

AD-A059 509 DELAWARE UNIV NEWARK DEPT OF STATISTICS AND COMPUTER--ETC F/G 12/1  
A DUALITY THEOREM FOR PHASE TYPE QUEUES.(U)

JUN 78 V RAMASWAMI, M F NEUTS

AFOSR-77-3236

UNCLASSIFIED

TR-78/11

AFOSR-TR-78-1143

NL

1 of 1  
AD  
A059 509

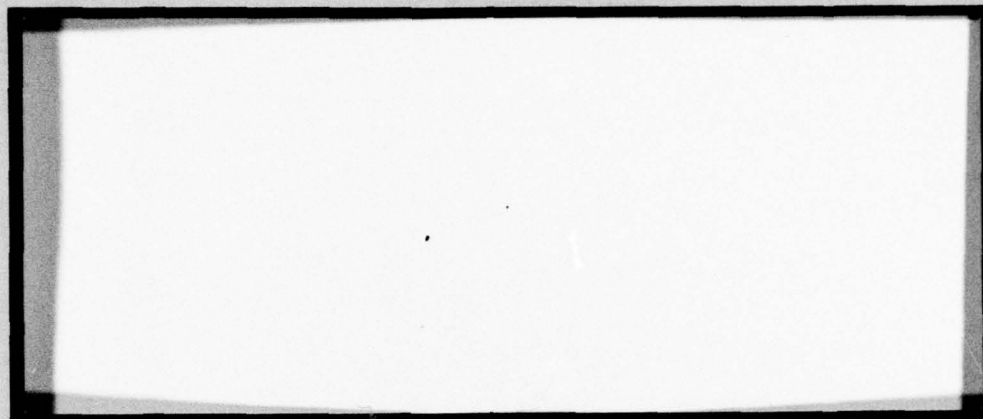


END  
DATE  
FILMED  
-12-78  
DDC

LEVEL II

2

AD A059509



Department of  
STATISTICS AND COMPUTER SCIENCE

DDC FILE COPY

DDC  
RECEIVED  
SEP 25 1978  
D

78 07 20 152

UNIVERSITY OF DELAWARE  
Newark, Delaware 19711

ACCESSION BY	
DTIC	White Section <input checked="" type="checkbox"/>
DDC	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	<input type="checkbox"/>
BY	
DISTRIBUTION/AVAILABILITY CODES	
Dist. AVAIL. and/or SPECIAL	
A	

**LEVEL II**

2

6 A DUALITY THEOREM FOR PHASE TYPE QUEUES,

by

10 V./Ramaswami

Purdue University and University of Delaware

Marcel F./Neuts

University of Delaware

15 AFOSR-77-3236

16 2304

17 A5

18 AFOSR

14 TR-78/11

19 TR-78-1143

Department of Statistics  
and  
Computer Science

Technical Report, No. 78/11

D D C

RECEIVED  
SEP 25 1978  
REGULATED  
D

9

11 Jun 1978

12 25 p.

1. Research supported by AFOSR-72-2350C at the Department of Statistics, Purdue University and AFOSR-77-3236 at the Department of Statistics & Computer Science, University of Delaware, Newark, Delaware 19711.
2. Research supported by AFOSR-77-3236 at the Department of Statistics & Computer Science, University of Delaware, Newark, Delaware 19711.

78 07 20 152

DISTRIBUTION STATEMENT A  
Approved for public release;  
Distribution Unlimited

391 807

mt

#### ABSTRACT

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 hyper-exponential as special cases and form a versatile class with a number of interesting closure properties.

#### KEY WORDS

Queueing theory, duality, phase type distributions

## 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/1 and 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 hyper-exponential (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 (c.f. references in [5] and [8]).

