LIMITS FOR QUEUES AS THE WAITING ROOM GROWS

by

Daniel P. Heyman
Bell Communications Research
Red Bank, NJ 07701

Ward Whitt
AT&T Bell Laboratories
Murray Hill, NJ 07974

May 11, 1988

ABSTRACT

We study the convergence of finite-capacity open queueing systems to their infinite-capacity counterparts as the capacity increases. Convergence of the transient behavior is easily established in great generality provided that the finite-capacity system can be identified with the infinite-capacity system up to the first time that the capacity is exceeded. Convergence of steady-state distribution is more difficult; it is established here for single-facility models such as GI/GI/c/n with c servers, $n-c$ extra waiting space and the first-come first-served discipline, in which all arrivals finding the waiting room full are lost without affecting future arrivals, via stability properties of generalized semi-Markov processes.

1. Introduction

Consider an open queueing system with capacity n. When n is very large, we expect that the standard descriptive stochastic processes, such as the number of customers in the system at time t for $t \geq 0$, and their limiting steady-state distributions are very close to their counterparts in the same system with infinite capacity. Indeed, for simple models such as the M/M/c/n queue (c servers and $n - c$ extra waiting spaces) for which the steady-state distributions can be displayed explicitly, convergence of the steady-state distribution as $n \to \infty$ is easily verified (provided that the infinite-capacity model is stable). We establish convergence results here that do not depend on explicit expressions for the quantities of interest. We also give examples to show that some care is needed.

In Section 2 we establish very strong convergence (total variation) in great generality for the stochastic processes representing the transient behavior, provided that we can represent the finite-capacity system up to the first time that the capacity is exceeded in terms of the infinite-capacity system. The real difficulty is obtaining convergence of the limiting steady-state distributions. In Section 4 we obtain a convergence result for the steady-state distributions of single facilities with c servers working in parallel and the first-come first-served (FCFS) queue discipline, in which all arrivals finding the system full are lost without affecting future arrivals. In addition to being of interest for its own sake, this result illustrates how similar results can be obtained for other models. Our approach is to apply continuity or stability results for Markov processes and generalized semi-Markov processes (GSMPs), as in Whitt (1980b). In particular, we apply Lemma 1 there. In order to establish Theorems 2 and 3 of Whitt (1980b), the GSMPs were assumed to have finite state space. Since the GSMPs here do not have finite state space, we obtain corresponding results here via a stochastic order bound (Corollary 2 to Theorem 1).

There is considerable related literature. The limits here express a form of model stability, continuity or robustness; see Chapter 3 of Franken, König, Arndt and Schmidt (1981), Chapter 8 of Stoyan (1983), Chapter 4 of Borovkov (1984), Brandt and Lisek (1983), and Whitt (1980a, b). Perhaps more closely related is the literature about approximating countable-state Markov chains by finite-state Markov chains; see Seneta (1981) and references there. In that context, the desired conclusion is that the steady-state distributions in the finite-state chains converge to the steady-state distribution of the infinite-state chain as the size of the state space grows. The classical paper by Ledermann and Reuter (1954) establishes limits of this kind for birth-and-death processes. While there are several results of this type, there seem to be many open problems. Somewhat related, at least in spirit, is the convergence of a sequence of closed Markovian queueing networks to a related open queueing network as the population grows; see Sections 5 and 8 of Whitt (1984).

2. Convergence of Transient Behavior

Let the stochastic process $[X_\infty , Y_\infty] \equiv \{ [X_\infty (t) , Y_\infty (t)] : t \geq 0 \}$ describe an open queueing system with infinite capacity. (In this section the queueing system can be very general, e.g., a multi-class open queueing network.) The random variable $X_\infty (t)$ represents the number of customers in the system at time t, and the random variable $Y_\infty (t)$ represents other aspects of interest at time t, such as residual interarrival times and service times. The random variable $Y_\infty (t)$ might contain supplementary variables to make $[X_\infty , Y_\infty]$ a Markov process, but need not. We assume that $[X_\infty (t) , Y_\infty (t)]$ takes values in a complete separable metric space (which usually would be Euclidean space, but need not be) and that the sample paths of $[X_\infty , Y_\infty]$ are RCLL (right-continuous with left limits).

