THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 58, No. 3, March 1979
Printed in U.S.A.

Multiqueue Systems with Nonexhaustive Cyclic Service

By P. J. KUEHN

(Manuscript received July 14, 1978)

Queuing models with cyclic-type service are applicable for performance studies of polling mechanisms in data communication and switching systems or cyclic scheduling algorithms in real time computers. This paper provides an approximate analysis of the multiqueue system $M^{[X]}/G/1$ with batch Poisson input, general service times, general overhead (switchover) times, and a single server operating under a cyclic strategy with nonexhaustive service of queues. Based on a new concept of conditional cycle times, the generating function of the stationary probabilities of state, the Laplace-Stieltjes transforms of the delay distributions, and the mean waiting times are derived explicitly for each queue through an imbedded Markov chain approach and an independence assumption. The approximate analytic results are validated by computer simulations. Besides this analysis, a stability criterion is derived for the general case of GI/G/1 systems with cyclic priority service. The paper concludes with a number of studies of the behavior of cyclic queues discovering interesting properties such as the dependence of cycle times and waiting times on the arrival and service process types and on the efficiency of cyclic priorities.

I. INTRODUCTION

Cyclic service is a frequently used mechanism for the information transfer between peripheral units and their centralized control opposed to asynchronous or synchronous interrupt mechanisms. In a cyclic service operation, the centralized control scans the peripheral units in a cyclic sequence. At each peripheral unit, the queue of waiting items (user or control data) is served either completely ("exhaustive service") or up to a specified maximum number of transferred items per scan ("nonexhaustive service") until the centralized device switches over to the succeeding unit within the cycle sequence. Examples of this type of operation are found in data communications systems (polling, asynchronous multiplexing), telephone switching systems (device scanning), and certain I/O mechanisms of real-time computers. The performance of these cyclic service mechanisms is of considerable interest for traffic engineering, namely with respect to throughput and resource utilizations, delays, unbalanced load, overload behavior, and the influence of various statistical properties of the traffic.

In the sequel, we refer to the general cyclic queuing model shown in Fig. 1. There are g arrival groups of "customers" and their corresponding waiting lines (queues). Customers of group j arrive according to a general independent (GI) arrival process with probability distribution function (pdf) $A_j(t) = P\{T_{Aj} \leq t\}$, where T_{Aj} denotes the random variable of the interarrival time in queue j, $j = 1, 2, \cdots, g$, and $\lambda_j = 1/ET_{Aj}$ defines the arrival rate of customers in queue j. Special cases of the GI arrival processes are: D (deterministic), M (Markovian), E_k (Erlangian order k), or H_2 (hyperexponential order 2). In case of batch arrivals, the arrival process is defined by both the random interarrival time T_{Bj} of batches and the random batch size K_j, $j = 1, 2, \cdots, g$. The batch size in queue j is given by its probability distribution $q_{jk} = P\{K_j = k\}$, $k = 0, 1, \cdots$. The total arrival rate of j-customers λ_j and the arrival rate of batches λ_{Bj} are related to each other through $\lambda_j = \lambda_{Bj} \cdot EK_j$. For the special case of deterministic arrival processes in more

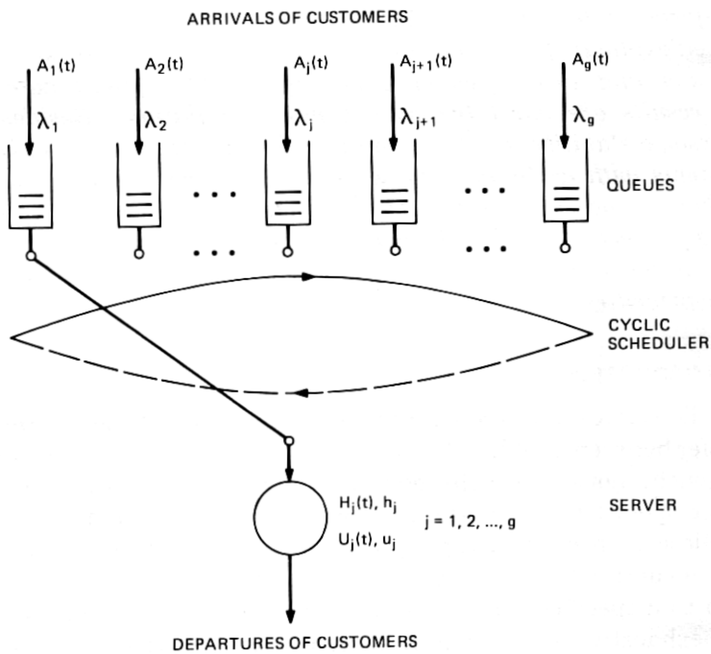

Fig. 1—Cyclic queuing model.

than one queue, a "phase shift parameter" is additionally needed which describes the relation between the periodic arrival patterns in those queues.

Similarly, customers of queue j receive a random service time T_{Hj} with pdf $H_j(t) = P\{T_{Hj} \leq t\}$ and mean $h_j = ET_{Hj}, j = 1, 2, \cdots, g$. Once the server has finished service at a particular queue j, it switches to the succeeding queue in a finite switchover (overhead) time T_{Uj} with pdf $U_j(t) = P\{T_{Uj} \leq t\}$ and mean $u_j = ET_{Uj}, j = 1, 2, \cdots, g$.

Finally, the general nonexhaustive cyclic operation of the server may be specified by a sequence ("cycle") $\{i_1, i_2, \cdots, i_l\}$, where $i_k \in (1, 2, \cdots, g)$ denotes the number of that queue which is served in kth position within the cycle (l = cycle length). The sequence $\{i_1, i_2, \cdots, i_l\}$ is repeated in a cyclic manner. If there is no customer to serve from the i_kth queue, the server switches over to the i_{k+1}st queue (modulo l). An example of this general (mixed) cyclic sequence for $g = 3$ queues is $\{1, 2, 1, 3, 1, 2\}$, where $l = 6$. In this case, an overhead phase is inserted after every queue visit. Cyclic schedules with different frequencies of visits at the various queues within a cycle will also be referred to as "cyclic priority service." An important special case of cyclic priority service is obtained when all visits at a particular queue within a cycle are clustered such that the server attends queue 1 successively up to l_1 times, queue 2 successively up to l_2 times, and so on ($l = l_1 + l_2 + \cdots + l_g$). In this case, an overhead occurs only when changing to another queue. Limiting cases of this schedule are cycles with $l_j = 1, j = 1, 2, \cdots, g$ ("ordinary cyclic service" $\{1, 2, \cdots, g\}$) or cycles with $l_j \gg 1, j = 1, 2, \cdots, g$ ("exhaustive cyclic service"). The queuing analysis in this paper is limited to the practical important case of ordinary cyclic service; for stability and simulation studies the more general (nonmixed) cyclic priority service will be considered.

Queues with cyclic service have received considerable attention in literature (see Refs. 1 to 16). Cyclic queues with exhaustive service with or without overhead have been treated in case of $M/G/1$ models[1-8] and in case of discrete arrival and service processes.[9] The case of nonexhaustive cyclic service involves considerable mathematical difficulties and has been treated rigorously only for $M/G/1$ models with two queues without overhead.[10,11] Because of the mathematical intractability of most cyclic queuing problems, several approximate methods were suggested.[12-16] The approximate methods usually rest on some simplifying assumptions such as the "independence assumption"[12] under which the stochastic processes within a particular queue are considered more or less independent of the processes within the other queues.

In Section II of this paper, we first derive a stability criterion for queues of the type $GI/G/1$ with cyclic priority service. Section III deals with the cycle time analysis for $GI/G/1$ queues in case of ordinary cyclic service. Sections IV and V present an analysis of the probabilities of state and the waiting times for cyclic queuing models of the types $M/G/1$ and $M^{[X]}/G/1$ with ordinary cyclic service and general overhead, respectively. In Section VI finally, we report various numerical results of the approximate analysis and of computer simulations for validation and qualitative performance studies. Some of those results discover new insight into the properties of cyclic queues and could have direct consequences for system engineering and future research as well.

II. STABILITY OF CYCLIC QUEUES

Contrary to most standard queuing problems, an obvious and simple criterion does not exist under which a queue in a cyclic queuing system stays stable. In the following sections, we develop a stability criterion for $GI/G/1$ multiqueue systems with a (nonmixed) cyclic priority service.

2.1 A stability criterion for queues with cyclic priority service

Following analogously to a common definition for stability in system theory, a queuing system will be called "stable" if for positive service times and finite input rates the average queue lengths are limited (note that a stationary queue is stable, whereas a stable queue need not necessarily be stationary). Additionally, we assume that all arrival and service processes are stationary so that the following reasoning can be based on average values *independent* of specific distributional assumptions.

Let T_C be the random cycle time, $c = ET_C$ the average cycle time, and $c_0 = u_1 + u_2 + \cdots + u_g$ the average of the cycle time under the condition that no customer is served during a cycle. The average number of arriving j customers during a cycle is $n_j = \lambda_j c$. In the stationary state of the system, the average number of arriving j customers equals the average number of served j customers, $j = 1, 2, \cdots, g$. Thus, we have

$$c = c_0 + \sum_{j=1}^{g} (\lambda_j c) h_j,$$

from which we find the result

$$c = \frac{c_0}{1 - \rho_0}, \tag{1}$$

where $\rho_0 = \rho_1 + \rho_2 + \cdots + \rho_g$ defines the total server utilization and $\rho_j = \lambda_j h_j$ is the server utilization by j customers only, $j = 1, 2, \cdots, g$. The result according to (1) has already been discovered for cyclic queues with exhaustive and ordinary cyclic service. To find the boundary of system stability, we proceed as follows: First, we state a stability criterion for a particular queue j under the condition of stability of the residual queues. This condition can always be achieved for sufficiently small arrival rates in the residual queues. The whole system is stable if and only if all individual stability conditions are satisfied simultaneously.

