AN $M/M/s$ QUEUE WITH SERVERS' VACATIONS*

YONATAN LEVY AND URI YECHIALI

Department of Statistics, Tel Aviv University

Abstract

In an $M/M/s$ queueing system a server that completes service and finds no waiting units in line leaves for a *vacation* of an exponentially distributed duration. At the end of the vacation the server returns to the main system. Two models are analysed. In the first, a server returning to an empty queue takes immediately another vacation. In the second, only a single vacation is taken each time. For model 1, formulas for the distribution of the number of busy servers and the mean number of units in system, L, are derived. Numerical calculations indicate that L is very closely a linear function of the mean vacation time. Finally it is shown that model 2 may be analysed similarly to model 1.

Résumé

Dans une système de file d'attente $M/M/s$ si, après avoir complété le service d'une unité, un serveur se retrouve devant une file d'attente vide, alors il se retire pour un congé d'une durée ayant une distribution exponentielle. Une fois son congé terminé, ce dernier revient au système. Deux modèles sont étudiés. Dans le premier modèle, si au retour du serveur la file d'attente est vide, ce dernier se retire pour un nouveau congé. Dans le second modèle, un serveur ne peut se permettre deux périodes de congé successives. Pour le premier modèle, on dérive des formules pour la distribution du nombre de serveurs au travail et pour le nombre moyen d'unités dans le système, L. Des expériences numériques indiquent que L varie presque linéairement avec la durée moyenne d'un congé. Finalement on démontre qu'on peut réaliser une analyse similaire du second modèle.

1 Introduction

In this work we study an $M/M/s$ queueing system where servers that become idle leave for a random period of time called *vacation*. These vacations may be utilized to perform additional work assigned to the servers. We consider two models. In the first, a server returning to an empty queue immediately takes another vacation. In the second, only a single vacation is taken each time, i.e. a server takes another vacation only after having served at least one unit. A similar framework was studied by the authors[1] for the $M/G/1$ queue where different analytic techniques were used.

Previous studies on the control of multi-server queues were concerned with variable service capacity. Yadin and Naor[5] extended works by Romani[4] and by Moder and Phillips[3] by varying the service capacity of the system as a function of the number of waiting units and the recent history of the system. Our models differ mainly from these studies in that the return of a server from a vacation is *independent* of the queue length. Thus, the control of the system's performance may be achieved by changing the length of the vacation period.

*Received 29 April 1975; revised 23 September 1975, and 16 December 1975.

2 Model 1

We consider an $M/M/s$ system with arrival rate λ and service intensity μ per server. When a server completes service and no units are waiting in the queue the server leaves for a vacation whose duration is an exponentially distributed random variable with finite mean $1/\theta$. If the server finds an empty queue at the end of a vacation he immediately takes another vacation, etc.

The process can be formulated as a continuous-time Markov chain with a state space $\{(k, i) : k = 0, 1, 2, \dots, s; i \geqslant k\}$, where k denotes the number of busy servers and i the number of units in the system.

The steady state equations of the process are:

$$\lambda p_{00} = \mu p_{11}, \tag{1a}$$

$$(\lambda + s\theta)p_{0i} = \lambda p_{0,i-1} \qquad (i = 1, 2, \dots), \tag{1b}$$

$$(\lambda + k\mu)p_{kk} = (s - k + 1)\theta p_{k-1,k} + k\mu p_{k,k+1} + (k + 1)\mu p_{k+1,k+1} \qquad (k = 1, 2, \dots, s - 1), \tag{1c}$$

$$(\lambda + s\mu)p_{ss} = \theta p_{s-1,s} + s\mu p_{s,s+1}, \tag{1d}$$

$$[\lambda + k\mu + (s - k)\theta]p_{ki} = (s - k + 1)\theta p_{k-1,i} + \lambda p_{k,i-1} + k\mu p_{k,i+1} \qquad (k = 1, 2, \dots, s; i > k). \tag{1e}$$

Let $P_{k\bullet}$ $(k = 0, 1, 2, \dots, s)$ be the probability that k servers are busy, i.e.

