

THE SINGLE SERVER SEMI-MARKOV QUEUE

J.H.A. de SMIT

Twente University of Technology, Enschede, The Netherlands

Received 20 May 1985

Revised 8 November 1985

A general model for the single server semi-Markov queue is studied. Its solution is reduced to a matrix factorization problem. Given this factorization, results are obtained for the distributions of actual and virtual waiting times, queue lengths both at arrival epochs and in continuous time, the number of customers during a busy period, its length and the length of a busy cycle. Two examples are discussed for which explicit factorizations have been obtained.

semi-Markov queue * matrix factorization * actual waiting time * virtual waiting time * queue length * busy period * busy cycle

1. Introduction

The single server semi-Markov queue is described as follows. Customers arrive at a single server at time epochs T_1, T_2, \dots ; with $T_1 = 0$. The interarrival times are denoted by $A_n = T_n - T_{n-1}$, $n = 2, 3, \dots$; and the service time of the n -th customer by S_n , $n = 1, 2, \dots$. The queue discipline is first-come-first-served. Let $(Y_n, n = 1, 2, \dots)$ be an irreducible aperiodic Markov chain with finite state space $\{1, 2, \dots, N\}$ and assume that, for all choices of n, x, y and j ,

$$\begin{aligned} P(A_{n+1} \leq x, S_n \leq y, Y_{n+1} = j | A_2, \dots, A_n, S_1, \dots, S_{n-1}, Y_1, \dots, Y_n) \\ = P(A_{n+1} \leq x, S_n \leq y, Y_{n+1} = j | Y_n), \end{aligned} \quad (1.1)$$

while the latter conditional probability does not depend on n . Hence the model is completely specified by the functions

$$G_{ij}(\phi, \psi) = E(\exp(-\phi A_{n+1} - \psi S_n) 1(Y_{n+1} = j) | Y_n = i),$$

$i = 1, 2, \dots, N; j = 1, 2, \dots, N$; where $1(A)$ is the indicator function of the event A , i.e. $1(A) = 1$ if A occurs and $1(A) = 0$, otherwise.

Let $G(\phi, \psi)$ be the $N \times N$ -matrix with elements $G_{ij}(\phi, \psi)$.

This model has been studied before by Arjas [2]. He gives a formal solution using a matrix generalization of an identity for random walks. Takács [25] has analysed the model by an algebraic method generalizing the algebraic solution for the ordinary single server queue, while in [24] he has studied a discrete version. The departure

process has been studied by McNickle [16]. Several authors have considered various special cases of the model. The case

$$G_{ij}(\phi, \psi) = \frac{\lambda}{\lambda + \phi} B_{ij}(\psi), \quad (1.2)$$

i.e. Poisson arrivals and semi-Markov service times, has been studied by Neuts [18], Çinlar [6] and Arjas [2]. They have obtained formal results which, however, are not numerically tractable. In [19] Neuts has treated a more general model also allowing for group arrivals and he describes a numerical method leading to results for the busy period, the queue length and the waiting time. The case

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) \frac{\mu}{\mu + \psi}, \quad (1.3)$$

i.e. semi-Markov arrivals and exponential services, has been studied by Çinlar [7] and Arjas [2]. In the latter paper the model (1.3) is treated as the dual case of (1.2). Both authors obtain formal solutions. Latouche [12] gives a numerical solution for the queue length for the special case

$$G_{ij}(\phi, \psi) = \prod_{k=i}^N \frac{\lambda_k}{\lambda_k + \phi} P_{ij} \frac{\mu}{\mu + \psi}.$$

In [14] Loynes has studied the actual waiting time for

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) B_{ij}(\psi),$$

where either A_{ij} or B_{ij} is a rational function and in [15] the virtual waiting time for

$$G_{ij}(\phi, \psi) = \frac{\lambda_i}{\lambda_i + \phi} B_{ij}(\psi).$$

Recently Burman and Smith [5] have given approximations for the M|G|1 queue with Markov modulated arrivals, i.e.

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) B(\psi),$$

in which $A(\phi) = (\phi I + \Lambda - Q)^{-1} \Lambda$, where Q is the infinitesimal generator of the modulating Markov chain and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$, with λ_i the arrival rate when the modulating chain is in state i . The more general model for the M|G|1 queue with Markov modulated arrivals and services, i.e.

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) B_i(\psi),$$

in which $A(\phi)$ is as given above, has been completely solved by Regterschot and de Smit [23] using the methods discussed in the present paper. Similarly a complete solution for

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) \frac{\mu_i}{\mu_i + \psi},$$

generalizing the ordinary GI|M|1 queue is given in de Smit and Regterschot [9].

A related model has been considered by Purdue [21] and a more general model by Ramaswami [22].

The distribution of the waiting time in a single server semi-Markov queue is identical to that of the maximum of the corresponding random walk defined on a finite Markov chain. The latter quantity has been studied extensively. The scalar case has been treated by Kemperman [11]. For the matrix case we refer to Miller [17], Presman [20], Arjas [1], Borovkov [4] and Arndt [3].

Formal results have also been obtained for models in which the underlying Markov process has a more general state space (see Kaspi [10] and the references in that paper).

