Journal of Applied Mathematics and Stochastic Analysis, **13**:4 (2000), 415-427.

COMPLEX-ANALYTIC AND MATRIX-ANALYTIC SOLUTIONS FOR A QUEUEING SYSTEM WITH GROUP SERVICE CONTROLLED BY ARRIVALS

LEV ABOLNIKOV
Loyola Marymount University
Department of Mathematics
Los Angeles, CA 91316 USA

ALEXANDER DUKHOVNY
San Francisco State University
Department of Mathematics
San Francisco, CA 94132 USA

(Received February, 2000; Revised October, 2000)

A bulk M/G/1 system is considered that responds to large increases (decreases) of the queue during the service act by alternating between two service modes. The switching rule is based on two "up" and "down" thresholds for total arrivals over the service act. A necessary and sufficient condition for the ergodicity of a Markov chain embedded into the main queueing process is found. Both complex-analytic and matrix-analytic solutions are obtained for the steady-state distribution. Under the assumption of the same service time distribution in both modes, a combined complex-matrix-analytic method is introduced. The technique of "matrix unfolding" is used, which reduces the problem to a matrix iteration process with the block size much smaller than in the direct application of the matrix-analytic method.

Key words: Arrival-Controlled Systems, Complex-Analytic Method, Matrix-Analytic Method.

AMS subject classifications: 60K25.

1. Introduction

In controlled queueing systems, it is usually assumed that parameters of the arriving and service processes may depend on some characteristics of the queueing process (usually the queue length) at certain moments of time. This assumption enables the system to change input and service parameters in order to keep important elements of the queueing process under control.

In particular, in [1, 2] the authors considered bulk queueing systems where the average number of arriving groups of customers per unit time, their random sizes, and the service capacity depend on queue length at some special moments of time. As such, those parameters can be changed if the length of the queue reaches an unacceptable level.

This approach seems to be very promising and productive under one indispensable condition: the information about arrival and queueing processes must be complete, reliable, and available at any time the system is functioning. However, it is sometimes difficult or even impossible to achieve that condition in real-life queueing situations, at least for certain periods of time. For instance, in queueing problems with "server's vacations", the information about the number of customers in the system may not be available when the server is "on vacation" and, therefore, in such problems the system can control the queueing process only during busy periods. To find out how effectively the system can control the queueing process under these conditions, it is necessary to analyze the corresponding queueing model with the help of queueing theory methods.

In the existing literature, there are many publications related to controlled queueing systems with state dependent parameters. (The most recent and complete survey of these methods is contained in [5]). Due to the importance in computer science applications, many articles are devoted to systems with a "vacationing" server, in which the server may not be available for service during some periods of time. (An excellent survey of hundreds of models of this kind is contained in [4].) In all of those papers, it was supposed that the system has complete information about the queueing process all the time. To the best of our knowledge, controlled queueing systems with incomplete information have not been considered in the literature.

In this article, a controlled queueing system is analyzed where the waiting line is not reliably observable. Thus, parameters of the service process and service discipline depend in a certain way only on the number of customers that arrive to the system during the busy period. The following is the general description of this queueing model:

> Customers arrive at a service station by a Poisson process and are served in groups. The service time has a general distribution corresponding to the current service mode. The service station has two available modes of work: in Mode 1, the service group size is n, in Mode 2 it is m, where $m > n$. Mode 2 being more "expensive", the station seeks to restrict its usage to "heavy arrival" situations by imposing "arrival thresholds". At the moment before completion of a service act, a decision is made on the next service mode: it switches "up" from 1 to 2 if the number of arrivals over the last service period reached or exceeded m; it switches "down" from 2 to 1 if the number of arrivals was less than a "down threshold" N, where $n \leq N \leq m$; otherwise the mode remains as it was.

In the next section, we analyze the queueing process in this system via an imbedded, two-dimensional Markov chain, and establish transition equations for its steady-state probabilities and the necessary and sufficient condition of ergodicity.

In Section 3, we obtain a system of equations for the generating functions of the steady-state probabilities and represent its solution in terms of a finite number of undetermined constants. In the framework of the classical complex-analytic method, the missing constants are found from a system of equations based on the ("Rouche") roots of a characteristic equation.

Under the assumption that the service time distribution is the same for both service modes, a "divisibility" method is introduced to set up a real-arithmetic system of equations for the missing constants. The equations follow from the fact that a polynomial involving the constants should be divisible by the "characteristic" polynomial that vanishes at the above mentioned "Rouche" roots. Three special cases are studied, in which explicit formulas are obtained for the generating function of the stationary queue-length distribution:

Case 1. $N = m$;
Case 2. $N = m - 1$, $n = 1$;
Case 3. $N = n = 1$.

In Section 4, we discuss possible ways of finding matrix-analytic solutions for this system. In one method, we replace the two-dimensional, imbedded Markov chain by a one-dimensional chain of the so-called block type. For this chain, the steady-state probabilities can be found by a matrix-iteration process (e.g., Neuts [9]) with matrix size $M = 2m$.