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

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

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

$$H(x) = 1 - \underline{\alpha} \exp(Tx) \underline{e}, \quad x \geq 0 \quad (2.1)$$

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,  $\mu_1, \dots, \mu_m$  respectively has representation

$$T = \begin{matrix} \underline{\alpha} = (1, 0, \dots, 0) \\ \left[ \begin{array}{cccccc} -\mu_1 & \mu_1 & 0 & \dots & 0 \\ & -\mu_2 & \mu_2 & \dots & 0 \\ & & \ddots & & \vdots \\ & & & & -\mu_m \end{array} \right] \end{matrix}$$

Example 2.4: The hyper-exponential distribution which is the mixture of  $m$  exponentials with parameters, say,  $\mu_1, \dots, \mu_m$  has a representation  $T = \text{diag}(-\mu_1, \dots, -\mu_m)$  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  $1, \dots, m$  and probabilities  $\alpha_1, \dots, \alpha_m$  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 c.d.f.  $H(\cdot)$ , the PH-distribution given by (2.1). Such a renewal process is called a Renewal Process of Phase Type (PH-renewal process) (Neuts [5]).

The above procedure also constructively defines a new Markov chain with state-space  $\{1, \dots, m\}$ , initial probability vector  $\underline{\alpha}$  and infinitesimal generator  $Q^* = T + T^0 A^0$ , where  $A^0 = \text{diag}(\alpha_1, \dots, \alpha_m)$  and  $T^0 = (\underline{T}^0, \dots, \underline{T}^0)$ . 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  $\underline{e}$  denote the stationary probability vector of the Markov chain  $Q^*$ , i.e., the unique (strictly positive) vector satisfying  $\underline{e}Q^* = \underline{0}$ ,  $\underline{e}\underline{e} = 1$ . It may be easily verified that  $\underline{e} = -\lambda \underline{\alpha} T^{-1}$ , where  $\lambda^{-1} = -\underline{\alpha} T^{-1} \underline{e}$  is the mean of  $H(\cdot)$ .

The  $m \times m$  matrices  $P(v, t)$ ,  $v \geq 0$ ,  $t \geq 0$ , defined in [5] are such that the entry  $P_{jj'}(v, t)$  is the conditional probability, given that the initial phase is  $j$ , that at time  $t$ , the  $Q^*$ -chain is in state  $j'$  and that  $v$  renewals have occurred in  $(0, t]$ . It is known [5] that these have generating function

$$\tilde{P}(z, t) = \sum_{v=0}^{\infty} z^v P(v, t) = \exp[(T + zT^0 A^0)t], \quad |z| \leq 1, t \geq 0.$$

We also recall [6] that, under the assumption  $Q^*$  is irreducible, the matrices  $P(v, t)$ ,  $v \geq 1$ ,  $t > 0$ , are all strictly positive. Further, it may be easily seen that  $P_{ij}(0, t) > 0$  for all  $t \geq 0$ ,  $1 \leq i \leq m$ .

### 3. The Phase Type Queues

#### a. The GI/PH/1 Queue:

Consider a GI/G/1 queue in which the service time c.d.f.  $H(\cdot)$  is of phase type with representation  $(\underline{\alpha}, T)$  and where the inter-arrival times are i.i.d., with a non-degenerate probability distribution  $F(\cdot)$ . Such a model, denoted by GI/PH/1, has been discussed in detail by Neuts [6], and for later use we quote the following results.



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

$$\tilde{Q}(x) = \begin{bmatrix} \tilde{B}_0(x) & \tilde{A}_0(x) & \dots & \dots & \dots \\ \tilde{B}_1(x) & \tilde{A}_1(x) & \tilde{A}_0(x) & \dots & \dots \\ \tilde{B}_2(x) & \tilde{A}_2(x) & \tilde{A}_1(x) & \tilde{A}_0(x) & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}, \quad x \geq 0,$$

where,

$$\tilde{A}_n(x) = \int_0^x P(n, t) dF(t), \quad n \geq 0, \quad x \geq 0,$$

$$\tilde{B}_n(x) = \sum_{\nu=n+1}^{\infty} \int_0^x P(\nu, t) dF(t) A^{\circ\circ}, \quad n \geq 0, \quad x \geq 0,$$

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

Since  $F(\cdot)$  is non-degenerate, we have (from the strict positivity of the matrices  $P(\nu, t)$  for  $\nu \geq 1$ ,  $t > 0$  and the positivity of the entries  $P_{ij}(0, t)$  for  $t \geq 0$  mentioned at the end of Section 2) that the matrices  $A_n \stackrel{\text{def}}{=} \tilde{A}_n(\infty)$ ,  $n \geq 1$ , are all strictly positive; further, the matrix  $A_0 \stackrel{\text{def}}{=} \tilde{A}_0(\infty)$  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 [6] that this queue is stable iff  $\rho = \underline{\theta} \underline{\beta} > 1$ , where  $\underline{\theta}$  is (also) the invariant probability vector of  $A \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} A_n$ , and  $\underline{\beta} = \sum_{n=1}^{\infty} n A_n \underline{e}$ .

