*The Annals of Probability*
1982, Vol. 10, No. 1, 265–269

# CLOSURE OF PHASE TYPE DISTRIBUTIONS UNDER OPERATIONS ARISING IN RELIABILITY THEORY

BY DAVID ASSAF AND BENNY LEVIKSON

*Hebrew University and University of Haifa*

Consider three basic operations arising in reliability theory: finite mixtures, finite convolutions, formation of coherent systems of independent components. In this paper it is shown that: (i) The class of all phase type distributions is closed under all three operations; (ii) the class of all phase type distributions having a representation in which the matrix is upper triangular is the smallest class which is closed under all three operations and contains all exponential distributions.

The paper includes a section of preliminaries in which the relevant material regarding phase type distributions and reliability theory is summarized.

**1. Introduction and summary.** As is well known to workers in reliability theory, several classes of distributions are closed under various reliability operations. Notably, the DFR class is closed under mixtures, the IFR class is closed under convolutions, and the IFRA class is closed both under convolutions and under the formation of coherent systems of independent components (see Barlow and Proschan (1975)—Chapter 4, and Block and Savits (1976)). It should be noted, however, that none of the above classes is closed under all three operations (mixtures, convolutions, and formation of coherent systems). Thus, for example, if the life distribution of each component is a mixture of exponentials (e.g., different manufacturers), then the life distribution of a coherent system composed of such components does not necessarily belong to any of the above classes.

In Section 3 it is shown that the class of all *phase type* (PH) distributions (introduced by Neuts (1975)) is closed under all three operations. Furthermore, the smallest class of distributions which contains all exponentials and is closed under finite mixtures, finite convolutions, and the formation of coherent systems, is characterized (Theorem 5) as the class of all PH distributions having an upper-diagonal-matrix representation.

Section 2 contains preliminary results and is composed of two parts. The first summarizes basic definitions and notation from reliability theory needed in this paper. The second part summarizes the basic ideas of the theory of PH distributions. In particular, all PH distributions possess an exponential matrix representation (see formula 2.3). Some advantages and applications of this representation to other fields are indicated.

**2. Preliminaries**

(a) *Basic Definitions and Notation from Reliability Theory.* Consider $n$ independent components (the dependent case is not considered in this paper), each of which is either functioning or not. Let

$$x_i = \begin{cases} 1 & i\text{th component is functioning} \\ 0 & \text{otherwise} \end{cases}$$

A system composed of these components is specified by a *structure function* $\phi$: $\{0, 1\}^n$

Received May 1, 1980; revised January 19, 1981.

*AMS* 1970 *subject classifications.* Primary 60K10; secondary 60J25.

*Key words and phrases.* Phase type distributions, Markov process, closure theorems in reliability theory, coherent systems.

$\rightarrow \{0, 1\}$ defined by

$$\phi(x_1, \cdots, x_n) = \begin{cases} 1 & \text{system is functioning} \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

We consider *monotone* systems, that is, systems for which $\phi(x_1, \cdots, x_n) \geq \phi(y_1, \cdots, y_n)$ whenever $x_i \geq y_i$ for all $1 \leq i \leq n$. A system is *coherent* if, in addition to monotonicity, it has no irrelevant components. (A component is irrelevant if it never affects functioning of the system.)

Denote the life distribution of the $i$th component by $F_i(t)$, then the life distribution of the system, $F_\phi(t)$, is given by

$$F_\phi(t) = 1 - h_\phi(\bar{F}_1(t), \cdots, \bar{F}_n(t)) \qquad (\bar{F} = 1 - F)$$

where $h_\phi: [0, 1]^n \rightarrow [0, 1]$ is defined by

$$h_\phi(p_1, \cdots, p_n) = E\phi(X_1, \cdots, X_n) \tag{2.2}$$

$X_1, \cdots, X_n$ are independent such that $X_i = 1$ with probability $p_i$ and $X_i = 0$ with probability $1 - p_i$.

