The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 843–850

FUNCTIONS OF MARKOV CHAINS[1]

BY R. V. ERICKSON

Michigan State University

1. Introduction and statement of results. Let $X = \{x_t, t \in \mathbf{T}, \mathbf{I}\}$ be a standard Markov chain with stationary transition probabilities, where the time set $\mathbf{T}$ is either $[0, \infty)$ (*continuous*) or $\{0, 1, \cdots\}$ (*discrete*) and the minimal state space $\mathbf{I}$ is countable and discrete. (The terminology here and below is that of Chung [2].) The above statements about X are abbreviated by saying that X is an M.c.; and if $\mathbf{I}$ is finite, X is termed a *finite* M.c.

Let f be a surjection of $\mathbf{I}$ to $\mathbf{J}$ and identify $f^{-1}(\alpha)$ with α in $\mathbf{J}$. The process $Y = (f, X) = \{y_t = f(x_t), t \in \mathbf{T}, \mathbf{J}\}$ is said to be a *function of an* M.c., and its finite joint distributions are given by

$$(1) \qquad P(y(t_1 + \cdots + t_k) = \alpha_k, k = 1, \cdots, n) = \mathbf{p}P(t_1)I_{\alpha 1} \cdots P(t_n)I_{\alpha n}\mathbf{1}$$

where $\mathbf{p} = (P(x_0 = i)), i \in \mathbf{I}$, is the initial distribution of X, $P(t) = (P(x_{s+t} = j \mid x_s = i))$, $i, j \in \mathbf{I}$, is the transition matrix of X, I_α is the matrix with (i, j) entry equal to one if $i = j \in f^{-1}(\alpha) \subset \mathbf{I}$ and zero otherwise, and $\mathbf{1}$ is the column vector all of whose components are one.

It is the purpose of this paper to determine necessary and sufficient conditions under which an arbitrary finite or countable state process is (equal in joint distribution to) a function of an M.c.

From (1) it is clearly necessary for such a process to possess the following property: A stochastic process $Y = \{y_t, t \in \mathbf{T}, \mathbf{J}\}$ is *matricial* if there exists a countable index set K, a real row vector $\mathbf{a} = (a_k)_{k \in K}$, a real matrix function $R(t) = (R_{ij}(t))_{i, j \in K}$, $t \in \mathbf{T}$, with $R(0)$ equal the identity matrix I, real matrices $A_\alpha = (A_{ij}(\alpha))_{i, j}\, \alpha \in \mathbf{J}$, and a real column vector $\mathbf{b} = (b_k)_{k \in K}$ such that the joint distributions of Y are given by

$$(2) \qquad P(y(t_1 + \cdots + t_k) = \alpha_k, k = 1, \cdots, n) = \mathbf{a}R(t_1)A_{\alpha 1} \cdots R(t_n)A_{\alpha n}\mathbf{b}, \qquad n \geqq 1.$$

(So that (2) is well defined when K is infinite, it is assumed that multiplication begins on the left.) The set $\mathscr{R} = \{\mathbf{a}, R(t), A_\alpha, \mathbf{b}\}$ is called a *representation* of Y, and *size* $\mathscr{R}$ is used to denote the cardinality of K.

Before stating the results of this paper we give one further definition which generalizes the notion of rank introduced by Gilbert [6]: Let $Y = \{y_t, t \in \mathbf{T}, \mathbf{J}\}$ be a stochastic process; for $(\mathbf{t}, \boldsymbol{\alpha}) = (t_1, \cdots, t_n, \alpha_1, \cdots, \alpha_n) \in \mathscr{T} \times \mathscr{J} = \{\phi, \phi\} \cup \bigcup_{n=1}^{\infty} (\mathbf{T}^n \times \mathbf{J}^n)$, let $p_Y(\mathbf{t}, \boldsymbol{\alpha}) = P(y(t_1 + \cdots + t_k) = \alpha_k, k = 1, \cdots, n)$ denote the finite joint distributions of Y, and set $p_Y(\phi, \phi) = 1$; for β in $\mathbf{J}$, define *rank* β to be the supremum of the ranks of the matrices $(p_Y(\mathbf{r}_i s_i \mathbf{t}_j, \boldsymbol{\alpha}_i \beta \gamma_j))_{i, j = 1, \cdots, k}$, where

Received February 6, 1969; revised September 18, 1969.

