J. Appl. Prob. **21**, 143–156 (1984)
Printed in Israel

THE ASYMPTOTIC BEHAVIOR OF QUEUES WITH TIME-VARYING ARRIVAL RATES

DANIEL P. HEYMAN* AND
WARD WHITT,* *Bell Laboratories*

Abstract

This paper discusses the asymptotic behavior of the $M_t/G/c$ queue having a Poisson arrival process with a general deterministic intensity. Since traditional equilibrium does not always exist, other notions of asymptotic stability are introduced and investigated. For the periodic case, limit theorems are proved complementing Harrison and Lemoine (1977) and Lemoine (1981).

PERIODIC QUEUE; NON-STATIONARY QUEUE; MULTISERVER QUEUE; PERIODIC POISSON PROCESS; REGENERATIVE PROCESS; WAITING TIME

0. Introduction and summary

The purpose of this paper is to contribute to the theory of queues with time-varying arrival rates. We assume that the arrival process is a Poisson process with a general deterministic intensity $\lambda(t)$. We are interested in periodic arrival processes, which have an 'embedded stationarity' and can be represented as stationary point processes with the proper initial conditions, but we are also interested in arrival processes that are fundamentally non-stationary, that cannot be put in the framework of Franken et al. (1981).

Most of the work on queues with time-varying arrival rates has been concerned with describing the time-dependent behavior of the queue. Early papers by Luchak (1956) and Clarke (1956) focused on solving the Kolmogorov equations for the queue-length process. Since then, considerable progress has been made by developing approximations for the time-dependent behavior; see Rothkopf and Oren (1979), Clark (1981), and Taafe (1982) for closure approximations; see Newell (1968), (1971), McClish (1979), Keller (1982) and Massey (1981) for asymptotic expansions; see these sources for earlier work.

Received 18 January 1983; revision received 20 April 1983.
* Postal address: Bell Laboratories, Crawfords Corner Road, Holmdel, NJ 07733, U.S.A.

The present paper is concerned with the asymptotic behavior as time increases. Related work for the $M_t/G/1$ queue with periodic Poisson arrival process is contained in Harrison and Lemoine (1977), Lemoine (1981) and Wolff (1982), §3. The results here were reported in Heyman and Whitt (1978).

We use stochastic comparisons here (Theorem 2.1 and Section 4). Related work on stochastic comparisons for queues with time-varying arrival rates is contained in Ross (1978), Rolski (1981), Heyman (1982) and Whitt (1981).

Our time-varying arrival rates are represented via a deterministic intensity $\lambda(t)$ for a Poisson process. One could also look at the Poisson process or another arrival process in a random environment; see Chapter 6 of Neuts (1981) and references there.

This paper is organized as follows. In Section 1 we present general definitions of asymptotic stability for stochastic processes to cover cases in which the arrival process is neither stationary nor periodic. In Section 2 we give an example to show that the obvious generalization of the stability conditions in Harrison and Lemoine (1977) are not sufficient without periodicity; then we establish sufficient conditions for general asymptotic stability of the Markovian $M_t/M/c$ queue having c servers and a Poisson arrival process with a general deterministic intensity $\lambda(t)$. In Section 3 we prove limit theorems for the $M_t/M/c$ system having a periodic Poisson arrival process, and possibly a finite waiting room. The restriction to exponential service times enables us to provide relatively simple proofs. In Section 4 we briefly indicate how to construct stationary versions in the periodic case, so that Section 3 can be put in the framework of Franken et al. In Section 5 we briefly indicate how the results of Section 3 can be obtained for non-exponential service times.

1. Asymptotic stability of stochastic processes

Our starting point is the standard notion of asymptotic stability for a stochastic process $X(t)$: convergence in distribution as $t \to \infty$. In this section we introduce concepts of asymptotic stability for stochastic processes that do not converge in distribution as $t \to \infty$. There obviously are many different kinds of asymptotic stability. Our concepts are based on the one-dimensional marginal distributions. For each $t \geqq 0$, let F_t be the c.d.f. of $X(t)$. Without stationarity or periodicity, it might seem that the c.d.f.'s F_t could wander all over so that they are never near any c.d.f. infinitely often. Fortunately, this is not the case. With the right framework, the space of c.d.f.'s is a compact metric space so that any sequence has a convergent subsequence. We focus on the subset of c.d.f.'s that F_t is near infinitely often as $t \to \infty$.

An appropriate framework involves the concept of vague convergence of probability measures or c.d.f.'s; see p. 79 of Chung (1974). We consider only real-valued stochastic processes and thus only c.d.f.'s on the real line R, but the concepts extend to more general spaces; see Chapter 7 of Bauer (1972).

