No. 1023

Processor-Sharing and Random-Service Queues with Semi-Markovian Arrivals

by

De-An Wu and Hideaki Takagi

February 2003

Processor-Sharing and Random-Service Queues with Semi-Markovian Arrivals

De-An Wu
Doctoral Program in Policy and Planning Sciences, University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305-8573, Japan
phone: +81-29-853-5167; fax: +81-29-853-5167
e-mail: tbu@sk.tsukuba.ac.jp

and

Hideaki Takagi
Vice President, University of Tsukuba
1-1-1 Tennoudai, Tsukuba-shi, Ibaraki 305-8577, Japan
phone: +81-29-853-2005; fax: +81-29-853-6310
e-mail: takagi@sk.tsukuba.ac.jp

Abstract

We consider single-server queues with exponentially distributed service times in which the arrival process is governed by a semi-Markov process (SMP). Two service disciplines, processor sharing (PS) and random service (RS), are investigated. We note that the sojourn time distribution of a type l customer who meets upon his arrival k customers present in the SMP/M/1/PS queue is identical with the waiting time distribution of a type l customer who meets upon his arrival $k+1$ customers present in the SMP/M/1/RS queue. The Laplace-Stieltjes transforms of the sojourn time distribution for an arbitrary customer in the SMP/M/1/PS queue and the waiting time distribution for an arbitrary customer in the SMP/M/1/RS queue are derived. We also consider a special case of the SMP in which the interarrival time distribution is determined only by the type of the next arrival.

Key words: Queues; semi-Markov arrival process; processor sharing; random service; sojourn time; waiting time

1 Introduction

We study queueing systems with a single server in which the arrival process is governed by a semi-Markov process (SMP). The service time follows exponential distribution and the capacity of the waiting room is infinite. Two service disciplines are considered: (i) processor sharing (PS), i.e., when there are k customers in the system each receives service at rate $1/k$, and (ii) random service (RS), i.e., when the server becomes available, the next customer to enter service is chosen at random among all waiting customers. The systems described above are denoted by SMP/M/1/PS and SMP/M/1/RS, respectively, throughout the paper.

The processor sharing discipline is the limiting case of the round robin discipline as the quantum of service time approaches 0; it is a reasonable service discipline in the performance modeling of computer and communication systems. Since Coffman, Muntz and Trotter [2] first analyzed an M/M/1/PS queue, several queueing systems with processor sharing service have been studied, for example, an M/G/1/PS queue [9, 13], a GI/M/1/PS queue [6, 10] and a GI/G/1/PS queue [11]. A survey of works on processor-sharing queues prior to 1987 is given by Yashkov [14], who cites many other references. To the best of our knowledge, however, there are no studies on queueing systems with semi-Markov arrivals and processor sharing discipline.

Ramaswami [10] finds the first two moments of the sojourn time distribution in a GI/M/1/PS queue (he notes an error in [2]). Cohen [4] points out that the sojourn time distribution of a customer who meets upon his arrival k customers present in a GI/M/1/PS queue is identical with the waiting time distribution of a customer who meets upon his arrival $k+1$ customers present in a GI/M/1/RS queue. We note that the same relation exists between the SMP/M/1/PS queue and the SMP/M/1/RS queue. This is the motivation that we analyze two queueing systems SMP/M/1/PS and SMP/M/1/RS together in this paper.

The rest of this paper is organized as follows. In Section 2 we define the semi-Markov arrival process and review some results about the queue size distribution before arrivals in SMP/M/1 queues. In Section 3, the Laplace-Stieltjes transform of the sojourn time distribution for an arbitrary customer in the SMP/M/1/PS queue is derived. The waiting time distribution of an arbitrary customer in the SMP/M/1/RS queue is analyzed in Section 4. A special case of the SMP is investigated in Section 5.

2 Preliminaries

2.1 Semi-Markov Arrival Process

The semi-Markov arrival process can be described as follows [1]. There are L types of customers numbered 1 through L. Customers arrive at time epochs $0 = T_0 < T_1 < T_2 < \dots$. Then $A_n := T_n - T_{n-1}$, $n > 1$, is the interarrival time, and we set $A_0 := 0$. Let $S^{(n)}$ denote the type of a customer arriving at epoch T_n. For a given sequence of arrival epochs, all interarrival times are mutually independent. It is assumed that A_{n+1} and $S^{(n+1)}$ depend only on $S^{(n)}$, i.e.,

$$P\{S^{(n+1)} = l, A_{n+1} \le t | S^{(0)}, \dots, S^{(n)}, A_1, \dots, A_n\} = P\{S^{(n+1)} = l, A_{n+1} \le t | S^{(n)}\}; \quad l = 1, \dots, L\,;\ t \ge 0. \quad (1)$$

