposable. Using the form of the matrices A_1, A_2, and A_3, one can make sure that the components of the vector $\vec{\pi}$ satisfying the system (3) and (4) are calculated from the formulas

$$\pi_i = \begin{cases} \dfrac{\beta}{\mu_1}\left(\dfrac{\beta+\mu_1}{\mu_1}\right)^{i-1}\pi_0, & 1 \leqslant i \leqslant s \\ \dfrac{\beta}{\mu_1}\left(\dfrac{\beta+\mu_1}{\mu_1}\right)^{s-1}\left(\dfrac{\beta+\mu_1}{\mu_2}\right)\pi_0, & s+1 \leqslant i \leqslant Q \\ \dfrac{\beta}{\mu_2}\left(\dfrac{\beta+\mu_1}{\mu_1}\right)^{i-(Q+1)}\left(\left(\dfrac{\beta+\mu_1}{\mu_1}\right)^{S-(i-1)} - 1\right)\pi_0, & Q+1 \leqslant i \leqslant S, \end{cases}$$

where

$$\pi_0 = \left(1 + \frac{\mu_2 - \mu_1}{\mu_2}\left(\left(\frac{\beta+\mu_1}{\mu_1}\right)^s - 1\right) + Q\frac{\beta}{\mu_2}\left(\frac{\beta+\mu_1}{\mu_1}\right)^s\right)^{-1}.$$

Having noticed now that $\vec{\pi}A_0\vec{1} = \lambda\sum_{i=1}^{S}\pi_i = \lambda(1-\pi_0)$ and

$$\vec{\pi}A_2\vec{1} = \mu_1\sum_{i=1}^{s}\pi_i + \mu_2\sum_{i=s+1}^{S}\pi_i = Q\beta\left(\frac{\beta+\mu_1}{\mu_1}\right)^s\pi_0,$$

and using then the inequality (2) and the explicit form of the probability π_0, after a series of simple transformations we establish that the stationary distribution $\vec{p}$ of the process $\Omega(t)$ exists if and only if

$$\lambda < \frac{Q\beta\left(\frac{\beta+\mu_1}{\mu_1}\right)^s}{\frac{\mu_2-\mu_1}{\mu_2}\left(\left(\frac{\beta+\mu_1}{\mu_1}\right)^s - 1\right) + Q\frac{\beta}{\mu_2}\left(\frac{\beta+\mu_1}{\mu_1}\right)^s}. \tag{5}$$

This implies that as follows from (5) for $\mu_1 = \mu_2$ the condition $\lambda < \mu_1$ is necessary and sufficient for existence of the stationary distribution, which coincides with the stationarity criterion for an ordinary queuing system $M/M/1/\infty$.

We demonstrate that the joint stationary distribution $p_{n,i}$ is representable in the multiplicative form. For that, we first dwell on determining the stationary distribution of the volume of inventory, provided that at the server the orders are processed instantaneously. Let $\hat{I}(t)$ be the volume of inventory at the instant t. The set of states of the process $\{\hat{I}(t),\ t \geqslant 0\}$ is $\{0, 1, \ldots, S\}$. One can show similar to what was done for the process $\{\tilde{I}(t),\ t \geqslant 0\}$ that the stationary distribution

$$r_i = \lim_{t\to\infty} \mathbf{P}\{\hat{I}(t) = i\}, \quad 0 \leqslant i \leqslant S,$$

of the process $\{\hat{I}(t),\ t \geqslant 0\}$ does exist. Without putting down the matrix of transition intensities, we immediately put down the global balance equations:

$$\begin{aligned} -\beta r_0 + \lambda r_1 &= 0, \\ -(\beta+\lambda) r_i + \lambda r_{i+1} &= 0, \quad 1 \leqslant i \leqslant s, \\ -\lambda r_i + \lambda r_{i+1} &= 0, \quad s+1 \leqslant i \leqslant Q-1, \\ \beta r_{i-Q} - \lambda r_i + \lambda r_{i+1} &= 0, \quad Q \leqslant i \leqslant S-1, \\ \beta r_s - \lambda r_S &= 0. \end{aligned}$$

Solution of this system with regard for the normalization condition $\sum_{i=0}^{S} r_i = 1$ is given by

$$r_i = \begin{cases} \dfrac{\beta}{\lambda}\left(\dfrac{\beta+\lambda}{\lambda}\right)^{i-1} r_0, & 1 \leqslant i \leqslant s \\ \dfrac{\beta}{\lambda}\left(\dfrac{\beta+\lambda}{\lambda}\right)^{s} r_0, & s+1 \leqslant i \leqslant Q \\ \dfrac{\beta}{\lambda}\left(\dfrac{\beta+\lambda}{\lambda}\right)^{s}\left(1 - \left(\dfrac{\beta+\lambda}{\lambda}\right)^{i-(Q+s+1)}\right) r_0, & Q+1 \leqslant i \leqslant S, \end{cases}$$

where $r_0 = \lambda^s / (\lambda^s + Q\beta(\beta+\lambda)^s)$.