We allow improper c.d.f.'s, which correspond to probability measures with total mass less than 1. A c.d.f. F is *proper* if and only if $F(x) \to 1$ as $x \to \infty$. A sequence of c.d.f.'s $\{F_n\}$ *converges vaguely* to a c.d.f. F if there exists a countable dense subset D of R such that

$$\lim_{n\to\infty} [F_n(b) - F_n(a)] = F(b) - F(a) \tag{1.1}$$

for all $a, b \in D$. It is significant that the space of all c.d.f.'s with this mode of convergence can be represented as a compact metric space. Hence, every sequence has a convergent subsequence. (See Chung and Bauer for more details.)

Let $\mathscr{L} \equiv \mathscr{L}(X)$ be the set of all c.d.f.'s that arise as vague limits of sequences $\{F_{t_k}\}$ with $t_k \to \infty$ as $k \to \infty$. Since the space of all c.d.f.'s is a compact metric space, $\mathscr{L}$ is non-empty and compact. It is natural to use $\mathscr{L}$ to describe the asymptotic behavior of the stochastic process $X(t)$ as $t \to \infty$.

Definition 1.1. The stochastic process $\{X(t), t \geqq 0\}$ is said to be:

(a) *convergent* if $\mathscr{L}$ contains a single element F and F is proper;

(b) *asymptotically periodic* if there exists $T > 0$ such that $\{X(nT + s), n \geqq 0\}$ is convergent as $n \to \infty$ for each s, $0 \leqq s < T$;

(c) *strongly stable* if F is proper for all $F \in \mathscr{L}$;

(d) *weakly stable* if F is proper for some $F \in \mathscr{L}$;

(e) *divergent* if $\mathscr{L}$ contains only a single element F and $F(x) = 0$ for all x.

Obviously (a)$\to$(b)$\to$(c)$\to$(d) in Definition 1.1. We consider stochastic processes that are weakly stable but not strongly stable, strongly stable but not asymptotically periodic, and asymptotically periodic but not convergent. An interesting direction for future research is to identify conditions for various properties of $\mathscr{L}$. For example, when is $\mathscr{L}$ connected? When are all elements of $\mathscr{L}$ absolutely continuous?

We close this section by giving a criterion for strong stability. Since vague convergence of proper c.d.f.'s to a proper limit is equivalent to weak convergence, strong stability can be expressed in terms of uniform tightness (p. 80 of Chung (1974)) as follows.

Proposition 1.1. A stochastic process $\{X(t), t \geqq 0\}$ is strongly stable if and only if for all $\varepsilon > 0$ there exist t_0, a and b, $-\infty < a < b < \infty$, such that

$$F_t(b) - F_t(a) > 1 - \varepsilon$$

for all $t \geqq t_0$.

As a consequence of Proposition 1.1, we can define proper bounding c.d.f.'s for $\mathscr{L}$ when X is strongly stable. Let F_L and F_U be the c.d.f.'s defined by

$$(1.2)\qquad \begin{aligned} F_L(x) &= \sup\{F(x): F \in \mathscr{L}\}, \qquad x \in R, \\ F_U(x) &= \inf\{F(x): F \in \mathscr{L}\}, \qquad x \in R. \end{aligned}$$

Since $F_L(x) \geqq F(x) \geqq F_U(x)$ for all x, F_L and F_U are lower and upper bounds on $\mathscr{L}$ in the usual stochastic order. (See (2.5).) The pair (F_L, F_U) may give a nice summary description of the asymptotic behavior of $X(t)$ as $t \to \infty$.

2. Queues with general arrival rates

Harrison and Lemoine (1977) analyzed the $M_t/G/1$ queue with periodic Poisson arrival process and general service times. They established conditions for the workload process (virtual waiting-time process) to be asymptotically periodic in the sense of Definition 1.1. It is natural to conjecture that the workload process and the queue-length process (by which we mean the number of customers in the system) would be strongly stable even if the arrival process is not periodic provided that $\lambda(t)$ is bounded and

$$(2.1)\qquad \limsup_{t \to \infty} t^{-1} \int_0^t \lambda(s) ds < \mu,$$

where $\lambda(t)$ is the deterministic arrival rate at time t and μ is the constant service rate ($\mu = 1/Ev$, where v is the service time). However, in the following examples we show that (2.1) is not sufficient for strong stability. In fact, under (2.1) it is still possible for the average queue length in $[0, t]$ to diverge to $+\infty$ as $t \to \infty$.