Let

$$a_{lm}(t) := P\{S^{(n+1)} = m, A_{n+1} \le t | S^{(n)} = l\}$$

be the probability that the arrival process moves from state l to state m in time t. We note that $a_{lm}(\infty)$ is the probability that the arrival of a type l customer is followed by the arrival of a type m customer. Let us introduce the Laplace-Stieltjes transform (LST) of $a_{lm}(t)$ by

$$\alpha_{lm}(s) := \int_0^\infty e^{-st} da_{lm}(t).$$

Hereafter we use matrix $\mathbf{A}(s) := \{\alpha_{lm}(s)\}$. It is noted that $\mathbf{A}(0) = \{a_{lm}(\infty)\}$ is a stochastic matrix; thus

$$\mathbf{A}(0)\mathbf{1} = \mathbf{1},$$

where $\mathbf{1} := [1, \dots, 1]^{\mathrm{T}}$ (T denotes transpose). If $\boldsymbol{\pi} := [\pi_1, \dots, \pi_L]$ is the stationary distribution of the stochastic matrix $\mathbf{A}(0)$, we have

$$\boldsymbol{\pi}\mathbf{A}(0) = \boldsymbol{\pi}, \quad \boldsymbol{\pi}\mathbf{1} = 1. \tag{2}$$

Without much loss of generality, we assume that the Markov chain $\{S^{(n)}; n = 0, 1, 2, \dots\}$ is ergodic. For a real number $s \geq 0$, if $\lambda_{\mathrm{M}}(s)$ denotes the eigenvalue of the matrix $\mathbf{A}(s)$ with the maximum absolute value, then [1, eq.(9)]

$$\alpha = -\frac{d}{ds}\lambda_{\mathrm{M}}(s)\bigg|_{s=0+} \tag{3}$$

is the mean interarrival time defined by

$$\alpha := \sum_{l=1}^{L} \pi_l \sum_{m=1}^{L} \int_0^\infty t da_{lm}(t). \tag{4}$$

2.2 Queue Size Distribution Immediately before Arrivals in SMP/M/1 Queues

Similar to GI/M/1 queues [10], due to the memoryless property of exponentially distributed service times, the queue size distribution immediately before arrivals in the SMP/M/1/PS queue is the same as that in the corresponding SMP/M/1 queue with FIFO service discipline. Here we review some results about the SMP/M/1/FIFO queue from [1], which will be used in studying the sojourn time distribution in the SMP/M/1/PS queue. Let μ be the service rate.

The following result is a special case of the theorem by Çinlar [1], which is fundamental to analyze SMP/M/1 queues.

Theorem 1 *The equation*

$$\det[z\mathbf{I} - \mathbf{A}(s + \mu - \mu z)] = 0 \tag{5}$$

has exactly L roots, $\gamma_1(s), \gamma_2(s), \dots, \gamma_L(s)$ within the unit circle $|z| = 1$ if $\Re(s) > 0$, where $\mathbf{I}$ denotes an $L \times L$ identity matrix.

This theorem is the matrix version of Lemma 1 in Takács [12, p.113] for a GI/M/1 queue, and it can be proved by application of permutation theory and Rouché's theorem [7, 8]. We denote the distinct roots of (5) by $\gamma^{(1)}(s), \gamma^{(2)}(s), \dots, \gamma^{(M)}(s)$ with $M \leq L$.

For analyzing SMP/M/1 queues, we impose the following assumption which is the same as the one in [1].

Assumption 1 *All the elementary divisors* [5, p.142] *of the matrix* $\mathbf{A}[s+\mu-\mu\gamma^{(i)}(s)]$ *corresponding to the eigenvalue* $\gamma^{(i)}(s)$ *are of the first degree,* $i=1,\ldots,M$.