In Section 2 we derive a system of Wiener-Hopf type equations the solution of which reduces to a factorization problem. The same factorization problem arises from the analysis in Arjas [2] which is based on an identity for the corresponding random walk defined on a Markov chain. Arjas gives a formal factorization in which the factors have a probabilistic interpretation related to the busy period. From his results one does not find numerical solutions. Our interest is in explicit factorizations leading to numerical algorithms such as those in [9] and [23]. Moreover, in Section 3 we show how the distributions of waiting times, queue lengths and busy period can be found once the factorization problem has been solved. The relationships are generalizations of results obtained by Cohen [8] for the ordinary single server queue. Finally in Section 4 we give examples of explicit factorizations.

The actual waiting time of the n -th customer is denoted by W_n , the virtual time at time t by W_t^* , the total number of customers in the system just before the arrival of the n -th customer by C_n , the total number of customers in the system at time t by C_t^* , the length of the n -th busy period by P_n , the length of the n -th busy cycle by L_n and the number of customers during the n -th busy period by M_n .

We shall write $x^+ = \max(0, x)$, $x^- = \min(0, x)$, while δ_{ij} is Kronecker's symbol, i.e. $\delta_{ii} = 1$, and $\delta_{ij} = 0$ for $i \neq j$. I is the $N \times N$ -identity matrix and $\mathbf{1}$ is the N -dimensional column vector with all components equal to 1.

2. Reduction to a factorization problem

Let $U_n = S_1 + \dots + S_n$, $n = 1, 2, \dots$; $U_0 = 0$; be the cumulative service time of the first n customers and assume that the system is initially empty, i.e. $W_1 = 0$. For $i = 1, 2, \dots, N$; $j = 1, 2, \dots, N$; $\text{Re } \phi \geq 0$, and $\text{Re } \eta \geq 0$, $\text{Re } \theta \geq 0$, $|r| < 1$, or $\text{Re } \eta \geq 0$, $\text{Re } \theta > 0$, $|r| \leq 1$, or $\text{Re } \eta > 0$, $\text{Re } \theta \geq 0$, $|r| \leq 1$, we define

$$Z_{ij}(r, \phi, \eta, \theta) = \sum_{n=1}^{\infty} r^n E(\exp(-\phi W_n - \eta T_n - \theta U_{n-1}) \mathbf{1}(Y_n = j) | Y_1 = i)$$

and

$$V_{ij}(r, \phi, \eta, \theta) = \sum_{n=1}^{\infty} r^{n+1} E((1 - \exp(\phi[W_n + S_n - A_{n+1}]^-)) \cdot \exp(-\eta T_{n+1} - \theta U_n) \mathbf{1}(Y_{n+1} = j) | Y_1 = i).$$

Let $Z(r, \phi, \eta, \theta)$ and $V(r, \phi, \eta, \theta)$ be the $N \times N$ -matrices with elements $Z_{ij}(r, \phi, \eta, \theta)$ and $V_{ij}(r, \phi, \eta, \theta)$, respectively. Then we obtain the following system of Wiener-Hopf-type equations.

Theorem 2.1. For $\operatorname{Re} \phi = 0$ and $\operatorname{Re} \eta \geq 0$, $\operatorname{Re} \theta \geq 0$, $|r| < 1$, or $\operatorname{Re} \eta \geq 0$, $\operatorname{Re} \theta > 0$, $|r| \leq 1$, or $\operatorname{Re} \eta > 0$, $\operatorname{Re} \theta \geq 0$, $|r| \leq 1$, we have

$$Z(r, \phi, \eta, \theta)(I - rG(\eta - \phi, \phi + \theta)) = rI + V(r, -\phi, \eta, \theta). \quad (2.1)$$

Proof. From the identity

$$\exp(-\phi x^+) + \exp(-\phi x^-) = \exp(-\phi x) + 1,$$

and the relationship

$$W_{n+1} = [W_n + S_n - A_{n+1}]^+, \quad n = 1, 2, \dots,$$

we have for $\operatorname{Re} \phi = 0$, $\operatorname{Re} \eta \geq 0$, $\operatorname{Re} \theta \geq 0$,

$$\begin{aligned} & E(\exp(-\phi W_{n+1} - \eta T_{n+1} - \theta U_n) 1(Y_{n+1} = j) | Y_1 = i) \\ &= E(\exp(-\phi [W_n + S_n - A_{n+1}]^+ - \eta T_{n+1} - \theta U_n) 1(Y_{n+1} = j) | Y_1 = i) \\ &= E(\exp(-\phi [W_n + S_n - A_{n+1}] - \eta T_{n+1} - \theta U_n) 1(Y_{n+1} = j) | Y_1 = i) \\ &\quad + E((1 - \exp(-\phi [W_n + S_n - A_{n+1}]^-)) \exp(-\eta T_{n+1} - \theta U_n) 1(Y_{n+1} = j) | Y_1 = i). \end{aligned} \quad (2.2)$$

Moreover

$$\begin{aligned} & E(\exp(-\phi [W_n + S_n - A_{n+1}] - \eta T_{n+1} - \theta U_n) 1(Y_{n+1} = j) | Y_1 = i) \\ &= \sum_{k=1}^N E(\exp(-\phi W_n - \eta T_n - \theta U_{n-1}) 1(Y_n = k) | Y_1 = i) \\ &\quad \cdot E(\exp(-(\eta - \phi)A_{n+1} - (\phi + \theta)S_n) 1(Y_{n+1} = j) | Y_n = k) \\ &= \sum_{k=1}^N E(\exp(-\phi W_n - \eta T_n - \theta U_{n-1}) 1(Y_n = k) | Y_1 = i) G_{kj}(\eta - \phi, \phi + \theta). \end{aligned} \quad (2.3)$$