By a direct substitution[2] in the equilibrium equations system (EES) $\vec{p}\mathcal{Q} = \vec{0}$ one can make sure that the joint stationary distribution $p_{i,j}$ is given by

$$p_{i,j} = \begin{cases} \dfrac{1}{\vartheta}\mathcal{C}_i(j)\left(\dfrac{\lambda}{\mu_1}\right)^i r_j, & i \geqslant 0,\ 0 \leqslant j \leqslant s \\ \dfrac{1}{\vartheta}\mathcal{C}_i(j)\left(\dfrac{\lambda}{\mu_2}\right)^i r_j, & i \geqslant 0,\ s+1 \leqslant j \leqslant S, \end{cases}$$

[2] We will not dwell in detail on this because of tedious calculations.

where ϑ is the normalizing factor and $\mathcal{C}_i(j)$ are constants defined by

$$\mathcal{C}_i(j) = \begin{cases} \mathcal{C}_0(0), & i = 0,\ 1 \leqslant j \leqslant S \\ \mathcal{C}_0(0), & i = 1,\ 0 \leqslant j \leqslant S \\ \mathcal{C}_0(0), & i \geqslant 2,\ 0 \leqslant j \leqslant s \\ w_i(j)\mathcal{C}_0(0), & i \geqslant 2,\ s+1 \leqslant j \leqslant S, \end{cases}$$

$$w_i(j) = \begin{cases} \left(\dfrac{\mu_2}{\mu_1}\right)^{i-1}, & i \geqslant 2,\ j = s+1 \\ 1, & i = 2,\ s+2 \leqslant j \leqslant Q \\ \mathcal{U}\left(h^s - \dfrac{\mu_2}{\mu_1}\right), & i = 2,\ j = Q+1 \\ \mathcal{V}\left(h^s - h^{j-(Q+2)}\left(\dfrac{\beta\mu_2 + \lambda\mu_1}{\lambda\mu_1}\right)\right), & i = 2,\ Q+2 \leqslant j \leqslant S \\ \dfrac{\mu_2}{\lambda}\left(\dfrac{\lambda+\mu_2}{\mu_2}w_2(j-1) - 1\right), & i = 3,\ s+2 \leqslant j \leqslant Q \\ \dfrac{\mu_2}{\lambda}\left(\dfrac{\lambda+\mu_2}{\mu_2}w_{i-1}(j-1) - w_{i-2}(j-2)\right), & i \geqslant 4,\ s+2 \leqslant j \leqslant Q \\ \mu_2^2\mathcal{U}\left(h^s\left(\dfrac{\lambda+\mu_2}{\lambda\mu_2^2}w_2(Q) - \dfrac{1}{\lambda\mu_2}\right) - \dfrac{1}{\mu_1^2}\right), & i = 3,\ j = Q+1 \\ \mu_2^{i-1}\mathcal{U}\left(h^s\left(\dfrac{\lambda+\mu_2}{\lambda\mu_2^{i-1}}w_{i-1}(Q) - \dfrac{w_{i-2}(Q)}{\lambda\mu_2^{i-2}}\right) - \dfrac{1}{\mu_1^{i-1}}\right), & i \geqslant 4,\ j = Q+1 \\ \dfrac{\mu_2^2}{\lambda}\mathcal{V}\left(h^s\dfrac{\lambda+\mu_2}{\mu_2^2}w_2(j-1) - h^{j-(Q+2)}\left(\dfrac{1}{\mu_2} - \dfrac{\beta}{\mu_1^2}\right)\right), & i = 3,\ Q+2 \leqslant j \leqslant S \\ \dfrac{\mu_2^{i-1}}{\lambda}\mathcal{V}\left(h^s\dfrac{\lambda+\mu_2}{\mu_2^{i-1}}w_2(j-1) - h^{j-(Q+2)}\left(\dfrac{w_{i-2}(j-1)}{\mu_2^{i-2}} - \dfrac{\beta}{\mu_1^{i-1}}\right)\right), & i \geqslant 4,\ Q+2 \leqslant j \leqslant S, \end{cases}$$

and

$$h = \frac{\beta}{\lambda} + 1, \quad \mathcal{U} = (h^s - 1)^{-1}, \quad \mathcal{V} = \left(h^s - h^{j-(Q+1)}\right)^{-1}.$$

Using the normalization condition $\vec{p}\,\vec{1} = 1$, we find an expression for the normalizing constant ϑ:

$$\vartheta = \left\{\frac{\mu_1}{\mu_1 - \lambda}h^s + \frac{\lambda+\mu_2}{\mu_2}\left(h^s\left(Q\frac{\beta}{\lambda} - 1\right) + 1\right)\right. \\ \left. + \frac{\beta}{\lambda}\sum_{i=2}^{\infty}\left(\frac{\lambda}{\mu_2}\right)^i\left(h^s\sum_{j=s+1}^{S}w_i(j) - \sum_{j=Q+1}^{S}h^{j-(Q+1)}w_i(j)\right)\right\}\left(1 + Q\frac{\beta}{\lambda}h^s\right)^{-1}\mathcal{C}_0(0).$$

Being capable to calculate the joint stationary distribution $p_{n,i}$, one can determine every possible stationary indices of control performance, which is the subject matter of the next section.

To conclude this section, we dwell in brief on determination of the joint stationary distribution in the system under consideration under the strategy (s, S). The infinitesimal transition matrix in this case has the same structure as the matrix $\mathcal{Q}$, that is, it is also a block tridiagonal matrix where only the elements A_{00} and A_1 different. In the case at hand, they are given by

$$(A_{00})_{ij} = \begin{cases} \beta, & j = S+1,\ 1 \leqslant i \leqslant s+1 \\ -\beta, & j = i,\ i = 1 \\ -(\lambda+\beta), & j = i,\ 2 \leqslant i \leqslant s+1 \\ -\lambda, & j = i,\ s+2 \leqslant i \leqslant S+1 \\ 0, & \text{otherwise,} \end{cases}$$

$$(A_1)_{ij} = \begin{cases} \beta, & j = S+1,\ 1 \leqslant i \leqslant s+1 \\ -\beta, & j = i,\ i = 1 \\ -(\lambda+\beta+\mu_1), & j = i,\ 2 \leqslant i \leqslant s+1 \\ -(\lambda+\mu_2), & j = i,\ s+2 \leqslant i \leqslant S+1 \\ 0, & \text{unless otherwise stated.} \end{cases}$$

