

ON SOME QUEUE LENGTH CONTROLLED STOCHASTIC PROCESSES¹

LEV ABOLNIKOV

*Department of Mathematics, Loyola Marymount University
Los Angeles, CA 90045, USA*

JEWGENI E. DSHALALOW

*Department of Applied Mathematics, Florida Institute of Technology
Melbourne, FL 32901, USA*

ALEXANDER M. DUKHOVNY

*Department of Mathematics, San-Francisco State University
San-Francisco, CA 94132, USA*

ABSTRACT

The authors study the input, output and queueing processes in a general controlled single-server bulk queueing system. It is supposed that inter-arrival time, service time, batch size of arriving units and the capacity of the server depend on the queue length.

The authors establish an ergodicity criterion for both the queueing process with continuous time parameter and the embedded process, study their transient and steady state behavior and prove ergodic theorems for some functionals of the input, output and queueing processes. The following results are obtained: Invariant probability measure of the embedded process, stationary distribution of the process with continuous time parameter, expected value of a busy period, rates of input and output processes and the relative speed of convergence of the expected queue length. Various examples (including an optimization problem) illustrate methods developed in the paper.

Key words: Controlled Input, Controlled Group Size, Controlled Service, Controlled Batch Size, Single-Server Queue, Queueing Process, Embedded Markov Chain, Semi-Regenerative Process, Semi-Markov Process, Ergodic Theorems, Input Process, Output Process, Optimization.

AMS Subject Classification: Primary 60K10, 60K25, Secondary 90B22, 90B25.

1. INTRODUCTION

Let $\xi(t)$ denote the number of units in a single-server queueing system at time $t \geq 0$ and let $\xi_n = \xi(t_n + 0)$, $n = 1, 2, \dots$, where t_n are successive moments of service completions. We assume here that the server capacity, service time and the input process depend upon the number of units in the system as follows. If at time $t_n + 0$ the number of units ξ_n is j then:

- within the random time interval (t_n, t_{n+1}) the input flow of units is a compound Poisson

¹Received: May 1990 Revised: October 1990

process with interarrival rate λ_j and with batch size distributed according to $\{\alpha_i^{(j)}; i = 1, 2, \dots\}$ and with mean $a^{[j]}$.

- the server takes for service the n th group of units of size $\min\{j, m(j)\}$ (where $m(k)$ is the server capacity, $k = 0, 1, \dots$), provided that $j > 0$. If $j = 0$ the server begins its *idle period* which ends as soon as a next group of units arrives at the system. This group is of random size $\alpha_i^{(0)}$ and the server takes $\min\{\alpha_i^{(0)}, m(0)\}$ units for the next service.

- the n th group is being served a random time with the distribution function $B_j \in \{B_0, B_1, \dots\}$ of general type with the finite first moment b_j .

Such a system supplements and generalizes the established class of single-server bulk queues with state dependent parameters, therefore, including a larger variety of real situations subject to stochastic control. Many authors, who treated queueing processes in special models of this class (pioneered by Finch [12]) (*cf.* [3-5, 7-9, 12-16]) analytically, were facing real problems to justify results obtained and to reduce a solution to an explicit form. Some other authors (*cf.* [11]) instead developed numerical algorithms. While there is no doubt about the need for queueing systems with broad stochastic control, it is also desirable to improve the analysis of existing results.

In the present paper the authors consider the queueing, input, output processes and some functionals of them in the above queueing system. After a formal description of all participating processes, the authors treat the embedded process $\{\xi_n\}$ and establish a necessary and sufficient criterion of its ergodicity. The transient and stationary distribution of $\{\xi_n\}$ and the stationary distribution of $\{\xi(t)\}$ are found in terms of their generating functions. In the last section, besides the queueing process, the authors study the input and output processes obtaining explicit formulas for expected rates of these processes and establishing several ergodic theorems. Various examples, including an optimization problem, demonstrate applications of the general results obtained in this article.

2. FORMAL DESCRIPTION OF THE SYSTEM

The stochastic process $\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, \xi(t); t \geq 0\} \rightarrow Q = \{0, 1, \dots\}$ or shortly $\{\xi(t)\}$ describes the number of units in a single-server queueing system at time t . Other processes are related to $\{\xi(t)\}$ as follows. Denote by

a) $\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, t_n; n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\} (t_0 = 0)\} \rightarrow (\mathbb{R}_+, \mathfrak{B}_+)$ (where $\mathfrak{B}_+ = \mathfrak{B}(\mathbb{R}_+)$ is the Borel σ -algebra) the point process of successive moments of service completions;

b) $\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, \xi_n = \xi(t_n +); n \in \mathbb{N}_0\} \rightarrow Q$ the embedded process over the moments of time $\{t_n\}$;

c) $\{\tau_s; s \in \mathbb{N}\}$ a stationary Poisson point process with rate λ describing the flow of groups of arriving customers;

d) $\{N(t); t \geq 0\}$ the associated compound Poisson counting process with $\alpha_s := N(\tau_{s+1}) - N(\tau_s)$ as the size of the s th arriving batch. $\{\alpha_s\}$ is assumed to be a lattice renewal process independent of $\{\xi(t)\}$ and $\{t_n\}$ and such that $a(z) := E[z^{\alpha_s}] = \sum_{i \geq 1} a_i z^i$, $z \in B(0,1) \cup \{-1,1\}$, (where $B(z_0, R)$ is an open ball in \mathbb{C} (with Euclidean norm) centered at z_0 and with radius R), $0 < a := E[\alpha_s] < \infty$, $s = 1, 2, \dots$;

e) $\{\tau_s^{(j)}; s \in \mathbb{N}\}$ the j th *partial input process* that will govern the flow of group of units on those intervals (t_n, t_{n+1}) when $\xi_n = j$, i.e. let $\{\hat{\tau}_s^{(j)}; s \in \mathbb{N}\}$ be a stationary Poisson point process with rate λ_j then

$$\tau_s^{(j)} = \sum_{n=0}^{\infty} \hat{\tau}_s^{(j)} I_{\{j\} \times (t_n, t_{n+1})} \circ (\xi_n, \hat{\tau}_s^{(j)}),$$

where I_A denotes the indicator function of a set A . The corresponding controlled arrival point process denoted by $\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, \tau_s^\xi; s \in \mathbb{N}\} \rightarrow (\mathbb{R}_+, \mathfrak{B}_+)$ is then the superposition of partial input processes $\{\tau_s^{(j)}\}$, $j = 0, 1, \dots$;

f) $\{N_j(t); t \geq 0\}$ the compound Poisson counting process associated with the above point process $\{\tau_s^{(j)}\}$. It should count only those units that arrive during the intervals (t_n, t_{n+1}) when $\xi_n = j$. More formally, let $\{\hat{N}_j(t)\}$ be a compound Poisson process with rate λ_j , the batch size distributed according to $\{a_i^{(j)}; i = 1, 2, \dots\}$ and the batch size mean $a^{[j]}$. Then

$$N_j(t) = \sum_{n=0}^{\infty} [\hat{N}_j(\min\{t_{n+1}, t\}) - \hat{N}_j(t_n)] I_{\{j\} \times (0, t]} \circ (\xi_n, t_n)$$

