ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2024, Vol. 85, No. 1, pp. 49–63.

STOCHASTIC SYSTEMS

Transient Behavior of a Two-Phase Queuing System with a Limitation on the Total Queue Size

V. M. Vishnevsky[*,a], **K. A. Vytovtov**[*,b], **and E. A. Barabanova**[*,c]

[*] *Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
e-mail: [a]*vishn@inbox.ru,* [b]*vytovtov_konstan@mail.ru,* [c]*elizavetaalexb@yandex.ru*

Received April 16, 2023
Revised November 15, 2023
Accepted December 21, 2023

Abstract—The transient mode of a two-phase queuing system with a Poisson input flow, exponential distribution of service time in each phase, and a limitation on the total buffer size of the two phases is considered. Nonstationary probabilities of system states are found using the Laplace transform. A numerical calculation and analysis of the system performance characteristics in transient mode with parameters corresponding to new-generation optical networks were carried out.

Keywords: two-phase queuing system, transient mode, Laplace transform

DOI: 10.31857/S0005117924010041

1. INTRODUCTION

Multiphase queuing systems (QSs), or so-called tandem networks, are widely used to describe the operation of telecommunication systems, in which the process of processing requests consists of several stages [1]. This class of systems includes, for example, multi-stage switching systems or a network of linear topology base stations. Moreover, the stationary operation mode of such systems is well studied both for the case of Poisson and correlated input flow. Let us note only some recent works on this topic [1–7].

In recent years, in addition to the study of the stationary mode of QSs, the study of the transient mode of their operation has continued. For example, an important problem in designing optical telecommunication networks with high information transfer rates is to study changes in the system performance characteristics over time and estimate the transition time in stationary mode after the system reboot process or a failure of service devices [8]. A similar situation arises when studying new-generation 5G and 6G networks [9]. Due to the relevance of this problem, in recent years the number of works devoted to the study of the transient operating mode of QSs and their non-stationary Markov models has increased [10–19]. One of the first works where such a problem for a two-phase QS with a Poisson input flow, an infinite buffer in the first phase, and a zero buffer in the second phase was considered is the paper of 1967 [20]. In further works, more complex systems are studied, such as systems with phase-type service time [11, 12] and various types of tandem networks [14, 15]. It should be noted that in most of these works, the authors do not provide final expressions that allow analyzing the performance characteristics of the QS in the transient mode, but use ready-made numerical methods of existing software packages. An analysis of the stability of non-stationary Markov processes with continuous time, describing the functioning of the main classes of QSs with non-stationary input flows, including those varying according to a sinusoidal law, was carried out in [10, 13, 18, 19]. In [21–23], the main performance characteristics of single-line and multi-line QSs with Poisson and correlated flows in transient mode are analyzed.

This work studies the non-stationary performance characteristics of a two-phase QS with a limitation on the total buffer size of the two phases. One example of a real system, the model of which is represented by this QS, is a car service station with two stages of service: diagnostics and repair. Cars queued for service at each stage are placed in a common parking lot with a certain number of parking spaces, which determines the limit on the total number of cars simultaneously located at the service station. The stationary mode of this QS was studied in [24]. There is no study of the non-stationary mode of such a system in the world literature, which determines the novelty of this article.

The structure of the article is as follows. Section 3 presents differential equations that describe the functioning of a two-phase QS, for the convenience of writing which new functions are introduced. Section 4 presents an expression for finding the probabilities of states of a two-phase QS, containing an auxiliary matrix, the elements of which are found using the Laplace transform apparatus. Section 5 provides expressions for finding the main performance indicators of a two-phase QS in transient mode. The numerical results of the study are presented in Section 6.

2. STATEMENT OF THE PROBLEM

A two-phase QS with one single-line service device on each phase is considered. The input flow is Poisson with intensity λ, and the time for servicing requests by devices of the first and second phases has an exponential distribution with intensities μ_1 and μ_2, respectively.

After completing the servicing of a request in the first phase, each request moves to the second phase. The number of requests in the first and second phases can take the values $n_1 = \overline{0, N}$, $n_2 = \overline{0, N}$, respectively, where N is the maximum number of requests in the system. In this case, a limitation is imposed on the total buffer size of the two phases of the system, such that $n_1 + n_2 \leqslant N$ at any time. A new request can enter the system only under the condition $n_1 + n_2 < N$ (Fig. 1).

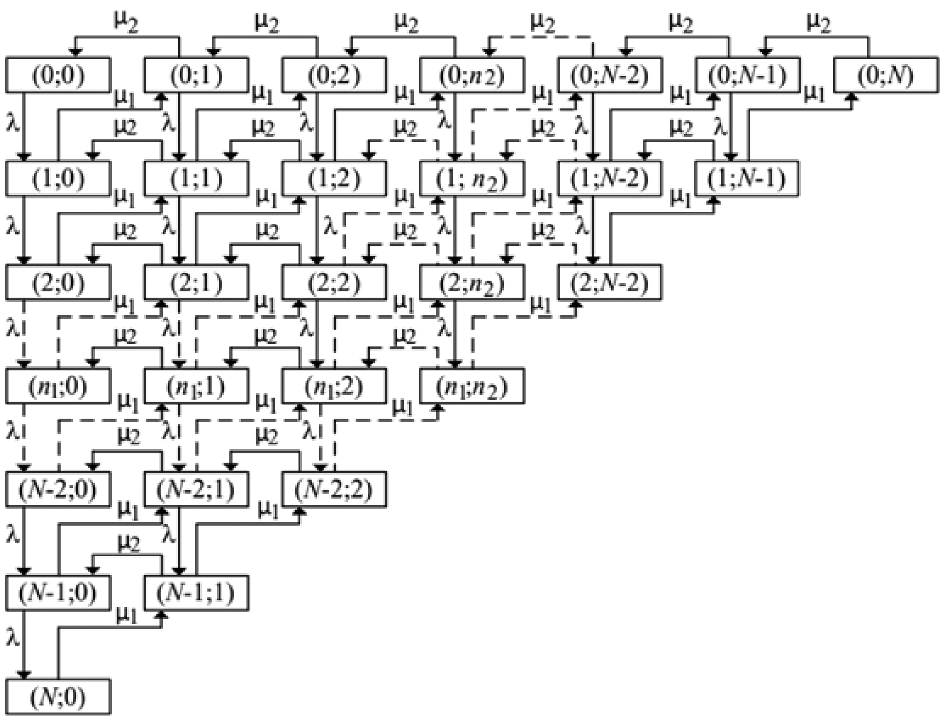

Fig. 1. System state graph.

The purpose of this work is to analyze the performance characteristics of the system described above in transient mode, such as transition time, loss probability, throughput, and the average number of requests in the system.

