The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 1020–1029

FUNCTIONS OF FINITE MARKOV CHAINS AND EXPONENTIAL TYPE PROCESSES

By Paul E. Boudreau

IBM, Research Triangle Park, North Carolina

1. Introduction and notation. When is an arbitrary random process, $Y(t)$, equal, in joint distribution, to a function of a Markov chain?

When is a function of a Markov chain, $f[X(t)]$, itself a Markov chain?

This paper is devoted to the above questions when $Y(t)$ is an exponential type process [13], p. 207, and the Markov chain, $X(t)$ is a basic Markov chain [13], p. 207. The structure of an exponential type process of order K [13], p. 208 is analyzed.[1] A necessary and sufficient condition for an exponential type process of order K to be a function of a basic Markov chain with K states (Theorem 3.1) and a necessary and sufficient condition for an exponential type process to be a Markov chain (Theorem 4.2) are established.

If $\Phi = \{i_1, i_2, \cdots, i_n\}$ is a finite sequence of states of a random process $Z(t)$ and $S = \{s_1, s_1 + s_2, \cdots, s_1 + s_2 + \cdots + s_r\}$ is a corresponding monotone sequence of times, then the pair $(\Phi; S)$ is termed a *sequence pair of length n* for the process $Z(t)$. We denote the joint probabilities by:

$$P_Z(\Phi; S) = \Pr[Z(\tau_j) = i_j \text{ for } 1 \leqq j \leqq n]$$

where $\tau_j = \sum_{i=1}^{j} s_j$.

2. The structure of an exponential type process. Let $Y(t)$ be an exponential type process of order K, with state space $\mathfrak{M} = \{1, 2, \cdots, M\}$. The joint probabilities for $Y(t)$ are given by:

$$P_Y(\Phi; S) = \mathbf{b}[\textstyle\prod_{j=1}^{n} e^{Ds_j} B(i_j)]\mathbf{c}', \tag{2.1}$$

where $\mathbf{b} = (b_k)$ is a K-vector of the form $b_1 = 1$, $b_k = 0$ or 1 for $2 \leqq k \leqq K$, $D = \operatorname{diag}\{0 = \nu_1, \nu_2, \cdots, \nu_k\}$ is a $K \times K$ diagonal matrix, $B(m) = (b_{\alpha\beta}(m))$ for $1 \leqq m \leqq M$ are the $K \times K$ matrices appearing in the definition of exponential type, and $\mathbf{c}' = (1, 0, \cdots, 0)'$ is transpose of the K-vector $(1, 0, \cdots, 0)$.

A set of M, $K \times K$ matrices $R(1), R(2), \cdots, R(M)$ is termed a set of *factor* matrices provided that $R(m)R(m) = R(m)$ for $1 \leqq m \leqq M$, $R(k)R(m)$ is the zero matrix whenever $k \neq m$ and $\sum_{m=1}^{M} R(m) = I$, the $K \times K$ identity matrix. The first result here is that the M matrices associated with an exponential type process of order K constitute a set of factor matrices.

Received 31 October 1966; revised 27 September 1967.

[1] The results reported herein are valid under more general definitions of exponential type process and basic Markov chain than those used by Leysieffer [13]. We only require that the ν's (eigenvalues) be distinct complex numbers with non-positive real parts and $\nu_1 = 0$. Also the restriction that all initial probabilities be non-zero is replaced with the trivial requirement that the state spaces be the "essential" state space. That is, we assume that if m is a state of the process $Z(t)$ then for some $t \geqq 0$, $\Pr[Z(t) = m] > 0$.

LEMMA 2.1. *Let $\nu_1, \nu_2, \cdots, \nu_K$ be distinct complex numbers, $n \geqq 1$ a fixed integer, and $b_{\alpha\beta}$ fixed constants. Assume that for all $0 \leqq s_1, s_1 + s_2, \cdots, s_1 + s_2 + \cdots + s_n$,*

$$\sum_{\alpha_1=1}^{K} \sum_{\alpha_2=1}^{K} \cdots \sum_{\alpha_n=1}^{K} b_{\alpha_1\alpha_2} b_{\alpha_2\alpha_3} \cdots b_{\alpha_n\beta}[\exp(\textstyle\sum_{j=1}^{n} \nu_{\alpha_j} s_j)] = 0 \tag{2.2}$$

then all coefficients in (2.2) vanish, that is $b_{\alpha_1\alpha_2} b_{\alpha_2\alpha_3} \cdots b_{\alpha_n\beta} = 0$ for all $1 \leqq \alpha_1, \alpha_2, \cdots, \alpha_n \leqq K$.

Fix an integer α, $1 \leqq \alpha \leqq K$ and a sequence pair $(\Phi; S)$. If all coefficients in (2.1) which involve any $\alpha_j = \alpha$ vanished, then (2.1) would not involve ν_α. Furthermore, if this were true for all sequence pairs, (2.1) would be a representation of the required type for $P_Y(\Phi; S)$ involving only $K - 1$ of the ν's and $Y(t)$ would not be of order K. This, rather trivial observation, leads to the following necessary condition for an exponential type process to be of order K.

