On the $M/G/1$ Queue with Rest Periods and Certain Service-Independent Queueing Disciplines

MICHEL SCHOLL
IRIA, Rocquencourt, France

LEONARD KLEINROCK
University of California, Los Angeles, California

(Received September 1977, revised August 1980; accepted October 1982)

The $M/G/1$ queue with rest periods and FCFS order of service was first studied by Miller. We give further results concerning the behavior of the delay under the FCFS discipline. We then solve for the second moment of the waiting time in an $M/G/1$ queue with rest periods and *R*andom *O*rder of *S*ervice (ROS). We finally solve for the Laplace-Stieltjes transform of the distribution function of the waiting time in an $M/G/1$ queue with rest periods and nonpreemptive LCFS order of service. The relationship found between the second moments of the waiting time in an $M/G/1$ queue *with rest periods* under the FCFS, ROS and LCFS disciplines is precisely that found by Takács for the $M/G/1$ queue *without rest periods*.

IN MOST QUEUEING systems models, an idle server will remain alert awaiting a new arrival and will commence service immediately upon the customer's arrival. However, in some physical systems an idle server will initiate some other uninterruptible task (such as a vacation, a coffee break, a telephone call, or a tool change), which we shall refer to as a "rest period"; after completing this task, the server returns and begins serving any backlog that may have accumulated during its absence. If the system is still empty upon the server's return, we assume that it initiates another (independent) rest period.

Of interest in this paper, then, is the study of delays in an $M/G/1$ queue with rest periods under the following queueing disciplines, which are independent (of any measure) of service time: *F*irst-*C*ome *F*irst-*S*erved (FCFS), *R*andom *O*rder of *S*ervice (ROS) and nonpreemptive *L*ast-*C*ome-*F*irst-*S*erved (LCFS).

If the order of service is independent (of any measure) of service time, then the *d*istribution *f*unction (df) of the number of customers in the system and the average waiting time can easily be shown to be independ-

Subject classification: 681 and 698 priority with rest periods.

Operations Research
Vol. 31, No. 4, July–August 1983

0030-364X/83/3104-0705 $01.25

ent of the discipline, whether there is a rest period or not. The approach for showing this statement in the case of an $M/G/1$ queue *with rest periods* is exactly the same as that used in a *regular* $M/G/1$ queue (Kleinrock [1976] pp. 112–113). In particular, this is true for FCFS, ROS and LCFS.

Section 2 sketches a simple derivation of the z-transform of the df of the total number of customers in system in an $M/G/1$ queue with rest periods. This result was first obtained by Miller [1964], who studied the FCFS discipline using another approach. Such a model has been partially used by Cooper [1970] to analyze a system of queues served in cyclic order. Analyses of this system, also called an $M/G/1$ queue with server's vacations or an $M/G/1$ queue with T-policies, have been studied by Levy and Yechiali [1975], Heyman [1977], Heyman and Sobel [1982], and Shanthikumar [1980].

We next show that the delay in an $M/G/1$ queue with rest periods under the FCFS discipline has the same distribution as the sum of the following two independent random variables: the time in system as if there were no rest periods plus an additional delay distributed as the residual life of the rest period.

Kingman [1962a] and Takács [1963] have previously studied and reported on the $M/G/1$ queue (*without rest periods*) with ROS. In Section 4 we derive the second moment of waiting time in an $M/G/1$ queue *with rest periods* and ROS.

The $M/G/1$ queue with LCFS order of service (*without rest periods*) has been extensively studied by Takács [1963], Vaulot [1954] and Riordan [1961]. Kleinrock [1976] has derived the *L*aplace-*S*tieltjes-*T*ransform (LST) of the *p*robability *d*ensity *f*unction (pdf) of the waiting time. Following almost exactly the same argument, in Section 5 we easily solve for the LST of the pdf of the waiting time in an $M/G/1$ queue *with rest periods* and LCFS order of service.

The relationship that we find in Section 6 between the second moments of waiting time for the three disciplines (FCFS, ROS and LCFS) in the case of an $M/G/1$ queue with rest periods is precisely that which has been found in Takács in the case of a regular $M/G/1$ queue.

When solving for the time in a priority queueing system under the Alternating Priority Discipline, Miller [1964] first introduced and studied the $M/G/1$ queue with rest periods and FCFS order of service. (In this discipline, customers belong to one among N classes. The server continues servicing customers of the same class (say i) until they are depleted and the server starts servicing customers of another class, followed by another class, and eventually again begins servicing customers of class i. For class i, the server's rest period is the time elapsing between successive visits to the class.)

The study of multiple access from a set of N data sources to a single packet-switched data communication channel is one among various applications of queueing models with rest periods. Scholl [1976] and Kleinrock and Scholl [1977] model several multiple access schemes as $M/G/1$ queueing systems with rest periods. *T*ime *D*ivision *M*ultiple *A*ccess (TDMA) (Kleinrock [1976], Scholl) is the simplest example. In TDMA, each data source generates fixed sized packets to be transmitted on a FCFS basis. The system assigns each data source a periodic sequence of time slots on the channel (the packet transmission time being equal to one slot). The channel slots are usually switched to users in a round-robin (i.e. cyclic) fashion. The system can be modeled by $M/D/1$ queues with a rest period with a service time equal to N slots (1 slot being the actual transmission time and $(N-1)$ slots assigned to the other $(N-1)$ users). The rest period is equal to N slots.