We note that the multiple eigenvalues are not ruled out. The matrix $\mathbf{A}[s+\mu-\mu\gamma^{(i)}(s)]$ may have elementary divisors not of the first degree if they correspond to other eigenvalue than $\gamma^{(i)}(s)$. If we denote by $\mathbf{g}_i(s)$ the left-hand eigenvector corresponding to the eigenvalue $\gamma_i(s)$, we have

$$\mathbf{g}_i(s)\{\gamma_i(s)\mathbf{I}-\mathbf{A}[s+\mu-\mu\gamma_i(s)]\}=0; \quad i=1,\ldots,L. \tag{6}$$

We collect all the eigenvectors corresponding to the eigenvalues $\gamma_1(s),\ldots,\gamma_L(s)$ into an $L\times L$ matrix $\mathbf{G}(s)$, i.e.,

$$\mathbf{G}(s):=[\mathbf{g}_1(s),\mathbf{g}_2(s),\ldots,\mathbf{g}_L(s)]^{\mathrm{T}}. \tag{7}$$

Let $\mathbf{\Gamma}(s)$ denote an $L\times L$ diagonal matrix with diagonal elements $\gamma_1(s),\gamma_2(s),\ldots,\gamma_L(s)$.

We next consider the Markov chain $\{(X^{(n)},S^{(n)});n=0,1,2,\ldots\}$, where $X^{(n)}$ denotes the number of customers seen by nth SMP arrival. The following result is cited form Çinlar [1, Theorem 5], which gives the queue size distribution immediately before arrivals in the SMP/M/1/FIFO queue. Thus it is also the queue size distribution immediately before arrivals in the SMP/M/1/PS queue.

Theorem 2 *Under Assumption 1, all states of the Markov chain* $\{(X^{(n)},S^{(n)});n=0,1,2,\ldots\}$ *are ergodic if* $\alpha\mu>1$. *In this case*

$$\varpi_k^{(l)}:=\lim_{n\to\infty}P\{X^{(n)}=k,S^{(n)}=l\}; \quad k=0,1,2,\ldots; \quad l=1,2,\ldots,L, \tag{8}$$

exist and are independent of the initial distribution. Letting $\boldsymbol{\varpi}_k:=[\varpi_k^{(1)},\varpi_k^{(2)},\ldots,\varpi_k^{(L)}]$ *we have*

$$\boldsymbol{\varpi}_k=\boldsymbol{\pi}\mathbf{G}^{-1}(\mathbf{I}-\mathbf{\Gamma})\mathbf{\Gamma}^k\mathbf{G}; \quad k=0,1,2,\ldots, \tag{9}$$

where $\mathbf{\Gamma}:=\mathbf{\Gamma}(0)$ *and* $\mathbf{G}:=\mathbf{G}(0)$.

We rewrite (9) in scalar form as

$$\varpi_k^{(m)}=\sum_{l=1}^{L}\beta_l\gamma_l^k g_{lm}; \quad k=0,1,2,\ldots; \quad m=1,2,\ldots,L, \tag{10}$$

where β_l is the lth element of the row vector

$$\boldsymbol{\beta}:=\boldsymbol{\pi}\mathbf{G}^{-1}(\mathbf{I}-\mathbf{\Gamma}), \tag{11}$$

$\gamma_l:=\gamma_l(0)$, and g_{lm} is the element of the matrix $\mathbf{G}$ at the lth row and the mth column.

3 Sojourn Time in the SMP/M/1/PS Queue

We now derive the sojourn time distribution in the SMP/M/1/PS queue. Let us focus on a *tagged* customer of type l who finds k other customers in the system upon his arrival. Let $S_k^{(l)}(t)$ denote the sojourn time distribution of this tagged customer. We define

$$A_{lm}(j,t) := \int_0^t \frac{(\mu x)^j}{j!} e^{-\mu x} da_{lm}(x) \tag{12}$$

as the probability that exactly j customers are served during the interarrival time of two successive arrivals of type l and type m customers when the interarrival time is less than t.

Lemma 1 *The functions $S_k^{(l)}(t)$ satisfy the following relations*

$$\begin{aligned} S_k^{(l)}(t) &= \sum_{m=1}^{L}\sum_{j=1}^{k+1} \frac{1}{k+1}\int_0^t [a_{lm}(\infty) - a_{lm}(x)] \frac{\mu(\mu x)^{j-1}}{(j-1)!} e^{-\mu x} dx \\ &+ \sum_{m=1}^{L}\sum_{j=0}^{k} \frac{k+1-j}{k+1} A_{lm}(j,t) * S_{k+1-j}^{(m)}(t); \quad l = 1,2,\ldots,L; \quad k = 0,1,2,\ldots, \end{aligned} \tag{13}$$

where $A_{lm}(j,t) * S_k^{(m)}(t)$ denotes the convolution of $A_{lm}(j,t)$ and $S_k^{(m)}(t)$.