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 inter-arrival times and service times in the latter. The PH/G/1 queue is a sub-class 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 c.d.f. is given by  $F(\cdot)$ . Defining  $\hat{\tau}_n$  to be the epoch of the  $n$ -th departure ( $\tau_0=0$ ), 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  $\{(\hat{\xi}_n, \hat{j}_n, \hat{\tau}_{n+1} - \hat{\tau}_n) : n \geq 0\}$  is a semi-Markov sequence with state space  $\{0, 1, \dots\} \times \{1, \dots, m\} \times [0, \infty)$  and transition matrix

$$\hat{Q}(x) = \begin{bmatrix} \tilde{C}_0(x) & \tilde{C}_1(x) & \tilde{C}_2(x) & \dots & \dots \\ \tilde{A}_0(x) & \tilde{A}_1(x) & \tilde{A}_2(x) & \dots & \dots \\ 0 & \tilde{A}_0(x) & \tilde{A}_1(x) & \dots & \dots \\ 0 & 0 & \tilde{A}_0(x) & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}, \quad x \geq 0,$$

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

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

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  $\rho = \lambda \mu < 1$ . Further,  $\rho = \lambda \mu$ , where  $\mu$  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  $\underline{1} = \{(1, j) : 1 \leq j \leq m\}$  to the set of states  $\underline{0} = \{(0, j) : 1 \leq j \leq m\}$  which are governed by the matrix  $\tilde{G}(x)$ ,  $x \geq 0$ , 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

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

we have

Theorem 3.2:

(i) The matrix  $G(\cdot)$  satisfies the matrix functional equation

$$G(s) = \sum_{n=0}^{\infty} A_n(s) G^n(s), \quad s \geq 0, \quad (3.3)$$

where

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

(ii) For  $s > 0$ , there exists a unique non-negative 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

$$G = \sum_{n=0}^{\infty} A_n G^n \quad (3.4)$$

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

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

$$\left. \begin{aligned} G_0(s) &= 0 \\ G_{n+1}(s) &= \sum_{v=0}^{\infty} A_v(s) G_n^v(s), \quad n \geq 0 \end{aligned} \right\} \quad (3.5)$$

Proof: All the results above follow from Theorem 2.2.11 in [8] which in turn is proven by Neuts [7] 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  $\underline{i+1} = \{(i+1, j) : 1 \leq j \leq m\}$  to the set  $\underline{i} = \{(i, j) : 1 \leq j \leq m\}$  for any  $i \geq 0$ .
2. Note that the  $i$ -th entry of  $G(s)\mathbf{e}$  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  $\{(n, 1), \dots, (n, m)\}$ . 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  $x \geq 0$ ,

$n, k \geq 1$ ,  $\tilde{D}_{n,n+k}(x)$  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

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

Proposition 4.1: For  $n \geq 1$ ,  $k \geq 2$ , we have

$$\tilde{D}_{n,n+k} = \tilde{D}_{n,n+1} * \tilde{D}_{n+1,n+2} * \dots * \tilde{D}_{n+k-1,n+k}, \quad (4.2)$$

and

$$D_{n,n+k} = D_{n,n+1} D_{n+1,n+2} \dots D_{n+k-1,n+k}, \quad (4.3)$$

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  $\underline{n+1}$ , ..., from  $\underline{n+k-1}$  to  $\underline{n+k}$ . Given any sequence  $j_1, \dots, j_{k-1} \in \{1, \dots, m\}$  denoting the phases at the epochs of such first entrances to  $\underline{n+1}, \dots, \underline{n+k-1}$ , 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  $n \geq 1$ ,  $s \geq 0$ ,

$$D_{n,n+1}(s) = A_0(s) + \sum_{v=1}^{n-1} A_v(s) D_{n-v+1,n+1}(s) + \sum_{v=n}^{\infty} A_v(s) A^{o \circ} D_{1,n+1}(s) \quad (4.5)$$

