

THE *BMAP/G/1* VACATION QUEUE WITH QUEUE-LENGTH DEPENDENT VACATION SCHEDULE

YANG WOO SHIN¹ and CHARLES E. M. PEARCE²

(Received 1 July 1996; revised 28 July 1997)

Abstract

We treat a single-server vacation queue with queue-length dependent vacation schedules. This subsumes the single-server vacation queue with exhaustive service discipline and the vacation queue with Bernoulli schedule as special cases. The lengths of vacation times depend on the number of customers in the system at the beginning of a vacation. The arrival process is a batch-Markovian arrival process (BMAP). We derive the queue-length distribution at departure epochs. By using a semi-Markov process technique, we obtain the Laplace-Stieltjes transform of the transient queue-length distribution at an arbitrary time point and its limiting distribution.

1. Introduction

Because of its applicability to the performance evaluation of computer, communication and manufacturing systems, the queue with server vacations has been the subject of extensive study over the last two decades. For detailed bibliographies on vacation models, the reader is referred to Doshi [2] and Takagi [12]. A number of different vacation models have been introduced. Vacation models are distinguished by their scheduling disciplines, that is, the rules determining when a service stops and a vacation begins. In the exhaustive service discipline, the server takes a vacation only when there are no customers in the system. In the nonexhaustive service discipline, a vacation may start even when customers are present in the system.

Most of the previous work on vacation queues assumes that customers arrive at the system in accordance with a stationary Poisson process. The *M/G/1* vacation queue with queue-length dependent vacation schedule and vacation times has been studied by Harris and Marchal [3]. This model was extended to its *M^X/G/1* version by Shin [11].

¹Department of Statistics, Changwon National University, 9 Sarimdong, Changwon 641 – 773, Korea

²Applied Mathematics Department, The University of Adelaide, Adelaide SA 5005, Australia

© Australian Mathematical Society, 1998, Serial-fee code 0334-2700/98

Lucantoni, Meier–Hellstern and Neuts [6] studied the exhaustive vacation system with Markovian arrival process (MAP) and provided algorithmically–tractable equations for the distributions of the waiting times at an arbitrary time and at arrival instants, as well as the queue length at an arbitrary time, at arrival instants and at departure instants. Machihara [9] considered the vacation queue with phase–type Markov renewal arrivals, semi–Markovian service times and semi–Markovian vacation times. Takine and Hasegawa [13] analysed the batch $SPP/G/1$ queue with multiple vacations and exhaustive service disciplines, using a supplementary variable technique.

In this paper, we treat a $BMAP/G/1$ vacation queue whose vacation schedule and the lengths of whose vacation times depend on the queue length of the system at the beginning of a vacation. The vacation schedule considered subsumes the exhaustive service discipline and the Bernoulli schedule. Following the general approach of Neuts [10] we derive the queue–length distribution at departure epochs, the transient queue–length distribution and its limiting distribution.

The paper is organized as follows. In Section 2, we set up the model and address an underlying Markov renewal process, the transition matrix of which is spatially inhomogeneous. This is sufficient for an analysis of the stationary queue–length distribution at departure epochs, which is effected in Section 3. To treat the transient queue–length distribution at an arbitrary time point, we need to first analyse the associated first–passage problem for the renewal process introduced in Section 2. This is done in Section 4. We then consider the queue length at an arbitrary time point in Section 5 and the corresponding stationary distribution in Section 6.

2. The model

2.1. Arrival process Arrivals to the system are according to a $BMAP$ with m phases and coefficient matrices $\{D_k, k \geq 0\}$. For a detailed definition and examples of $BMAPs$ we refer the reader to Lucantoni [4]. The matrix D_0 has negative diagonal elements, the matrices D_k ($k \geq 1$) are nonnegative and $D = \sum_{k=0}^{\infty} D_k$ is irreducible with stationary probability vector π . We assume that $D \neq D_0$, which ensures that the matrix transform $[sI - D_0]^{-1}$ exists for all s with $\text{Re } s \geq 0$. The arrival rate λ is given by $\lambda = \pi d$, where $d = \sum_{k=1}^{\infty} k D_k e$ and e is a conformable column vector with every component unity. In the derivation of moment formulæ, we assume that the matrix series $\sum_{k=0}^{\infty} k^2 D_k$ converges.

Let $P(n, t)$ be the matrix whose (i, j) entry is $P_{i,j}(n, t)$, the conditional probability that n arrivals occur in $(0, t]$ and the arrival phase is j at time t , given that the phase at time 0 is i . The matrix generating function $P^*(z, t) = \sum_{n=0}^{\infty} P(n, t) z^n$ ($|z| \leq 1$) of the

sequence of matrices $\{P(n, t)\}$ is then given by

$$P^*(z, t) = \exp[D(z)t] \quad \text{for } |z| \leq 1, t \geq 0, \tag{2.1}$$

where $D(z) = \sum_{k=0}^{\infty} z^k D_k$.