Proof. Our proof extends the method of Ramaswami [10]. Depending on the situation whether the tagged customer ends his service before or after the next arrival, we have

$$S_k^{(l)}(t) = {}_1S_k^{(l)}(t) + {}_2S_k^{(l)}(t), \tag{14}$$

where ${}_1S_k^{(l)}(t)$ is the probability that the tagged customer, being type l and finding k other customers present, ends his service before the next arrival and has a sojourn time less than t, and ${}_2S_k^{(l)}(t)$ is the probability that the tagged customer, being type l and finding k other customers present, ends his service after the next arrival and has a sojourn time less than t.

For ${}_1S_k^{(l)}(t)$, due to the memoryless property of the exponentially distributed service time all customers present at time x have the identical distribution for the residual sojourn time. If there is a departure in a short time interval $(x, x+\Delta x]$, each customer present at time x has the same chance to depart. Thus if at least j customers end their services before the next arrival, the probability that the tagged customer is the jth to leave the system is

$$\frac{1}{j}\frac{\binom{k}{j-1}}{\binom{k+1}{j}} = \frac{1}{k+1}.$$

Conditioning on the type of the next arrival and the number of departures before the next arrival we have

$$ {}_1S_k^{(l)}(t) = \sum_{m=1}^{L} \sum_{j=1}^{k+1} \frac{1}{k+1} \int_0^t [a_{lm}(\infty) - a_{lm}(x)] \frac{\mu(\mu x)^{j-1}}{(j-1)!} e^{-\mu x} dx. \tag{15} $$

For ${}_2S_k^{(l)}(t)$, the probability that the tagged customer is not one of the j customers who depart from the system before the next arrival is

$$ \frac{\binom{k}{j}}{\binom{k+1}{j}} = \frac{k+1-j}{k+1}. $$

Conditioning on the length of the interarrival time, the type of the next arrival, and the number of departures before the next arrival we obtain

$$ {}_2S_k^{(l)}(t) = \sum_{m=1}^{L} \sum_{j=0}^{k} \frac{k+1-j}{k+1} A_{lm}(j,t) * S_{k+1-j}^{(m)}(t). \tag{16} $$

Substituting (15) and (16) into (14) gives (13). □

Let us introduce the generating function of the LST of $S_k^{(l)}(t)$ by

$$ \sigma^{(l)}(z,s) := \sum_{k=0}^{\infty} \sigma_k^{(l)}(s) z^k; \quad l = 1,2,\ldots,L, \tag{17} $$

where

$$ \sigma_k^{(l)}(s) := \int_0^{\infty} e^{-st} dS_k^{(l)}(t); \quad k = 0,1,2,\ldots. \tag{18} $$

Letting the column vector $\boldsymbol{\sigma}(z,s) := [\sigma^{(1)}(z,s), \sigma^{(2)}(z,s), \ldots, \sigma^{(L)}(z,s)]^{\mathrm{T}}$ we have the following theorem:

Theorem 3 *The vector $\boldsymbol{\sigma}(z,s)$ satisfies the differential equation*

$$ [z\mathbf{I} - \mathbf{A}(s+\mu-\mu z)] \frac{\partial \boldsymbol{\sigma}(z,s)}{\partial z} + \boldsymbol{\sigma}(z,s) = \frac{\mu}{(1-z)(s+\mu-\mu z)} [\mathbf{A}(0) - \mathbf{A}(s+\mu-\mu z)]. \tag{19} $$

Proof. Let us introduce the notation

$$ \psi^{(l)}(z,s) := \sum_{k=0}^{\infty} (k+1) \sigma_k^{(l)}(s) z^k. \tag{20} $$

It is easy to verify that

$$ \psi^{(l)}(z,s) = \sigma^{(l)}(z,s) + z \frac{\partial \sigma^{(l)}(z,s)}{\partial z}. \tag{21} $$

Taking the LST of (13), multiplying by $(k+1)z^k$, and summing over $k = 0, 1, 2, \ldots$, we obtain

$$\begin{aligned} \psi^{(l)}(z,s) &= \frac{\mu}{(1-z)(s+\mu-\mu z)} \sum_{m=1}^{L} [\alpha_{lm}(0) - \alpha_{lm}(s+\mu-\mu z)] \\ &\quad + \frac{1}{z}\sum_{m=1}^{L} \alpha_{lm}(s+\mu-\mu z)[\psi^{(m)}(z,s) - \sigma^{(m)}(z,s)]. \end{aligned} \tag{22}$$