(b) *On Phase Type Distributions.* Consider an $(m + 1)$-state continuous parameter Markov chain having a single absorbing state (which, as a convention, is taken to be the state $m + 1$). Assume all non-absorbing states 1 through $m$ are transient and eventual absorption at $m + 1$, starting at any other state is certain. A process $\{V(t); t \geq 0\}$ of this type, $V(t)$ being the state at time $t$, will be called an absorbing-continuous-parameter-Markov-chain (abbreviated to ACPMC). Denote by $P = (p_{ij})$ the $(m + 1) \times (m + 1)$ *transition matrix* of the embedded Markov chain, and by $\lambda = (\lambda_1, \cdots, \lambda_{m+1})$, $\lambda_{m+1} = 0$ the *transition rate vector* (i.e., the vector of parameters of the exponential holding times). Thus, whenever in state $i$, $1 \leq i \leq m$, the process stays in $i$ for an exponential length of time (with mean $\lambda_i^{-1}$) and then moves to its new state $j (j \neq i, 1 \leq j \leq m + 1)$ with probability $p_{ij}$. (Naturally, $p_{ii} = 0$ for $i \neq m + 1$). We find it convenient to denote by $Q$ the $m \times m$ submatrix obtained from $P$ by deleting the last row and column. The *infinitesimal generator* of the process is the $(m + 1) \times (m + 1)$ matrix $G = \begin{pmatrix} A & \beta \\ \mathbf{0} & 0 \end{pmatrix}$ where $A = (a_{ij})$ is the $m \times m$ matrix given by

$$a_{ij} = \begin{cases} \lambda_i p_{ij} & i \neq j \\ -\lambda_i & i = j \end{cases}$$

and $\beta$ is the $m$ vector $\beta = -A\mathbf{e}$ ($\mathbf{e} = (1, \cdots, 1)'$).

Let $\alpha_i$ be the (initial) probability of starting at state $i$. Assume $\alpha_{m+1} = 0$ and denote $\alpha = (\alpha_1, \cdots, \alpha_m)$. The distribution of the time till absorption is the solution of a system of linear differential equations which yields

$$F(x) = 1 - \alpha \exp(xA)\mathbf{e}. \tag{2.3}$$

A probability distribution is said to be of (continuous) *phase type* (PH) if it has a representation as the time till absorption in an ACPMC. The notation PH will be used to denote both the phase type property and the class of all distributions possessing the PH property. It should be noted that the representation of a PH distribution is not unique (different processes may give rise to the same distribution of time till absorption).

The PH family is quite large and contains, for example, all exponential distributions (take $m = 1$, $A = (-\lambda)$ and $\alpha = (1)$), the generalized Erlang distributions and the hyperexponential distributions. The Laplace-Stieltjies transform and the moments of PH distributions all have convenient closed forms. PH distributions have been successfully applied in the theory of queues, renewal theory and the Galton-Watson process. For these as well as other properties and applications the reader is referred to the papers by Neuts (1975, 1977, 1978, 1979) and by Assaf (1982).

## 3. Closure Theorems.

THEOREM 1. (Neuts, 1975). *The class* PH *is closed under finite mixtures and finite convolutions.*

THEOREM 2. *The class* PH *is closed under the formation of coherent systems.*

The following lemmas are needed:

LEMMA 1. *Let* $V_1(t), \cdots, V_n(t)$ *be independent* ACPMC's, $V_i(t)$ *having states* 1 *through* $m_i + 1$, $1 \le i \le n$. *Then the vector process* $\mathbf{V}(t) = (V_1(t), \cdots, V_n(t))$, *with state space* $S = \{s = (v_1, \cdots, v_n); 1 \le v_i \le m_i + 1, 1 \le i \le n\}$ *is an* ACPMC.

PROOF. Let $V_i$ have initial probability vector $\boldsymbol{\alpha}_i$, transition rate vector $\boldsymbol{\lambda}_i$, and transition matrix $P_i$, $1 \le i \le n$. More specifically, let

$$\boldsymbol{\alpha}_i = (\alpha_{i,1}, \cdots, \alpha_{i,m_i}), \qquad \boldsymbol{\lambda}_i = (\lambda_{i1}, \cdots, \lambda_{i,m_i+1})(\lambda_{i,m_i+1} = 0),$$

and

$$P_i = (P_i(k, \ell)) \qquad 1 \le k, \quad \ell \le m_i + 1.$$

We now show that $\mathbf{V}(t)$ is an ACPMC with elements (indexed by *) given by initial probability vector $\boldsymbol{\alpha}^*$ with coordinates

$$\alpha^*(s) = \begin{cases} \prod_{i=1}^n \alpha_{i,v_i} & \text{if } s = (v_1, \cdots, v_n) \text{ with } 1 \le v_i \le m_i \text{ all } 1 \le i \le n \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

and by transition rate vector given by its coordinates as

$$\lambda^*(s) = \textstyle\sum_{i=1}^n \lambda_{i,v_i} \qquad \text{for} \quad s = (v_1, \cdots, v_n) \quad 1 \le v_i \le m_i + 1 \tag{3.2}$$

(Here we use the convention $\lambda_{i,}\ m_i + 1 = 0$ for all $1 \le i \le n$).