[1] Work partially done at the University of Michigan towards the degree of Ph.D.

the supremum is taken over all $k \geqq 1$ and all $(\mathbf{r}_i s_i \mathbf{t}_j, \boldsymbol{\alpha}_i \beta \boldsymbol{\gamma}_j) \in \mathscr{T} \times \mathscr{J}$ for fixed $\beta \in \mathbf{J}$, where $(\mathbf{rt}, \boldsymbol{\alpha\beta})$ denotes $(r_1, \cdots, r_m, t_1, \cdots, t_n, \alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_n)$ if $(\mathbf{r}, \boldsymbol{\alpha}) = (r_1, \cdots, r_m, \alpha_1, \cdots, \alpha_m)$ and $(\mathbf{t}, \boldsymbol{\beta}) = (t_1, \cdots, t_m, \beta_1, \cdots, \beta_n)$. The *rank of the process* Y, *rank* Y, is defined as $\sum_{\beta \in J}$ rank β. (States of rank zero are never entered and are omitted from discussion.)

The first theorem is the key to extending Dharmadhikari's results [3], [4].

THEOREM 1. *Every state of a countable state process* Y *has finite rank* iff Y *is matricial relative to some representation* $\mathscr{R} = \{\mathbf{a}, R(t), A_\alpha, \mathbf{b}\}$ *for which rank* A_α *is finite for each state* α. *In particular,* Y *has finite rank* iff Y *is matricial relative to a representation of finite size.*

Dharmadhikari's techniques can now be applied to prove the following.

THEOREM 2. *Let* Y *be a process with discrete time set and countable state space. If each state has finite rank, then* Y *is a function of an* M.c., *and a state has a finite pre-image if it is polyhedral* (as defined below).

At the present time we are unable to prove a version of Theorem 2 for continuous time processes because of difficulties in constructing continuous parameter semigroups. We can give, however, necessary and sufficient conditions for a process to be a function of a finite M.c. with discrete or continuous time parameter. These conditions, which allow us to obtain nonnegative solutions to certain equations, are phrased in terms of *polyhedral cones*, that is, cones which are generated by nonnegative linear combinations of a finite number of given vectors. Given any representation $\mathscr{R} = \{\mathbf{a}, R(t), A_\alpha, \mathbf{b}\}$, denote by *row* $\mathscr{R}$ the set of row vectors $\{\mathbf{a}R(t_1)A_{\alpha_1} \cdots R(t_n)A_{\alpha_n}R(s)\}$, where $t_i, s \in \mathbf{T}$, $\alpha_i \in \mathbf{J}$, $i = 1, \cdots, n$, $n \geqq 1$. $\mathscr{R}$ is said to be *polyhedral relative to* $\mathscr{C}$ if there exist finitely many vectors $\{c_i\} \subset$ span row $\mathscr{R}$ such that for $\mathscr{C} =$ cone $\{c_i\}$ (= cone generated by $\{c_i\}$) we have (i) $\mathbf{a} \in \mathscr{C}$, (ii) $\mathscr{C}R(t) \subset \mathscr{C}$ for each $t \in \mathbf{T}$, (iii) $\mathscr{C}A_\alpha \subset \mathscr{C}$ for each $\alpha \in \mathbf{J}$, and (iv) $\mathbf{cb} > 0$ for all $\mathbf{c} \in \mathscr{C} \backslash \{0\}$. A state α *is polyhedral* iff cone $[(\text{row } \mathscr{R})A_\alpha]$ is contained in a cone $\mathscr{C}_\alpha \subset$ span $[(\text{row } \mathscr{R})A_\alpha]$ generated by finitely many vectors such that $\mathscr{C}_\alpha R(t) A_\alpha \subset \mathscr{C}_\alpha$ and $\mathbf{cb} > 0$ for all $\mathbf{c} \in \mathscr{C}_\alpha \backslash \{0\}$ (see Theorem 2 above). (When size $\mathscr{R}$ is finite and the A_α's form a supplementary set of orthogonal projections, it is clear that if each state is polyhedral, then $\mathscr{R}$ is polyhedral relative to cone $\{c_\alpha{}^i A_\alpha \mid c_\alpha{}^i$ generate $\mathscr{C}_\alpha, \alpha \in \mathbf{J}\}$.) If Y is matricial relative to $\mathscr{R}$, the preceding adjectives will often be applied to Y.