At the same time, under the assumption that the service time distribution is the same for both modes, we apply the technique of "*matrix unfolding*" for a power series to the complex-analytic formulas of Section 3, and construct a matrix-analytic solution where the block size is reduced to $M = m + n - N$.

2. System Description, Basic Notations and Equations

Customers arrive to the queueing system by a Poisson process with a rate λ. The server operates continuously, and has two modes of work: in Mode 1, the service group size is n, in Mode 2 it is m. When the mode is r, the service time distribution function has Laplace transform $G_r(s)$. The mode switching rule described in the introduction is formalized as follows.

Let η_j be the service mode of the $(j+1)$st service act starting at the moment t_j. At $t_j - 0$, a decision is made on η_j. If η_{j-1} was 1, and the number of arrivals α_j over the jth service period was below m, then η_j is set to 1; if $\alpha_j \geq m$, the mode switches "up" to 2. If η_{j-1} was 2 and $\alpha_j < N$, where $n \leq N \leq m$, the service mode switches down to 1; otherwise, it remains the same.

Let Q_j be the queue length at $t_j + 0$. Clearly, $\{(Q_j, \eta_j)\}$ is a Markov chain, and its transitions can be formalized as follows:

$$\eta_{j+1} = \begin{cases} 1 & \text{if } ((\eta_j = 1 \wedge \alpha_j < m) \vee (\eta_j = 2 \wedge \alpha_j < N)), \\ 2 & \text{if } ((\eta_j = 1 \wedge \alpha_j \geq m) \vee (\eta_j = 2 \wedge \alpha_j \geq N)), \end{cases} \tag{2.1}$$

$$Q_{j+1} = \begin{cases} \max(0, Q_j + \alpha_j - n) & \text{if } \eta_{j+1} = 1, \\ \max(0, Q_j + \alpha_j 0m) & \text{if } \eta_{j+1} = 2. \end{cases} \tag{2.2}$$

By construction, the conditional probability generating function (pgf) of α_j, $j = 0, 1, \ldots$, given that $\eta_j = r$, is $K^r(z) = \sum_{i=0}^{\infty} k_i^r z^i = G_r(\lambda - \lambda z)$. For an arbitrary power series $W(z) = \sum_0^{\infty} w_i z^i$, we will denote(for an integer $s > 1$) $W_s^-(z) = \sum_{i=0}^{s-1} w_i z^i; W_s^+(z) = \sum_{i=s}^{\infty} w_i z^{i-s}$. Now, all transition probabilities a_{ir}^{ls} of the chain $\{(Q_j, \eta_j)\}$ from state (i, r) to state (l, s) can be written in terms of coefficients of $K^1(z)$ and $K^2(z)$ (see Appendix).

As usual, the transition relations for the steady-state probabilities p_{js}, $j = 0, 1, \ldots$,

$s = 1, 2$, (if they exist) are:

$$p_{js} = \sum_{i,r} p_{ir} a_{ir}^{js}; \quad \sum_{i,r} p_{ir} = 1. \tag{2.3}$$

In terms of generating functions $P_r(z) \equiv \sum_{i=0}^{\infty} p_{ir} z^i$, $r = 1, 2$, relations (2.3) can be written as follows:

$$P_1(z) = P_1(z) z^{-n} K_m^{1-}(z) + P_2(z) z^{-n} K_N^{2-}(z) + q_1(z), \tag{2.4}$$

$$P_2(z) = P_1(z) K_m^{1+}(z) + P_2(z) z^{N-m} K_N^{2+}(z) + q_2(z), \tag{2.5}$$

where

$$q_1(z) = \sum_{r=1}^{2} \sum_{i=0}^{n-1} p_{ir} \sum_{j=0}^{n-i-1} k_j^r (z^n - z^{i+j}),$$

and

$$q_2(z) = \sum_{i=0}^{m-N-1} p_{i2} \sum_{j=N}^{m-i-1} k_j^2 (z^{m-N} - z^{i+j}).$$

Or, in the vector form,

$$[P_1(z), P_2(z)][I - -H(z)] = [z^{-n} q_1(z), z^{N-m} q_2(z)], \tag{2.6}$$

where

$$H(z) = \begin{pmatrix} z^{-n} K_m^{1-}(z) & K_m^{1+}(z) \\ z^{-n} K_N^{2-}(z) & z^{N-m} K_N^{2+}(z) \end{pmatrix}. \tag{2.7}$$

Theorem 2.1: *The chain $\{(Q_j, \eta_j)\}$ has a steady-state distribution if and only if*

$$\delta = \pi_1(n - K^{1\prime}(1)) + \pi_2(m - K^{2\prime}(10) > 0, \tag{2.8}$$

where $\pi_1 = P_1(1)$ and $\pi_2 = P_2(1)$ are the steady-state probabilities that the system is in Mode 1 or 2:

$$\pi_1 = K_N^{2-}(1)[K_N^{2-}(1) + K_m^{1+}(1)]^{-1}; \; \pi_2 = K_m^{1+}(1) p K_N^{2-}(1) + K_m^{1+}(1)]^{-1}. \tag{2.9}$$

Proof: The transition relations (2.1) and (2.2) show that the chain $\{(Q_j, \eta_j)\}$ is irreducible and aperiodic, and from (2.2) it follows that it is also a so-called Markov controlled random walk (see Dukhovny [6]), with the jumps depending on the controlling parameter η. Formula (2.7) gives $H(z)$, the matrix generating function of its jumps.