(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

inter-arrival interval. Also let us denote by  $S$  the event that the process enters  $n+1$  at or before time  $x$  and that the phase at which it enters  $n+1$  is  $j$ . Then by a simple probabilistic argument, we have

$$P(S, N=v | \xi_0=n, J_0=i) =$$

$$\begin{cases} \int_0^x P_{ij}(0,u) dF(u), & \text{if } v=0 \\ \int_0^x \sum_{k=1}^m P_{ik}(v,u) \tilde{D}_{n-v+1, n+1}(k, j; x-u) dF(u), & \text{if } 1 \leq v \leq n-1 \\ \int_0^x \sum_{v=n}^{\infty} \sum_{\ell=1}^m P_{i\ell}(v,u) \sum_{k=1}^m \alpha_k \tilde{D}_{1, n+1}(k, j; x-u) dF(u), & \text{if } v=n, \end{cases}$$

whence

$$\tilde{D}_{n, n+1}(i, j; x) =$$

$$\begin{aligned} & \int_0^x P_{ij}(0,u) dF(u) + \sum_{v=1}^{n-1} \int_0^x \sum_{k=1}^m P_{ik}(v,u) \tilde{D}_{n-v+1, n+1}(k, j; x-u) dF(u) \\ & + \sum_{v=n}^{\infty} \int_0^x \sum_{\ell=1}^m P_{i\ell}(v,u) \sum_{k=1}^m \alpha_k \tilde{D}_{1, n+1}(k, j; x-u) dF(u). \end{aligned}$$

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),

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

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

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

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

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

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  $n, k \geq 1$ ,  $D_{n,n+k}(0)$  is stochastic. Also for  $n \geq 2$ ,  $k \geq 1$ ,  $D_{n,n+k}(0)$  is strictly positive.

Proof: In view of (4.3) it suffices to prove the results for  $k=1$ . Let  $n \geq 1$  and assume as inductive hypothesis that  $D_{1,2}(0), \dots, D_{n,n+1}(0)$  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

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

where

$$M = A_1 + \sum_{v=2}^n A_v D_{n+2-v,n+1}(0) + (A - \sum_{v=0}^n A_v)A^{\circ\circ}D_{1,n+1}(0)$$

(with the second term on the right being 0 when  $n=1$ ).

Now  $M \geq A_1 \gg 0$ , and  $M\underline{e} = \underline{e} - A_0\underline{e} \ll \underline{e}$ . Thus the strictly positive matrix  $M$  has spectral radius  $\rho$  less than 1. That

$D_{n+1,n+2}(0) \gg 0$ , follows now from (4.10) by writing it as

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

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

$$\begin{aligned} (1-n)\underline{y}D_{n+1,n+2}(0)\underline{e} &= \underline{y}A_0\underline{e} \\ &= \underline{y}(\underline{e} - M\underline{e}) = (1-n) \end{aligned}$$

and since  $n < 1$ ,  $\underline{y}D_{n+1,n+2}(0)\underline{e} = 1$ . 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

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

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

- a)  $D_{n,n+k}(s)\underline{e} \leq \underline{e}$
- b)  $D_{n,n+1}(s)\underline{e} \leq \phi(s)\underline{e}$
- c)  $D_{n,n+k}(s)\underline{e} \leq \{\phi(s)\}^k \underline{e}$

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  $n \rightarrow \infty$ , the matrices  $D_{n,n+1}(s)$  describing the first passage times from  $n$  to  $n+1$  in the GI/PH/1 queue converge to the matrix  $G(s)$ , where  $G(s)$  is the unique non-negative solution of the non-linear matrix functional equation

$$G(s) = \sum_{v=0}^{\infty} A_v(s) G^v(s) \quad (5.2)$$

and which describes the first passage times from  $1$  to  $0$  in the dual PH/G/1 queue. Further the limit  $G(0+)$  is stochastic iff  $\rho \leq 1$ , 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  $\overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)$  and  $\underline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)$  denote the matrices whose respective  $(i,j)$ -th entries are  $\overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(i,j;s)$  and  $\underline{\lim}_{n \rightarrow \infty} D_{n,n+1}(i,j;s)$ . From Proposition 4.11(b) it is clear that these matrices are strictly substochastic; further,

$$0 \leq \underline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s) \leq \overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s) \quad (5.3)$$