With the formalism like in d) we have: $\alpha_s^{(j)} := N(\tau_{s+1}^{(j)}) - N(\tau_s^{(j)})$ denotes the size of s th arrival batch of units on those intervals (t_n, t_{n+1}) when $\xi_n = j$. Unlike $\{\alpha_s\}$, the process $\{\alpha_s^\xi\}$ depends upon $\{\xi_n\}$ and $\{t_n\}$. But we assume that, given ξ_n , the random variables $\alpha_{\theta_n}^{(\xi_n)}$, $\alpha_{\theta_n+1}^{(\xi_n)}$, $\alpha_{\theta_n+2}^{(\xi_n)}$, ..., $\alpha_{\theta_n+1-1}^{(\xi_n)}$, are conditionally independent (on (t_n, t_{n+1})), where $\theta_n = \inf\{s \in \mathbb{N}: \tau_s^\xi \geq t_n\}$. Denote $a_j(z) := E[z^{\alpha_s^{(j)}}] = \sum_{i \geq 1} a_i^{(j)} z^i$, $z \in B(0,1) \cup \{-1,1\}$, $0 < a^{[j]} := E[z^{\alpha_s^{(j)}}] < \infty$, $s = 1, 2, \dots$;

g) $\{N^\xi(t)\}$ the arrival counting process associated with the controlled point process $\{\tau_s^\xi\}$. $\{N^\xi(t)\}$ is the sum of all partial counting processes $\{N_j(t)\}$, i.e. $N^\xi(t) = \sum_{j \geq 0} N_j(t)$.

The Poisson processes introduced in c) and d) will be considered as special cases of the corresponding controlled processes described in e)-g). In this event, the associated superscripts will be dropped.

Units arrive at the system in accordance with processes introduced in e)-g) and they are

placed in a line in a waiting room of an infinite capacity. Arriving groups of units obey the FIFO discipline. Within any group the order is arbitrary. At n th moment of servicing time following t_n , the server takes for service the n th group of units of size $M(\xi_n)$ defined as

$$(2.1) \quad M(\xi_n) = \begin{cases} \min\{\hat{N}_0(\tau_{\theta_n}^\xi), m(0)\}, & \xi_n = 0 \\ \min\{\xi_n, m(\xi_n)\}, & \xi_n > 0 \end{cases}$$

where $m(\cdot)$ denotes the capacity of the server. Observe that the server takes units for service at time $t_n + 0$, on a busy period, and at time $\tau_{\theta_n}^{(0)} + 0$, at the end of a corresponding idle period. The service time σ_n of the n th batch is distributed according to B_{ξ_n} picked from a given sequence $\{B_0, B_1, \dots\}$ of distribution functions each of which is concentrated on \mathbb{R}_+ and with the properties $B_i(0+) = 0$, $B_i(+\infty) = 1$, $b_i := E[\sigma_n | \xi_{n-1} = i] \in (0, \infty)$, $i = 0, 1, \dots$. The service times σ_n are supposed to be conditionally independent given ξ_n , $n = 1, 2, \dots$.

We introduce another notion. Let $\{C(t)\}$ be the counting process associated with the point process $\{t_n\}$. Then, the first group of units that arrives in interval $(t_{C(t)}, t_{C(t)+1})$ will be of size $N^\xi(\tau_{\vartheta_t}^\xi)$, and it arrives at time $\tau_{\vartheta_t}^\xi$, where $\vartheta_t := \inf\{s \in \mathbb{N} : \tau_s^\xi \geq t_{C(t)}\}$. Then, as in (2.1), we can similarly define the size of group which the server takes for service in interval $(t_{C(t)}, t_{C(t)+1})$:

$$(2.2) \quad M(\zeta(t)) = \begin{cases} \min\{\hat{N}_0(\tau_{\vartheta_t}^\xi), m(0)\}, & \zeta(t) = 0 \\ \min\{\zeta(t), m(\zeta(t))\}, & \zeta(t) > 0 \end{cases}$$

where $\zeta(t)$ denotes the value of the queueing process $\xi(t)$ at time $t_{C(t)}$.

If μ denotes the initial measure of the process $\{\xi(t)\}$ (or $\{\xi_n\}$), we will use for μ the unit mass ϵ_x to emphasize that the process started from a state $\{x\}$. The corresponding conditional probabilities and expectations are denoted by P^x and E^x .

3. EMBEDDED PROCESS $\{\xi_n\}$

Let ν_n be the number of units arriving at the system during the service in (t_n, t_{n+1}) . Obviously, for every $n = 0, 1, \dots$, the elements of the process $\{\xi_n\}$ are connected by the following relation:

$$\xi_{n+1} = \begin{cases} [\xi_n - m(\xi_n)]^+ + \nu_n, & \xi_n > 0 \\ (\alpha_{\theta_n}^{(\xi_n)} - m(0))^+ + \nu_n, & \xi_n = 0, \end{cases}$$

(where $(u)^+ := \sup\{u, 0\}$). Due to the nature of the arrival process, it follows then that $\{t_n\}$ is a sequence of stopping times relative to the canonic filtering $\sigma(\xi(u); u \leq t)$ (t -past σ -algebra of the process $\xi(t)$) and that $\{\Omega, \mathcal{F}, (P^x)_{x \in Q}, (\xi_n, t_n): n = 0, 1, \dots\} \rightarrow (Q \times \mathbb{R}_+, \mathcal{B}(Q \times \mathbb{R}_+))$ is a Markov renewal process (that will appear again in section 6). $\{\xi_n\}$ is a homogeneous Markov chain whose transi-

tion probability matrix is denoted by $A = (a_{ij}; i, j \in \mathbb{N}_0)$. We assume that for some N (which may be arbitrarily large)

$$\begin{aligned} \lambda_j &= \lambda, a_j(z) = a(z), j = N + 1, N + 2, \dots, \\ m(N + 1) &= m(N + 2) = \dots = :m (\leq N + 1) \end{aligned}$$

and

$$B_{N+1} = B_{N+2} = \dots = :B$$

Under the above restrictions, the transition probability matrix A of the Markov chain $\{\xi_n\}$ is reduced to a so-called $\Delta_{m,N}$ -matrix

$$(3.1) \quad A = (a_{ij}; i, j \in Q; a_{ij} = k_{j-i+m}, i > N, j \geq i - m; a_{ij} = 0, i > N, j < i - m)$$

studied earlier by Abolnikov and Dukhovny [2], where the values k_{j-i+m} of the corresponding entries a_{ij} can be determined from formula (3.3) given below.

Let $K(z) := \sum_{j \geq 0} k_j z^j$ and $A_i(z) = \sum_{j \geq 0} a_{ij} z^j$. The evaluation of A_i results

$$(3.2) \quad A_0(z) = \left[\sum_{i=1}^{m(0)} a_i^0 (1 - z^{i-m(0)}) + z^{-m(0)} a_0(z) \right] K_0(z) = K_0(z) \sum_{i \geq 1} a_i^0 z^{(i-m(0))^+}$$

$$(3.2a) \quad A_i(z) = z^{(i-m(i))^+} K_i(z), i = 1, 2, \dots, N,$$

$$(3.2b) \quad A_i(z) = z^{i-m} K(z), i = N + 1, N + 2, \dots,$$

$$(3.3) \quad K_i(z) = \beta_i(\lambda_i - \lambda_i a_i(z)), i = 0, 1, \dots, K(z) = \beta(\lambda - \lambda a(z)),$$

where $\beta_i(\theta) := \int_0^\infty \exp(-\theta x) B_i(dx), i = 0, 1, \dots, \beta(\theta) = \int_0^\infty \exp(-\theta x) B(dx), Re(\theta) \geq 0$.