Under the condition that all queues $\nu \neq j$ are stable, queue j approaches the stability boundary as $n_j \to l_j$; this corresponds to a maximum arrival rate $\lambda_{j\,\max}$ at the margin $n_j = l_j$ and an average cycle length c_j^*:

$$\lambda_{j\,\max} = \frac{l_j}{c_j^*}, \qquad \text{where} \qquad c_j^* = \frac{c_0 + l_j h_j}{1 - \rho_0 + \rho_j}. \tag{2a}$$

Thus, the system is stable if for all queues

$$\lambda_j < \lambda_{j\,\max} = \frac{l_j}{c_0 + l_j h_j} \cdot (1 - \rho_0 + \rho_j), \quad j = 1, 2, \cdots, g, \tag{2b}$$

are fulfilled simultaneously. In a similar way, criteria of partial stability can be stated in cases where some queues are saturated (a saturated queue i contributes to the average cycle time by $l_i \cdot h_i$). Finally, it should be noted that the average cycle time c stays always stable since $c \leq c_0 + l_1 h_1 + \cdots + l_g h_g$.

2.2 Examples

To further explore the stability criterion, consider the example of $g = 2$ queues. From (2b) we find the following relationships between λ_1 and λ_2:

$$\lambda_1 < \frac{l_1}{c_0 + l_1 h_1} \cdot (1 - \lambda_2 h_2), \qquad \lambda_2 < \frac{l_2}{c_0 + l_2 h_2} \cdot (1 - \lambda_1 h_1), \tag{3a, b}$$

where $c_0 = u_1 + u_2$. These relationships are shown graphically by two marginal lines in Fig. 2 with the intersection

$$\lambda_{jo} = \frac{l_j}{c_0 + l_1 h_1 + l_2 h_2}, \qquad j = 1, 2. \tag{4}$$

The absolute stable region is below the hatched area when both individual criteria (3a), (3b) are fulfilled simultaneously. For $\lambda_1 < \lambda_{1o}$ queue 2 always saturates first, whereas for $\lambda_2 < \lambda_{2o}$ queue 1 saturates first. At the intersection $(\lambda_{1o}, \lambda_{2o})$, both queues saturate simultaneously. Similarly, for $\lambda_1 > \lambda_{1o}$, $\lambda_2 > \lambda_{2o}$, both queues are saturated (absolute unstable region). Within the intermediate regions $\lambda_1 > l_1 \cdot (1 - \lambda_2 h_2)/(c_0 + l_1 h_1)$, $\lambda_2 < \lambda_{2o}$ (or $\lambda_2 > l_2 \cdot (1 - \lambda_1 h_1)/(c_0 + l_2 h_2)$, $\lambda_1 < \lambda_{1o}$), queue 1 (or queue 2) is saturated, whereas queue 2 (or queue 1) is stable; in these regions of partial stability, the cyclic queuing system

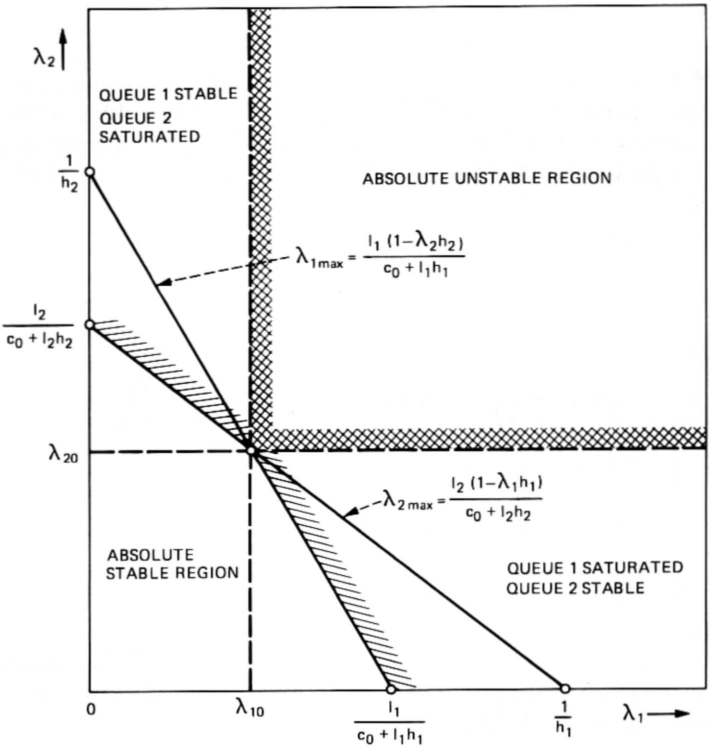

Fig. 2—Stability regions for a $GI/G/1$ queuing system with 2 queues and cyclic priority service.

can be considered as consisting only from the stable queue where the contribution of the unstable queue to the cycle time affects as an increased cycle time overhead of l_1 (or l_2) consecutive service times T_{H1} (or T_{H2}). Furthermore, starting from any point within the absolute stable region and increasing λ_1 and λ_2 simultaneously, we state that the queue with the greater λ_j/l_j ratio reaches saturation first, *independent* of the service and overhead time parameters. This statement differs from many other queuing stability criteria.

Finally, we discuss briefly two important special cases of the above example. The first special case is that of ordinary cyclic service ($l_1 = l_2 = 1$). The intersection in Fig. 2 falls on the median $\lambda_1 = \lambda_2$. This means that the queue with the greater arrival rate always saturates first. This result was already found by M. Eisenberg[11] for $M/G/1$ systems without overhead. Additionally, the average number of j customers served during a cycle is identical with the probability α_j that the server meets at least one customer in queue j:

$$\alpha_j = \lambda_j c = \frac{c_0}{1 - \rho_0} \cdot \lambda_j, \qquad j = 1, 2. \tag{5}$$

The second special case is that of exhaustive cyclic service (l_1, $l_2 \gg 1$). In this case, both marginal lines fall together on $\lambda_1 h_1 + \lambda_2 h_2 = 1$. The stability criterion is simply $\lambda_1 h_1 + \lambda_2 h_2 = \rho_1 + \rho_2 = \rho_0 < 1$. Approaching the stability margin, both queues saturate simultaneously independent of c_0.

III. CYCLE TIME ANALYSIS

In this section, we consider multiqueue systems of the type $GI/G/1$ with overhead and ordinary cyclic service. Since the queuing analysis in Sections IV and V is based on the knowledge of cycle time, we briefly discuss a known result and then develop an improved approximate solution for the pdf of the cycle time.

3.1 Cycle time analysis by Hashida and Ohara

The exact solution for the pdf of the cycle time T_C is still unknown, except for the mean c in (1). Based on the probabilities α_j in (5) and the approximation assumption of independence, Hashida and Ohara[13] gave the following expression for the Laplace-Stieltjes transform (LST) of the cycle time pdf $C(t) = P\{T_C \leq t\}$:

$$\phi_C(s) = \prod_{i=1}^{g} \phi_{Ui}(s) \cdot \prod_{i=1}^{g} (\alpha_i \phi_{Hi}(s) + [1 - \alpha_i]). \qquad (6)$$

In (6), $\phi_C(s) = \int_{0-}^{\infty} e^{-st} dC(t)$ defines the LST of $C(t)$; similarly, $\phi_{Ui}(s)$ and $\phi_{Hi}(s)$ denote the LSTs of $U_i(t)$ and $H_i(t)$, respectively. The expression (6) follows directly when considering T_C as a sum of independent random variables.

From (6), the exact mean cycle time c follows straightforwardly and agrees with (1). However, it was found by intensive simulations (some of them are given in Section VI) that (6) underestimates the cycle time variance and, herewith, also the mean waiting times. For this reason, we shall now improve the cycle time analysis by introduction of a new concept of "conditional cycle times."

3.2 Conditional cycle times

The basic idea of the queuing analysis in Ref. 13 and in Sections IV and V of this paper is the description of the queue length of a particular queue j at the scan instant by an imbedded Markov chain. The influence of all queues $\nu \neq j$ on the considered queue j will be expressed only through the cycle time. The cycle time T_C is the time interval between two successive scan instants of a queue (say, j). However, a particular realization of T_C clearly depends on whether a j customer is served or not in a cycle. Therefore, we introduce two conditional cycle times $T_{C_j'}$ and $T_{C_j''}$, with respect to the considered queue j for cycles without or with a service time contribution to the cycle by a customer of queue j, respectively. The corresponding cycles are denoted by $\mathbf{C}_j'$

and C_j''. Since $ET_{C_j''} > ET_{C_j'}$, it is more likely that after a long cycle another long one is followed, and vice versa. Thus, the concept of conditional cycle times increases the cycle time variance through a reduction of the independence assumption by explicit consideration of some history of a current cycle.