3. CONSTRUCTION OF DIFFERENTIAL EQUATIONS DESCRIBING THE FUNCTIONING OF A TWO-PHASE QS WITH A LIMITATION ON THE TOTAL BUFFER SIZE

The Markov process describing the operation of the considered QS consists of $R=\frac{1}{2}\left(N^2+3N+2\right)$ states of the system $S\left(n_1,n_2,t\right)$, when n_1 requests are served in the first phase, and n_2 requests are served in the second phase at time t, where $n_1+n_2 \leqslant N$ (Fig. 1). The system of differential equations for such a QS has the form:

$$\begin{gathered}
\frac{dP\left(0,0,t\right)}{dt}=-\lambda P\left(0,0,t\right)+\mu_2 P(0,1,t), (n_1,n_2=0);\\
\frac{dP\left(0,n_2,t\right)}{dt}=-(\lambda+\mu_2)P\left(0,n_2,t\right)+\mu_2 P(0,n_2+1,t)+\mu_1 P(0,n_2-1,t),\\
(n_1=0,n_2=\overline{1,N-1});\\
\frac{dP\left(0,N,t\right)}{dt}=-\mu_2 P\left(0,N,t\right)+\mu_1 P(1,N-1,t), (n_1=0,n_2=N);\\
\frac{dP\left(n_1,0,t\right)}{dt}=-\left(\lambda+\mu_2\right)P\left(n_1,0,t\right)+\mu_2 P(n_1,1,t)+\lambda P\left(n_1-1,0,t\right),\\
(n_1=\overline{1,N-1},n_2=0);\\
\frac{dP\left(N,0,t\right)}{dt}=-\mu_1 P\left(N,0,t\right)+\lambda P(N-1,0,t), (n_1=N,n_2=0);\\
\frac{dP\left(n_1,n_2,t\right)}{dt}=-\left(\lambda+\mu_1+\mu_2\right)P\left(n_1,n_2,t\right)+\mu_2 P(n_1,n_2+1,t)\\
+\mu_1 P(n_1+1,n_2-1,t)+\lambda P(n_1-1,n_2,t), (n_1,n_2>0,n_1+n_2<N);\\
\frac{dP\left(n_1,n_2,t\right)}{dt}=-\left(\mu_1+\mu_2\right)P\left(n_1,n_2,t\right)+\mu_1 P(n_1+1,n_2-1,t)\\
+\lambda P(n_1-1,n_2,t), (n_1,n_2>0,n_1+n_2=N).
\end{gathered}\tag{1}$$

It should be noted that the well-known approach to constructing a system of differential equations, which involves the use of various forms of writing the equations for different permissible values of n_1 and n_2, is very inconvenient for calculating and analyzing the characteristics of the QS in the transient mode. For the convenience of further solution and analysis of system (1), we introduce the functions

$$\upsilon_1\left(x,M\right)=\frac{|x-M+0.5|+x-M+0.5}{2\left|x-M+0.5\right|},\tag{2}$$

$$\upsilon_2\left(x,K\right)=\frac{|K-x-0.5|+K-x-0.5}{2\left|K-x-0.5\right|},\tag{3}$$

where $M=\overline{0,N}$, $K=\overline{0,N}$. Then system (1) can be written in the form

$$\begin{aligned}
\frac{dP\left(n_1,n_2,t\right)}{dt}&=-\left[\lambda\upsilon_2\left(n_1+n_2,N-1\right)+\mu_1\upsilon_1\left(n_1,1\right)+\mu_2\upsilon_1\left(n_2,1\right)\right]P\left(n_1,n_2,t\right)\\
&+\mu_1\upsilon_1\left(n_2,1\right)\upsilon_2\left(n_1+n_2,N\right)P\left(n_1+1,n_2-1,t\right)+\mu_2\upsilon_2\left(n_1+n_2,N-1\right)\\
&\times P\left(n_1,n_2+1,t\right)+\lambda\upsilon_1\left(n_1,1\right)\upsilon_2\left(n_1+n_2,N\right)P\left(n_1-1,n_2,t\right),
\end{aligned}\tag{4}$$

where $n_1 = \overline{0, N}$, $n_2 = \overline{0, N}$, $n_1 + n_2 \leqslant N$. The described system can be represented in matrix form:

$$\frac{d\vec{P}(t)}{dt} = \mathbf{A}\vec{P}(t), \tag{5}$$

where $\mathbf{A}$ is the matrix of coefficients of the system of differential Eqs. (4), $\vec{P}(t) = \{P(n_1, n_2, t)\}^{\mathrm{T}}$ is the column vector of system state probabilities. To construct the matrix $\mathbf{A}$, we additionally introduce the function

$$\vartheta(n_k, n_l) = (N+1)\, n_k + n_l - \frac{n_k(n_k - 1)}{2} + 1, \tag{6}$$

transforming the number of requests n_k, n_l in the first and second buffer, respectively, into the number of a column or row of this matrix. A brief description of functions (2), (3) and (6) is given in the Appendix. Then the elements of the matrix of system (5), located on the main diagonal, are written in the form

$$A_{\vartheta(n_1,n_2),\vartheta(n_1,n_2)} = -\left[\lambda \upsilon_2(n_1 + n_2, N) + \mu_1 \upsilon_1(n_1, 1) + \mu_2 \upsilon_1(n_2, 1)\right]. \tag{7}$$

The remaining non-zero elements are determined by the relations

$$\begin{aligned} &A_{\vartheta(n_1,n_2),\vartheta(n_3,n_4)} = \mu_1 \upsilon_1(n_2, 1)\, \upsilon_2(n_1 + n_2, N+1); \\ &A_{\vartheta(n_1,n_2),\vartheta(n_1,n_5)} = \mu_2 \upsilon_2(n_1 + n_2, N); \\ &A_{\vartheta(n_1,n_2),\vartheta(n_6,n_2)} = \lambda \upsilon_1(n_2, 1)\, \upsilon_2(n_1 + n_2, N+1). \end{aligned} \tag{8}$$

Here $n_1 = \overline{0, N}$, $n_2 = \overline{0, N}$, $n_3 = n_1 + 1$, $n_4 = n_2 - 1$, $n_5 = n_2 + 1$, $n_6 = n_1 - 1$. The remaining elements $A_{i,j}$ of the matrix $\mathbf{A}$ in (5) are equal to zero. The new function (6) is also necessary for the ordered construction of a column vector of system state probabilities at time t in equation (5). Indeed, in terms of n_k and n_l it has the form

$$\begin{aligned} \vec{P}(t) = \{&p(0,0,t), \ldots, p(0,N,t), p(1,0,t), \ldots, \\ &p(1, N-1, t), \ldots\ldots, p(N-1, 0, t), p(N-1, 1, t), p(N, 0, t)\}^{\mathrm{T}}, \end{aligned} \tag{9}$$

where T is the transposition operator. However, to correctly solve (5), it is necessary to use not a two-dimensional array of numbers n_k and n_l when indicating the state of the system, but a sequence number from 1 to $R = \frac{1}{2}(N^2 + 3N + 2)$. To do this, using function (6), we finally obtain

$$\vec{P}(t) = \{P(1,t), P(2,t), P(3,t), P(4,t), \ldots, P(\vartheta(n_k, n_l), t), \ldots, P(\vartheta(N, 0), t)\}^{\mathrm{T}}, \tag{10}$$

where $P(\vartheta(n_k, n_l), t)$ corresponds to $p(n_k, n_l, t)$ in (9).

Thus, using the functions $\vartheta(n_k, n_l)$, $\upsilon_1(x, M)$, $\upsilon_2(x, K)$ makes it possible to construct a matrix of coefficients in (5) in general form for any number of requests N.