The following two theorems were established by Abolnikov and Dukhovny [2]:

3.1 Theorem. *Let $\{\xi_n\}$ be an irreducible aperiodic Markov chain with the transition probability matrix A in the form of a $\Delta_{m,N}$ -matrix (3.1). $\{\xi_n\}$ is recurrent-positive if and only if*

$$(3.4) \quad \lim_{z \rightarrow 1: z \in B(0,1)} \frac{d}{dz} A_i(z) < \infty, i = 0, 1, \dots, N,$$

and

$$(3.5) \quad \lim_{z \rightarrow 1: z \in B(0,1)} \frac{d}{dz} K(z) < m.$$

The condition (3.5) is equivalent to

$$(3.5a) \quad \rho := \lambda a b < m$$

3.2 Theorem. *Under the condition of (3.5) the function $z^m - K(z)$ has exactly m roots that belong to the closed unit ball $\bar{B}(0,1) = \{z \in \mathbb{C}: \|z\| \leq 1\}$. The roots lying on the boundary $\partial B(0,1)$ of $B(0,1)$ are simple and all of them are r th roots of 1 for some r .*

Introduce the following notation:

$$\begin{aligned} p_i^{(r)} &= P\{\xi_r = i\}, i = 0, 1, \dots, P^{(r)} = \{p_0^{(r)}, p_1^{(r)}, \dots\}, r = 0, 1, \dots, \\ P^{(0)}(z) &= \sum_{j \geq 0} p_j^{(0)} z^j, u_i(x) = \sum_{r \geq 0} p_i^{(r)} x^r, U(x) = \{u_0(x), u_1(x), \dots\}, \end{aligned}$$

$$w(x, z) = \sum_{j \geq 0} u_j(x) z^j, \quad |x| < 1, \quad |z| \leq 1,$$

$$p_i = \lim_{r \rightarrow \infty} p_i^{(r)}, \quad P(z) = \sum_{j \geq 0} p_j z^j, \quad P = \{p_0, p_1, \dots\}.$$

We put some well-known relationships between vectors P , $P^{(r)}$, $U(x)$ and matrix A that will be used in the next section:

$$(3.6) \quad P^{(r+1)} = P^{(r)} A$$

As $\{\xi_n\}$ is ergodic (under the conditions in Theorem 3.1) the invariant probability measure P is a unique solution of the matrix equation

$$(3.7) \quad Y = YA, \quad Y1 = 1, \quad Y \in (l^1, \|\cdot\|)$$

Taking into account equations (3.6) and (3.7) we have

$$(3.8) \quad U(x) = xU(x)A + P^{(0)}$$

In the following section we will treat the generating functions $w(x, z)$ and $P(z)$.

4. TRANSIENT AND STATIONARY PROBABILITIES OF PROCESS (ξ_n)

4.1 Theorem. *The generating function $w(x, z)$ of the transient probabilities of the Markov chain $\{\xi_n\}$ with the transition probability $\Delta_{m, N}$ -matrix $A = (a_{ij})$ satisfies the following equations*

$$(4.1) \quad w(x, z) = \frac{x \sum_{i=0}^N u_i(x) [z^m A_i(z) - z^i K(z)] + z^m P^{(0)}(z)}{z^m - xK(z)}, \quad \|z\| \leq 1, \quad |x| < 1$$

$$(4.2) \quad \lim_{z \rightarrow \phi_s(x)} \frac{\partial^k}{\partial z^k} \left[\sum_{i=0}^N u_i(x) [z^i - xA_i(z)] - P^{(0)}(z) \right] = 0, \quad k = 0, 1, \dots, r_s - 1, \quad s = 1, \dots, S,$$

where $\phi_s(x)$ are the roots of function $z^{N+1} - m(z^m - xK(z))$ inside $B(0, 1)$ with their multiplicities r_s such that $\sum_{s=1}^S r_s = N + 1$. The system of equations (4.2) has a unique solution u_0, \dots, u_N for all x such that $0 < |x| < 1$.

Proof. From equation (3.8) we get (4.1) after elementary transformations. Then, by Rouché's theorem, for any fixed x from interval $(-1, 1)$ the function $z \mapsto z^{N+1} - xz^{N+1} - mK(z)$ has exactly $N + 1$ roots inside the unit ball $B(0, 1)$. Rewriting (4.1) in form

$$\left[w(x, z) - \sum_{i=0}^N u_i(x) z^i \right] z^{-(N+1)} = \frac{- \sum_{j=0}^N u_j(x) [z^j - xA_j(z)] + P^{(0)}(z)}{z^{N+1} - m[z^m - xK(z)]}$$

and observing that the function in the left-hand side of the last equation is analytic in $\bar{B}(0, 1)$ for every fixed $x \in (-1, 1)$ we get (4.2).

It only remains to prove that the system of equations (4.2) has a unique solution. Assume that $\mathbf{v} = (v_0, \dots, v_N)$ is another solution of (4.2). Consider w in (4.1) as an operator applied to a vector-function $\mathbf{u} = (u_0, \dots, u_N)$ (in notation $w(\mathbf{u})[x, z]$). $w(\mathbf{v})[x, z]$ is clearly analytic in z within the unit ball $B(0, 1)$ and continuous on the boundary $\partial B(0, 1)$ for every $|x| < 1$. As in Theorem 1,

Dukhovny [s], it can be shown that the coefficients $v_i^*(x)$ of the expansion of $w(v)$ in Maclaurin series in powers of z form a sequence V^* which is an element of $(l^1, \|\cdot\|)$ for every fixed $|x| < 1$. Now because of (4.1) $V^* = (v_0^*, v_1^*, \dots)$ satisfies (3.8) and due to (4.2) $v_i = v_i^*, i = 0, \dots, N$ (but its components certainly differ from $u_i, i = 1, \dots, N$). Thus, we have two different solutions U and V^* of (4.2) in l^1 which yields the following contradiction

$$\|U - V^*\| = \|x(U - V^*)A\| < \|(U - V^*)A\| \leq \|U - V^*\| \|A\| \leq \|U - V^*\| \quad \square$$

4.2 Theorem. *The invariant probability measure P of the Markov chain (ξ_n) with the transition probability $\Delta_{m,N}$ -matrix A exists if and only if $\rho < m$, where $\rho = \lambda ab_{N+1}$. Under this condition, the generating function $P(z)$ of the invariant probability measure P satisfies the following relations:*

$$(4.3) \quad P(z) = \frac{\sum_{i=0}^N p_i [z^m A_i(z) - z^i K(z)]}{z^m - K(z)}$$

$$(4.4) \quad \sum_{i=0}^N p_i \frac{d^k}{dz^k} [A_i(z) - z^i] \Big|_{z=z_s} = 0, \quad k = 0, \dots, r_s - 1, \quad s = 1, \dots, S$$

$$(4.5) \quad \sum_{i=0}^N p_i [A_i'(1) - i + m - \rho] = m - \rho$$

where $A_i(z)$ satisfy formulas (3.2)-(3.3) and z_s are the roots of the function $z^{N+1} - m[z^m - K(z)]$ inside the ball $B(0,1)$ not equal 1 with the multiplicities r_s , such that $\sum_{s=1}^S r_s = N$. The system of equations (4.4)-(4.5) has a unique solution p_0, \dots, p_N .