Inserting (2.3) into (2.2) yields for $\operatorname{Re} \phi = 0$, and $\operatorname{Re} \eta \geq 0$, $\operatorname{Re} \theta \geq 0$, $|r| < 1$, or $\operatorname{Re} \eta \geq 0$, $\operatorname{Re} \theta > 0$, $|r| \leq 1$, or $\operatorname{Re} \eta > 0$, $\operatorname{Re} \theta \geq 0$, $|r| \leq 1$,

$$\begin{aligned} Z_{ij}(r, \phi, \eta, \theta) &= r\delta_{ij} + r \sum_{k=1}^N Z_{ik}(r, \phi, \eta, \theta) G_{kj}(\eta - \phi, \phi + \theta) \\ &\quad + V_{ij}(r, -\phi, \eta, \theta), \end{aligned}$$

which proves the theorem.

The system of equations (2.1) can be solved whenever the symbol $I - rG(\eta - \phi, \phi + \theta)$ can be factorized, i.e. if two matrices $H^+(r, \phi, \eta, \theta)$ and $H^-(r, \phi, \eta, \theta)$ exist such that for $\operatorname{Re} \phi = 0$,

$$I - rG(\eta - \phi, \phi + \theta) = H^+(r, \phi, \eta, \theta)H^-(r, \phi, \eta, \theta) \quad (2.4)$$

and

for $\text{Re } \phi > 0$, $H^+(r, \phi, \eta, \theta)$ is analytic and for $\text{Re } \phi \geq 0$ it is bounded and continuous and its inverse exists and is also bounded; (2.5)

for $\text{Re } \phi < 0$, $H^-(r, \phi, \eta, \theta)$ is analytic and for $\text{Re } \phi \leq 0$ it is bounded and continuous and its inverse exists and is also bounded. (2.6)

With a standard Wiener-Hopf argument we then find the following solution of (2.1).

Theorem 2.2. *If a factorization according to (2.4)–(2.6) exists then we have, for $\text{Re } \phi \geq 0$ and $\text{Re } \eta \geq 0$, $\text{Re } \theta \geq 0$, $|r| < 1$, or $\text{Re } \eta > 0$, $\text{Re } \theta > 0$, $|r| \leq 1$, or $\text{Re } \eta > 0$, $\text{Re } \theta \geq 0$, $|r| \leq 1$,*

$$Z(r, \phi, \eta, \theta) = r[I - rG(\eta, \theta)]^{-1}H^+(r, 0, \eta, \theta)[H^+(r, \phi, \eta, \theta)]^{-1}. \quad (2.7)$$

Proof. From (2.1) and (2.4) we have for $\text{Re } \phi = 0$,

$$Z(r, \phi, \eta, \theta)H^+(r, \phi, \eta, \theta) = [rI + V(r, -\phi, \eta, \theta)][H^-(r, \phi, \eta, \theta)]^{-1}. \quad (2.8)$$

The left-hand side of (2.6) is analytic for $\text{Re } \phi > 0$ and it is bounded and continuous for $\text{Re } \phi \geq 0$; the right-hand side is analytic for $\text{Re } \phi < 0$ and bounded and continuous for $\text{Re } \phi \leq 0$. By analytic continuation we can define an entire function which is equal to the left-hand side of (2.6) for $\text{Re } \phi \geq 0$ and equal to the right-hand side for $\text{Re } \phi \leq 0$. But this entire function is bounded and hence by Liouville's theorem it is a constant. So for $\text{Re } \phi \geq 0$

$$\begin{aligned} Z(r, \phi, \eta, \theta)H^+(r, \phi, \eta, \theta) &= Z(r, 0, \eta, \theta)H^+(r, 0, \eta, \theta) \\ &= r[I - rG(\eta, \theta)]^{-1}H^+(r, 0, \eta, \theta), \end{aligned}$$

which proves the theorem.

Theorem 2.2 has been proved by Arjas [2] in a slightly different form and using a different method. Moreover, he has given a factorization in which the factors have a probabilistic interpretation related to the busy period. For the number of customers during the first busy period we have

$$M_1 = \inf\{n > 0 \mid U_n < T_{n+1}\}$$

for the length of the first busy period

$$P_1 = U_{M_1}$$

and for the length of the first busy cycle

$$L_1 = T_{M_1+1}.$$

The following factorization is given in [2]

$$H^+(r, \phi, \eta, \theta) = [\hat{G}(r, \eta - \phi, \phi + \theta)]^{-1}, \quad (2.9)$$

where

$$\hat{G}_{ij}(r, \eta, \theta) = \sum_{n=0}^{\infty} r^n E(\exp(-\eta T_{n+1} - \theta U_n) 1(M_1 > n, Y_{n+1} = j) | Y_1 = i)$$

and

$$H^-(r, \phi, \eta, \theta) = I - \hat{H}(r, \eta - \phi, \phi + \theta), \quad (2.10)$$

where

$$\hat{H}_{ij}(r, \eta, \theta) = E(r^{M_1} \exp(-\eta L_1 - \theta P_1) 1(Y_{M_1+1} = j) | Y_1 = i).$$

It is easily checked that this factorization satisfies (2.4)–(2.6). Moreover such a factorization is unique up to a constant matrix, so any factorization with $\lim_{\phi \rightarrow \infty} H^+(r, \phi, \eta, \theta) = I$ is identical to the one given above. Expressions (2.9) and (2.10) do not lead to explicit numerically tractable results. As in the scalar case (the ordinary GI|G|1 queue) such results can only be obtained in special cases. In Section 4 we give two examples.