$$p_{k\bullet} = \sum_{i=k}^{\infty} p_{ki}.$$

Equations for $p_{k\bullet}$ can be found as follows: Summation of equations (1b) for $i = 1, 2, \dots$ together with equation (1a) yields

$$s\theta(p_{0\bullet} - p_{00}) = \mu p_{11}. \tag{2}$$

For every $k = 1, 2, \dots, s - 1$, summation of equations (1e) over $i = k + 1$, $k + 2, \dots$ together with equation (1c) gives inductively

$$(s - k)\theta(p_{k\bullet} - p_{kk}) = (k + 1)\mu p_{k+1,k+1} \qquad (k = 0, 1, \dots, s - 1). \tag{3a}$$

Solving for $p_{k\bullet}$ gives the form

$$p_{k\bullet} = \frac{(k + 1)\mu}{(s - k)\theta} p_{k+1,k+1} + p_{kk} \qquad (k = 0, 1, \dots, s - 1). \tag{3b}$$

Equations (3b) together with

$$\sum_{k=0}^{s} p_{k\bullet} = 1 \tag{4}$$

and equation (1a) give a set of $(s + 2)$ linear equalities in $(2s + 2)$ variables (the variables being p_{kk} and $p_{k\bullet}$ for $k = 0, 1, \dots, s$).

Since the number of variables exceeds the number of equations, the set of equalities is not sufficient to yield a unique solution. We thus have to employ an indirect method that will provide us with the information necessary for the solution of the model.

Our approach will be to define partial generating functions and to exploit their properties and the information concentrated in them.

Similar techniques were used previously by Mitrani and Avi-Itzhak[2] and by Yechiali.[6] Those works emphasized the existence of a set of equations that may give a solution to the problem, but did not find explicit formulas for the unknown probabilities. Moreover, a proof that the set of equations derived via the generating functions is an independent one is not presented in the above mentioned works.

In the present work, we explicitly derive a unique solution to the set of probabilities $p_{k\bullet}$.

2.1 *Generating functions*

For each k define the partial generating function

$$G_k(z) = \sum_{i=k}^{\infty} p_{ki} z^i \qquad (k = 0, 1, \ldots, s; |z| \leqslant 1). \tag{5}$$

Multiplying the ith equation of (1b) by z^i, summing over all i, and adding (1a) yield

$$[\lambda(1 - z) + s\theta]G_0(z) = \mu p_{11} + s\theta p_{00}. \tag{6}$$

Similarly, for each $k = 1, 2, \ldots, s - 1$, by multiplying the corresponding equation of (1e) by z^i and adding (1c) multiplied by z^k we obtain

$$[\lambda z(1 - z) - k\mu(1 - z) + (s - k)\theta z]G_k(z) - (s - k + 1)\theta z G_{k-1}(z) =$$
$$(k + 1)\mu p_{k+1,k+1} z^{k+1} + [(s - k)\theta z - k\mu] p_{kk} z^k - (s - k + 1)\theta z p_{k-1,k-1} z^{k-1}$$
$$(k = 1, 2, \ldots, s - 1). \tag{7}$$

Replacing $(k + 1)\mu p_{k+1,k+1}$ in equations (6) and (7) by the left-hand side of (3) gives

$$[\lambda(1 - z) + s\theta]G_0(z) = s\theta p_{0\bullet}, \tag{8a}$$

$$[\lambda z(1 - z) - k\mu(1 - z) + (s - k)\theta z]G_k(z) - (s - k + 1)\theta z G_{k-1}(z) =$$
$$(s - k)\theta p_{k\bullet} z^{k+1} - (s - k + 1)\theta p_{k-1,\bullet} z^k \quad (k = 1, 2, \ldots, s - 1). \tag{8b}$$

In a similar way we can find for $k = s$,

$$[\lambda z(1 - z) - s\mu(1 - z)]G_s(z) - \theta z G_{s-1}(z) = -\theta p_{s-1,\bullet} z^s. \tag{9}$$

That is, equation (8b) holds for $k = s$ as well.

For $k = 0, 1, 2, \ldots, s$ let

$$f_k(z) = \lambda z(1 - z) - k\mu(1 - z), \tag{10a}$$

$$h_k(z) = (s - k)\theta z. \tag{10b}$$

Equations (8) and (9) are now rewritten as

$$[f_0(z) + h_0(z)]G_0(z) = h_0(z) p_{0\bullet} \tag{11a}$$

$$[f_k(z) + h_k(z)]G_k(z) - h_{k-1}(z)G_{k-1}(z) = h_k(z) p_{k\bullet} z^k - h_{k-1}(z) p_{k-1,\bullet} z^{k-1}$$
$$(k = 1, 2, \ldots, s). \tag{11b}$$

The set (11) enables us to calculate the $G_k(z)$'s recursively if the $(s+1)$ probabilities $p_{k\bullet}$ are known. Thus we have to find $(s+1)$ additional equations in the $(s+1)$ unknown $p_{k\bullet}$'s.