2.2. Vacation schedule Let $L \geq 0$ be a specified integer. If after a service completion there are $n \geq 0$ customers in the system, then with probability ν_n another service starts and with the complementary probability $1 - \nu_n$ a vacation of length V_n starts. Here $\nu_0 = 0$. Further, the random variables V_n have a common distribution for $n > L$ and we write generically V_{L+1} for a random variable with this distribution. If after a vacation-completion epoch there are $n \geq 1$ customers in the system, a service starts immediately, while if there are no customers in the system, another vacation starts (with length V_0). We denote by $\hat{V}_n(t) = P(V_n \leq t)$ the distribution function of the random variable V_n ($0 \leq n \leq L + 1$).

The vacation schedule described above subsumes the exhaustive service discipline ($\nu_0 = 0$ and $\nu_n = 1$ for $n \geq 1$) and the Bernoulli schedule ($\nu_0 = 0$, while $\nu_n = \nu$ and the random variables V_n have a common distribution for $n > 0$).

2.3. Service times The distribution function for a service time is denoted by $\hat{B}(x)$ and its Laplace-Stieltjes transform by $B(s)$. We write μ for the mean service time. Suppose that just after a service completion $n \geq 1$ customers are present. We represent the distribution function of the time to the completion of the next service by \hat{B}_n for $1 \leq n \leq L$ and by \hat{B}_{L+1} for $n > L$. These “effective conditional service-time distribution functions” are then given by

$$\hat{B}_n(t) = \nu_n \hat{B}(t) + (1 - \nu_n) \hat{V}_n * \hat{B}(t) \quad (1 \leq n \leq L + 1),$$

where as usual $*$ denotes convolution. We write μ_n for the mean of the distribution given by \hat{B}_n ($1 \leq n \leq L + 1$). We put $\rho = \lambda \mu_{L+1}$ and throughout the paper assume that $\rho < 1$, so that the process possesses proper stationary behaviour.

2.4. The renewal process We are now in a position to address the basic renewal process. Let $X(t)$ and $J(t)$ denote respectively the number of customers in the system and the arrival phase at time t . We denote by τ_n ($n \geq 0$) the instants of successive departures from the system, with $\tau_0 = 0$. By X_n and J_n we signify respectively the number of customers in the system and the phase of the arrival process immediately after τ_n . Then $\{X_n, J_n, \tau_n - \tau_{n-1}, n \geq 1\}$ is a Markov renewal sequence with transition

probability matrix $\hat{Q}(x)$ given by

$$\hat{Q}(x) = \begin{pmatrix} \hat{A}_{0,0}(x) & \hat{A}_{0,1}(x) & \hat{A}_{0,2}(x) & \hat{A}_{0,3}(x) & \cdots & \cdots \\ \hat{A}_{1,0}(x) & \hat{A}_{1,1}(x) & \hat{A}_{1,2}(x) & \hat{A}_{1,3}(x) & \cdots & \cdots \\ 0 & \hat{A}_{2,0}(x) & \hat{A}_{2,1}(x) & \hat{A}_{2,2}(x) & \cdots & \cdots \\ 0 & 0 & \hat{A}_{3,0}(x) & \hat{A}_{3,1}(x) & \cdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & & \\ 0 & \cdots & \hat{A}_{L,0}(x) & \hat{A}_{L,1}(x) & \hat{A}_{L,2}(x) & \cdots \\ 0 & \cdots & 0 & \hat{A}_{L+1,0}(x) & \hat{A}_{L+1,1}(x) & \cdots \\ 0 & \cdots & 0 & 0 & \hat{A}_{L+1,0}(x) & \cdots \\ \vdots & \cdots & \vdots & \vdots & \vdots & \end{pmatrix}.$$

Here $\hat{A}_{k,n}(x)$ is the $m \times m$ matrix whose (i, j) entry is the probability, given a departure at time 0 which left k customers in the system and the arrival process in phase i , that the next departure occurs no later than time x , that it leaves the arrival process in phase j , and that during that time there are n arrivals.

The matrices $\hat{A}_{k,n}(x)$ are given by

$$\begin{aligned} \hat{A}_{k,n}(x) &= \int_0^x P(n, t) d\hat{B}_k(t) \quad (1 \leq k \leq L + 1, n \geq 0), \\ \hat{A}_{0,n}(x) &= \sum_{k=0}^{\infty} \sum_{i=1}^{n+1} \int_0^x d\hat{V}_0^{(k)}(u) e^{D_0 u} \int_{v=0}^{x-u} P(i, v) d\hat{V}_0(v) \\ &\quad \times \int_{w=0}^{x-u-v} P(n - i + 1)(w) d\hat{B}(w), \end{aligned} \tag{2.2}$$

where $\hat{V}_0^{(k)}$ is the k -fold convolution of $\hat{V}_0(t)$ with itself.

The joint summand and integrand on the right-hand side of (2.2) corresponds to there being k vacations with no arrivals, the k th vacation ending at time u , the next vacation being of length v and i customers arriving during that vacation, the first service time of the busy period being of length w and $n - i + 1$ customers arriving during that service.

We define $\hat{W}_n(k, x)$ as the $m \times m$ matrix whose (i, j) entry is the conditional probability, given a vacation begins at time 0 with n customers in the system and the phase of the arrival process is i , that the end of the vacation occurs no later than time x , that it ends with the arrival process in phase j , and that during the vacation there are k arrivals.