4. STATE PROBABILITIES OF A TWO-PHASE QS IN A TRANSIENT MODE

To connect the system state probabilities at time t with the probabilities of system states at some initial time t_0, we introduce the matrix $\mathbf{L}$, the order of which is one greater than the order of the fundamental matrix of the system of equations (4) and such that

$$\vec{P}(t) = \mathbf{L}(t - t_0)\, \vec{P}(t_0), \tag{11}$$

where $\vec{P}(t) = \{P(\vartheta(n_1, n_2), t)\}^{\mathrm{T}}$ is the vector column of system state probabilities at time t.

Let us apply the direct Laplace transform to the system of equations (5):

$$\int_0^\infty e^{-st}\frac{d\vec{P}\,(t)}{dt}dt = \int_0^\infty e^{-st}\mathbf{A}\vec{P}\,(t)dt. \tag{12}$$

Then the elements of the matrix $\mathbf{L}\,(t-t_0)$ are determined by the following theorem.

Theorem 1. *The elements of the matrix $\mathbf{L}\,(t-t_0)$ of a two-phase QS with a limitation on the total buffer size of two phases N, described by the system of equations (4), have the form*

$$L_{l,j}\,(t-t_0) = \sum_{k=1}^{R}(-1)^{l+j}\,\frac{\Delta_{j,l}(s_k)}{\left.\frac{d\Delta(s)}{ds}\right|_{s=s_k}}\exp\left(s_k\,(t-t_0)\right), \tag{13}$$

where $\Delta(s)$ is the determinant of the matrix $\mathbf{B}=\mathbf{A}-s\mathbf{I}$, $\mathbf{A}$ is the coefficient matrix in (5), $\mathbf{I}$ is the unit diagonal matrix, $s=\alpha+i\beta$ is the independent variable in the complex domain, $i=\sqrt{-1}$, $\Delta_{li}(s)$ is the determinant of the minor element B_{li} of the matrix $\mathbf{B}$, s_k is the kth root of the polynomial $\Delta\,(s)$ in the case when all its roots are simple, $R=(N^2+3N+2)/2$ is the number of roots of the polynomial $\Delta\,(s)$, equal to the number of differential equations in system (4).

Proof. Considering that $\int_0^\infty e^{-st}\left(\frac{d\vec{P}(t)}{dt}\right)dt = s\vec{P}\,(s)-\vec{P}\,(0)$, where $\vec{P}\,(0)$ is the column vector of initial conditions, and also that in this case $\mathbf{A}$ is a constant matrix, let us carry out the transformations

$$s\vec{P}\,(s)-\vec{P}\,(0)=\mathbf{A}\vec{P}\,(s)\ \Rightarrow\ \mathbf{A}\vec{P}\,(s)-s\vec{P}\,(s)=-\vec{P}\,(0)\ \Rightarrow\ (\mathbf{A}-s\mathbf{I})\,\vec{P}\,(s)=-\vec{P}\,(0)\,. \tag{14}$$

As a result, we obtain a system of linear inhomogeneous algebraic equations

$$\mathbf{B}\vec{P}\,(s)=-\vec{P}\,(0) \tag{15}$$

with constant coefficients. Taking into account (4), system (15) can be written in the form

$$\begin{aligned}&-\left[\lambda\upsilon_2\,(n_1+n_2,N-1)+\mu_1\upsilon_1\,(n_1,1)+\mu_2\upsilon_1\,(n_2,1)+s\right]P\,(n_1,n_2,s)\\&+\mu_1\upsilon_1\,(n_2,1)\,\upsilon_2\,(n_1+n_2,N)\,P\,(n_1+1,n_2-1,s)+\mu_2\upsilon_2\,(n_1+n_2,N-1)\,P\,(n_1,n_2+1,s)\\&+\lambda\upsilon_1\,(n_1,1)\,\upsilon_2\,(n_1+n_2,N)\,P\,(n_1-1,n_2,s)=P\,(n_1,n_2,0)\,,\end{aligned} \tag{16}$$

where $n_1=\overline{0,N},\ n_2=\overline{0,N},\ n_1+n_2\leqslant N$. Then, in accordance with (7) and (8), the non-zero elements of the matrix $\mathbf{B}$ are written as

$$\begin{aligned}&B_{\vartheta(n_1,n_2),\vartheta(n_1,n_2)}(s)=-\left[\lambda\upsilon_2\,(n_1+n_2,N)+\mu_1\upsilon_1\,(n_1,1)+\mu_2\upsilon_1\,(n_2,1)+s\right];\\&B_{\vartheta(n_1,n_2),\vartheta(n_3,n_4)}(s)=\mu_1\upsilon_1\,(n_2,1)\,\upsilon_2\,(n_1+n_2,N+1)\,;\\&B_{\vartheta(n_1,n_2),\vartheta(n_1,n_5)}=\mu_2\upsilon_2\,(n_1+n_2,N)\,;\\&B_{\vartheta(n_1,n_2),\vartheta(n_6,n_2)}(s)=\lambda\upsilon_1\,(n_2,1)\,\upsilon_2\,(n_1+n_2,N+1)\,.\end{aligned} \tag{17}$$

To find images of elements of the matrix $\mathbf{L}$ it is necessary to use linearly independent initial conditions. These conditions are:

$$P\,(N_1,N_2,0)=1(N_1=\overline{0,N},N_2=\overline{0,N},N_1{+}N_2\leqslant N);P\,(n_1,n_2,0)=0(n_1\neq N_1,n_2\neq N_2). \tag{18}$$

Solutions to system (16) for $P\,(n_1,n_2,0)=1\ (n_1=n_2=0)$, $P\,(n_1,n_2,0)=0\ (n_1=\overline{1,N}$, $n_2=\overline{1,N}$, $n_1+n_2\leqslant N)$ give the first column of images of the elements of the transformation matrix, the solution to system (15) for $P(0,1,0)=1$ and the rest $P(n_1,n_2,0)=0$ give the second column of images of the elements of the transformation matrix. Similarly, all columns of images of elements of the transformation matrix $\mathbf{L}(s-s_0)$ are found. To obtain an image of the matrix $\mathbf{L}$, it is advisable to use the Cramer method. In accordance with this method, the elements of the matrix, which are linearly independent solutions (16), are fractions of the form

$$L_{l,j}(s-s_0) = (-1)^{l+j}\frac{\Delta_{j,l}(s)}{\Delta(s)}, \tag{19}$$

where $\Delta(s)$ is the determinant of the matrix $\mathbf{B}$, $\Delta_{jl}(s)$ is the minor of the element B_{jl} of the matrix $\mathbf{B}$. Now consider the inverse Laplace transform. First of all, we note that the image of the element of the probability transformation matrix (19) is a proper fraction

$$L_{l,j}(s-s_0) = (-1)^{l+j}\frac{\Delta_{j,l}(s)}{\Delta(s)} = (-1)^{l+j}\frac{a_n s^n + a_{n-1}s^{n-1} + \ldots + a_2 s^2 + a_1 s + a_0}{b_m s^m + b_{m-1}s^{m-1} + \ldots + b_2 s^2 + b_1 s + b_0}. \tag{20}$$

Moreover, $n < m$, since the numerator is the determinant of the algebraic complement of the matrix element whose determinant is in the denominator. Then the fraction in (20) can be factorized

$$L(s) = \frac{\Delta_{j,l}(s)}{\Delta(s)} = A_1\frac{1}{s-s_1} + A_2\frac{1}{s-s_2} + \ldots + A_m\frac{1}{s-s_m} = \sum_{k=1}^{m} A_k\frac{1}{s-s_k}. \tag{21}$$