LEMMA 2.2. *If $Y(t)$ is an exponential type process of order K, then for each integer μ, $1 \leqq \mu \leqq K$ there is an integer $n \geqq 1$, a sequence of subscripts $1 \leqq \alpha_1, \alpha_2, \cdots, \alpha_k = \mu, \cdots, \alpha_n \leqq K$ and a sequences of states $1 \leqq i_1, i_2, \cdots, i_n \leqq M$ such that*

$$b_{\alpha_1} b_{\alpha_1\alpha_2}(i_1) \cdots b_{\alpha_{k-1}\mu}(i_{k-1}) b_{\mu\alpha_{k+1}}(i_k) \cdots b_{\alpha_n 1}(i_n) \neq 0. \tag{2.3}$$

Note that Lemma 2.2 insures that for given integers $1 \leqq \gamma, \mu \leqq K$, we can choose states and subscripts such that $b_{\alpha_1} b_{\alpha_1\alpha_2}(i_1) \cdots b_{\alpha_{k-1}\gamma}(i_{k-1}) \neq 0$ and $b_{\mu\alpha_{k+1}}(i_k) \cdots b_{\alpha_n 1}(i_n) \neq 0$.

THEOREM 2.3. *Let $Y(t)$ be an exponential type process of order K. Then the matrices $B(1), \cdots, B(M)$ form a set of factor matrices.*

PROOF. Let $i_1, i_2, \cdots, i_l = m, \cdots, i_n$ be a sequence of states and $\tau_1, \tau_2, \cdots, \tau_l, \cdots, \tau_n$ be a corresponding sequence of times, where $\tau_j = \sum_{l=1}^{j} s_l$. We have the following marginal probability relation:

$$\sum_{m=1}^{M} \Pr[Y(\tau_j) = i_j \text{ for } 1 \leqq j \leqq n] = \Pr[Y(\tau_j) = i_j \text{ for } 1 \leqq j \neq l \leqq n]. \tag{2.4}$$

Using (2.1) and letting $A = (a_{\gamma\mu}) = \sum_{m=1}^{M} B(m)$, the above equation becomes:

$$\begin{aligned} &\mathbf{b}[\textstyle\prod_{j=1}^{l-1} e^{Ds_j} B(i_j)] e^{Ds_l} A [\prod_{h=l+1}^{n} e^{Ds_h} B(i_h)]\mathbf{c}' \\ &\qquad = [\textstyle\prod_{j=1}^{l-1} e^{Ds_j} B(i_j)] e^{D(s_l + s_{l+1})} B(i_{l+1}) [\prod_{h=l+2}^{n} e^{Ds_h} B(i_h)]\mathbf{c}'. \end{aligned} \tag{2.5}$$

Expanding both sides of (2.5) and using the linear independence to compare coefficients in the resulting expanded form we obtain the identity:

$$\begin{aligned} &b_{\alpha_1} b_{\alpha_1\alpha_2}(i_1) \cdots b_{\alpha_{l-1}\gamma}(i_{l-1}) a_{\gamma\mu} b_{\mu\alpha_{l+2}}(i_{l+1}) \cdots b_{\alpha_n 1}(i_n) \\ &\qquad = \delta_{\gamma\mu} b_{\alpha_1} b_{\alpha_1\alpha_2}(i_1) \cdots b_{\alpha_{l-1}\mu}(i_{l-1}) b_{\mu\alpha_{l+2}}(i_{l+1}) \cdots b_{\alpha_n 1}(i_n), \end{aligned} \tag{2.6}$$

where $\gamma = \alpha_l$, $\mu = \alpha_{l+1}$. Since $Y(t)$ is of order K, we can invoke Lemma 2.2 to conclude that $a_{\gamma\mu} = \delta_{\gamma\mu}$ for $1 \leqq \gamma, \mu \leqq K$. That is, $A = I$.

Let $\tau_1, \tau_2, \cdots, \tau_l, \tau_{l+1}, \cdots, \tau_n$ be a sequence of times and consider the joint probability.

$$\begin{aligned} &\Pr[Y(\tau_j) = i_j \quad \text{for} \quad 1 \leqq j \leqq l - 1, \quad Y(\tau_l) = k, \quad Y(\tau_{l+1}) = m, \\ &\qquad Y(\tau_h) = i_h \quad \text{for} \quad l + 2 \leqq h \leqq n]. \end{aligned} \tag{2.7}$$

If $k \neq m$ all the probabilities expressed by (2.7) must be zero when $\tau_l = \tau_{l+1}$ (i.e. when $s_{l+1} = 0$). Utilizing this fact and essentially repeating the previous argument yields the fact that $B(k)B(m)$ is the zero matrix when ever $k \neq m$. These two results can now be easily applied to show that $B(m)B(m) = B(m)$, and the result follows.

Suppose for a given exponential type process $Y(t)$ of order K, we have a second representation of the form (2.1):

$$P_Y(\Phi; S) = \mathbf{a}[\textstyle\prod_{j=1}^{n} e^{Fs_j} A(i_j)]\mathbf{c}' \tag{2.8}$$

where $F = \text{diag}\ \{0 = \omega_1, \omega_2, \cdots, \omega_K\}$ with $\omega_1, \omega_2, \cdots, \omega_K$ distinct complex numbers and $\text{Re}\ (\omega_j) < 0$ for $j > 1$. Writing (2.1) and (2.8) in expanded form and using the linear independence and Lemma 2.2 one can easily show that the set $\{0 = \omega_1, \omega_2, \cdots, \omega_K\}$ equals the set $\{0 = \nu_1, \nu_2, \cdots, \nu_K\}$. The conclusion is that the ν's are unique and the matrix D is essentially unique. If, for example, we insist that the ν's are ordered in some fashion (say lexicographically) then D is unique. Consequently a simple reordering procedure (via an appropriate permutation matrix) allows us to assume $D = F$ in equation (2.8). The next result concerns the uniqueness of the representation (2.1). The proof is based on the linear independence and repeated application of Lemma 2.2 to both the b's and the a's.