Let $W_n(z, s)$ be the double transform of $\hat{W}_n(k, x)$, that is,

$$W_n(z, s) = \sum_{k=0}^{\infty} z^k \int_0^{\infty} e^{-sx} d\hat{W}_n(k, x) \quad \text{for } |z| \leq 1, \text{ Re } s \geq 0,$$

and put $W_n(z) = W_n(z, 0)$ and $W_n = W_n(1, 0)$. Then from the definition of $P(n, t)$ and (2.1), we deduce that

$$W_n(z, s) = \int_0^\infty e^{-[sI - D(z)]x} d\hat{V}_n(x) \equiv V_n(sI - D(z)) \quad (1 \leq n \leq L + 1).$$

We shall make use of the transform matrices

$$A_{k,n}(s) = \int_0^\infty e^{-st} d\hat{A}_{k,n}(t), \quad A_k(z, s) = \sum_{j=0}^\infty z^j A_{k,n}(s)$$

and for notational convenience set

$$A_{k,n} = A_{k,n}(0), \quad A_k = A_k(1, 0), \quad A_k(z) = A_k(z, 0).$$

Arguments analogous to those in Lucantoni, Meier–Hellstern and Neuts [6] lead readily to the relations

$$A_k(z, s) = [v_k + (1 - v_k)W_k(z, s)]A(z, s) \quad (1 \leq k \leq L + 1), \quad (2.3)$$

$$A_0(z, s) = \frac{1}{z} [I - W_0(0, s)]^{-1} [W_0(z, s) - W_0(0, s)]A(z, s),$$

where $A(z, s) = \int_0^\infty \exp[-(sI - D(z))x] d\hat{B}(x)$.

In the following sections we shall make some use of mean values. Let α_k denote the row vector whose i -th entry is the mean number of arrivals during an effective service, conditional on that service having begun with k customers present and the arrival process in phase i . Then for $1 \leq k \leq L + 1$,

$$\alpha_k = \left. \frac{dA_k(z, 0)}{dz} \right|_{z=1} e.$$

Direct calculation from (2.3) yields

$$\alpha_k = [v_k I + (1 - v_k)W_k] \alpha + (1 - v_k)W'_k(1)e \quad (1 \leq k \leq L + 1),$$

where

$$\alpha = \lambda e + (A - I)(e\pi - D)^{-1}d,$$

$$W'_k(1)e = \lambda E(V_k)e + (W_k - I)(e\pi - D)^{-1}d.$$

3. The stationary queue length at departure epochs

The transition probability matrix with respect to departure epochs is $\hat{Q}(\infty)$. We represent the stationary probability vector \mathbf{p} of $\hat{Q}(\infty)$ in the partitioned form $\mathbf{p} = (\mathbf{p}_0, \mathbf{p}_1, \dots)$, where each \mathbf{p}_i is an m -vector. The eigenvalue equations for \mathbf{p}_i can be expanded as

$$\mathbf{p}_i = \begin{cases} \mathbf{p}_0 A_{0,i} + \sum_{j=1}^{i+1} \mathbf{p}_j A_{j,i-j+1}, & \text{if } 0 \leq i < L \\ \mathbf{p}_0 A_{0,i} + \sum_{j=1}^L \mathbf{p}_j A_{j,i-j+1} + \sum_{j=L+1}^{i+1} \mathbf{p}_j A_{L+1,i-j+1}, & \text{if } i \geq L. \end{cases}$$

If $\mathbf{p}(z) = \sum_{i=0}^{\infty} \mathbf{p}_i z^i$, these relations may be expressed as

$$\mathbf{p}(z)[zI - A_{L+1}(z)] = \mathbf{p}_0 A_0(z)(1 - z) + \sum_{i=0}^L \mathbf{p}_i z^i [A_i(z) - A_{L+1}(z)]. \tag{3.1}$$

We wish to find the unknown vectors $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_L$ so that the generating function $\mathbf{p}(z)$ is completely determined. This we achieve in Lemma 1 below. For $k \geq 1, x \geq 1$, let $\hat{G}^{(r)}(k, x)$ be the $m \times m$ matrix the (j, j') entry of which is the probability that the first passage from state $(L + i + r, j)$ to state $(L + i, j')$ ($1 \leq j, j' \leq m$) occurs in exactly k transitions and takes no more than time x , with $(L + i, j')$ being the first state visited in level $L + i$. For convenience we set $\hat{G}(k, x) := \hat{G}^{(1)}(k, x)$.

By a first-passage argument (Neuts [10]), the joint transform matrix $G(z, s)$ defined by

$$G(z, s) = \sum_{k=1}^{\infty} \int_0^{\infty} e^{-sx} d\hat{G}^{(1)}(k, x) z^k \quad (|z| \leq 1, \text{Re } s \geq 0)$$

satisfies the nonlinear matrix equation

$$G(z, s) = z \sum_{k=0}^{\infty} A_{L+1,k}(s) G^k(z, s).$$

Further, if $D(G(z, s)) = \sum_{k=0}^{\infty} D_k G^k(z, s)$, then $G(z, s)$ also satisfies

$$G(z, s) = z \int_0^{\infty} e^{-[sI - D(G(z,s))]x} d\hat{B}_{L+1}(x) \equiv z B_{L+1}(sI - D(G(z, s)))$$

(cf. Lucantoni and Neuts [7]).

