The Annals of Probability
1980, Vol. 8, No. 5, 974–985

A DUALITY THEOREM FOR PHASE TYPE QUEUES[1]

BY V. RAMASWAMI AND MARCEL F. NEUTS

Purdue University and University of Delaware

A duality theorem of Heathcote exhibiting a relationship between first passage times of the queue length in the GI/M/1 queue and the busy period of its dual M/G/1 queue is generalized to the phase type queues GI/PH/1 and PH/G/1. The phase type distributions include a number of well-known models such as generalized Erlang and hyperexponential as special cases and form a versatile class with a number of interesting closure properties.

1. Introduction. The problem, discussed in this paper, arose in the context of a larger investigation into the nature of the busy period of subclasses of the GI/G/1 queues, which are denoted by the symbols PH/G/1 and GI/PH/1 and which are defined below. We attempted i.a. to generalize the well-known duality result of Heathcote [2], which relates the limit distribution of the times between points of increase for the maximum queue length process in the GI/M/1 queue to the distribution of the busy period of the dual M/G/1 queue. This result implies in particular that in an unstable GI/M/1 queue the maximum queue length process grows approximately like the counting process of a renewal process.

The duality theorem of Heathcote carries over to the GI/PH/1and PH/G/1 queues, but the technical difficulties of the proofs are substantially greater than in [2]. Several deeper properties of the matrix transforms used in the study of these queues are needed, and the methods of the present paper may be of independent interest. Our main theorem implies in particular that the maximum queue length process in an unstable GI/PH/1 queue grows approximately like the counting process of an appropriately defined Markov renewal process.

2. Phase type distributions and phase type renewal processes.

(a). *Phase type distributions*. A (continuous) *probability distribution of phase type*, introduced by M. F. Neuts [4], is any continuous probability distribution on $[0, \infty)$ which is obtainable as the distribution of the time till absorption in a continuous-time finite state space Markov chain with a single absorbing state into which absorption is certain. The class of such distributions includes a number of well-known particular cases such as generalized Erlang and hyperexponential (i.e., a mixture of a finite number of exponentials) distributions and due to its interesting closure properties [4] constitutes a versatile class with properties especially useful in the algorithmic solution of several queueing models (cf. references in [6] and [8]).

Received May 14, 1979.

[1]Research supported by AFOSR-72-2350C at the Depart. Statist., Purdue Univ. and AFOSR-77-3236 at the Depart. Statist. and Computer Sci., Univ. Delaware.

AMS 1970 *subject classifications.* Primary 62K25.

Key words and phrases. Queueing theory, duality, phase type distributions.

To be specific, consider a Markov chain with state space $\{1,\cdots,m,m+1\}$, initial probability vector $(\underline{\alpha},\alpha_{m+1})$ and infinitesimal generator

$$Q=\begin{bmatrix} T & \underline{T}^{\circ} \\ \underline{0} & 0 \end{bmatrix},$$

where $\underline{\alpha}=(\alpha_1,\cdots,\alpha_m)$, T is a nonsingular $m\times m$ matrix with $T_{ii}<0$ and $T_{ij}\geqslant 0$ for $i\neq j$, and $\underline{T}^{\circ}\geqslant\underline{0}$ is an m-vector satisfying $T\underline{e}+\underline{T}^{\circ}=\underline{0}$ with $\underline{e}'=(1,\cdots,1)$. For such a Markov chain the probability distribution of X, the time till absorption in $m+1$, is given by the cdf,

$$H(x) = 1-\underline{\alpha}\exp(Tx)\underline{e}, \qquad x\geqslant 0. \tag{2.1}$$

Definition 2.2. Any probability distribution $H(\cdot)$ obtained as above is called a *phase type distribution* (PH-distribution), and the pair $(\underline{\alpha},T)$ is *a representation* of $H(\cdot)$.

Remark. To avoid uninteresting complications, in the sequel we shall assume $\alpha_{m+1}=0$ so that $H(\cdot)$ does not have an atom at 0.

Example 2.3. The *generalized Erlang* distribution which is the convolution of m independent exponential distributions with parameters, say, $\mu_1,\cdots,\mu_m$ respectively has representation

$$\underline{\alpha} = (1,0,\cdots,0)$$

$$T=\begin{bmatrix} -\mu_1 & \mu_1 & 0 & \cdots & 0 \\ & -\mu_2 & \mu_2 & \cdots & 0 \\ & & \ddots & & \\ & & & & -\mu_m \end{bmatrix}$$

Example 2.4. The *hyperexponential* distribution which is the mixture of m exponentials with parameters, say, $\mu_1,\cdots,\mu_m$ has a representation $T=\operatorname{diag}(-\mu_1,\cdots,-\mu_m)$ with the components of $\underline{\alpha}$ giving the respective mixture ratios.

(b). *Renewal processes of phase type*. Each time the Markov chain Q becomes absorbed in the state $(m+1)$, restart it by performing a multinomial trial with possible outcomes $1,\cdots,m$ and probabilities $\alpha_1,\cdots,\alpha_m$ to pick a new "initial state". Considering each absorption into the state $m+1$ as a renewal, we obtain a renewal process for which the time between any two successive renewals has cdf $H(\cdot)$, the PH-distribution given by (2.1). Such a renewal process is called a *renewal process of phase type* (PH-renewal process) (Neuts [6]).