Example 2.1. For simplicity, we first consider a deterministic fluid-flow queueing model, as in Oliver and Samuel (1962) and Newell (1971). Let there be a single server with constant service rate μ. Let the arrival rate $\lambda(t)$ be either b or 0, where $\mu < b$. In particular, let

$$(2.2)\qquad \lambda(t) = \begin{cases} b, & \tau_k \leqq t \leqq \tau_k + dk, \\ 0, & \text{otherwise,} \end{cases}$$

where $\tau_k = 1 + 2 + \cdots + k = k(k+1)/2$, $d < 1$ and $db < \mu$, so that

$$\bar{\lambda} \equiv \lim_{t \to \infty} t^{-1} \int_0^t \lambda(s) ds = db < \mu$$

and (2.1) is satisfied. Note that by choosing d sufficiently small, the traffic intensity $\rho = \bar{\lambda}/\mu$ can be made arbitrarily small.

Let $x(t)$ be the content or 'queue length' at time t; it grows at rate $b - \mu$ when $\lambda(t) = b$ and declines at rate μ when $\lambda(t) = 0$ and $x(t) > 0$. The deterministic processes $\lambda(t)$ and $x(t)$ are depicted in Figure 1, from which it is clear that $\lambda(t)$ is strongly stable but not asymptotically periodic and $x(t)$ is weakly stable but not strongly stable. For this deterministic flow example, $x(\tau_k) = 0$ for all k and $x(\tau_k + dk) = (b - \mu)dk \to \infty$ as $k \to \infty$, so indeed $x(t)$ is weakly stable but not strongly stable. Moreover, since the integral of $x(t)$ in the interval $(\tau_{k-1}, \tau_k]$, which is the area of the triangle, is $d^2k^2(b^2 - b\mu)/2\mu$,

$$\liminf_{t\to\infty} t^{-1} \int_0^t x(s)ds \geqq \lim_{n\to\infty} \frac{1}{\tau_n} \sum_{k=1}^{n} d^2k^2(b^2 - b\mu)/2\mu = \infty. \tag{2.3}$$

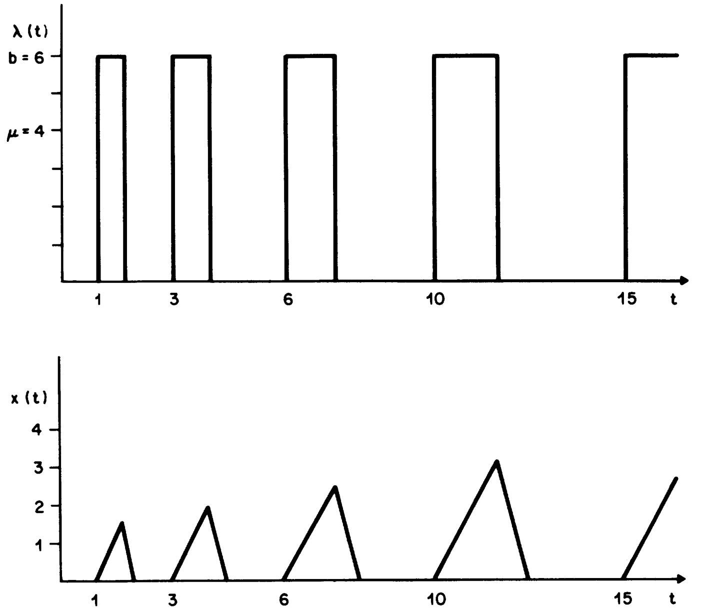

Figure 1. The arrival rate $\lambda(t)$ and the queue-length process $x(t)$ for the deterministic fluid-flow model in Example 2.1.

Example 2.2. Consider the Markovian $M_t/M/1$ queue with deterministic arrival rate $\lambda(t)$ and service rate μ as given in Example 2.1. Let $X(t)$ be the queue-length process at time t (number of customers in the system). As in Example 2.1, the overall traffic intensity $\rho = \bar{\lambda}/\mu$ can be arbitrarily small, but the time-dependent traffic intensity $\rho_t = \lambda(t)/\mu$ is strictly greater than 1 and constant during the intervals $[\tau_k, \tau_k + dk]$, $k \geqq 1$. Focusing on these intervals alone, we can apply a heavy-traffic limit theorem, Theorem 3.1 of Iglehart and Whitt (1970), to establish that

$$(dk)^{-1/2}[X(\tau_k + dk) - X(\tau_k) - (b - \mu)dk]$$

converges in distribution as $k \to \infty$ to a proper normal distribution. Hence, $X(\tau_k + dk) \to \infty$ as $k \to \infty$, so that $X(t)$ is not strongly stable. With some additional work, it is also possible to show that $X(\tau_k) \to 0$ as $k \to \infty$, so that $X(t)$ is weakly stable.

We now establish a positive result.

Theorem 2.1. The queue-length process $X(t)$ in an $M_t/M/c$ queue with general deterministic arrival rate $\lambda(t)$ is strongly stable and all c.d.f.'s in the limit set $\mathscr{L}$ have finite moments of all orders if there exist positive numbers t_0, ε and T such that

$$\int_{t_0+nT}^{t_0+(n+1)T} \lambda(s)ds \leqq (c\mu - \varepsilon)T \tag{2.4}$$

for all $n \geqq 0$.