Our most important assumption concerns the way the finite-capacity systems are related to the infinite-capacity system. *We assume that the system with capacity n can be constructed in terms of the infinite-capacity system up to the first time that the population exceeds n.* Let $T_{\infty n}$ be the first passage time defined by

$$T_{\infty n} = \inf \{ t \geq 0 : X_\infty(t) > n \} , \quad n \geq 1 . \tag{2.1}$$

We assume that $\{ [X_\infty(t), Y_\infty(t)] : 0 \leq t < T_{\infty n} \}$ is a representation (has the same probability distribution) of the corresponding stochastic process associated with the n-capacity system, say $[X_n, Y_n] \equiv \{ [X_n(t), Y_n(t)] : 0 \leq t < T_{nn} \}$, where T_{nn} is the first time that the capacity would have been exceeded. So far, we have said nothing about $[X_n(t), Y_n(t)]$ for $t \geq T_{nn}$. However, the representation assumption implies that the transient behavior of $[X_n, Y_n]$ converges to the transient behavior of $[X_\infty, Y_\infty]$ in a very strong sense.

The RCLL property implies that for each positive t

$$\sup_{0 \leq s \leq t} \{ X_\infty(s) \} < \infty \quad \text{w.p.1} , \tag{2.2}$$

p. 110 of Billingsley (1968) and p. 70 of Whitt (1980a), which in turn guarantees that

$$\lim_{n \to \infty} T_{\infty n} = \infty \quad \text{w.p.1} . \tag{2.3}$$

For the special construction above, w.p.1, for all $t > 0$ there exists n_0 such that

$$[X_n(s), Y_n(s)] = [X_\infty(s), Y_\infty(s)] , \quad 0 \leq s \leq t , \tag{2.4}$$

for all $n \geq n_0$. (Of course, n_0 depends on the sample path as well as t.) This implies convergence in total variation for the probability distributions, i.e., for each $t > 0$,

$$\lim_{n \to \infty} \sup_A \left| P([X_n, Y_n] \in \pi_t^{-1}(A)) - P([X_\infty, Y_\infty] \in \pi_t^{-1}(A)) \right| = 0 , \tag{2.5}$$

where A is a measurable subset of the function space $D[0, t]$ of RCLL functions on $[0, t]$ and π_t is the projection map from the function space $D[0, \infty)$ onto $D[0, t]$, defined by $\pi_t(x)(s) = x(s)$, $0 \leq s \leq t$. (See Chapter 3 of Billingsley (1968), Section 2 of Whitt (1980a), and Chapter 3 of Ethier and Kurtz (1986).) An elementary consequence of (2.5) is that

$$\lim_{n \to \infty} P(X_n(t) = k) = P(X_\infty(t) = k) \tag{2.6}$$

for all t and k. Another consequence is

$$[X_n, Y_n] \Rightarrow [X_\infty, Y_\infty] \quad \text{as} \quad n \to \infty \quad \text{in} \quad D[0, \infty) , \tag{2.7}$$

where $\Rightarrow$ denotes convergence in distribution (weak convergence) with the Skorohod J_1 topology on $D[0, \infty)$; i.e.,

$$\lim_{n \to \infty} E\, f([X_n, Y_n]) = E\, f([X_\infty, Y_\infty]) \tag{2.8}$$

for all continuous bounded real-valued functions on $D[0, \infty)$.

While we have established convergence of the processes (transient behavior) as $n \to \infty$ in great generality, we have yet to treat stationary or limiting distributions. The following example illustrates some of the difficulties.

