



## Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations

S. W. Fuhrmann; Robert B. Cooper

*Operations Research*, Volume 33, Issue 5 (Sep. - Oct., 1985), 1117-1129.

Stable URL:

<http://links.jstor.org/sici?sici=0030-364X%28198509%2F10%2933%3A5%3C1117%3ASDITMQ%3E2.0.CO%3B2-F>

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*Operations Research* is published by INFORMS. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/informs.html>.

---

*Operations Research*  
©1985 INFORMS

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2003 JSTOR

# Stochastic Decompositions in the $M/G/1$ Queue with Generalized Vacations

S. W. FUHRMANN

*AT&T Bell Laboratories, Holmdel, New Jersey*

ROBERT B. COOPER

*Florida Atlantic University, Boca Raton, Florida*

(Received April 1984; accepted September 1984)

This paper considers a class of  $M/G/1$  queueing models with a server who is unavailable for occasional intervals of time. As has been noted by other researchers, for several specific models of this type, the stationary number of customers present in the system at a random point in time is distributed as the sum of two or more independent random variables, one of which is the stationary number of customers present in the standard  $M/G/1$  queue (i.e., the server is always available) at a random point in time. In this paper we demonstrate that this type of decomposition holds, in fact, for a very general class of  $M/G/1$  queueing models. The arguments employed are both direct and intuitive. In the course of this work, moreover, we obtain two new results that can lead to remarkable simplifications when solving complex  $M/G/1$  queueing models.

---

**T**HIS PAPER considers a class of  $M/G/1$  queueing models with a server who is unavailable over occasional intervals of time. The times when the server is unavailable may correspond to times when the server is working at other queues (as in priority queueing models), scanning for new work (a typical aspect of many telecommunications systems), or doing maintenance.

As has been noted in the past for several specific models of this type, the following property holds.

***$M/G/1$  DECOMPOSITION PROPERTY.** The (stationary) number of customers present in the system at a random point in time is distributed as the sum of two or more independent random variables, one of which is the (stationary) number of customers present in the corresponding standard  $M/G/1$  queue (i.e., the server is always available) at a random point in time.*

This type of decomposition was first observed by Gaver (1962), and subsequently by Miller (1964), Cooper (1970), Levy and Yechiali (1975), Shanthikumar (1980), Scholl and Kleinrock (1983), Ali and Neuts (1984),

*Subject classification:* 681 queues, 691 feedback.

Federgruen and Green (1984), Neuts and Ramalhoto (1984) and Ott (1984). The more recent results (in particular, those of Ali and Neuts, Neuts and Ramalhoto, and Ott) demonstrated that this type of decomposition holds in more diverse settings than was previously realized.

In this paper we demonstrate that the  $M/G/1$  Decomposition Property holds, in fact, for a very general class of  $M/G/1$  queueing models. The arguments employed are both direct and intuitive. In the course of this work, moreover, we obtain two new results (Propositions 2 and 3) that can lead to remarkable simplifications when solving complex  $M/G/1$  queueing models.

We note in passing that an analogous type of decomposition also occurs in certain related  $GI/G/1$  models. See Gelenbe and Iasnogorodski (1980), Servi (1984), Doshi (1985), and Keilson and Servi (1985).

In the next section we define the class of models to be considered in this paper. Section 2 states and derives the results.

## 1. THE $M/G/1$ QUEUE WITH GENERALIZED VACATIONS

This section details the definitions, notations, and assumptions used in this paper.

We consider a class of  $M/G/1$  queueing models with a server who is unavailable for occasional intervals of time. We call a time interval when the server is either unavailable or idle a *vacation*. Thus, at all times the server is either *busy* or is *on vacation*; also, a vacation necessarily begins every time the system becomes empty. Note that the standard  $M/G/1$  queue (i.e., the server is always available) fits this structure, with all vacations corresponding to idle periods.