In the following proof we use the standard stochastic order. We say that one random variable Y_1 is *stochastically less than or equal to* another random variable Y_2, and write $Y_1 \leqq_{st} Y_2$, if for all x

$$P(Y_1 > x) \leqq P(Y_2 > x). \tag{2.5}$$

Proof. We demonstrate the desired results by constructing other processes that are easier to analyze and that are stochastically greater than or equal to the given queue-length process $X(t)$. First, we show that the c-server system is appropriately dominated by a 1-server system, so that it suffices to let $c = 1$. Second, we show for the 1-server system that $X(t)$ at an arbitrary time t is appropriately dominated by $X(t_0 + nT)$ for $t_0 + nT \leqq t < t_0 + (n+1)T$, so that it suffices to consider the embedded sequence $\{X(t_0 + nT), n = 1, 2, \cdots\}$. Third, we show that the embedded sequence $\{X(t_0 + nT), n = 1, 2, \cdots\}$ is appropriately dominated by a stationary sequence associated with a periodic arrival process. Finally, we show that a minor transformation of this stationary sequence satisfies the recursive definition of the delay sequence in a $GI/G/1$ queue, so that we can complete the proof by applying known criteria for stability and finite moments of the delay in a $GI/G/1$ queue, as contained in Lemoine (1976).

For any integer c, let $X_c(t)$ be the queue-length process in an $M_t/M/c$ queue with individual service rate μ/c and a given Poisson arrival process with deterministic intensity $\lambda(t)$. Since the c-server system has the same departure rate as the 1-server system when all c servers are busy,

$$X_c(t) \leqq_{st} c + X_1(t), \qquad t \geqq 0, \tag{2.6}$$

so that it suffices to consider $c = 1$. A rigorous proof of (2.6) can be given using a construction like the one in the proof of Theorem 1 of Sonderman (1979); i.e., $X_c(t)$ is constructed on the same space with $X_1(t)$ and the two processes are given the same departure points whenever $X_c(t) \geqq c$. A related result is Theorem 4 of Yu (1974). Henceforth, we assume that $c = 1$.

Next, for every $t \in (t_0 + nT, t_0 + (n+1)T]$ and every sample point,

$$X(t) \leqq X(t_0 + nT) + A_n, \tag{2.7}$$

where A_n is the number of arrivals in the interval $(t_0 + nT, t_0 + (n+1)T]$, which is independent of $X(t_0 + nT)$. By (2.4), $A_n \leqq_{\text{st}} B$ where B is a Poisson random variable with mean $(c\mu - \varepsilon)T$. Hence, without loss of generality, we focus on the embedded sequence $\{X(t_0 + nT), n \geqq 0\}$.

We now bound $X(t)$ by the queue length $Y(t)$ in a modified system. We obtain $Y(t)$ by not letting the A_n arrivals in $(t_0 + nT, t_0 + (n+1)T]$ begin service until $t_0 + (n+1)T$. For each t and each sample point, $X(t) \leqq Y(t)$. Hence, it suffices to consider the embedded sequence $\{Y(t_0 + nT), n = 1, 2, \cdots\}$.

Now let

$$Y_n = Y(t_0 + nT) - A_{n-1}, \qquad n \geqq 1, \tag{2.8}$$

so that the sequence $\{Y_n, n \geqq 1\}$ is distributed as

$$Y_{n+1} = (Y_n + A_{n-1} - C_n)^+, \qquad n \geqq 1, \tag{2.9}$$

where $\{C_n\}$ is an i.i.d. sequence of Poisson variables with mean μT and $(x)^+ = \max\{x, 0\}$. Since $A_n \leqq_{\text{st}} B_n$, where B_n is Poisson with mean $(\mu - \varepsilon)T$, $Y_n \leqq_{\text{st}} Z_n$ for each n, where

$$Z_{n+1} = (Z_n + B_{n-1} - C_n)^+, \qquad n \geqq 1, \tag{2.10}$$

$B_{n-1} - C_n$ is independent of Z_n and $\{(B_{n-1} - C_n), n \geqq 1\}$ is i.i.d.

Finally, Z_n in (2.10) has the same structure as the waiting time of the nth customer in a $GI/G/1$ queue. Since B_n and C_n are Poisson and $E(B_n - C_n) = -\varepsilon T$, Z_n converges in distribution as $n \to \infty$ to a random variable Z with finite moments of all orders; see Lemoine (1976). Let Z'_n be Z_n with initial variable $Z'_0 = 0$. From (2.10), $Z_n \leqq_{\text{st}} Z'_n + Z_0$ for all n. Since Z'_n increases stochastically to Z as $n \to \infty$, in general $Z_n \leqq_{\text{st}} Z + Z_0$ where Z is independent of Z_0. Here $X(t_0)$ plays the role of Z_0 and $X(t_0) \leqq_{\text{st}} A(t_0)$, where $A(t_0)$ is the Poisson number of arrivals in $[0, t_0]$. Hence, Z_n is stochastically dominated for all n by a random variable with finite moments of all orders.