Let $C_j'(t)$ and $C_j''(t)$ be the pdfs of the conditional cycle times $T_{C_j'}$ and $T_{C_j''}$, and α_{ji}', α_{ji}'' the corresponding probabilities for the service of an i customer ($i \neq j$) during a conditional cycle without or with a j service, respectively. Then, it follows by similar reasoning as for (6):

$$\phi_{C_j'}(s) = \prod_{i=1}^{g} \phi_{Ui}(s) \cdot \prod_{i \neq j} (\alpha_{ji}' \phi_{Hi}(s) + [1 - \alpha_{ji}']), \tag{7a}$$

$$\phi_{C_j''}(s) = \prod_{i=1}^{g} \phi_{Ui}(s) \cdot \prod_{i \neq j} (\alpha_{ji}'' \phi_{Hi}(s) + [1 - \alpha_{ji}'']) \cdot \phi_{Hj}(s). \tag{7b}$$

For the (unconditional) cycle time T_C we find from the law of total probabilities

$$\phi_C(s) = (1 - \alpha_j)\phi_{C_j'}(s) + \alpha_j \phi_{C_j''}(s). \tag{8}$$

With $c_j' = ET_{C_j'}$ and $c_j'' = ET_{C_j''}$ we state the conditional cycle time balances:

$$c_j' = c_0 + \sum_{i \neq j} \alpha_{ji}' h_i, \tag{9a}$$

$$c_j'' = c_0 + \sum_{i \neq j} \alpha_{ji}'' h_i + h_j, \tag{9b}$$

$$c = (1 - \alpha_j) c_j' + \alpha_j c_j''. \tag{9c}$$

Similarly, as in (5), we assume

$$\alpha_{ji}' = \lambda_i c_j', \tag{9d}$$

$$\alpha_{ji}'' = \lambda_i c_j'', \qquad i \neq j. \tag{9e}$$

Inserting (9d), (9e) in (9a), (9b), we find

$$c_j' = \frac{c_0}{1 - \rho_0 + \rho_j}, \tag{10a}$$

$$c_j'' = \frac{c_0 + h_j}{1 - \rho_0 + \rho_j}. \tag{10b}$$

Note that the exact value of c in (1) follows from (9c) and (10a), (10b). It should also be mentioned that the solution (10a), (10b) holds only as long as $\alpha_{ji}'' \leq 1$. This condition is always fulfilled in case of symmetrical load ($\lambda_1 = \lambda_2 = \cdots = \lambda_g$, $h_1 = h_2 = \cdots = h_g$). In case of higher unsymmetrical loads, it can indeed happen that $\alpha_{ji}'' > 1$ so that α_{ji}'' can no longer be interpreted as probability; this difficulty can be overcome by a suitable limitation of α''_{ji} by 1 (i.e., queue i always contributes a service time to the conditional cycle time $T_{C''_j}$).

With (9d), (9e) and (10a), (10b), the conditional cycle time pdfs $C'_j(t)$ and $C''_j(t)$ are completely defined. The mean conditional cycle times c'_j and c''_j are given by (10a), (10b). For the variances, we find from (7a), (7b):

$$\text{VAR } T_{C'_j} = \sum_{i=1}^{g} \text{VAR } T_{Ui} + \sum_{i \neq j} (\alpha'_{ji} h_i^{(2)} - \alpha'^2_{ji} \cdot h_i^2), \qquad (11a)$$

$$\text{VAR } T_{C''_j} = \sum_{i=1}^{g} \text{VAR } T_{Ui} + \sum_{i \neq j} (\alpha''_{ji} h_i^{(2)} - \alpha''^2_{ji} \cdot h_i^2) + \text{VAR } T_{Hj}, \qquad (11b)$$

where $h_i^{(2)} = ET_{Hi}^2$ denotes the ordinary second moment of T_{Hi}. The cycle time variance is finally given by

$$\text{VAR } T_C = (1 - \alpha_j) \cdot [\text{VAR } T_{C'_j} + c'^2_j] + \alpha_j \cdot [\text{VAR } T_{C''_j} + c''^2_j] - c^2. \qquad (11c)$$

The second moments of the cycle times follow from the definition $\text{VAR } T_C = c^{(2)} - c^2$.

IV. QUEUING ANALYSIS OF *M/G/1* SYSTEMS

Based on the concept of the conditional cycle times of Section III, a queuing analysis is given for multiqueue systems of the type $M/G/1$ with general overhead times and ordinary cyclic service by means of an imbedded Markov chain. Basically, the derivation follows the approach of Hashida and Ohara.[13]

4.1 Probabilities of state

For an exact analysis, the state of the system at a time t has to be defined such that all past history is summarized in it so that the future development of the system state process is completely determined from it. In the present case, the system state could be described by a vector $\{N_1(t), N_2(t), \cdots, N_g(t), I(t), X_0(t)\}$, where $N_j(t)$ defines the number of waiting customers in queue j, $j = 1, 2, \cdots, g$, $I(t)$ points to the present location of the server within the cycle, and $X_0(t)$ specifies the age of the current service (or overhead) phase of the server. An exact analysis on this base seems not to be feasible.

In the following analysis, only the state N_j of a particular queue j is considered. Moreover, the analysis does not apply to continuous time but is restricted to a set of special points, namely the scan instants (or departure instants) of the considered queue j. The time intervals between the scan instants of queue j are the conditional cycle times $T_{C'_j}$ and $T_{C''_j}$; the influence of all the other queues on the queue length process in queue j is completely expressed by those cycle times. Although the following imbedded Markov chain solution is formally exact, the analysis approach is approximate since $T_{C'_j}$ $T_{C''_j}$ are assumed to be independent and identically distributed (iid) variables; the expressions for their pdf are only approximations, too.

The outlined method only renders results for the particular queue j under consideration. In the case of unsymmetrical systems, the procedure must be repeated for the other queues, too. For ease of reading, we suppress the subscript j in the following treatment, i.e., we write $\lambda, h, \rho, c', c'', \cdots$ instead of $\lambda_j, h_j, \rho_j, c'_j, c''_j \cdots$.

4.1.1 State distribution at scan instants

We assume that the queuing sytem is in the stationary state. Let N be the number of waiting customers at the server arrival instant (scan instant) of a particular queue. We are interested in the stationary distribution

$$p_n = P\{N = n\}, \qquad n = 0, 1, 2, \cdots. \tag{12}$$

Because of the memoryless property of the arrival process, the system state of the considered queue forms an imbedded Markov chain at the discrete set of scan instants (renewal points). The stationary distribution satisfies the equation (see Ref. 17, pp. 167–174):

$$p_n = p_0 \cdot p_{0n} + \sum_{m=1}^{n+1} p_m \cdot p_{mn}, \qquad n = 1, 2, \cdots, \tag{13a}$$

where the transition probabilities p_{mn} are given by

$$p_{mn} = \begin{cases} \displaystyle\int_{t=0-}^{\infty} e^{-\lambda t} \cdot \frac{(\lambda t)^{n-m+1}}{(n-m+1)!} \cdot dC''(t), & m > 0 \\ \\ \displaystyle\int_{t=0-}^{\infty} e^{-\lambda t} \cdot \frac{(\lambda t)^{n}}{n!} \cdot dC'(t), & m = 0. \end{cases} \tag{13b}$$

Together with the normalizing condition

$$\sum_{n=0}^{\infty} p_n = 1, \tag{13c}$$

the stationary probabilities of state at the scan instants are completely determined by the set of equations (13a), (13b), and (13c). Introducing the probability generating function of the state distribution p_n, $n = 0, 1, 2, \cdots$,

$$G(x) = \sum_{n=0}^{\infty} p_n x^n, \tag{14}$$

we obtain after some algebraic manipulations

$$G(x) = p_0 \cdot \frac{x\phi_{C'}(z) - \phi_{C''}(z)}{x - \phi_{C''}(z)}, \qquad \text{where} \qquad z = \lambda(1 - x). \tag{15}$$

Note that $G(x)$ is completely expressed by p_0 and the LSTs of the two conditional cycle times. Using the identity $G(1) = 1$, we find from (15) through evaluation of $\lim_{x \to 1} G(x)$ by L'Hospital's rule

$$p_0 = \frac{1 - \lambda c''}{1 - \lambda(c'' - c')} = 1 - \alpha. \tag{16}$$

The latter identity can be shown by using equations (10a), (10b), and (5).

The expected number of waiting customers at the scan instant follows from

$$EN = \frac{d}{dx} G(x)\bigg|_{x=1}.$$

This results in

$$EN = p_0 \cdot \lambda \cdot \frac{\lambda c'^{(2)} \cdot (1 - \lambda c'') + c'(\lambda^2 c''^{(2)} + 2 - 2\lambda c'')}{2(1 - \lambda c'')^2}, \tag{17}$$

where $c'^{(2)} = ET^2_{C'}$, and $c''^{(2)} = ET^2_{C''}$.

4.1.2 State distribution at departure instants

Let N^* be the number of waiting customers within the considered queue which are left behind by a departing customer of that queue with distribution

$$p_n^* = P\{N^* = n\}, \qquad n = 0, 1, 2, \cdots \tag{18}$$

and generating function

$$G^*(x) = \sum_{n=0}^{\infty} p_n^* x^n. \tag{19}$$

The probability p_n^* can be expressed through the probability of having m customers at the scan instant given that the considered queue is not empty, $p_m/(1 - p_0)$, and the probability of $n - m + 1$ new arrivals in that queue during the subsequent service time of one customer. Hence,

$$p_n^* = \sum_{m=1}^{n+1} \frac{p_m}{1 - p_0} \cdot \int_{t=0-}^{\infty} e^{-\lambda t} \cdot \frac{(\lambda t)^{n-m+1}}{(n - m + 1)!} \cdot dH(t), \qquad n = 0, 1, 2, \cdots. \tag{20}$$

Substituting (20) in (19) and interchanging the order of summation and integration, we find

$$G^*(x) = \frac{1}{1-p_0} \cdot \frac{G(x) - p_0}{x} \cdot \phi_H(z), \tag{21}$$

where

$$z = \lambda(1 - x).$$

Therefore, it follows for the expected number of customers at the departure epoch:

$$EN^* = \frac{d}{dx} G^*(x)\bigg|_{x=1} = \frac{EN}{1-p_0} - 1 + \rho. \tag{22}$$