The above procedure also constructively defines a new Markov chain with state-space $\{1,\cdots,m\}$, initial probability vector $\underline{\alpha}$ and infinitesimal generator $Q^*=T+T^{\circ}A^{\circ}$, where $A^{\circ}=\operatorname{diag}(\alpha_1,\cdots,\alpha_m)$ and $T^{\circ}=(\underline{T}^{\circ},\cdots,\underline{T}^{\circ})$. This Markov chain describes the "phase" of the system and is of considerable importance. In [4] it is shown that one may, without loss of generality, assume that the representation $(\underline{\alpha}, T)$ of $H(\cdot)$ is so chosen as to make Q^* irreducible, and we shall henceforth assume that this is indeed the case.

We let $\underline{\theta}$ denote the stationary probability vector of the Markov chain Q^*, i.e., the unique (strictly positive) vector satisfying $\underline{\theta}Q^* = \underline{0}$, $\underline{\theta}\underline{e} = 1$. It may be easily verified that $\underline{\theta} = -\lambda\underline{\alpha}T^{-1}$, where $\lambda^{-1} = -\underline{\alpha}T^{-1}\underline{e}$ is the mean of $H(\cdot)$.

The $m \times m$ matrices $P(\nu, t), \nu \geqslant 0, t \geqslant 0$, defined in [6] are such that the entry $P_{jj'}(\nu, t)$ is the conditional probability, given that the initial phase is j, that at time $t+$, the Q^*-chain is in state j' and that ν renewals have occurred in $(0, t]$. It is known [6] that these have generating function

$$\tilde{P}(z,t) \equiv \Sigma_{\nu=0}^{\infty} z^{\nu} P(\nu,t) = \exp[(T + zT^{\circ}A^{\circ})t], |z| \leqslant 1, t \geqslant 0.$$

We also recall [7] that, under the assumption Q^* is irreducible, the matrices $P(\nu, t), \nu \geqslant 1, t > 0$, are all strictly positive. Further, it may be easily seen that $P_{ii}(0, t) > 0$ for all $t \geqslant 0, 1 \leqslant i \leqslant m$.

3. The phase type queues.

(a). *The* GI/PH/1 *queue*. Consider a GI/G/1 queue in which the service time cdf $H(\cdot)$ is of phase type with representation $(\underline{\alpha}, T)$ and where the interarrival times are i.i.d., with a nondegenerate probability distribution $F(\cdot)$. Such a model, denoted by GI/PH/1, has been discussed in detail by Neuts [7], and for later use we quote the following results.

Let τ_n denote the epoch of the nth arrival, ξ_n the size of the system at τ_n+ and J_n the phase of the service at τ_n+. Then $\{(\xi_n, J_n, \tau_n - \tau_{n-1})\colon n \geqslant 0\}$, where $\tau_0 = \tau_{-1} = 0$, defines a semi-Markov sequence with state space $\{1, 2, \cdots\} \times \{1, \cdots, m\} \times [0, \infty)$ and transition matrix $\tilde{Q}(x)$ given by

$$\tilde{Q}(x) = \begin{bmatrix} \tilde{B}_0(x) & \tilde{A}_0(x) & \cdots & \cdots & \cdots \\ \tilde{B}_1(x) & \tilde{A}_1(x) & \tilde{A}_0(x) & \cdots & \cdots \\ \tilde{B}_2(x) & \tilde{A}_2(x) & \tilde{A}_1(x) & \tilde{A}_0(x) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \qquad x \geqslant 0,$$

where,

$$\tilde{A}_n(x) = \textstyle\int_0^x P(n,t)\, dF(t), \qquad n \geqslant 0, x \geqslant 0,$$

$$\tilde{B}_n(x) = \Sigma_{\nu=n+1}^{\infty} \textstyle\int_0^x P(\nu,t)\, dF(t) A^{\circ\circ}, \qquad n \geqslant 0, x \geqslant 0,$$

where $A^{\circ\circ}$ is the $m \times m$ matrix each of whose rows is $\underline{\alpha}$.

Since $F(\cdot)$ is nondegenerate, we have (from the strict positivity of the matrices $P(\nu, t)$ for $\nu \geqslant 1, t > 0$ and the positivity of the entries $P_{ii}(0, t)$ for $t \geqslant 0$ mentioned at the end of Section 2) that the matrices $A_n \equiv \tilde{A}_n(\infty), n \geqslant 1$, are all strictly positive; further, the matrix $A_0 \equiv \tilde{A}_0(\infty)$ is such that each of its diagonal entries is positive. This entails that the embedded Markov chain with transition probability matrix $\tilde{Q}(\infty)$—with possibly some of the states $(1,j)$ removed—is irreducible and aperiodic. We shall conclude our introduction to the GI/PH/1 queue by recalling

[7] that this queue is stable iff $\rho = \underline{\theta}\underline{\beta} > 1$, where $\underline{\theta}$ is (also) the invariant probability vector of $A \equiv \Sigma_{n=0}^{\infty} A_n$, and $\underline{\beta} = \Sigma_{n=1}^{\infty} n A_n \underline{e}$.