In the next section, we present the mathematical model of these queueing systems with rest periods.

1. THE MODEL

Consider an $M/G/1$ queue with unlimited storage. Let λ be the (Poisson) arrival process intensity, and let $B^*(s)$ denote the Laplace-Stieltjes transform LST of the pdf of the service time with first and second moments denoted by $E(x)$ and $E(x^2)$ respectively. After completing a customer's service, the server will select another customer in queue (if any) according to a given order (queueing discipline) and will begin service immediately. This $M/G/1$ queue model is modified as follows:

If there are no customers in queue waiting for service, the server, becoming idle for lack of work, will withdraw from the system for a rest period of duration T_0 drawn from an arbitrary pdf with LST $P^*(s)$ and first and second moments $E(T_0)$ and $E(T_0^2)$, respectively. At the end of the rest period, the server will return and begin to serve the customers that have accumulated during its absence according to the same queueing discipline. If there is no backlog, the server will take another rest period which starts immediately. The rest periods are identically distributed and independent of each other and of the arrival and service processes.

Below, we consider the queueing disciplines, FCFS, ROS and LCFS. We first study the df of the total number of customers in an $M/G/1$ queue with rest periods. This is independent of the queueing discipline when the order of service is independent of service time, so we can assume FCFS without loss of generality within this class.

2. NUMBER IN SYSTEM IN THE *M*/G/1 QUEUE WITH REST PERIODS

The main derivation is that of $Q(z)$, the z-transform of the df of the number of customers in the system, $\tilde{q}$, just after instants of a customer's departure. Miller shows that the solution at these imbedded Markov points provides the solution for all points in time under steady state conditions.

2.1. The Generating Function of the Number in Queue

THEOREM (first established by Miller; see also Heyman and Sobel, Levy and Yechiali, and Shanthikumar). *$Q(z)$ is given by:*

$$Q(z) = [B^*(\lambda - \lambda z)/\lambda E(T_0)] \cdot [(1 - \rho)(1 - P^*(\lambda - \lambda z))/(B^*(\lambda - \lambda z) - z)] \tag{1}$$

where

$$\rho \triangleq \lambda E(x). \tag{2}$$

In addition, we have

$$P\,[\textit{no customers waiting for service just after departure instant}] \triangleq p_0 = Q(0) = (1 - \rho)(1 - P^*(\lambda))/\lambda E(T_0). \tag{3}$$

The method of proof for this theorem is based on the familiar "imbedded Markov chain" approach, introduced by Kendall [1951], which has become a standard approach in queueing theory (for example, Kleinrock [1975] p. 174). Let us define the following sequences of random variables related to C_n, the nth customer entering the system at time τ_n.

x_n = service time for C_n
q_n = number of customers left behind by the department of C_n.

The sequence q_n forms an imbedded Markov chain. The solution at the departure points will also provide the solution for all points in time (see Miller). Equation 1 is obtained by solving for $P(z) \triangleq E(z^{\tilde{q}}) = \lim_{n\to\infty} E(z^{q_n})$. (See Scholl for a detailed proof.)

Equation 1 can be rewritten as

$$Q(z) = Q_{M/G/1}(z) V_0(z) \tag{4}$$

where $Q_{M/G/1}(z)$ represents the z-transform of the df of the number of customers in a regular $M/G/1$ queue (Kleinrock [1975] Equation 5.86), viz.

$$Q_{M/G/1}(z) = B^*(\lambda - \lambda z)((1 - \rho)(1 - z)/(B^*(\lambda - \lambda z) - z)) \tag{5}$$

and $V_0(z)$ represents the z-transform of the df of the number of arrivals during a time interval distributed as the residual life of a rest period

$$V_0(z) = (1 - P^*(\lambda - \lambda z))/E(T_0)(\lambda - \lambda z). \tag{6}$$

As is well known (Kleinrock [1975] p. 197), the z-transform of the df of the number of (Poisson) arrivals $\tilde{r}$, arriving at a rate λ during a random time interval with LST denoted by $T^*(s)$, is $E(z^{\tilde{r}}) = T^*(\lambda - \lambda z)$. Using this result in (6), we have

$$V_0(z) = C_0^*(\lambda - \lambda z) \tag{7}$$

where $C_0^*(s)$ is the LST of the pdf of the residual life of a rest period, and

$$C_0^*(s) = (1 - P^*(s))/sE(T_0). \tag{8}$$

Thus the df of the number of customers in the system behaves as the convolution of the two following random variables:

a. Number in system in a regular $M/G/1$ queue with the same arrival and service processes
b. Number of arrivals during a time interval distributed as the residual life of the rest period.