3. Waiting time, queue length and busy period

Assuming that we have been able to factorize $I - rG(\eta - \phi, \phi + \theta)$ according to (2.4)–(2.6) and that we have found the solution for $Z(r, \phi, \eta, \theta)$ given by (2.7), we can obtain results for the distributions of several quantities of interest. Let the row vector $\pi = (\pi_1, \pi_2, \dots, \pi_N)$ denote the stationary distribution of the Markov chain, then π is determined by

$$\pi G(0, 0) = \pi \quad (3.1)$$

and

$$\pi \mathbf{1} = 1.$$

Let

$$\alpha_{ij} = E(A_{n+1} 1(Y_{n+1} = j) | Y_n = i) = -\frac{\partial}{\partial \phi} G_{ij}(\phi, 0) \Big|_{\phi=0},$$

$$\beta_{ij} = E(S_n 1(Y_{n+1} = j) | Y_n = i) = -\frac{\partial}{\partial \psi} G_{ij}(0, \psi) \Big|_{\psi=0},$$

and $\bar{\alpha}$ and $\bar{\beta}$ the $N \times N$ -matrices with elements α_{ij} and β_{ij} , respectively, then the expected steady state interarrival time is given by

$$\alpha = \pi \bar{\alpha} \mathbf{1},$$

and the expected steady state service time by

$$\beta = \pi \bar{\beta} \mathbf{1}.$$

The traffic intensity is defined by

$$\rho = \beta / \alpha,$$

assuming that $\alpha < \infty$ and $\beta < \infty$.

Arjas [2] has shown that, for $\rho < 1$, (W_n, Y_n) converges weakly to a proper random vector (W, Y) as $n \rightarrow \infty$. Since $((C_n, Y_n), n = 1, 2, \dots)$ is regenerative with the same regeneration epochs as $((W_n, Y_n), n = 1, 2, \dots)$ it follows that, for $\rho < 1$, (C_n, Y_n) converges weakly to a proper random vector (C, Y) as $n \rightarrow \infty$. Similarly for $\rho < 1$ and A_n having a non-arithmetic marginal distribution it is shown that W_t^* and C_t^* converge to proper random variables W^* and C^* respectively as $t \rightarrow \infty$.

Actual waiting time

Assume $\rho < 1$ and write

$$Z_j(\phi) = E(\exp(-\phi W)1(Y = j))$$

and $Z(\phi) = (Z_1(\phi), \dots, Z_N(\phi))$, then using Abel's theorem we have from (2.7) for $\text{Re } \phi \geq 0$,

$$Z(\phi) = \pi H^+(1, 0, 0, 0)[H^+(1, \phi, 0, 0)]^{-1}. \tag{3.2}$$

Virtual waiting time

For $\text{Re } \eta > 0, \text{Re } \phi \geq 0, i = 1, \dots, N$; we denote

$$Z_i^*(\eta, \phi) = \int_0^\infty \exp(-\eta t) E(\exp(-\phi W_t^*) | Y_1 = i) dt,$$

and $Z^*(\eta, \phi)$ is the column vector with components $Z_i^*(\eta, \phi)$. Let N_t be the number of customers arriving during $[0, t]$, then we have for the virtual waiting time at time t

$$W_t = [W_{N_t} + S_{N_t} + T_{N_t} - t]^+. \tag{3.3}$$

Using the identity

$$\exp(-\phi x^+) = \frac{1}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} \frac{\phi}{\phi - \xi} \exp(-\xi x), \quad 0 < \text{Re } \xi < \text{Re } \phi, \tag{3.4}$$

where the path of integration is a line parallel to the imaginary axis, we have

$$\begin{aligned} & E(\exp(-\phi W_t^*) | Y_1 = i) \\ &= \frac{1}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} \frac{\phi}{\phi - \xi} \sum_{j=1}^N \sum_{n=1}^\infty \int_0^t \exp(\xi(t-u)) \\ &\quad \cdot E(\exp(-\xi S_n)1(A_{n+1} > t-u) | Y_n = j) \\ &\quad \cdot d_u E(\exp(-\xi W_n)1(T_n \leq u, Y_n = j) | Y_1 = i). \end{aligned} \tag{3.5}$$

The interchange of the order of integration and of summation that we have applied here is justified by absolute convergence using Fubini's theorem. For $0 < \text{Re } \xi < \text{Re } \phi$ and $\text{Re } \eta > 0$ we thus obtain

$$\begin{aligned} Z_i^*(\eta, \phi) &= \frac{1}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} \frac{\phi}{\phi-\xi} \sum_{j=1}^N \int_0^\infty \exp(-(\eta-\xi)t) \\ &\quad \cdot E(\exp(-\xi S_n) 1(A_{n+1} > t) | Y_n = j) dt Z_{ij}(1, \xi, \eta, 0) \\ &= \frac{1}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} \frac{\phi}{\phi-\xi} \frac{1}{\eta-\xi} \sum_{k=1}^N [Z(1, \xi, \eta, 0) \\ &\quad \cdot [G(0, \xi) - G(\eta-\xi, \xi)]]_{ik}. \end{aligned} \quad (3.6)$$