Lemma 5.4: Define

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

$$M_{n+1}(s) = A_0(s) + \sum_{v=1}^n A_v(s) \prod_{k=n-v+2}^{n+1} M_k(s), \quad n \geq 1$$

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  $n \geq 1$ ,  
(ii)  $M_n(s)$  is entrywise non-decreasing in  $n$ ,  
(iii)  $\lim_{n \rightarrow \infty} M_n(s) = G(s)$ , 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  $\sum_{v=0}^{\infty} A_v(s) \underline{e} < \underline{e}$ , there exist constants  $w(s) > 0$  and  $\delta(s) > 0$  such that  $w(s) + \delta(s) < 1$ ,  $\sum_{v=1}^{\infty} A_v(s) \underline{e} \leq w(s) \underline{e}$  and  $A_0(s) \underline{e} \leq \delta(s) \underline{e}$ . Now,

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

whence

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

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

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

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

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

$$\begin{aligned} M_{n+2}(s) \underline{e} &= \sum_{v=0}^{\infty} \{A_1(s) + A_2(s)M_{n+1}(s) + \dots + A_{n+1}(s)M_2(s) \dots M_{n+1}(s)\}^v A_0(s) \underline{e} \\ &\leq \frac{\delta(s)}{1-w(s)} \underline{e} < \underline{e} \end{aligned}$$

by using the substochastic of  $M_2(s), \dots, M_{n+1}(s)$  assumed as

inductive hypothesis.

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

$$\begin{aligned} M_{n+1}(s) &= A_0(s) + A_1(s)M_{n+1}(s) + \dots + A_n(s)M_2(s) \dots M_{n+1}(s) \\ &\leq \sum_{v=0}^n A_v(s)M_{n+1}^v(s) \\ &\leq \sum_{v=0}^{\infty} A_v(s)M^v(s) \end{aligned}$$

Letting  $n \rightarrow \infty$  this yields

$$M(s) \leq \sum_{v=0}^{\infty} A_v(s)M^v(s) \quad (5.5)$$

Also

$$M_{n+m}(s) \geq A_0(s) + A_1(s)M_{n+m}(s) + \dots + A_n(s)M_{m+1}(s) \dots M_{m+n}(s)$$

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

$$M(s) \geq \sum_{v=0}^n A_v(s)M^v(s), \text{ for all } n \geq 0.$$

Since the above inequality holds for every  $n$ ,

$$M(s) \geq \sum_{v=0}^{\infty} A_v(s)M^v(s) \quad (5.6)$$

Now, by (5.5) and (5.6),

$$M(s) = \sum_{v=0}^{\infty} A_v(s)M^v(s),$$

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:

$$\lim_{n \rightarrow \infty} D_{n,n+1}(s) \geq G(s) \quad (5.8)$$

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

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

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

$$R(s) = \overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)$$

Then

$$R(s) \leq \sum_{v=0}^{\infty} A_v(s) R^v(s)$$

Proof: For sufficiently large  $n$  and fixed  $N < n-1$ , we have,

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

and letting  $n \rightarrow \infty$ ,

$$R(s) \leq A_0(s) + \sum_{v=1}^N A_v(s) R^v(s) + \overline{\lim}_{n \rightarrow \infty} \sum_{v=N+1}^{n-1} A_v(s) D_{n-v+1,n+1}(s),$$

for,

$$\overline{\lim}_{n \rightarrow \infty} \sum_{v=n}^{\infty} A_v(s) A^{\circ \circ} D_{1,n+1}(s) = 0.$$