In the remainder of the paper $G(z, s)$ will be used in the derivation of a number of quantities of interest. However an inspection of final formulæ relating to queue lengths will show that these depend only on $G(1, s)$ and $G := G(1, 0)$. The efficient determination of these is considered in Lucantoni, Chaudhry and Whitt [5].

It is well-known (see Neuts [10]) that the matrix G has the following probabilistic interpretation. For $\nu \geq 1$ and $i \geq 0$, the (j, k) entry of G^ν is the probability that the Markov chain $\{(X_n, J_n), n \geq 0\}$ with transition probability matrix $\hat{Q}(\infty)$ eventually visits level $L + i$ by entering the specific state $(L + i, k)$, given that it starts in the state $(L + i + \nu, j)$. We note also that since $\rho < 1$, the matrix G is stochastic and its invariant vector \mathbf{g} satisfies $\mathbf{g}G = \mathbf{g}$ and $\mathbf{g}\mathbf{e} = 1$. This leads to the following result.

LEMMA 1. Let $(\mathbf{X}^*, \mathbf{J}^*) = \{(X_n^*, J_n^*), n \geq 0\}$ denote the censored Markov chain obtained by embedding $\{(X_n, J_n), n \geq 0\}$ at the epochs when it visits the set of states $\{(i, j) : 0 \leq i \leq L, 1 \leq j \leq m\}$. Then the transition probability matrix Q^* of $(\mathbf{X}^*, \mathbf{J}^*)$ is given by

$$Q^* = \begin{pmatrix} A_{0,0} & A_{0,1} & A_{0,2} & \cdots & A_{0,L-1} & \bar{A}_{0,L} \\ A_{1,0} & A_{1,1} & A_{1,2} & \cdots & A_{1,L-1} & \bar{A}_{1,L} \\ 0 & A_{2,0} & A_{2,1} & \cdots & A_{2,L-2} & \bar{A}_{2,L-1} \\ 0 & 0 & A_{3,0} & \cdots & A_{3,L-3} & \bar{A}_{3,L-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{L,0} & \bar{A}_{L,1} \end{pmatrix},$$

If the invariant probability vector of Q^* is, in partitioned form, $\mathbf{x} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_L]$, where \mathbf{x}_i is an m -vector, then the vectors \mathbf{p}_i take the form

$$\mathbf{p}_i = c\mathbf{x}_i \quad (0 \leq i \leq L), \tag{3.2}$$

where

$$c = (1 - \rho) / \left\{ U_x(1)[I - A_{L+1} + \mathbf{e}\boldsymbol{\pi}]^{-1}\boldsymbol{\alpha}_{L+1} + \mathbf{x}_0\mathbf{e} + \sum_{i=0}^L \mathbf{x}_i(\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_{L+1}) \right\} \tag{3.3}$$

and $U_x(1) = \sum_{i=0}^L \mathbf{x}_i(A_i - A_{L+1})$.

PROOF. That the transition probability matrix of $(\mathbf{X}^*, \mathbf{J}^*)$ is given by Q^* follows from the probabilistic interpretation of G . Relation (3.2) ensues from the fact that the vector $(\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_L)$ is an eigenvector of Q^* corresponding to eigenvalue unity.

Define

$$U_x(z) = \mathbf{x}_0(z - 1)A_0(z) + \sum_{i=0}^L z^i \mathbf{x}_i[A_i(z) - A_{L+1}(z)].$$

On setting $z = 1$ in (3.1) and adding $p(1)e\pi = \pi$ to both sides, we have since $I - A_{L+1} + e\pi$ is nonsingular that

$$p(1) = (\pi + U(1))[I - A_{L+1}(1) + e\pi]^{-1}. \tag{3.4}$$

From (3.1) and (3.2), we derive

$$\frac{p(z)[zI - A_{L+1}(z)]e}{z - 1} = c \left[x_0e + \sum_{i=0}^L x_i z^i \frac{[A_i(z) - A_{L+1}(z)]e}{z - 1} \right].$$

Letting $z \rightarrow 1$ gives

$$p(1)(e - \alpha_{L+1}) = c \left[x_0e + \sum_{i=0}^L (\alpha_i - \alpha_{L+1}) \right]. \tag{3.5}$$

Relation (3.3) follows from (3.4), (3.5) and the fact that $\pi\alpha_{L+1} = 1 - \rho$, and we are done.