We arrived at the result, which is not obvious, by examining our solution in (4); note that we have not proven it directly. In addition, by taking the first derivative of (4) at $z = 1$, we obtain the expected number in system:

$$\bar{q} = \lambda E(x) + \lambda^2 E(x^2)/(2(1 - \rho)) + \lambda E(T_0^2)/2E(T_0). \tag{9}$$

The sum of the first two terms of the right hand side of (9) represents the expected number in system in a regular $M/G/1$ queue (the Pollaczek-Khinchin formula; see Kleinrock [1975] p. 187).

By Little's result [1961], the expected time of a customer in the system (waiting time plus service time), denoted by $E(T)$, is

$$E(T) = E(x) + \lambda E(x^2)/(2(1 - \rho)) + E(T_0^2)/2E(T_0). \tag{10}$$

Recall that when the order of service is independent of service time, the df of the total number of customers in the system, and thus the expected time in the system (delay), are both independent of the queueing discipline.

In particular, (1) and (10) hold for FCFS as well as for ROS and LCFS.

2.2. Server's Busy Fraction and System's Busy Fraction

The proportion of time the server is busy servicing customers is easily shown to be $\rho = \lambda E(x)$ (see Miller). This result is not surprising. It is precisely the busy fraction of an $M/G/1$ queue with the same arrival intensity and the same expected service time. Thus, the busy fraction does not depend on the duration of the rest period. If the latter is long, then the (server's) busy period is long, and there are fewer rest periods per unit of time. Furthermore, as long as the order of service conserves work, i.e. as long as there is neither creation nor destruction of work (service requirement) (Kleinrock [1976]), the busy fraction (indeed, the busy period) is independent of the order of service (see Scholl).

Recalling that p_0 denotes the probability that the system is idle, i.e. there are no customers waiting for service or being served, we have in a regular $M/G/1$ queue

$$P[\text{system idle}] = P[\text{server idle}] = 1 - \rho$$

while in an $M/G/1$ queue with rest periods the system's busy fraction is larger than the server's busy fraction. In this case, p_0 is given by (3) and we have

$$p_0 = P[\text{system idle}] < P[\text{server idle}] = 1 - \rho.$$

3. LST OF THE pdf OF TIME IN SYSTEM FOR FCFS

From (1) we may easily derive the LST $S^*(s)$ of the pdf of time in the system, $\tilde{s}$, where $\tilde{s}$ is defined as the time interval from a customer's arrival instant until his service completion. We follow the simple approach used in Kleinrock ([1975] Section 5.7) for solving for the LST of the time in system from the z-transform of the df of number in the system in a (regular) $M/G/1$ queue and an FCFS order of service. Observe (from 2.1) that $S^*(\lambda - \lambda z)$ is the z-transform of the df of the number of customers arriving during the time spent in the system by an arbitrary customer. But, this number of customers is precisely the number of customers that arrive after a given "tagged" customer and that are left behind by his departure. Therefore, we may write

$$Q(z) = S^*(\lambda - \lambda z). \tag{11}$$

By making the change of variable $s = \lambda - \lambda z$, we have from (1)

$$S^*(s) = (((1 - \rho)sB^*(s))/s - \lambda + \lambda B^*(s))((1 - P^*(s))/sE(T_0)). \tag{12}$$

The first factor is the LST of the pdf of time that a customer spends in the system in a regular $M/G/1$ queue with the same arrival and service processes (Kleinrock [1975]). The second factor is the LST of the pdf of the residual life of a rest period (see (8)).

From (12) we have

$$E[\tilde{s}] = E[T] = E(\tilde{s}_{M/G/1}) + E[\text{residual life of a rest period}] \tag{13}$$

where $E(\tilde{s}_{M/G/1})$ is the expected time in the system in an $M/G/1$ queue (Pollaczek-Khinchin mean value formula (Kleinrock [1975])). Obviously,

(13) reduces to (10). Also, the expected waiting time in queue, denoted by W_{FCFS}, is

$$W_{\text{FCFS}} = \lambda E(x^2)/(2(1 - \rho)) + E(T_0^2)/2E(T_0). \tag{14}$$

The second moment $E[w^2_{\text{FCFS}}]$ of the waiting time is

$$E[w^2_{\text{FCFS}}] = E[w^2_{M/G/1}] + 2[\lambda E(x^2)/2(1 - \rho)][E(T_0^2)/2E(T_0)] + [E(T_0^3)/3E(T_0)].$$

Kleinrock ([1975], Equation 5.114) gives an expression of the second moment $E[w^2_{M/G/1}]$ of the waiting time in an $M/G/1$ queue. A simple calculation finally gives

$$E[w^2_{\text{FCFS}}] = [\lambda E(x^2)/(1 - \rho)]W_{\text{FCFS}} + [\lambda E(x^3)/3(1 - \rho)] + [E(T_0^3)/3E(T_0)]. \tag{15}$$

From the form of $S^*(s)$ given in (12), we see that $\tilde{s}$ has the same distribution as a random variable which is the sum of the following two independent random variables (we have not been able to prove this result directly):

- The time in system as if there were no rest periods, plus
- An additional delay distributed as the residual life of the rest period.