Remark. The distribution of Z, the limit of Z_n in (2.10), is a stochastic upper bound to F_U in (1.2).

3. Periodic arrival processes

We now consider the $M_t/M/c$ queue with periodic Poisson arrival process. We assume the waiting room is infinite, but all the results in this section hold for finite waiting rooms, for which strong stability is of course trivial. Let the length of a period be 1. Thus, the deterministic arrival rate $\lambda(t)$ satisfies $\lambda(t+n) = \lambda(t)$ for every $t > 0$ and positive integer n. The average arrival rate is given by

$$\bar{\lambda} = \int_0^1 \lambda(t)dt.$$

We assume that $\bar{\lambda} < c\mu$. Theorem 2.1 implies that $X(t)$ is strongly stable, but now we can get more.

Much about $X(t)$ can be deduced from the embedded process X_n defined by

(3.1) $$X_n = X(s+n), \qquad n = 0, 1, \cdots$$

for any fixed $s \in [0, 1)$.

Lemma 3.1. The embedded sequence $\{X_n\}$ in (3.1) is an irreducible aperiodic positive-recurrent Markov chain with stationary transition probabilities.

Proof. Clearly X_n is a Markov chain with stationary transition probabilities satisfying

$$P(X_{n+1} = j \mid X_n = i) > 0$$

for $j = i+1, i$ and $i-1$ with $i, j \geqq 0$. Consequently, X_n is irreducible and aperiodic. Hence, X_n converges in distribution to a possibly improper limit as $n \to \infty$; p. 389 of Feller (1968). However, since $X(t)$ is strongly stable by Theorem 2.1, X_n is strongly stable, so that the limit must be proper. By Feller, X_n is positive recurrent.

As an immediate consequence of Lemma 3.1, we obtain the following results about $X(t)$.

Theorem 3.1. The queue-length process $X(t)$ is asymptotically periodic with period 1.

Define epochs ξ_k by

(3.2) $$\begin{aligned} \xi_0 &= \inf\{n : X(n) = 0,\ n = 0, 1, \cdots\} \\ \xi_{k+1} &= \inf\{n : X(n) = 0,\ n > \xi_k\}, \qquad k \geqq 1. \end{aligned}$$

Theorem 3.2. The process $X(t)$ is regenerative with regeneration epochs ξ_k in (3.2) satisfying $E\xi_k < \infty$, $k \geqq 0$.

Proof. Set $s = 0$ in the definition of X_n in (3.1). Visits to state 0 (or any other state) for the Markov chain X_n are regeneration points for $X(t)$. Since the Markov chain is positive recurrent, all first-passage times have finite mean.

Lemma 3.1 and Theorem 3.2 are of interest in themselves, e.g., to obtain expressions for the limiting behavior in terms of cycles and to justify the regenerative method of simulation (Crane and Lemoine (1977)), but they also can be used to establish analogs of Theorems 3.1 and 3.2 for related continuous-time processes such as the amount of work (unexpired service time) in the system at time t and the number of customers in the queue (excluding customers in service).

We now apply Theorem 3.2 to obtain stronger asymptotic stability results for the discrete-time processes describing the system as seen by arrivals. Let T_n be the epoch of the nth arrival; let Q_n be the queue length (number in system) seen by the nth arrival; and let D_n be the delay of the nth customer before entering service. The random variable Q_n can be defined as

$$Q_n = X(T_n -), \qquad n \geqq 1. \tag{3.3}$$

For each n, the random variable D_n is distributed as ($\stackrel{d}{=}$)

$$D_n \stackrel{d}{=} U_1 + \cdots + U_{(Q_n - c+1)^+}, \qquad n \geqq 1, \tag{3.4}$$

where $\{U_n\}$ is an i.i.d. sequence of exponential random variables with mean $(c\mu)^{-1}$.

It is important to note that, in general, $\{X(T_n), n \geqq 1\}$ need not be strongly stable when $X(t)$ is strongly stable and $T_n \to \infty$ because the T_n are random. However, we get this and even more.

Theorem 3.3. The processes Q_n and D_n in (3.3) and (3.4) are regenerative with finite expected regeneration cycles and are convergent.