It was proved in [6] that the condition

$$\left.\frac{d}{dz}\right|_{z=1} \det[I - H(z)] > 0 \tag{2.10}$$

is necessary and sufficient for the ergodicity of such a random walk. Condition (2.10) can also be written as

$$\pi_1[H'_{11}(1) + H'_{12}(1)] + \pi_2[H'_{21}(1) + H'_{22}(1)] < 0. \tag{2.11}$$

It follows from equation (2.6) that

$$[\pi_1, \pi_2][I - H(1)] = 0, \quad \text{with } \pi_1 + \pi_2 = 1. \tag{2.12}$$

Now, we use (2.7) to derive (2.9) from (2.12), and use (2.9) to derive (2.8) from (2.11). □

3. Complex-Analytic Approach to Stationary Probabilities

In the framework of the traditional complex-analytic approach, a formula for the generating function of the steady-state probabilities often contains a finite number of unknown constants. To determine the constants, a system of equations is obtained based on the fact that some expression containing the constants has to vanish at the roots of a "characteristic" equation (i.e., "Rouche roots").

Here, on the strength of (2.7), (2.6) can be written as

$$[P_1(z), P_2(z)]\begin{pmatrix} z^n - K_m^{1-}(z) & -z^{m-N}K_m^{1+}(z) \\ -K_N^{2-}(z) & z^{m-N} - K_N^{2+}(z) \end{pmatrix} = [q_1(z), q_2(z)]. \tag{3.1}$$

Solving the system (3.1), we obtain

$$P_1(z) = \Delta_1/\Delta; \quad P_2(z) = \Delta_2/\Delta, \tag{3.2}$$

for all z such that $\Delta \neq 0$, where

$$\Delta = z^{m+n-N}\det[I - H(z)] = z^{m+N} - \widehat{K}(z),$$

$$\widehat{K}(z) = z^{m-N}K_m^{1-}(z) + z^nK_N^{2+}(z) + z^{m-N}K_m^{1+}(z)K_N^{2-}(z) - K_m^{1-}(z)K_N^{2+}(z), \tag{3.3}$$

$$\Delta_1 = \{q_1(z)[z^{m-N} - K_N^{2+}(z)] + q_2(z)K_N^{2-}(z)\}, \tag{3.4}$$

and

$$\Delta_2 = \{q_1(z)z^{m-N}K_m^{1+}(z) + q_2(z)[z^n - K_m^{1-}(z)]\}. \tag{3.5}$$

Therefore, by the definition of $q_1(z)$ and $q_2(z)$, the formulas in (3.2) contain $m+n-N$ unknown constants p_{i1}, $i = 0,\ldots,n-1$, and p_{i2}, $i = 0,\ldots,m-N-1$.

Theorem 3.1: *Under the condition of ergodicity* (2.8), Δ *has* $m+n-N$ *roots in* $|z| \leq 1$ *(we denote the roots by* $\zeta_0 = 1, \zeta_1,\ldots,\zeta_{m+n-N-1}$*). Assuming that the roots are distinct, the constants* p_{i1}, $i = 0,\ldots,n-1$, *and* p_{i2}, $i = 0,\ldots,m-N-1$, *form a unique solution of the system of equations:*

$$q_1(z)[z^{m-N} - K_N^{2+}(z)] + q_2(z)K_N^{2-}(z) = 0, \quad \forall z = \zeta_1,\ldots,\zeta_{m+n-N-1} \tag{3.6}$$

$$\sum_{r=1}^{2}\sum_{i=0}^{n-1}p_{ir}\sum_{j=0}^{n-i-1}k_j^r(n-i-j) + \sum_{i=0}^{m-n-1}p_{i2}\sum_{j=N}^{m-i-1}k_j(m-i-j) = \delta. \tag{3.7}$$

(In the case when, for some t the multiplicity of ζ_t is $s > 1$, (3.6) would include vanishing derivatives of the left-hand side at ζ_t up to the order of $s-1$.)