**ASSUMPTION 1.** *Customers arrive to the system according to a Poisson process of rate  $\lambda > 0$  and have service times given by a general distribution function  $H(\cdot)$ . The service times of different customers are independent of each other and are independent of the arrival process. In addition, each service time is independent of the sequence of vacation periods that precede that service time. We let  $\hat{H}(\cdot)$  denote the L.S.T. (Laplace-Stieltjes transform) of  $H(\cdot)$ ,  $h$  denote the mean of  $H(\cdot)$ , and  $\rho = \lambda h$  denote the server utilization.*

**ASSUMPTION 2.** *All customers arriving to the system are eventually served. Thus, the system has an infinite queueing capacity and  $\rho < 1$ . Moreover, customers do not balk, defect, or renege from the system.*

**ASSUMPTION 3.** *Customers are served in an order that is independent of their service times.*

**ASSUMPTION 4.** *Service is nonpreemptive. That is, once selected for service, a customer is served to completion in a continuous manner.*

*Remark 1.* Many of the standard queueing disciplines satisfy Assumptions 3 and 4, including FIFO (First In First Out), LIFO (Last In First Out), and SIRO (Service In Random Order). Moreover, any such queueing discipline will lead to the same distribution for the number of customers in the system. Thus, when deriving results for the number of customers in the system, we can, without loss of generality, select any queueing discipline that satisfies Assumptions 3 and 4.

**ASSUMPTION 5.** *The rules that govern when the server begins and ends vacations do not anticipate future jumps of the Poisson arrival process. Here the notion of "anticipate" is as defined in Wolff (1982).*

We call a queueing system that satisfies Assumptions 1 through 5 an  $M/G/1$  Queue with Generalized Vacations. For brevity, we will refer to such a queueing system in this paper as a *vacation system* or a *vacation model*. We say that a vacation model has the property of *exhaustive service* in case each time that the server becomes available, he works in a continuous manner until the system becomes empty.

Before proceeding further, we give some illustrative examples of vacation models. Note that a comprehensive list of models that classify as vacation models (i.e., satisfy Assumptions 1 through 5) and have appeared in the open literature would be very large indeed.

*Example 1.* The  $M/G/1$  Queue with Vacations. In this  $M/G/1$  queueing model, each time that the system becomes empty, the server begins a vacation of random length. If the server returns from vacation to find one or more customers waiting, he works until the system is empty (exhaustive service), then begins another vacation. If the server returns from a vacation to find no customers waiting, he begins another vacation immediately. The lengths of vacations are independent of the arrival process. This model (and variations thereof) has been studied by Miller, Cooper (1970), Levy and Yechiali, Heyman (1977), Ott (1979), Shanthikumar (1980), Scholl and Kleinrock, Levy and Kleinrock (1985), and others.

*Example 2.* The  $N$ -Policy. In this  $M/G/1$  queueing model, each time that the system becomes empty, the server waits until exactly  $N$  customers are waiting ( $N$  is some fixed positive integer), then works continuously until the system is again empty (exhaustive service). Note that in this model, the length of a vacation depends on the arrival process. This model was studied by Yadin and Naor (1963) and by Heyman (1968), who showed that the model possesses certain optimal properties.

The following three examples are of models *without* exhaustive service.

*Example 3.* The  $M/G/1$  Queue with Gated Vacations. This is the same model as Example 1, with the following proviso. Now when the server

returns from a vacation, the server accepts only those customers who were waiting when the server returned, deferring the service of subsequent arrivals until after the next vacation. One can imagine that when the server returns from vacation, a gate closes behind the last waiting customer, and the server will serve only those customers in front of the gate before departing on another vacation. Variations of this model have been analyzed by Takács (1977), Cooper (1981, p. 267), and by Ali and Neuts.

*Example 4. Limited Service Queueing Models.* This  $M/G/1$  queueing model places an upper bound, say  $k$ , on the number of customers that the server will serve per visit to the queue. In one variation of this model, if the server returns from vacation to find  $j$  customers waiting, the server will give service to  $\min(j, k)$  customers before again going on vacation. In another variation, the server works until either  $k$  customers have been served consecutively or the system becomes empty, then goes on another vacation.

*Example 5. An  $M/G/1$  Queueing Model in which the Server Must Search for Customers.* Neuts and Ramalhoto studied an  $M/G/1$  queueing model in which the server must search out and find customers before serving them. More explicitly, each time that a service is completed, the server begins a *seek phase* (in our terminology, a vacation). The length of a seek phase depends on the number of customers in the system in the following way. A seek in progress at time  $t$  ends during the time interval  $(t, t + dt)$  with probability  $j\sigma dt$ , provided there are  $j$  customers in the system at time  $t$ .