The stationary distribution exists only when

$$(\lambda-\mu_1)\left(\left(\frac{\beta+\mu_1}{\mu_1}\right)^s - 1\right) < (\mu_2-\lambda)(S-s)\frac{\beta}{\mu_2}\left(\frac{\beta+\mu_1}{\mu_1}\right)^s.$$

By denoting by $p^*_{n,i}$, $n \geqslant 0$, $0 \leqslant i \leqslant S$, the stationary probability that on the queue and server there are n orders, and the volume of inventory is i, one can show by substituting into EES that the probabilities $p^*_{n,i}$ are given by

$$p^*_{i,j} = \begin{cases} \mathcal{D}_i(j)\left(\dfrac{\lambda}{\mu_1}\right)^i \phi_j, & 0 \leqslant j \leqslant s \\ \mathcal{D}_i(j)\left(\dfrac{\lambda}{\mu_2}\right)^i \phi_j, & s+1 \leqslant j \leqslant S, \end{cases} \qquad n \geqslant 0,$$

where the constants $\mathcal{D}_i(j)$ and ϕ_j are given by

$$\phi_j = \begin{cases} \dfrac{\beta}{\lambda}\left(\dfrac{\beta+\lambda}{\lambda}\right)^{j-1}\phi_0, & 1 \leqslant j \leqslant s \\ \dfrac{\beta}{\lambda}\left(\dfrac{\beta+\lambda}{\lambda}\right)^{s}\phi_0, & s+1 \leqslant j \leqslant S \\ \left(\dfrac{\lambda}{\beta+\lambda}\right)^s\left(1+(S-s)\dfrac{\beta}{\lambda}\right)^{-1}, & j = 0, \end{cases}$$

$$\mathcal{D}_i(j) = \begin{cases} \mathcal{D}_0(0), & i = 0,\ 1 \leqslant j \leqslant S \\ \mathcal{D}_0(0), & i = 1,\ 0 \leqslant j \leqslant S \\ \mathcal{D}_0(0), & i \geqslant 2,\ 0 \leqslant j \leqslant s \\ \psi_i(j)\mathcal{D}_0(0), & i \geqslant 2,\ s+1 \leqslant j \leqslant S, \end{cases}$$

$$\psi_i(j) = \begin{cases} \left(\dfrac{\mu_2}{\mu_1}\right)^{i-1}, & i \geqslant 2,\ j = s+1 \\ 1, & i \geqslant 2,\ s+i \leqslant j \leqslant S \\ \left(\dfrac{\mu_2}{\mu_1}\right)^{i-3}\dfrac{\mu_2}{\lambda}\left(\left(\dfrac{\lambda+\mu_2}{\mu_2}\right)\dfrac{\mu_2}{\mu_1} - 1\right), & i \geqslant 3,\ j = s+2 \\ \dfrac{\mu_2}{\lambda}\left(\dfrac{\lambda+\mu_2}{\mu_2}\psi_{i-1}(j-1) - \psi_{i-2}(j-1)\right), & i \geqslant 4,\ s+3 \leqslant j \leqslant S. \end{cases}$$

Finally, the value of the constant $\mathcal{D}_0(0)$ is established from the normalization condition:

$$\mathcal{D}_0(0) = \left(1 + (S-s)\frac{\beta}{\lambda}\right)\left(\frac{\mu_1}{\mu_1 - \lambda} + \frac{\beta}{\lambda}\left((S-s)\left(1 + \frac{\lambda}{\mu_2}\right) + \sum_{i=2}^{\infty}\sum_{j=s+1}^{S}\left(\frac{\lambda}{\mu_2}\right)^i \psi_i(j)\right)\right)^{-1}.$$

4. SOME STATIONARY CONTROL PERFORMANCE INDICES UNDER THE STRATEGY (s, Q)

4.1. Time of Sojourn in the System

We first calculate the stationary distribution $W(x)$ of the time of sojourn in the system of an arbitrary order in terms of the Laplace–Stieltjes transform (LST) $\tilde{W}(z) = \int_0^\infty e^{-zx} dW(x)$. For that purpose we find the LST $\tilde{W}_{ij}(z) = \int_0^\infty e^{-zx} dW_{ij}(x)$ of the stationary distribution $W_{ij}(x)$ of the time of sojourn in the system of an order finding j orders on the queue and in the server, the volume of inventory being equal to i. Since under the Poisson flow the probability that in the stationary mode an order arriving to the system finds it in the state (j, i) is equal to $p_{j,i}$, the LST $\tilde{W}(z)$ is established by averaging $\tilde{W}_{ij}(z)$ over the distribution $p_{j,i}$:

$$\tilde{W}(z) = \sum_{j=0}^{\infty}\sum_{i=1}^{S} p_{j,i}\tilde{W}_{ij}(z).$$

The system sojourn time of an order consists of two addends, the waiting time for servicing and the time proper of servicing. Having noticed that the servicing intensity depends on the inventory volume and that also the processes of inventory replenishment, arrival of new order and their servicing are independent and have exponential distribution, one can easily establish the following recurrent formulas for determination of the LST $\tilde{W}_{ij}(z)$:

$$\tilde{W}_{i0}(z) = \begin{cases} \dfrac{\beta}{\alpha\mu_2 + \beta + z}\dfrac{\mu_2}{\mu_2 + z} + \dfrac{\alpha\mu_2}{\alpha\mu_2 + \beta + z}, & 1 \leqslant i \leqslant s \\ \dfrac{\mu_2}{\mu_2 + z}, & s+1 \leqslant i \leqslant S, \end{cases}$$

$$\tilde{W}_{ij}(z) = \begin{cases} \dfrac{\beta}{\alpha\mu_2 + \beta + z}\dfrac{\mu_2}{\mu_2 + z}\tilde{W}_{Q,j-1}(z) + \dfrac{\alpha\mu_2}{\alpha\mu_2 + \beta + z}\dfrac{\beta}{\beta + z}\tilde{W}_{Q,j-1}(z), & i = 1,\ j \geqslant 1 \\ \dfrac{\beta}{\alpha\mu_2 + \beta + z}\dfrac{\mu_2}{\mu_2 + z}\tilde{W}_{Q+i-1,j-1}(z) + \dfrac{\alpha\mu_2}{\alpha\mu_2 + \beta + z}\tilde{W}_{i-1,j-1}(z), & 2 \leqslant i \leqslant s,\ j \geqslant 1 \\ \dfrac{\mu_2}{\mu_2 + z}\tilde{W}_{i-1,j-1}(z), & s+1 \leqslant i \leqslant S,\ j \geqslant 1. \end{cases}$$

$\tilde{W}_{i,j}(z)$ should be calculated beginning from $j = 0$ successively over all i. Using the established expressions for $\tilde{W}_{ij}(z)$ and $\tilde{W}(z)$, one can calculate recurrently the instants of stationary distribution of the time of sojourn in the system. For example, the stationary mean time of order sojourn $w = -\tilde{W}'(0)$ in the system is given by

$$w = \sum_{j=0}^{\infty}\sum_{i=1}^{S} p_{j,i}\tilde{W}_{ij}, \tag{6}$$

where the values of $\tilde{W}_{ij}$ are calculated from

$$\tilde{W}_{i0} = \begin{cases} \dfrac{1}{\mu_2+\beta} + \dfrac{\beta}{\alpha\mu_2+\beta}\dfrac{1}{\mu_2}, & 1 \leqslant i \leqslant s \\ \dfrac{1}{\mu_2}, & s+1 \leqslant i \leqslant S, \end{cases}$$

$$\tilde{W}_{ij} = \begin{cases} \dfrac{\beta}{\alpha\mu_2+\beta}\dfrac{1}{\mu_2} + \dfrac{\alpha\mu_2}{\alpha\mu_2+\beta}\dfrac{1}{\beta} + \dfrac{1}{\alpha\mu_2+\beta} + \tilde{W}_{Q,j-1}, & i=1,\ j \geqslant 1 \\ \dfrac{1}{\alpha\mu_2+\beta} + \dfrac{\beta}{\alpha\mu_2+\beta}\left(\dfrac{1}{\mu_2} + \tilde{W}_{Q+i-1,j-1}\right) + \dfrac{\alpha\mu_2}{\alpha\mu_2+\beta}\left(\dfrac{1}{\beta} + \tilde{W}_{i-1,j-1}\right), & 2 \leqslant i \leqslant s,\ j \geqslant 1 \\ \dfrac{1}{\mu_2} + \tilde{W}_{i-1,j-1}, & s+1 \leqslant i \leqslant S,\ j \geqslant 1. \end{cases}$$

The calculations of $\tilde{W}_{i,j}$ should by carried out successively for all i beginning from $j = 0$. The values of w can be established from other considerations, namely from the Little formula.[3] Let $\mathbf{E}[N]$ be the mean number of the queued orders and those in the server, and λ_A be the mean number of orders admitted to the system in time unit, that is,

$$\mathbf{E}[N] = \sum_{n=1}^{\infty}\sum_{i=0}^{S} np_{n,i}, \quad \lambda_A = \lambda\sum_{n=0}^{\infty}\sum_{i=1}^{S} p_{n,i}.$$

Then, the mean time of order sojourn in the system is equal to $w = \mathbf{E}[N]/\lambda_A$.

Let us write out formulas to calculate some other indices of the control performance. The mean number $\mathbf{E}[N, I \leqslant i]$ of orders in the system where the volume of inventory does not exceed i is equal to

$$\mathbf{E}[N, I \leqslant i] = \sum_{n=1}^{\infty}\sum_{j=0}^{i} np_{n,j}.$$

Correspondingly, the mean number $\mathbf{E}[N, I > i]$ of orders in system where the volume of inventory exceeds i is equal to $\mathbf{E}[N, I > i] = \mathbf{E}[N] - \mathbf{E}[N, I \leqslant i]$. The intensity β^* of entering orders for replenishment of inventory and the mean number of orders λ_L lost in a time unit because of lack of product obey the formulas

$$\beta^* = \mu_2\sum_{n=1}^{\infty} p_{n,s+1}, \quad \lambda_L = \lambda - \lambda_A = \lambda\sum_{n=0}^{\infty} p_{n,0}.$$

The mean number $\mathbf{E}[N, I > 0]$ of queued orders and those on server when the product inventory is positive is given by

$$\mathbf{E}[N, I > 0] = \sum_{n=1}^{\infty}\sum_{i=1}^{S} np_{n,i},$$

and the mean number $\mathbf{E}[N, I = 0]$ of orders queued and in server where the product inventory is zero is equal to $\mathbf{E}[N, I = 0] = \mathbf{E}[N] - \mathbf{E}[N, I > 0]$. The mean volume $\mathbf{E}[I]$ of product inventory is given by

$$\mathbf{E}[I] = \sum_{n=0}^{\infty}\sum_{i=1}^{S} ip_{n,i}.$$

[3] Its validity for the given system can be verified by using, for example, Theorem 6.1 of [18].