To find the coefficients A_k, multiply both sides of (21) by $(s-s_1)$ and get

$$L(s) = \frac{\Delta_{j,l}(s)}{\Delta(s)}(s-s_1) = A_1 + (s-s_1)\sum_{k=2}^{m} A_k\frac{1}{s-s_k}. \tag{22}$$

The right-hand side of (22) at $s \to s_1$ is equal to A_1, since $s-s_1 \to 0$. The left side represents the uncertainty $0/0$, since the factor $s-s_1$ is present in both the numerator and the denominator. Let us reveal this uncertainty using L'Hopital's rule and obtain the left-hand side in the form

$$\lim_{x\to x_1}\frac{\Delta_{j,l}(s)}{\Delta(s)}(s-s_1) = \lim_{x\to x_1}\frac{\Delta_{j,l}(s) + (s-s_1)\frac{d\Delta_{j,l}(s)}{ds}}{\frac{d\Delta(s)}{ds}} = \frac{\Delta_{j,l}(s_1)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_1}}. \tag{23}$$

Taking into account (22) and (23), we obtain

$$A_1 = \frac{\Delta_{j,l}(s_1)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_1}}. \tag{24}$$

Similarly, we find the kth coefficient in (22) as

$$A_k = \frac{\Delta_{j,l}(s_k)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_k}}. \tag{25}$$

Thus, expression (21) takes the form

$$L(s) = \sum_{k=1}^{m}\frac{\Delta_{j,l}(s_k)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_k}}\cdot\frac{1}{s-s_k}. \tag{26}$$

Applying the inverse Laplace transform to (26) and carrying out mathematical transformations

$$L(t)=\frac{1}{2\pi i}\int\limits_{\sigma-i\infty}^{\sigma+i\infty}\sum_{k=1}^{m}\frac{\Delta_{j,l}(s_k)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_k}}\frac{exp(st)ds}{s-s_k}=\frac{1}{2\pi i}\sum_{k=1}^{m}\frac{\Delta_{j,l}(s_k)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_k}}\int\limits_{\sigma-i\infty}^{\sigma+i\infty}\frac{exp(st)ds}{s-s_k}$$

$$=\sum_{k=2}^{m}\frac{\Delta_{j,l}(s_k)}{\left[\frac{d\Delta(s)}{ds}\right]|_{s=s_k}}exp(s_k t),\quad (27)$$

we obtain an expression for the original from image (26) in the form (12). The theorem has been proven.

By substituting (27) and expression (11), we can find the probabilities of states of a two-phase QS in the transition mode under given initial conditions. These expressions make it possible to calculate and analyze the performance indicators of the system under consideration at an arbitrary moment of time t in both transient and stationary modes: the time before the system enters stationary mode, the probability of losses, throughput, and the number of requests served in each phase.

5. PERFORMANCE INDICATORS OF A TWO-PHASE QS IN TRANSIENT MODE

5.1. Transition Time

The transition time is the time during which the QS goes into stationary mode. In accordance with [22], the time of the transition mode is determined by the smallest absolute value of the real part of the pole of the state probability images:

$$\tau_{tr}=\frac{k}{\alpha_{\min}}. \quad (28)$$

Here $\forall\alpha_j\in\Gamma$: $(\Gamma=\alpha_j,\alpha_j\geqslant\alpha_{\min}\Longrightarrow\alpha_j=\alpha_{\min})$ and $k>0$, $k\in R$. The value of k is selected based on the formulation of a specific problem. It was shown in [22] that the transition mode can be considered completed when $k=(3\div 5)$.

5.2. Probability of Losses

Since the maximum number of requests in the system is $n_1+n_2=N$, all requests in the states $(i,N-i)$, $i=\overline{0,N}$ will be lost. Considering that the presence of requests in the specified states are independent events, the sum of the probabilities of these states at time t

$$P_{loss}(t)=\sum_{i=0}^{N}P(i,N-i,t)=\sum_{i=0}^{N}P(\vartheta(i,N-i),t) \quad (29)$$

determines the probability of loss of requests.

5.3. Throughput

Since expression (29) determines the resulting probability of requests being lost in the system, it is obvious that requests entering the system in any other states will be serviced. Then the throughput at time t in the transition mode is equal to

$$A(t)=[1-P_{loss}(t)]\lambda=\left[1-\sum_{i=0}^{N}P(i,N-i,t)\right]\lambda=\left[1-\sum_{i=0}^{N}P(\vartheta(i,N-i),t)\right]\lambda. \quad (30)$$

Since the throughput actually represents the intensity of the system servicing the requests received by it, then in the time dt the system services $A(t)\,dt$ requests. Consequently, during the transition mode the number of requests served is equal to

$$Z_{service\ tr} = \int\limits_{t_0}^{t_0+\tau_{tr}} \lambda \left[1 - \sum_{i=0}^{N} P\left(\vartheta\left(i, N-i\right), t\right)\right] dt, \tag{31}$$

and the number of lost requests is

$$Z_{loss\ tr} = \int\limits_{t_0}^{t_0+\tau_{tr}} \lambda \sum_{i=0}^{N} P\left(\vartheta\left(i, N-i\right), t\right) dt. \tag{32}$$

Thus, the sum of (31) and (32) gives the number of requests received during the transition mode $\lambda\tau_{tr}$, which confirms the correctness of the obtained relationships.

5.4. Number of Requests Served at Each Phase in Transition Mode

Let $P\left(n_1, n_2, t\right)$ be the probability of finding n_1 requests in the first phase and n_2 requests in the second phase at time t, then the number of requests in the first phase, provided that the system is in the state $\left(n_1, n_2\right)$, is equal to $n_1 P\left(n_1, n_2, t\right)$. Summing $n_1 P\left(n_1, n_2, t\right)$ over all possible states, we obtain the average number of requests in the first phase at time t as

$$Z_{phase1}(t) = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \left[n_1 P\left(n_1, n_2, t\right)\right] = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \left[n_1 P\left(\vartheta\left(n_1, n_2\right), t\right)\right], \tag{33}$$

where $n_1 + n_2 \leqslant N$. Similarly, the average number of requests in the second phase at time t in the transition mode is equal to

$$Z_{phase2}(t) = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \left[n_2 P\left(n_1, n_2, t\right)\right] = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \left[n_2 P\left(\vartheta\left(n_1, n_2\right), t\right)\right], \tag{34}$$

where $n_1 + n_2 \leqslant N$. Then the average number of requests in the system in the transition mode will be

$$Z(t) = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \left[\left(n_1 + n_2\right) P\left(n_1, n_2, t\right)\right] = \sum_{n_1=0}^{N} \sum_{n_2=0}^{N} \left[\left(n_1 + n_2\right) P\left(\vartheta\left(n_1, n_2\right), t\right)\right], \tag{35}$$

where $n_1 + n_2 \leqslant N$.

6. NUMERICAL STUDY OF A TWO-PHASE QS TRANSIENT MODE

Let us consider the transient mode of a two-phase QS operation, which adequately describes the operation of a switch in an all-optical network. In the presented numerical experiment, the values $\lambda = 8 \cdot 10^6$ packets/s, $\mu_1 = 15 \cdot 10^6$ packets/s, $\mu_2 = 10 \cdot 10^6$ packets/s $(\lambda < \mu_2 < \mu_1)$ correspond to the actual characteristics of modern optical networks [25]. Here n_1 is the number of packets in the first servicing phase, n_2 is the number of packets in the second servicing phase, $N = n_1 + n_2 = 4$ is the maximum number of packets in the system. The small buffer size in this numerical example is determined by the technical limitations of modern optical devices.