A first equation is obviously

$$\sum_{k=0}^{s} p_{k\bullet} = 1. \tag{12}$$

A second equation will be

$$\sum_{k=1}^{s} kp_{k\bullet} = \lambda/\mu. \tag{13}$$

To show (13) we proceed as follows: summation of the first k equations of (11) $(k = 0, 1, 2, \dots, s)$ gives

$$G_0(z) = h_0(z)p_{0\bullet}/(f_0(z) + h_0(z)). \tag{14a}$$

$$G_k(z) = [h_k(z)p_{k\bullet}z^k - \sum_{i=0}^{k-1} f_i(z)G_i(z)]/(f_k(z) + h_k(z))$$

$$(k = 1, 2, \dots, s-1). \tag{14b}$$

$$f_s(z)G_s(z) = -\sum_{k=0}^{s-1} f_k(z)G_k(z). \tag{14c}$$

Substituting (10a) for $f_k(z)$ and dividing by $(1-z)$, equation (14c) yields

$$\sum_{k=0}^{s} (k\mu - \lambda z)G_k(z) = 0. \tag{15}$$

Since $G_k(1) = p_{k\bullet}$, equations (12) and (15) imply that

$$\sum_{k=0}^{s} kp_{k\bullet} = \lambda/\mu.$$

That is, the average number of busy servers is λ/μ, as in the ordinary $M/M/s$ system. In other words, the average load carried by each server is $\lambda/s\mu$ per unit time, and as long as $\lambda/\mu < s$ the system will reach steady state.

To find the remaining $(s-1)$ equations let

$$A(z) = \left\| \begin{matrix} f_0(z)+h_0(z) & 0 & 0 & \cdots & 0 & 0 \\ -h_0(z) & f_1(z)+h_1(z) & 0 & \cdots & 0 & 0 \\ 0 & -h_1(z) & f_2(z)+h_2(z) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & f_{s-1}(z)+h_{s-1}(z) & 0 \\ 0 & 0 & 0 & \cdots & -h_{s-1}(z) & f_s(z)+h_s(z) \end{matrix} \right\|$$

$$\bar{g}(z) = \begin{bmatrix} G_0(z) \\ G_1(z) \\ \vdots \\ G_s(z) \end{bmatrix}, \quad \bar{b}(z) = \begin{bmatrix} b_0(z) \\ b_1(z) \\ \vdots \\ b_s(z) \end{bmatrix}$$

where $b_0(z) = h_0(z)p_{0\bullet}$

and $\qquad b_k(z) = h_k(z)p_{k\bullet}z^k - h_{k-1}(z)p_{k-1,\bullet}z^{k-1} \qquad (k = 1, 2, \ldots, s).$

Equation (11) when written in matrix form is

$$A(z)\bar{g}(z) = \bar{b}(z), \ |z| \leqslant 1. \tag{16}$$

Equation (16) is a set of $(s + 1)$ linear equations and for all values of z where $A(z)$ is non-singular Cramer's law implies

$$|A(z)|G_k(z) = |A_k(z)| \qquad (k = 0, 1, \ldots, s), \tag{17}$$

where $|A|$ stands for the determinant of matrix A, and the matrix $A_k(z)$ is obtained by replacing the kth column of $A(z)$ with $\bar{b}(z)$. Since $|A(z)|$ is a continuous function of z, equation (17) holds for every $|z| \leqslant 1$. Also, since for any $z \in [0, 1]$, $G_k(z) > 0$, then every z which is a root of $|A(z)|$ is also a root of $|A_k(z)|$.

Proposition 1

The polynomial $|A(z)|$ has exactly $(s - 1)$ distinct roots in the interval $(0, 1)$.