Proof. Let η_k be the index of the first customer to arrive after ξ_k, defined in (3.2). Clearly, the η_k are regeneration points for Q_n and D_n. Moreover, the regeneration epochs do not occur on a periodic discrete set because clearly $P(\eta_{k+1} - \eta_k = 1) > 0$. It remains to show that $E\eta_k < \infty$. Let $A(t)$ be the number of arrivals in $[0, t]$. By the law of large numbers, $n^{-1}A(n) \to \bar{\lambda}$ as $n \to \infty$, so that

$$\frac{A(\xi_k)}{k} = \frac{A(\xi_k)}{\xi_k} \cdot \frac{\xi_k}{k} \to \bar{\lambda} E(\xi_1 - \xi_0) < \infty$$

as $k \to \infty$. On the other hand, by the law of large numbers again, with summands which may have infinite mean, Theorem 5.4.2 of Chung (1974),

$$\frac{A(\xi_k)}{k} = k^{-1} \sum_{j=1}^{k} [A(\xi_j) - A(\xi_{j-1})] \to E[A(\xi_1) - A(\xi_0)]$$

as $k \to \infty$. Since $\eta_k = A(\xi_k) + 1$,

$$E(\eta_k - \eta_{k-1}) = E(A(\xi_k) - A(\xi_{k-1})) = \bar{\lambda} E(\xi_1 - \xi_0) < \infty.$$

Let D^*, Q^* and X_s^* be random variables having the limiting distribution of D_n, Q_n and $X(s+n)$ as $n \to \infty$, respectively, which exist by Theorems 3.3 and 3.1. By §3 of Wolff (1982), we have the following relation between the distributions of Q^* and X_s^*.

Theorem 3.4. The distributions of Q^* and X_s^* are related by

$$P(Q^* \leqq x) = \bar{\lambda}^{-1} \int_0^1 P(X_s^* \leqq x)\lambda(s)ds. \tag{3.5}$$

Theorem 3.5. The limiting random variables X_s^*, Q^* and D^* have finite means of all orders. The distribution of D^* and Q^* are related by

$$\int_0^\infty e^{-xt} dP(D^* \leqq x) = \sum_{k=0}^\infty \left(\frac{c\mu}{t + c\mu}\right)^k P((Q^* - c + 1)^+ = k). \tag{3.6}$$

Proof. By Theorem 2.1, X_s^* has finite moments of all orders for all $s \in [0, 1)$. By (3.5), the distribution of Q^* is a mixture of the distributions of X_s^*, so that Q^* also has finite moments of all orders. By (3.4),

$$D^* \stackrel{d}{=} U_1 + \cdots + U_{(Q^* - c + 1)^+}, \tag{3.7}$$

where $\{U_n\}$ is a sequence of i.i.d. exponential random variables independent of Q^*. Hence, D^* has finite moments of all orders, which can be obtained from Q^* via the Laplace transform (3.6).

We now obtain various relations between customer averages and time averages. Let $C(t)$ be the number of occupied servers at time t.

Theorem 3.6. With probability 1,

$$\lim_{t \to \infty} t^{-1} \int_0^t X(s)ds = \bar{\lambda}(ED^* + \mu^{-1}) \tag{3.8}$$

and

$$\lim_{t \to \infty} t^{-1} \int_0^t C(s)ds = \bar{\lambda}/\mu. \tag{3.9}$$

Proof. By Stidham (1974) or Theorem 11.4 of Heyman and Sobel (1982),

$$\lim_{t \to \infty} t^{-1} \int_0^t [X(s) - C(s)]ds = \lim_{t \to \infty} t^{-1} \int_0^t \lambda(s)ds \cdot \lim_{n \to \infty} n^{-1} \sum_{i=1}^n D_i$$

and

$$\lim_{t \to \infty} t^{-1} \int_0^t X(s)ds = \lim_{t \to \infty} t^{-1} \int_0^t \lambda(s)ds \cdot \lim_{n \to \infty} n^{-1} \sum_{i=1}^n (D_i + V_i)$$

where V_n is the exponential service time of the nth arriving customer having mean μ^{-1}, provided the two limits on the right exist in each case. By assumption, the average arrival rate converges to $\bar{\lambda}$. By Theorems 3.3 and 3.5, plus the law of large numbers, e.g., Theorem 6.4 of Heyman and Sobel (1982), the other limits on the right are as stated.

Let $W(t)$ be the workload (i.e., unexpired service time) in the system at time t, which is distributed as

$$W(t) \stackrel{d}{=} V_1 + \cdots + V_{X(t)} \tag{3.10}$$

for each t, where $\{V_n\}$ is a sequence of i.i.d. exponential service times with mean μ^{-1}. The following extends Brumelle's (1971) result for $GI/G/c$ queues.

Theorem 3.7. With probability 1,

$$\lim_{t\to\infty} t^{-1} \int_0^t W(s)ds = \bar{\lambda}(\mu^{-1}ED^* + \mu^{-2}).$$

Proof. Apply the generalization of $L = \lambda W$, $H = \lambda G$, as contained in Heyman and Stidham (1980) or p. 408 of Heyman and Sobel (1982). Let the basic integrable real-valued function $f_n(t)$ be the remaining service time of customer n at time t, as shown in Figure 11.4 of Heyman and Sobel. The rest of the argument is as in Examples 11.11. and 11.12 of Heyman and Sobel.