On the other hand, EN^* equals the expected number of customers which have arrived during the sojourn (waiting + service) time of the departing customer (for this, consider as an example the queue discipline FCFS, first come, first-served). Hence, $EN^* = \lambda \cdot (w + h)$, where $w = ET_W$ denotes the average waiting time in the considered queue (service being excluded). Solving for w, we find with (22)

$$w = \frac{1}{\lambda} \cdot \left[\frac{EN}{1-p_0} - 1\right]. \tag{23}$$

4.2 Delay analysis

For the following derivation, the queue discipline FCFS is assumed. Let T_W be the waiting time which an arbitrary customer of the considered queue (in the following denoted by "test customer") has to undergo with pdf $W(t)$ and LST $\phi_W(s)$. Through an analogous reasoning as in the previous section, p_n^* can alternatively be considered as the distribution of the number of arriving customers during the sojourn time T_S of the test customer. Since $T_S = T_W + T_H$ and since T_W and T_H are independent of each other, the pdf of T_S is the convolution of $W(t)$ and $H(t)$, symbolized by $W(t) \circledast H(t)$. Hence,

$$p_n^* = \int_{t=0-}^{\infty} e^{-\lambda t} \cdot \frac{(\lambda t)^n}{n!} \cdot d(W(t) \circledast H(t)), \qquad n = 0, 1, 2, \cdots. \tag{24}$$

Applying (19) in (24), we find $G^*(x) = \phi_W(z) \cdot \phi_H(z)$, where $z = \lambda(1 - x)$, which, with (21), finally results in

$$\phi_W(s) = \frac{1 - \lambda c''}{c'} \cdot \frac{1 - \phi_{C'}(s)}{s - \lambda \cdot [1 - \phi_{C''}(s)]}. \tag{25}$$

From (25) we find for the mean waiting time

$$w = -\frac{d}{ds}\phi_W(s)\bigg|_{s=0} = \frac{c'^{(2)}}{2c'} + \frac{\lambda c''^{(2)}}{2(1 - \lambda c'')}. \tag{26}$$

Equation (26) reveals that the mean waiting time depends basically on the first and second moments of the conditional cycle times. Note also that (26) agrees with (23) when the corresponding results for p_0, EN, c', c'', $c'^{(2)}$, and $c''^{(2)}$ from eqs. (16), (17), (10a), (10b), and (11a), (11b), respectively, are inserted.

It may be mentioned that the result (26) can also be derived directly through the application of renewal theory and Little's law: An arriving test customer of the considered queue meets either a cycle $\mathbf{C}'$ or $\mathbf{C}''$ in progress. Since the arrival process is a Markovian process, the probabilities of meeting a cycle $\mathbf{C}'$ or $\mathbf{C}''$ is simply the weighted ratio of frequencies, i.e. $(1-\alpha)\cdot(c'/c)$ or $\alpha\cdot(c''/c)$, respectively. According to our approximation assumption, the conditional cycle times $T_{C'}$ and $T_{C''}$ are iid-variables. Thus, the average residual cycle times are $c'^{(2)}/2c'$ and $c''^{(2)}/2c''$ according to renewal theory (see, for example, Ref. 17, pp. 158–161). The average waiting time w consists of the average residual cycle time and the product of the mean cycle time c'' and the average number L of customers met at the arrival instant of the test customer; the latter one can be expressed through Little's law (see, for example, Ref. 17, pp. 156–158) through $L=\lambda\cdot w$. The average waiting time w can now be balanced as

$$w=(1-\alpha)\cdot\frac{c'}{c}\cdot\frac{c'^{(2)}}{2c'}+\alpha\cdot\frac{c''}{c}\cdot\frac{c''^{(2)}}{2c''}+(\lambda w)c''. \qquad (27)$$

Solving for w, we yield precisely the result (26) from (27).

The pdf $W(t)$ can be obtained by the inversion of (25) either through a partial fraction expansion (in case of rational LSTs), by the numerical inversion technique of D. Jagerman,[18] or by an approximation using the ordinary first and second moments.[19]

Finally, we mention that (25) includes the exact result for the limiting case $g=1$ of a single cyclic queue with overhead. Furthermore, another limit with zero overhead can be derived from (25); this case represents the worst case with respect to the approximation accuracy (see Section VI).

V. QUEUING ANALYSIS OF $M^{[X]}/G/1$ SYSTEMS

In this section, the solution of Section IV for single Poisson arrivals (M) is generalized to batch Poisson arrivals ($M^{[X]}$) in every queue. The analysis follows analogously to Section IV; i.e., we consider all processes with respect to a particular queue j. Again, the subscript j will be suppressed for ease of reading.

5.1 Probabilities of state

5.1.1 Arrival process

Customers of the considered queue arrive in batches of size K with distribution

$$q_k = P\{K = k\}, \qquad k = 0, 1, 2, \cdots \tag{28a}$$

and probability generating function

$$Q(x) = \sum_{k=0}^{\infty} q_k x^k. \tag{28b}$$

The interarrival times of batches are exponentially distributed with mean $1/\lambda_B = EK/\lambda$, where λ and λ_B are the arrival rates of customers and batches within the considered queue, respectively.

Let $N_B(t)$ be the number of batch arrival instants in $(0, t)$ with distribution

$$P\{N_B(t) = n\} = \frac{(\lambda_B t)^n}{n!} \cdot e^{-\lambda_B t}, \qquad n = 0, 1, 2, \cdots \tag{29a}$$

and probability generating function

$$g(x, t) = \sum_{n=0}^{\infty} \frac{(\lambda_B t)^n}{n!} \cdot e^{-\lambda_B t} \cdot x^n = e^{-\lambda_B t(1-x)}. \tag{29b}$$

Finally, let $N_A(t)$ be the total number of customers arriving at the considered queue in $(0, t)$. Then, the probability generating function of the distribution $P\{N_A(t) = k\}$, $k = 0, 1, 2, \cdots$, is given by

$$h(x, t) = \sum_{k=0}^{\infty} P\{N_A(t) = k\} \cdot x^k = g(Q(x), t) = e^{-\lambda_B t[1-Q(x)]}. \tag{30}$$

5.1.2 State distribution at scan instants

Since the cycle time approximation of Section III holds for $GI/G/1$ cyclic queues, the same pdfs $C(t)$, $C'(t)$, and $C''(t)$ can be used for batch arrival processes. As in Section 4.1., let p_n be the stationary probability of state for n customers waiting within the considered queue at the scan instants. The transition probabilities p_{mn} in (13a) are now

$$p_{mn} = \begin{cases} \displaystyle\int_{t=0-}^{\infty} \sum_{\nu=0}^{\infty} P\{N_B(t) = \nu\} \\ \quad \cdot P\{N_A(t) = n - m + 1 \,|\, N_B(t) = \nu\} \cdot dC''(t), & m > 0 \\ \\ \displaystyle\int_{t=0-}^{\infty} \sum_{\nu=0}^{\infty} P\{N_B(t) = \nu\} \\ \quad \cdot P\{N_A(t) = n \,|\, N_B(t) = \nu\} \cdot dC'(t), & m = 0. \end{cases} \tag{31}$$

The probabilities of state p_n are completely determined by (13a), (31), and (13c). The application of the generating function results finally in the same expression for $G(x)$ as in (15), however, with $z = \lambda_B \cdot [1 - Q(x)]$, where $Q(x)$ is defined by (28a), (28b). Also, for p_0 the identical result is obtained as in (16). Further results can easily be derived analogously as in Section 4.1.

5.1.3 State distribution at departure instants

Using the same definitions for p_n^* and $G^*(x)$ as in Section 4.1.2, we find, instead of (20),

$$p_n^* = \sum_{m=1}^{n+1} \frac{p_m}{1 - p_0} \cdot \int_{t=0-}^{\infty} \sum_{\nu=0}^{\infty} P\{N_B(t) = \nu\} \cdot P\{N_A(t) = n - m + 1 \mid N_B(t) = \nu\} \, dH(t), \qquad n = 0, 1, 2, \cdots . \quad (32)$$

This again results in the same expression for $G^*(x)$ as in (21) with $z = \lambda_B \cdot [1 - Q(x)]$, from which further results could be derived analogously.

5.2 Delay analysis

Following the method outlined in Secton 4.2, p_n^* is also the distribution of the number of arriving customers during the sojourn time of a test customer of that queue. The number N^* of customers left behind in the considered queue by the departing test customer is now built up from *two* components:

$$N^* = N_1^* + N_2^*,$$

where

N_1^* = the number of customers that had arrived together with the test customer in one batch but that were *behind* the test customer

N_2^* = the number of customers that had arrived in *subsequently* arriving batches during the sojourn time T_S of the test customer.

Let $r_n = P\{N_1^* = n\}$, $n = 0, 1, 2, \cdots$, be the probability that the departing test customer leaves n customers behind which had arrived together with the test customer in one batch. The test customer arrived in a batch of size $K = k$ with probability (see Ref. 20)

$$q_k^* = \frac{kq_k}{EK}, \quad k = 1, 2, 3, \cdots . \quad (33)$$

The test customer is first, second, $\cdots$, kth in the batch of size k with probability $1/k$. Thus, q_k^*/k defines the probability that the test customer arrived in a batch of size k in $(k-n)$th position, $n = 0, 1, 2, \cdots, k-1$. Then,

$$r_n = \sum_{k=n+1}^{\infty} \frac{q_k^*}{k} = \frac{1}{EK} \cdot \sum_{k=n+1}^{\infty} q_k, \qquad n = 0, 1, 2, \cdots, \tag{34}$$

and probability generating function

$$R(x) = \sum_{n=0}^{\infty} r_n x^n = \frac{1}{EK} \cdot \frac{1 - Q(x)}{1 - x}. \tag{35}$$

Now, we can establish the relation between p_n^* and T_S analogously as in Section 4.2:

$$p_n^* = \int_{t=0-}^{\infty} \left[\sum_{\mu=0}^{n} P\{N_1^* = \mu\} \cdot \sum_{\nu=0}^{\infty} P\{N_2^*(t) = n - \mu \mid N_B(t) = \nu\} \cdot P\{N_B(t) = \nu\} \right] \cdot d(W(t) \circledast H(t)), \qquad n = 0, 1, 2, \cdots. \tag{36}$$

In (36), the bracket term expresses the probability of new arrivals within a sojourn time of length t through consideration of all principal possibilities of batch configurations of the departing test customer.