Substituting (21) into (22) yields

$$\begin{aligned} z\frac{\partial \sigma^{(l)}(z,s)}{\partial z} + \sigma^{(l)}(z,s) &= \frac{\mu}{(1-z)(s+\mu-\mu z)} \sum_{m=1}^{L} [\alpha_{lm}(0) - \alpha_{lm}(s+\mu-\mu z)] \\ &\quad + \sum_{m=1}^{L} \alpha_{lm}(s+\mu-\mu z)\frac{\partial \sigma^{(m)}(z,s)}{\partial z}. \end{aligned} \tag{23}$$

Rewriting (23) in matrix form gives (19). □

Let $\sigma(s)$ denote the LST of the sojourn time for an arbitrary customers. It follows that

$$\sigma(s) = \sum_{m=1}^{L}\sum_{k=0}^{\infty} \varpi_k^{(m)} \sigma_k^{(m)}(s). \tag{24}$$

Substituting (10) into (24) yields

$$\begin{aligned} \sigma(s) &= \sum_{m=1}^{L}\sum_{k=0}^{\infty}\sum_{l=1}^{L} \beta_l g_{lm} \gamma_l^k \sigma_k^{(m)}(s) = \sum_{l=1}^{L}\sum_{m=1}^{L} \beta_l g_{lm} \sigma^{(m)}(\gamma_l, s) \\ &= \sum_{l=1}^{L} \beta_l \mathbf{g}_l \boldsymbol{\sigma}(\gamma_l, s), \end{aligned} \tag{25}$$

where $\mathbf{g}_l$ is the lth row of the matrix $\mathbf{G}$. Thus the mean sojourn time σ of an arbitrary customer is given by

$$\sigma = -\sum_{l=1}^{L} \beta_l \mathbf{g}_l \left[\frac{\partial}{\partial s}\boldsymbol{\sigma}(\gamma_l, s)\right]_{s=0+}. \tag{26}$$

Theorem 4 *The mean sojourn time of an arbitrary customer in the SMP/M/1/PS queue is given by*

$$\sigma = \frac{1}{\mu}\boldsymbol{\pi}\mathbf{G}^{-1}(\mathbf{I} - \boldsymbol{\Gamma})^{-1}\mathbf{G}\mathbf{1}. \tag{27}$$

Proof. Let us introduce the column vector

$$\mathbf{v}_l(s) := \left[\frac{\partial}{\partial z}\boldsymbol{\sigma}(z,s)\right]_{z=\gamma_l}; l = 1, 2, \ldots, L.$$

Evaluating both sides of (19) at $z = \gamma_l$ yields

$$[\gamma_l \mathbf{I} - \mathbf{A}(s + \mu - \mu\gamma_l)]\mathbf{v}_l(s) + \boldsymbol{\sigma}(\gamma_l, s) = \frac{\mu}{(1-\gamma_l)(s+\mu-\mu\gamma_l)}[\mathbf{A}(0) - \mathbf{A}(s+\mu-\mu\gamma_l)]. \tag{28}$$

Differentiating (28) with respect to s and taking the limit as s approaching 0+ gives

$$[\gamma_l \mathbf{I} - \mathbf{A}(\mu - \mu\gamma_l)]\mathbf{v}'_l(0) + \left[\frac{\partial}{\partial s}\boldsymbol{\sigma}(\gamma_l, s)\right]_{s=0+} = -\frac{[\mathbf{A}(0) - \mathbf{A}(\mu-\mu\gamma_l)]}{\mu(1-\gamma_l)^3}, \tag{29}$$

where we have used

$$\mathbf{v}_l(0) = \left[\frac{\partial}{\partial z}\boldsymbol{\sigma}(z, 0)\right]_{z=\gamma_l} = \left[\frac{d}{dz}\left(\frac{1}{1-z}\right)\right]_{z=\gamma_l}\mathbf{1} = \frac{1}{(1-\gamma_l)^2}\mathbf{1}. \tag{30}$$

Recall that $\mathbf{g}_l$ is the left-hand eigenvector of the matrix $\mathbf{A}(\mu - \mu\gamma_l)$ corresponding to the eigenvalue γ_l. It follows that

$$\mathbf{g}_l[\gamma_l \mathbf{I} - \mathbf{A}(\mu - \mu\gamma_l)] = 0.$$

Multiplying (29) by $\mathbf{g}_l$ on the left gives

$$-\mathbf{g}_l\left[\frac{\partial}{\partial s}\boldsymbol{\sigma}(\gamma_l, s)\right]_{s=+0} = \frac{\mathbf{g}_l\mathbf{1}}{\mu(1-\gamma_l)^2}. \tag{31}$$