THEOREM 2.4. *Let $Y(t)$ be a process of exponential type with order K. Let* $\mathbf{b}$, $B(1), \cdots, B(M)$, D, e^{Dt} *be as in* (2.1). *Suppose* $\mathbf{a} = (a_1, a_2, \cdots, a_K)$ *is a K-vector with $a_j = 0$ or 1 and $A(m) = (a_{\theta\mu}(m))$ for $1 \leqq m \leqq M$ are $K \times K$ matrices, such that for any sequence pair* $(\Phi; S)$:

$$P_Y(\Phi; S) = \mathbf{a}[\textstyle\prod_{j=1}^{n} e^{Ds_j} A(i_j)]\mathbf{c}'. \tag{2.9}$$

Then

(i) $\mathbf{b} = \mathbf{a}$,

(ii) $b_{\theta\mu}(m) = 0$ *if and only if* $a_{\theta\mu}(m) = 0$,

(iii) $b_{\theta\theta}(m) = a_{\theta\theta}(m)$ *for all m and θ*,

(iv) *if* $b_\theta = b_\mu = 1$, *then* $b_{\theta\mu}(m) = a_{\theta\mu}(m)$ *for all m*,

(v) *if $b_{\theta\mu}(m)$ and $b_{\theta\mu}(k)$ are non-zero, then*

$$b_{\theta\mu}(m)/a_{\theta\mu}(m) = b_{\theta\mu}(k)/a_{\theta\mu}(k).$$

Let us point out that an immediate consequence of part (iv) of Theorem 2.4 is that the representation (2.1) is unique if $\mathbf{b} = (1, 1, \cdots, 1)$ and the ν's are ordered lexicographically.

We now introduce the concept of an antecedent matrix for a set of factor matrices and establish their existence and connection with the problem of calculating the joint probabilities of exponential type process.

DEFINITION 2.5. Let $B(1), \cdots, B(M)$ be a set of $K \times K$ factor matrices. A $K \times K$ matrix C is termed an *antecedent matrix* for the set $\{B(1), \cdots, B(m)\}$ provided that

(i) C is non-singular,

(ii) $C\mathbf{c}' = \Psi'$, where $\Psi = (1, 1, \cdots, 1)$,

(iii) $CB(m)C^{-1} = E_C(m)$ is a diagonal matrix whose diagonal entries are 0's and 1's for all $1 \leqq m \leqq M$.

For $1 \leqq m \leqq M$, let $r_m = r[B(m)]$ = the rank of the matrix $B(m)$, $\sigma_m = \sum_{k=1}^{m} r_k$ and

$$E(m) = \text{diag}\{0, 0, \cdots, 0, 1, 1, \cdots, 1, 0, \cdots, 0\},$$

$$(\text{1's in the } \tau_m\text{th through } \sigma_m\text{th positions})$$

where $\tau_1 = 1$ and $\tau_m = \sigma_{m-1} + 1$ for $m > 1$.

THEOREM 2.6. *Let $Y(t)$, K, $B(1), \cdots, B(M)$ be as in Theorem 2.3. Then there exists an antecedent matrix for $B(1), \cdots, B(M)$.*

PROOF. The existence of a non-singular matrix C such that $CB(m)C^{-1} = E(m)$ for $1 \leqq m \leqq M$ is a standard result concerning a set of indempotent, pairwise orthogonal matrices which sum to the identity. We will not prove this, rather we will use this together with the fact that $\mathfrak{M}$ is the essential state space of $Y(t)$ to construct such a C with the important additional property $C\mathbf{c}' = \psi'$.

Consider a fixed m, $1 \leqq m \leqq M$. Since $\mathfrak{M}$ is the essential state space, there exists an $s \geqq 0$ such that

$$0 < P_Y(m; s) = \mathbf{b}e^{Ds}B(m)\mathbf{c}' = \mathbf{b}e^{Ds}(b_{11}(m), b_{21}(m), \cdots, b_{K1}(m))'.$$

Thus the first column of $B(m)$ is not the zero vector and $r_m \geqq 1$. It is clear that $\sum_{m=1}^{M} r[E(m)] = r[\sum_{m=1}^{M} E(m)]$. Choosing C such that $CB(m)\, C^{-1} = E(m)$, yields $r_m = r[E(m)]$ and $C[\sum_{m=1}^{M} B(m)]C^{-1} = \sum_{m=1}^{M} E(m)$. By Theorem 2.3, $\sum_{m=1}^{M} B(m) = I$, hence $\sum_{m=1}^{M} E\,(m) = I$. We conclude that

$$(2.10) \qquad r_m \geqq 1, \qquad K = \sum_{m=1}^{M} r_m \geqq \sum_{m=1}^{M} 1 = M.$$

Since the rank of $B(m)$ is r_m, r_m columns of $B(m)$ form a linearly independent set. The first column of $B(m)$ is not zero, hence it can be taken as one of the vectors in such a linearly independent set. Let us first assume that columns one through r_m are linearly independent. We define

$$(2.11) \qquad \mathbf{w}_{mj} = (b_{1j}(m), b_{2j}(m), \cdots, b_{Kj}(m))' \quad \text{for} \quad 2 \leqq j \leqq K$$

and

$$(2.12) \qquad \mathbf{w}_{m1} = B(m)\mathbf{c}' - \sum_{j=2}^{r_m} \mathbf{w}_{mj}.$$

Our assumption is that $B(m)\mathbf{c}'$, $\mathbf{w}_{m2}$, $\mathbf{w}_{m3}$, $\cdots$, $\mathbf{w}_{mr_m}$ constitute a linearly independent set. If columns $1, k_2, \cdots, k_{r_m}$ of $B(m)$ form a linearly independent set rather than columns $1, 2, \cdots, r_m$, take $\mathbf{w}_{m2} = k_2$nd column, $\cdots$, $\mathbf{w}_{mr_m} = k_{r_m}$th column instead in (2.11) and suitably modify the following argument.