Finally, the mean number of replenishments of the product inventory in a unit time is

$$\beta \sum_{n=0}^{\infty} \sum_{i=0}^{s} p_{n,i}.$$

4.2. Selection of the Value of α

We recall that according to the assumptions made at the instant when the product inventory volume becomes equal to s an order for inventory replenishment is placed immediately. The time of replenishment is distributed exponentially with the parameter β. If at the instant of vanishing inventory, that is, becoming zero before the instant of its replenishment, the arriving orders will be lost for some time. Since as long as the volume of inventory does not exceed s, the order service intensity μ_2 dops to $\alpha\mu_2$, reduction in the value of α reduces also the rate of inventory consumption. Various criteria can be suggested to select α. It seems only natural to select, for example, an α such that the probability of losing at least one order is minimal. The present subsection proposes a method to calculate this probability. We notice that since $0 < \alpha \leqslant 1$, the probability of losing one order is the smaller, the smaller α. Therefore, that value of α that is closest to zero and technically possible is the quasi-optimal for the selected criterion. However, with reduction of α the mean time of sojourn of the w order in the system grows. Therefore, the proposed criterion is not conceptual. The situation can be corrected by introducing in the criterion an additional upper constraint on w.

Now we proceed to calculating the probability of losing at least one order. Assume that at an arbitrary time instant τ an order is placed for inventory replenishment. Denote by ξ the time before the inventory replenishment and by η, the time required to service s orders. Introduce the following probabilities:

$$a_{ij} = \mathbf{P}\{\xi < \eta | I(\tau) = i, N(\tau) = j\}, \quad 1 \leqslant i \leqslant S, \quad j \geqslant 0.$$

Then, provided that at the instant of preparation of an application for inventory the number of orders queued and staying on the server was j, $j \geqslant 0$, the probability that the instant of inventory replenishment comes prior to that of inventory consumption is equal to a_{sj}. To calculate the probabilities a_{sj}, $j \geqslant 0$, it is necessary to consider apart two cases of $j \geqslant s$ and $0 \leqslant j < s$.

Let $j \geqslant s$, that is, the number of queued orders and those on the server at the instant of sending order for inventory replenishment is greater than or equal to s. Obviously, $a_{s,s} = a_{s,s+1} = a_{s,s+2} = \ldots.$ Consider the Markov random process $\{X(t),\ t \geqslant 0\}$ of the set of states $\{0 \leqslant i \leqslant s-1\} \bigcup \{r\} \bigcup \{e\}$. The state $X(t) = i$, $0 \leqslant i \leqslant s-1$ implies that i orders were serviced, the state $X(t) = r$ implies that inventory replenishment took place, and the state $X(t) = e$ implies that serviced were exactly s orders. Then, the matrix of transition intensities of the process $\{X(t), t \geqslant 0\}$ is representable as

$$\begin{pmatrix} -(\alpha\mu_2+\beta) & \alpha\mu_2 & 0 & & 0 & \beta \\ 0 & -(\alpha\mu_2+\beta) & \alpha\mu_2 & & 0 & \beta \\ & & \ddots & \ddots & & \vdots \\ & & & -(\alpha\mu_2+\beta) & \alpha\mu_2 & \beta \\ 0 & 0 & \ldots & \ldots & 0 & 0 \\ 0 & 0 & \ldots & \ldots & 0 & 0 \end{pmatrix} = \begin{pmatrix} T & \vec{t}^e & \vec{t}^r \\ \mathbf{0} & \vec{0} & \vec{0} \end{pmatrix},$$

with the $s \times s$ matrix T. Let at the initial instant the process be at the state 0, that is, $X(0) = 0$. Then, by introducing the vector of initial states $\vec{\gamma} = (1, 0, \ldots, 0)$ of dimension s, the probability $a_{s,s}$

can be calculated using the formula $-\gamma T^{-1}\vec{t}^r$ and after some simple transformations reduced to the form

$$a_{s,s} = \frac{\beta}{\alpha\mu_2+\beta}\sum_{n=0}^{s-1}\left(\frac{\alpha\mu_2}{\alpha\mu_2+\beta}\right)^n.$$

The last formula has a transparent probabilistic sense. One can readily demonstrate that the rest of the probabilities $a_{i,i}$, $1 \leqslant i \leqslant s$ are given by

$$a_{i,i} = \frac{\beta}{\alpha\mu_2+\beta}\sum_{n=0}^{i-1}\left(\frac{\alpha\mu_2}{\alpha\mu_2+\beta}\right)^n, \quad 1 \leqslant i \leqslant s. \tag{7}$$

Obviously, the probability that the instant of completion of serving s orders comes before the instant of inventory replenishment is given by

$$1 - a_{s,s} = -\gamma T^{-1}\vec{t}^e = \left(\frac{\alpha\mu_2}{\alpha\mu_2+\beta}\right)^s.$$

As the result, the probability $\pi_{\geqslant s}$ that at least one order is lost if at the instant of placing an order for inventory replenishment there were $j \geqslant s$ orders for inventory on the queue and server

$$\pi_{\geqslant s} = \left(\frac{\alpha\mu_2}{\alpha\mu_2+\beta}\right)^s \frac{\lambda}{\lambda+\beta}.$$