*Example 6. Queues Served in Cyclic Order.* In these models (also called *Queues with Alternating Priority and Polling Models*), the server visits a set of queues in a fixed cyclic order. The vacations associated with any particular queue correspond to times when the server is visiting (*polling*) the other queues. Models with exhaustive service and gated service have been studied by Cooper and Murray (1969), Cooper (1970), Eisenberg (1972), and Hashida (1972), to name a few. Models with limited service (i.e., the server will serve at most  $k$  customers per visit at each queue) are more difficult to analyze and few results exist; see Nomura and Tsukamoto (1978), Eisenberg (1979), Kuehn (1979), Takagi (1983), Fuhrmann (1984b), and Watson (1984).

*Example 7. Static Priority Systems.* Suppose there are  $N$  priority classes of customers and Class  $i$  customers ( $1 \leq i \leq N$ ) have priority (either preemptive or nonpreemptive) over classes  $i + 1, \dots, N$ . See Jaiswal (1968). Again, the vacations seen by any particular customer class correspond to times when the server is working on the other customer classes or is idle. While the results in this paper are oriented

primarily toward nonpreemptive systems, certain conclusions can also be drawn about preemptive systems. For example, suppose that Class  $i$  customers have preemptive resume priority over Class  $j$  customers. First note that Class  $j$  customers have no effect whatever on the service received by Class  $i$  customers. Second, it is a simple matter to appropriately "inflate" the service times of Class  $j$  customers to account for the interruptions triggered by the arrivals of higher priority customers. Thus, for purposes of analysis, the service to Class  $j$  customers can be viewed as being nonpreemptive, but with appropriately longer service times. When these systems are viewed in this manner, the results in this paper apply.

Let  $A_n$  denote the number of customers that arrive during the  $n$ th vacation; let  $Z_n$  denote the number of customers already present when the  $n$ th vacation began. Thus, when the server returns from the  $n$ th vacation there will be  $A_n + Z_n$  customers in the system. We assume that  $\{(A_n, Z_n), n = 1, 2, \dots\}$  is a stationary, ergodic sequence, and let  $A$  and  $Z$  denote the generic random variables for  $\{A_n\}$  and  $\{Z_n\}$ , respectively. That is,

$$P(A_n = k) = P(A = k)$$

$$P(Z_n = k) = P(Z = k)$$

for all  $n, k$ . We let  $\alpha(\cdot)$  denote the p.g.f. (probability generating function) of  $A$  and let  $\zeta(\cdot)$  denote the p.g.f. of  $Z$ . Note that a vacation system has exhaustive service in case  $\zeta(z) \equiv 1$ . We also let

$\psi(\cdot)$  = the p.g.f. for the stationary distribution of the number of customers that a random departing customer leaves behind in the vacation system.

$\pi(\cdot)$  = the p.g.f. for the stationary distribution of the number of customers that a random departing customer leaves behind in the corresponding standard M/G/1 queue (i.e., the server is always available).

As is well-known (see, e.g., Cooper (1981, p. 210)), the function  $\pi(\cdot)$  is given by

$$\pi(z) = \frac{(1 - \rho)(1 - z)\tilde{H}(\lambda - \lambda z)}{\tilde{H}(\lambda - \lambda z) - z}.$$

A principal goal of this paper is to find general expressions for the function  $\psi(\cdot)$ .

*Remark 2.* The functions  $\psi(\cdot)$  and  $\pi(\cdot)$  are also the p.g.f.s for the number of customers present at a random point in time in the vacation system and in the standard M/G/1 queueing system, respectively. This

result follows from a theorem by Burke (see Cooper 1981, p. 187) and the fact that Poisson arrivals see time averages (see Wolff).

Suppose that  $I_0$  denotes some group of customers in a vacation system. For the purposes of this paper, the set  $I_0$  will correspond to either a single customer or to the set of customers who arrive during some vacation. Now let  $I_1$  denote the set of customers who arrive to the system while members of  $I_0$  are being served. We call  $I_1$  the *first generation offspring* of  $I_0$ . In general, let  $I_k$  for  $k \geq 2$  denote the set of customers who arrive to the system while members of  $I_{k-1}$  are being served. We call  $I_k$  the *kth generation offspring* of  $I_0$ . Finally, we call the set  $\cup_{k=0}^{\infty} I_k$  the *ancestral line* of  $I_0$ .