It is easily verified that $\mathbf{w}_{m1}, \mathbf{w}_{m2}, \cdots, \mathbf{w}_{mr_m}$ form a linearly independent set of K-vectors. We repeat the above procedure for each $1 \leqq m \leqq M$, and obtain M sets of vectors: $\mathfrak{A}_m = \{\mathbf{w}_{m1}, \mathbf{w}_{m2}, \cdots, \mathbf{w}_{mr_m}\}$, $1 \leqq m \leqq M$. Thus for each $1 \leqq m \leqq M$, $\mathfrak{A}_m$ is a linearly independent set of r_m vectors and

$$B(m)\mathbf{c}' = \sum_{j=1}^{r_m} \mathbf{w}_{mj}.$$

From (2.10) it is clear that $\mathfrak{A} = \bigcup_{m=1}^{M} \mathfrak{A}_m$ contains K vectors.

From the fact that the $B(m)$'s sum to the identity, we obtain

$$\sum_{m=1}^{M} \sum_{j=1}^{r_m} \mathbf{w}_{mj} = \mathbf{c}'. \tag{2.13}$$

Fix $1 \leqq k, m \leqq K$ and $1 \leqq j \leqq K$. For $j \geqq 2$, $\mathbf{w}_{kj}$ is the jth column of $B(k)$. By Theorem 2.3, $B(m)B(k) = \delta_{mk}B(k)$ hence $B(m)\mathbf{w}_{kj} = \delta_{mk}\mathbf{w}_{kj}$. If $j = 1$, we have

$$\begin{aligned} B(m)\mathbf{w}_{k1} &= B(m)[B(k)\mathbf{c}' - \textstyle\sum_{j=2}^{r_m} \mathbf{w}_{kj}] \\ &= \delta_{mk}B(k)\mathbf{c}' - \textstyle\sum_{j=2}^{r_m} \delta_{mk}\mathbf{w}_{kj} = \delta_{mk}\mathbf{w}_{k1}. \end{aligned}$$

We conclude that

$$B(m)\mathbf{w}_{kj} = \delta_{mk}\mathbf{w}_{kj} \quad \text{for} \quad 1 \leqq k, \quad m \leqq M, \quad 1 \leqq j \leqq K. \tag{2.14}$$

Suppose that

$$\sum_{k=1}^{M} \sum_{j=1}^{r_k} a_{kj}\mathbf{w}_{kj} = \mathbf{0}', \tag{2.15}$$

a_{kj} complex constants, where $\mathbf{0}$ is the zero K-vector.

If we multiply both sides of (2.15) by $B(m)$ and utilize (2.14) we obtain

$$\sum_{j=1}^{r_m} a_{mj}\mathbf{w}_{kj} = \mathbf{0}'$$

for each $1 \leqq m \leqq M$. However $\mathfrak{A}_m$ is a set of linearly independent vectors, and we conclude that $a_{mj} = 0$ for $1 \leqq j \leqq r_m$. Since this is true for each $1 \leqq m \leqq M$, we have shown that $\mathfrak{A}$ is a linearly independent set of K, K-vectors.

We define the $K \times K$ matrix $C = (c_{ij})$ by:

$$C^{-1} = (\mathbf{w}_{11}, \mathbf{w}_{12}, \cdots, \mathbf{w}_{1r_1}, \cdots, \mathbf{w}_{m1}, \mathbf{w}_{m2}, \cdots, \mathbf{w}_{mr_m}, \cdots, \mathbf{w}_{M1}, \mathbf{w}_{M2}, \cdots, \mathbf{w}_{Mr_M}). \tag{2.16}$$

We have shown that C^{-1} is non-singular. It is a straightforward exercise to show that $CB(m)C^{-1} = E(m)$ for $1 \leqq m \leqq M$. The fact that $C\mathbf{c}' = \boldsymbol{\psi}'$ follows immediately from (2.13) and C is an antecedent matrix.

The next theorem asserts that any exponential type process of order K is a function of what one might call a "pseudo-Markov chain" on K points, which results when the non-negativity condition on a stochastic matrix is relaxed. The proof will not be included here.

THEOREM 2.7. *Let $Y(t)$ be an exponential type process or order K. If C is an antecedent matrix for the set* $\{B(1), \cdots, B(M)\}$, then

(i) $Ce^{D^t}C^{-1}\boldsymbol{\psi}' = \boldsymbol{\psi}'$,

(ii) $\mathbf{b}C^{-1}\boldsymbol{\psi}' = 1$,

(iii) *for all sequence pairs* $(\Phi; S)$:

$$\mathrm{P}_Y(\Phi; S) = \mathbf{q}C[\textstyle\prod_{j=1}^{n} Q_C(s_j)E_C(i_j)]\boldsymbol{\psi}' \tag{2.17}$$

and

(iv) $Q_C'(0)$, *the derivative at zero, has distinct eigenvalues* $0 = \nu_1, \nu_2, \cdots, \nu_k$, *where* $\mathbf{q}_C = \mathbf{b}C^{-1}$, $Q_C(t) = Ce^{D^t}C^{-1}$ *and* $E_C(m) = CB(m)C^{-1}$.

3. Necessary and sufficient condition for an exponential type process to be a function of a basic Markov chain. Let $Y(t)$ be an exponential type process or order K. For each non-zero integer j, $2 \leqq j \leqq K$, Lemma 2.2 guarantees the existence of a sequence of subscripts $j = \alpha_1, \alpha_2, \cdots, \alpha_n$ and a corresponding sequence of states $i_1, i_2, \cdots, i_n$ such that

$$b_{j\alpha_2}(i_1)b_{\alpha_2\alpha_3}(i_2) \cdots b_{\alpha_n 1}(i_n) \neq 0. \tag{3.1}$$

If we insist that the length, n, of such a sequence be minimal, it is clear that $j = \alpha_1, \alpha_2, \cdots, \alpha_n, 1$ are distinct. We will say $2 \leqq j \leqq K$ has *power* q if q is the minimum length of a sequence satisfying (3.1). Clearly the power cannot exceed $K - 1$. It will not be necessary to define the power of 1.