Using (11) and (31) in (26) gives (27). □

4 Waiting Time in the SMP/M/1/RS Queue

We proceed to analyze the SMP/M/1/RS queue. In this system the service time follows exponential distribution with rate μ, the arrival process is governed by a semi-Markov process described in Section 2, and the service discipline is random. The random service discipline is described as follows: at the end of a service the next customer to be served is selected at random among all the customer present in the queue. Since the order of service does not influence the unfinished work of the system, the queue size distribution is independent of the order of service. Hence, the queue size distribution immediately before arrivals in the SMP/M/1/RS queue is identical with that in the corresponding SMP/M/1/FIFO queue.

Let $W_k^{(l)}(t)$ denote the waiting time distribution of a *tagged* customer of type l who finds $k+1$ other customers in the system upon his arrival. We then have the following lemma.

Lemma 2 *The functions $W_k^{(l)}(t)$ satisfy the relations*

$$\begin{aligned} W_k^{(l)}(t) &= \sum_{m=1}^{L}\sum_{j=1}^{k+1}\frac{1}{k+1}\int_0^t [a_{lm}(\infty) - a_{lm}(x)]\frac{\mu(\mu x)^{j-1}}{(j-1)!}e^{-\mu x}dx \\ &+ \sum_{m=1}^{L}\sum_{j=0}^{k}\frac{k+1-j}{k+1}A_{lm}(j,t) * W_{k+1-j}^{(m)}(t); \quad l = 1,2,\ldots,L; \quad k = 0,1,2,\ldots, \end{aligned} \tag{32}$$

where $A_{lm}(j,t)$ *is defined in (12), and* $*$ *denotes the convolution.*

Proof. The proof is similar to that for Lemma 1 in Section 3. □

Remark Comparing Lemma 1 and Lemma 2 we note that the sojourn time distribution for a type l customer who meets upon his arrival k customers present in the SMP/M/1/PS queue is identical with the waiting time distribution of a type l customer who meets upon his arrival $k+1$ customer present in the SMP/M/1/RS queue. This is an extension of the relation between the GI/M/1/PS queue and the GI/M/1/RS queue mentioned by Cohen [4].

Let us introduce the generating function of the LST of $W_k^{(l)}(t)$ by

$$w^{(l)}(z,s) := \sum_{k=0}^{\infty} w_k^{(l)}(s) z^k,$$

where

$$w_k^{(l)}(s) := \int_0^{\infty} e^{-st} dW_k^{(l)}(t).$$

With the same method as deriving (19) we get the following theorem:

Theorem 5 *The vector* $\boldsymbol{w}(z,s) := [w^{(1)}(z,s), w^{(2)}(z,s), \ldots, w^{(L)}(z,s)]^{\mathrm{T}}$ *satisfies the differential equation*

$$[z\mathbf{I} - \mathbf{A}(s+\mu-\mu z)]\frac{\partial \boldsymbol{w}(z,s)}{\partial z} + \boldsymbol{w}(z,s) = \frac{\mu}{(1-z)(s+\mu-\mu z)}[\mathbf{A}(0) - \mathbf{A}(s+\mu-\mu z)]. \tag{33}$$

Let $w(s)$ denote the LST of the waiting time distribution for an arbitrary customer. Conditioning on the number of customers present in the system immediately before the arrival of the arbitrary customer and his type, we obtain

$$w(s) = \sum_{m=0}^{L} \varpi_0^{(m)} + \sum_{m=1}^{L}\sum_{k=1}^{\infty} \varpi_k^{(m)} w_{k-1}^{(m)}(s). \tag{34}$$

Using (10) in (34) gives

$$\begin{aligned} w(s) &= \sum_{m=1}^{L}\sum_{l=1}^{L} \beta_l g_{lm} + \sum_{m=1}^{L}\sum_{k=0}^{\infty}\sum_{l=1}^{L} \beta_l g_{lm} \gamma_l^{k+1} w_k^{(m)}(s) \\ &= \sum_{l=1}^{L}\sum_{m=1}^{L} \beta_l g_{lm} + \sum_{l=1}^{L}\sum_{m=1}^{L} \beta_l g_{lm} \gamma_l w^{(m)}(\gamma_l, s) \\ &= \sum_{l=1}^{L} \beta_l \mathbf{g}_l \mathbf{1} + \sum_{l=1}^{L} \beta_l \gamma_l \mathbf{g}_l \boldsymbol{w}(\gamma_l, s). \end{aligned} \tag{35}$$