Introducing r_n and $R(x)$ from (34) and (35) and applying (19) on (36), we find after some intermediate calculations

$$G^*(x) = \phi_W(z) \cdot \phi_H(z) \cdot R(x), \tag{37}$$

where

$$z = \lambda_B[1 - Q(x)].$$

Equating both expressions in (21) and (37) yields the final result

$$\phi_W(s) = \frac{1 - \lambda c''}{\lambda c'} \cdot \frac{1 - \phi_{C'}(s)}{\phi_{C''}(s) - x} \cdot \frac{1}{R(x)}, \tag{38a}$$

where $x = f(s)$ the solution of

$$s = \lambda_B[1 - Q(x)]. \tag{38b}$$

From (38a), (38b), we find for the mean waiting time of a customer

$$w = \left(\frac{c'^{(2)}}{2c'} + \frac{\lambda c''^{(2)}}{2(1 - \lambda c'')} \right) + \frac{c''}{2(1 - \lambda c'')} \cdot \left[\frac{EK^2}{EK} - 1 \right]. \tag{39}$$

Note that the mean waiting time consists of two terms; the first term is identical with that of an $M/G/1$ system [see (26)], whereas the second term expresses the influence of *batch* arrivals.

Analogously to Section 4.2, the mean waiting time can be derived directly. Let $w(i)$ be the mean conditional waiting time of a customer who is ith in his batch. For $w(1)$, a similar balance can be stated as in (27):

$$w(1) = (1-\alpha)\cdot\frac{c'}{c}\cdot\frac{c'^{(2)}}{2c'} + \alpha\cdot\frac{c''}{c}\cdot\frac{c''^{(2)}}{2c''} + (\lambda w)c''. \tag{40a}$$

The relationship between $w(1)$ and $w(i)$ is

$$w(i) = w(1) + (i-1)\cdot c''. \tag{40b}$$

The mean waiting time w, irrespective of the test customer's position within the batch, follows by averaging over the conditional waiting times. Thus, with (33) and (40b), we have

$$w = \sum_{k=1}^{\infty} \frac{kq_k}{EK}\cdot\frac{1}{k}\cdot\sum_{i=1}^{k} w(i) = w(1) + \frac{c''}{2}\cdot\left[\frac{EK^2}{EK} - 1\right]. \tag{41a}$$

Inserting (41a) in (40a) and solving for $w(1)$ yields

$$w(1) = \frac{c'^{(2)}}{2c'} + \frac{\lambda c''^{(2)}}{2(1-\lambda c'')} + \frac{\lambda c''^{(2)}}{2(1-\lambda c'')}\cdot\left[\frac{EK^2}{EK} - 1\right]. \tag{41b}$$

The mean waiting time w is completely determined with (41a), (41b) and agrees with (39).

Finally, we give the explicit results for w in the case of two special batch size distributions. For *constant batch size k*, i.e., $q_i = \delta(i, k)$, we find

$$w|_{M^{[X]}/G/1} = w|_{M/G/1} + \frac{k-1}{2}\cdot\frac{c''}{1-\lambda c''}. \tag{42a}$$

In the case of *geometrically distributed batch sizes*, which are defined by $q_i = q^i\cdot(1-q)$, $i = 0, 1, \cdots$, and $q = (EK+1)/EK$, the result is

$$w|_{M^{[X]}/G/1} = w|_{M/G/1} + EK\cdot\frac{c''}{1-\lambda c''}. \tag{42b}$$

The expressions (42a), (42b) demonstrate at first the increase of the waiting time through the batch Poisson arrival process compared to the pure Poisson arrival process and, second, the increase of w through geometrically distributed batches against constant batches.

VI. NUMERICAL RESULTS

In this section, the results of the approximate analysis are validated by computer simulations. Further results are given to show various properties of cyclic queuing systems.

6.1 Cycle time variance for ordinary cyclic service

Since the mean cycle time c according to (1) is always exact, the approximation accuracy can be judged in a first step by the cycle time variance VAR T_C (note that even the pdf $C(t)$ of the cycle time would not be sufficient for a complete validation since successive cycle times are not independent of each other; for a more complete validation, some covariance measure should be considered, too). We expect very good accuracy for low traffic (since the independence assumption is asymptotically exact for zero arrival rates) as well as for heavy traffic (since each of the queues contributes in the limit with a full service time to the cycle so that the cycle times become independent of each other again).

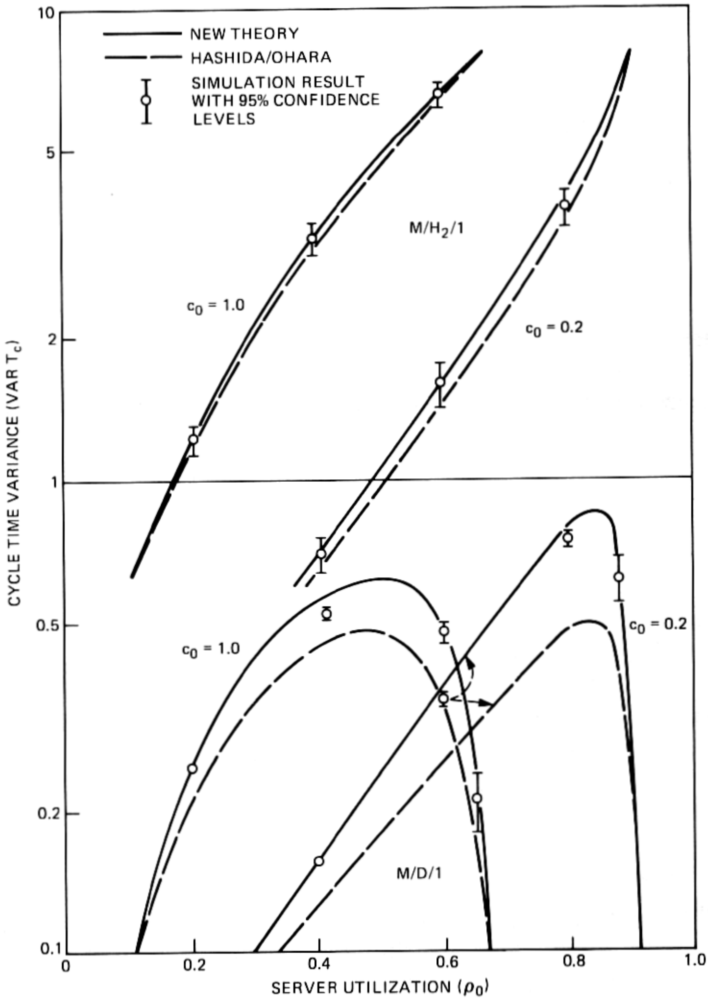

Fig. 3—Accuracy of cycle time variances

Parameters: g = 2 symmetrical queues $M/D/1$ and $M/H_2/1$
c_H = 2.0 coefficient of variation of service times for $M/H_2/1$
$h_1 = h_2 = 1$ average service times
c_0 = 0.2 and 1.0, constant overhead.

Table I—Cycle time variance for various $GI/M/1$ queues with ordinary cyclic service

Parameters: $g = 10$ symmetrical queues; $h_j = 1$, $u_j = 0.1$ (0.5), $\lambda_j = \lambda$, $j = 1, 2, \cdots, 10$; Constant overhead times; $D/M/1$: Equal phase shift between arrival instants; $H_2/M/1$: Interarrival time coefficient of variation $c_A = 2.0$; $M^{[X]}/M/1$: Constant batch size $k = 4$.

Parameters		VAR T_C (simulation)					VAR T_C (approx.)
ρ_0	c_0	$D/M/1$	$E_4/M/1$	$M/M/1$	$H_2/M/1$	$M^{[X]}/M/1$	$GI/M/1$
0	1.0	0	0	0	0	0	0
0.4	1.0	1.83	1.88	1.91	1.88	1.66	1.38
0.6	1.0	6.60	5.38	5.14	5.03	4.61	3.25
0.8	1.0	21.90	15.40	14.50	12.90	11.00	9.20
0.909	1.0	10.00	10.00	10.00	10.00	10.00	10.00
0	5.0	0	0	0	0	0	0
0.2	5.0	2.23	2.77	2.78	2.81	2.64	2.38
0.4	5.0	11.10	7.85	7.59	7.28	6.66	5.80
0.6	5.0	21.70	13.00	11.30	10.70	10.40	10.30
0.667	5.0	10.00	10.00	10.00	10.00	10.00	10.00

In Fig. 3, VAR T_C is shown dependent on the server utilization ρ_0 in case of $g = 2$ symmetrical queues of the types $M/D/1$ and $M/H_2/1$, each with two cases of constant overhead. As expected, the cycle time variance depends largely on the pdfs of the service and overhead times. The solid curves of the new approximation with the concept of conditional cycle times compare generally better with the simulation than the previous theory by Hashida and Ohara,[13] especially for low overhead. From a large number of computer simulations for $M/G/1$ systems, we made the following qualitative observations:

(*i*) The cycle time variance accuracy decreases with increasing number of queues and increasing service time variance.

(*ii*) The cycle time variance accuracy increases with increasing overhead and for approaching the low or heavy traffic region.