Proof

Clearly,

$$|A(z)| = \prod_{k=0}^{s} [f_k(z) + h_k(z)].$$

$|A(z)|$ is a polynomial of degree $2(s + 1)$ whose roots will be denoted by $\{z_l(k) | k = 0, 1, \ldots, s; l = 1, 2\}$, where $z_1(k)$ and $z_2(k)$ $(z_1(k) < z_2(k))$ are the roots of the quadratic equation

$$f_k(z) + h_k(z) = 0. \tag{18}$$

For $k = 0$, equation (18) means $\lambda z(1 - z) + s\theta z = 0$, and its roots are $z_1(0) = 0$ and $z_2(0) = (\lambda + s\theta)/\lambda > 1$. For $k = s$, equation (18) gives $\lambda z(1 - z) - s\mu(1 - z) = 0$, the roots of which are $z_1(s) = 1$ and $z_2(s) = s\mu/\lambda > 1$. We now show that for $k = 1, 2, \ldots, s - 1$, $0 < z_1(k) < 1$, $z_2(k) > 1$ and $z_1(k) \neq z_1(j)$ for $k \neq j$. For $0 < k < s$ equation (18) takes the form

$$\lambda z^2 - [(\lambda + k\mu) + (s - k)\theta]z + k\mu = 0. \tag{18$'$}$$

The roots of (18′) are

$$z_{1,2}(k) = [(\lambda + k\mu) + (s - k)\theta \pm \sqrt{Q_k}]/(2\lambda), \tag{19}$$

where

$$Q_k = (\lambda - k\mu)^2 + (s - k)^2\theta^2 + 2(\lambda + k\mu)(s - k)\theta. \tag{20}$$

For $\lambda > 0$, $\mu > 0$, $(s - k)\theta > 0$ the following inequalities hold

$$Q_k \begin{cases} > [(\lambda - k\mu) + (s - k)\theta]^2 & (21) \\ > [(k\mu - \lambda) + (s - k)\theta]^2 & (22) \\ < [(\lambda + k\mu) + (s - k)\theta]^2 & (23) \end{cases}$$

Since $z_2(k)$ and $z_1(k)$ correspond, respectively, to the plus and minus signs in (19), (21) implies that $z_2(k) > [\lambda + (s - k)\theta]/\lambda > 1$. By (22), $z_1(k) < [\lambda + k\mu + (s - k)\theta - (k\mu - \lambda + (s - k)\theta)]/(2\lambda) = 1$, and, by (23), $z_1(k) > 0$. Thus, we have shown that $|A(z)|$ has exactly $(s - 1)$ roots in the open interval (0, 1). It remains only to show that they are distinct.

For every $k = 1, 2, \ldots, s - 1$ the roots $z_1(k)$ and $z_2(k)$ satisfy

$$z_1(k)z_2(k) = k\mu/\lambda, \tag{24}$$

$$z_1(k) + z_2(k) = (\lambda + k\mu + (s - k)\theta)/\lambda = 1 + s\theta/\lambda + k(\mu - \theta)/\lambda. \tag{25}$$

If, for some $0 < k \neq j < s, z_1(k) = z_1(j) = z_0$ then, from (24), $z_2(k)/z_2(j) = k/j$. On the other hand $z_2(k)/z_2(j)$ can also be calculated using (25).

$$\frac{z_2(k)}{z_2(j)} = \frac{1 + \frac{s\theta}{\lambda} + k\frac{\mu - \theta}{\lambda} - z_0}{1 + \frac{s\theta}{\lambda} + j\frac{\mu - \theta}{\lambda} - z_0}. \tag{26}$$

This expression can be equal to k/j if and only if $1 + s\theta/\lambda - z_0 = 0$, which is impossible since $z_0 < 1$.

With this contradiction the proof is complete.

Let $z_j = z_1(j)$ $(j = 1, 2, \ldots, s - 1)$. Equation (17) implies

$$|A_k(z_j)| = 0, \qquad k = 1, \ldots, s; j = 1, 2, \ldots, s - 1. \tag{27}$$

Note that we delete $k = 0$ since $|A_0(z_j)| \equiv 0$ for any $\bar{b}(z)$.

However, for each z_j, the s homogeneous equations of (27) differ from each other only by a constant multiplier.

This follows since, by (17), for any pair $1 \leqslant i \neq k \leqslant s$ and for any $|z| \leqslant 1$, $|A_i(z)|/|A_k(z)| = G_i(z)/G_k(z)$. That is, the equation $|A_i(z_j)| = 0$ may be obtained by multiplying the equation $A_k(z_j) = 0$ by the positive constant $G_i(z_j)/G_k(z_j)$.