From (2.1) we have for $0 \leq \text{Re } \phi \leq \text{Re } \eta$, $\text{Re } \eta > 0$,

$$\begin{aligned} Z(1, \phi, \eta, 0) [G(0, \phi) - G(\eta - \phi, \phi)] \\ = Z(1, \phi, \eta, 0) [G(0, \phi) - I] + I + V(1, -\phi, \eta, 0). \end{aligned} \quad (3.7)$$

Inserting (3.7) into (3.6) and applying contour integration we find for $\text{Re } \eta > 0$, $\text{Re } \phi \geq 0$,

$$\begin{aligned} Z^*(\eta, \phi) = \frac{1}{\eta} \left[I - \frac{\eta}{\eta - \phi} Z(1, \phi, \eta, 0) [I - G(0, \phi)] - \frac{\phi}{\phi - \eta} Z(1, \eta, \eta, 0) \right. \\ \left. \cdot [I - G(0, \eta)] \right] 1. \end{aligned} \quad (3.8)$$

This result was already obtained by Takács [25] using an algebraic method. Since for $\text{Re } \eta > 0$,

$$Z(1, 0, \eta, 0) = [I - G(\eta, 0)]^{-1},$$

we see that

$$\alpha \lim_{\eta \downarrow 0} \eta Z_{ij}(1, 0, \eta, 0) = \pi_j$$

by showing that the left-hand side satisfies the equations (3.1). Consequently for $\rho < 1$ we have

$$\lim_{\eta \downarrow 0} \eta Z_{ij}(1, 0, \eta, 0) = \frac{1}{\alpha} Z_j(0),$$

and with (2.7)

$$\lim_{\eta \downarrow 0} \eta Z_{ij}(1, \eta, \eta, 0) = \frac{1}{\alpha} Z_j(0) \quad \text{and} \quad \lim_{\eta \downarrow 0} \eta Z_{ij}(1, \phi, \eta, 0) = \frac{1}{\alpha} Z_j(\phi).$$

For $\rho < 1$ and A_n non-arithmetic we then obtain from (3.8), for $\text{Re } \phi \geq 0$,

$$E(\exp(-\phi W^*)) = 1 - \rho + \frac{1}{\alpha\phi} Z(\phi) [I - G(0, \phi)] 1, \quad (3.9)$$

which generalizes the corresponding result for the ordinary single server queue given by Cohen [8].

Number of customers in the system at arrival epochs

Since

$$\{C_n \leq j\} = \Omega, \quad n = 1, 2, \dots, j+1,$$

and

$$\{C_{n+j+1} \leq j\} = \{T_n + W_n + S_n < T_{n+j+1}\}, \quad n = 1, 2, \dots,$$

where Ω is the sure event, we find

$$\begin{aligned} & \sum_{n=1}^{\infty} r^n E(s^{C_n} 1(Y_n = k) | Y_1 = i) \\ &= r[I - rsG(0, 0)]_{ik}^{-1} + r(1-s) \sum_{l=1}^N \sum_{j=0}^{\infty} (rs)^j \\ & \cdot \sum_{n=1}^{\infty} P(T_n + W_n + S_n < T_{n+j+1}, Y_n = l, Y_{n+j+1} = k | Y_1 = i). \end{aligned} \quad (3.10)$$

Let

$$R_{ik}(r, s) = \sum_{n=1}^{\infty} r^n E(s^{C_n} 1(Y_n = k) | Y_1 = i)$$

and let $R(r, s)$ be the $N \times N$ -matrix with elements $R_{ik}(r, s)$. We assume that there exists a $\delta > 0$ such that $G(\phi, \psi)$ can be continued analytically to the region $\text{Re } \phi > -\delta$ for $\text{Re } \psi \geq 0$; moreover we assume that $\lim_{|\phi| \rightarrow \infty} G(-\phi, \phi) = 0$. The latter assumption implies that the conditional distribution functions $P(S_n - A_{n+1} \leq x | Y_n = i, Y_{n+1} = j)$ are continuous and that for all n, i, j, l and k ,

$$P(T_n + W_n + S_n = T_{n+j+1}, Y_n = l, Y_{n+j+1} = k | Y_1 = i) = 0.$$

Using the inversion formula for Laplace-Stieltjes transforms we have

$$\begin{aligned} & P(T_n + W_n + S_n < T_{n+j+1}, Y_n = l, Y_{n+j+1} = k | Y_1 = i) \\ &= \frac{1}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} E(\exp(-\xi(W_n + S_n + T_n - T_{n+j+1})) 1(Y_n = l, Y_{n+j+1} = k) | Y_1 = i) \\ &= \frac{1}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} E(\exp(-\xi W_n) 1(Y_n = l) | Y_1 = i) \sum_{m=1}^N G_{lm}(-\xi, \xi) [G(-\xi, 0)]_{mk}. \end{aligned}$$

Inserting this into (3.10) yields for $|r| < 1, |s| < 1$,

$$\begin{aligned} R(r, s) &= r[I - rsG(0, 0)]^{-1} + \frac{r(1-s)}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} Z(r, \xi, 0, 0) G(-\xi, \xi) \\ & \cdot [I - rsG(-\xi, 0)]^{-1}, \end{aligned} \quad (3.11)$$

with $\text{Re } \xi > 0$ and so small that $\text{sp}(rsG(-\xi, 0)) < 1$, where $\text{sp}(M)$ is the spectral radius of the matrix M .