The transition matrix of a continuous time finite M.c. is a uniformly continuous semigroup. To insure that a process Y of finite rank is matricial relative to some representation $\mathscr{R} = \{\mathbf{a}, R(t), A_\alpha, \mathbf{b}\}$ for which $R(t)$ is a uniformly continuous semigroup, the following is sufficient: Say that Y has *continuous distributions* if $p_Y(\mathbf{r}s\mathbf{t}, \boldsymbol{\alpha}\beta\boldsymbol{\gamma})$ is continuous as a function of s in $\mathbf{T}$ for all $\beta \in \mathbf{J}$ and all $(\mathbf{rt}, \boldsymbol{\alpha\gamma}) \in \mathscr{T} \times \mathscr{J}$. This is automatically satisfied if T is discrete.

THEOREM 3. *Let* Y *be a process with finite state space and discrete or continuous time set.* Y *is a function of a finite* M.c. iff Y *is a polyhedral process with continuous distributions and finite rank. In particular,* Y *is a function of a finite* M.c. *if* Y *has continuous distributions and each state has rank one or two.*

COROLLARY. *Let Y be as in Theorem* 3. *Then Y is a finite* M.c. iff *Y has continuous distributions and each state has rank one.*

For discrete time set the "if" part of Theorem 3 is proved by Dharmadhikari [4] and the "only if" part by Heller [8]. Fox and Rubin [7] show, also for the discrete time case, that finite rank alone does not imply the result, for they construct a (countable state) M.c. X and a function f such that $Y = (f, X)$ has rank four and is a function of no finite M.c. Our Theorem 2 shows that every such example of this phenomenon must be of this type.

Section 6 gives results similar to Fox and Rubin [7] and Dharmadhikari [5] on splitting states of a process into Markovian states.

2. Proof of Theorem 1. The proof of Theorem 1 is patterned on a construction used by Dharmadhikari [3] for discrete **T**. Assume first that rank $\beta = n_\beta < \infty$ for each $\beta \in \mathbf{J}$. Then there exist sequences (hereafter fixed) $(\mathbf{r}_{\beta i} s_{\beta i}, \boldsymbol{\alpha}_{\beta i}\beta)$ and $(\mathbf{t}_{\beta j}, \gamma_{\beta j}) \in \mathscr{T} \times \mathscr{J}$, $i, j = 1, \cdots, n_\beta$, such that the $n_\beta \times n_\beta$ matrix

$$P_\beta = (p_Y(\mathbf{r}_{\beta i} s_{\beta i} \mathbf{t}_{\beta j}, \boldsymbol{\alpha}_{\beta i} \beta \gamma_{\beta j}))$$

has rank n_β. Define $\mathrm{row}_\beta(\mathbf{r}s, \boldsymbol{\alpha}\beta)$ as the $1 \times n_\beta$ row vector with entries $p_Y(\mathbf{r}s\mathbf{t}_{\beta j}, \boldsymbol{\alpha}\beta\gamma_{\beta j})$, and $\mathrm{col}_\beta\,(\mathbf{t}, \gamma)$ as the $n_\beta \times 1$ column vector with entries $p_Y(\mathbf{r}_{\beta i} s_{\beta i}\mathbf{t}, \boldsymbol{\alpha}_{\beta i}\beta\gamma)$, and $P_{\beta\delta}(u)$ as the $n_\beta \times n_\delta$ matrix with rows $\mathrm{row}_\delta(\mathbf{r}_{\beta i} s_{\beta i} u, \boldsymbol{\alpha}_{\beta i}\beta\delta)$, $i, j = 1, \cdots, n_\beta$, $\beta, \delta \in \mathbf{J}$, $(\mathbf{r}, \boldsymbol{\alpha}), (\mathbf{t}, \gamma) \in \mathscr{T} \times \mathscr{J}$, $s \in \mathbf{T}$. Note that $P_{\beta\delta}(u) = (p_Y(\mathbf{r}_{\beta i} s_{\beta i} u \mathbf{t}_{\delta j}, \boldsymbol{\alpha}_{\beta i}\beta\delta\gamma_{\delta j}))$ so that $P_\beta = P_{\beta\beta}(0)$ and that $P_{\beta\delta}(0) = 0$ for $\beta \neq \delta$. For each $\beta \in \mathbf{J}$ the matrix