Example 2.1. Let the infinite capacity system be a simple M/M/1 queue with traffic intensity $\rho < 1$. Let the associated n-capacity system be the modification in which the system closes down, i.e., empties immediately with all service times set equal to 0, the instant an arriving customer finds n customers in queue. Clearly the representation assumption above holds, so that (2.2)-(2.8) hold, but the limiting distributions do not converge. ■

3. General Multi-Server FCFS Loss Systems

Now assume that the infinite-capacity queueing system is a single facility with unlimited waiting space, c servers working in parallel and the FCFS queue discipline. Let the stochastic behavior be specified by a sequence $\{(u_k, v_k) : k \geq 1\}$ of ordered pairs of nonnegative random variables, where u_k represents the interarrival time between the $(k-1)^{\text{st}}$ and k^{th} arrival, and v_k is the service time of the k^{th} arrival, but we make *no* independence or common-distribution assumptions. Hence, we have an A/A/c/∞ model (A for arbitrary, instead of G for general stationary or GI for renewal). Let the system start out empty at time 0. Other initial conditions can be introduced through the basic sequence $\{(u_k, v_k)\}$; i.e., if $u_k = 0$ for $1 \le k \le K$ with $u_{K+1} > 0$, then there are K customers in the system at time 0.

Let the capacity-n system be defined in terms of the infinite capacity system by letting $n - c$ be the size of the waiting room and stipulating that arrivals that find the waiting room full are lost without affecting future arrivals or service times (an A/A/c/n system). We use the same sequence $\{(u_k, v_k)\}$, so that we represent the capacity-n system in terms of the infinite-capacity system as assumed in Section 2. Hence the limits described there apply here. Now we want to say what happens *after* T_{nn}. With the additional structure in this section, we can conclude that the infinite-capacity system serves as a stochastic bound for the finite-capacity systems. We use the fact that all customers admitted to the capacity-n system have the same service times as their counterparts in the infinite-capacity system.

Theorem 1. *With the special construction,* $X_n(t) \le X_\infty(t)$ *for all* n *and* t.

Proof. For any sample path, we can represent $\{X_n(t) : t \ge 0\}$ in terms of $\{X_\infty(t) : t \ge 0\}$ in two steps: first, by replacing the service times of all customers that would be lost by 0 and, second, by not counting these customers. Since customers with 0 service times do not affect the time in system of other customers, the second step is clearly consistent with the claim. For the first step, we use known monotonicity properties for A/A/c/∞ systems, in particular, Theorem 8 and the following remark in Whitt (1981): Making the service times smaller can only reduce $X_\infty(t)$. To see this directly, recall that

$$D_n = U_n + W_n + v_n , \tag{3.1}$$

where D_n is the departure epoch, $U_n = u_1 + \cdots + u_n$ is the arrival epoch, and W_n the waiting time before beginning service of the n^{th} arrival. Decreasing some of the service times v_n causes D_n to decrease or remain the same, because U_n is unchanged and W_n was shown to be a nondecreasing function of $(v_1, \ldots, v_{n-1})$ by Kiefer and Wolfowitz (1955). Since all arrival epochs are the same and all departure epochs are ordered, all queue lengths are ordered. ■

Recall that a set S of probability measures on a complete separable metric space is tight if for each $\varepsilon > 0$ there exists a compact subset K such that $P(K) > 1 - \varepsilon$ for all P in S; pp. 9, 37 of Billingsley. Tightness guarantees that every sequence from S has a weak convergent subsequence (with a proper limit). In the terminology of Section 1 of Heyman and Whitt (1984), we give conditions for the stochastic process $\{X_n(t) : t \geq 0\}$ and the family of stochastic processes $\{X_n(t) : t \geq 0, n \geq 0\}$ to be *strongly stable*.