Proof. The necessary and sufficient condition for the ergodicity of (ξ_n) follows from Theorem 4.1, where the condition (3.4) is obviously satisfied and the condition $\rho < m$ follows directly from (3.7). Now taking into account (3.4) and (3.5a) from equation

$$\begin{aligned} E^x[z^{\xi_{n+1}}] &= E^x[z^{(\alpha_{\theta_n}^{(\xi_n)} - m(0))^+} z^{\nu_n} | \xi_n = 0] + \sum_{j \geq 1} E^x[z^{(\xi_n - m(\xi_n))^+} z^{\nu_n} | \xi_n = j] P^x\{\xi_n = j\} \\ &= \sum_{j \geq 0} A_j(z) P^x\{\xi_n = j\} \quad \text{we obtain formula (4.3).} \end{aligned}$$

Proceeding similarly as in Proof of Theorem 4.1 we get equations (4.4) and (4.5). As in Theorem 4.1, we can show that the non-uniqueness of the solution of the system of equations (4.4)-(4.5) would imply the non-uniqueness of the invariant probability measure P which is impossible after we meet the ergodicity condition. □

4.3 Example. We did not discuss how to evaluate the roots of the functions in Theorems 4.1 and 4.2 assuming that in general they can be found by various numerical methods. However, for a wide class of special systems it is possible to find the generating function $P(z)$ in an explicit form, avoiding numerical procedures. As an illustration to Theorems 4.1 and 4.2, consider a special case of the bulk queueing system under the assumption that the server takes for service all units available in the system provided it does not exceed N and a service is not initiated right after an idle period. In other words, assume that $m(j) = j, j = 1, \dots, N, m(j) = m, j = N + 1, N + 2, \dots$. The capacity of the

server at the beginning of a busy period $m(0)$ remains under the general assumptions. In addition, assume that $B_i = B_1$, $\lambda_i = \lambda_1$ and $a_i(z) = a_1(z)$, $i = 1, \dots, N$. Then, in (3.2a) $A_i(z)$ are reduced to $K_1(z)$, $i = 1, \dots, N$.

Let $\hat{K}_1(z)$ and $\hat{A}_0(z)$ be N th degree interpolating polynomials of $K_1(z)$ and $A_0(z)$, respectively, taking on the values of $K_1(z)$ and $A_0(z)$ and all its derivatives up to $(r_s - 1)$ st order at $z = z_s$, $s = 1, \dots, S$, and at $z = 1$. Then, denoting $c = p_1 + \dots + p_N$ we have

$$\sum_{j=0}^N p_j [A_j(z) - z^j] = p_0 A_0(z) + c K_1(z) - \sum_{j=0}^N p_j z^j$$

and from (4.4) it follows that

$$(4.6) \quad \sum_{j=0}^N p_j z^j = p_0 \hat{A}_0(z) + c \hat{K}_1(z)$$

Substituting (4.6) into (4.3) we obtain

$$(4.7) \quad P(z) = p_0 \hat{A}_0(z) + \hat{K}_1(z) + z^m \frac{p_0 [A_0(z) - \hat{A}_0(z)] + c [K_1(z) - \hat{K}_1(z)]}{z^m - K(z)}$$

To find the unknown constants p_0 and c we set up a system of two equations following the conditions $P(1) = 1$ and $P(0) = p_0$ and taking into account (4.6) and (4.7):

$$(4.8) \quad p_0 = p_0 \hat{A}_0(0) + c \hat{K}_1(0)$$

$$(4.9) \quad 1 = p_0 \left(1 + \frac{A_0'(1) - \hat{A}_0'(1)}{m - \rho} \right) + c \left(1 + \frac{\rho_1 - \hat{K}_1'(1)}{m - \rho} \right)$$

The system (4.8)-(4.9) along with (4.6) is equivalent to the system (4.4)-(4.5) in Theorem 4.2, thus it has a unique solution when $\rho < m$. Observe that the form of the polynomials $\hat{A}_0(z)$ and $\hat{K}_1(z)$ depends upon a relation between $m(0)$, N and m . For instance, if a function f is analytic at zero and thus $f(z) = \sum_{j \geq 0} f_j z^j$, and if \hat{f} is its interpolating polynomial defined in the same way as \hat{K}_1 and \hat{A}_0 for K_1 and A_0 , respectively, then \hat{f} can be represented in the form

$$\hat{f}(z) = f_0 + f_1 z + \dots + f_{N-m} z^{N-m} + z^{N+1-m} \hat{f}^*(z),$$

where $\hat{f}^*(z)$ is the interpolating polynomial of the series $f^*(z) = \sum_{j \geq N+1-m} f_j z^{j+m-N-1}$. It takes on the values of f^* and all its derivatives up to $(r_s - 1)$ st order at $z = z_s$, $s = 1, \dots, S$, and at $z = 1$, where z_s are the roots of $z^m - K(z)$ in $\bar{B}(0,1)$.

Let $m = 1$. Then

$$\hat{A}_0(z) = \sum_{i=0}^{N-1} a_{0i} z^i + z^N \sum_{i \geq N} a_{0i}, \quad \hat{K}_1(z) = \sum_{j=0}^{N-1} k_j^1 + z^N \sum_{j \geq N} k_j^1, \quad k_j^1 = D_j K_1(z),$$

where D_i is defined as

$$D_i \phi = \frac{1}{i!} \lim_{x \rightarrow 0} \frac{d^i}{dx^i} \phi(x).$$

From (4.6) we have in this case

$$p_i = p_0 a_{0i} + c k_i^1, \quad i = 0, 1, \dots, N-1; \quad p_N = p_0 \sum_{i \geq N} a_{0i} + c \sum_{i \geq N} k_i^1$$

while (4.9) yields

$$p_0^{-1} = 1 + \frac{\sum_{j > N} a_{0i}(i-N)}{m-\rho} + \frac{(1-a_{00})}{k_0^1} \left[1 + \frac{\sum_{j > N} (i-N)k_i^1}{m-\rho} \right]$$

Consider the system with $N = 1$ and $m = 2$. In this case the function $z^2 - K(z)$ has a unique root $z_1 \in B(0,1)$, thus

$$\hat{A}_0(z) = 1 + (z - 1)[A_0(z_1) - 1](z_1 - 1)^{-1}$$

$$\hat{K}_1(z) = 1 + (z - 1)[K_1(z_1) - 1](z_1 - 1)^{-1}$$

and $c = p_1$. The probabilities p_0 and p_1 can be found from (4.8) and (4.9):

$$p_1 = p_0 \frac{A_0(z_1) - 1}{z_1 - K_1(z_1)}$$

$$p_0^{-1} = 1 + (m - \rho)^{-1} \left(A_0'(1) - \frac{A_0(z_1) - 1}{z_1 - 1} \right) + \frac{A_0(z_1) - 1}{z_1 - K_1(z_1)} \left[1 + (m - \rho)^{-1} \left(\rho_1 - \frac{K_1(z_1) - 1}{z_1 - 1} \right) \right]$$

Another interesting particular case arises when $a_{0i} = 0, i > m(0)$, and $K_1(z) = K_0(z) = A_0(z)$. Then (4.6) turns into

$$\sum_{j=0}^N p_j z^j = \hat{K}_0(z)(p_0 + c)$$

while (4.7) reduces to

$$(4.10) \quad P(z) = (p_0 + c) \frac{z^m K_0(z) - K(z) \hat{K}_0(z)}{z^m - K(z)}$$

Now $p_0 + c$ is obtained from $P(1) = 1$:

$$(4.11) \quad p_0 + c = (m - \rho) \left(m - \rho + \rho_0 - \hat{K}_0'(1) \right)^{-1}.$$

Pollaczek-Khintchine formula follows from (4.10) and (4.11) for $K_0 = K, N = 0, m = 1$.