Let now $0 \leqslant j \leqslant s-1$, that is, at the instant of order placement for inventory replenishment there were on queue and server less than s orders. Calculation of the probabilities a_{sj} is started from the case $j = s-1$. Using the formula of total probability, we can write down

$$a_{s,s-1} = \frac{\beta}{\lambda+\mu_2\alpha+\beta} + \frac{\lambda}{\lambda+\mu_2\alpha+\beta}a_{s,s} + \frac{\mu_2\alpha}{\lambda+\mu_2\alpha+\beta}a_{s-1,s-2}, \tag{8}$$

$$a_{s-1,s-2} = \frac{\beta}{\lambda+\mu_2\alpha+\beta} + \frac{\lambda}{\lambda+\mu_2\alpha+\beta}a_{s-1,s-1} + \frac{\mu_2\alpha}{\lambda+\mu_2\alpha+\beta}a_{s-2,s-3}, \tag{9}$$

$$\cdots$$

$$a_{2,1} = \frac{\beta}{\lambda+\mu_2\alpha+\beta} + \frac{\lambda}{\lambda+\mu_2\alpha+\beta}a_{2,2} + \frac{\mu_2\alpha}{\lambda+\mu_2\alpha+\beta}a_{1,0}, \tag{10}$$

$$a_{1,0} = \frac{\beta}{\lambda+\beta} + \frac{\lambda}{\lambda+\beta}a_{1,1}. \tag{11}$$

We notice that expressions for the probabilities $a_{i,i}$, $1 \leqslant i \leqslant s$ were established previously (see (7)). The system of linear algebraic Eqs. (8)–(11) is solved recurrently. By denoting

$$d = \mu_2\alpha/(\lambda+\mu_2\alpha+\beta),$$

solution is representable as

$$a_{k+1,k} = \frac{1}{\lambda+\mu_2\alpha+\beta}\sum_{i=2}^{k+1}(\lambda a_{i,i}+\beta)\,d^{k+1-i} + \frac{\lambda a_{1,1}+\beta}{\lambda+\beta}d^k, \quad 0 \leqslant k \leqslant s-1.$$

Therefore, the probability $a_{s,s-1}$ that the instant of inventory replenishment comes before the instant of consuming the inventory, provided that at the instant of preparation of the order for inventory replenishment the number of orders queued and on server was $(s-1)$, is equal to

$$a_{s,s-1} = \frac{1}{\lambda+\mu_2\alpha+\beta}\sum_{i=2}^{s}(\lambda a_{i,i}+\beta)\,d^{s-i} + \frac{\lambda a_{1,1}+\beta}{\lambda+\beta}d^{s-1}.$$

Table 1. Probabilities of losing at least one order for $(S, s, \lambda, \mu_2, \beta) = (15, 7, 3, 10, 4)$ and various values of α

α	$\pi_{\geqslant s}$	π_0	π_4
1	0.0474	0.0007	0.0149
0.9	0.0381	0.0006	0.0128
0.8	0.0293	0.0005	0.0105
0.75	0.0251	0.0005	0.0094
0.7	0.0211	0.0004	0.0082
0.6	0.0140	0.0003	0.0059
0.5	0.0082	0.0002	0.0038
0.4	0.0039	0.0001	0.0020
0.3	0.0013	0.0001	0.0008
0.25	0.0006	0	0.0004
0.2	0.0002	0	0.0001
0.1	0	0	0

The rest of the probabilities $a_{s,s-2}, a_{s,s-3}, \ldots, a_{s,0}$ can be obtained in a similar manner. The total formula for calculation of $a_{s,j}$, $1 \leqslant j \leqslant s-1$, is given by

$$a_{k+i,k} = \frac{1}{\lambda + \mu_2\alpha + \beta} \sum_{n=2}^{k+1} (\lambda a_{n-1+i,n} + \beta)\, d^{k+1-n} + \frac{\lambda a_{i,1} + \beta}{\lambda + \beta} d^k, \quad 0 \leqslant k \leqslant s-i, \quad 1 \leqslant i \leqslant s.$$

Calculation of $a_{k+i,k}$ is carried out beginning from $i = 1$ sequentially in k, $0 \leqslant k \leqslant s-i$.

The probability π_j that at least one order is lost if at the instant of placing the order for inventory replenishment on queue and server there were $0 \leqslant j \leqslant s-1$ orders is given by

$$\pi_j = (1 - a_{s,j}) \frac{\lambda}{\lambda + \beta}.$$

Table 1 shows the results of calculating the probabilities π_0, π_4 and $\pi_{\geqslant s}$ for different values of α under fixed initial parameters. As was expected, the more orders are on queue and server at the instant of placing the order for inventory replenishment, the higher the probability of losing at least one order. At that, with lower rate of servicing, the loss probability decreases and tends to zero.

5. NUMERICAL RESULTS

This section presents some results of numerical calculations of the basic indices of control performance under the strategies (s, Q) and (s, S), as well as the numerical analysis of the problem of optimization of a possible variant of cost functional under the strategy (s, Q). Table 2 shows the

Table 2. Values of the mean number of orders on queue and server $\mathbf{E}[N]$, mean volume $\mathbf{E}[I]$ of inventory, and the intensity λ_L of losing orders under various values of the intensities of arrival of new orders λ and $(\alpha, \beta, \mu_2, s, Q, S) = (0.1; 3; 15; 7; 8; 15)$