THEOREM 3.1. *Let $Y(t)$ be an exponential type process with order K and state space $\mathfrak{M} = \{1, 2, \cdots, M\}$. Let $\mathbf{b}, B(1), \cdots, B(M), D, e^{Dt}$ be as in* (2.1).

If there exists an antecedent matrix C such that

(i) *$Q_C(t) = Ce^{Dt}C^{-1}$ is a probability matrix, and*

(ii) *$\mathbf{q}_C = \mathbf{b}C^{-1}$ is a probability vector,*

then there is a basic Markov chain $X(t)$ with state space $\mathfrak{K} = \{1, 2, \cdots, K\}$ and a function f mapping $\mathfrak{K}$ onto $\mathfrak{M}$, such that $Y(t)$ and $f[X(t)]$ are equal in joint distribution and conversely.

PROOF. For simplicity, we will drop the subscript C and write $Q(t) = (q_{ij}(t)) = Ce^{Dt}C^{-1}$, $\mathbf{q} = (q_1, q_2, \cdots, q_K) = \mathbf{b}C^{-1}$ and $E(m) = (e_{ij}(m)) = CB(m)C^{-1}$ for $1 \leqq m \leqq M$. First we assume C is an antecedent matrix satisfying (i) and (ii). From the corresponding properties of the $B(m)$'s it is clear that the $E(m)$'s form a set of factor matrices, $r[E(m)] \geqq 1$ and $\sum_{m=1}^{M} r[E(m)] = K$. Define a Markov chain $X(t)$ with state space $\mathfrak{K}$, transition matrix $Q(t)$ and initial probability vector $\mathbf{q}$. By Theorem 2.7 $Q'(0)$ has distinct eigenvalues $0 = \nu_1, \nu_2, \cdots, \nu_k$, hence $X(t)$ is a basic Markov chain. For $1 \leqq m \leqq M$, define $\mathfrak{K}_m = \{k: k \in \mathfrak{K} \text{ and } e_{kk}(m) = 1\}$. Using the properties of the $E(m)$'s it is readily verified that the sets $\mathfrak{K}_m$ are non-empty, pairwise disjoint and that $\mathfrak{K} = \bigcup_{m=1}^{M} \mathfrak{K}_m$. Define a mapping f of $\mathfrak{K}$ onto $\mathfrak{M}$ by $f 1_{\mathfrak{K}_m} = m$, that is $f^{-1}(m) = \mathfrak{K}_m$. Using the Markovian property of $X(t)$ it is easily shown that for any sequence pair $(\Phi; S)$ for the process $f[X(t)]$

$$P_{f[X]}(\Phi; S) = \mathbf{q}[\textstyle\prod_{j=1}^{n} Q(s_j)E(i_j)]\boldsymbol{\psi}'. \tag{3.2}$$

Referring to equation (2.17), it is clear that $Y(t)$ and $f[X(t)]$ are equal in joint distribution.

Conversely assume $X(t)$ is a basic Markov chain with state space $\mathfrak{K}$, initial probability vector $\mathbf{p} = (p_1, p_2, \cdots, p_K)$, and transition matrix $P(t) = (p_{ij}(t))$. Furthermore there is a function f, mapping $\mathfrak{K}$ onto $\mathfrak{M}$, such that $Y(t)$ and $[X(t)]$ are equal in joint distribution.

Since $Y(t)$ is an exponential type process of order K, we have a representation for its joint probabilities of the form:

$$P_Y(\Phi; S) = \mathbf{b}[\textstyle\prod_{j=1}^{n} e^{Ds_j}B(i_j)]\mathbf{c}'. \tag{3.3}$$

We must show that there exists an antecedent matrix C for the set

$\{B(1), \cdots, B(M)\}$ such that $Ce^{Dt}C^{-1}$ is a transition matrix and $\mathbf{q} = \mathbf{b}C^{-1}$ is a probability vector. Actually we will prove a slightly stronger result, namely that there is an antecedent matrix C^* such that $P(t) = C^*e^{Dt}C^{*-1}$ and $\mathbf{p} = \mathbf{b}C^{*-1}$.

Let $0 = \omega_1, \omega_2, \cdots, \omega_K$ be the eigenvalues of $P'(0)$, the derivative at zero matrix, and $G(1), \cdots, G(M)$ be the $K \times K$ diagonal matrices defined by $G(m) = (g_{ij}(m)) = (\delta_{ij}\delta_{f(j)m})$. There exists a non-singular matrix $C = (c_{ij})$ such that

$$C^{-1}P'(0)C = F = \text{diag}\{0 = \omega_1, \omega_2, \cdots, \omega_K\} \quad \text{and} \quad P(t) = Ce^{Ft}C^{-1}.$$

Since $P'(0)\boldsymbol{\psi}' = \mathbf{0}'$ and $\omega_1 = 0$, we may assume that the columns of C are normalized such that the first column of C is $\boldsymbol{\psi}'$ and $a_j = \sum_{i=1}^{K} p_i c_{ij}$ is zero or one for $j > 1$.