$$\begin{pmatrix} P_\beta & \mathrm{col}_\beta(\mathbf{t}, \gamma) \\ \mathrm{row}_\beta\,(\mathbf{r}s, \boldsymbol{\alpha}\beta) & p_Y(\mathbf{r}s\mathbf{t}, \boldsymbol{\alpha}\beta\gamma) \end{pmatrix}$$

has rank n_β. Hence there is a unique row vector $\mathbf{a}(\mathbf{r}s, \boldsymbol{\alpha}\beta)$ such that (i) $\mathbf{a}(\mathbf{r}s, \boldsymbol{\alpha}\beta) = \mathrm{row}_\beta\,(\mathbf{r}s, \boldsymbol{\alpha}\beta) P_\beta{}^{-1}$, and (ii) $\mathbf{a}(\mathbf{r}s, \boldsymbol{\alpha}\beta)\,\mathrm{col}_\beta\,(\mathbf{t}, \gamma) = p_Y(\mathbf{r}s\mathbf{t}, \boldsymbol{\alpha}\beta\gamma)$. If we set $R_{\alpha\beta}(u) = P_{\alpha\beta}(u) P_\beta{}^{-1}$, it follows from (i), (ii) and the definition of $P_{\alpha\beta}(u)$ that $R_{\alpha\beta}(u)\,\mathrm{col}_\beta\,(\mathbf{t}, \gamma) = \mathrm{col}_\alpha\,(u\mathbf{t}, \beta\gamma)$, and by induction (iii) $p_Y(ru\mathbf{t}, \alpha\beta\gamma) = \mathbf{a}(r, \alpha)\,\mathrm{col}_\alpha\,(u\mathbf{t}, \beta\gamma) = \mathbf{a}(r, \alpha) R_{\alpha\beta}(u) R_{\beta, \gamma 1}(t_1) \cdots R_{\gamma(n-1), \gamma n}(t_n)\,\mathrm{col}_{\gamma n}\,(\phi, \phi)$, where $(\mathbf{t}, \gamma) = (t_1, \cdots, t_n, \gamma_1, \cdots, \gamma_n)$. If we let $\mathbf{a}$ be the row vector with blocks $\mathbf{a}(0, \alpha)$, let $R(u)$ and $P(u)$ be the matrices with blocks $R_{\alpha\beta}(u)$ and $P_{\alpha\beta}(u)$, let I_α be the matrix with the identity in block (α, α) and zero elsewhere, and let $\mathbf{b}$ be the column vector with blocks $\mathrm{col}_\beta\,(\phi, \phi)$, we have, setting $r = 0$ and summing over α in (iii): (iv) $p_Y(u\mathbf{t}, \beta\gamma) = \mathbf{a}R(u) I_\beta\, R(t_1) I_{\gamma 1} \cdots R(t_n) I_{\gamma n}\mathbf{b}$ for all $(u\mathbf{t}, \beta\gamma) \in \mathscr{T} \times \mathscr{J}$, and $R(u) = P(u)P^{-1}$, where $P = P(0)$. This is obvious when **J** is finite so that Y has finite rank, and can be shown to hold when **J** is countable using well-known results concerning associativity of infinite matrices (see Kemeny, Snell, Knapp [10]). Hence Y is matricial relative to $\mathscr{R} = \{\mathbf{a}, R(t), I_\alpha, \mathbf{b}\}$. Since the converse of each statement of Theorem 1 is obvious, this completes the proof.