We define *vacation customers* to be customers who arrive while the server is on vacation. Note that for any customer, say  $\mathcal{E}$ , there corresponds a unique vacation customer, say  $\mathcal{A}$ , such that  $\mathcal{E}$  is in the ancestral line of  $\mathcal{A}$  (with  $\mathcal{E} = \mathcal{A}$  if and only if  $\mathcal{E}$  is a vacation customer).  $\mathcal{A}$  is called the *ancestor* of  $\mathcal{E}$ .

### 2. DECOMPOSITION RESULTS

As has been known for some time (see Cooper 1970), the  $M/G/1$  Decomposition Property holds for vacation systems with exhaustive service. More explicitly,

$$\psi(z) = \frac{1 - \alpha(z)}{\alpha'(1)(1 - z)} \pi(z). \tag{1}$$

(This equation can be derived in a simple, direct, and intuitive manner. See Fuhrmann 1984a.) In the current paper, we demonstrate that the  $M/G/1$  Decomposition Property holds for *any* vacation system. We begin with the following result.

**PROPOSITION 1.** *Consider a random (tagged) customer. Let  $\mathcal{A}$  denote the ancestor of the tagged customer, and let  $I_0$  denote the set of vacation customers who arrived during the same vacation to which  $\mathcal{A}$  arrived (in particular,  $\mathcal{A} \in I_0$ ). Let  $I$  denote the ancestral line of  $I_0$ , and let  $X$  denote the number of members of  $I$  who are present in the system immediately after the tagged customer departs. Then  $X$  has p.g.f.*

$$\frac{1 - \alpha(z)}{\alpha'(1)(1 - z)} \pi(z). \tag{2}$$

*Proof.* Let

- $t_j$  = time at which the  $j$ th service to a member of  $I$  begins;
- $\tau_j$  = time at which the  $j$ th service to a member of  $I$  completes;
- $X_0$  = the number of members in the set  $I_0$ ; and
- $X_j$  = the number of members of  $I$  that are present in the system immediately after  $\tau_j$  for  $j \geq 1$ .

Note that given  $X_0$ , the vacation discipline plays no role whatever in determining the distributions of the random variables  $\{X_j\}$ . In fact, over each interval  $[t_j, \tau_j]$  the number of members of  $I$  present in the system increases by a random variable with p.g.f.  $\tilde{H}(\lambda - \lambda z)$  (corresponding to the number of customers who arrive during a random service time) and decreases by one (corresponding to the customer who is served). Outside of the intervals  $[t_j, \tau_j]$ , the population  $I$  does not change. Consequently, all vacation policies that lead to the same distribution for the random variable  $A$  will have the same distribution for the random variable  $X$ . In particular, all such vacation policies will have the same distribution for  $X$  as does an exhaustive service discipline. But for an exhaustive service discipline,  $X$  corresponds to the *total number* of customers that the tagged customer leaves behind, and so has p.g.f. given by Equation 1, which is the same as (2).

By itself, Proposition 1 is of very limited utility, since it gives the distribution of only a special subpopulation of the customers in the system. Nonetheless, Proposition 1 illustrates an important principle and will play a key role in our following development.

**ASSUMPTION 6.** *The random variables  $A_n$  and  $Z_n$  for  $n = 1, 2, \dots$  are independent. In words, the number of customers that arrive during a vacation is independent of the number of customers present in the system when the vacation began.*

**PROPOSITION 2.** *Consider a vacation system that also satisfies Assumption 6. For this system,*

$$\psi(z) = \zeta(z) \frac{1 - \alpha(z)}{\alpha'(1)(1 - z)} \pi(z). \tag{3}$$

*Proof.* Without loss of generality, we assume a nonpreemptive LIFO queueing discipline (see Remark 1). Consider a random (*tagged*) customer, and let  $\mathcal{A}$  denote the tagged customer's ancestor. We call the vacation to which  $\mathcal{A}$  arrived the *tagged vacation*, and let  $I_0$  denote the population that arrived during the tagged vacation.

Now let  $P_1$  denote the population that was present when the tagged vacation began, and let  $N_1$  denote the number of customers in the population  $P_1$ . Assumption 6 implies that the random variable  $N_1$  has p.g.f.  $\zeta(\cdot)$ .

Let  $P_2$  denote the population of customers that are both:

- Members of the ancestral line of  $I_0$ , and
- Present in the system when the tagged customer departs.

Let  $N_2$  denote the number of customers in the population  $P_2$ . Proposition 1 immediately implies that the random variable  $N_2$  has p.g.f. given by (2).