We continue by defining, for $1 \leqq m \leqq M$,

$$A(m) = (a_{\theta\mu}(m)) = C^{-1}G(m)C,$$

to get the representation:

$$P_Y(\Phi; S) = \mathbf{a}[\textstyle\prod_{j=1}^{n} e^{Fs_j}A(i_j)]\mathbf{c}' \tag{3.4}$$

where $\mathbf{a} = (1, a_2, \cdots, a_K)$. Note that:

$$a_{\theta\mu}(m) = \textstyle\sum_{f(k)=m} c^{\theta k}c_{k\mu},$$

where $C = (c_{ij})$ and $C^{-1} = (c^{ij})$.

According to the argument preceding Theorem 2.4, it is no loss of generality to assume $F = D$ and

$$P_Y(\Phi; S) = \mathbf{a}[\textstyle\prod_{j=1}^{n} e^{Ds_j}A(i_j)]\mathbf{c}'. \tag{3.5}$$

Note that if $b_j = 1$ for all j, the result follows from part (iv) of Theorem 2.4. Comparing (3.3) and (3.5) yields the identity:

$$b_{\alpha_1}b_{\alpha_1\alpha_2}(i_1) \cdots b_{\alpha_n 1}(i_n) \equiv a_{\alpha_1}a_{\alpha_1\alpha_2}(i_1) \cdots a_{\alpha_n 1}(i_n). \tag{3.6}$$

We partition the set $\mathcal{K}$, using the power concept, as follows:

$$\mathcal{K}_1 = \{\alpha : \alpha \in \mathcal{K} \quad \text{and} \quad b_\alpha = 1\},$$

$$\mathcal{K}_l = \{\alpha : \alpha \in \mathcal{K}, b_\alpha = 0, \quad \text{and} \quad \alpha \quad \text{is of power} \quad l\} \quad \text{for} \quad 2 \leqq l \leqq k.$$

Clearly the $\mathcal{K}_l$'s are pairwise disjoint and $\mathcal{K} = \bigcup_{l=1}^{K} \mathcal{K}_l$. For $1 \leqq l \leqq K$, let $\mathcal{K}_l = \{\theta_{1l}, \theta_{2l}, \cdots, \theta_{k_l l}\}$. For each $\theta_{jl} \in \mathcal{K}_l$, $l \geqq 2$, let $\theta_{jl} = \alpha_{1jl}, \alpha_{2jl}, \cdots, \alpha_{ljl}$ be a sequence of length l and $i_1, i_2, \cdots, i_l$ be a sequence of states of $\mathfrak{M}$ such that

$$b_{\alpha_{1jl}\alpha_{2jl}}(i_1)b_{\alpha_{2jl}\alpha_{3jl}}(i_2) \cdots b_{\alpha_{ljl}1}(i_l) \neq 0.$$

If $h_1, h_2, \cdots, h_l$ is a second such sequence of states we know from Theorem 2.4 part (v)) that

$$b_{\alpha_{1jl}\alpha_{2jl}}(i_1) \cdots b_{\alpha_{ljl}1}(i_l)(a_{\alpha_{1jl}\alpha_{2jl}}(i_1) \cdots a_{\alpha_{ljl}1}(i_l))^{-1}$$
$$= b_{\alpha_{1jl}\alpha_{2jl}}(h_1) \cdots b_{\alpha_{ljl}1}(h_l)(a_{\alpha_{1jl}\alpha_{2jl}}(h_1) \cdots a_{\alpha_{ljl}1}(h_l))^{-1}.$$

For simplicity we will drop the states and denote this common ratio by:

$$(3.7) \qquad d_{jl} = b_{\theta_{jl}\alpha_{2jl}} b_{\alpha_{2jl}\alpha_{3jl}} \cdots b_{\alpha_{ljl}1}(a_{\theta_{jl}\alpha_{2jl}} a_{\alpha_{2jl}\alpha_{3jl}} \cdots a_{\alpha_{ljl}1})^{-1} \neq 0.$$

If $\theta = \theta_{j1} \varepsilon \mathcal{K}_1$, $b_\theta = 1$, by (3.6) $b_{\theta_{j1}\alpha_2}(i_1) \cdots b_{\alpha_n 1}(i_n) = a_{\alpha_{j1}\alpha_2}(i_1) \cdots a_{\alpha_n 1}(i_n)$, and whenever $b_{\theta_{j1}\alpha_2}(i_1) \cdots b_{\alpha_n 1}(i_n) \neq 0$,

$$(3.8) \qquad d_{j1} = b_{\theta_{j1}\alpha_2} \cdots b_{\alpha_n 1}(a_{\theta_{j1}\alpha_2} \cdots a_{\alpha_n 1})^{-1} = 1.$$

Define a $K \times K$ matrix $C^* = (c^*_{\phi\gamma})$ by

$$(3.9) \qquad c^*_{\phi\gamma} = (d_{jl})^{-1} c_{\phi\theta_{jl}} \quad \text{if} \quad \gamma = \theta_{jl} \varepsilon \mathcal{K}_l,\ l \geqq 1,$$

where $C = (c_{\phi\gamma})$ is the matrix we started with. It is readily verified that C^* is non-singular and $C^{*-1} = (c_*^{\phi\gamma})$ where $c_*^{\phi\gamma} = d_{jr} c^{\theta_{jr}\gamma}$ if $\phi = \theta_{jr} \varepsilon \mathcal{K}_r$. From the corresponding properties of C and the definition of C^*, we have $C^{*-1}P'(0)C^* = D$, $C^* e^{Dt} C^{*-1} = P(t)$ and $\mathbf{b} = \mathbf{a} = \mathbf{p}C = \mathbf{p}C^* = \mathbf{a}^*$. Moreover, since $b_1 = 1$, $1 \varepsilon \mathcal{K}_1$ and $C^*\mathbf{c}' = C\mathbf{c}' = \boldsymbol{\psi}'$. As before, we define $A^*(m) = (a^*_{\theta\mu}(m)) = C^{*-1}G(m)C^*$, and obtain:

$$(3.10) \qquad P_Y(\Phi; S) = \mathbf{a}^*[\textstyle\prod_{j=1}^n e^{Ds_j} A^*(i_j)]\mathbf{c}',$$

where $a^*_{\theta\mu}(m) = \sum_{f(j)=m} c_*^{\theta j} c^*_{j\mu}$.