Corollary 1. *If $\{X_\infty(t) : t \geq 0\}$ is tight, then $\{X_n(t) : t \geq 0, n \geq 0\}$ is tight, so that every sequence $\{X_{n_k}(t_k) : k \geq 0\}$ has a weak convergent subsequence.*

We say that one real-valued random variable X_1 is stochastically less than or equal to another, and write $X_1 \leq_{st} X_2$, if $P(X_1 > t) \leq P(X_2 > t)$ for all t; see Chapter 1 of Stoyan (1983).

Corollary 2. *(a) If $X_\infty(t) \Rightarrow X_\infty(\infty)$ and $X_n(t) \Rightarrow X_n(\infty)$ as $t \to \infty$, then*

$$X_n(\infty) \leq_{st} X_\infty(\infty) .$$

(b) If (a) holds for all n, then $\{X_n(\infty) : n \geq 1\}$ is tight. (Every subsequence has a weakly convergent sub-subsequence.)

It is of course also of interest to compare the finite-capacity systems for different capacity sizes. We would like to conclude that $X_{n_1}(t) \leq X_{n_2}(t)$ when $n_1 < n_2 < \infty$ with the special construction, but this is not true in general, as we show below. However, it is possible to show that the epoch of the k^{th} admitted arrival and the k^{th} departure (not the departure epoch of the k^{th} arrival) occur sooner in the system with larger capacity. This is verified by a minor modification of Theorem 1 of Sonderman (1979).

Example 3.1. To see that we need not have $X_1(t) \leq X_2(t)$ for all t or

$$\bar{X}_1 \equiv \lim_{t \to \infty} t^{-1} \int_0^t X_1(s)\,ds \leq \lim_{t \to \infty} t^{-1} \int_0^t X_2(s)\,ds \equiv \bar{X}_2 , \tag{3.2}$$

consider a D/A/1/∞ model in which $u_k = 1$, $v_{2k} = 2$ and $v_{2k+1} = \varepsilon$ for all k and a small positive ε ($0 < \varepsilon < 1/2$). For the D/A/1/1 model constructed from it, $X_1(t) = 1$ for all $t \geq 2$, so that $\bar{X}_1 = 1$ in (3.2). For the D/A/1/2 model, $X_2(t) = 2$ for $3 \leq t < 4$; $X_2(t) = 1$ for $1 \leq t < 1+\varepsilon$, $2 \leq t < 3$ and $4 \leq t < 4+\varepsilon$; $X_2(t) = 0$ for $0 \leq t < 1$, $1+\varepsilon \leq t < 2$ and $4 + \varepsilon \leq t < 5$. Moreover, the form of $X_2(t)$ in the interval $[1+4k,\ 5+4k]$ is independent of k, so that $\bar{X}_2 = (3+2\varepsilon)/4$ in (3.2). For $\varepsilon < 1/2$, $\bar{X}_2 < \bar{X}_1$.

As given, the D/A/1/∞ model above is not stable, but it can easily be made stable by inserting periodic blocks of 0 service times. This would reduce both $\bar{X}_1$ and $\bar{X}_2$, but leave $\bar{X}_2 < \bar{X}_1$. Moreover, as constructed, the processes $X_n(t)$ do not converge in distribution as $t \to \infty$. To obtain such convergence, we can perturb the model above slightly. In particular, consider the GI/A/1/∞ model in which u_k is uniformly distributed in $[1-\delta,\ 1+\delta]$ for very small δ and

$$P(v_{2k} = 2 \ \text{ and } \ v_{2k+1} = \varepsilon \ \text{ for all } \ k) = P(v_{2k} = \varepsilon \ \text{ and } \ v_{2k+1} = 2 \ \text{ for all } \ k) = 1/2 .$$

Then $X_1(t) \Rightarrow X_1(\infty)$ and $X_2(t) \Rightarrow X_2(\infty)$ with $E[X_1(\infty)] > E[X_2(\infty)]$. ■

4. Multiple GI/GI Classes with Heterogeneous Servers

In this section we introduce additional structure to the A/A/c/n loss models of Section 3 in order to establish convergence of the steady-state distributions as $n \to \infty$. The essential idea is to represent the loss model as a generalized semi-Markov process (GSMP) and apply Lemma 1 of Whitt (1980b), which we state below. To be concrete, we specify a specific model, which is a generalization of the classical GI/GI/c/n loss model.