Proof: It was proved in Dukhovny [6] that under the condition of ergodicity (2.8), the number of roots of $\det[I - H(z)]$ in the region $|z| < 1$ is equal to the number of its poles in the same region less 1 (both numbers counted with multiplicities). By (2.7), the only pole of $\det[I - H(z)]$ in $|z| < 1$, is $z = 0$, with multiplicity $m + n - N$. Thus, the number of roots of $\det[I - H(z)]$ and, therefore, of Δ in $|z| < 1$, is $m + n - N - 1$. Since $P_1(z)$ is analytic in $|z| < 1$, by Δ_1 has to vanish at those roots, and (3.4) yields (3.6). Also, (3.7) follows from the normalizing condition $P_1(1) + P_2(1) = 1$.

Under the ergodicity condition (2.8), the system (3.6)-(3.7) has a unique solution. Indeed, one solution is provided by the stationary probabilities. If there was any other solution, the formulas (3.2)-(3.5) would have provided a new absolutely summable solution of the transition equations (2.3), which is impossible under the ergodicity condition. □

Remark: Naturally, once the missing constants are found that make Δ_1 vanish at the roots of Δ, it will also make Δ_2 vanish at those roots, which guarantees that $P_2(z)$ is also analytic in $|z| < 1$. In fact, (3.5) could have been equivalently used in place of (3.4) to generate the equations of (3.6).

It turns out that under the assumption that the service time distribution is the same in both modes, a different method can be used to obtain the missing constants, thereby avoiding using the roots of Δ directly. Under this assumption,

$$K^1(z) = K^2(z) = K(z) = \sum_{j=0}^{\infty} k_j z^j,$$

$$q_1(z) = \sum_{i=0}^{n-1} p_i \sum_{j=0}^{n-i-1} k_j(z^n - z^{i+j}),\ q_2(z) = \sum_{i=0}^{m-N-1} p_2 \sum_{j=N}^{m-i-1} k_j(z^{m-N} - z^{i+j}),$$

where $p_i = p_{i1} + p_{i2}$;

$$\delta = \pi_1 n + \pi_2 m - K'(1),$$

$$\pi_1 = K_N^-(1)[1 - h(1)]^{-1}, \pi_2 = K_m^+(1)[1 - h(1)]^{-1},$$

where $h(z) = \sum_{i=N}^{m-1} k_i z^{i-n}$.

Now, adding (3.4) and (3.5) we obtain the generating function of the stationary queue length distribution $P(z) = P_1(z) + P_2(z)$, which is given by

$$P(z) = \Delta^{-1}\{q_1(z)[z^{m-N} - h(z)] + q_2(z)[z^n - z^N h(z)]\}. \tag{3.8}$$

We denote

$$z^{m+m-N} - g(z) \equiv \prod_{t=0}^{m+n-N-1} (z - \zeta_t), \tag{3.9}$$

where $g(z) = \sum_{j=0}^{m+n-N-1} g_j z^j$. Since the left-hand side of (3.6) is a polynomial, (3.6) is equivalent to the divisibility of that polynomial by $z^{m+n-N} - g(z)$.

Corollary 1: *Under the assumption of the same service time distribution, if the* *ergodicity condition* (2.8) *is true, the constants* $p_i = p_{i1} + p_{i2}$, $i = 0, \ldots, n-1$, *and* p_{i2}, $i = 0, \ldots, m - N - 1$, *form a unique solution of the system of equations*:

$$q_1(z)[z^{m-N} - h(z)] + q_2(z)[z^n - z^N h(z)] = 0 \ \ mod[z^{m+n-N} - g(z)], \quad (3.10)$$

$$\sum_{i=0}^{n-1} p_i \sum_{j=0}^{n-i-1} k_j(n-i-j) + \sum_{i=0}^{m-n-1} p_{i2} \sum_{j=N}^{m-i-1} k_j(m-i-j) = \delta. \quad (3.11)$$

In general, to implement (3.10), the remainder of division of its left-hand side by $z^{m+n-N} - g(z)$ must be obtained, and all its coefficients must be equated to 0. The resulting form of (3.10) is

$$\sum_{i=0}^{n-1} p_i \sum_{j=0}^{n-i-1} k_j\{(z^n - z^{i+j})[z^{m-N} - h(z)]\mathrm{mod}[z^{m+n-N} - g(z)]\}$$
$$+ \sum_{i=0}^{m-N-1} p_{i2} \sum_{j=N}^{m-i-1} k_j\{(z^{m-N} - z^{i+j})[z^n - z^N h(z)]\mathrm{mod}[z^{m+n-N} - g(z)]\} \equiv 0.$$

The "divisibility" approach involves a system of equations with real coefficients (as opposed to (3.6)). Also, it is obviously insensitive to repeating roots.

For some important special cases, under the assumption of the same service time distribution, it is possible to present the solution in an explicit closed form.

Special Case $N = m$: A simple way to organize switching is to use the "expensive" mode 2 *only* when the arrival over a service act is m or more, that is to set $N = m$. Then $h(z) \equiv q_2(z) \equiv 0$, and (2.9) shows that $\pi_1 = K_m^-(1), \pi_2 = K_m^+(1)$.