To analyze the performance characteristics of the considered QS, first of all, the matrix $\mathbf{A}$ is constructed in accordance with (7) and (8), and then the matrix $\mathbf{B}$ is constructed in accordance

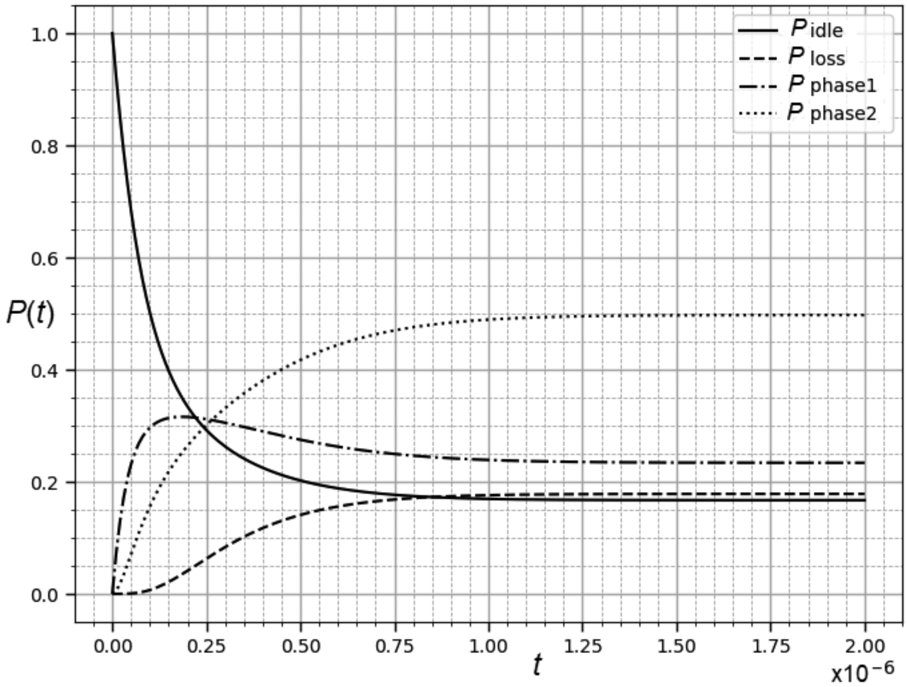

Fig. 2. Dependence of system state probabilities on time in transition mode.

with (17). Next, the elements of the matrix $\mathbf{L(t)}$ are written in accordance with (13). To do this, we find the poles of functions that describe the elements of the matrix $\mathbf{L(s-s_0)}$ in terms of the Laplace transform: $s_0 = 0$, $s_1 = -1.2 \cdot 10^7$, $s_2 = -2.8 \cdot 10^7$, $s_{3,4} = -3.9 \cdot 10^7 \pm i1.9 \cdot 10^7$, $s_{5,6} = -3.2 \cdot 10^7 \pm i9.9 \cdot 10^7$, $s_{7,8} = -2.8 \cdot 10^7 \pm i1.2 \cdot 10^7$, $s_{9,10} = -5.1 \cdot 10^7 \pm i1.4 \cdot 10^7$, $s_{11,12} = -1.5 \cdot 10^7 \pm i6.0 \cdot 10^7$, $s_{13,14} = -2.2 \cdot 10^7 \pm i4.0 \cdot 10^7$. One of these poles is zero, all others have a negative real part. This indicates the presence of a stationary mode in the system. Moreover, 12 of the 15 poles are pairwise complex conjugate, which indicates the oscillatory nature of the probabilities of states in the transition mode. Indeed, the exponent of a complex number in (12) is a combination of trigonometric functions in accordance with Euler's formula.

Studying the poles of state probability images also allows one to calculate the time constant using formula (28) $\tau = 1/|\alpha_{\min}| = 1/5\,138\,202.473908113 = 1.9462 \cdot 10^{-7}$ s and transition time $\tau_{tr} = 5\tau = 9,731 \cdot 10^{-7}$ s.

The dependence of state probabilities on time for the case under consideration is presented in Fig. 2. The figure shows: $P_{idle}(t)$ is the probability that the system is free; $P_{phase1}(t)$, $P_{phase2}(t)$ are dependencies of the probabilities of the states of finding requests only in the first and only in the second phases of service, respectively; $P_{loss}(t)$ is the probability of losses calculated in accordance with (29).

From Fig. 2 it can be seen that the time of the transition mode, calculated from (28), corresponds to the time of reaching the stationary mode according to the state probability graphs. The oscillatory nature of the transition mode is clearly visible from the dependence of the probability of finding requests in the first phase of service $P_{phase1}(t)$ (Fig. 3). Note that the probabilities of states in a stationary mode, obtained by the authors using the proposed approach, are equal to the stationary probabilities calculated using a well-known technique [24]. Indeed, from Fig. 2, it is clear that $\pi_{idle} = 0.167$, $\pi_{loss} = 0.172$, $\pi_{phase1} = 0.24$, $\pi_{phase2} = 0.49$, which corresponds to the stationary probabilities calculated using formulas (6) and (7) presented in [24].

Next, the performance indicators of the considered QS are calculated in accordance with Section 5 of this work.

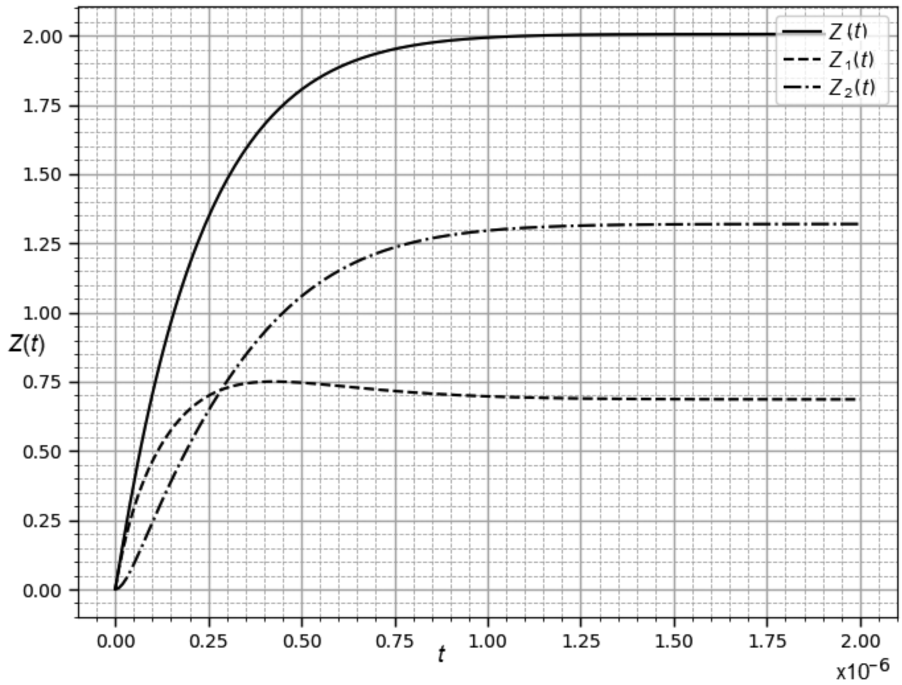

Fig. 3. Dependence of the average number of requests in each phase on time.