(b). *The* PH/G/1 *queue.* By PH/G/1 we denote a GI/G/1 queue in which the arrival process is a PH-renewal process. Note that this queue is the dual of the GI/PH/1 queue defined above in that it may be obtained by reversing the roles of interarrival times and service times in the latter. The PH/G/1 queue is a subclass of the more general N/G/1 queues studied by V. Ramaswami [8], and we shall state below some basic properties concerning the PH/G/1 queue by particularizing results obtained in [8].

Let us assume that the arrival process is the PH-renewal process with representation $(\underline{\alpha}, T)$ and that the service time cdf is given by $F(\cdot)$. Defining $\hat{\tau}_n$ to be the epoch of the nth departure ($\hat{\tau}_0 = 0$), and $\hat{\xi}_n$ and $\hat{J}_n$ to be respectively the queue length (i.e., the number of customers in the system) and the phase of the arrival process at $\tau_n +$, it is easily seen that $\{(\hat{\xi}_n, \hat{J}_n, \hat{\tau}_{n+1} - \hat{\tau}_n)\colon n \geqslant 0\}$ is a semi-Markov sequence with state space $\{0, 1, \cdots\} \times \{1, \cdots, m\} \times [0, \infty)$ and transition matrix

$$\hat{Q}(x) = \begin{bmatrix} \tilde{C}_0(x) & \tilde{C}_1(x) & \tilde{C}_2(x) & \cdots & \cdots \\ \tilde{A}_0(x) & \tilde{A}_1(x) & \tilde{A}_2(x) & \cdots & \cdots \\ 0 & \tilde{A}_0(x) & \tilde{A}_1(x) & \cdots & \cdots \\ 0 & 0 & \tilde{A}_0(x) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}, \qquad x \geqslant 0,$$

where $\tilde{A}_n(\cdot), n \geqslant 0$, are as defined earlier in (a), and

$$\tilde{C}_n(x) = \int_0^x \exp(Ty) \cdot T^\circ A^\circ \cdot \tilde{A}_n(x-y)\, dy, \qquad n \geqslant 0.$$

It is easily seen that the Markov chain defined by $\hat{Q}(\infty)$ is irreducible and aperiodic. We recall from [8] that the PH/G/1 queue is stable iff $\rho = \underline{\theta}\underline{\beta} < 1$. Further, $\rho = \lambda\mu$, where μ is the mean of $F(\cdot)$, is the traffic intensity of the PH/G/1 queue.

Basic to the discussion of the PH/G/1 queue are the first passage times from the set of states $\underline{1} = \{(1,j)\colon 1 \leqslant j \leqslant m\}$ to the set of states $\underline{0} = \{(0,j)\colon 1 \leqslant j \leqslant m\}$ which are governed by the matrix $\tilde{G}(x), x \geqslant 0$, whose (j,j')th entry is the probability that starting at $(1,j)$, the process enters the set $\underline{0}$ for the first time at or before time x by visiting the state $(0,j')$. Defining the Laplace-Stieltjes transform

$$G(s) = \int_0^\infty e^{-sx}\, d\tilde{G}(x), \qquad s \geqslant 0, \tag{3.1}$$

we have

THEOREM 3.2. (i) *The matrix* $G(\cdot)$ *satisfies the matrix functional equation*

$$G(s) = \Sigma_{n=0}^{\infty} A_n(s) G^n(s), \qquad s \geqslant 0, \tag{3.3}$$

where

$$A_n(s) = \int_0^\infty e^{-sx}\, d\tilde{A}_n(x) \qquad s \geqslant 0.$$

(ii) *For* $s > 0$, *there exists a unique nonnegative matrix* $G(s)$ *which satisfies* (3.3). *For* $s > 0, G(s)$ *is analytic and can be written in the form* (3.1) *where the entries of* $\tilde{G}(\cdot)$ *are (defective) probability mass-functions.*

(iii) *The matrix* $G = G(0+)$, *defined by continuity, is substochastic, and is the minimal solution in the class of substochastic matrices of the matrix functional equation*

$$G = \sum_{n=0}^{\infty} A_n G^n. \tag{3.4}$$

(iv) *If* $\rho \leqslant 1$, *then* G *is stochastic. If* $\rho > 1$, *then at least one component of* $G\underline{e}$ *is less than one.*

(v) *For any* $s \geqslant 0, G(s)$ *is strictly positive and is the monotone limit of the nondecreasing sequence* $\{G_n(s)\}_0^{\infty}$ *defined by*

$$(3.5)\qquad G_0(s) = 0$$

$$G_{n+1}(s) = \sum_{\nu=0}^{\infty} A_\nu(s) G_n^\nu(s), \qquad n \geqslant 0.$$

PROOF. All the results above follow from Theorem 2.2.11 in [8] which in turn is proved by Neuts [5] in a more general context in the analysis of Markov renewal processes with $\hat{Q}(\cdot)$ of the form given above.

REMARKS.

1. In view of the structure of $\hat{Q}(\cdot)$, we have that $\tilde{G}(\cdot)$ also describes first passages from the set $\underline{i+1} = \{(i+1,j)\colon 1 \leqslant j \leqslant m\}$ to the set $\underline{i} = \{(i,j)\colon 1 \leqslant j \leqslant m\}$ for any $i \geqslant 0$.
2. Note that the ith entry of $G(s)\underline{e}$ is the Laplace-Stieltjes transform of the busy period starting with one customer and in phase i.
3. Equation (3.3) generalizes Takács' equation for the M/G/1 queue.