When the pdf of the rest period is such that any rest period duration is less than or equal to any service time, this additional delay t_1 is the time elapsing between the arrival instant of the first customer (the customer who initiates the busy period) and the end of the rest period. (As an example, we might consider an $M/D/1$ "slotted" queueing system defined as follows. The time axis is slotted; any customer's service must start at the beginning of a slot and lasts exactly one slot. Such a system can be modeled by an $M/D/1$ queue with a deterministic rest period, whose length is equal to one slot (see Scholl).) Clearly, t_1 is distributed as the residual life of the rest period, and all customers of a busy period incur a time in the system larger than what they would incur if there were no rest period, by an amount equal to t_1.

However, this is no longer true if a rest period can last longer than a service time. Below, we give two counterexamples in which the additional delay is not the same for all users of a busy period and is not the residual life t_1 of the rest period upon arrival of the first customer, although this additional delay *is distributed* as the residual life of the rest period.

Our first counterexample is illustrated in Figure 1, where we plot the unfinished work, $U(t)$, versus time. Six customers arrive during the rest period. However, if there were no rest period, we would observe three

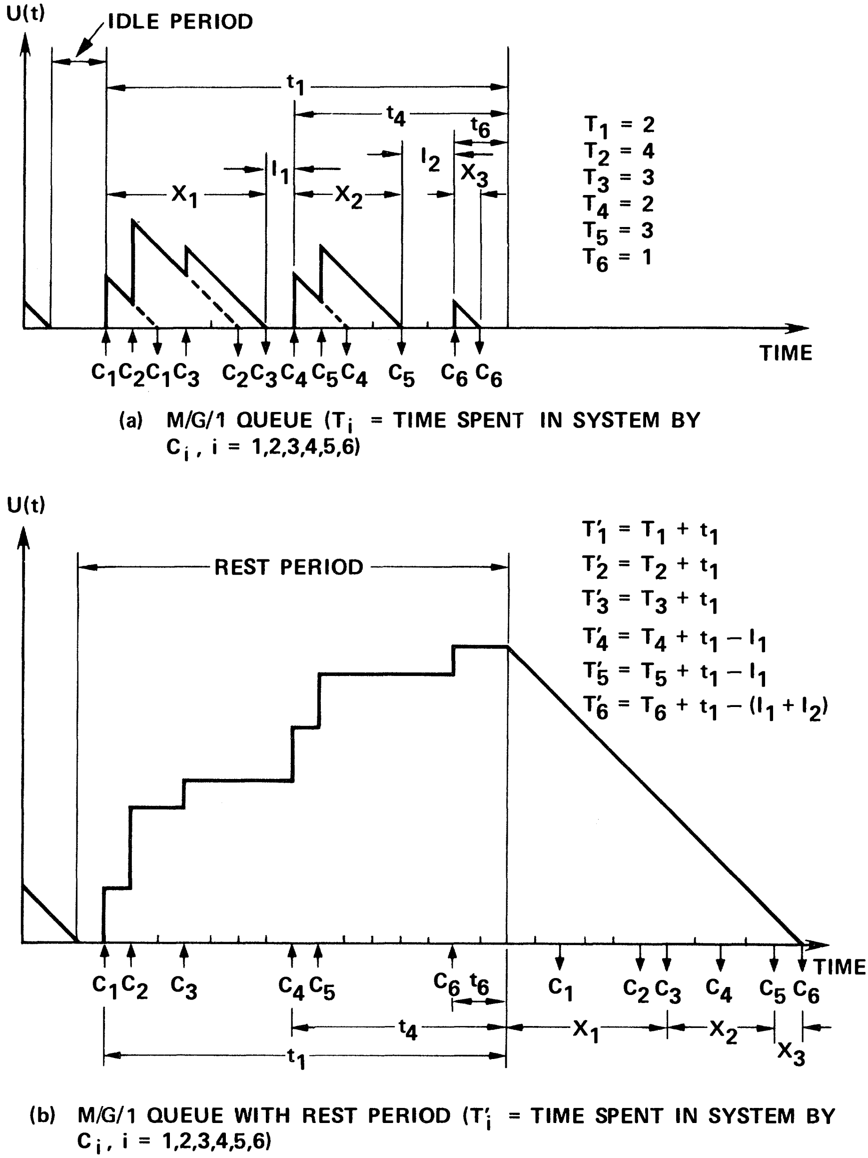

Figure 1. $M/G/1$ queue: first counterexample.

busy periods, respectively initialized by customers C_1, C_4 and C_6. The waiting time in the system (with rest period) of those customers in the first busy period (C_1, C_2, C_3) increases by the unexpired rest period when C_1 arrives (t_1). Clearly this is no longer true for C_4, C_5 and C_6. C_4 and C_5 incur an additional delay equal to the unexpired rest period upon C_4's arrival (t_4) *plus* the amount of backlog (in seconds of service) accumulated before C_4's arrival (x_1). C_6 incurs an additional delay equal to t_6 *plus* the amount of backlog accumulated before his arrival ($x_1 + x_2$).