For $\rho < 1$ denote

$$R_k(s) = E(s^{C_n} 1(Y = k))$$

and let $R(s) = (R_1(s), \dots, R_N(s))$, then using Abel's theorem we have from (3.11), for $|s| < 1$,

$$R(s) = \frac{1-s}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} Z(\xi) G(-\xi, \xi) [I - sG(-\xi, 0)]^{-1}, \tag{3.12}$$

with $\text{Re } \xi > 0$ and so small that $\text{sp}(sG(-\xi, 0)) < 1$. If $G(\phi, \psi)$ is a rational function of ϕ it will be possible to evaluate the above integral by considering a closed contour in the right half plane. If on the other hand $G(\phi, \psi)$ is a rational function of ψ it is shown, using Theorem 1 in [3], that $Z(\phi)$ is a rational function of ϕ so that the integral can be evaluated by considering a closed contour in the left half plane.

Number of customers in the system in continuous time

For $j = 0, 1, \dots$, we have

$$\{C_i^* \leq j\} = \{t < T_{j+1}\} \cup \bigcup_{n=1}^{\infty} \{T_n + W_n + S_n < t, T_{n+j} \leq t < T_{n+j+1}\}.$$

Let

$$R_i^*(\eta, s) = \int_0^{\infty} \exp(-\eta t) E(s^{C_i^*} | Y_1 = i) dt,$$

and let $R^*(\eta, s)$ be the column vector with components $R_i^*(\eta, s)$, $i = 1, 2, \dots, N$. Assume again that there exists a $\delta > 0$, such that $G(\phi, \psi)$ can be continued analytically to the region $\text{Re } \phi > -\delta$, for $\text{Re } \psi \geq 0$, and that $\lim_{|\phi| \rightarrow \infty} G(-\phi, \phi) = 0$. Then analogous to the previous derivations we find for $\text{Re } \eta > 0$, $|s| < 1$,

$$\begin{aligned} R^*(\eta, s) &= \frac{1}{\eta} [I - (1-s)(I - sG(\eta, 0))^{-1}] \mathbf{1} \\ &\quad + \frac{1-s}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} \frac{1}{\eta - \xi} Z(1, \xi, \eta, 0) G(\eta - \xi, \xi) G(\eta - \xi, 0)^{-1} \\ &\quad \cdot [I - sG(\eta - \xi, 0)]^{-1} [I - G(\eta - \xi, 0)] \mathbf{1}, \end{aligned} \tag{3.13}$$

with $0 < \text{Re } \xi < \text{Re } \eta$. For $\rho < 1$ and A_n non-arithmetic we have from (3.13), for $|s| < 1$,

$$\begin{aligned} E(s^{C^*}) &= \frac{1-s}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} \frac{1}{\alpha\xi} Z(\xi) G(-\xi, \xi) G(-\xi, 0)^{-1} \\ &\quad \cdot [I - sG(-\xi, 0)]^{-1} [G(-\xi, 0) - I] \mathbf{1}, \end{aligned} \tag{3.14}$$

where $\text{Re } \xi > 0$ and so small that $\text{sp}(sG(-\xi, 0)) < 1$.

Busy period

We have defined

$$\hat{H}_{ij}(r, \eta, \theta) = E(r^{M_1} \exp(-\eta L_1 - \theta P_1) \mathbf{1}(Y_{M_1+1} = j) | Y_1 = i).$$

From the renewal equation we have

$$Z(r, \infty, \eta, \theta) = rI + Z(r, \infty, \eta, \theta)\hat{H}(r, \eta, \theta),$$

so that for $|r| < 1, \operatorname{Re} \eta \geq 0, \operatorname{Re} \theta \geq 0$, or $|r| \leq 1, \operatorname{Re} \eta > 0, \operatorname{Re} \theta \geq 0$, or $|r| \leq 1, \operatorname{Re} \eta \geq 0, \operatorname{Re} \theta > 0$,

$$\hat{H}(r, \eta, \theta) = I - rZ(r, \infty, \eta, \theta)^{-1}. \tag{3.15}$$

From (2.7) we see that if $\lim_{\phi \rightarrow \infty} H^+(r, \phi, \eta, \theta) = I$, this is in agreement with (2.10).

4. Two examples

In two recent papers ([9] and [23]) two special cases of the single server semi-Markov queue have been studied and explicit factorizations have been obtained. Here we give some results from those papers.

A semi-Markov queue with exponential service times

(See de Smit and Regterschot [9].)