Thus, (27) yields only *one* equation for each z_j $(j = 1, 2, \ldots, s - 1)$, and together with (12) and (13) we have a set of $s + 1$ linear equations in the $s + 1$ unknowns $p_{k\bullet}$ $(k = 0, 1, 2, \ldots, s)$.

It now remains only to show that this set yields a unique solution.

For each z_j let $|A_s(z_j)| = 0$ represent the single equation just mentioned. By expanding $|A_s(z_j)|$ progressively by the $(j + 1)$th column, we obtain

$$|A_s(z_j)| = \sum_{i=0}^{j} \{b_i(z_j) \prod_{k=0}^{i-1} [f_k(z_j) + h_k(z_j)] \prod_{k=i}^{s-1} h_k(z_j)\} = 0$$
$$(j = 1, 2, \ldots, s - 1). \tag{28}$$

Since $b_i(z_j)$ is a linear combination of $p_{i\bullet}$ and $p_{i-1,\bullet}$, we can derive, by some elementary operations, the following recursive formula:

$$p_{j\bullet} = \sum_{i=0}^{j-1} f_i(z_j)D_{i,j-1}(z_j)p_{i\bullet}/h_j(z_j) \qquad (j = 1, 2, \ldots, s - 1), \tag{29}$$

where

$$D_{i,j}(z) = \prod_{k=i}^{j} \frac{(s-k)\theta}{f_k(z) + h_k(z)}.$$

This set of recursive equations, together with (13), determines *uniquely* the probabilities $p_{1\bullet}, p_{2\bullet}, \ldots, p_{s\bullet}$ as a function of $p_{0\bullet}$. Then $p_{0\bullet}$ is obtained from (12).

Moreover, it is easy to see that if $\lambda/\mu = s$ then $p_{s\bullet} = 1$.

Given the $p_{j\bullet}$'s, the generating functions $G_j(z)$'s are explicitly calculated from (14).

We also note that from (1b) we have

$$p_{0i} = \frac{s\theta}{\lambda + s\theta}\left(\frac{\lambda}{\lambda + s\theta}\right)^i p_{0\bullet}. \tag{30}$$

That is, the number of customers in the system when all servers are in vacation is geometrically distributed.

2.2 *Mean number of units in system*

Let $p_{\bullet i}$ denote the probability that i units are present in the system. Thus

$$p_{\bullet i} = \sum_{k=0}^{s} p_{ki} \qquad (i = 0, 1, 2, \ldots), \tag{31}$$

where

$$p_{ki} = 0 \text{ for } i < k.$$

The generating function of the number of units in system is given by

$$G(z) = \sum_{i=0}^{\infty} p_{\bullet i} z^i, \qquad |z| \leqslant 1. \tag{32}$$

Clearly, by (5),

$$G(z) = \sum_{k=0}^{s} G_k(z). \tag{33}$$

The mean number of units in system, L, is given by

$$L = G'(1) = \sum_{k=0}^{s} G_k'(1). \tag{34}$$

Differentiating (14a), (14b), and (15) yields respectively

$$G_0'(1) = \frac{\lambda}{s\theta} p_{0\bullet} \tag{35a}$$

$$G_k'(1) = \frac{\sum_{i=0}^{k} (\lambda - i\mu) p_{i\bullet}}{(s-k)\theta} + k p_{k\bullet} \qquad (k = 1, 2, \ldots, s-1), \tag{35b}$$

$$G_s'(1) = \frac{\lambda + \sum_{i=0}^{s-1} (\lambda - i\mu) p_{i\bullet}}{s\mu - \lambda} \tag{35c}$$

By substituting in (34) the terms of $G_k'(1)$ from (35) for $k = 0, 1, \ldots, s$ we obtain, after algebraic manipulations,

$$L = \frac{\rho}{1-\rho} + \frac{1}{1-\rho}\sum_{k=0}^{s-1}\left(1 - \frac{k}{s}\right)\left(\frac{\lambda - k\mu}{\theta} + k\right)p_{k\bullet}, \tag{36}$$

where

$$\rho = \lambda/(s\mu).$$

2.3 *An example:* $s = 2$

When the number of servers is $s = 2$ the set (29) yields a single equation, namely

$$p_{1\bullet} = \frac{\lambda(1 - z_1)}{\lambda z_1(1 - z_1) + 2\theta z_1} 2p_{0\bullet}.$$

z_1 is derived from (19) as

$$z_1 = \frac{1}{2\lambda}\{(\lambda + \mu + \theta) - [(\lambda + \mu + \theta)^2 - 4\lambda\mu]^{1/2}\}.$$