More generally, following Heyman [private communication], we may observe that, in a busy period of the system with rest period, any customer C incurs an additional delay equal to the residual life of the rest period upon the first customer's arrival *less* his "lost idle period," if any, when the lost idle period of customer C is defined as the total time the system would have been idle between C_1's arrival and C's arrival if there were no rest period: C_4 and C_5 have their waiting time increased by $t_1 - I_1$, while C_6 incurs an additional delay equal to $t_1 - (I_1 + I_2)$.

Obviously, this additional delay is nonnegative: the fact that C arrives in the busy period initialized by C_1 implies t_1 is larger than C's lost idle period. In Figure 1, we assumed that customers arrived before the end of the rest period. The above observation is, of course, also true for customers arriving after the end of the rest period.

Figure 2 illustrates an example where three customers are served in one busy period of the system with rest period. If there were no rest period, we would observe two busy periods (Figure 2(a)).

When we introduce a rest period, the first busy period is shifted on the time axis by t_1 seconds, where t_1 is the unexpired rest period when the first customer C_1 arrives. C_1 and C_2 incur an additional delay t_1, while C_3, who initiates a busy period (in the system without rest period: Figure 2(a)), incurs an additional delay γ equal to t_1 less his lost idle period I_1.

4. SECOND MOMENT OF THE WAITING TIME WITH ROS

Below we give the first and second moments W_{ROS} and $E[w_{\text{ROS}}^2]$ of the waiting time w in an $M/G/1$ queue with rest period and ROS. Recall that (1) gives the z-transform of the df of the number in system and that the expected waiting time must be the same as for FCFS (see (14)).

Our notation ROS designates that upon completion of service, or upon return from a rest period, the server chooses a customer to be served at random from among all customers present in the queue.

Consider a tagged customer K entering the system. Let its waiting time be $\tilde{w}$, let w_1 denote the waiting time of K if the server is busy, and let w_2 denote its waiting time if the server takes a rest period. If the server is busy (with probability ρ; see Section 2.2), K must wait a time u_1 before the current customer's service completion and then a time v_1:

$$w_1 = u_1 + v_1. \tag{16}$$

If the server is idle (with probability $1 - \rho$), K must wait a time u_2 until the end of the rest period, and then a time v_2:

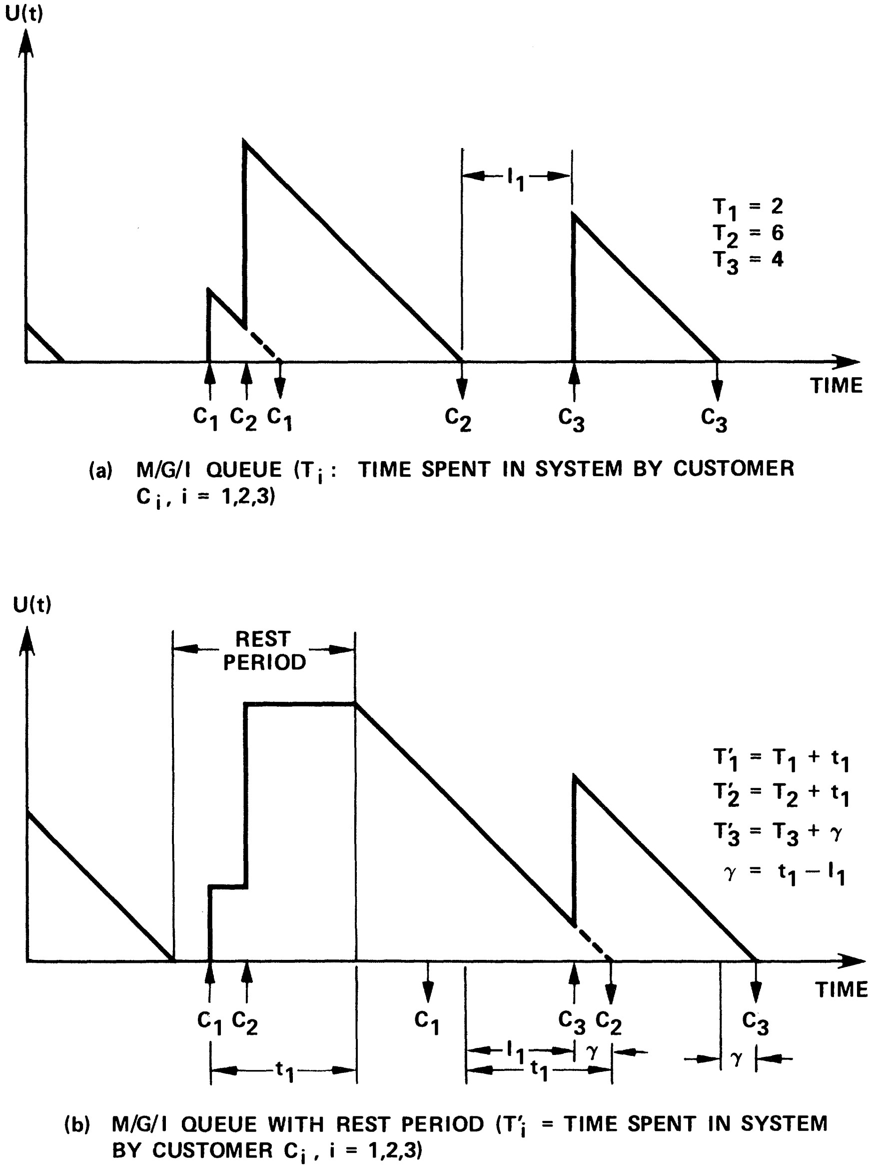

Figure 2. $M/G/1$ queue: second counterexample.