(*iii*) Observations (*i*) and (*ii*) apply to the new and old theory; the concept of conditional cycle times, however, yields generally a better accuracy.

Since the approximation for the pdf of the cycle time is independent of the arrival process type, it is interesting to know how the actual cycle time variance depends on various process types. For comparison, five different $GI/M/1$ systems with $g = 10$ queues (the accuracy is generally better for $g < 10$), two cases of overhead, and five cases of load have been considered (see Table I). Summarizing, we make the following observations:

(*i*) The cycle time variance depends indeed on the arrival process type. This dependence decreases, however, as the load approaches the low or the heavy traffic regions.

(*ii*) For medium loads, the cycle time variance may *decrease* as the arrival process peakedness increases.

(*iii*) The approximation generally underestimates the true cycle time variance. The accuracy increases with the overhead, the peakedness of the arrival process, and as the load approaches the low or heavy traffic region.

At first sight, observation (*ii*) is counterintuitive and surprising since the mean waiting time generally *increases* with the arrival process peakedness (see, for example, Fig. 9). However, regular arrival patterns may result in very short and very long cycles since many idle cycles could be produced after a service until the next arrival occurs.

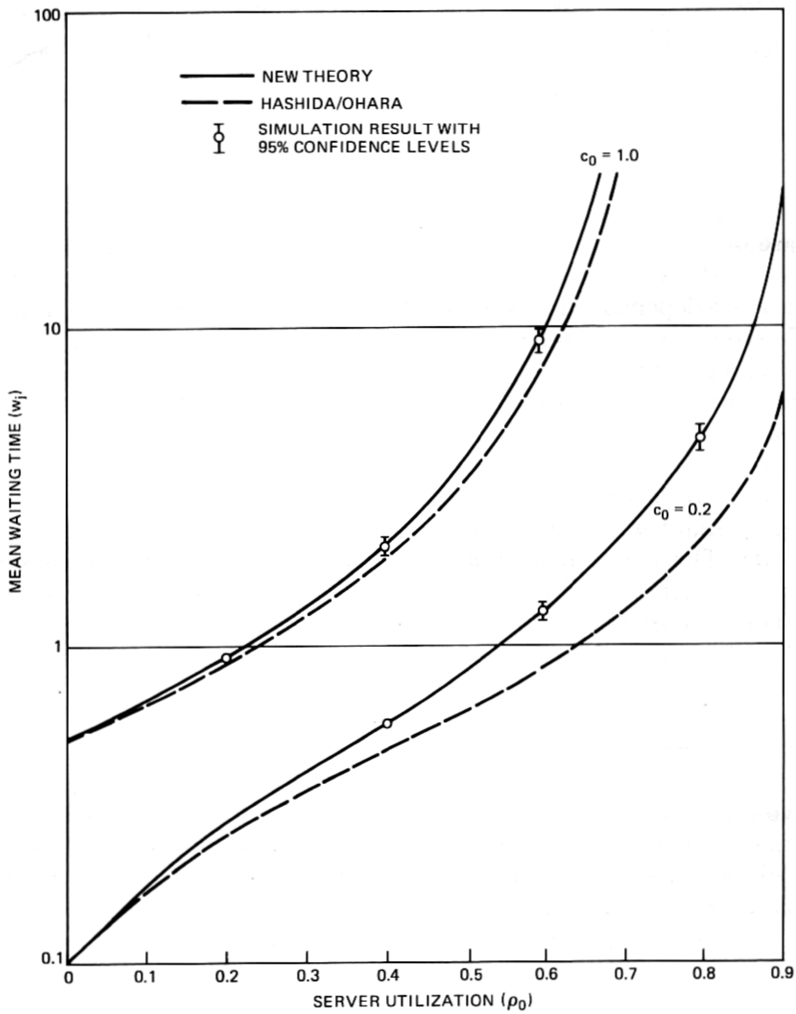

Fig. 4—Accuracy of mean waiting times for cyclic queuing systems $M/D/1$
Parameters: $g = 2$ symmetrical queues
$h_1 = h_2 = 1$ average service times
$c_0 = 0.2$ and 1.0, constant overhead.

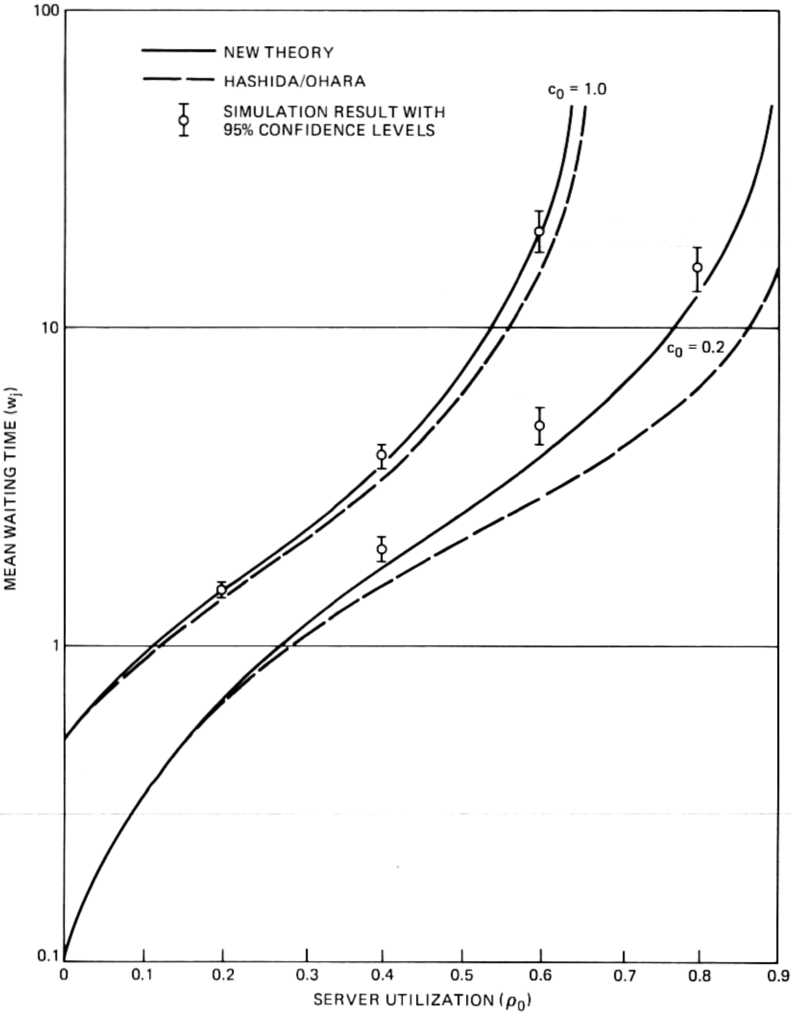

Fig. 5—Accuracy of mean waiting times for cyclic queuing systems $M/H_2/1$
Parameters: g = 2 symmetrical queues
$h_1 = h_2 = 1$ average waiting times
c_0 = 0.2 and 1.0, constant overhead
c_H = 2.0 coefficient of variation of service times.

On the contrary, batch arrivals may stabilize the cycle time since many cycles consist of one service time and the overhead only. Although these characteristics depend largely on the parameter combination, they indicate some interesting effects which may be important for applications and theory as well.

6.2 Mean waiting time

Since the mean waiting times in (26) are basically dependent on the first two moments of the conditional cycle times, we can expect the same accuracy trends as for the cycle time variance. Figures 4 and 5 show results for systems of the type $M/D/1$ and $M/H_2/1$ with two symmetrical queues for low and high overhead. The accuracy for $M/D/1$ is excellent, whereas for $M/H_2/1$ and low overhead the mean waiting time is underestimated. In any case, the new approach yields a better accuracy compared to Ref. 13, which results from the conditional cycle time concept.

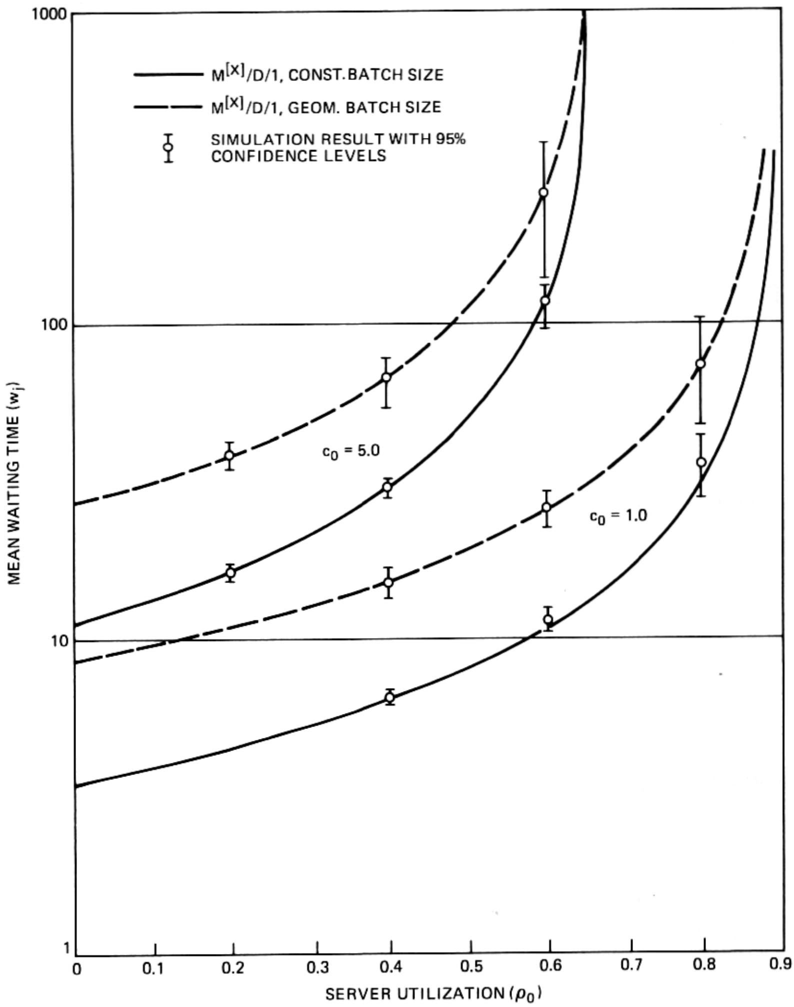

Fig. 6—Accuracy of mean waiting times for cyclic queuing systems $M^{[X]}/D/1$

Parameters: g = 10 symmetrical queues
h_j = 1 average service time, $j = 1, 2, \cdots, 10$
c_0 = 1.0 and 5.0, constant overhead
EK = 4 constant (average) batch size.