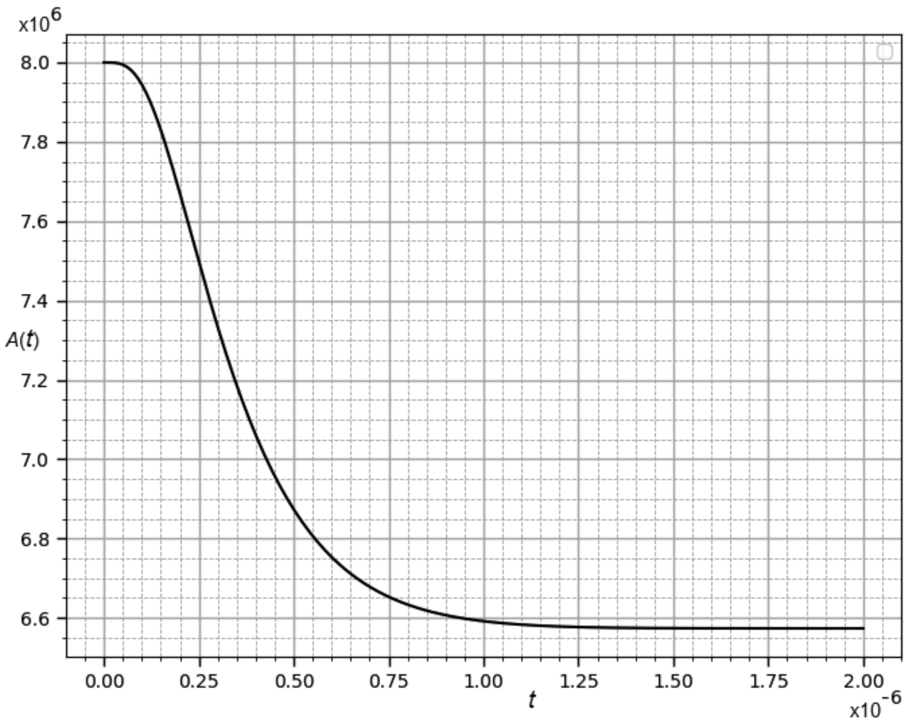

Fig. 4. Dependence of system throughput on time in transition mode.

Figure 4 shows the dependence of the system throughput on time in the transient mode, calculated in accordance with (30). The system throughput at the initial time is equal to $8 \cdot 10^6$ packets/s and decreases to a stationary value of $6.62 \cdot 10^6$ packets/s. Studying changes in the throughput of an all-optical switch in transient mode makes it possible to obtain more accurate estimates of its

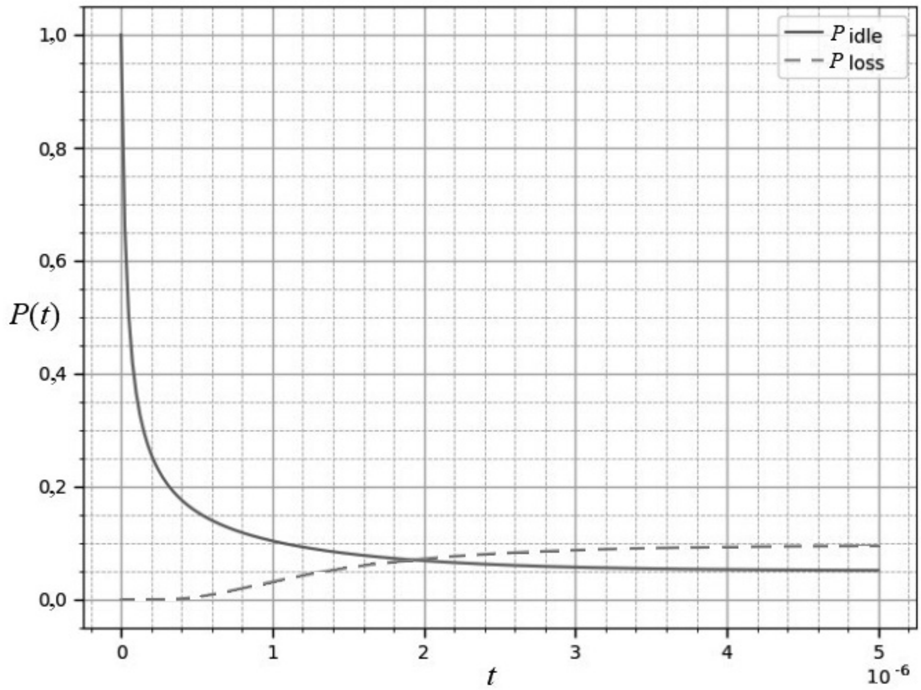

Fig. 5. Dependence of system state probabilities on time in transition mode.

performance, taking into account possible switch reboots when changing information transmission routes in all-optical networks.

Figure 5 shows the time dependence of the number of packets in the first and second phases, as well as the total number of packets in the system in the transient mode, calculated in accordance with (33)–(35). It can be seen that until the moment $t = 0.28 \cdot 10^{-7}$ s the number of packets in the first phase exceeds the number of packets in the second phase. At the same time, in stationary mode, the average number of packets in the first phase is less than in the second phase of service, which is obvious, since $\mu_1 > \mu_2$. Considering that the number of requests in the first and second phases of the QS under study corresponds to the number of packets processed in the first and second stages of the all-optical switch [8], the results obtained make it possible to estimate the degree of filling of the switch buffers during the transition mode.

As the buffer size increases, the size of the matrices in (12) increases, which requires additional computational resources. Figure 4 shows the calculation of a two-phase system with a buffer volume of $N = 15$: the probability of losses $P_{loss}(t)$ and the probability that the system is empty, $P_{idle}(t)$. The graph shows that with an increase in the buffer size, the probability of losses in the stationary mode decreased — $\pi_{loss} = 0.1$, and the time of the transition mode increased — $\tau_{tr} = 4 \cdot 10^{-7}$ s.

7. CONCLUSION

In this paper, the transient mode of a two-phase QS with a Poisson input flow, an exponential law of distribution of service time in each phase and a limitation on the total buffer size of two phases is considered and analyzed. Previously, the non-stationary mode of such a system was not considered in the world literature. However, it is of interest for various applications, in particular in the design of all-optical network switches. It should be noted that the study of the non-stationary mode of an all-optical switch allows a more accurate assessment of its performance metrics, which differ significantly from stationary values due to the high information transfer rate of all-optical networks [8].

A system of differential equations describing the functioning of this QS is presented, the solution of which is written using the Laplace transform. The characteristics of system performance in transient mode, such as the probability of losses, throughput, the average number of serviced requests, and the transition time, were obtained. Obviously, as the buffer size increases, obtaining numerical solutions to the characteristics of a two-phase QS with a limited buffer is a computationally intensive task and requires the use of high-performance computing systems or the use of approaches based on simulation modeling and machine learning [26].