$$w_2 = u_2 + v_2 \tag{17}$$

$$W_{\mathrm{ROS}} = \rho E(w_1) + (1 - \rho)E(w_2) \tag{18}$$

$$E[w^2_{\mathrm{ROS}}] = \rho E(w_1{}^2) + (1 - \rho)E(w_2{}^2). \tag{19}$$

The derivation of the first and second moments of w_1 and w_2 follows closely the arguments of Kingman [1962a]. Scholl shows that we have:

$$E(w_1) = (1/(2-\rho))[(2+\rho)(E(x^2)/2E(x)) + \rho((\lambda E(x^2)/2(1-\rho)) + (E(T_0^2)/2E(T_0)))] \quad (20)$$

$$\begin{aligned} E(w_1^2) = {} & 2\lambda E(x^3)/(3\rho(1-\rho)(2-\rho)) \\ & + [\lambda E(x^2)]^2/(\rho(1-\rho)^2(2-\rho)) \\ & + (2p^2/((2-\rho)(3-2\rho)))(E(T_0^3)/3E(T_0)) \\ & + \lambda E(x^2)(E(T_0^2)/2E(T_0))(2/((2-\rho)^2(3-2\rho))) \\ & \cdot[2+(2-\rho)^2/(1-\rho)] \end{aligned} \quad (21)$$

$$E(w_2) = [(2+\rho)/(2-\rho)][E(T_0^2)/2E(T_0)] \quad (22)$$

$$\begin{aligned} E(w_2^2) = {} & (E(T_0^3)/E(T_0))[⅓ \\ & + \rho/(2-\rho) + 2\rho^2/(2-\rho)(3-2\rho)] \\ & + \lambda E(x^2)[E(T_0^2)/E(T_0)]((6-\rho)/(2-\rho)^2(3-2\rho)). \end{aligned} \quad (23)$$

Substituting (20) and (22) in (18), we verify, as we might expect, that the expected waiting time is the same as in an $M/G/1$ queue with rest periods and an FCFS order of service (see (14)):

$$W_{\mathrm{ROS}} = \lambda E(x^2)/2(1-\rho) + E(T_0^2)/2E(T_0).$$

Knowing the second moment of the waiting time in both cases (server idle or busy), we can finally obtain the second moment of waiting time. Substituting (21) and (23) into (19), we have, after some algebraic manipulations:

$$\begin{aligned} E[w_{\mathrm{ROS}}^2] = {} & [2/(2-\rho)][\lambda E(x^3)/3(1-\rho) \\ & + \lambda E(x^2)/(1-\rho)(\lambda E(x^2)/2(1-\rho) \\ & + E(T_0^2)/2E(T_0)) + E(T_0^3)/3E(T_0)]. \end{aligned} \quad (24)$$

Comparing (24) and (15), which gives the second moment $E[w_{\mathrm{FCFS}}^2]$ of the waiting time in an $M/G/1$ queue with rest period and FCFS order of serivce, we have the following relationship:

$$E[w_{\mathrm{ROS}}^2] = E[w_{\mathrm{FCFS}}^2]/[1-(\rho/2)]. \quad (25)$$

This is precisely the result found by Takács relating the second moments of the waiting time in a regular (no rest period) ROS $M/G/1$ queue and in a regular FCFS $M/G/1$ queue.

5. TIME IN SYSTEM IN AN *M/G/1* QUEUE WITH REST PERIODS AND LCFS ORDER OF SERVICE

(We consider only the nonpreemptive case: A new arrival who finds the server busy cannot preempt the customer in service and must wait at least until the current service completion. Although straightforward, the preemptive resume LCFS discipline, which is not independent (of any measure) of service time, will not be studied.)

Kleinrock ([1976], p. 118) derives the LST of the pdf of the waiting time in an $M/G/1$ queue without rest periods and with nonpreemptive LCFS order of service. Following almost exactly the same argument, we solve for the LST, $W^*_{\mathrm{LCFS}}(s)$, of the pdf of the waiting time in an $M/G/1$ queue with rest periods and nonpreemptive LCFS order of service. Recall that the z-transform of the df of the number in system is given by (1) and that the expected waiting time is given by (14) (both independent of the order of service).

A new arrival finds the server either busy (with probability ρ; see Section 2.2) or idle (with probability $(1-\rho)$).

$$W^*_{\mathrm{LCFS}}(s) = E[e^{-sw} \mid \text{server busy upon arrival}]\rho + E[e^{-sw} \mid \text{server idle upon arrival}](1-\rho). \tag{26}$$

If the server is busy, the arrival waits only for the customer found in service, and its delay is due to arrivals that enter the system after it does.