The submatrix $Q^*$ of $P^*$ (see Section 2(b)) is given by its elements as $Q^*(s, s') = 0$ if $s$ and $s'$ differ in more than one coordinate and by

$$Q^*(s, s') = \frac{\lambda_{i,v}}{\lambda^*(s)} P_i(v, v') \tag{3.3}$$

for $s = (v_1, \cdots, v_{i-1}, v, v_{i+1}, \cdots, v_n)$ and $s' = (v_1, \cdots, v_{i-1}, v', v_{i+1}, \cdots, v_n)$.

The absorbing state of the process $\mathbf{V}(t)$ is clearly the state $(m_1 + 1, \cdots, m_n + 1)$.

Verification of these results is based on standard techniques and is thus only indicated. (3.1) follows directly from the independence of $V_1(0), \cdots, V_n(0)$. Now, whenever in state $s = (v_1, \cdots, v_n)$, a transition to a new state occurs as soon as transition occurs in one of the coordinates (the probability of a transition in two or more coordinates at the same time is zero). Thus, the time until transition is the minimum of $n$ independent exponential r.v.'s and thus has an exponential distribution with parameter given by (3.2). Recall that if $X_1, \cdots, X_n$ are independent such that each $X_i$ is exponential with parameter $\lambda_i$, then

$$P(X_i < X_j \text{ for all } j \neq i \mid \operatorname{Min}(X_1, \cdots, X_n) = t) = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}.$$

The probability of a direct transition from $s = (v_1, \cdots, v_{i-1}, v, v_{i+1}, \cdots, v_n)$ to $s' = (v_1, \cdots, v_{i-1}, v', v_{i+1}, \cdots, v_n)$ is thus the probability of transition in the $i$th coordinate, given by $\frac{\lambda_{i,v}}{\lambda^*(s)}$, multiplied by the probability of transition from $v$ to $v'$ in the $i$th process, given by $P_i(v, v')$. Hence (3.3) follows. □

LEMMA 2. *Let $V(t)$ be an* ACPMC *with states* 1 *through* $m + 1$, *initial probability vector* $\alpha$, *transition rate vector* $\lambda$ *and transition submatrix* $Q$. *Let* $E \subset \{1, 2, \cdots, m + 1\}$ *be a set such that* $m + 1 \in E$ *and* $\alpha_i = 0$ *for all* $i \in E$. *Then the distribution of the hitting time of* $E$ *is of* PH.

PROOF. Let $D = E - \{m + 1\}$. Define a new ACPMC with $k + 1$ states ($k = m - |D|$) whose elements are obtained from $\alpha$, $\lambda$ and $Q$ by deleting all coordinates of $\alpha$ and $\lambda$, and all rows and columns of $Q$ corresponding to states in $D$. The hitting time of $E$ in the original process is the same as the time till absorption in the new ACPMC and thus has a PH distribution. □

PROOF OF THEOREM 2. Let $\phi$ be the structure function of the coherent system. Let the life distribution of the $i$th component be of PH with the notation of Lemma 1. For the vector process $\mathbf{V}(t) = (V_1(t), \cdots, V_n(t))$ and its state space $S$, define the transformation

$$T: S \rightarrow \{0, 1\}^n$$

by

$$T(v_1, \cdots, v_n) = (\epsilon_1, \cdots, \epsilon_n)$$

where $\epsilon_i = 1$ if $1 \leq v_i \leq m_i$ and $\epsilon_i = 0$ if $v_i = m_i + 1$. Let $E \subset S$ be the set defined by

$$E = \{s \in S; \phi(T(s)) = 0\}. \tag{3.4}$$

Clearly $(m_1 + 1, \cdots, m_n + 1) \in E$ and $\alpha^*(s) = 0$ for all $s \in E$ (assuming naturally that all components, as well as the system, are initially functioning). From Lemma 2 it follows that the distribution of the hitting time of $E$ is of PH. By the definition of $E$ it follows that the hitting time of $E$ is the same as the failure time of the system. Thus the distribution of the time until failure of the system is of PH. □

REMARK. Suppose a coherent system whose components are PH is replaced by a new system whenever it fails. Then, the limiting distribution of the age of the current system, as well as its remaining life distribution, are of PH. This follows from Theorem 2 of this paper and from Theorem 8 of Neuts (1975).