FUNDING

This work was supported by the Russian Science Foundation, project no. 23-29-00795. https://rscf.ru/en/project/23-29-00795/

APPENDIX

Formally, eliminating certain terms in equations (4) and preserving the remaining ones can be done using the Heaviside function. However, this function is essentially logical, not analytical, and, therefore, does not allow one to write down an expression for the probabilities of system states in a general form. In particular, when using it in program code, it is necessary to organize additional loops. Therefore, to enable a compact analytical representation of the system of equations (4), the analytical function was introduced

$$\sigma_1(x, x_0) = \frac{|x - x_0| + x - x_0}{2|x - x_0|}. \tag{A.1}$$

Thus, the function limiting from below the permissible states of the system has the form

$$\upsilon_1(x, M) = \frac{|x - M + 0.5| + x - M + 0.5}{2|x - M + 0.5|}. \tag{A.2}$$

For example, for $M = 0$ the function $\upsilon_1(x, M)$ has the form shown in Fig. 6.

A shift of 0.5 along the time axis was chosen due to the fact that otherwise, in the state $x = M$ of the system, this function would be indefinite, and its derivative would tend to infinity at this point. Similarly with (A.1), we introduce the function

$$\sigma_2(x, x_0) = \frac{|x_0 - x| + x_0 - x}{2|x_0 - x|}. \tag{A.3}$$

Thus, the function that limits from above the permissible states of the system can be written in the form

$$\upsilon_2(x, K) = \frac{|K - x - 0.5| + K - x - 0.5}{2|K - x - 0.5|}, \tag{A.4}$$

where $K = \overline{0, N}$ is the state of the system. For $K = 4$, the function $\upsilon_2(x, M)$ has the form shown in Fig. 7.

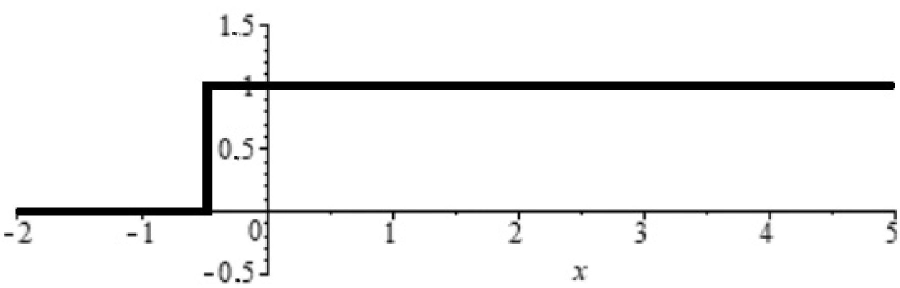

Fig. 6. Function $\upsilon_1(x, 0)$.

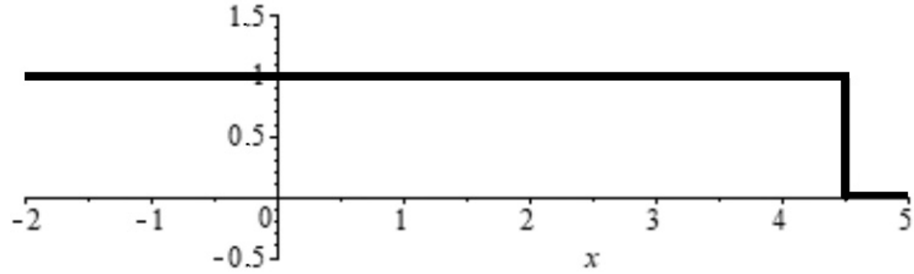

Fig. 7. Function $\upsilon_2(x, 4)$.

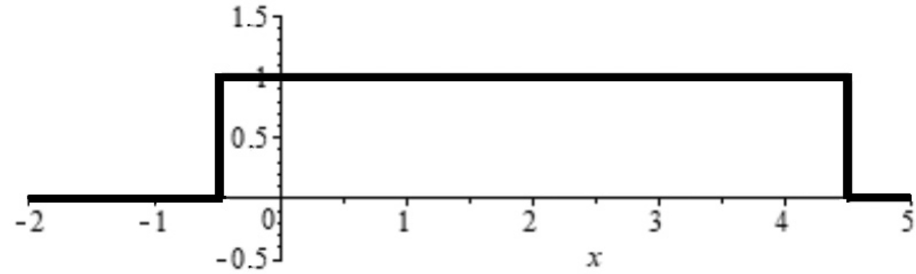

Fig. 8. Function $\upsilon(x, 0, 4)$.

Obviously, the function limiting the permissible range of values from the smallest M to the largest K takes the form

$$\begin{aligned}\upsilon(x, M, N) &= \upsilon_1(x, M)\,\upsilon_2(x, N) \\ &= \frac{(|x - M + 0.5| + x - M + 0.5)\,(|K - x - 0.5| + K - x - 0.5)}{4\,|x - M + 0.5|\,|K - x - 0.5|}. \end{aligned} \tag{A.5}$$

For example, with $M = 0$ and $K = 4$ it has the form shown in Fig. 8.

In relation to the problem being solved, x can take the values n_1, n_2, $n_1 + n_2$, etc. The advantage of functions (A.2), (A.4) and (A.5) is the absence of conditions. However, it should be noted that such conditions still exist when the module is expanded. However, despite the fact that these functions do not speed up the calculation process, they allow for analytical study of the resulting expressions and simplify the program code.

To find the function $\vartheta(n_k, n_l)$ (see (6)), which transforms a pair of numbers n_k, n_l, characterizing the state of the system, into the column number of the matrix $\mathbf{A}$, let us analyze the following pattern for $N = 4$: for $n_k = 0$ the values of n_l change from 0 to N, and the values of $\vartheta(n_k, n_l)$ change from 1 up to $N + 1$; for $n_k = 1$ the values of n_l change from 0 to $N - 1$, and the values of $\vartheta(n_k, n_l)$ change from $N + 2$ to $2N + 1$; for $n_k = 2$ the values of n_l change from 0 to $N - 2$, and the values of $\vartheta(n_k, n_l)$ change from $2N + 2$ to $3N$; for $n_k = 3$ the values of n_l change from 0 to $N - 3$, and the values of $\vartheta(n_k, n_l)$ change from $3N + 1$ to $4N - 2$; for $n_k = 4$ we have $n_l = 0$ and $\vartheta(n_k, n_l) = 4N - 1$. Therefore, the expression for $\vartheta(n_k, n_l)$ must contain the term $n_k(N + 1)$, as well as the term n_l. Thus, for $n_k = 0$:

$$\vartheta(0, n_l) = (N + 1)\,n_k + n_l + 1 = (N + 1)\,n_k + n_l + 0 \cdot (-0.5) + 1, \tag{A.6}$$

for $n_k = 1$:

$$\vartheta(1, n_l) = (N + 1)\,n_k + n_l + 1 = (N + 1)\,n_k + n_l - 1 \cdot 0 + 1, \tag{A.7}$$

for $n_k = 2$:

$$\vartheta\,(2, n_l) = (N+1)\,n_k + n_l + 0 = (N+1)\,n_k + n_l - 2\cdot 0.5 + 1, \tag{A.8}$$

for $n_k = 3$:

$$\vartheta\,(3, n_l) = (N+1)\,n_k + n_l - 2 = (N+1)\,n_k + n_l - 3\cdot 1 + 1, \tag{A.9}$$

for $n_k = 4$:

$$\vartheta\,(4, n_l) = (N+1)\,n_k + n_l - 5 = (N+1)\,n_k + n_l - 4\cdot 1.5 + 1, \tag{A.10}$$

for $n_k = m$:

$$\vartheta\,(m, n_l) = (N+1)\,m + n_l - (m+1) = (N+1)\,m + n_l - m\frac{m-1}{2} + 1. \tag{A.11}$$

Thus, expressions (A.6)–(A.11) connect a pair of numbers to the corresponding ordinal number of the element in the row (column) of the coefficient matrix. It is easy to check that relation (6) is valid for any N.