Therefore the mean waiting time w of an arbitrary customer is given by

$$w = -\sum_{l=1}^{L} \beta_l \gamma_l \mathbf{g}_l \left[\frac{\partial}{\partial s}\boldsymbol{w}(\gamma_l, s)\right]_{s=0+}. \tag{36}$$

Theorem 6 *The mean waiting time of an arbitrary customer in the SMP/M/1/RS queue is given by*

$$w = \frac{1}{\mu}\boldsymbol{\pi}\mathbf{G}^{-1}(\mathbf{I}-\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}\mathbf{G}\mathbf{1}. \tag{37}$$

Proof. With the same method as deriving (31) we obtain

$$-\mathbf{g}_l\left[\frac{\partial}{\partial s}\boldsymbol{w}(\gamma_l, s)\right]_{s=0+} = \frac{\mathbf{g}_l\mathbf{1}}{\mu(1-\gamma_l)^2}; \quad l = 1, 2, \ldots, L. \tag{38}$$

Using (11) and (38) in (36) gives (37). □

5 Special Semi-Markovian Arrival Processes

In this section we consider the sojourn time in the SMP/M/1/PS queue and the waiting time in the SMP/M/1/RS queue in a special case of the semi-Markov arrival process. Namely, we assume that the interarrival time distribution is determined only by the type of the next arrival, i.e., $\mathbf{A}(s) = \mathbf{1}[\alpha_1(s), \alpha_2(s), \ldots, \alpha_L(s)]$, where

$$\alpha_l(s) := \int_0^\infty e^{-st} dP\{S^{(n+1)} = l, A_{n+1} \le t | S^{(n)}\}; \quad l = 1, 2, \ldots, L.$$

It is then easy to verify that

$$\det[\lambda\mathbf{I} - \mathbf{A}(s)] = \lambda^{L-1}[\lambda - \alpha(s)]$$

and that $\lambda[\lambda - \alpha(s)]$ is the minimal polynomial [5, p.89] of $\mathbf{A}(s)$, where

$$\alpha(s) := \int_0^\infty e^{-st} dP\{A_{n+1} \le t | S^{(n)}\} = \sum_{l=1}^{L} \alpha_l(s).$$

Let us first consider the sojourn time in the SMP/M/1/PS queue for this special semi-Markov arrival process. If Assumption 1 is satisfied, the roots of the equation $\det[z\mathbf{I} - \mathbf{A}(s + \mu - \mu z)] = 0$ are $\gamma_1(s) = \gamma_2(s) = \cdots = \gamma_{L-1}(s) = 0$ and $\gamma_L(s) = \gamma(s)$, where $\gamma(s)$ is the root of the equation $z - \alpha(s + \mu - \mu z) = 0$ within the unit circle $|z| = 1$. Here we write γ for $\gamma(0)$. Therefore $\boldsymbol{\Gamma}$ becomes a diagonal matrix with the diagonal elements $0, 0, \ldots, 0, \gamma$. We note that the left-hand eigenvector of $\mathbf{A}(\mu - \mu\gamma)$ corresponding to the eigenvalue γ is

$$g_L = [\alpha_1(\mu - \mu\gamma), \alpha_2(\mu - \mu\gamma), \ldots, \alpha_L(\mu - \mu\gamma)],$$

which is the last row of the matrix $\mathbf{G}$. The last column of the matrix $\mathbf{G}^{-1}$ is $(1/\gamma)\mathbf{1}$, which is the right-hand eigenvector corresponding to γ. It follows that