The model is given by

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) \frac{\lambda_i}{\lambda_i + \psi}.$$

It is shown that for $|r| < 1, \operatorname{Re} \eta \geq 0, \operatorname{Re} \theta \geq 0$, or $|r| \leq 1, \operatorname{Re} \eta > 0, \operatorname{Re} \theta \geq 0$, or $|r| \leq 1, \operatorname{Re} \eta \geq 0, \operatorname{Re} \theta > 0$, or $r = 1, \eta = 0, \theta = 0$ and $\rho < 1, \det(I - rG(\eta - \phi, \phi + \theta))$, considered as a function of ϕ , has exactly N zeros $\mu_1(r, \eta, \theta), \dots, \mu_N(r, \eta, \theta)$, in the left half plane $\operatorname{Re} \phi < 0$. We then have a factorization according to (2.4)-(2.6) in which

$$H^+(r, \phi, \eta, \theta) = K(r, \phi, \eta, \theta)^{-1},$$

where K has the form

$$K(r, \phi, \eta, \theta) = I + C \operatorname{diag}\left(\frac{1}{\phi - \mu_1(r, \eta, \theta)}, \dots, \frac{1}{\phi - \mu_N(r, \eta, \theta)}\right)B$$

with the $N \times N$ -matrices B and C depending on r, η and θ but not on ϕ . The k -th row B_k of B is a non-zero vector satisfying

$$B_k = B_k rG(\eta - \mu_k(r, \eta, \theta), \mu_k(r, \eta, \theta) + \theta) \quad \text{and} \quad C = L^{-1},$$

where

$$L_{ij} = \frac{B_{ij}}{\theta + \lambda_j + \mu_i(r, \eta, \theta)}.$$

If $\rho < 1$ we have with the general results of the previous section for $\operatorname{Re} \phi \geq 0$,

$$Z(\phi) = \pi K(1, 0, 0, 0)^{-1} K(1, \phi, 0, 0)$$

and if moreover A_n is non-arithmetic,

$$E(\exp(-\phi W^*)) = 1 - \rho + \frac{1}{\alpha} Z(\phi) \operatorname{diag}\left(\frac{1}{\phi + \lambda_1}, \dots, \frac{1}{\phi + \lambda_N}\right) \mathbf{1}.$$

Similarly from (3.12) and (3.14) we have obtained (somewhat more complicated) explicit results for $R(s)$ and $E(s^{C^*})$. It turns out that the distributions of W and W^* are mixtures of exponentials and that the distributions of C and C^* are mixtures of geometric distributions.

The queue $M|G|1$ with Markov modulated arrivals and services

(See Regterschot and de Smit [23].)

The model is given by

$$G_{ij}(\phi, \psi) = A_{ij}(\phi) B_i(\psi)$$

in which

$$A(\phi) = (\phi I + \Lambda - Q)^{-1} \Lambda,$$

where Q is the infinitesimal generator of an irreducible Markov chain with state space $\{1, 2, \dots, N\}$ and $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_N)$, with $\lambda_i > 0$ the arrival rate when the Markov chain is in state i . Denote $B(\phi) = \operatorname{diag}(B_1(\phi), \dots, B_n(\phi))$. The matrix $\Lambda - Q$ has N eigenvalues ν_1, \dots, ν_N which all lie in the right half plane $\operatorname{Re} \phi > 0$. For $|r| < 1, \operatorname{Re} \eta \geq 0, \operatorname{Re} \theta \geq 0$, or $|r| \leq 1, \operatorname{Re} \eta > 0, \operatorname{Re} \theta \geq 0$, or $|r| \leq 1, \operatorname{Re} \eta \geq 0, \operatorname{Re} \theta > 0$, $\det(I - rG(\eta - \phi, \phi + \theta))$, considered as a function of ϕ , has exactly N zeros $\mu_1(r, \eta, \theta), \dots, \mu_N(r, \eta, \theta)$ in the right half plane $\operatorname{Re} \phi > 0$, for $r = 1, \eta = 0, \theta = 0$ and $\rho < 1$ one of these zeros, say $\mu_1(1, 0, 0)$, is equal to 0 while the others remain in the right half plane $\operatorname{Re} \phi > 0$. Let the row vector R_i be a left eigenvector of $\Lambda - Q$ corresponding to the eigenvalue ν_i and R the $N \times N$ -matrix the i -th row of which is R_i . Let D^j be a non-zero column vector satisfying

$$D^j = rG(\eta - \mu_j(r, \eta, \theta), \mu_j(r, \eta, \theta) + \theta) D^j$$

and D the $N \times N$ -matrix the j -th column of which is D^j . Define the $N \times N$ -matrix S by

$$S_{ij} = \frac{1}{\nu_i + \eta - \mu_j(r, \eta, \theta)} R_i \Lambda D^j$$

and

$$C = S^{-1} R \Lambda.$$

Then we have a factorization according to (2.4)-(2.6) in which

$$H^-(r, \phi, \eta, \theta) = K(r, \phi, \eta, \theta)^{-1},$$

where

$$K(r, \phi, \eta, \theta) = I - D \operatorname{diag}\left(\frac{1}{\phi - \mu_1(r, \eta, \theta)}, \dots, \frac{1}{\phi - \mu_N(1, \eta, \theta)}\right) C.$$

Note again that D and C depend on r , η and θ but not on ϕ . If $\rho < 1$ we have from the results of the previous section for $\text{Re } \phi \geq 0$,

$$Z(\phi) = \pi H^+(0) [H^+(\phi)]^{-1},$$

where for $\phi \neq 0$.

$$H^+(\phi) = (I - B(\phi)A(-\phi))K(1, \phi, 0, 0)$$

and

$$H^+(0) = (I - P) \left[I + D \text{diag} \left(0, \frac{1}{\mu_2(1, 0, 0)}, \dots, \frac{1}{\mu_N(1, 0, 0)} \right) C \right] - (\bar{\beta} - \bar{\alpha})D^1 C_1,$$

with C_1 the first row of C and $P = A(0) = (\Lambda - Q)^{-1}A$; and moreover

$$E(\exp(-\phi W^*)) = \frac{1}{\alpha} Z(\phi) \Lambda^{-1} \mathbf{1},$$

where $\alpha = \pi \Lambda^{-1} \mathbf{1}$.