4. First passage times in a GI/PH/1 queue. Consider the GI/PH/1 queue defined in Section 3a. By level $\underline{n}$ we denote the set of states $\{(n,1),\cdots,(n,m)\}$. The principal objects of study of this paper are the first passage times from $\underline{n}$ to $\underline{n+k}$ in the GI/PH/1 queue; these will be the subject matter of this section. In the sequel, we let for $x \geqslant 0$, $n, k \geqslant 1$, $\tilde{D}_{n,n+k}(x)$ denote the $m \times m$ matrix whose (i,j)th entry $\tilde{D}_{n,n+k}(i,j;x)$ is the conditional probability that the process enters $\underline{n+k}$ for the first time at or before x by visiting $(n+k,j)$ given that it starts at level $\underline{n}$ and in phase i. We also define the Laplace-Stieltjes transform

$$D_{n,n+k}(s) = \int_0^{\infty} e^{-sx}\, d\tilde{D}_{n,n+k}(x), \qquad s \geqslant 0.$$

PROPOSITION 4.1. *For* $n \geqslant 1, k \geqslant 2$, *we have*

$$\tilde{D}_{n,n+k} = \tilde{D}_{n,n+1} * \tilde{D}_{n+1,n+2} * \cdots * \tilde{D}_{n+k-1,n+k}, \tag{4.2}$$

and

$$D_{n,n+k} = D_{n,n+1} D_{n+1,n+2} \cdots D_{n+k-1,n+k}, \tag{4.3}$$

where $*$ *in* (4.2) *denotes matrix convolution.*

PROOF. The transition from $\underline{n}$ to $\underline{n+k}$ can occur only along a path of first passages from $\underline{n}$ to $\underline{n+1}, \cdots,$ from $\underline{n+k-1}$ to $\underline{n+k}$. Given any sequence

$j_1, \cdots, j_{k-1} \varepsilon \{1, \cdots, m\}$ denoting the phases at the epochs of such first entrances to $\underline{n+1}, \cdots, \underline{n+k-1}$, the duration between these first entrance times are conditionally independent. This is an immediate consequence of the Markov property at transition epochs for the Markov renewal process defined by $\tilde{Q}(\cdot)$. Hence (4.2). (4.3) follows immediately.

THEOREM 4.4. *For* $n \geqslant 1, s \geqslant 0$,

$$(4.5) \quad D_{n,n+1}(s) = A_0(s) + \Sigma_{\nu=1}^{n-1} A_\nu(s) D_{n-\nu+1,n+1}(s) + \Sigma_{\nu=n}^{\infty} A_\nu(s) A^{\circ\circ} D_{1,n+1}(s)$$

(*where the second term in the right side is taken to be* 0 *when* $n = 1$).

PROOF. Let N denote the number of departures in the first interarrival interval. Also let us denote by $\mathbb{S}$ the event that the process enters $\underline{n+1}$ at or before time x and that the phase at which it enters $\underline{n+1}$ is j. Then by a simple probabilistic argument, we have

$$P(\mathbb{S}, N = \nu \mid \xi_0 = n, J_0 = i) = \begin{cases} \int_0^x P_{ij}(0,u)\, dF(u), & \text{if } \nu = 0 \\ \int_0^x \Sigma_{k=1}^m P_{ik}(\nu,u) \tilde{D}_{n-\nu+1,n+1}(k,j;x-u)\, dF(u), & \text{if } 1 \leqslant \nu \leqslant n-1 \\ \int_0^x \Sigma_{\nu=n}^{\infty} \Sigma_{l=1}^m P_{il}(\nu,u) \Sigma_{k=1}^m \alpha_k \tilde{D}_{1,n+1}(k,j;x-u)\, dF(u), & \text{if } \nu = n, \end{cases}$$

whence

$$\tilde{D}_{n,n+1}(i,j;x) = \int_0^x P_{ij}(0,u)\, dF(u) + \Sigma_{\nu=1}^{n-1} \int_0^x \Sigma_{k=1}^m P_{ik}(\nu,u) \tilde{D}_{n-\nu+1,n+1}(k,j;x-u)\, dF(u) + \Sigma_{\nu=n}^{\infty} \int_0^x \Sigma_{l=1}^m P_{il}(\nu,u) \Sigma_{k=1}^m \alpha_k \tilde{D}_{1,n+1}(k,j;x-u)\, dF(u).$$

Computing the Laplace-Stieltjes transforms of these and putting the result in matrix notation, we get (4.5).

From the irreducibility of $\tilde{Q}(\cdot)$ it is clear that the matrix $D_{n,n+1}(0)$ should be stochastic. That this is indeed the case is verified below by directly using (4.5).