3. Equivalent representations. Two representations of the same process are said to be *equivalent*. The interrelations between equivalent representations provide the key to the proofs of Theorem 2 and Theorem 3.

Given a representation $\mathscr{R} = \{\mathbf{a}, R(t), A_\alpha, \mathbf{b}\}$ of finite size, define col $\mathscr{R}$ to be the set of column vectors $\{R(t_1)A_{\alpha 1}\cdots R(t_n)A_{\alpha n}\mathbf{b}\}$ where $t_i \in \mathbf{T}$, $\alpha_i \in \mathbf{J}$, $i = 1, \cdots, n$, $n \geqq 1$, let $N(\mathscr{R}) = \{\mathbf{c} \in \text{span col}\,\mathscr{R} \mid (\text{row}\,\mathscr{R})\mathbf{c} = 0\}$, let $M(\mathscr{R}) = N(\mathscr{R})^{\perp}$ (in span col $\mathscr{R}$) have basis $\{v_1, \cdots, v_d\}$, let V be the matrix with $\mathbf{v}_j$ in column j, and let $V\downarrow\mathscr{R}$ denote the set $\{\mathbf{a}V, V^*R(t)V, V^*A_\alpha V, V^*\mathbf{b}\}$ where V^* is any left inverse of V with $V^*N(\mathscr{R}) = \mathbf{0}$.

Statement (i) of the following proposition guarantees that a function of a finite M.c. is polyhedral, thus showing that Dharmadhikari's conditions are necessary as well as sufficient. Statement (ii) shows that $\dim M(\mathscr{R}) = \text{rank } Y$ is an invariant for the class of (equivalent) representations of Y. The third and fourth statements extend the results of Leysieffer [11] and Boudreau [1].

PROPOSITION. *Let Y be matricial relative to representations $\mathscr{R}$ and $\mathscr{R}' = \{\mathbf{a}', R'(t), A_\alpha', \mathbf{b}'\}$ of finite size. Then*

(i) *$V\downarrow\mathscr{R}$ is a representation of Y, and if $\mathscr{R}$ is polyhedral relative to $\mathscr{C}$, $V\downarrow\mathscr{R}$ is polyhedral relative to $\mathscr{C}V$,*

(ii) *rank $Y = \dim M(\mathscr{R}) = \dim M(\mathscr{R}') = \dim M(V\downarrow\mathscr{R}) =$ size $V\downarrow\mathscr{R}$,*

(iii) *if size $\mathscr{R}$ = rank Y, then $R(s+t) \equiv R(s)R(t)$, $R(t)\mathbf{b} \equiv \mathbf{b}$, and the matrices A_α form a supplementary set of orthogonal projections; if Y also has continuous distributions, then $R(t) = \exp(tR)$, where $R = \lim_{t\to 0+}(R(t)-R(0))/t$ when $\mathbf{T}$ is continuous, and $R(n) = [R(1)]^n$ when $\mathbf{T}$ is discrete, and*

(iv) *if size $\mathscr{R}$ = size $\mathscr{R}'$ = rank Y, there is a nonsingular matrix H such that $\mathscr{R}' = \{\mathbf{a}H, H^{-1}R(t)H, H^{-1}A_\alpha H, H^{-1}\mathbf{b}\}$.*