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

$$R(s) \leq \sum_{v=0}^{\infty} A_v(s) R^v(s),$$

since,

$$\lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \sum_{v=N+1}^{n-1} A_v(s) D_{n-v+1, n+1}(s) = 0$$

as is seen from the fact that

$$\sum_{v=N+1}^{n-1} A_v(s) D_{n-v+1, n+1}(s) \underline{e} \leq \sum_{v=N+1}^{\infty} A_v(s) \underline{e} \rightarrow 0 \text{ as } N \rightarrow \infty.$$

Corollary 5.10: Denoting the maximal eigenvalue of a non-negative matrix  $C$  by  $sp(C)$ , we have that

$$sp[\overline{\lim} D_{n, n+1}(s)] \leq sp[G(s)] \quad (5.11)$$

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  $R(s) = \overline{\lim}_{n \rightarrow \infty} D_{n, n+1}(s)$ , and let  $\underline{x}(s) \gg 0$  be a right-eigenvector of  $R(s)$  associated with  $\eta(s)$ . Then

$$\begin{aligned} \eta(s) \underline{x}(s) &= R(s) \underline{x}(s) \\ &\leq \sum_{v=0}^{\infty} A_v(s) R^v(s) \underline{x}(s) && \text{by Lemma 5.9} \\ &= \sum_{v=0}^{\infty} A_v(s) \eta^v(s) \underline{x}(s) \end{aligned}$$

implying that

$$A^*(\eta(s), s) \underline{x}(s) \geq \eta(s) \underline{x}(s) \quad (5.12)$$

where

$$A^*(z, s) = \sum_{v=0}^{\infty} A_v(s) z^v, \quad 0 \leq z \leq 1$$

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

$$\xi(n(s), s) \geq n(s) \quad (5.13)$$

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  $t \geq 0$ . Also  $G(s)$  is the unique solution of equation (5.2) whence it follows that  $\text{sp}[G(s)]$  is the unique solution  $z_0$  of the equation