LEMMA 4.6. $D_{1,2}(0)$ *is stochastic.*

PROOF. By (4.5),

$$(4.7) \qquad D_{1,2}(0) = A_0 + (A - A_0) A^{\circ\circ} D_{1,2}(0)$$

or,

$$A^{\circ\circ} D_{1,2}(0) \underline{e} = A^{\circ\circ} A_0 \underline{e} + A^{\circ\circ}(A - A_0) A^{\circ\circ} D_{1,2}(0) \underline{e}.$$

Letting $A^{\circ\circ} D_{1,2}(0) \underline{e} = u \underline{e}$, this yields

$$u = (\underline{\alpha} A_0 \underline{e}) + u - (\underline{\alpha} A_0 \underline{e}) \underline{u},$$

and since $\underline{\alpha} A_0 \underline{e} > 0$, $u = 1$. Thus

$$A^{\circ\circ} D_{1,2}(0)\underline{e} = \underline{e}. \tag{4.8}$$

Now, multiplying (4.7) by $\underline{e}$ and using (4.8), we get

$$D_{1,2}(0)\underline{e} = A_0\underline{e} + (A - A_0)\underline{e} = \underline{e}.$$

THEOREM 4.9. *For all $n, k \geqslant 1$, $D_{n,n+k}(0)$ is stochastic. Also for $n \geqslant 2$, note $k \geqslant 1$, $D_{n,n+k}(0)$ is strictly positive.*

PROOF. In view of (4.3) it suffices to prove the results for $k = 1$. Let $n \geqslant 1$ and assume as inductive hypothesis that $D_{1,2}(0), \cdots, D_{n,n+1}(0)$ are all stochastic. We now show that $D_{n+1,n+2}(0)$ is stochastic.

Using (4.5) and (4.3) it is easily seen that

$$[I - M] D_{n+1,n+2}(0) = A_0, \tag{4.10}$$

where

$$M = A_1 + \Sigma_{\nu=2}^{n} A_\nu D_{n+2-\nu,n+1}(0) + (A - \Sigma_{\nu=0}^{n} A_\nu) A^{\circ\circ} D_{1,n+1}(0)$$

(with the second term on the right being 0 when $n = 1$). Now $M \geqslant A_1 \gg 0$, and $M\underline{e} = \underline{e} - A_0\underline{e} \ll \underline{e}$. Thus the strictly positive matrix M has spectral radius η less than 1. That $D_{n+1,n+2}(0) \gg 0$, follows now from (4.10) by writing it as

$$D_{n+1,n+2}(0) = \Sigma_{\nu=0}^{\infty} M^\nu A_0.$$

Let $\underline{y} \gg \underline{0}$ be such that $\underline{y}M = \eta\underline{y}$, $\underline{y}\underline{e} = 1$. Then from (4.10), we have

$$(1 - \eta)\underline{y} D_{n+1,n+2}(0)\underline{e} = \underline{y} A_0 \underline{e}$$

$$= \underline{y}(\underline{e} - M\underline{e}) = (1 - \eta)$$

and since $\eta < 1$, $\underline{y} D_{n+1,n+2}(0)\underline{e} = 1$. Thus $D_{n+1,n+2}(0)\underline{e} = \underline{e}$, and the proof is complete by mathematical induction.

Before concluding this section we list some simple but useful results as

PROPOSITION 4.11. *Let*

$$\phi(s) = \int_0^\infty e^{-sx} dF(x), \qquad s \geqslant 0.$$

For all $s \geqslant 0$, $n, k \geqslant 1$,

(a) $D_{n,n+k}(s)\underline{e} \leqslant \underline{e}$

(b) $D_{n,n+1}(s)\underline{e} \leqslant \phi(s)\underline{e}$

(c) $D_{n,n+k}(s)\underline{e} \leqslant \{\phi(s)\}^k \underline{e}$.

PROOF. (a) is immediate from the definition of $D_{n,n+k}(\cdot)$. (b) is got by applying (a) in (4.5). Finally (c) is got by applying (b) in (4.3).

5. The duality theorem. In this section we prove the following duality theorem which generalizes the result of Heathcote [2] to the phase type queues, discussed in Section 3.

THEOREM 5.1. (*Duality theorem*): *Let* $s > 0$ *be fixed. As* $n \to \infty$, *the matrices* $D_{n,n+1}(s)$ *describing the first passage times from* $\underline{n}$ *to* $\underline{n+1}$ *in the* GI/PH/1 *queue converge to the matrix* $G(s)$, *where* $G(s)$ *is the unique nonnegative solution of the nonlinear matrix functional equation*

$$G(s) = \Sigma_{\nu=0}^{\infty} A_\nu(s) G^\nu(s) \tag{5.2}$$

and which describes the first passage times from $\underline{1}$ *to* $\underline{0}$ *in the dual* PH/G/1 *queue. Further the limit* $G(0+)$ *is stochastic* iff $\rho \leqslant 1$, i.e., iff *the* GI/PH/1 *queue is unstable.*

The lengthy proof of Theorem 5.1 is accomplished in several stages. In the ensuing discussion $s > 0$ is assumed to be fixed. Also we let $\limsup_{n\to\infty} D_{n,n+1}(s)$ and $\liminf_{n\to\infty} D_{n,n+1}(s)$ denote the matrices whose respective (i,j)th entries are $\limsup_{n\to\infty} D_{n,n+1}(i,j;s)$ and $\liminf_{n\to\infty} D_{n,n+1}(i,j;s)$. From Proposition 4.11(b) it is clear that these matrices are strictly substochastic; further,

$$0 \leqslant \liminf_{n\to\infty} D_{n,n+1}(s) \leqslant \limsup_{n\to\infty} D_{n,n+1}(s). \tag{5.3}$$

LEMMA 5.4. *Define*

$$M_1(s) = A_0(s)$$

$$M_{n+1}(s) = A_0(s) + \Sigma_{\nu=1}^{n} A_\nu(s) \Pi_{k=n-\nu+2}^{n+1} M_k(s), \qquad n \geqslant 1$$

where the matrix product is formed by taking the terms in increasing order of the index k. *Then*