Using a delay cycle analysis (Kleinrock [1976] p. 11), we have

$$E[e^{-sw} \mid \text{server busy upon arrival}] = G_0{}^*(s + \lambda - \lambda G^*(s)) \tag{27}$$

where $G_0{}^*(s)$ is the LST of the residual life of the service time upon arrival and $G^*(s)$ is the LST of the pdf of the busy period in an $M/G/1$ queue. $G^*(s)$ is given by the following functional equation (Kleinrock [1975] Equation 5.137):

$$G^*(s) = B^*(s + \lambda - \lambda G^*(s)). \tag{28}$$

Thus we have

$$E[e^{-sw} \mid \text{server busy upon arrival}] = [1 - G^*(s)]/[s + \lambda - \lambda G^*(s)]E(x). \tag{29}$$

If the server is idle, the new arrival waits until the end of the rest period, at which instant it begins service only if no other arrival occurred in the interim. Consequently,

$$E[e^{-sw} \mid \text{server idle upon arrival}] = C_0{}^*(s + \lambda - \lambda G^*(s))$$

where $C_0{}^*(s)$ is the LST of the pdf of the residual life of the rest period upon arrival (see (8)). Therefore,

$$E[e^{-sW} \mid \text{server idle upon arrival}] = [1 - P^*(s + \lambda - \lambda G^*(s))]/[s + \lambda - \lambda G^*(s)]E(T_0). \tag{30}$$

Substituting (29) and (30) into (26), we have the expression for $W^*(s)$ that we were seeking:

$$W^*(s) = (\rho[1 - B^*(s + \lambda - \lambda G^*(s))]/([s + \lambda - \lambda G^*(s)]E(x))) + ((1 - \rho)[1 - P^*(s + \lambda - \lambda G^*(s))]/([s + \lambda - \lambda G^*(s)]E(T_0))). \tag{31}$$

By taking the first and second derivatives of (31) at $s = 0$, we obtain the following expressions:

$$W_{\text{LCFS}} = \lambda E(x^2)/2(1 - \rho) + E(T_0{}^2)/2E(T_0) \tag{32}$$

$$E[w^2_{\text{LCFS}}] = [\lambda E(x^2)/(1 - \rho)^2]W_{\text{LCFS}} + [\lambda E(x^3)/3(1 - \rho)^2] + [E(T_0{}^3)/3E(T_0)](1/(1 - \rho)). \tag{33}$$

As expected, $W_{\text{LCFS}} = W_{\text{FCFS}}$ (see (14)), but the second moment $E[w^2_{\text{LCFS}}]$ is larger than $E[w^2_{\text{FCFS}}]$. Comparing (33) to (15), we finally have

$$E[w^2_{\text{LCFS}}] = E[w^2_{\text{FCFS}}]/(1 - \rho). \tag{34}$$

It is surprising to find that this result holds for the $M/G/1$ queue (*without rest periods*) (see Takács) as well as for the $M/G/1$ queue (*with rest periods*).

6. DISCUSSION

We have studied the $M/G/1$ queue with rest periods under three queueing disciplines independent (of any measure) of the service time (FCFS, ROS, LCFS). In the FCFS case (Section 3), we showed that the time that a customer spends in the system has the same distribution as the sum of the two random variables:

a. The time in system in an $M/G/1$ queue with same arrival and service processes and FCFS order of service and
b. A time interval distributed as the residual life of a rest period.

As we stated earlier, the *df of number in system* as well as the *expected waiting time* obtained under the FCFS discipline are the same under the ROS and LCFS order of service (Section 2). We solved for the second moment of waiting time in the ROS case (Section 4) and for the LST of waiting time in the LCFS case (Section 5). While the expected waiting time is independent of the order of service, the second moment of the waiting time is significantly affected by the queueing discipline. From (25) and (34) we have the following relationship:

$$E[w^2_{\text{FCFS}}] = [1 - (\rho/2)]E[w^2_{\text{ROS}}] = (1 - \rho)E[w^2_{\text{LCFS}}]. \tag{35}$$

Equation (35) holds for the regular $M/G/1$ queue (see Takács) as well as for the $M/G/1$ queue with rest periods. Table I summarizes the results of the waiting time moment calculations for both the regular $M/G/1$ queue and the $M/G/1$ queue with rest periods. In both cases, FCFS discipline leads to the smallest second moment of the waiting time, while the LCFS discipline leads to the largest one.

Kingman [1962b] showed that under an arbitrary work-conserving, nonpreemptive queueing discipline (independent of the service time), the variance of the waiting time in a $G/G/1$ queue is not less than that given by the FCFS discipline. Vasicek [1977] showed the following more general result:

THEOREM (Vasicek). *The expected value of any convex function of the waiting time (such as the variance) in a general single-server queue under a general queueing discipline (independent of service time) does not exceed that under the LCFS discipline.*