λ	Strategy (s, Q)			Strategy (s, S)		
	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L
4	0.6604	10.3905	9.6477×10^{-5}	0.6116	9.9907	2.6392×10^{-4}
5	1.0709	10.1786	2.3277×10^{-4}	0.9620	9.7931	5.8555×10^{-4}
6	1.7307	9.9831	5.2460×10^{-4}	1.4938	9.6162	0.0011
7	2.8822	9.8002	0.0013	2.3567	9.4570	0.0020
8	5.0723	9.6336	0.0028	3.8746	9.3161	0.0038
9	9.5550	9.5006	0.0050	6.7654	9.1996	0.0064

Table 3. Values of the mean number $\mathbf{E}[N]$ of orders on queue and server, mean volume of inventory $\mathbf{E}[I]$, and intensity λ_L of lost orders under various values of service intensity μ_2 and $(\alpha, \beta, \lambda, s, Q, S) = (0.1; 3; 6; 7; 8; 15)$

μ_2	Strategy (s, Q)			Strategy (s, S)		
	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L
11	3.3399	10.0785	5.1419×10^{-4}	5.5761	9.5803	0.0017
12	2.6808	10.0497	4.4167×10^{-4}	4.1438	9.5392	**0.0016**
13	2.2516	10.0248	$\mathbf{4.2652}\times10^{-4}$	3.2934	9.5066	0.0017
14	1.9517	10.0028	4.5693×10^{-4}	2.7410	9.4797	0.0018
15	1.7307	9.9831	5.2460×10^{-4}	2.3567	9.4570	0.0020
16	1.5611	9.9654	6.2479×10^{-4}	2.0751	9.4376	0.0024

Table 4. Values of the mean number $\mathbf{E}[N]$ of orders on queue and on server, mean volume $\mathbf{E}[I]$ of inventory, and intensity λ_L of lost orders under various values of intensity of inventory replenishment β and $(\alpha, \lambda, \mu_2, s, Q, S) = (0.1; 5; 10; 7; 8; 15)$

β	Strategy (s, Q)			Strategy (s, S)		
	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L
2	3.7928	9.7141	0.0013	3.0021	9.3840	0.0019
2.5	2.6808	10.0497	3.6806×10^{-4}	2.2896	9.6739	5.4854×10^{-4}
3	2.1565	10.2754	1.2508×10^{-4}	1.9226	9.8722	1.9090×10^{-4}
3.5	1.8633	10.4379	4.9690×10^{-5}	1.7056	10.0170	7.5970×10^{-5}
4	1.6802	10.5608	2.2620×10^{-5}	1.5649	10.1277	3.3840×10^{-5}
4.5	1.5565	10.6573	1.1574×10^{-5}	1.4675	10.2153	1.6615×10^{-5}

Table 5. Values of the mean $\mathbf{E}[N]$ number of orders on queue and server, mean volume $\mathbf{E}[I]$ of inventory, and intensity λ_L lost orders under various values of α and $(\lambda, \beta, \mu_2, s, Q, S) = (4, 1, 7, 7, 8, 15)$

α	Strategy (s, Q)			Strategy (s, S)		
	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L	$\mathbf{E}[N]$	$\mathbf{E}[I]$	λ_L
0.2	11.3777	8.4804	0.0443	6.2383	8.5352	0.0553
0.3	7.2463	8.1600	0.1113	4.1143	8.4738	0.1162
0.4	4.7245	7.9475	0.1817	2.9964	8.4597	0.1710
0.5	3.3075	7.8120	0.2410	2.3688	8.4675	0.2146
0.6	2.5004	7.7229	0.2863	1.9853	8.4838	0.2482
0.7	2.01139	7.6605	0.3201	1.7332	8.5024	0.2741

results of calculations of the mean number of orders queued and on the server, mean size of inventory, and intensity of order losses under various values of the arriving flow.

Just as we expected, with increased intensity of order arrival λ one observes a faster growth of the mean number of orders on queue and server under the strategy (s, Q) than under the strategy (s, S). At that, with increase in λ the mean size of inventory reduces.

Table 3 shows the results of calculations of the same characteristics but at variation only of the service intensity μ_2. The data indicate that the intensity of losses of orders is always lower for the strategy (s, Q). The same regularity is observed if instead of the service intensity μ_2 varied is the intensity of inventory replenishment β (see Table 4). Notably, this regularity is not retained with variation of α (see Table 5): the order loss intensity under greater values of λ turns out to be smaller under the strategy (s, S) than the strategy (s, Q).

Table 6. Values of the functional $\mathcal{F}(\alpha)$ under different α

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
$\mathcal{F}(\alpha)$	154.4116	151.2455	150.1005	149.4784	149.0754	148.7881	148.5706	**148.3992**	148.5803	148.6452

Consider the cost functional as the characteristic of control performance under the strategy (s, Q). For that, we first make a guessmate about the nature of costs related with system operation. Let in the system at hand the following kinds of costs are possible:

C, costs arising at storing one order per time unit;

C_0, nonrecurring costs related with ordering an inventory replenishment;

C_1, costs of getting one unit of the product;

C_2, costs of storing one product unit per time unit;

C_3, costs arising at irretrievable loss of one order.

Let now in the considered system under the strategy (s, Q) and certain values of the initial parameters one has to select a value of α for which minimized is the functional $\mathcal{F}(\alpha)$ obeying the expression

$$\mathcal{F}(\alpha) = C\left(\frac{\mathbf{E}[N, I \leqslant s]}{\alpha} + \mathbf{E}[N, I > s]\right) + (C_0 + QC_1)\beta^* + C_2\mathbf{E}[I] + C_3\lambda_L.$$

It seems impossible to solve this problem analytically. However, by using the results obtained in the preceeding sections one can determine numerically quasi-optimal values of α. Table 6 compiles the values of the functional $\mathcal{F}(\alpha)$ under various values of α and $(\lambda, \beta, \mu_2, s, Q, C, C_0, C_1, C_2, C_3) = (2, 3, 5, 5, 15, 10, 500, 25, 2, 20)$. By analyzing the results obtained, one can conclude that for the given values the function has at least one extremum (minimum). Consequently, use of the strategy (s, Q) with two modes of server operation may be justified depending on the functional $\mathcal{F}(\alpha)$ because it enables one to improve control performance by varying α.