4. Stationary versions

As given, the continuous-time processes $X(t)$ and $W(t)$ and the discrete-time processes Q_n and D_n are not stationary. Also the arrival process $A(t)$ is not a stationary point process. However, it is easy to construct stationary versions, so that the periodic case can be put in the framework of Franken et al. (1981). First, if

$$\Lambda(t) = \lambda(t + \theta), \qquad t \geqq 0, \tag{4.1}$$

where θ is a random variable uniformly distributed on $[0, 1]$, then the intensity is a stationary ergodic process; e.g., p. 616 of Rolski (1981). The associated arrival process

$$A'(t) = \Pi\left(\int_0^t \Lambda(s)ds\right), \qquad t \geqq 0, \tag{4.2}$$

where Π is a Poisson process with unit intensity independent of θ, is a stationary point process. The associated queue-length process $X'(t) = X(t + \theta)$, $t \geqq 0$ where $X'(0) \stackrel{d}{=} X_0^*$ and $A'(t)$ in (4.2) is the arrival process, is the stationary version of $X(t)$.

For the associated embedded sequence Q_n, the stationary point process $A'(t)$ in (4.2) is replaced by its synchronous or Palm version (the sequence of interarrival times is stationary); see Franken et al. (1981). With this arrival sequence and with initial distribution Q^*, the process Q_n becomes stationary. For example, Theorems 3.6 and 3.7 can be obtained from these stationary versions; p. 106 of Franken et al.

5. General service times

Extensions are not difficult for non-exponential service times, but we do not try for maximum generality in this direction. First, for service-time distributions that are stochastically less than or equal to the exponential distribution, stability in the setting of Theorem 2.1 and the regenerative structure in the setting of Theorem 3.2 follow for $X(t)$ and the related processes by stochastic dominance; Theorem 8 of Whitt (1981). Second, for service-time distributions of phase type as in Chapter 2 of Neuts (1981), Theorem 3.2 can be extended by considering the Markov chain obtained by appending to the variable X_n in (3.1) supplementary variables indicating the phase of each customer in service. For special phase-type distributions, we give a quick proof. Suppose the service-time distribution is a finite mixture of convolutions of a single exponential distribution. Let V_n be the number of phases in the system at epoch $s+n$. When the number of phases exceeds cm, where m is the maximum number of phases in a service time, then we know all c servers are busy. Hence, if $\bar{\lambda} < c\mu$ as in Section 3, then there is a positive ε and an integer k_0 such that

$$E(V_{n+1} - V_n \mid V_n = k) < -\varepsilon \tag{5.1}$$

for all $k \geqq k_0$. As a consequence, we can apply Pakes (1969) to establish strong stability for the chain V_n. Thus Lemma 3.1 and the other results in Section 3 hold for V_n. For X_n defined by (3.1), we have $X_n \leqq V_n$ for each sample point, so that everything carries over to X_n as well.

Even though the result above does not cover all service-time distributions with finite mean, it is close because even these special phase-type distributions are dense in the family of all distributions on the real line in the usual topology of weak convergence. Moreover, for general service-time distributions that are stochastically less than or equal to such a phase-type distribution, the previously mentioned stochastic dominance can be applied again.

More generally, one can consider the vector-valued Markov chain obtained by appending to X_n in (3.1) supplementary variables indicating the residual service time of each customer in service. As in the case of the standard $GI/G/c$ model, it is perhaps easiest to analyze the Kiefer–Wolfowitz workload vector, obtained by assigning the customers and their service times upon arrival to the server who will eventually serve them; see Whitt (1982) and references there. Here we would look at the embedded sequence obtained by considering the times $s + n$, $n \geqq 1$. This embedded workload sequence can be analyzed using the generalization of Pakes (1969) criterion (5.1) in Tweedie (1975).

Of course, the analysis gets much more involved with these supplementary variables, taking us beyond the relatively elementary methods used in this paper. For general results about convergence in distribution, with independence relaxed so that there no longer are embedded Markov chains or regenerative structure, the methods in Franken et al. (1981) and references there seem appropriate.

References

BAUER, H. (1972) *Probability Theory and Elements of Measure Theory.* Holt, Rinehart and Winston, New York.

BRUMELLE, S. L. (1971) On the relation between customer and time averages in queues. *J. Appl. Prob.* **8**, 508–520.

CHUNG, K. L. (1974) *A Course in Probability Theory,* 2nd edn. Academic Press, New York.