5. CONTINUOUS TIME PARAMETER PROCESS $\{\xi(t)\}$

Recall that the successive instants of the service completions $\{t_n; n \in \mathbb{N}_0\}$ are stopping times for $\{\xi(t)\}$ and that $\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, (\xi_n, t_n): n = 0, 1, \dots\} \rightarrow (Q \times \mathbb{R}_+, \mathfrak{B}(Q \times \mathbb{R}_+))$ is a Markov renewal process. In other words, t_n 's are the moments of "Markov regeneration" of $\{\xi(t)\}$ and thus it is a *semi-regenerative processes* relative to the point process $\{t_n\}$. The stationary probability measure π of $\{\xi(t)\}$ exists under the same ergodicity conditions as the invariant probability measure P for the embedded Markov chain (ξ_n) (as was stated in Theorem 3.1).

Introduce the following probabilities:

$$\delta_{ik}^0(t) = P^0\{\xi(t) = k \mid t_1 > \tau_1^{(0)} + t, \hat{N}_0(\tau_1^{(0)}) = i\} = P^0\{N^\xi(t) = k - i \mid \sigma_1 > t\}$$

$$\delta_{jk}(t) = P^j\{\xi(t) = k \mid t_1 > t\} = P^j\{N^\xi(t) = k - j \mid \sigma_1 > t\}, j > 0.$$

Clearly,

$$E^0[z^{\xi(t)} \mid t_1 > \tau_1^{(0)} + t, \hat{N}_0(\tau_1^{(0)}) = i] = \sum_{k \geq i} z^{k-i} \delta_{ik}^0(t) = \exp[-\lambda_0(1 - a_0(z))t]$$

$$E^j[z^{\xi(t)} \mid t_1 > t] = \sum_{k \geq j} z^{k-j} \delta_{jk}(t) = \exp[-\lambda_j(1 - a_j(z))t], j > 0.$$

For each j and k denote $K^j(k, t) := P^j\{\xi(t) = k, t_1 > t\}$. We will call the matrix

$$(K^j(k, t); (j, k) \in Q \times Q)$$

a *semi-regenerative kernel*. Obviously,

$$K^j(k, t) = \begin{cases} e^{-\lambda_0 t}, & j = k = 0 \\ \sum_{i=1}^k a_i^0 \int_0^t \lambda_0 \exp[-\lambda_0(t-u)] \delta_{ik}^0(u) (1 - B_0(u)) du, & j = 0, k > 0 \\ \delta_{jk}(t) [1 - B_j(t)], & 0 \leq j \leq k \\ 0, & 0 \leq k < j, j > 0, \end{cases}$$

Therefore, the generating function of the j th row of the semi-regenerative kernel is

$$(5.1) \quad \sum_{k \geq 0} K^0(k, t) z^k = e^{-\lambda_0 t} [1 + \lambda_0 a_0(z) \int_0^t e^{\lambda_0 t a_0(z) u} (1 - B_0(u)) du]$$

$$(5.1a) \quad \sum_{k \geq 0} z^k K^j(k, t) = z^j (1 - B_j(t)) \exp[-\lambda_j (1 - a_j(z)) t], \quad j > 0.$$

The direct integrability of each element of the semi-regenerative kernel $(K^j(k, t), j, k = 0, 1, \dots)$ suffices for the existence of the steady state probabilities $\pi_k, k = 0, 1, \dots$, of the queueing process $\{\xi(t)\}$ in formula (5.2) below. We may apply the Main Convergence Theorem for semi-regenerative processes (cf. Çinlar [6]). Denoting by H the integrated semi-regenerative kernel $(\int_0^\infty K^j(k, t) dt, (j, k) \in Q \times Q)$ we have the equation for stationary probability measure π of the process $\{\xi(t)\}$ in matrix form

$$(5.2) \quad \pi = \frac{PH}{P\beta}$$

where $\beta = (b_0 + \frac{1}{\lambda_0}, b_1, b_2, \dots)^T, b_j = E^j[\sigma_1] = \int_0^\infty x B_j(dx), j = 0, 1, \dots$. Observe that $P\beta$ can be interpreted as the expected stationary “inter-service” time.

Now from (5.1) and (5.1a) we find the generating function $h_j(z)$ of the j th row of the integrated semi-regenerative kernel H :

$$(5.3) \quad h_0(z) = \frac{1}{\lambda_0} + a_0(z) \frac{1 - \beta_0(\lambda_0 - \lambda_0 a_0(z))}{\lambda_0(1 - a_0(z))}$$

$$(5.3a) \quad h_j(z) = z^j \frac{1 - \beta_j(\lambda_j - \lambda_j a_j(z))}{\lambda_j(1 - a_j(z))}, \quad j > 0$$

Finally, using (5.3) and (5.3a) we obtain from (5.2) a relation between the generating function $\pi(z)$ of the stationary probability measure π and the generating function $P(z)$ of the invariant probability measure P .

$$(5.4) \quad \pi(z) = \sum_{k \geq 0} \pi_k z^k = \frac{1}{P\beta} \left\{ p_0 \frac{1}{\lambda_0} K_0(z) + \sum_{j=0}^N p_j z^j (g_j(z) - g(z)) + g(z) P(z) \right\}$$

where

$$(5.5) \quad g_j(z) = \frac{1 - K_j(z)}{\lambda_j(1 - a_j(z))}, \quad j = 0, 1, \dots, N,$$

$$(5.6) \quad g_j(z) = g(z) = g_{N+1}(z), \quad j = N+1, N+2, \dots$$

and

$$(5.7) \quad P\beta = \sum_{j=0}^N (b_j - b) p_j + b + p_0 \frac{1}{\lambda_0},$$

where $N + 1$ probabilities p_j in the right-hand side of (5.7) are determined from equations (4.4-4.5). The above results can be summarized as the following statement.

5.1 Theorem. *The steady state of the queueing process $\{\xi(t); t \geq 0\}$ exists if and only if $\rho < m$ and, under this condition, the limiting distribution in form of the generating function π satisfies formulas (5.4-5.7).*

We introduce the following notions.

5.2 Definitions.

(i) Let $\bar{\rho} = (\rho_0 + a^{[0]}, \rho_1, \rho_2, \dots)^T$. Then we call $P\bar{\rho}$ the (stationary expected) capacity of the system.

(ii) Let $M_j = E^x[M(\xi_n) | \xi_n = j]$ be the expected number of units taken for service given that after a preceding completion of service there were j units in the system. If $\bar{M} = (M_0, M_1, \dots)^T$ then the value $P\bar{M}$ gives the stationary expected number of units taken for service. We call $P\bar{M}$ the (stationary expected) capacity of the server.