Now, equations (3.10) yield that $q_1(z) = c[z^n - g(z)]$, where c is found from the condition $P(1) = 1$. Formulas (3.2)-(3.5) yield a complete closed-form solution:

$$P(z) = \frac{[\pi_1 n + \pi_2 m - K'(1)][z^n - g(z)]}{[n - g'(1)][z^n - K_m^-(z) - z^n K_m^+(z)]},$$

$$P_1(z) = P(z)[1 - K_m^+(z)], \ \ P_2(z) = P(z)K_m^+(z).$$

Special Case $N = m - 1, n = 1$: In this case, $h(z) = k_{m-1}$, $q_1(z) = p_0 k_0(z-1)$, and $q_2(z) = p_{02}k_{m-1}(z-1)$.

Since $m + n - N = 2$, Δ has only one root ζ in $|z| < 1$, so $z^2 - g(z) = (z-1)(z-\zeta)$.

The system of equations (3.6)-(3.7) reduces to

$$p_0 k_0(\zeta - k_{m-1}) + p_{02}k_{m-1}(\zeta - \zeta^{m-1}k_{m-1}) = 0. \quad (3.12)$$

$$p_0 k_0 + p_{02}k_{m-1} = \delta = \pi_1 + \pi_2 m - K'(1). \quad (3.13)$$

Solving the system (3.12)-(3.13), we obtain:

$$p_{0k} = \delta(\zeta - \zeta^{m-1}k_{m-1})/k_{m-1}(1 - \zeta^{m-1}),$$

$$p_{02}k_{m-1} = \delta(k_{m-1} - \zeta)/k_{m-1}(1 - \zeta^{m-1}).$$

Now, (3.8) yields a closed-form result for $P(z)$:

$$P(z) = \frac{\delta(z-1)}{\Delta}[z - \frac{\zeta - \zeta^{m-1}k_{m-1}}{1-\zeta^{m-1}} - \frac{k_{m-1} - \zeta}{1-\zeta^{m-1}}z^{m-1}].$$

Special Case $N = n = 1$: Suppose the server is allowed to return to the (single service) Mode 1, from Mode 2, only if no customers arrived during the preceding service act. In this case, $q_1(z) = p_0k_0(z-1)$. Therefore, (3.8) reduces to

$$P(z)\Delta^{-1}\{p_0k_0(z-1)[z^{m-1} - h(z)] + zq_2(z)[1-h(z)]\}. \tag{3.14}$$

Under the assumptions of the case, it is possible to present the result in a closed form with two easily determined constants.

By the definition of $h(z)$, $1 - h(z) \neq 0$ in $|z| \leq 1$. Therefore, function $f(z) = [z^{m-1} - h(z)]/[1-h(z)]$ is *regular* (i.e., analytic inside and continuous on the boundary) in $|z| \leq 1$. Let $f^*(z)$ be an $(m-1)$st degree polynomial such that $\phi(z) = [f(z) - f^*(z)]/[z^m - g(z)]$ is regular $|z| \leq 1$. (That is, $f^*(z)$ interpolates values and, possibly, respective derivatives of $f(z)$ at the m roots of $z^m - g(z)$ in $|z| \leq 1$.)

Theorem 3.3: *When* $N = n = 1$,

$$P(z) = \Delta^{-1}[z^m - g(z)][1-h(z)][p_0k_0(z-1)\phi(z) + c], \tag{3.15}$$

$$c = [\pi_1 + \pi_2 m - K'(1)][m - g'(1)]^{-1}, \tag{3.16}$$

$$p_0k_0 = cg(0)/f^*(0). \tag{3.17}$$

Proof: Let us rewrite the right-hand side of (3.14) as

$$\Delta^{-1}[1-h(z)]\{p_0k_0(z-1)\phi(z)[z^m - g(z)] + p_0k_0(z-1)f^*(z) + zq_2(z)\}. \tag{3.18}$$

Since both sides of (3.14) must be regular in $|z| \leq 1$, it follows that the polynomial $p_0k_0(z-1)f^*(z) + zq_2(z)$ must vanish at the roots of Δ in $|z| \leq 1$. Since the degree of the polynomial is m, it can therefore be expressed as $c[z^m - g(z)]$. Hence we achieve (3.15). Evaluating (3.15) at $z = 0$ and $z = 1$, we obtain (3.16) and (3.17). $\square$

Clearly, the success of the "divisibility" method depends on how efficiently and accurately we can compute coefficients of $z^{m+n-N} - g(z)$. When $m+n-N$ is not large, one can find the roots of Δ in $|z| < 1$ and expand the product $\prod_{t=0}^{m+n-N-1}(z-\zeta_t)$. When $m+n-N$ is large, root-finding entails certain difficulties such as possible loss of (complex arithmetic) accuracy or handling of multiple roots. (We refer the reader, for example, to Chaudhry, et. al [3] for the discussion of root-finding.) As an alternative, a real-arithmetic method is presented below to find $g(z)$ using matrix iterations.