Together with (12) and (13) we obtain

$$p_{0\bullet} = \alpha[\lambda z_1(1 - z_1) + 2\theta z_1](1 - \rho),$$
$$p_{1\bullet} = \alpha 2\lambda(1 - z_1)(1 - \rho),$$
$$p_{2\bullet} = \alpha\lambda(2 - z_1)(1 - z_1)\rho,$$

where $\alpha = [\lambda(1 - z_1^2) + 2\theta z_1]^{-1}$.

The mean number of units in system is given by

$$L = \frac{\rho}{1-\rho} + \alpha\lambda[\lambda(1 - z_1) + \theta]/\theta.$$

2.4 *Numerical results*

Numerical calculations were performed by the authors to obtain values of L, the mean number of units in system, for values of s running up to 15, and for various values of the offered load, λ/μ, and of $1/\theta$. (The calculations were performed on a CDC 6600 computer of Tel-Aviv University; 164 seconds of CP time were required, for example, to calculate 18,000 distinct values of L for the

TABLE 1

MEAN NUMBER OF UNITS IN SYSTEM FOR VARIOUS VALUES OF s, ρ, AND $1/\theta$

	$s = 5$			$s = 10$			$s = 15$		
$1/\theta$ \ ρ	0.2	0.6	0.9	0.2	0.6	0.9	0.2	0.6	0.9
0.0	1.001	3.354	11.362	2.000	6.152	15.019	3.000	9.072	18.924
0.1	1.028	3.498	11.634	2.026	6.325	15.425	3.026	9.252	19.425
0.2	1.054	3.628	11.875	2.052	6.479	15.768	3.051	9.413	19.841
1.0	1.250	4.500	13.500	2.250	7.500	18.000	3.250	10.500	22.500
2.0	1.450	5.450	15.314	2.488	8.598	20.417	3.492	11.675	25.337
4.0	1.921	7.202	18.748	2.948	10.601	24.908	3.960	13.822	30.556
6.0	2.351	8.872	22.085	3.396	12.498	29.224	4.418	15.853	35.540
8.0	2.774	10.503	25.380	3.837	14.345	33.461	4.869	17.830	40.417
10.0	3.194	12.110	28.652	4.273	16.162	37.654	5.316	19.774	45.233