REFERENCES

1. Dudin, A.N., Klimenok, V.I., and Vishnevsky, V.M., *Methods to Study Queuing Systems with Correlated Arrivals*, Berlin/Heidelberg: Springer, 2020.
2. Rohit Singh Tomar and Dr.R.K. Shrivastav, Three Phases of Service For A Single Server Queueing System Subject To Server Breakdown And Bernoulli Vacation, *Int. J. Math. Trend. Techn. (IJMTT)*, 2020, vol. 66 (5), pp. 124–136.
3. Murat Sagir and Vedat Saglam, Optimization and analysis of a tandem queueing system with parallel channel at second station, *Communications in Statistics — Theory and Methods*, 2022, vol. 51, no. 21, pp. 1–14.
4. Sudhesh, R. and Vaithiyanathan, A., Stationary analysis of infinite queueing system with two — stage network server, *RAIRO-Oper. Res.*, 55, 2021, pp. 2349–2357.
5. Al-Rawi, Z.R. and Al Shboul, K.M.S., A Single Server Queue with Coxian-2 Service and One-Phase Vacation (M/C-2/M/1 Queue), *Open J. Appl. Sci.*, 2021, vol. 11, no. 6, pp. 766–774.
6. Serite Ozkar, Two-commodity queueing-inventory system with phase-type distribution of service times, *Annals of Operations Research*, 2022. https://link.springer.com/article/10.1007/s10479-022-04865-3
7. Anastasia Galileyskaya, Ekaterina Lisovskaya, Michele Pagano, and Svetlana Moiseeva, Two-Phase Resource Queueing System with Requests Duplication and Renewal Arrival Process, *LNCS*, 2020, 12563, pp. 350–364.
8. Barabanova, E.A., Vytovtov, K.A., Vishnevsky, V.M., and Podlazov, V.S., High-capacity strictly non-blocking optical switches based on new dual principle, *J. Phys.: Conf. Ser.*, 2021, vol. 2091, no. 1. https://iopscience.iop.org/article/10.1088/1742-6596/2091/1/012040
9. Ivanova, D., Adou, Y., Markova, E., Gaidamaka, Y., and Samouylov, K., Mathematical Framework for Mixed Reservation- and Priority-Based Traffic Coexistence in 5G NR Systems, *Mathematics*, 2023, vol. 11, no. 4. https://doi.org/10.3390/math11041046
10. Zeifman, A.I., Razumchik, R.V., Satin, Y.A., and Kovalev, I.A., Ergodicity bounds for the markovian queue with time-varying transition intensities, batch arrivals and one queue skipping policy, *Appl. Math. Comput.*, 2021, vol. 395, p. 125846.
11. Kempa Wojciech, M. and Paprocka Iwona, Transient behavior of a queueing model with hyper-exponentially distributed processing times and finite buffer capacity, *Sensors*, 2022, vol. 22, no. 24. https://doi.org/10.3390/s22249909

12. Rubino, G., *Transient analysis of Markovian queueing systems: a survey with focus on closed forms and uniformization / Queueing Theory 2: Advanced Trends*, Wiley-ISTE: Hoboken, NJ,. USA, 2021, pp. 269–307.

13. Zeifman, A., Korolev, V., and Satin, Y., Review Two Approaches to the Construction of Perturbation Bounds for Continuous-Time Markov Chains, *Mathematics*, 2020, vol. 8. https://doi.org/10.3390/math8020253

14. Sita Rama Murthy, M., Srinivasa Rao, K., Ravindranath, V., and Srinivasa Rao, P., Transient Analysis of K-node Tandem Queuing Model with Load Dependent Service Rates, *Int. J. Engin. Techno.*, 2018, vol. 7, no. 3.31, pp. 141–149.

15. Suhasini, A.V.S., Rao K. Srinivasa, and Reddy, P.R.S., Transient analysis of tandem queueing model with nonhomogenous poisson bulk arrivals having statedependent service rates, *Int. J. Advanc. Comput. Math. Sci.*, 2012, vol. 3, no. 3, pp. 272–289.

16. Neelam Singla and Garg P.C., Transient and Numerical Solution of a Feedback Queueing System with Correlated Departures, *Amer. J. Numer. Anal.*, 2014, vol. 2, no. 1, pp. 20–28.

17. Shyam Sundar Sah and Ram Prasad Ghimire, Transient Analysis of Queueing Model, *J. Inst. Engin.*, 2015, vol. 11, no. 1, pp. 165–171.

18. Zeifman, A.I., On the Nonstationary Erlang Loss Model, *Autom. Remote Control*, 2009, vol. 70, no. 12, pp. 2003–2012.

19. Kovalev, I.A., Satin, Ya.A., Sinitsina, A.V., and Zeifman, A.I., On one approach to estimating the convergence rate of non-stationary Markov models of queuing systems, *Inform. and its application*, 2022, vol 16, no. 3, pp. 75–82.

20. Prabhu, N.U., Transient Behaviour of a Tandem Queue, *Management Science*, 1967, vol. 13, no. 9, pp. 631–639. https://doi.org/10.1287/mnsc.13.9.631

21. Vyshnevsky, V.M., Vytovtov, K.A., Barabanova, E.A., and Semenova, O.V., Analysis of an $MAP/M/1/N$ queue with periodic and non-periodic piecewise constant input rate, *Mathematics*, 2022, vol. 10, no. 10. https://www.mdpi.com/2227-7390/10/10/1684

22. Vishnevsky, V., Vytovtov, K., Barabanova, E., and Semenova, O., Transient behavior of the $MAP/M/1/N$ queuing system, *Mathematics*, 2021, vol. 9, no. 2559. https://doi.org/10.3390/math9202559

23. Vytovtov, K.A., Barabanova, E.A., and Vishnevsky, V.M., Modeling and Analysis of Multi-channel Queuing System Transient Behavior for Piecewise-Constant Rates, *LNCS*, 2023, vol. 13766, pp. 397–409.

24. Jackson, R.R.P., Queueing Systems with Phase Type Service, *Oper. Res. Soc.*, 1954, vol. 5, no. 4, pp. 109–120.

25. Zhuravlev, A.P., Ryumshin, K.Yu., Atakischev, O.I., Titenko, E.A., and Titenko, M.A., Modulation parameters of modern communication systems, *T-Comm: telecommunications and transport.*, 2023, vol. 17, no. 7, pp. 13–20.

26. Vishnevsky, V.M. and Semyonova, O.V., *Metody mashinnogo obucheniya dlya issledovaniya stokhasticheskikh modelei tsiklicheskogo oprosa v shirokopolosnykh besprovodnykh setyakh* (Machine learning methods for studying stochastic models of cyclic polling in broadband wireless networks), Moscow: IPU, RAS, 2023.

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