Now let  $P$  denote the population of customers that are present when the tagged customer departs. We argue that, due to LIFO queueing discipline,  $P$  is precisely the union of the two sets  $P_1$  and  $P_2$ . That  $P_1 \subset P$  follows from the LIFO queueing discipline. That  $P_2 \subset P$  follows from the definition of  $P_2$ . Thus,  $(P_1 \cup P_2) \subset P$ . Now let  $\mathcal{A}^*$  denote any vacation customer who arrives after the tagged vacation but before the tagged customer is served. Consider the time interval that begins when  $\mathcal{A}^*$  arrives and ends when the last member of the ancestral line of  $\mathcal{A}^*$  departs from the system. Because of the LIFO queueing discipline, no member of the ancestral line of  $I_0$  will be served during this time interval. Therefore, when service to the tagged customer (or for that matter, to any member of the ancestral line of  $I_0$ ) begins, all customers present in the system are members of either  $P_1$  or  $P_2$ . Moreover, any customers who arrive while the tagged customer is served will be members of  $P_2$ . Thus,  $P \subset (P_1 \cup P_2)$  and, therefore,  $P = P_1 \cup P_2$ .

It is clear that the populations  $P_1$  and  $P_2$  are disjoint. Moreover, Assumption 6 and the LIFO queueing discipline imply that the related random variables  $N_1$  and  $N_2$  are independent. Thus,  $\psi(\cdot)$  is simply the product of the p.g.f.s for  $N_1$  and  $N_2$ , and hence is given by Equation 3.

Note that for models with exhaustive service,  $\zeta(z) \equiv 1$ , and Equation 3 reduces to Equation 2. Thus, Proposition 2 can be viewed as a generalization of the result for exhaustive service.

Besides illustrating that the  $M/G/1$  Decomposition Property holds under more general settings, Proposition 2 can be highly useful in solving specific models. Consider the  $M/G/1$  Queue with Gated Vacations (Example 3). We define

$$R^{(1)}(z) = \tilde{H}(\lambda - \lambda z)$$

$$R^{(k)}(z) = R^{(1)}(R^{(k-1)}(z)), \quad k \geq 2.$$

(Readers familiar with concepts of branching processes will recognize that  $R^{(k)}(\cdot)$  is the p.g.f. for the number of  $k$ th generation offspring of a random customer.) It is not difficult to show that  $\zeta(z) = \prod_{k=1}^{\infty} \alpha(R^{(k)}(z))$ , which is Equation 4 of Ali and Neuts. Proposition 2 then yields immediately that

$$\psi(z) = \left\{ \prod_{k=1}^{\infty} \alpha(R^{(k)}(z)) \right\} \left\{ \frac{1 - \alpha(z)}{\alpha'(1)(1 - z)} \right\} \pi(z)$$

which is essentially Equation 17 of Ali and Neuts.

The definition of gated service implies that customers are served in an order quite different from LIFO (the queueing discipline used in the proof of Proposition 2). Since customers are served nonpreemptively and in an order independent of their service times, however, Proposition 2 nonetheless applies to this model. See Remark 1.

Proposition 2 employed a rather strong independence assumption (Assumption 6) to obtain the M/G/1 Decomposition Property in the particular form of Equation 3. As the next result demonstrates, however, Assumptions 1 through 5 alone lead to the M/G/1 Decomposition Property.

**PROPOSITION 5.** *Consider any vacation system. We define  $\chi(\cdot)$  to be the p.g.f. for the number of customers in the system at a random point in time when (given that) the server is on vacation. Then the functions  $\psi(\cdot)$ ,  $\chi(\cdot)$ , and  $\pi(\cdot)$  are related by*

$$\psi(z) = \chi(z)\pi(z). \tag{4}$$

*Proof.* Without loss of generality, we assume a nonpreemptive LIFO queueing discipline (see Remark 1). Consider a random (tagged) customer, and let  $\mathcal{A}$  denote the tagged customer's ancestor. Let  $P_1$  denote the population of customers already present in the system when  $\mathcal{A}$  arrived, and let  $N_1$  denote the number of customers in the population  $P_1$ . Since Poisson arrivals see time averages (see Wolff; note the relevance of Assumption 5), the variable  $N_1$  has p.g.f.  $\chi(\cdot)$ .