In particular, we assume that arrivals come from m independent sources, with the interarrival times and service times associated with each source coming from independent sequences of i.i.d. (independent and identically distributed) random variables. The overall arrival process thus is a superposition of renewal processes. As before, let customers be assigned to servers according to the FCFS discipline. Let server j work at a constant rate r_j, which also may differ from server to server. Let the mean service requirement and mean interarrival time for source i be τ_i and λ_i^{-1}, respectively, which we assume are finite and positive. Let

$$\alpha = \sum_{i=1}^{m} \lambda_i \tau_i \tag{4.1}$$

represent the total *offered load*. Let

$$\beta = \sum_{j=1}^{c} r_j \tag{4.2}$$

be the *maximum long-run processing rate* and let

$$\rho = \alpha / \beta \tag{4.3}$$

be the *traffic intensity*.

The stability condition for the infinite-capacity system is obviously $\rho < 1$, but we do not know general sufficient conditions for convergence of $X_\infty(t)$ to a proper limit as $t \to \infty$, so we will assume convergence. (For the single-class model, i.e., the standard GI/GI/c/∞ model, general sufficient conditions are given in Theorem 2.3(a) of Whitt (1972) and Corollary 2.8, p. 252, of Asmussen (1987).) In general, we make a conjecture.

Conjecture. *If* $\rho < 1$ *and the* m *interarrival-time distribution are all non-lattice, then* $X_\infty(t) \Rightarrow X_\infty(\infty)$ *as* $t \to \infty$ *where* $X_\infty(\infty)$ *is proper.*

With this structure, we can represent the A/A/c/n models, with $n \le \infty$, as GSMPs as in Whitt (1980b). The discrete state of the GSMP, say $V_n(t)$, consists of two components: $X_n(t)$, the number of customers in the system, and a random vector of source indices of all customers in the queue (in order of arrival, which coincides with the order in which they will enter service). (Since the number of different states is countable, the state can be coded as a positive integer, as in Whitt (1980b).)

To complete the GSMP, let $Y_n(t) \equiv [Y_{n1}(t), \ldots, Y_{n(m+c)}(t)]$ be an $(m+c)$-dimensional *vector of clock times*. For $1 \leq i \leq m$, $Y_{ni}(t)$ is the time until the next arrival from the i^{th} source at time t; for $m+1 \leq i \leq m+c$, $Y_{ni}(t)$ is the remaining service time for the customer in service at the $(i-m)^{\text{th}}$ server at time t. The process $\{[V_n(t), Y_n(t)] : t \geq 0\}$ is the associated continuous-time Markov process (CTMP) of the GSMP V_n.

In this setting we are able to establish convergence of the steady-state distributions, provided that they always exist as unique proper limits.

Theorem 2. *If $[V_n(t), Y_n(t)] \Rightarrow [V_n(\infty), Y_n(\infty)]$ as $t \to \infty$ for all proper initial conditions $[V_n(0), Y_n(0)]$ and for all n, $n \leq \infty$, then*

$$[V_n(\infty), Y_n(\infty)] \Rightarrow [V_\infty(\infty), Y_\infty(\infty)] \quad \text{as} \quad n \to \infty ,$$

so that

$$X_n(\infty) \Rightarrow X_\infty(\infty) \quad \text{as} \quad n \to \infty .$$

To prove Theorem 2, we apply Lemma 1 of Whitt (1980b), which we restate here.