4. Matrix-Analytic Approach

In its common framework, using the matrix-analytic approach, one has to arrange the states of the Markov chain in point into a one-dimensional sequence so that the matrix of transition probabilities would have a characteristic M/G/1-type structure.

Namely, there would be some integer M, such that if one partitions the transition matrix into $M \times M$ matrix blocks, the blocks, starting from some block-row:

(1) would be the same along every direction parallel to the main diagonal, and

(2) all blocks below the first subdiagonal would be zero matrices.

Denoting by K_i, $i = 0, 1, \ldots$, the non-zero blocks in the repeating block-rows, a so called *power bounded* matrix G is sought that satisfies the *characteristic* matrix equation:

$$G = K(G) = \sum_{i=0}^{\infty} K_i G^i. \tag{4.1}$$

Such a matrix G is usually (approximately) obtained by successive iterations in (4.1):

$$G_{j+1} = K(G_j),\ G_0 = 0.$$

Based on the elements of G, blocks of the steady-state probabilities are then found one-by-one. We refer the reader to Neuts [9] for the fundamentals of the matrix-analytic method. In Gail, et. al [8], a formula was developed that expresses G in terms of its first row in a way that leads to faster computation of G.

To implement this method in our case, we replace the two-dimensional chain $\{(Q_j, \eta_j)\}$ by a one-dimensional chain $\{\nu_j\}$. The states are mapped as follows: $(0,1) \to 0$, $(0,2) \to 1$, $(1,1) \to 2$, $(1,2) \to 3$, etc.; that is, $(Q_j, \eta_j) \to \nu_j = 2Q_j + \eta_j - 1$.

Respectively, transition probabilities of the chain $\{\nu_j\}$ are given by:

$$P(\nu_{j+1} = v \mid \nu_j = u\} = a_{ir}^{ls},$$

where $i = [u/2]$, $r = u - 2i + 1$, $l = [v/2]$, $s = v - 2i + 1$.

Now, the formulas for a_{ir}^{ls} given in the Appendix demonstrate that the transition matrix for the chain $\{\nu_j\}$ is of block M/G/1-type, with the block size being $M = 2m$. From the same formulas we obtain the elements of matrices K_i, and find the matrix G by successive iterations in (4.1). For brevity, we have omitted the standard technical details of the computation process.

At the same time, under the assumption of the same service time distribution, based on the formulas developed in Section 3, the solution can be obtained by a matrix-analytic method with a significantly smaller block size . To accomplish this, we apply the technique of *matrix unfolding* for a power series that was introduced in Dukhovny [7].

Definition: Let $f(z) = \sum_{-\infty}^{\infty} f_i z^i \in W$. This represents the Wiener algebra of Laurent series with absolutely summable coefficients. A matrix unfolding of dimension M (M-unfolding) of $f(z)$ is a matrix series

$$U\{f(z)\} = \sum_{-\infty}^{\infty} U_j\{f(z)\} x^j, \quad |x| = 1, \tag{4.2}$$

where the coefficients of x^j are $M \times M$ matrices whose (r,s) elements are equal to f_{jM+s-r}, $r, s = 1, 2, \ldots, M$.

For example, if $f(z)$ is a polynomial of degree M,

$$U\{f(z)\} = L_f + R_f x, \tag{4.3}$$

$$L_f = \begin{pmatrix} f_0 & f_1 & \cdots & f_{M-1} \\ 0 & f_0 & \cdots & f_{M-2} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & f_0 \end{pmatrix}, \quad R_f = \begin{pmatrix} f_M & 0 & \cdots & 0 \\ f_{M-1} & f_M & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ f_1 & \cdots & f_{M-1} & f_M \end{pmatrix}. \tag{4.4}$$

It is easy to see that if $f(z) \in W$, $U\{f(z)\}$ absolutely converges $\forall x$, $|x| = 1$. Also, if $f(z) \in W^+$, (that is, $f_j = 0$, $\forall j \leq 0$) $U\{f(z)\}$ absolutely converges $\forall x$, $|x| \leq 1$; if $f(z) \in W^-$, (that is, $f_j = 0, \forall j > 0$) $U\{f(z)\}$ absolutely converges $\forall x$, $|x| \geq 1$.

Lemma 1: *The M-unfolding of the product of two functions belonging to W is the product of the M-unfoldings of the factors.*

The proof of Lemma 1 follows directly from (4.2).

Definition: Let $\|\cdot\|_1$ be the norm of a matrix equal to the maximum of the sums of the absolute values of the elements of every row. Matrix $D = (d_{rs})$ will be called *power bounded* if $\| D^m \|_1 \leq \text{cost}$, $j = 1, 2, \ldots$.