DEFINITION. A PH distribution is called Triangular Phase Type (TPH) if it has a representation in which the matrix $Q$ is strictly upper triangular.

As for the term PH, we adapt the convention that TPH denotes both the property and its corresponding class. It should be noted that TPH is the subclass of PH for which absorption is certain within a bounded number of transitions.

THEOREM 3. TPH *is the smallest class which contains all exponential distributions and is closed under finite mixtures and convolutions.*

PROOF. An exponential distribution is clearly TPH. By the constructions of Neuts (1975) for mixtures and convolutions for PH distributions in general, it follows that triangularity is preserved by these operations as well. To complete the proof we now show that TPH is the *smallest* class with the above property. Let $i_1 < i_2 \leq \cdots \leq i_K < m + 1$ define a *path* in which the process starts at state $i_1$, moves to $i_2$ etc., until it reaches the absorbing state $m + 1$ after $K$ transitions. The lifetime of such a path is distributed as a sum of $K$ independent r.v.'s $X(i_1) + X(i_2) + \cdots + X(i_K)$ where $X(i_j)$ is exponential with parameter $\lambda_{i_j}$. Since, by triangularity, there is a finite number of possible paths, it follows that the time till absorption can be represented as a finite mixture of finite sums of independent exponential r.v.'s □

THEOREM 4. *The class* TPH *is closed under the formation of coherent systems.*

PROOF. Following the construction and notation in Theorem 2, it is sufficient to show that *triangularity* is preserved. Arrange the states $s = (v_1, \cdots, v_n)$ by a lexicographic ordering. Since a one stage transition is possible only from $(v_1, \cdots, v_{i-1}, v, v_{i+1}, \cdots, v_n)$ to $(v_1, \cdots, v_{i-1}, v', v_{i+1}, \cdots, v_n)$ for some $1 \le i \le n$ and some $v'$, and since $v' \ge v$ by triangularity of the representations for each component, it follows that transitions of the system are possible only to higher states. Thus $Q^*$ is triangular. Now, any matrix which is obtained from a triangular matrix by deleting corresponding rows and columns (as the procedure of Lemma 2 requires) is also triangular. Thus the representation of the system is by a triangular matrix. □

An immediate result of Theorems 3 and 4 is

THEOREM 5. TPH *is the smallest class containing all exponential distributions which is closed under finite mixtures, finite convolutions and the formation of coherent systems.*

**Acknowledgement.** The authors would like to thank a referee for his comments, which resulted in an improved presentation of Theorem 2.

## REFERENCES

[1] ASSAF, D. (1982). Renewal decisions for phase-type distributions. To appear in *Oper. Res*
[2] BARLOW, R. E., and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing. Probability Models.* Holt Rinehart and Winston, New York.
[3] BLOCK, H. W., and SAVITS, T. H. (1976). The IFRA Closure Problem. *Ann. Probability* 4 1030–1032.
[4] NEUTS, M. F. (1975). Probability distributions of phase-type. *Liber Amicorum Professor Emeritus H. Florin.* University of Louvain, 173–206.
[5] NEUTS, M. F. (1977). The M/G/1 Queue with several types of customers and change-over times. *Adv. Appl. Probability* **9**, 604–644.
[6] NEUTS, M. F. (1978). Renewal Processes of phase-type. *Naval Res. Logist. Quart.* **25** 445–454.
[7] NEUTS, M. F. (1979). A versatile Markovian point process. *J. Appl. Probability* **16** 764–779.

HEBREW UNIVERSITY
JERUSALEM
ISRAEL

UNIVERSITY OF HAIFA
MOUNT CARMEL
HAIFA 31 999,
ISRAEL