Now let  $P_2$  denote the population of customers in the ancestral line of  $\mathcal{A}$  that are present when the tagged customer departs from the system, and let  $N_2$  denote the number of customers in the population  $P_2$ . Reasoning along the lines of Proposition 1 leads to the conclusion that  $N_2$  has p.g.f.  $\pi(\cdot)$ .

Now observe that (the arguments are the same as in the proof of Proposition 2) because of the LIFO queueing discipline, the population of customers present when the tagged customer departs is precisely the union of the two disjoint sets  $P_1$  and  $P_2$ , and moreover, the random variables  $N_1$  and  $N_2$  are independent. Thus,  $\psi(\cdot)$  is the product of the p.g.f.s for  $N_1$  and  $N_2$  and hence is given by Equation 4.

The relationship between Propositions 2 and 3 is that, given Assumptions 1-6,

$$\chi(z) = \zeta(z) \frac{1 - \alpha(z)}{\alpha'(1)(1 - z)}. \tag{5}$$

This equation, in fact, can be argued directly.

As a concrete illustration of Proposition 3, consider Example 5, i.e., the M/G/1 model in which the server must seek out and find customers before serving them. This model was studied by Neuts and Ramalhoto, who determined by direct analytical methods that (for this specific model)

$$\chi(z) = \exp\left(-\frac{\lambda}{\sigma} \int_z^1 \frac{1 - u}{\tilde{H}(\lambda - \lambda u) - u} du\right)$$

and  $\psi(z) = \chi(z)\pi(z)$ .

Proposition 3 points out that the latter equation holds under much more general conditions.

*Remark 3.* The results presented thus far also apply to more general  $M/G/1$  models. Consider for example, models with *feedback of customers*. That is, at each service completion epoch, the departing customer is instantaneously replaced by a group of  $k$  customers with probability  $g_k$  for  $k = 0, 1, 2, \dots$  and let  $G(z) = \sum_{k=0}^{\infty} g_k z^k$ . Now let  $R(\cdot)$  denote the p.g.f. for the total number of customers that replace a served customer (i.e., both the customers who arrive during the service time and the customers who arrive coincidentally with the served customer's departure epoch). Then  $R(z) = \tilde{H}(\lambda - \lambda z)G(z)$ . Let the functions  $\alpha(\cdot)$ ,  $\zeta(\cdot)$ ,  $\psi(\cdot)$ ,  $\pi(\cdot)$ , and  $\chi(\cdot)$  be defined as before, relative to this feedback system. In particular,

$$\pi(z) = \frac{(1 - \rho)(1 - z)R(z)}{R(z) - z}$$

where now  $\rho = \lambda h + G'(1)$ . (Of course, if  $G(z) \equiv 1$ , then the above equation reduces to the expression given for  $\pi(z)$  in Section 1.) Then Propositions 1, 2, and 3 still hold. In fact, the arguments leading to these results are essentially unchanged. These results also extend to various other  $M/G/1$  models, e.g., to models in which customers arrive at a different rate during vacation periods, as in Shogan (1979) and Shanthikumar (1982).

We conclude this section with some comments on waiting times. (Here, *waiting time* = *queueing time* + *service time*). We begin with the following definition.

**ASSUMPTION 7.** *The waiting time of a customer is independent of the part of the arrival process that occurs after the customer's arrival epoch.*

The reader should note that Assumption 7 does not hold for many of the vacation models we have considered. First, queueing disciplines other than FIFO will generally not satisfy Assumption 7 (if the queueing discipline is not FIFO, then on at least some occasions, one customer waits while another customer who arrived at a later time is served). Even with a FIFO queueing discipline, Assumption 7 does not hold for Examples 2 and 5, but *does* hold for Examples 1, 3, and 4.

Let  $W(\cdot)$  denote the distribution function of the waiting time of a random customer in a vacation system with a FIFO queueing discipline, and let  $\tilde{W}(\cdot)$  denote the L.S.T. of  $W(\cdot)$ . Let  $W_1(\cdot)$  and  $\tilde{W}_1(\cdot)$  denote the analogous functions for the corresponding standard  $M/G/1$  queue.