(i) $M_n(s)$ *is strictly substochastic for every* $n \geqslant 1$,

(ii) $M_n(s)$ *is entrywise nondecreasing in* n,

(iii) $\lim_{n\to\infty} M_n(s) = G(s)$, *where* $G(s)$ *is the unique solution of* (5.2); *the limit is taken entrywise.*

PROOF. It is obvious that $M_1(s)$ is strictly substochastic. Now, since $\Sigma_{\nu=0}^{\infty} A_\nu(s)\underline{e} \ll \underline{e}$, there exist constants $w(s) > 0$ and $\delta(s) > 0$ such that $w(s) + \delta(s) < 1$, $\Sigma_{\nu=1}^{\infty} A_\nu(s)\underline{e} \leqslant w(s)\underline{e}$ and $A_0(s)\underline{e} \leqslant \delta(s)\underline{e}$. Now,

$$M_2(s) = [I - A_1(s)]^{-1} A_0(s) = \Sigma_{\nu=0}^{\infty} A_1^\nu(s) A_0(s),$$

whence

$$M_2(s)\underline{e} \leqslant \frac{\delta(s)}{1 - w(s)}\underline{e} \ll \underline{e}.$$

In other words, $M_2(s)$ is also strictly substochastic, and clearly $M_2(s) \geqslant M_1(s)$.

Now, assume as inductive hypotheses that $M_k(s) \leqslant M_{k+1}(s)$ and $M_{k+1}(s)\underline{e} \ll \underline{e}$, for $1 \leqslant k \leqslant n$. We have,

$$\begin{aligned} &M_{n+2}(s) \\ &= [I - \{A_1(s) + A_2(s)M_{n+1}(s) + \cdots + A_{n+1}(s)M_2(s)\cdots M_{n+1}(s)\}]^{-1} A_0(s) \\ &\geqslant [I - \{A_1(s) + A_2(s)M_n(s) + \cdots + A_n(s)M_2(s)\cdots M_n(s)\}]^{-1} A_0(s) \\ &= M_{n+1}(s). \end{aligned}$$

The inequality above is obtained by writing the inverse as a power series, dropping the term $A_{n+1}(s)M_2(s)\cdots M_{n+1}(s)$, and using the induction hypothesis $M_k(s) \leqslant M_{k+1}(s)$, $1 \leqslant k \leqslant n$. Furthermore,

$$\begin{aligned} &M_{n+2}(s)\underline{e} \\ &= \Sigma_{\nu=0}^{\infty}\{A_1(s) + A_2(s)M_{n+1}(s) + \cdots + A_{n+1}(s)M_2(s)\cdots M_{n+1}(s)\}^{\nu}A_0(s)\underline{e} \\ &\leqslant \frac{\delta(s)}{1-w(s)}\underline{e} \ll \underline{e} \end{aligned}$$

by using the substochasticity of $M_2(s),\cdots, M_{n+1}(s)$ assumed as inductive hypothesis.

Having proved statements (i) and (ii), at this point these also imply the existence of $\lim_{n\to\infty}M_n(s) = M(s)$. So all that we need to prove is that $M(s) = G(s)$. We have,

$$\begin{aligned} M_{n+1}(s) &= A_0(s) + A_1(s)M_{n+1}(s) + \cdots + A_n(s)M_2(s)\cdots M_{n+1}(s) \\ &\leqslant \Sigma_{\nu=0}^{n}A_\nu(s)M_{n+1}^{\nu}(s) \\ &\leqslant \Sigma_{\nu=0}^{\infty}A_\nu(s)M^{\nu}(s). \end{aligned}$$

Letting $n \to \infty$ this yields

$$M(s) \leqslant \Sigma_{\nu=0}^{\infty}A_\nu(s)M^{\nu}(s). \tag{5.5}$$

Also

$$M_{n+m}(s) \geqslant A_0(s) + A_1(s)M_{n+m}(s) + \cdots + A_n(s)M_{m+1}(s)\cdots M_{m+n}(s)$$

yields upon letting $m \to \infty$, that

$$M(s) \geqslant \Sigma_{\nu=0}^{n}A_\nu(s)M^{\nu}(s) \qquad \text{for all} \quad n \geqslant 0.$$

Since the above inequality holds for every n,

$$M(s) \geqslant \Sigma_{\nu=0}^{\infty}A_\nu(s)M^{\nu}(s). \tag{5.6}$$

Now, by (5.5) and (5.6),

$$M(s) = \Sigma_{\nu=0}^{\infty}A_\nu(s)M^{\nu}(s),$$

and it follows that $M(s) = G(s)$ by appealing to the uniqueness of the solution to (5.2) stated in Theorem 3.2(ii).