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

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

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

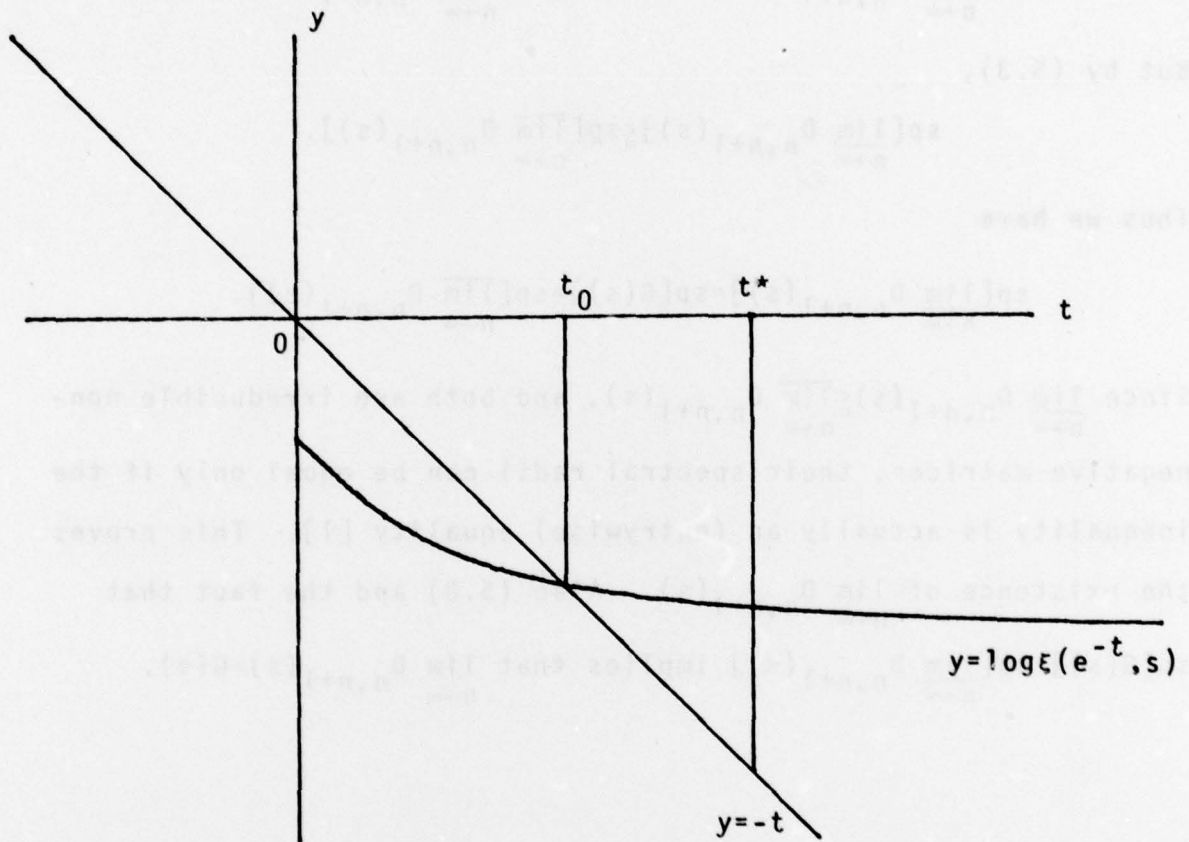


Figure 1

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

Proof of Theorem 5.1: By Formulas (5.8) and (5.11), we have

$$\text{sp}[\lim_{n \rightarrow \infty} D_{n,n+1}(s)] \geq \text{sp}[G(s)] \geq \text{sp}[\overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)]$$

But by (5.3),

$$\text{sp}[\lim_{n \rightarrow \infty} D_{n,n+1}(s)] \leq \text{sp}[\overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)].$$

Thus we have

$$\text{sp}[\lim_{n \rightarrow \infty} D_{n,n+1}(s)] = \text{sp}[G(s)] = \text{sp}[\overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)].$$

Since  $\lim_{n \rightarrow \infty} D_{n,n+1}(s) \leq \overline{\lim}_{n \rightarrow \infty} D_{n,n+1}(s)$ , and both are irreducible non-negative matrices, their spectral radii can be equal only if the inequality is actually an (entrywise) equality [1]. This proves the existence of  $\lim_{n \rightarrow \infty} D_{n,n+1}(s)$ . Also (5.8) and the fact that  $\text{sp}[G(s)] = \text{sp}[\lim_{n \rightarrow \infty} D_{n,n+1}(s)]$  implies that  $\lim_{n \rightarrow \infty} D_{n,n+1}(s) = G(s)$ .



BIBLIOGRAPHY

- [1] Gantmacher, F. R. (1959)  
The Theory of Matrices (vol. 1)  
Chelsea, New York, NY
- [2] Heathcote, C. R. (1965)  
On the Maximum of the Queue GI/M/1  
J. Appl. Prob., 2, pp 206-214
- [3] Kingman, J. F. C. (1961)  
A Convexity Property of Positive Matrices  
Quart. J. Math., 12, pp 283-284
- [4] Neuts, M. F. (1975)  
Probability Distributions of Phase Type  
--in Liber Amicorum Professor Emeritus H. Florin--  
Dept. of Math., University of Louvain, Belgium, pp 173-206
- [5] Neuts, M. F. (1976)  
Renewal Processes of Phase Type  
Nav. Res. Log. Quart (forthcoming)
- [6] Neuts, M. F. (1976)  
Markov Chains with Applications in Queueing Theory  
with have a Matrix-geometric Invariant Probability Vector  
Adv. Appl. Prob. (forthcoming)
- [7] Neuts, M. F. (1976)  
Moment Formulas for the Markov Renewal Branching Process  
Adv. Appl. Prob., 8, pp 690-711
- [8] Ramaswami, V. (1978)  
The N/G/1 Queue and its Detailed Analysis  
Tech. Report #78/1, Dept. of ST & CS, University of  
Delaware, Newark, DE 19711

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 78-1143	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A DUALITY THEOREM FOR PHASE TYPE QUEUES	5. TYPE OF REPORT & PERIOD COVERED Interim	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) V. Ramaswami and Marcel F. Neuts	8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3236	
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Delaware Department of Statistics & Computer Science Newark, Delaware 19711	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5	
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332	12. REPORT DATE June 1978	
	13. NUMBER OF PAGES 23	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Queueing theory, duality, phase type distributions		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  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 hyper-exponential as special cases and form a versatile class with a number of interesting closure properties.		