CLARK, G. M. (1981) Use of Polya distributions in approximate solutions to nonstationary $M/M/s$ queues. *Commun. ACM* **24**, 206–217.

CLARKE, A. B. (1956) A waiting line process of Markov type *Ann. Math. Statist.* **27**, 452–459.

CRANE, M. A. AND LEMOINE, A. J. (1977) *An Introduction to the Regenerative Method of Simulation Analysis.* Springer-Verlag, New York.

FELLER, W. (1968) *An Introduction to Probability Theory and its Applications.* Vol. I, 3rd edn. Wiley, New York.

FRANKEN, P., KÖNIG, D., ARNDT, U. AND SCHMIDT, V. (1981) *Queues and Point Processes.* Akademie-Verlag, Berlin.

HARRISON, J. M. AND LEMOINE, A. J. (1977) Limit theorems for periodic queues. *J. Appl. Prob.* **14**, 566–576.

HEYMAN, D. P. (1982) On Ross's conjectures about queues with non-stationary Poisson arrivals. *J. Appl. Prob.* **19**, 245–249.

HEYMAN, D. P. AND SOBEL, M. J. (1982) *Stochastic Models in Operations Research,* Volume 1. McGraw-Hill, New York.

HEYMAN, D. P. AND STIDHAM, S., JR. (1980) The relation between customer and time averages in queues. *Operat. Res.* **28**, 983–984.

HEYMAN, D. P. AND WHITT, W. (1978) Queues with time-varying arrivals. *ORSA/TIMS New York Conference Bulletin,* May 1978, WPA16.2, 159.

IGLEHART, D. L. AND WHITT, W. (1970) Multiple channel queues in heavy traffic, I. *Adv. Appl. Prob.* **2**, 150–177.

KELLER, J. B. (1982) Time-dependent queues. *SIAM Rev.* **24**, 401–412.

LEMOINE, A. J. (1976) On random walks and stable $GI/G/1$ queues. *Math. Operat. Res.* **1**, 159–164.

LEMOINE, A. J. (1981) On queues with periodic Poisson input. *J. Appl. Prob.* **18**, 889–900.

LUCHAK, G. (1956) The solution of the single-channel queueing equations characterized by a time-dependent Poisson-distributed arrival rate and a general class of holding times. *Operat. Res.* **4**, 711–732.

MASSEY, W. A. (1981) Non-Stationary Queues. Ph.D. Dissertation, Department of Mathematics, Stanford University.

MCCLISH, D. (1979) Queues and Stores with Non-Homogeneous Input. Ph.D. Dissertation, Department of Statistics, The University of North Carolina.

NEUTS, M. F. (1981) *Matrix-Geometric Solutions in Stochastic Models, An Algorithmic Approach.* The Johns Hopkins University Press, Baltimore.

NEWELL, G. F. (1968) Queues with time-dependent arrival rates, I, II and III. *J. Appl. Prob.* **5**, 436–451, 579–606.

NEWELL, G. F. (1971) *Applications of Queueing Theory.* Chapman and Hall, London.

OLIVER, R. M. AND SAMUEL, A. H. (1962) Reducing letter delays in post offices. *Operat. Res.* **10**, 839–892.

PAKES, A. G. (1969) Some conditions for ergodicity and recurrence of Markov chains. *Operat. Res.* **17**, 1058–1061.

ROLSKI, T. (1981) Queues with non-stationary input stream: Ross's conjecture. *Adv. Appl. Prob.* **13**, 603–618.

ROSS, S. M. (1978) Average delay in queues with non-stationary Poisson arrivals. *J. Appl. Prob.* **15**, 602–609.

ROTHKOPF, M. H. AND OREN, S. S. (1979) A closure approximation for the nonstationary $M/M/s$ queue. *Management Sci.* **25**, 522–534.

SONDERMAN, D. (1979) Comparing multi-server queues with finite waiting rooms, II: different number of servers. *Adv. Appl. Prob.* **11**, 448–455.

STIDHAM, S., JR. (1974) A last word on $L = \lambda W$. *Operat. Res.* **22**, 417–421.

TAAFE, M. (1982) Approximating Nonstationary Queues. Ph.D. Dissertation, Department of Industrial and Systems Engineering, The Ohio State University.

TWEEDIE, R. L. (1975) Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Proc. Appl.* **3**, 385–403.

WHITT, W. (1981) Comparing counting processes and queues. *Adv. Appl. Prob.* **13**, 207–220.

WHITT, W. (1982) Existence of limiting distributions in the $GI/G/s$ queue. *Math. Operat. Res.* **7**, 88–94.

WOLFF, R. W. (1982) Poisson arrivals see time averages. *Operat. Res.* **30**, 223–231.

YU, O. S. (1974) Stochastic bounds for heterogeneous-server queues with Erlang service times. *J. Appl. Prob.* **11**, 785–796.