6. CONCLUSIONS

The main distinction of the considered exponential control system of inventory with positive service time from the existing systems lies in the assumption that the server can operate in two modes, at normal speed and reduced by α^{-1}-fold times. The current mode depends on the volume of inventory. Under the strategies (s, Q) and (s, S), the joint stationary distribution of the inventory volume and the number of orders on system is represented in the multiplicative form. If some cost functional is used in the system as control, then variation of α may improve the performance of control.

Despite the fact that analysis of the basic stationary system characteristics was given in detail, the question of finding interesting characteristics such as the stationary distribution of the occupation period, joint stationary distribution of the occupation period and the number of serviced/lost orders was left aside. This problem, as well as that of generalizing the results obtained to the case of multiserver systems and processes controlled by the Markov chain seems to be an interesting subject for future research.

ACKNOWLEDGMENTS

The work of D. Shajin and A. Krishnamoorthy was supported by the State of Kerala Council for Science, Technology and Environment, no. 001/KESS/2013/CSTE. The work of B. Benny was supported by the program of development of the Department and Commission for University Grants at the Government of India, grant no. F.FIP/12th Plan/KLKE003TF05.

REFERENCES

1. Sigman, K. and Simchi-Levi, D., Light Traffic Heuristic for an M/G/1 Queue with Limited Inventory *Ann. Oper. Res.* 1992, vol. 40, pp. 371–380.
2. Melikov, A.A. and Molchanov, A.A., Stock Optimization in Transport/Storage, *Cybernet. Syst. Anal.*, 1992, vol. 28, no. 3, pp. 484–487.
3. Schwarz, M., Sauer, C., Daduna, H., Kulik, R., and Szekli, R., M/M/1 Queueing Systems with Inventory, *Queueing Syst. Theory Appl.*, 2006, vol. 54, pp. 55–78.
4. Saffari, M., Asmussen, S., and Haji, R., The M/M/1 Queue with Inventory, Lost Sale and General Lead Times, *Queueing Syst. Theory Appl.*, 2013, vol. 75, pp. 65–77.
5. Schwarz, M., Wichelhaus, C., and Daduna, H., Product Form Models for Queueing Networks with an Inventory, *Stochast. Models*, 2007, vol. 23, no. 4, pp. 627–663.
6. Krishnamoorthy, A., Lakshmy, B., and Manikandan, R., A Survey on Inventory Models with Positive Service Time, *OPSEARCH*, 2011, vol. 48, no. 2, pp. 153–169.
7. Anbazhagan, N., Vigneshwaran, B., and Jeganathan, K., Stochastic Inventory System with Two Types of Services, *Int. J. Adv. Appl. Math. Mech.*, 2014, vol. 2, no. 1, pp. 120–127.
8. Krishnamoorthy, A., Benny, B., and Shajin, D., A Revisit to Queueing-Inventory System with Reservation, Cancellation and Common Life Time, *OPSEARCH*, 2017, vol. 54, no. 2, pp. 336–350.
9. Shajin, D., Benny, B., Deepak, T., and Krishnamoorthy, A., A Relook at Queueing Inventory Model with Reservation, Cancellation and Common Life Time, *Commun. Appl. Anal.*, 2016, vol. 20, no. 4, pp. 545–574.
10. Krishnamoorthy, A. and Shajin, D., Stochastic Decomposition in Retrial Queueing Inventory System, in *Proc. 11th Int. Conf. on Queueing Theory and Network Applications*, 2016, pp. 1–4.
11. Baek, J.W. and Moon, S.K., The M/M/1 Queue with a Production Inventory System and Lost Sales, *Appl. Math. Comput.*, 2014, vol. 233, pp. 534–544.
12. Krenzler, R. and Daduna, H., Loss Systems in a Random Environment: Steady-State Analysis, *Queueing Syst. Theory Appl.*, 2014, vol. 80, no. 1–2, pp. 127–153.
13. Krishnamoorthy, A., Shajin, D., and Lakshmy, B., Product Form Solution for Some Queueing-Inventory Supply Chain Problem, *OPSEARCH*, 2016, vol. 53, no. 1, pp. 85–102.
14. Krishnamoorthy, A., Manikandan, R., and Shajin, D., Analysis of Multi-Server Queueing-Inventory System, *Adv. Oper. Res.*, 2015, vol. 2015, pp. 1–16.
15. Krishnamoorthy, A., Manikandan, R., and Lakshmy, B., A Revisit to Queueing-Inventory System with Positive Service Time, *Ann. Oper. Res.*, 2013, vol. 233, no. 1, pp. 221–236.
16. Shnurkov, P.V. and Mel'nikov, R.V., Analysis of the Problem of Continuous-Product Inventory Control under Deterministic Lead Time, *Autom. Remote Control*, 2008, vol. 69, no. 10, pp. 1734–1751.
17. Neuts, M.F., *Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach*, Baltimore: The Johns Hopkins Univ. Press, 1981.
18. El-Taha, M. and Stidham, S., *Sample-Path Analysis of Queueing Systems*, New Jersey: Springer, 1999.

This paper was recommended for publication by M.F. Karavai, a member of the Editorial Board