**PROPOSITION 4.** *Consider a vacation system with a FIFO queueing discipline that also satisfies Assumption 7. Define  $F(s) = \chi(1 - s/\lambda)$ . Then,*

$$\tilde{W}(s) = F(s)\tilde{W}_1(s). \tag{6}$$

*Proof.* Under a FIFO queueing discipline, the customers that a departing customer leaves behind are precisely those customers who arrived during the departing customer's waiting time. Assumption 7 then implies that

$$\psi(z) = \int_0^\infty \exp(-\lambda t(1-z)) dW(t)$$

so that  $\psi(\cdot)$  and  $\tilde{W}(\cdot)$  are related by  $\psi(z) = \tilde{W}(\lambda - \lambda z)$  and  $\tilde{W}(s) = \psi(1 - s/\lambda)$ . Of course,  $\pi(\cdot)$  and  $\tilde{W}_1(\cdot)$  are related in a totally analogous manner. Equation 6 now follows from Proposition 3.

For some models, the function  $F(\cdot)$  can be given an interesting interpretation. In particular, we consider the *M/G/1 Queue with Vacations* (Example 1) and the *M/G/1 Queue with Gated Vacations* (Example 3). For these models we define

- $V(\cdot)$  = the distribution function for the length of a random vacation.
- $S(\cdot)$  = the distribution function for the length of a random sojourn by the server. (A sojourn by the server can be viewed as a time interval that separates successive vacations.)

We let  $\tilde{V}(\cdot)$  and  $\tilde{S}(\cdot)$  denote the L.S.T of  $V(\cdot)$  and  $S(\cdot)$ , respectively. For both models,  $\chi(z)$  is given by Equation 5, and  $\alpha(z) = \tilde{V}(\lambda - \lambda z)$ . For the *M/G/1 Queue with Vacations*,  $\zeta(z) \equiv 1$ , and therefore  $F(s) = [1 - \tilde{V}(s)]/vs$ , where  $v$  denotes the mean of  $V(\cdot)$ . That is, for the *M/G/1 Queue with Vacations*,  $F(\cdot)$  is the L.S.T. for the forward recurrence time of a vacation. For the *M/G/1 Queue with Gated Vacations*, note that due to the way the gating mechanisms works,  $\zeta(z) = S(\lambda - \lambda z)$ . Consequently, for the *M/G/1 Queue with Gated Vacations*,  $F(\cdot)$  is the product of two functions:

- The L.S.T. of the forward recurrence time of a vacation.
- The L.S.T. of a random sojourn by the server.

### 3. SUMMARY

This paper demonstrates that the *M/G/1 Decomposition Property* holds for all *M/G/1* models of a quite general form (Proposition 3). The analogous decomposition of waiting times was demonstrated in a more restrictive setting (Proposition 4).

Propositions 2 and 3 in this paper can lead to remarkable simplifications when solving complex *M/G/1* queueing models. Specific illustrations of this are given in the paper (Examples 3 and 5). The authors are currently preparing two additional papers (Fuhrmann and Cooper 1985) and Fuhrmann (1984b) that apply Propositions 2 and 3 to certain models of queues served in cyclic order (Example 6).

## ACKNOWLEDGMENT

The authors thank Y. Levy and J. G. Shanthikumar for their helpful comments.