TABLE I

MOMENTS OF THE WAITING TIME

Moments of the waiting time	M/G/1 queue with Rest Period	M/G/1 queue without Rest Period
$E\{W\}$	$\frac{\lambda E(x^2)}{2(1-\rho)} + \frac{E(T_0^2)}{2E(T_0)}$	$\frac{\lambda E(x^2)}{2(1-\rho)}$
$E\{W_{FCFS}^2\}$	$\frac{\lambda E(x^2)}{1-\rho}\left[\frac{\lambda E(x^2)}{2(1-\rho)} + \frac{E(T_0^2)}{2E(T_0)}\right] + \frac{\lambda E(x^3)}{3(1-\rho)} + \frac{E(T_0^3)}{3E(T_0)}$	$\frac{\lambda E(x^2)}{1-\rho}\left(\frac{\lambda E(x^2)}{2(1-\rho)}\right) + \frac{\lambda E(x^3)}{3(1-\rho)}$
$E\{W_{ROS}^2\}$	$\frac{2}{2-\rho}\left\{\frac{\lambda E(x^2)}{1-\rho}\left[\frac{\lambda E(x^2)}{2(1-\rho)} + \frac{E(T_0^2)}{2E(T_0)}\right] + \frac{\lambda E(x^3)}{3(1-\rho)} + \frac{E(T_0^3)}{3E(T_0)}\right\}$	$\frac{2}{2-\rho}\left\{\frac{\lambda E(x^2)}{1-\rho}\left(\frac{\lambda E(x^2)}{2(1-\rho)}\right) + \frac{\lambda E(x^3)}{3(1-\rho)}\right\}$
$E\{W_{LCFS}^2\}$	$\frac{1}{1-\rho}\left\{\frac{\lambda E(x^2)}{1-\rho}\left[\frac{\lambda E(x^2)}{2(1-\rho)} + \frac{E(T_0^2)}{2E(T_0)}\right] + \frac{\lambda E(x^3)}{3(1-\rho)} + \frac{E(T_0^3)}{3E(T_0)}\right\}$	$\frac{1}{1-\rho}\left\{\frac{\lambda E(x^2)}{1-\rho}\left(\frac{\lambda E(x^2)}{2(1-\rho)}\right) + \frac{\lambda E(x^3)}{3(1-\rho)}\right\}$

It is easy to extend this result to a $G/G/1$ queue with rest period (see Scholl).

ACKNOWLEDGMENTS

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract MDA 903-77-C-0272.

REFERENCES

COOPER, R. B. 1970. Queues Served in Cyclic Order: Waiting Times. *Bell Syst. Tech. J.* **49,** 399–413.

HEYMAN, D. P. 1977. The *T*-Policy for the *M/G/1* Queue. *Mgmt. Sci.* **23,** 775–778.

HEYMAN, D. P., AND M. J. SOBEL. 1982. *Stochastic Models in Operations Research.* McGraw-Hill, New York.

HEYMAN, D. P. Private communication.

KENDALL, D. 1951. Some Problems in the Theory of Queues. *J. Roy. Statist. Soc.* Ser. B, **13,** 151–185.

KINGMAN, J. 1962a. On Queues in Which Customers Are Served in Random Order. *Proc. Cambridge Phil. Soc.* **58,** 79–91.

KINGMAN, J. 1962b. The Effect of Queue Discipline on Waiting Time Variance. *Proc. Cambridge Phil. Soc.* **58,** 163–164.

KLEINROCK, L. 1975. *Queueing Systems, Vol. I: Theory.* Wiley Interscience, New York.

KLEINROCK, L. 1976. *Queueing Systems, Vol. II: Computer Applications.* Wiley Interscience, New York.

KLEINROCK, L., AND M. SCHOLL. 1977. Packet Switching in Radio Channels: New Conflict-Free Multiple Access Schemes for a Small Number of Data Users. In *Proceedings of the International Conference on Communications,* pp. 22.1-105 to 22.1–111, Chicago (June).

LEVY, Y., AND U. YECHIALI. 1975. Utilization of Idle Time in an $M/G/1$ Queueing System. *Mgmt. Sci.* **22,** 202–211.

LITTLE, J. 1961. A Proof of the Queueing Formula $L = \lambda W$. *Opns. Res.* **9,** 383–387.

MILLER, L. W. 1964. Alternating Priorities in Multi-Class Queues, Ph.D. dissertation, Cornell University, Ithaca, N.Y.

RIORDAN, J. 1961. Delays for Last-Come-First-Served Service and the Busy Period. *Bell Syst. Tech. J.* **40,** 785–793.

SCHOLL, M. 1976. Multiplexing Techniques for Data Transmission over Packet Switched Radio Systems, Computer Science Department, University of California, Los Angeles, UCLA-ENG-76123 (December). (Also published as Ph.D. dissertation.)

SHANTHIKUMAR, J. G. 1980. Some Analyses on the Control of Queues Using Level Crossing of Regenerative Processes. *J. Appl. Prob.* **17,** 814–821.

TAKÁCS, L. 1963. Delay Distributions for One Line with Poisson Input, General Holding Times, and Various Orders of Service. *Bell Syst. Tech. J.* **42,** 487–503.

VASICEK, O. A. 1977. An Inequality for the Variance of Waiting Time Under a General Queueing Discipline. *Opns. Res.* **25,** 879–884.

VAULOT, E. 1954. Délais d'attente des appel téléphoniques dans l'orde inverse de leur arrivée. *Comp. Rend. Acad. Sci.* **238,** 1188–1189.