Now we derive the mean queue size at departure epochs. Define

$$U(z) = p_0(z - 1)A_0(z) + \sum_{i=0}^L z^i p_i(A_i(z) - A_{L+1}(z)).$$

Differentiation of (3.1) yields

$$p'(z)(zI - A(z)) + p(z)(I - A'_{L+1}(z)) = U'(z). \tag{3.6}$$

Setting $z = 1$ and adding $p'(1)e\pi$ to both sides give

$$p'(1) = p'(1)e\pi + [U'(1) - p(1)(I - A'_{L+1}(1))](I - A_{L+1} + e\pi)^{-1}. \tag{3.7}$$

Differentiation of (3.6) at $z = 1$ and postmultiplication by e provide

$$p'(1)\alpha_{L+1} = p'(1)e - \frac{1}{2}[p(1)A''_{L+1}(1) + U''(1)]e.$$

On postmultiplying (3.7) by α_{L+1} , we derive that the stationary mean queue length is

$$p'(1)e = \frac{1}{2(1 - \rho)} \left[p(1)A''_{L+1}(1)e + U''(1)e + 2[U' - p(1)(I - A'_{L+1})](I - A_{L+1}(1) + e\pi)^{-1} \right].$$

4. Hitting times

In this section we consider first-passage times from level i to level $i - 1$, those from level $i - 1$ to level i and the recurrence times for level i . These will be employed in Section 5.

4.1. First-passage times from level i to level $i - 1$ For $i > i' \geq 0, k \geq 1, x \geq 0$, let $\hat{G}^{[i,i']}(k, x)$ be the $m \times m$ matrix whose (j, j') entry is the probability that the first passage from state (i, j) to state (i', j') ($1 \leq j, j' \leq m$) occurs in exactly k transitions and takes time no more than x and that (i', j') is the first state visited in level i' . For notational simplicity we write $\hat{G}^{[i]}(k, x)$ for $\hat{G}^{[i+1,i]}(k, x)$. Of course $\hat{G}^{[i]}(k, x)$ has for each $i \geq L$ the common value $\hat{G}(k, x)$ introduced in the previous section. We define the transform matrix $G^{[i,i']}(z, s)$ of $\hat{G}^{[i,i']}(k, x)$ by

$$G^{[i,i']}(z, s) = \sum_{k=1}^{\infty} \int_0^{\infty} e^{-sx} d\hat{G}(k, x) z^k, \quad \text{for } |z| \leq 1, \text{Re } s \geq 0$$

and let $G^{[i,i']}(1, 0) = G^{[i,i']}(1, 0)$.

By conditioning on the time and destination of the first transition, we derive from the law of total probability the recursive formula

$$G^{[k]}(z, s) = zA_{k+1,0}(s) + z \sum_{n=1}^{\infty} A_{k+1,n}(s)G^{[k+n,k]}(z, s) \quad \text{for } 0 \leq k < L, \quad (4.1)$$

where

$$G^{[k+n,k]}(z, s) = G^{[k+n-1]}(z, s)G^{[k+n-2]}(z, s) \cdots G^{[k]}(z, s) \\ = \begin{cases} [G(z, s)]^{k+n-L} \prod_{i=k}^{L-1} \downarrow G^{[i]}(z, s) & \text{if } k+n > L \\ \prod_{i=k}^{k+n-1} \downarrow G^{[i]}(z, s) & \text{if } k+n \leq L. \end{cases}$$

Here and subsequently $\prod_{i=k}^j \downarrow C_k$ denotes the matrix product $C_j C_{j-1} \cdots C_k$. Likewise

we shall write $\prod_{i=k}^j \uparrow C_k$ for the product $C_k C_{k+1} \cdots C_j$.

The solution of (4.1) may be effected as follows. First we extend the definition of $\bar{A}_{i,L-i+1}$ to

$$\bar{A}_{i,L-i+1}(z, s) := \sum_{k=0}^{\infty} A_{i,L-i+1+k}(s)G^k(z, s) \quad (1 \leq i \leq L).$$

In terms of this, (4.1) can be expressed as

$$G^{[k]}(z, s) = z \sum_{n=0}^{L-k-1} A_{k+1,n}(s) \prod_{r=k}^{k+n-1} \downarrow G^{[r]}(z, s) + z \bar{A}_{k+1,L-k}(z, s) \prod_{r=k}^{L-1} \downarrow G^{[r]}(z, s) \tag{4.2}$$

for $1 \leq k \leq L$, where as subsequently we interpret the empty product as being an identity matrix.

The matrix Q^* is irreducible and hence $\sum_{i=1}^{L-k} A_{k,i} + \bar{A}_{k,L-k+1}$ is strictly substochastic for $1 \leq k \leq L$. Thus, as each $G^{[r]}(0)$ is stochastic, we must have that

$$\sum_{n=1}^{L-k-1} A_{k+1,n}(0) \prod_{r=k+1}^{k+n-1} \downarrow G^{[r]}(1, 0) + \bar{A}_{k+1,L-k}(1, 0) \prod_{r=k+1}^{L-1} \downarrow G^{[r]}(1, 0) \quad (0 \leq k < L)$$

is strictly substochastic and so has spectral radius less than unity. Accordingly, for $|z| \leq 1, \text{Re } s \geq 0$,

$$z \sum_{n=1}^{L-k-1} A_{k+1,n}(s) \prod_{r=k+1}^{k+n-1} \downarrow G^{[r]}(z, s) + z \bar{A}_{k+1,L-k}(z, s) \prod_{r=k+1}^{L-1} \downarrow G^{[r]}(z, s)$$

has spectral radius less than unity and

$$I - z \sum_{n=1}^{L-k-1} A_{k+1,n}(s) \prod_{r=k+1}^{k+n-1} \downarrow G^{[r]}(z, s) - z \bar{A}_{k+1,L-k}(z, s) \prod_{r=k+1}^{L-1} \downarrow G^{[r]}(z, s)$$

is invertible for $0 \leq k < L$.