## REFERENCES

- ALI, O. M. E., AND M. F. NEUTS. 1984. A Service System with Two Stages of Waiting and Feedback of Customers. *J. Appl. Prob.* **21**, 404–413.
- COOPER, R. B. 1970. Queues Served in Cyclic Order: Waiting Times. *Bell Syst. Tech. J.* **49**, 399–413.
- COOPER, R. B. 1981. *Introduction to Queueing Theory*, Ed. 2. North-Holland (Elsevier), New York.
- COOPER, R. B., AND G. MURRAY. 1969. Queues Served in Cyclic Order. *Bell Syst. Tech. J.* **48**, 675–689.
- DOSHI, B. T. 1985. Stochastic Decomposition in a  $GI/G/1$  Queue with Vacations (to appear in *J. Appl. Prob.*).
- EISENBERG, M. 1972. Queues with Periodic Service and Changeover Time. *Opns. Res.* **20**, 440–451.
- EISENBERG, M. 1979. Two Queues with Alternating Service. *Siam J. Appl. Math.* **20**, 287–303.
- FEDERGRUEN, A., AND L. GREEN. 1984. Queueing Systems with Service Interruptions (submitted for publication).
- FUHRMANN, S. W. 1984a. A Note on the  $M/G/1$  Queue with Server Vacations. *Opns. Res.* **32**, 1368–1373.
- FUHRMANN, S. W. 1984b. Symmetric Queues Served in Cyclic Order (to appear in *Opns. Res. Lett.*).
- FUHRMANN, S. W., AND R. B. COOPER. 1985. Application of Decomposition Principle in  $M/G/1$  Vacation Model to Two Continuum Cyclic Queueing Models (Especially Token-Ring LANs). *AT&T Tech. J.* **64**, 1091–1099.
- GAVER, D. P., JR. 1962. A Waiting Line with Interrupted Service, Including Priorities. *J. Roy. Statist. Soc.* **B24**, 73–90.
- GELLENBE, E., AND R. IASNOGORODSKI. 1980. A Queue with Server of Walking Type. *Ann. Inst. Henri Poincaré* **16**, No. 1, 63–73.
- HASHIDA, O. 1972. Analysis of Multiqueue. *Rev. Electric. Commun. Lab.* **20**, 189–199.
- HEYMAN, D. P. 1968. Optimal Operating Policies for  $M/G/1$  Queueing Systems. *Opns. Res.* **16**, 362–382.
- HEYMAN, D. P. 1977. The  $T$ -Policy for the  $M/G/1$  Queue. *Mgmt. Sci.* **23**, 775–778.
- JAISSWAL, N. K. 1968. *Priority Queues*. Academic Press, New York.
- KEILSON, J., AND L. SERVI. 1985. Oscillating Random Walk Models for  $GI/G/1$  Vacation Systems with Bernoulli Schedules (submitted to *J. Appl. Prob.*).
- KUEHN, P.J. 1979. Multiqueue Systems with Nonexhaustive Cyclic Service. *Bell Syst. Tech. J.* **58**, 671–698.
- LEVY, H., AND L. KLEINROCK. 1985. A Queue with Starter: Delay Analysis (to appear in *Opns. Res.*).
- LEVY, Y., AND U. YECHIALI. 1975. Utilization of Idle Time in an  $M/G/1$  Queueing System. *Mgmt. Sci.* **22**, 202–211.

- MILLER, L. W. 1964. Alternating Priorities in Multi-Class Queues. Ph.D. dissertation, Cornell University, Ithaca, N.Y.
- NEUTS, M. F., AND M. F. RAMALHOTO. 1984. A Service Model in which the Server is Required to Search for Customers. *J. Appl. Prob.* **21**, 157–166.
- NOMURA, M., AND K. TSUKAMOTO. 1978. Traffic Analysis on Polling Systems. *Transact. Inst. Electr. Commun. Engr. Japan* **J61-B**, 600–607.
- OTT, T. J. 1979. A Note on the  $M/G/1$  Queue with Vacation Times (unpublished work).
- OTT, T. J. 1984. On the  $M/G/1$  Queue with Additional Inputs. *J. Appl. Prob.* **21**, 129–142.
- SCHOLL, M., AND L. KLEINROCK. 1983. On the  $M/G/1$  Queue with Rest Periods and Certain Service-Independent Queueing Disciplines. *Opns. Res.* **31**, 705–719.
- SERVI, L. 1984.  $D/G/1$  Queue with Vacation (submitted for publication).
- SHANTHIKUMAR, J. G. 1980. Some Analyses of the Control of Queues Using Level Crossing of Regenerative Processes. *J. Appl. Prob.* **17**, 814–821.
- SHANTHIKUMAR, J. G. 1982. Analysis of a Single Server Queue with Time and Operation-Dependent Server Failures. *Advan. Mgmt. Stud.* **1**, 339–359.
- SHOGAN, A. W. 1979. A Single Server Queue with Arrival Rate Dependent on Server Breakdowns. *Naval Res. Logist. Quart.* **26**, 487–497.
- TAKÁCS, L. 1977. A Queueing Model with Feedback. *R.A.I.R.O.* **11**, 345–354.
- TAKAGI, H. 1983. Mean Message Waiting Time in  $N$  Symmetric Queues with Nonexhaustive Cyclic Service Discipline (unpublished work).
- WATSON, K. S. 1984. Performance Evaluation of Cyclic Service Strategies—a Survey. In *Proceedings of the Tenth International Symposium on Computer Performance*. North-Holland, Amsterdam.
- WOLFF, R. W. 1982. Poisson Arrivals See Time Averages. *Opns. Res.* **30**, 223–231.
- YADIN, M., AND P. NAOR. 1963. Queueing Systems with a Removable Service Station. *Opnl. Res. Quart.* **14**, 393–405.