PROOF. Statement (i) follows since $VV^* \mid M(\mathscr{R}) = id$, $VV^* \mid N(\mathscr{R}) = 0$ and since A_α and $R(t)$ map row $\mathscr{R}$ into row $\mathscr{R}$: For example, for $\mathbf{r} \in \text{row}\,\mathscr{R}$, $\mathbf{c} = \mathbf{m}+\mathbf{n} \in \text{col}\,\mathscr{R} = M(\mathscr{R}) \oplus N(\mathscr{R})$, and $\mathbf{c}' = A_\alpha VV^*\mathbf{c} = A_\alpha\mathbf{m} = \mathbf{m}'+\mathbf{n}' \in \text{col}\,\mathscr{R}$, we have $(\mathbf{r}A_\alpha)\mathbf{c} = (\mathbf{r}A_\alpha)\mathbf{m} = \mathbf{r}A_\alpha VV^*\mathbf{c} = \mathbf{r}\mathbf{m}' = \mathbf{r}VV^*\mathbf{c}' = \mathbf{r}VV^*A_\alpha VV^*\mathbf{c}$. The second part of (i) is proved similarly and uses the fact that $\mathscr{C} \subset \text{row}\,\mathscr{R}$. Given $(\mathbf{t}, \boldsymbol{\alpha}) = (t_1, \cdots, t_n, \alpha_1, \cdots, \alpha_n) \in \mathscr{T} \times \mathscr{J}$, let $R(\mathbf{t}, \boldsymbol{\alpha}) = R(t_1)A_{\alpha 1}\cdots R(t_n)A_{\alpha n}$, let $R'(\mathbf{t}, \boldsymbol{\alpha}) = R'(t_1)A'_{\alpha 1}\cdots R'(t_n)A'_{\alpha n}$, and let $R(\phi, \phi)$ and $R'(\phi, \phi)$ denote appropriate identity matrices. Choose a basis $\mathbf{u}_i = \mathbf{a}R(\mathbf{r}_i, \boldsymbol{\alpha}_i)R(s_i)$ for span row $\mathscr{R}$, a basis $\mathbf{w}_j = R(\mathbf{t}_j, \boldsymbol{\beta}_j)\mathbf{b}$ for span col $\mathscr{R}$, and define $\mathbf{u}_i' = \mathbf{a}'R'(\mathbf{r}_i, \boldsymbol{\alpha}_i)R'(s_i)$ and $\mathbf{w}_j' = R'(\mathbf{t}_j, \boldsymbol{\beta}_j)\mathbf{b}'$. Decompose $\mathbf{w}_j = \mathbf{m}_j + \mathbf{n}_j$, $\mathbf{m}_j \in M(\mathscr{R})$, and $\mathbf{n}_j \in N(\mathscr{R})$, and decompose $\mathbf{w}_j' = \mathbf{m}_j' + \mathbf{n}_j'$ similarly in $M(\mathscr{R}') \oplus N(\mathscr{R}')$. Let U be the matrix with rows $\mathbf{u}_i$, W and M the matrices with columns $\mathbf{w}_j$ and $\mathbf{m}_j$, respectively, and define U', W' and M' in a like fashion. Clearly rank $UM \leqq$ rank M; if $\mathbf{m} = \sum a_j\mathbf{m}_j$ and $U\mathbf{m} = \mathbf{0}$, then $\mathbf{m} \in M(\mathscr{R}) \cap N(\mathscr{R}) = \mathbf{0}$ so that rank UM = rank M. If $\mathscr{R}$ and $\mathscr{R}'$ are equivalent, then $UM = UW = U'W' = U'M'$, so that $\dim M(\mathscr{R}) = \text{rank } M \leqq \text{rank } M' \leqq \dim M(\mathscr{R}')$. Reversing the roles of $\mathscr{R}$ and $\mathscr{R}'$, we have the second and third equalities in (ii). If $\mathscr{R}$ now denotes the representation constructed in the proof of Theorem 1, then rank Y = size $\mathscr{R} \geqq \dim M(\mathscr{R})$ = size $V\downarrow\mathscr{R} = d$, say. Since the $d \times d$ matrices $V^*I_\alpha V$, $\alpha \in J$, may be assumed to form a supplementary set of orthogonal projections (just choose the $\mathbf{v}_j$'s so that they form an orthonormal basis for the summands of $M(\mathscr{R}) = \sum_\alpha \oplus I_\alpha M(\mathscr{R})$) $d = \sum_{\alpha \in J}$ rank $V^*I_\alpha V$. Also, the matrices considered in defining rank α have entries $\mathbf{a}R(\mathbf{r}_i, \boldsymbol{\beta}_i)V(V^*I_\alpha V)V^*R(\mathbf{t}_j, \boldsymbol{\gamma}_j)\mathbf{b}$ so that rank $\alpha \leqq$ rank $V^*I_\alpha V$, whence $d \geqq$ rank Y, and (ii) follows. We prove (iii) first for the representation constructed in Theorem 1. again denoted by $\mathscr{R}$. For $\mathscr{R}$ the matrices $A_\alpha = I_\alpha$ automatically form a supplementary set of orthogonal projections. The identities $\sum_{\beta \in J} p_Y(\mathbf{r}st\mathbf{u}, \alpha\beta\delta\gamma) \equiv p_Y(\mathbf{r}[s+t]\mathbf{u}, \alpha\delta\gamma)$ and $\sum_{\beta \in J} p_Y(\mathbf{r}s, \alpha\beta) \equiv p_Y(\mathbf{r}, \alpha)$ then imply