Lemma 1. *If P_n is an invariant probability measure for a CTMP $\{\xi_n(t) : t \geq 0\}$ for each $n \geq 1$, and $\xi_n \Rightarrow \xi$ in $D[0, \infty)$ whenever $\xi_n(0) \Rightarrow \xi(0)$, then any weak convergent limit of a subsequence of $\{P_n\}$ is an invariant probability measure for $\{\xi(t) : t \geq 0\}$.*

To apply Lemma 1 to prove Theorem 2, we use the next two lemmas, which we prove after we prove Theorem 2.

Lemma 2. *The family of steady-state random elements $\{[V_n(\infty), Y_n(\infty)] : n \geq 1\}$ is tight.*

Lemma 3. *Let* $[V_n(0), Y_n(0)]$ *serve as (proper) initial conditions for the CTMP* $[V_n, Y_n]$ *for* $n \leq \infty$. *If* $[V_n(0), Y_n(0)] \Rightarrow [V_\infty(0), Y_\infty(0)]$ *as* $n \to \infty$, *then*

$$[V_n, Y_n] \Rightarrow [V, Y] \quad in \quad D[0, \infty) \quad as \quad n \to \infty .$$

Proof of Theorem 2. The assumption in Theorem 2 implies that $[V_n(\infty), Y_n(\infty)]$ is the unique invariant random element for the CTMP $[V_n, Y_n]$ for $n \leq \infty$. By Lemma 2 and Prohorov's theorem (p. 35 of Billingsley), the sequence $\{[V_n(\infty), Y_n(\infty)] : n \geq 1\}$ is relatively compact: every subsequence has a further subsequence converging to a proper limit. We will show that any such limit must be $[V_\infty(\infty), Y_\infty(\infty)]$, which will complete the proof; see Theorem 2.3 of Billingsley. Hence, suppose $[V_{n'}(\infty), Y_{n'}(\infty)] \Rightarrow [V, Y]$ for some limit $[V, Y]$ and some subsequence with index n'. By Lemmas 1 and 3, $[V, Y]$ must be an invariant random element of $[V_\infty \; Y_\infty]$. Hence, $[V, Y] = [V_\infty(\infty), Y_\infty(\infty)]$. ■

Proof of Lemma 2. To establish tightness, it suffices to show that the marginal components $\{V_n(\infty) : n \geq 1\}$ and $\{Y_{nj}(\infty) : n \geq 1\}$, $1 \leq j \leq m + c$, are tight separately, e.g., Problem 6, p. 41, of Billingsley. For $\{V_n(\infty)\}$, it suffices to show that for any $\varepsilon > 0$ there exists k such that

$$P(X_n(\infty) > k) \leq \varepsilon \quad \text{for all} \quad n ,$$

but this holds because $X_n(\infty)$ is stochastically bounded by $X_\infty(\infty)$ as a consequence of Corollary 2(a) to Theorem 1 here, i.e.,

$$P(X_n(\infty) > k) \leq P(X_\infty(\infty) > k) \quad \text{for all} \quad n$$

and any individual random element is tight; Theorem 1.4 of Billingsley.

The clock times $\{Y_{nj}(\infty)\}$ can be dominated by equilibrium residual lifetime variables associated with the interarrival times and service times, in a manner similar to the proof of Theorem 2 of Whitt (1980b). (The proof there implicitly assumes that events never cease to be scheduled before the clock runs down and the deterministic speed of clock j in state s, is independent of s, both of which hold here.) Since the component arrival processes are renewal processes, for $j \le m$ the steady-state clock times $Y_{nj}(\infty)$ is distributed as the equilibrium life variable associated with the interarrival time from source j, independent of n, which is proper because the interarrival times have finite positive means. For the service times $(j > m)$,

$$P(Y_{nj}(\infty) > k) = \lim_{t \to \infty} t^{-1} \int_0^t P(Y_{nj}(u) > k)\, du \,, \tag{4.4}$$

but the long-run average on the right in (4.4) is made only larger by deleting idle periods of the server. The resulting average with all idle periods deleted coincides with the tail of the equilibrium life variable associated with the service time in question, independent of n, which is proper because the service times have finite positive mean. ■