LEMMA 5.7.

$$\liminf_{n\to\infty} D_{n,n+1}(s) \geqslant G(s). \tag{5.8}$$

PROOF. It suffices to show that $D_{n,n+1}(s) \geqslant M_n(s)$ for all $n \geqslant 1$ where $\{M_n(s)\}$ are as in Lemma 5.4. Clearly, $D_{1,2}(s) \geqslant M_1(s)$. Assume as inductive hypothesis that $D_{n,n+1}(s) \geqslant M_n(s)$ for $n = 1,\cdots, k-1$. Then

$$\begin{aligned} D_{k,k+1}(s) &= \left[I - \{A_1(s) + A_2(s)D_{k-1,k}(s) + \cdots + A_{k-1}(s)D_{2,k}(s)\}\right]^{-1}A_0(s) \\ &\geqslant \left[I - \{A_1(s) + A_2(s)M_{k-1}(s) + \cdots + A_{k-1}(s)M_2(s)\cdots M_{k-1}(s)\}\right]^{-1}A_0(s) \\ &= M_k(s), \end{aligned}$$

where the inequality is obtained by using the induction hypothesis and (4.3) in the series expansion of the inverse. The proof is now complete by mathematical induction.

LEMMA 5.9. *Let*

$$R(s) = \limsup_{n\to\infty} D_{n,n+1}(s).$$

Then

$$R(s) \leqslant \Sigma_{\nu=0}^{\infty} A_\nu(s) R^\nu(s).$$

PROOF. For sufficiently large n and fixed $N < n - 1$, we have,

$$\begin{aligned} D_{n,n+1}(s) &= A_0(s) + \Sigma_{\nu=1}^{n-1} A_\nu(s) D_{n-\nu+1,n+1}(s) \\ &\quad + \Sigma_{\nu=n}^{\infty} A_\nu(s) A^{\circ\circ} D_{1,n+1}(s) \text{ by (4.5)} \\ &= A_0(s) + \Sigma_{\nu=1}^{N} A_\nu(s) D_{n-\nu+1,n+1}(s) + \Sigma_{\nu=N+1}^{n-1} A_\nu(s) D_{n-\nu+1,n+1}(s) \\ &\quad + \Sigma_{\nu=n}^{\infty} A_\nu(s) A^{\circ\circ} D_{1,n+1}(s), \end{aligned}$$

and letting $n \to \infty$,

$$R(s) \leqslant A_0(s) + \Sigma_{\nu=1}^{N} A_\nu(s) R^\nu(s) + \limsup_{n\to\infty} \Sigma_{\nu=N+1}^{n-1} A_\nu(s) D_{n-\nu+1,n+1}(s),$$

for,

$$\limsup_{n\to\infty} \Sigma_{\nu=n}^{\infty} A_\nu(s) A^{\circ\circ} D_{1,n+1}(s) = 0.$$

Now letting $N \to \infty$, we get

$$R(s) \leqslant \Sigma_{\nu=0}^{\infty} A_\nu(s) R^\nu(s),$$

since,

$$\limsup_{N\to\infty} \limsup_{n\to\infty} \Sigma_{\nu=N+1}^{n-1} A_\nu(s) D_{n-\nu+1,n+1}(s) = 0$$

as is seen from the fact that

$$\Sigma_{\nu=N+1}^{n-1} A_\nu(s) D_{n-\nu+1,n+1}(s) \underline{e} \leqslant \Sigma_{\nu=N+1}^{\infty} A_\nu(s) \underline{e} \to \underline{0} \text{ as } N \to \infty.$$

COROLLARY 5.10. *Denoting the maximal eigenvalue of a nonnegative matrix C by* sp(C), *we have that*

$$\text{sp}[\limsup D_{n,n+1}(s)] \leqslant \text{sp}[G(s)]. \tag{5.11}$$

PROOF. By (5.3), (5.8) and the strict positivity of $G(s)$, note that $R(s) \gg 0$. Let $\eta(s)$ be the spectral radius of $R(s) = \limsup_{n\to\infty} D_{n,n+1}(s)$, and let $\underline{x}(s) \gg \underline{0}$ be a right-eigenvector of $R(s)$ associated with $\eta(s)$. Then

$$\begin{aligned} \eta(s)\underline{x}(s) &= R(s)\underline{x}(s) \\ &\leqslant \Sigma_{\nu=0}^{\infty} A_\nu(s) R^\nu(s) \underline{x}(s) \quad \text{by Lemma 5.9} \\ &= \Sigma_{\nu=0}^{\infty} A_\nu(s) \eta^\nu(s) \underline{x}(s) \end{aligned}$$

implying that

$$A^*(\eta(s), s)\underline{x}(s) \geqslant \eta(s)\underline{x}(s) \tag{5.12}$$

where

$$A^*(z,s) = \Sigma_{\nu=0}^{\infty} A_\nu(s) z^\nu, \qquad 0 \leqslant z \leqslant 1.$$

Let $\xi(z,s)$ be the Perron-Frobenius eigenvalue of $A^*(z,s)$. Now (5.12) implies that

(5.13) $$\xi(\eta(s),s) \geqslant \eta(s).$$

Now, for fixed $s > 0$, a theorem due to J. F. C. Kingman [3] implies that the function $\log \xi(e^{-t},s)$ is convex and decreasing for $t \geqslant 0$. Also $G(s)$ is the unique solution of equation (5.2) whence it follows that sp$[G(s)]$ is the unique solution z_0 of the equation

$$\xi(z,s) = z, \qquad 0 < z < 1,$$

so that $z_0 = e^{-t_0}$, where

$$\log[\xi(e^{-t_0},s)] = -t_0, \qquad t_0 > 0.$$

Setting $s = e^{-t^*}$ in equation (5.13), we obtain $\log \xi(e^{-t^*},s) \geqslant -t^*$ which upon consideration of the graph in Figure 1, implies that $t^* \geqslant t_0$. This clearly is equivalent to the inequality sp$[R(s)] \leqslant$ sp$[G(s)]$, which we set out to prove.