The following statement will be used throughout the remainder of this paper:

5.3 Theorem. *Given the equilibrium condition $\rho < m$, the capacity of the system and the capacity of the server are equal.*

Proof. From (2.1) we have by elementary probability arguments:

$$(5.8) \quad M_j = \begin{cases} \sum_{i=1}^{m(0)} i a_i^0 + \sum_{i=m(0)+1}^{\infty} m(0) a_i^0, & j = 0 \\ \min(j, m(j)), & j = 1, \dots, N \\ m, & j > N. \end{cases}$$

Now the statement follows from (4.3) by using the identity $P(1) = 1$. □

5.4 Examples.

(i) It follows from (5.4-5.6) that

$$(5.9) \quad \pi_0 = \frac{p_0}{\lambda_0 P \beta}$$

(ii) Let \mathfrak{B} denote the expected value of a busy period in the equilibrium. Taking into account that the expected length of an idle period is $\frac{1}{\lambda_0}$, we obtain another expression for π_0 :

$$\pi_0 = \frac{\frac{1}{\lambda_0}}{\frac{1}{\lambda_0} + \mathfrak{B}}$$

Then, from the last equation and (5.9) we get a formula for the expected length of a busy period in the steady state

$$(5.10) \quad \mathfrak{B} = \frac{P\beta}{p_0} - \frac{1}{\lambda_0}$$

(iii) If the input stream is independent of the states of the system we have from (4.3) and (5.4-5.6):

$$(5.11) \quad \begin{aligned} & \lambda[1 - a(z)]P\beta \pi(z) \\ & = p_0 K_0(z)[1 - a(z)] + P(z)(1 - z^m) + \sum_{j=0}^N p_j [z^m A_j(z) - z^j K_j(z)] \end{aligned}$$

(iv) Under the additional assumptions as those in Example 4.3, we can use p_0 and $\sum_{j=0}^N p_j$ (evaluated there for various special cases) needed for (5.10) and (5.11). This also takes place in (vi), formula (5.15) below.

(v) Under the condition of (iii) assume that $m(0) = 1$, i.e. after an idle period the server is in a "warm-up mode" by reducing its initial capacity. Then, from (5.8) we have $M_0 = 1$ and by Theorem 5.3 it follows that

$$(5.12) \quad P\bar{p} = \lambda a P\beta = P\bar{M} = p_0 + \sum_{j=1}^N p_j \min(j, m(j)) + m \sum_{j \geq N+1} p_j$$

(vi) Assuming in addition $m(j) = 1$, $j = 0, 1, 2, \dots$, (i.e. there is no batch service) we obtain from (5.12)

$$(5.13) \quad P\beta = \frac{1}{\lambda a}$$

Equations (5.11), (5.13) and Theorem 5.3 lead to the following elegant formula

$$(5.14) \quad \pi(z) = \frac{a(1-z)P(z)}{1-a(z)}$$

Recall that the ergodicity condition in this case is $\rho < 1$. This is a formula of Pollaczek-Khintchine type. It reduces to the "Pollaczek-Khintchine relation" when $a(z) = z$ shown by Schäl [16] for a similar system with state dependent service (note that Schäl did not give a formula for $\pi(z) = P(z)$ but just stated this relation). Formula (5.10) for the expected value of a busy period is then reduced to

$$(5.15) \quad \mathfrak{B} = \frac{1 - ap_0}{\lambda ap_0}$$

(vii) Assuming further that the group size is distributed geometrically, $a_i = pq^{i-1}$, $i = 1, 2, \dots$, we obtain from (5.14)

$$\pi(z) = P(z) \frac{1 - qz}{p}$$

that enables us to obtain an explicit relation between π and P :

$$p \pi_k = \begin{cases} p_0, & k = 0 \\ p_k - qp_{k-1}, & k \geq 1 \end{cases}$$

6. CONVERGENCE THEOREMS FOR SOME FUNCTIONALS OF INPUT, OUTPUT AND QUEUEING PROCESSES

In this section we will treat the input, output and queueing processes that are special in cont-

rolled systems. Observe that for a standard compound Poisson process $\{N(t)\}$ with rate λ and batch size average a its renewal function $E[N(t)]$ equals λat . In our case, since the rates of the input process $\{N^\xi(t)\}$ are constant only within each random interval (t_n, t_{n+1}) for every $j \in Q$, what will the corresponding function $E[N^\xi(t)]$ be for the input process controlled by the queueing process $\{\xi(t)\}$? The output rate of processed units is another valuable characteristics of the system. As regards the queueing process $\{\xi(t)\}$ we will be interested in the relative speed of convergence (or divergence) of its mean for large t .

First we will introduce some auxiliary processes and their functionals. For the Markov renewal process $\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, (\xi_n, t_n): n = 0, 1, \dots\} \rightarrow (Q \times \mathbb{R}_+, \mathfrak{B}(Q \times \mathbb{R}_+))$ consider for each x and j the Markov renewal function

$$t \mapsto R^x(j, t) = E^x \left[\sum_{n=0}^{\infty} I_{\{j\} \times [0, t]} \circ (\xi_n, t_n) \right]$$

which gives the expected number of entrances of the embedded Markov chain (ξ_n) in state $\{j\}$ during time interval $[0, t]$ given that the process started from state $\{x\}$. Another process of interest (which we introduced in section 2) is the minimal semi-Markov process

$$\{\Omega, \mathfrak{F}, (P^x)_{x \in Q}, (\zeta(t): t \geq 0)\} \rightarrow (\mathbb{R}_+, \mathfrak{B}_+)$$

associated with the above Markov renewal process. The process $\{\zeta(t)\}$ is right continuous with almost every path as a simple function on any compact interval. We formulate and prove below some limit theorems for functionals of the Markov renewal process, queueing process and semi-Markov process. Note that these results hold true for general semi-regenerative processes.

6.1 Lemma. *If $\rho < m$ the limits in (i-iii) exist and satisfy the given below formulas:*

- (i) $\lim_{t \rightarrow \infty} \frac{1}{t} R^x(j, t) = \frac{p_j}{P\beta}$
- (ii) $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t P^x \{\xi(u) = j\} du = \pi_j$
- (iii) $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t P^x \{\zeta(u) = j\} du = \frac{p_j b_j}{P\beta} I_{\{1, 2, \dots\}}(j) + \frac{p_0(b_0 + \frac{1}{\lambda_0})}{P\beta} I_{\{0\}}(j)$

Proof:

(i) Let $\gamma_0 = \inf\{n > 0: \xi_n = j, \xi_0 = x\}$, $\gamma_k = \inf\{n > \gamma_{k-1}: \xi_n = j, \xi_0 = x\}$. Then $\{t_{\gamma_k}; k \geq 0\}$ is a delayed renewal process (embedded in the point process $\{t_n\}$). Clearly, $\{t_{\gamma_k}; k \geq 0\}$ is recurrent if and only if $\rho < m$. Then it follows from Çinlar [6] and due to $\lim_{t \rightarrow \infty} P\{\gamma_0 \leq t\} = 1$ for $\rho < m$ that

$$\lim_{t \rightarrow \infty} \frac{R^x(j, t)}{t} = \mu_j,$$

where μ_j is the inverse of the expected time between two subsequent returns to state $\{j\}$. On the other hand, from Markov renewal theory it is known that

$$\mu_j = \frac{p_j}{P\beta}$$

and the statement (i) then follows.