Suppose a series $f(z) = \sum_0^\infty f_i z^i \in W$. A matrix "substitution" can be defined for its M-unfolding, where x is replaced by a power bounded matrix D of the same size M. It is easy to see that the following matrix series converge absolutely:

$$U\{f(z)\}\big|_D = \sum_{j=0}^{\infty} U_j\{f(z)\} D^j. \tag{4.5}$$

For an arbitrary $(M-1)$st degree polynomial $g(z)$, denoted by C its $M \times M$ companion matrix whose first superdiagonal consists of 1s, the Mth row elements are $g_0, \ldots, g_{M-1}$, and all the other elements are zeros. It is a standard fact of the theory of matrices that the eigenvalues of C coincide with the roots of $z^M - g(z)$ (with the same multiplicities). Let $G = C^M$.

The proofs of the following Lemmas 2 and 3 are given in the Appendix.

Lemma 2: *For an arbitrary $(M-1)$st degree polynomial $g(z)$,*

$$(I - R_g)G = L_g. \tag{4.6}$$

Lemma 3: *If all roots of $z^M - g(z)$ belong to $|z| \leq 1$, and the roots belonging to $|z| = 1$ are simple, then G is power bounded.*

Now, let $M = m + n - N$, let $g(z)$ be the polynomial defined by (3.9), and use its companion matrix C to define matrix $G = C^M$. By (4.6), the first row of G consists of the coefficients of $g(z)$. Having found G, we complete and solve the system of equations (3.10)-(3.11), thus completing the formulas for the steady-state probabilities.

Let $\widehat{K}(z) = z^{m-N} K_m^-(z) + z^n K_N^+(z) - h(z)K(z)$ and let $\widehat{K}_j = U_j\{\widehat{K}(z)\}$, $j = 0, 1, \ldots$, be the coefficients of the M-unfolding of $\widehat{K}(z)$.

Theorem 4.1: *Matrix G is a unique power bounded solution of the characteristic equation*

$$G = \sum_{j=0}^{\infty} \widehat{K}_j G^j. \tag{4.7}$$

Proof: Consider the function

$$\phi(z) = \Delta[z^{m+n-N} - g(z)]^{-1}. \tag{4.8}$$

A routine analysis shows that under the condition of ergodicity (2.8), both $\phi(z)$ and $[\phi(z)]^{-1}$ are analytic in $|z| < 1$, continuous on $|z| = 1$, and belong to W. To find G, let us rewrite (4.8) as:

$$\Delta = z^{m+n-N} - \widehat{K}(z) = \phi(z)[z^{m-n+1} - g(z]. \tag{4.9}$$

Using $M = m+n-N$, substitute G into M-unfoldings of both sides of (4.9) as defined in (4.5). On the strength of (4.6), we obtain:

$$G - \sum_{j=0}^{\infty} \widehat{K}_j G^j = \sum_{j=0}^{\infty} U_j\{\phi(z)\}[-L_g + (I - R_g)G] = 0. \tag{4.10}$$

Therefore, by Lemma 3 and (4.10), G is a power bounded solution of (4.7).

Any other power bounded solution of (4.7) must coincide with G. Indeed, suppose $\widehat{G}$ is a power bounded solution of (4.7). Rewrite (4.8) as

$$[\phi(z)]^{-1}[z^{m+n-N} - \widehat{K}(z)] = z^{m+n-N} - g(z). \tag{4.11}$$

Again, substitute $D = \widehat{G}$ into the M-unfoldings of both sides of (4.11) as defined in (4.5). The left hand side of (4.11) belongs to W by construction, so the resulting matrix series in the left hand side will converge absolutely. Since $\widehat{G}$ is assumed to be a solution of (4.7), we obtain from (4.11) that

$$0 = -Lg + (I - R_g)\widehat{G},$$

so by (4.6), $\widehat{G} = G$. Therefore, G is the unique power bounded solution of (4.7). □

Similar to (4.1), we compute G by an iteration process in (4.7), setting $G_0 = 0$, with $G_{j+1} = \widehat{K}(G_j)$. However, the block size M has now reduced from $2m$ to $m+n-N$. The benefit can be immediately seen, for example, in the special case (see Section 3) where $N = m-1$, $n = 1$; here M reduces from $2m$ to just 2.

Appendix

Transition Probabilities of the Chain $\{(Q_j, \eta_j)\}$: Analyzing transition relations (2.1) and (2.2), one can see that

$$a_{01}^{01} = \sum_{j=0}^{n} k_j^1, \quad a_{01}^{l1} = \begin{cases} k_{l+n}^1 & \text{if } 0 < l < m-n, \\ 0 & \text{if } l \geq m-n \end{cases} \qquad a_{01}^{l2} = k_{l+m}^1,$$

$$a_{i1}^{l1} = \begin{cases} k_{l-i+n}^1 & \text{if } 0 < l < m+i-n, \\ 0 & \text{if } l \geq m+i-n, \end{cases}, \quad a_{i1}^{l2} = \begin{cases} k_{l-i+m}^1 & \text{if } 0 < i \leq l, \\ 0 & \text{if } 0 < l < i, \end{cases}$$

$$a_{02}^{01} = \sum_{j=0}^{N} k_j^2, \quad a_{i2}^{l1} = \begin{cases} k_{l-i+n}^2 & \text{if } 0 < l < i+N-n, \\ 0 & \text{if } l \geq i+N-n, \end{cases}$$

$$a_{i2}^{02} = \begin{cases} \sum_{j=N+i}^{m} k_j^2 & \text{if } i+N \leq m, \\ 0 & \text{if } i+N > m, \end{cases} \qquad a_{i2}^{l2} = \begin{cases} k_{l-i+m}^2 & \text{if } l \geq i-m+N, \\ 0 & \text{if } l < i-m+N. \end{cases}$$