PROOF OF THEOREM 5.1. By Formulas (5.8) and (5.11), we have

$$\text{sp}[\liminf_{n\to\infty} D_{n,n+1}(s)] \geqslant \text{sp}[G(s)] \geqslant \text{sp}[\limsup_{n\to\infty} D_{n,n+1}(s)].$$

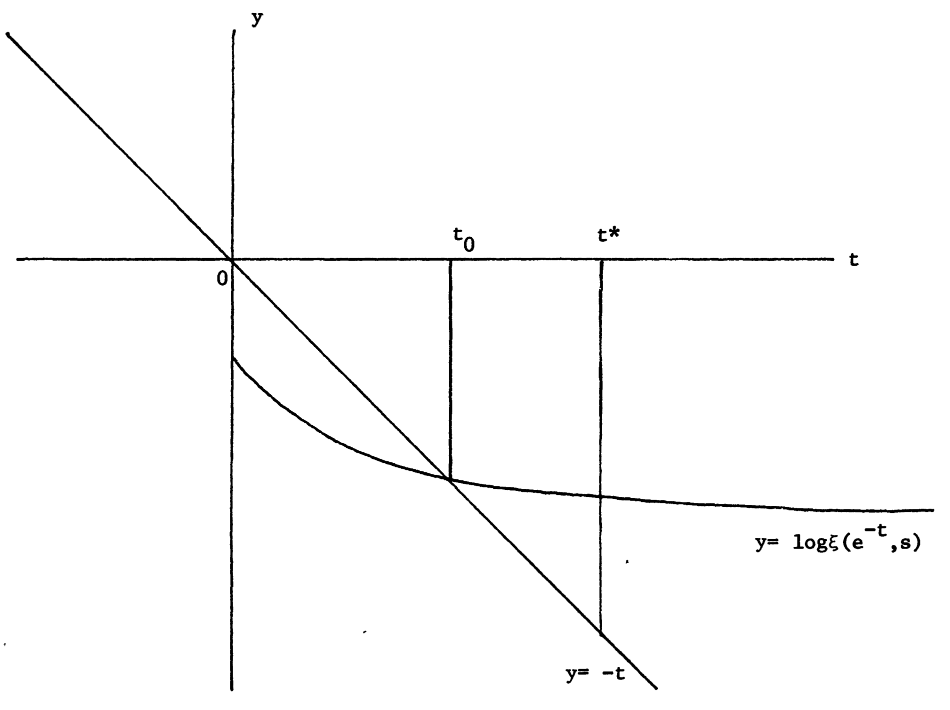

FIG. 1

But by (5.3),

$$\mathrm{sp}[\liminf_{n\to\infty} D_{n,n+1}(s)] \leqslant \mathrm{sp}[\limsup_{n\to\infty} D_{n,n+1}(s)].$$

Thus we have

$$\mathrm{sp}[\liminf_{n\to\infty} D_{n,n+1}(s)] = \mathrm{sp}[G(s)] = \mathrm{sp}[\limsup_{n\to\infty} D_{n,n+1}(s)].$$

Since $\liminf_{n\to\infty} D_{n,n+1}(s) \leqslant \limsup_{n\to\infty} D_{n,n+1}(s)$, and both are irreducible nonnegative matrices, their spectral radii can be equal only if the inequality is actually an (entrywise) equality [1]. This proves the existence of $\lim_{n\to\infty} D_{n,n+1}(s)$. Also (5.8) and the fact that $\mathrm{sp}[G(s)] = \mathrm{sp}[\liminf_{n\to\infty} D_{n,n+1}(s)]$ implies that $\lim_{n\to\infty} D_{n,n+1}(s) = G(s)$.

REFERENCES

[1] Gantmacher, F. R. (1959). *The Theory of Matrices, Vol.* 1. Chelsea, New York.

[2] Heathcote, C. R. (1965). On the maximum of the queue GI/M/1. *J. Appl. Probability* **2** 206–214.

[3] Kingman, J. F. C. (1961). A convexity property of positive matrices. *Quart. J. Math* **12** 283–284.

[4] Neuts, M. F. (1975). Probability distributions of phase type. In *Liber Amicorum Professor Emeritus H. Florin*. 173–206. Dept. Math., Univ. Louvain, Belgium.

[5] Neuts, M. F. (1976). Moment formulas for the Markov renewal branching process. *Advances Appl. Probability* **8**, 690–711.

[6] Neuts, M. F. (1978a). Renewal processes of phase type. *Naval Res. Logist. Quart.* **25** 445–454.

[7] Neuts, M. F. (1978b). Markov chains with applications in queueing theory which have a matrix-geometric invariant probability vector. *Advances Appl. Probability* **10** 185–212.

[8] Ramaswami, V. (1978). The N/G/1 queue and its detailed analysis. *Advances Appl. Probability* **12** 222–261.

Department of Mathematical Sciences
Drexel University
Philadelphia, PA 19104

Department of Mathematical Sciences
501 Kirkbride Office Bldg.
University of Delaware
Newark, Delaware 19711