Hence (4.2) may be solved explicitly by a downwards recursion *via*

$$G^{[L-1]}(z, s) = z A_{L,0}(s) [I - z \bar{A}_{L,1}(z, s)]^{-1} \quad (|z| \leq 1, \text{Re } s \geq 0),$$

$$G^{[k]}(z, s) = z A_{k+1,0}(s) \left[I - z \sum_{n=1}^{L-k-1} A_{k+1,n}(s) \prod_{r=k+1}^{k+n-1} \downarrow G^{[r]}(z, s) - z \bar{A}_{k+1,L-k}(z, s) \prod_{r=k+1}^{L-1} \downarrow G^{[r]}(z, s) \right]^{-1} \quad \text{for } 0 \leq k < L - 1.$$

4.2. First-passage times from level $i - 1$ to level i We denote by $\hat{H}^{[n]}(k, x)$ the $m \times m$ matrix whose (j, j') entry is the conditional probability that the Markov renewal process, starting in state $(n - 1, j)$ ($1 \leq j \leq m$), reaches state (n, j') ($1 \leq j' \leq m$) after exactly $k \geq 1$ transitions, taking no more than time $x \geq 0$. We define the joint transform matrix

$$H^{[n]}(z, s) := \sum_{k=1}^{\infty} z^k \int_0^{\infty} e^{-sx} d\hat{H}^{[n]}(k, x) \quad \text{for } |z| \leq 1, \text{Re } s \geq 1.$$

By an application of the law of total probability conditioning on the first transition of the Markov renewal process, we have the formulæ

$$\begin{aligned}
 H^{[1]}(z, s) &= [I - zA_{0,0}(s)]^{-1}z \left[\sum_{j=1}^{\infty} A_{0,j}(s)G^{[j,1]}(z, s) \right], \\
 H^{[i]}(z, s) &= [I - zA_{i-1,0}(s)H^{[i-1]}(z, s) - zA_{i-1,1}(s)]^{-1} \\
 &\quad \times z \sum_{j=2}^{\infty} A_{i-1,j}(s)G^{[i+j-2,i]}(z, s) \quad \text{for } i > 1.
 \end{aligned}$$

4.3. Recurrence times for level i Next we consider the return-time distribution of level i . For $i \geq 0, k \geq 1$ and $x \geq 0$, define $\hat{K}^{[i]}(k, x)$ to be the $m \times m$ matrix whose (j, j') entry ($1 \leq j, j' \leq m$) is the conditional probability that the Markov renewal process, starting in state (i, j) , returns to level i for the first time in exactly k transitions, taking time no more than time x , by hitting the state (i, j') . The joint transform matrix of matrix $\hat{K}^{[i]}(k, x)$ is defined by

$$K^{[i]}(z, s) = \sum_{k=1}^{\infty} z^k \int_0^{\infty} e^{-sx} d\hat{K}^{[i]}(k, x) \quad \text{for } |z| \leq 1, \text{Re } s \geq 1.$$

A first-passage argument shows that the transforms $K^{[i]}(z, s)$ are given by

$$\begin{aligned}
 K^{[0]}(z, s) &= zA_{0,0}(s) + z \sum_{j=1}^{\infty} A_{0,j}(s)G^{[j,0]}(z, s), \\
 K^{[i]}(z, s) &= zA_{i,0}(s)H^{[i]}(z, s) + z \sum_{j=1}^{\infty} A_{i,j}(s)G^{[i+j-1,i]}(z, s) \quad (i \geq 1).
 \end{aligned}$$

5. The queue length at an arbitrary time

In this section we derive the transient queue-length distribution at time t and its limiting distribution. This is accomplished by a classical argument based on Markov renewal processes. Consider the continuous-parameter process $\{(X(t), J(t)), t \geq 0\}$ and fix the initial state $X(0) = i_0, J(0) = j_0$.

Let $\hat{M}(t)$ be the matrix renewal function whose generic component $\hat{M}_{(i_0, j_0), (i, j)}(t)$ denotes the conditional expected number of visits to the state (i, j) ($i \geq 0, 1 \leq j \leq m$) in the interval $[0, t]$, given that $X(0) = i_0, J(0) = j_0$. We use $\hat{M}_k(t)$ to denote the k th row vector $\left(\hat{M}_{(i_0, j_0), (k, 1)}(t), \hat{M}_{(i_0, j_0), (k, 2)}(t), \dots, \hat{M}_{(i_0, j_0), (k, m)}(t) \right)$ of $\hat{M}(t)$ and introduce the transforms

$$M_k(s) = \int_0^{\infty} e^{-st} d\hat{M}_k(t), \quad M(z, s) = \sum_{k=0}^{\infty} z^k M_k(s)$$

for $\text{Re } s \geq 0$ and $|z| \leq 1$.