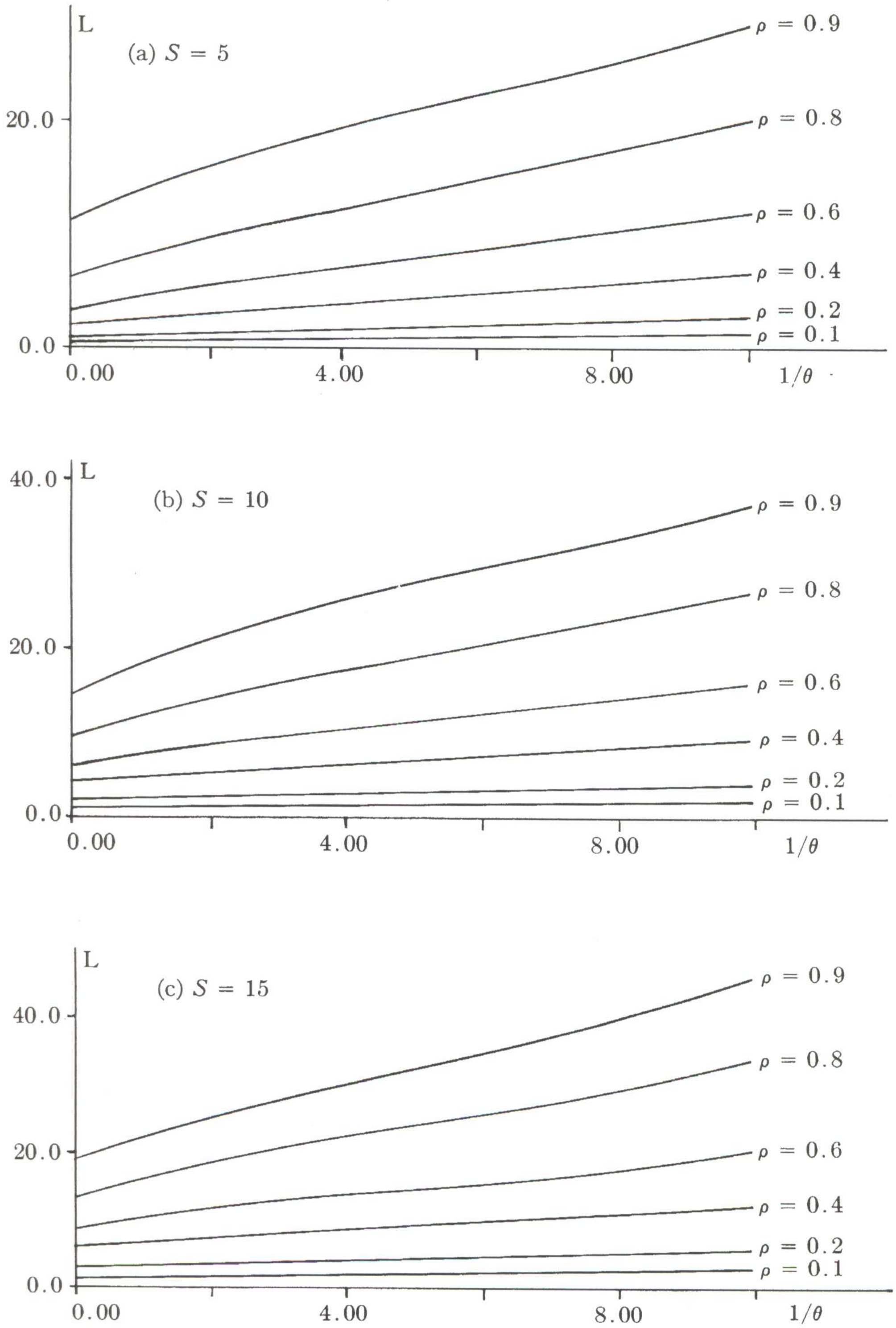

FIG. 1. The mean value of units in system, L, as a function of the mean vacation time, $1/\theta$, for various values of ρ.

following values of the parameters: $s = 5, 10, 15$; $\rho = 0.1, 0.2, 0.4, 0.6, 0.8, 0.9$; and $1/\theta = 0.01\ (0.01)\ 10.00$). Some selective results are exhibited in table 1. The numerical results show that L is closely a linear function of the mean vacation time for all values of λ/μ and s. Figure 1(a, b, c) demonstrates this near-linear relation.

3 Model 2

We now consider a variation of the model studied in Section 2. As before, the underlying structure is an $M/M/s$ system with servers' vacations. However, a server now takes only a single vacation at a time. When he returns to the main system he starts serving immediately if there are units waiting in the queue. If the queue is empty he waits until his turn to serve comes.

In the sequel we use the same notation as for the first model, i.e., the arrival rate is λ, the service rate per station is μ, and a vacation time is an exponentially distributed random variable with mean $1/\theta$.

The state space is now the set $\{(k, i) : i,k = 0, 1, 2, \ldots\}$ where this time k denotes the number of servers available for service – busy or idle – and i, as before, denotes the number of units in the system.

The balance equations in this case are

$$\text{(a)}\quad (\lambda + s\theta)p_{00} = \mu p_{11}, \tag{37}$$

$$\text{(b)}\quad (\lambda + s\theta)p_{0i} = \lambda p_{0,i-1} \qquad (i = 1, 2, \ldots),$$

$$\text{(c)}\quad (\lambda + (s-k)\theta)p_{k0} = \mu p_{k+1,1} + (s-k+1)\theta p_{k-1,0} \quad (k = 1, 2, \ldots, s-1),$$

$$\text{(d)}\quad (\lambda + (s-k)\theta + i\mu)p_{ki} = (i+1)\mu p_{k+1,i+1} + (s-k+1)\theta p_{k-1,i} + \lambda p_{k,i-1} + \delta_{ik} k\mu p_{k,k+1} \qquad (i = 1, 2, \ldots, k;\ k = 1, 2, \ldots, s-1),$$

$$\text{(e)}\quad (\lambda + (s-k)\theta + k\mu)p_{ki} = (s-k+1)\theta p_{k-1,i} + \lambda p_{k,i-1} + k\mu p_{k,i+1} \qquad (i = k+1, k+2, \ldots;\ k = 1, 2, \ldots, s-1),$$