$$\mathbf{G}^{-1}(\mathbf{I}-\boldsymbol{\Gamma})^{-1}\mathbf{G}\mathbf{1}$$

$$
\begin{aligned}
&= \begin{bmatrix} \cdot & \cdot & \dots & 1/\gamma \\ \cdot & \cdot & \dots & 1/\gamma \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & 1/\gamma \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/(1-\gamma) \end{bmatrix} \\
&\quad \times \begin{bmatrix} \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1(\mu-\mu\gamma) & \alpha_2(\mu-\mu\gamma) & \dots & \alpha_L(\mu-\mu\gamma) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} 1+\frac{\alpha_1(\mu-\mu\gamma)}{1-\gamma} & \frac{\alpha_2(\mu-\mu\gamma)}{1-\gamma} & \dots & \frac{\alpha_L(\mu-\mu\gamma)}{1-\gamma} \\ \frac{\alpha_1(\mu-\mu\gamma)}{1-\gamma} & 1+\frac{\alpha_2(\mu-\mu\gamma)}{1-\gamma} & \dots & \frac{\alpha_L(\mu-\mu\gamma)}{1-\gamma} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_1(\mu-\mu\gamma)}{1-\gamma} & \frac{\alpha_2(\mu-\mu\gamma)}{1-\gamma} & \dots & 1+\frac{\alpha_L(\mu-\mu\gamma)}{1-\gamma} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\
&= \frac{1}{1-\gamma}\mathbf{1}, \qquad (39)
\end{aligned}
$$

where we have used $\mathbf{G}^{-1}\mathbf{G} = \mathbf{I}$. Substituting (39) into (27) gives

$$\sigma = \frac{1}{\mu(1-\gamma)}. \qquad (40)$$

We remark that if the number of types of customers is one for which the LST of the interarrival time distribution is $\alpha(s)$, then (40) reduces to the mean sojourn time in a GI/M/1/PS queue given by Equation (8) in [10, p.440].

We also look at the waiting time in the SMP/M/1/RS queue for the special semi-Markov arrival process as described above. In the same way as deriving (39), we get

$$\mathbf{G}^{-1}(\mathbf{I}-\mathbf{\Gamma})^{-1}\mathbf{\Gamma}\mathbf{G}\mathbf{1} = \frac{\gamma}{1-\gamma}\mathbf{1}. \qquad (41)$$

Using (41) in (37) yields

$$w = \frac{\gamma}{\mu(1-\gamma)}. \qquad (42)$$

Hence, if there is a single type of customers for which the LST of the interarrival time distribution is $\alpha(s)$, (42) is reduced to the mean waiting time in the GI/M/1/RS queue. This queue is treated in Cohen's book [3, p.443], but he does not comment on this reduction in [4].

References

[1] E. Çinlar, "Queues with semi-Markovian arrivals," *Journal of Applied Probability*, Vol.4, No.2, pp.365–379, August 1967.

[2] E. G. Coffman, R. R. Muntz and H. Trotter, "Waiting time distributions for processor-sharing systems," *Journal of the Associate for Computing Machinery*, Vol.17, No.1, pp.123–130, January 1970.

[3] J. W. Cohen, *The Single Server Queue*, North-Holland Publishing Company, Amsterdam, 1982.

[4] J. W. Cohen, "On processor sharing and random service," *Journal of Applied Probability*, Vol.21, No.4, p.937, December 1984.

[5] F. R. Gantmacher, *The Theory of Matrices, Volume 1*, AMS Chelsea Publishing, Providence, Rhode Island, 2000.

[6] D. Jagerman and B. Sengupta, "The GI/M/1 processor-sharing queue and its heavy traffic analysis," *Communications in Statistics–Stochastic Models*, Vol.7, No.3, pp.379–395, September 1991.

[7] R. M. Loynes "Stationary waiting-time distribution for single server queue," *The Annals of Mathematical Statistics*, Vol.33, No.4, pp.1323–1339, December 1962.

[8] M. F. Neuts, "The single server queue with Poisson input and semi-Markov service Times," *Journal of Applied Probability*, Vol.3, No.1, pp.202–230, June 1966.

[9] T. J. Ott, "The sojourn time distribution in the M/G/1 queue with processor-sharing," *Journal of Applied Probability*, Vol.21, No.2, pp.360–378, June 1984.

[10] V. Ramaswami, "The sojourn time in the GI/M/1 queue with processor sharing," *Journal of Applied Probability*, Vol.21, No.2, pp.445–450, June, 1984.

[11] B. Sengupta, "An approximation for the sojourn-time distribution for the GI/G/1 processor sharing discipline queue," *Communications in Statistics–Stochastic Models*, Vol.8, No.1, pp.379–395, March 1991.

[12] L. Takács, *Introduction to the Theory of Queues*, Oxford University Press, New York, 1962.

[13] S. F. Yashkov, "A derivation of response time distribution for an M/G/1 processor-sharing queue," *Problems of Control and Information Theory*, Vol.12, pp.133–148, 1983.

[14] S. F. Yashkov, "Processor-sharing queues: some progress in analysis," *Queueing Systems*, Vol.2, No.1, pp.1–17, June 1987.