Now we determine $M(z, s)$ and $M_k(s)$ ($0 \leq k \leq L$). From the classical theory of Markov renewal process (see, for example, Çinlar [1, Chapt. 10]) we know that

$$\hat{M}_i(t) = \begin{cases} \delta_{i_0,i} e_{j_0} \Delta(t) + \hat{M}_0 * \hat{A}_{0,i}(t) + \sum_{k=1}^{i+1} \hat{M}_k * \hat{A}_{k,i+1-k}(t) & \text{for } 0 \leq i \leq L - 1 \\ \delta_{i_0,i} e_{j_0} \Delta(t) + \hat{M}_0 * \hat{A}_{0,i}(t) + \sum_{k=1}^L \hat{M}_k * \hat{A}_{k,i+1-k}(t) \\ \quad + \sum_{k=L+1}^{i+1} \hat{M}_k * \hat{A}_{L+1,i+1-k}(t) & \text{for } i \geq L. \end{cases} \tag{5.1}$$

Here $\delta_{i,j}$ is, as usual, the Kronecker delta, e_{j_0} ($1 \leq j_0 \leq m$) is the m -vector whose j_0 component is unity and whose other components are all zero, and Δ is the unit step function, taking values unity for nonnegative arguments and zero for negative arguments.

On taking Laplace transforms in (5.1) and forming generating functions, we derive

$$M(z, s) = \left[z^{i_0+1} e_{j_0} - (1 - z)M_0(s)A_0(z, s) + \sum_{j=0}^L z^j M_j(s)(A_j(z, s) - A_{L+1}(z, s)) \right] \times [zI - A_{L+1}(z, s)]^{-1}. \tag{5.2}$$

By the theory of delayed renewal processes, we readily evaluate each $M_k(s)$, in terms of quantities determined in the previous section, as

$$M_k(s) = \begin{cases} e_{j_0} G^{[i_0,k]}(1, s)[I - K^{[k]}(1, s)]^{-1} & \text{if } i_0 > k \\ e_{j_0}[I - K^{[k]}(1, s)]^{-1} & \text{if } i_0 = k \\ e_{j_0} \prod_{l=i_0+1}^k \uparrow H^{[l]}(1, s)[I - K^{[k]}(1, s)]^{-1} & \text{if } i_0 < k. \end{cases}$$

The transient joint distribution of the queue length and the arrival phase is given by the conditional probabilities

$$\hat{q}(k, j; t) := P(X(t) = k, J(t) = j | X(0) = i_0, J(0) = j_0) \quad (k \geq 0, 1 \leq j \leq m, t \geq 0).$$

Let $\hat{q}_k(t)$ be the m -vector with components $\hat{q}(k, j; t)$ ($1 \leq j \leq m$). By conditioning on the state of the Markov renewal process at the epoch of the last departure before time t , we find from the law of total probability that

$$\hat{q}_0(t) = \int_0^t d\hat{M}_0(u) e^{D_0(t-u)}, \tag{5.3a}$$

$$\begin{aligned}
 \hat{q}_j(t) &= \sum_{n=0}^{\infty} \int_0^t d\hat{M}_0(u) \int_0^{t-u} d\hat{V}_0^{(n)}(v) e^{D_0 v} P(j, t-u-v) (1 - \hat{V}_0(t-u-v)) \\
 &+ \sum_{n=0}^{\infty} \int_0^t dM_0(u) \int_0^{t-u} d\hat{V}_0^{(n)}(v) e^{D_0 v} \\
 &\times \sum_{k=1}^j \int_0^{t-u-v} d\hat{V}_0(\tau) P(k, \tau) P(j-k, t-u-v-\tau) (1 - \hat{B}(t-u-v-\tau)) \\
 &+ \sum_{k=1}^j \int_0^t d\hat{M}_k(u) P(j-k, t-u) (1 - \hat{B}_k(t-u)), \quad j \geq 1. \tag{5.3b}
 \end{aligned}$$

The first, second and third terms in (5.3b) correspond respectively to the cases where t falls during a vacation after the system becomes empty, during the first service of the first busy period after the system becomes empty, and during the second or later service time (including vacation time, if any) of a busy period. Taking Laplace transforms in (5.3), we have

$$\begin{aligned}
 q_0(s) &= M_0(s)(sI - D_0)^{-1}, \\
 q_j(s) &= M_0(s)[I - W_0(0, s)]^{-1} \int_0^{\infty} e^{-st} P(j, t) (1 - \hat{V}_0(t)) dt \\
 &+ M_0(s)[I - W_0(0, s)]^{-1} \sum_{k=1}^j \int_0^{\infty} e^{-st} P(k, t) d\hat{V}_0(t) \\
 &\times \int_0^{\infty} e^{-st} P(j-k, t) (1 - \hat{B}(t)) dt \\
 &+ \sum_{k=1}^j M_k(s) \int_0^{\infty} e^{-st} P(j-k, t) (1 - \hat{B}_k(t)) dt \quad (j \geq 1).
 \end{aligned}$$