From (3.12) we have, for $|s| < 1$,

$$R(s) = \frac{1-s}{2\pi i} \int_{-i\infty+0}^{i\infty+0} \frac{d\xi}{\xi} Z(\xi) B(\xi) (-\xi I + (1-s)A - Q)^{-1} \Lambda$$

and, with (3.14),

$$E(s^{C^*}) = \frac{1}{\alpha} R(s) \Lambda^{-1} \mathbf{1}.$$

The above expressions can be used for calculating $P(C = 0, Y = i)$, $P(C^* = 0)$ and the moments of C and C^* (see [23]).

References

- [1] E. Arjas, On a fundamental identity in the theory of semi-Markov processes, *Adv. Appl. Prob.* 4 (1972) 258-270.
- [2] E. Arjas, On the use of a fundamental identity in the theory of semi-Markov queues, *Adv. Appl. Prob.* 4 (1972) 271-284.
- [3] K. Arndt, On the distribution of the supremum of a random walk on a Markov chain, in: A.A. Borovkov, ed., *Advances in Probability Theory, Limit Theorems and Related Problems*, Optimization Software (New York, 1984).
- [4] K.A. Borovkov, Stability theorems and estimates of the rate of convergence of the components of factorizations for walks defined on Markov chains, *Theory Probab. Appl.* 25 (1980) 325-334.
- [5] D.Y. Burman and D.R. Smith, An asymptotic analysis of a queueing system with Markov-modulated arrivals, Bell Laboratories, Holmdel, NJ, 1983.
- [6] E. Çinlar, Time dependence of queues with semi-Markov services, *J. Appl. Prob.* 4 (1967) 356-364.
- [7] E. Çinlar, Queues with semi-Markov arrivals, *J. Appl. Prob.* 4 (1967) 365-379.
- [8] J.W. Cohen, *The Single Server Queue* (North-Holland, Amsterdam, 1982).
- [9] J.H.A. de Smit and G.J.K. Regterschot, A semi-Markov queue with exponential service times, *Proceedings of the International Symposium on Semi-Markov Processes* (Plenum, 1986).

- [10] H. Kaspi, Storage processes with Markov additive input and output, *Math. Oper. Res.* 9 (1984) 424–440.
- [11] J.H.B. Kemperman, *The Passage Problem for a Stationary Markov Chain* (The University of Chicago Press, 1961).
- [12] G. Latouche, *An exponential semi-Markov process with applications to queueing theory*, Laboratoire d'Informatique Théorique, Université Libre de Bruxelles, 1983.
- [13] R.M. Loynes, The stability of a queue with non-independent interarrival and service times, *Proc. Camb. Philos. Soc.* 58 (1962) 497–520.
- [14] R.M. Loynes, Stationary waiting time distributions for single server queues, *Ann. Math. Statist.* 33 (1962) 1323–1339.
- [15] R.M. Loynes, A continuous time treatment of certain queues and infinite dams, *J. Austral. Math. Soc.* 2 (1962) 484–498.
- [16] D.C. McNickle, The number of departures from a semi-Markov queue, *J. Appl. Prob.* 11 (1974) 825–828.
- [17] H.D. Miller, A matrix factorization problem in the theory of random variables defined on a finite Markov chain, *Proc. Cambridge Philos. Soc.* 58 (1962) 497–520.
- [18] M.F. Neuts, The single server queue with Poisson input and semi-Markov service times, *J. Appl. Prob.* 3 (1966) 202–230.
- [19] M.F. Neuts, Some explicit formulas for the steady-state behaviour of the queue with semi-Markov service times, *Adv. Appl. Prob.* 9 (1977) 141–157.
- [20] È.L. Presman, Factorization methods and boundary problems for sums of random variables given on Markov chains, *Math. USSR-Izv.* 3 (1969) 815–852.
- [21] P. Purdue, A queue with Poisson-input and semi-Markov service times: Busy period analysis, *J. Appl. Prob.* 12 (1975) 353–357.
- [22] V. Ramaswami, The $N|G|1$ queue and its detailed analysis, *Adv. Appl. Prob.* 12 (1980) 222–261.
- [23] G.J.K. Regterschot and J.H.A. de Smit, The queue $M|G|1$ with Markov modulated arrivals and services (1986), to appear in *Math. Oper. Res.*
- [24] L. Takács, A storage process with semi-Markov input, *Adv. Appl. Prob.* 7 (1975) 830–844.
- [25] L. Takács, A Banach space of matrix functions and its application in the theory of queues, *Sankhya Ser. A* 38 (1976) 201–211.