(ii) By Çinlar [6] we have

$$P^x\{\xi(u) = k\} = \sum_{j \geq 0} \int_0^u R^x(j, ds) K^j(k, t)$$

resulting

$$\int_0^t P^x\{\xi(u) = k\} du = \sum_{j \geq 0} R^x(j, \cdot) * g(t)$$

where $g(t) := \int_0^t K^j(k, v) dv$ is a non-decreasing continuous function and symbol “*” denotes the convolution operator. Then, it follows that

$$\lim_{t \rightarrow \infty} \frac{R^x(j, \cdot) * g(t)}{R^x(k, t)} = \frac{p_j}{p_k} \int_0^\infty K^j(k, u) du$$

Now applying (i) and formula (5.2) we finally obtain

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t P^x\{\xi(u) = j\} du = \lim_{t \rightarrow \infty} \frac{\sum_{j \geq 0} R^x(j, \cdot) * g(t)}{R^x(k, t)} \frac{R^x(k, t)}{t} = \pi_k$$

(iii) The statement follows directly from Çinlar [6]. \square

Now we will consider the input process whose both interarrival rate and group size depend on the Markov renewal process $\{\xi_n, t_n\}$. An approach is based on the treatment of an auxiliary process that is stochastically equivalent to the input process $\{N^\xi(t)\}$. Note that the direct treatment of the input process (in the form defined in section 2) is not effective. Since the parameters of the input process $\{N^\xi(t)\}$ are constant only within each random interval (t_n, t_{n+1}) , for every $j \in Q$ we need to know the total length of all those random intervals (the whole or in parts) within $[0, t]$ on which the queueing process $\xi(t)$ takes on value j . Denote this stochastic process by $T_t^{\{j\}}$. Observe that interval $[0, t]$ can be restored by summing $T_t^{\{j\}}(\omega)$ over $j \in Q$ for any fixed ω .

If $\{\hat{N}_j(t)\}$ is a compound Poisson process with rate λ_j and batch size average $a^{[j]}$, as introduced in section 2, then the process $\{\hat{N}_j(T_t^{\{j\}})\}$ gives the total number of units arrived in $[0, t]$ and it is stochastically equivalent to the partial input process $\{N_j^\xi(t)\}$. Thus, the processes $\{N^\xi(t)\}$ and $\{\sum_{j \geq 0} \hat{N}_j(T_t^{\{j\}})\}$ are stochastically equivalent.

We will treat the process $\{N^\xi(t)\}$ and then evaluate the expected number of units arriving in time interval $(0, t]$. Let $\Phi^x(j, t, v) := P^x\{T_t^{\{j\}} \leq v\}$. Then, the functional

$$(6.1) \quad T^x(j, t) = \int_0^t v \Phi^x(j, t, dv)$$

gives the total expected time of the semi-Markov process $\{\zeta(t)\}$ spent in state $\{j\}$ during time interval $(0, t]$ (or in other words, this is the time that the server spends in mode $\{j\}$ processing units in time interval $(0, t]$).

6.2 Lemma. *The following relation holds true:*

$$(6.2) \quad E^x[N^\xi(t)] = \sum_{j \geq 0} \lambda_j a^{[j]} T^x(j, t)$$

Proof. We have from (6.1):

$$\begin{aligned} E^x[\hat{N}_j(T_t^{\{j\}})] &= \int_0^t E^x[\hat{N}_j(T_t^{\{j\}}) | T_t^{\{j\}} = v] \Phi^x(j, t, dv) \\ &= \lambda_j a^{[j]} \int_0^t v \Phi^x(j, t, dv) = \lambda_j a^{[j]} T^x(j, t) \end{aligned}$$

and, as mentioned,

$$N^\xi(t) = \sum_{j \geq 0} \hat{N}_j(T_t^{\{j\}}) \quad P^x\text{-a.s.}$$

that gives the total number of units arrived in $(0, t]$. Now the statement follows by the Monotone Convergence Theorem. \square

One of the problems we set at the beginning of this section was to find $\lim_{t \rightarrow \infty} \frac{E^x[\xi(t)]}{t}$. By a direct computation, it can be shown that for $\rho < m$ the value of $\lim_{t \rightarrow \infty} E^x[\xi(t)]$ is a function of the second moment of service time that need not be finite. In the latter case, it is not obvious with what speed $\lim_{t \rightarrow \infty} E^x[\xi(t)]$ gets to infinity. We will show that, even if it diverges, it gets slower to infinity than with the unit speed. To solve this problem we will need to analyze a functional of the output process which is also of interest.

6.3 Theorem. *Let $S(t)$ denote the total number of served units in time interval $[0, t]$. Then for $\rho < m$ the units are processed with the expected rate equals the ratio of the capacity of the system and the capacity of the inter-service time:*

$$(6.3) \quad \lim_{t \rightarrow \infty} \frac{E^x[S(t)]}{t} = \frac{P\bar{\rho}}{P\bar{\beta}}$$

Proof. The functional $S(t)$ of the output process giving the total number of completely processed units in time interval $[0, t]$ obviously satisfies the following relation

$$\begin{aligned} S(t) &= \sum_{n=0}^{\infty} M(\xi_n) I_{\{0\} \times [0, t]} \circ (\xi_n, t_n) + \sum_{j \geq 1} M(j) \sum_{n=0}^{\infty} I_{\{j\} \times [0, t]} \circ (\xi_n, t_n) - M(\zeta(t)) \\ &\quad - M(\xi(0)) \end{aligned}$$

(as the total number of all units taken for service less the last batch of units that is being in service), where by (2.2) we have the equivalent expression for $M(\zeta(t))$:

$$M(\zeta(t)) = \min\{\hat{N}_0(\tau_{\vartheta_t}^\xi), m(0)\} I_{\{0\}} \circ \zeta(t) + \min(\zeta(t), m(\zeta(t))) I_{\{1, 2, \dots\}} \circ \zeta(t).$$

The conditional expectation of $S(t)$, given the initial measure ϵ_x can be obtained in form of the following expression

$$(6.4) \quad E^x[S(t)] = \sum_{j \geq 0} M(j) R^x(j, t) - \sum_{j \geq 0} M(j) P^x\{\zeta(t) = j\} - M(x)$$

due to routine probability arguments and by the Monotone Convergence Theorem. Applying Lemma 6.1 (i) to (6.4) and observing that the second sum in (6.4) is finite for each t we find that the units are processed with the rate

$$\lim_{t \rightarrow \infty} \frac{E^x[S(t)]}{t} = \sum_{j \geq 0} M(j) \frac{p_j}{P\bar{\beta}} = \frac{P\bar{M}}{P\bar{\beta}}$$

Finally, equation (6.3) is due to Theorem 5.3. \square

6.4 Theorem. *The expected rate of the input flow is*

$$(6.5) \quad \lim_{t \rightarrow \infty} \frac{E^x[N^\xi(t)]}{t} = \frac{P\bar{\rho}}{P\beta}$$