After some routine calculation, we have from (5.2) that

$$\begin{aligned}
 q(z, s) &= \sum_{j=0}^{\infty} z^j q_j(s) \\
 &= \left[M(z, s)(I - A_{L+1}(z, s)) + M_0(s)(1-z)A_0(z, s) \right. \\
 &\quad \left. + \sum_{k=0}^L M_k(s)(A_{L+1}(z, s) - A_k(z, s))z^k \right] [sI - D(z)]^{-1} \tag{5.4} \\
 &= \left[z^{i_0+1} e_{j_0} [I - A_{L+1}(z, s)] - (1-z)^2 M_0(s)A_0(z, s) \right. \\
 &\quad \left. + (1-z) \sum_{k=0}^L M_k(s)(A_k(z, s) - A_{L+1}(z, s))z^k \right] \\
 &\quad \times [zI - A_{L+1}(z, s)]^{-1} [sI - D(z)]^{-1}.
 \end{aligned}$$

6. The stationary distribution $q_j = \lim_{t \rightarrow \infty} q_j(t)$

We have from Markov renewal theory that

$$\lim_{s \rightarrow 0} sM_k(s) = p_k/E,$$

where

$$E = \sum_{i=0}^L p_i \left[-\frac{\partial A_i(1, s)}{\partial s} \Big|_{s=0} \right] e + \left\{ p(1) - \sum_{i=0}^L p_i \right\} \left[-\frac{\partial A_{L+1}(1, s)}{\partial s} \Big|_{s=0} \right] e.$$

Thus we have from (5.2) and (3.1) that

$$\lim_{s \rightarrow 0} sM(z, s)(zI - A_{L+1}(z, s)) = \frac{1}{E} p(z)[zI - A_{L+1}(z)] \tag{6.1}$$

and from (5.4) and (6.1) that

$$q(z)(-D(z)) = \frac{1}{E} p(z)(1 - z).$$

On differentiating, setting $z = 1$ and postmultiplying by e , we have from the relations $q(1) = \pi$, $\pi D'(1)e = \lambda$ and $p(1)e = 1$ that $E = 1/\lambda$. Thus $q(z)$ and $p(z)$ are connected via

$$q(z)D(z) = \lambda p(z)(z - 1).$$

Comparing the coefficients of z^i shows that the vectors q_i and p_i are related by

$$q_0 = \lambda p_0(-D_0)^{-1},$$

$$q_i = \left[\sum_{j=0}^{i-1} q_j D_{i-j} - \lambda(p_{i-1} - p_i) \right] (-D_0)^{-1} \quad \text{for } i \geq 1.$$

From (5.4) the moments of the queue length distribution at arbitrary time can be expressed in terms of the moments of $p(z)$. Following a procedure of Lucantoni [4], the first moment of $q(z)$ is given by

$$q'(1) = q'(1)e\pi + [\lambda p(1) - \pi D'(1)](e\pi + D)^{-1},$$

where

$$q'(1)e = p'(1)e - \frac{1}{2\lambda} \pi d_2 + \left[\frac{\pi D'(1)}{\lambda} - p(1) \right] (e\pi + D)^{-1} d$$

and $d_2 = D''(1)e$.

Acknowledgement

The work of the first-named author was supported by a KOSEF Research Fellowship.

References

- [1] E. Çinlar, *Introduction to Stochastic Processes* (Prentice-Hall, Englewood Cliffs, N. J., 1975).
- [2] H. T. Doshi, "Queueing systems with vacations – a survey", *Queueing Syst.* **1** (1986) 29–66.
- [3] C. M. Harris and W. G. Marchal, "State dependence in $M/G/1$ server-vacation models", *Oper. Res.* **36** (1988) 560–565.
- [4] D. L. Lucantoni, "New results on the single server queue with a batch Markovian arrival process", *Stoch. Models* **7** (1991) 1–46.
- [5] D. L. Lucantoni, G. L. Choudhury and W. Whitt, "The transient $BMAP/G/1$ queue", *Stoch. Models* **10** (1994) 145–182.
- [6] D. L. Lucantoni, K. S. Meier-Hellstern and M. F. Neuts, "A single server queue with server vacations and a class of non-renewal arrival processes", *Adv. Appl. Prob.* **22** (1990) .
- [7] D. L. Lucantoni and M. F. Neuts, "Simpler proofs of some properties of the fundamental period of the $MAP/G/1$ queue", *J. Appl. Prob.* **31** (1994) 235–243.
- [8] D. L. Lucantoni and M. F. Neuts, "Some steady state distributions for the $MAP/SM/1$ queue", *Stoch. Models* **10** (1994) 575–598.
- [9] F. Machihara, "A queue with vacation depending on services", in *Symposium on Performance Models for Information Communication Networks* (ed. O. Hashida), (Univ. of Tsukuba, Japan, 1995), 25–37.
- [10] M. F. Neuts, *Structured stochastic matrices of $M/G/1$ type and their applications* (Marcel Dekker, New York, 1989).
- [11] Y. W. Shin, "Time-dependent analysis of the $M^X/G/1$ vacation queue with queue length dependent vacation schedule", submitted for publication.
- [12] H. Takagi, *Queueing analysis: a foundation of performance evaluation, Volume 1: vacation and priority systems, Part 1* (North-Holland, Amsterdam, 1991).
- [13] T. Takine and T. Hasegawa, "A batch $SPP/G/1$ queue with multiple vacations and exhaustive service discipline", *Telecomm. Syst.* **1** (1993) 195–215.