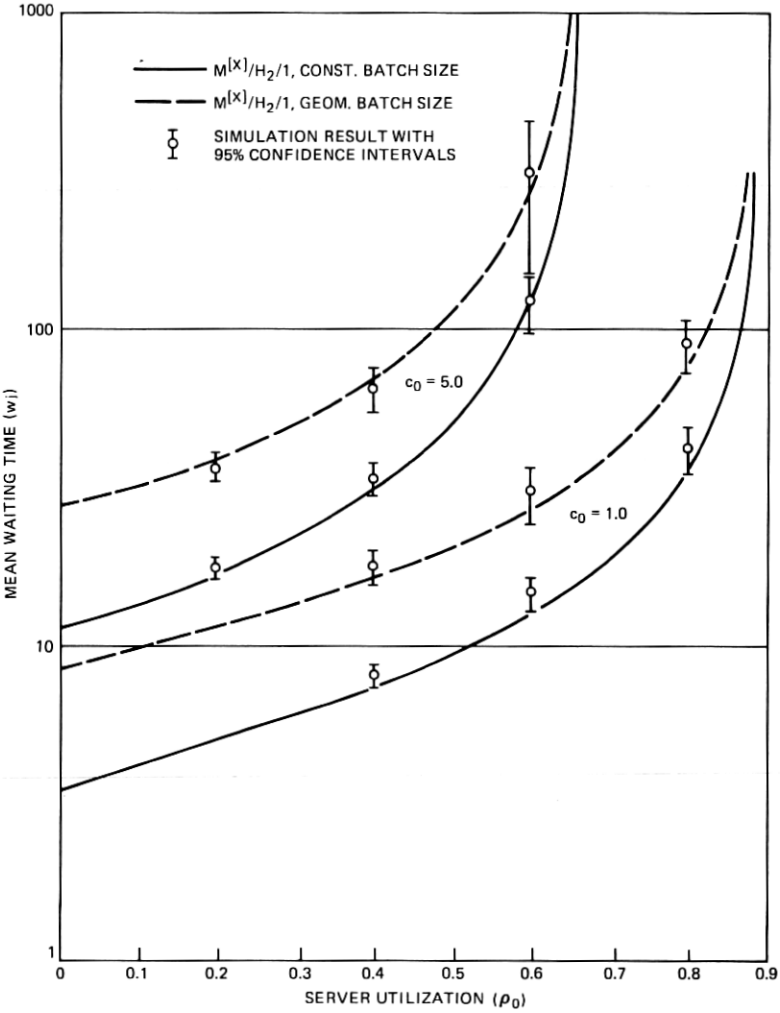

Fig. 7—Accuracy of mean waiting times for cyclic queuing systems $M^{[X]}/H_2/1$
Parameters: g = 10 symmetrical queues
h_j = 1 average service time, $j = 1, 2, \cdots, 10$
c_0 = 1.0 and 5.0, constant overhead
EK = 4 constant (average) batch size
c_H = 2.0 coefficient of variation of service times.

Figures 6 and 7 show the results for $g = 10$ symmetrical queues for systems $M^{[X]}/D/1$ (Fig. 6) and $M^{[X]}/H_2/1$ (Fig. 7), each with constant or geometrically distributed batch sizes, low and high overhead. All cases of batch arrival processes show an excellent accuracy. Many other validations have also shown that the accuracy is far less dependent on the parameters g, c_0, or G compared to single Poisson arrivals. This results from the fact that the cycle time analysis yields the best accuracy in case of batch arrivals; also, the contribution of the batch arrivals to the mean waiting time w dominates the expression (39) for larger batch sizes.

Another study on the influence of the service process type G and arrival process type GI on the mean waiting time w in case of ordinary cyclic service is shown in Figs. 8 and 9 for zero, low, and high overhead. The $M/G/1$ curves with overhead are analytic results according to (26), whereas the $GI/M/1$ curves are simulation results; the results for

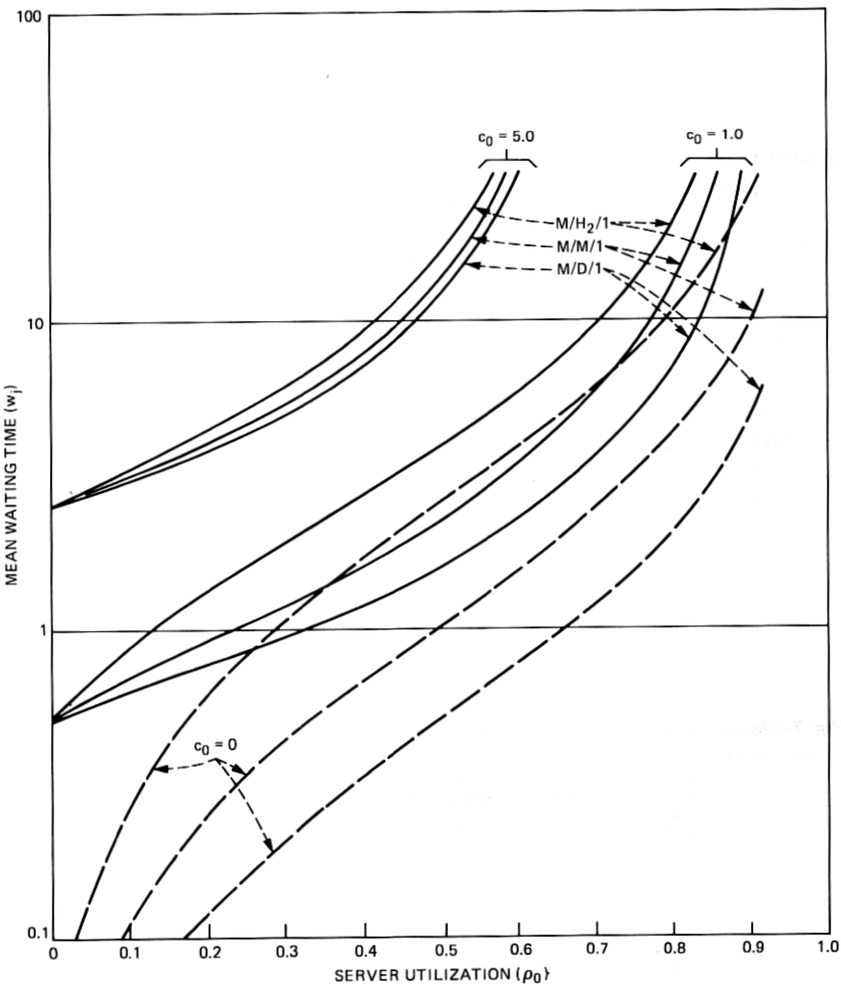

Fig. 8—Influence of service process type for cyclic queuing systems $M/G/1$
Parameters: g = 10 symmetrical queues
h_j = 1 average service time, $j = 1, 2, \cdots, 10$
c_0 = 0, 1.0, 5.0, constant overhead
Systems $M/D/1$, $M/M/1$, $M/H_2/1$ ($c_H = 2.0$).

zero overhead are exact and have been drawn from standard queuing tables by the author.[19] The main conclusions from Figs. 8 and 9 are:

(*i*) For $M/G/1$ systems with ordinary cyclic service, the influence of the service process decreases with increasing overhead.

(*ii*) For $GI/M/1$ systems with ordinary cyclic service, the influence of the arrival process does not remarkably decrease or may even increase with increasing overhead (see also Figs. 6 and 7 for batch arrivals).