We now verify that $A^*(m) = B(m)$ for all $1 \leqq m \leqq M$. Note that:

$$(3.11) \qquad a^*_{\theta\mu}(m) = (d_{jr}/d_{kl}) a_{\theta_{jr}\theta_{kl}} \quad \text{if} \quad \theta = \theta_{jr} \varepsilon \mathcal{K}_r \quad \text{and} \quad \mu = \theta_{kl} \varepsilon \mathcal{K}_l.$$

By Theorem 2.4 $b_{\theta\theta}(m) = a_{\theta\theta}(m)$, $b_{\theta\mu}(m) = 0$ if and only if $a_{\theta\mu}(m) = 0$ and $b_{\theta\mu}(m) = a_{\theta\mu}(m)$ if $\theta, \mu \varepsilon \mathcal{K}_1$.

From (3.8), (3.11) and the above it is clear that we have reduced the problem to showing that $a^*_{\theta\mu}(m) = b_{\theta\mu}(m)$ when $\theta \neq \mu$, $\theta = \theta_{jr} \varepsilon \mathcal{K}_r$, $\mu = \theta_{kl} \varepsilon \mathcal{K}_l$ and $b_{\theta\mu}(m) \neq 0$. In this case we have

$$(3.12) \qquad b_{\beta_1} b_{\beta_1\beta_2} \cdots b_{\beta_h\theta_{jr}} b_{\theta_{jr}\theta_{kl}} b_{\theta_{kl}\alpha_{2kl}} \cdots b_{\alpha_{lkl}1}$$
$$= a_{\beta_1} a_{\beta_1\beta_2} \cdots a_{\beta_h\theta_{jr}} a_{\theta_{jr}\theta_{kl}} a_{\theta_{kl}\alpha_{2kl}} \cdots a_{\alpha_{lkl}1} \neq 0,$$

where, if $l = 1$, $b_{\theta_{k1}\alpha_{2k1}} \cdots b_{\alpha_{1k1}} = b_{\theta_{k1}\alpha_2} \cdots b_{\alpha_n 1} \neq 0$. Hence

$$(3.13) \qquad b_{\theta_{jr}\theta_{kl}}(m)(a_{\theta_{jr}\theta_{kl}}(m))^{-1} = a_{\beta_1} a_{\beta_1\beta_2} \cdots a_{\beta_h\theta_{jr}} (b_{\beta_1} b_{\beta_1\beta_2} \cdots b_{\beta_h\theta_{jr}})^{-1}$$
$$\cdot a_{\theta_{kl}\alpha_{2kl}} \cdots a_{\alpha_{lkl}1} (b_{\theta_{kl}\alpha_{2kl}} \cdots b_{\alpha_{lkl}1})^{-1}.$$

Now,

$$b_{\beta_1} b_{\beta_1\beta_2} \cdots b_{\beta_h\theta_{jr}} b_{\beta_{jr}\alpha_{2jr}} \cdots b_{\alpha_{rjr}1} \equiv a_{\beta_1} a_{\beta_1\beta_2} \cdots a_{\beta_h\theta_{jr}} a_{\theta_{jr}\alpha_{2jr}} \cdots a_{\alpha_{rjr}} \neq 0,$$

and we see that

$$b_{\theta_{jr}\theta_{kl}}(m)(a_{\theta_{jr}\theta_{kl}}(m))^{-1} = d_{jr}/d_{kl}.$$

Therefore $b_{\theta_{jr}\theta_{kl}}(m) = a^*_{\theta_{jr}\theta_{kl}}(m)$, and $A^*(m) = B(m)$.

We now have a non-singular matrix C^* with the properties (i) $C^{*-1}P'(0)C^* = D$, (ii) $C^* e^{Dt} C^{*-1} = P(t)$, (iii) $\mathbf{p} = \mathbf{b}C^{*-1}$, (iv) $C^*\mathbf{c}' = \boldsymbol{\psi}'$, (v) $C^*B(m)C^{*-1}$

$= G(m)$ for $1 \leqq m \leqq M$. By (iv) and (v), C^* is an antecedent matrix for the $B(m)$'s. By (ii) $C^* e^{Dt} C^{*-1}$ is a transition matrix and by (iii) $\mathbf{b}C^{*-1}$ is a probability vector.

4. Necessary and sufficient condition for an exponential type process to be Markovian.

LEMMA 4.1. *Let $Y(t)$ be an exponential type process of order K. Let C be the antecedent matrix defined by* (2.16). *If $K = M$, then* (i) $\mathbf{q} = \mathbf{b}C^{-1}$ *is a probability vector and $Q(t) = Ce^{Dt}C^{-1}$ is a probability matrix.*

PROOF. Recall that the rank, r_m, of $B(m)$ is at least one and $\sum_{m=1}^{M} r_m = K$. Thus when $K = M$, $r_m = 1$, $E(m) = \operatorname{diag}\{0, \cdots, 0, 1, 0, \cdots, 0\}$ (1 in the mth position) and $C^{-1} = (B(1)\mathbf{c}', B(2)\mathbf{c}', \cdots, B(M)\mathbf{c}')$. Now, $\Pr[Y(0) = i] = \mathbf{b}B(i)\mathbf{c}' = \mathbf{b}C^{-1}E(i)C\mathbf{c}' = \mathbf{q}E(i)\boldsymbol{\psi}' = q_i$. Therefore $q_i \geqq 0$ and by Theorem 2.7 $\mathbf{q}\boldsymbol{\psi}' = 1$, hence $\mathbf{q}$ is a probability vector.