$$\text{(f)}\quad (\lambda + i\mu)p_{si} = \theta p_{s-1,i} + (1-\delta_{i0})\lambda p_{s,i-1} \qquad (i = 0, 1, \ldots, s-1),$$

$$\text{(g)}\quad (\lambda + s\mu)p_{si} = \theta p_{s-1,i} + \lambda p_{s,i-1} + s\mu p_{s,i+1} \qquad (i = s, s+1, \ldots),$$

where
$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

As in model 1, we cannot solve the above system of equations directly and we shall have to use the generating functions technique. Thus, for each $k = 0, 1, \ldots, s$ we define

$$G_k(z) = \sum_{i=0}^{\infty} p_{ki} z^i.$$

From equations (37a, b) we get

$$[\lambda(1-z) + s\theta]G_0(z) = \mu p_{11}. \tag{38}$$

From equations (37, c, d, e) one derives

$$[\lambda z(1-z) - k\mu(1-z) + (s-k)\theta z]G_k(z) - (s-k+1)\theta z G_{k-1}(z)$$
$$= \mu \sum_{i=1}^{k+1} i p_{k+1,i} z^i - \mu z \sum_{i=1}^{k} i p_{ki} z^i - k\mu(1-z) \sum_{i=0}^{k} p_{ki} z^i$$
$$(k = 1, 2, \ldots, s-1). \quad (39)$$

Equations (37f, g) yield

$$[\lambda z(1-z) - s\mu(1-z)]G_s(z) - \theta z G_{s-1}(z)$$
$$= -\mu z \sum_{i=1}^{s} i p_{si} z^i - s\mu(1-z) \sum_{i=0}^{s} p_{si} z^i. \quad (40)$$

If we write equations (38), (39), and (40) in matrix form as was done in Section 2 (i.e. $A(z)\bar{g}(z) = \bar{b}(z)$), it is readily seen that although the vector $\bar{b}(z)$ is somewhat different, the matrix $A(z)$ is *exactly* the same one as in model 1. Hence, proposition 1 applies here too, and using (27) we have $(s-1)$ equations in the unknowns p_{ki} ($k = 1, 2, \ldots, s$; $i = 0, 1, \ldots, k$). From equations (37a), (37b), (37c), (37d), and (37f) we derive $\frac{1}{2}s(s+1)+1$ equations in the $(s+1)(s+2)/2$ unknowns p_{ki} ($k = 0, 1, \ldots, s$; $i = 0, 1, \ldots, k$). Letting $z = 1$ in (38), (39), and (40) we obtain other $(s+1)$ equations with (some of) the above probabilities and with additional $(s+1)$ probabilities $p_{k\bullet}$ ($k = 0, 1, \ldots, s$). Adding the equation

$$\sum_{k=0}^{s} p_{k\bullet} = 1$$

we altogether have a system of $[(s-1) + (\frac{1}{2}s(s+1)+1) + (s+1) + 1] = (s+1)(s+4)/2$ equations in $[(s+1)(s+2)/2 + (s+1)] = (s+1) \times (s+4)/2$ unknowns.

Solving for these probabilities and substituting in (38), (39), and (40), we obtain the generating functions $G_k(z)$ for $k = 0, 1, \ldots, s$.

The mean number of units in system may now be found in a similar way as in model 1. Also, result (30) applies here as well.

References

(1) Y. Levy and U. Yechiali, "Utilization of idle time in an $M/G/1$ queueing system," Management Sci., vol. 22, 1975, 202.

(2) I.L. Mitrani and B. Avi-Itzhak, "A many-server queue with service interruptions," Operations Res., vol. 16, 1968, 628.

(3) J.J. Moder and C.R. Phillips, Jr., "Queueing with fixed and variable channels," Operations Res., vol. 10, 1962, 218.

(4) J. Romani, "un Modelo de la Teoria de la Colas con Numero Variable de Canales," Trabajos Estadistica, vol. 8, 1957, 175.

(5) M. Yadin and P. Naor, "On queueing systems with variable service capacities," Naval Res. Logistics Quart., vol. 14, 1967, 43.

(6) U. Yechiali, "A queueing-type birth-and-death process defined on a continuous-time Markov chain," Operations Res., vol. 21, 1973, 604.