Proof. From the definition of $T_t^{\{j\}}$ it also follows that

$$(6.6) \quad T_t^{\{j\}} = \int_{0+}^t I_{\{j\}} \circ \zeta(u) du = \int_0^t I_{\{j\}} \circ \zeta(u) du.$$

Applying Fubini's theorem in (6.6) we get another representation for T^x :

$$(6.7) \quad T^x(j, t) = \int_0^t P^x\{\zeta(u) = j\} du.$$

The statement follows from Lemmas 6.2 and 6.1 (iii) applied to (6.7). \square

6.5 Corollary. *For $\rho < m$ the expected number of units in the system in equilibrium is either finite or diverges slower than with the unit speed. In other words, $\lim_{t \rightarrow \infty} \frac{E^x[\xi(t)]}{t} = 0$.*

Proof. Observing that the number of units in the system at time t is $\xi(t) = \xi(0) + N^\xi(t) - S(t)$ the statement follows from Theorems 6.3 and 6.4. \square

6.6 Examples.

(i) A trivial special case follows from Theorem 6.4 when the input stream is independent of the queue. The counting process $\{N^\xi(t)\}$ is now the compound Poisson process $\{N(t)\}$ with parameters (λ, a) and the expected rate then should be λa . On the other hand, from (5.12) we have $P\bar{\rho} = \lambda a P\beta$ which yield the same result λa in the right-hand side of (6.5). Note that by Theorem 6.3, λa is also the expected rate of the number of processed units.

(ii) As an application of the above ergodic theorems, we consider the following optimization problem. Let c_1, c_2, c_3, r be real-valued Borel-measurable functions that represent the following cost rates:

$c_1(k)$ denotes the total expenses due to the presence of k units in the system per unit time;

$c_2(j)$ denotes the expenses of the service act of type j per unit time [observe that the decision to "apply a certain distribution function B_j " when the system accumulated j units, will be affected by the cost function c_2 that is usually inverse proportionally to the service rates];

c_3 is the penalty for every idle period per unit time;

r denotes the reward for each completely processed unit per unit time.

In connection with the above cost functions, we introduce the following functionals (whenever the explicit integral is given, the integration with respect to Lebesgue measure is understood):

$F_1[c_1, \xi](x, t) = E^x[\int_0^t c_1(\xi(u)) du]$ denotes the expected expenses due to the presence of all customers in the system in time interval $[0, t]$ given that initially x units were present;

$F_2[c_2, \zeta](x, t) = E^x[\int_0^t c_2(\zeta(u))du]$ denotes the expected expenses of all service acts in time interval $[0, t]$ given that initially x units were present in the system (in particular, the system is penalized by function c_2 , for the total time being idle in interval $[0, t]$ forcing the system to reduce the interarrival rate $\frac{1}{\lambda_0}$);

$c_3 R^x(0, t)$ denotes the expected expenses for all idle periods in time interval $[0, t]$ (in addition, c_3 penalizes the system for too many idle periods thereby forcing the system to reduce their number by taking "preventive measures").

Now by Fubini's Theorem and Lemma 6.1 (ii) we have

$$\lim_{t \rightarrow \infty} \frac{1}{t} F_1[c_1, \xi](x, t) = \sum_{j \geq 0} c_1(j) \pi_j = \pi c_1$$

as the expected cost rate due to the presence of all units in the system. Similarly, by Fubini's Theorem and Theorem 6.2 (iii) we have

$$\lim_{t \rightarrow \infty} \frac{1}{t} F_2[c_2, \zeta](x, t) = \frac{1}{P\beta} \sum_{j \geq 0} c_2(j) p_j b_j = \frac{P\overline{\beta c_2}}{P\beta}$$

(where $\overline{\beta c_2}$ denotes $(b_0 c_2(0), b_1 c_2(1), \dots)^T$) as the expected rate for all service acts over infinite horizon. By Lemma 6.1 (i) the penalty rate for idle periods is

$$\lim_{t \rightarrow \infty} c_3 \frac{1}{t} R^x(0, t) = \frac{c_3 p_0}{P\beta}$$

Finally, the expected gain of the system per unit time is derived from Theorem 6.4:

$$\lim_{t \rightarrow \infty} r \frac{1}{t} S(t) = r \frac{P\overline{\rho}}{P\beta}$$

The objective function is then

$$\frac{1}{P\beta} (r P\overline{\rho} - P\overline{\beta c_2} - c_3 p_0) - \pi c_1$$

REFERENCES

- [1]. Abolnikov, L.M., Investigation of a class of discrete Markov processes, *Eng. Cybernetics* (Izvestia of the Soviet Acad. Sci.), 15, No.2, 51-63, 1977.
- [2]. Abolnikov, L.M. and Dukhovny A.M., Necessary and sufficient conditions for the ergodicity of Markov chains with transition $\Delta_{m,n}$ ($\Delta'_{m,n}$) matrix, *Journ. Appl. Math. Simul.*, 1, No. 1, 13-24, 1987.
- [3]. Abolnikov, L.M. and Dshalalow J., Feedback queueing systems; duality principle and optimization, *Autom. and Remote Control* (Izvestia of the Soviet Acad. Sci.), 39, No.1, 11-20, 1978.
- [4]. Bachary, E., Kolesar, P., Multilevel bulk service queue, *Op. Res.*, 20, No.2, 1972.

- [5]. Chaudhry, M.L. and Templeton, J.G.C., *A First Course in Bulk Queues*, John Wiley and Sons, New York, 1983.
- [6]. Çinlar, E., *Introduction to Stochastic Processes*, Prentice Hall, Englewood Cliffs, N.J., 1975.
- [7]. Deb, R., Rajat, K., Optimal control of bulk queues with compound Poisson arrivals and batch service, *Op. Res.*, **21**, No.4, 227-245, 1984.
- [8]. Deb, R., Serfoso, R., Optimal control of batch service queues, *Advances in Appl. Probab.*, **5**, 340-361, 1973.
- [9]. Delbrouck, L., A feedback queueing system with batch arrivals, bulk service and queue-dependent service time, *Journ. Assoc. Comp. Mach.*, **17**, No.2, 314-323, 1970.
- [10]. Dukhovny, A.M., Markov chains with Quasitoeplitz transition matrix, *Journ. Appl. Math. Simul.*, **2**, No.1, 71-82, 1989.
- [11]. Federgruen, A., Tijms, H.C., Computation of the stationary distribution of the queue size in an $M | G | 1$ queueing system with variable service rate, *Journ. Appl. Probab.*, **17**, 515-522, 1980.
- [12]. Finch, P., A probability limit theorem with application to a generalization of queueing theory, *Acta Math. Acad. Sci. Hung.*, **10**, No. 3-4, 317-319, 1959.
- [13] Harris, C., Queues with state dependent stochastic service rates, *Op. Res.*, **14**, 117-130, 1967.
- [14] Mattias, M.B., A queueing system with a queue length dependent service, *Revista de Cienc. Mat.*, Ser. A, **3**, 19-32, 1972.
- [15] Shachbazov, A., A queueing system with warm-up and service time dependent on queue length, *Izvestia Acad. Sci. of Az.SSR*, Ser. Phys.-Techn. and Math. Sci., No.1, 32-35, 1974.
- [16] Schäl, M., The analysis of queues with state-dependent parameters by Markov renewal processes, *Advances in Appl. Prob.*, **3**, 155-175, 1971.