Also by Theorem 2.7 $Q(t)\boldsymbol{\psi}' = \boldsymbol{\psi}'$, hence to verify that $Q(t) = (q_{ij}(t))$ is a probability matrix we need only show that $q_{ij}(t) \geqq 0$. Fix $i, j \varepsilon \mathfrak{M}$, since $K = M$ and $\mathfrak{M}$ is the essential state space there exists $s \geqq 0$ such that:

$$(4.1)\quad 0 < \Pr[Y(s) = i] = \mathbf{b}e^{Ds}B(i)\mathbf{c}' = \mathbf{b}C^{-1}Ce^{Ds}C^{-1}E(i)C\mathbf{c}' = \mathbf{q}Q(s)E(i)\boldsymbol{\psi}' = \textstyle\sum_{l=1}^{K} q_l q_{li}(s).$$

For all $t \geqq 0$

$$(4.2)\quad 0 \leqq \Pr[Y(s) = i,\ Y(s+t) = j] = \mathbf{b}e^{Ds}B(i)e^{Dt}B(j)\mathbf{c}' = \mathbf{q}Q(s)E(i)Q(t)E(j)\boldsymbol{\psi}' = q_{ij}(t)\textstyle\sum_{l=1}^{K} q_l q_{li}(s).$$

It is clear from (4.1) and (4.2) that $q_{ij}(t) \geqq 0$.

THEOREM 4.2. *Let $Y(t)$ be an exponential type process of order K with state space $\mathfrak{M}$. Then $Y(t)$ is a Markov chain if and only if $K = M$.*

PROOF. If $K = M$, by Lemma 4.1 there is an antecedent matrix C such that $Ce^{Dt}C^{-1}$ and $\mathbf{b}C^{-1}$ are stochastic. Hence by Theorem 2.7, $Y(t)$ is equal in joint distribution to a basic Markov chain and consequently is a basic Markov chain itself.

Assume now that $Y(t)$ is a Markov chain. Let $P(t)$ and $\mathbf{p}$ be its $M \times M$ transition matrix and initial probability vector, respectively. From the Markov assumption we have for all sequence pairs:

$$(4.3)\quad P_Y(\Phi; S) = \mathbf{p}\left[\textstyle\prod_{j=1}^{n} P(s_j)E(i_j)\right]\boldsymbol{\psi}',$$

where $E(i) = \operatorname{diag}\{0, \cdots, 0, 1, 0, \cdots, 0\}$ with the 1 in the ith coordinate.

Let R denote the number of distinct elementary divisors of the derivative at zero matrix, $P'(0)$. Since we have shown previously that $K \geqq M$, we have $R \leqq M \leqq K$. Letting J denote the Jordan canonical form of $P'(0)$, C a nonsingular matrix such that $C^{-1}P'(0)C = J$, and introducing $P(t) = Ce^{Jt}C^{-1}$ into (4.3) yields a representation for $P_Y(\Phi; S)$ involving R distinct parameters in the exponents. Since $Y(t)$ is an exponential type process or order K, we also have a representation of the form (2.1). Using the linear independence to compare coefficients in these two representations one can easily show that the representation obtained from (4.3) is of the form required in the definition of exponential type. Since $Y(t)$ is of order K, we conclude that $R \geqq K$. Hence $R = M = K$, and the result follows.

As a final remark, it is clear that an exponential type process or order K is a Markov chain if and only if it is a basic Markov chain.

Acknowledgment. The author wishes to express his sincere gratitude to Professor A. B. Clarke (Western Michigan University) who suggested these problems and was a most inspiring counselor.

REFERENCES

[1] BARTLETT, M. S. (1956). *An Introduction to Stochastic Processes*. Cambridge Univ. Press.

[2] BLACKWELL, D. and KOOPMANS, L. (1957). On the identifiability problem for functions of finite Markov chains. *Ann. Math. Statist.* **28** 1011–1015.

[3] BURKE, C. J. and ROSENBLATT, M. (1958). A Markovian function of a Markov chain. *Ann. Math. Statist.* **29** 1111–1122.

[4] CHUNG, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, New York.

[5] DHARMADHIKARI, S. W. (1963). Sufficient conditions for a stationary process to be a function of a finite Markov chain. *Ann. Math. Statist.* **34** 1033–1041.

[6] FELLER, W. (1957). *An Introduction to Probability Theory and its Applications*, **1**. Wiley, New York.

[7] GANTMACHER, F. R. (1959). *The Theory of Matrices*, **1**. Chelsea.

[8] GILBERT, E. J. (1959). On the identifiability problem for functions of finite Markov chains. *Ann. Math. Statist.* **30** 688–697.

[9] JACOBSON, N. (1951). *Lectures in Abstract Algebra*, **2**. Van Nostrand, New York.

[10] KAPLAN, W. (1958). *Ordinary Differential Equations*. Addison-Wesley, New York.

[11] KENDALL, D. G. (1951). Some problems in the theory of queues. *J. Roy. Statist. Soc. Ser. B* **13** 151–185.

[12] KENDALL, D. G. (1953). Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain. *Ann. Math. Statist.* **24**.

[13] LEYSIEFFER, F. W. (1967). Functions of finite Markov chains. *Ann. Math. Statist.* **38** 206–212.

[14] ROSENBLATT, M. (1959). Functions of a Markov process that are Markovian. *J. Math. Mech.* **8** 585–596.