Proof of Lemma 3. Section 2 establishes Lemma 3 for the special in which $[Y_n(0), Y_n(0)] = [V_\infty(0), Y_\infty(0)]$ for all n. To treat these more general initial conditions, let $[V_\infty^n, Y_\infty^n]$ be the CTMP $[V_\infty, Y_\infty]$ initialized by $[V_n(0), Y_n(0)]$, $n \ge 1$. As in Section 2, we can construct $[V_n, Y_n]$ in terms of $[V_\infty^n, Y_\infty^n]$ up until the random time

$$T_n^n = \inf \{t \ge 0 : X_\infty^n(t) > n\} .$$

Since $[V_n(0), Y_n(0)]$ converges to a proper limit, it is evident that T_n^n converges to ∞ in probability, i.e.,

$$d([V_n, Y_n], [V_\infty^n, Y_\infty^n]) \xrightarrow{p} 0 ,$$

where d is a metric inducing the usual topology on $D[0, \infty)$. Hence, by Theorem 4.1 of Billingsley, it suffices to show that $[V_\infty^n, Y_\infty^n] \Rightarrow [V_\infty, Y_\infty]$, i.e., to show that the single CTMP $[V_\infty, Y_\infty]$ is continuous in the initial conditions, but such continuity is established in Theorem 1 in Whitt (1980b). ■

REFERENCES

[1] ASMUSSEN, S. 1987. *Applied Probability and Queues,* Wiley, New York.

[2] BILLINGSLEY, P. 1968. *Convergence of Probability Measures,* Wiley, New York.

[3] BOROVKOV, A. A. 1984. *Asymptotic Methods in Queueing Theory,* Wiley, Chichester.

[4] BRANDT, A. and B. LISEK. 1983. On the continuity of G/GI/m queues. *Math Operationsforsch. Statist.*, Ser. Statist. 12, 577-587.

[5] ETHIER, S. N. and T. G. KURTZ. 1986. *Markov Processes, Characterization and Convergence,* Wiley, New York.

[6] HEYMAN, D. P. and W. WHITT. 1984. The asymptotic behavior of queues with time-varying arrival rates. *J. Appl. Prob.* 21, 143-156.

[7] KIEFER, J. and J. WOLFOWITZ. 1955. On the theory of queues with many servers. *Trans. Amer. Math. Soc.* 78, 1-18.

[8] LEDERMANN, W. and G. E. H. REUTER. 1954. Spectral theory for the differential equations of simple birth and death processes. *Phil. Trans. Roy. Soc. London* A246, 321-369.

[9] SENETA, E. 1981. Computing the stationary distribution for infinite markov chains, in *Large Scale Matrix Problems,* A. Björck, R. J. Plemmons and H. Schneider (eds.), North Holland, New York.

[10] SONDERMAN, D. 1979. Comparing multi-server queues with finite waiting rooms, I: same number of servers. *Adv. Appl. Prob.* 11, 439-447.

[11] STOYAN, D. 1983. *Comparison Methods for Queues and Other Stochastic Models,* Wiley, Chichester.

[12] WHITT, W. 1972. Embedded renewal processes in the GI/G/s queue. *J. Appl. Prob.* 9, 650-658.

[13] WHITT, W. 1980a. Some useful functions for functional limit theorems. *Math. Opns. Res.* 5, 67-85.

[14] WHITT, W. 1980b. Continuity of generalized semi-Markov processes. *Math. Opns. Res.* 5, 494-501.

[15] WHITT, W. 1981. Comparing counting processes and queues. *Adv. Appl. Prob.* 13, 207-220.

[16] WHITT, W. 1984. Open and closed models for networks of queues. *AT&T Bell Lab. Tech. J.* 63, 1911-1979.