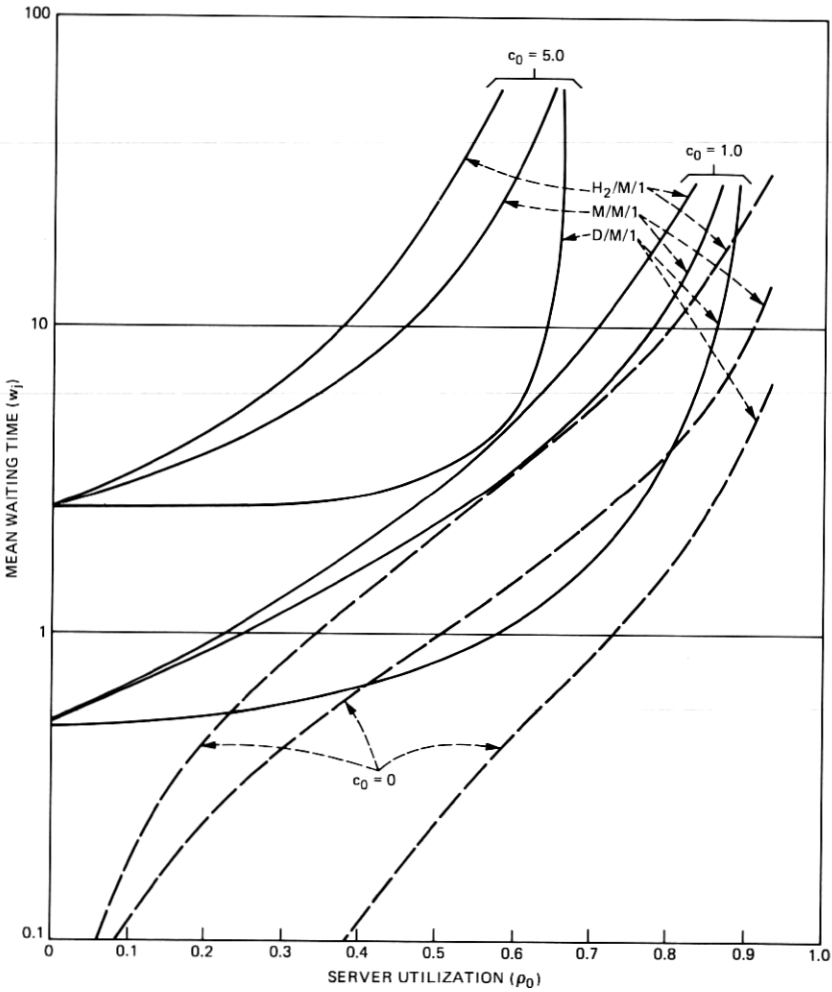

Fig. 9—Influence of arrival process type for cyclic queuing systems $GI/M/1$
Parameters g = 10 symmetrical queues
h_j = 1 average service time, $j = 1, 2, \cdots, 10$
c_0 = 0, 1.0, 5.0, constant overhead
Systems $D/M/1$, $M/M/1$, $H_2/M/1$ ($c_A = 2.0$).

These properties are most important for applications since they show that the results are much more sensitive to arrival processes than to service processes and that the usual approximation of arrival processes by Poisson processes may result in a quite dramatic error in the performance estimation. Therefore, future analytic studies on cyclic queuing systems should aim more to the generalization of arrival processes.

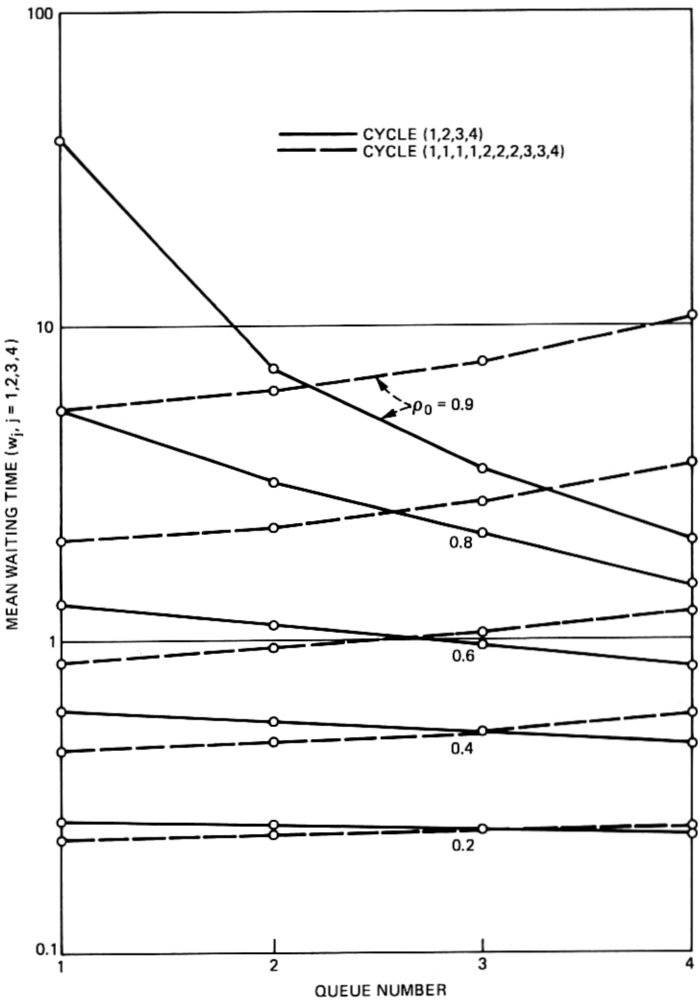

Fig. 10—Unbalanced load performance of cyclic queuing systems $M/D/1$
Parameters: $g = 4$ queues
$h_j = 1, j = 1, 2, 3, 4$, average service times
$u_j = 0.05, j = 1, 2, 3, 4$, constant overhead
$\lambda_1 : \lambda_2 : \lambda_3 : \lambda_4 = 4:3:2:1$ arrival rate ratios.

6.3 Equity of service for unbalanced load

So far, we have concentrated on cases of ordinary cyclic service. The final study shows how cyclic priority service can be used to achieve equity of service in cases of unbalanced load. In this case, queue 1 is served up to l_1 times, queue 2 up to l_2 times, $\cdots$, queue g up to l_g times within a cycle, so that $l_1, l_2, \cdots, l_g$ could be considered as cycle priorities. Figure 10 demonstrates the use of cyclic priority service in case of unbalanced load in a system $M/D/1$ with $g = 4$ queues and arrival rate ratios $\lambda_1:\lambda_2:\lambda_3:\lambda_4 = 4:3:2:1$. In case of ordinary cyclic service with the cycle $\{1, 2, 3, 4\}$, the unbalanced load produces also unbalanced waiting times with increasing absolute load ρ_0. The dashed curves show the result of cyclic priority service where $l_1:l_2:l_3:l_4 = \lambda_1:\lambda_2:\lambda_3:\lambda_4$ with the cycle $\{1, 1, 1, 1, 2, 2, 2, 3, 3, 4\}$. For small ρ_0, both schedules do not remarkably differ in performance. In case of higher ρ_0, the unbalanced load effects can be compensated for by a cyclic priority service.

VII. CONCLUSION

This paper provides a new approximate analysis for cyclic queuing systems $M^{[X]}/G/1$ with batch Poisson arrivals, general service and overhead times, and ordinary cyclic service. The method allows a relatively easy evaluation of numerical results. The accuracy of the method has been validated by computer simulations. In addition to the analysis method, a new stability criterion for systems $GI/G/1$ with general cyclic service is developed. A number of traffic studies are reported revealing more insight in the traffic performance of cyclic queuing systems.

VIII. ACKNOWLEDGMENTS

The author wishes to thank J. M. Holtzman and H. Zucker for many stimulating discussions. The constructive comments by two reviewers are greatly appreciated.

REFERENCES

1. B. Avi-Itzhak, W. L. Maxwell, and L. W. Miller, "Queuing With Alternating Priorities," Opns. Res., *13* (1965), pp. 306–318.
2. L. Takács, "Two Queues Attended by a Single Server," Opns. Res., *16* (1968), pp. 639–650.
3. R. B. Cooper and G. Murray, "Queues Served in Cyclic Order," B.S.T.J., *48,* No. 3 (March 1969), pp. 675–689.
4. R. B. Cooper, "Queues Served in Cyclic Order: Waiting Times," B.S.T.J., *49,* No. 3 (March 1970), pp. 399–413.
5. J. S. Sykes, "Simplified Analysis of an Alternating-Priority Queuing Model with Setup Times," Opns. Res., *18* (1970), pp. 1182–1192.
6. M. Eisenberg, "Two Queues with Changeover Times" Opns. Res., *19* (1971), pp. 386–401.
7. M. Eisenberg, "Queues with Periodic Service and Changeover Time," Opns. Res., *20* (1972), pp. 440–451.

8. O. Hashida, "Analysis of Multiqueue," Review of the Electr. Comm. Laboratories, Nippon Telegraph and Telephone Public Corp., *20* (1972), pp. 189–199.
9. T. S. Brodetskaya, "Capacity-Time Analysis of a Cyclic-Service Queuing System." Kibernetica, No. 2 (1975), pp. 64–68 (in Russian). Translation: New York: Plenum, 1976.
10. S. S. Nair, "Alternating Priority Queues with Non-Zero Switch Rule," Comp. and Opns. Res., *3* (1976), pp. 337–346.
11. M. Eisenberg, "Two Queues With Alternating Service," unpublished work.
12. M. A. Leibowitz, "An Approximate Method for Treating a Class of Multiqueue Problems," IBM Journal, *5* (1961), pp. 204–209.
13. O. Hashida and K. Ohara,, "Line Accomodation Capacity of a Communication Control Unit," Review of the Electr. Comm. Laboratories, Nippon Telegraph and Telephone Public Corporation, *20* (1972), pp. 231–239.
14. S. Ekberg, "Strategies for a Telephone Central Marker Handling Different Queuing Jobs," Proceedings TIMS XX, Tel Aviv (1973), pp. 523–528.
15. P. Kuehn and M. Langenbach-Belz, "Effectiveness of Device Servicing Strategies in Real Time Systems," Fourth Annual Conference of the German Informatics Society, Berlin, October 9–12, 1974, Springer-Verlag, 1975, pp. 463–472. Bell Laboratories Translation Series No. TR 77-72.
16. S. Halfin, "An Approximate Method for Calculating Delays for a Family of Cyclic-Type Queues," B.S.T.J., *54,* No. 10 (December 1975), pp. 1733–1754.
17. R. B. Cooper, *Introduction to Queuing Theory,* New York: MacMillan, 1972.
18. D. Jagerman, "An Inversion Technique for the Laplace Transform with Application to Approximation," B.S.T.J., *57,* No. 3 (March 1978), pp. 669–710.
19. P. J. Kuehn, "Tables on Delay Systems," Institute of Switching and Data Technics, University of Stuttgart, 1976.
20. P. J Burke, "Delays in Single-Server Queues with Batch Input," Opns. Res., *23* (1975), pp. 830–833.