Proof of Lemma 2: Suppose, at first, that all roots of $z^M - g(z)$ are simple, and denote them by ζ_j, $j = 1,\ldots,M$. Let $\bar{w}_j = (1, \zeta_j, \ldots \zeta_j^{M-1})^T$ and let Ω be the Vandermonde matrix whose jth column is $\bar{w}_j$. By the definition of C, $C\bar{w}_j = \zeta_j \bar{w}_j$, so

$$G\bar{w}_j = \zeta_j^M \bar{w}_j, \; G\Omega = \Omega[\text{diag}(\zeta_j^M, j = 1,\ldots,m)]. \tag{A.1}$$

On the strength of (A.1), since det $\Omega \neq 0$, (4.6) is equivalent to

$$\Omega[\text{diag}(\zeta_j^M, j = 1,\ldots,M)] = R_g \Omega[\text{diag}(\zeta_j^M, j = 1,\ldots,M)] + L_g \Omega. \tag{A.2}$$

An (r,s) element of the left-hand side of (A.2) is ζ_s^{r-1+M}. An (r,s) element of the right-hand side is $\zeta_s^{r-1} g(\zeta_s) = \zeta_s^{r-1+M}$, since ζ_s is a root of $z^M - g(z)$.

If some of the roots of $z^M - g(z)$ repeat, then it is always possible to find an $(M-1)$st degree polynomial $\widehat{g}(z)$ arbitrarily close to $g(z)$ such that $z^M - \widehat{g}(z)$ would have only simple roots, and (4.7) would hold. Since the elements of matrices in (4.6) are continuous functions of the coefficients of $g(z)$, it means that (4.6) holds for every $g(z)$. □

Proof of Lemma 3: Let $\{\lambda_j\}$ be the eigenvalues of G. Since $G = C^m$, these eigenvalues are Mth powers of the roots of $z^M - g(z)$. Under the assumptions of Lemma 3, it means that all the eigenvalues belong to the unit disk, and those eigenvalues that belong to the boundary $|z| = 1$ are simple. Let us represent $G = W \Lambda W^{-1}$, where $\Lambda = \text{diag}\{\Lambda_j\}$; Λ_j are the Jordan cells for each of the λ_j. Since $G^i = W\Lambda W^{-1}$ and, under the assumptions of the lemma, $\| \Lambda^i \|_1 \to 1$ as $i \to \infty$, we conclude that G is power bounded. □

References

[1] Abolnikov, L.M., Dshalalow, J.H., and Dukhovny, A.M., On stochastic processes in a multilevel control queueing system, *Stoch. Anal. and Appl.* **10**:2 (1992), 155-179.

[2] Abolnikov, L.M. and Dukhovny, A.M., Markov chains with transition delta-matrix: Ergodicity conditions, invariant probability measures and applications, *J. Appl. Math. and Stoch. Anal.* **5**:1 (1992), 83-98.

[3] Chaudhry, M.L., Harris, C.M. and Marchal, W.G., Robustness of rootfinding in single-server models, *ORSA J. on Computing* **2**:3 (1990), 273-286.

[4] Doshi, B.T., Single-serve queues with vacations, *Stoch. Anal. of Comput. and Commun. Sys.*, Elsevier Science, North Holland (1990), 217-265.

[5] Dshalalow, J.H., Queueing systems with state dependent parameters, In: *Fronters in Queueing: Models and Applications* (ed. by Dshalalow, J.H.), CRC Press, Boca Raton, FL (1997), 61-116.

[6] Dukhovny, A.M., Complex analytic factorization and partial indices of matrix functions associated with Markov controlled random walks, *Stoch. Models* **12**:4 (1996), 683-697.

[7] Dukhovny, A.M., Matrix-geometric solutions for bulk GI/M/1 systems with unbounded arrival groups, *Stoch. Models* **15**:3 (1999), 547-559.

[8] Gail, H.R., Hantler, S.L., and Taylor, B.A., Non-skip-free M/G/1 and G/M/1 type Markov chains, *Adv. in Appl. Probab.* **29**:3 (1997), 733-758.

[9] Neuts, M.F., *Matrix-Geometric Solutions in Stochastic Models*, The John Hopkins University Press 1981.