$$[*] \qquad U[R(s)R(t)-R(s+t)]W \equiv 0 \quad \text{and} \quad U[R(s)-I]\mathbf{b} \equiv \mathbf{0}.$$

Since rank $Y = $ size $\mathscr{R} = \dim M(\mathscr{R})$ implies that dim span row $\mathscr{R} = $ size $\mathscr{R} = $ dim span col $\mathscr{R}$, U and W are nonsingular and the first part of (iii) holds for $\mathscr{R}$. The second statement is now immediate for $\mathscr{R}$: since $R(s) = P(s)P(0)^{-1}$, if Y has continuous distributions $R(s)$ is a semigroup uniformly continuous in $\mathbf{T}$ (see Hille–Phillips [9]). If $\mathscr{R}'$ is any other representation of Y with size $\mathscr{R}' = $ rank Y, then $UI_\alpha R(0)I_\beta W = U'A_\alpha'R'(0)A_\beta'W'$ and $UI_\alpha W = U'A_\alpha'W'$. But $R(0)$ and $R'(0)$ are identity matrices so that $UW = U'A'W' = U'A'A'W'$, where $A' = \sum_{\alpha \in J} A_\alpha'$. U and W, and thus U', W', and A', are nonsingular, which implies $A'A' = A' = I$ and $A_\alpha' = H^{-1}I_\alpha H$, where $H = WW'^{-1}$. Identity [*] now holds for $\mathscr{R}'$, and this completes the proof of (iii). Statement (iv) now follows from the relations $\mathbf{a}W = \mathbf{a}'W'$, $UR(t)W = U'R'(t)W'$, and $U\mathbf{b} = U'\mathbf{b}'$, which hold since $\mathscr{R}$ and $\mathscr{R}'$ are equivalent and R and R' are semigroups.

Before proving Theorem 2 and Theorem 3, it is necessary to rephrase a condition in the definition of a polyhedral process.

Lemma. *Let R be a real $d \times d$ matrix and let $\mathscr{C} = \operatorname{cone}\{\mathbf{c}_1, \cdots, \mathbf{c}_n\} \subset \mathbf{R}^d$ (row vectors). Then $\mathscr{C} \exp(tR) \subset \mathscr{C}$ for $t \geqq 0$ iff there is a $t_0 > 0$ such that $\mathscr{C}(I+tR) \subset \mathscr{C}$ for $0 \leqq t \leqq t_0$.*

Proof. Let $\mathscr{C}^* = \{\mathbf{c}^* \mid \mathbf{c}\mathbf{c}^* \geqq 0, \text{all } \mathbf{c} \in C\} = \operatorname{cone}\{\mathbf{c}_1^*, \cdots, \mathbf{c}_{n^*}^*\}$ (column vectors) be the dual (or polar cone) of $\mathscr{C}$ so that $\mathscr{C} = \mathscr{C}^{**}$. Then $\mathbf{u} \in \mathscr{C}$ iff $\mathbf{u}\mathbf{c}_j^* \geqq 0, j = 1, \cdots, n^*$. The proof is completed by noting that $\mathbf{c}_i \exp(tR)\mathbf{c}_j^* = \mathbf{c}_i(I+tR)\mathbf{c}_j^* + o(t)$ as $t \to 0$, $i = 1, \cdots, n, j = 1, \cdots, n^*$.

4. Proof of Theorem 3. Let Y be a polyhedral process with continuous distributions and finite rank, and let $\mathscr{R}$ denote the representation constructed in Theorem 1. From the results of Section 3 we may assume that $\mathscr{R}$ is polyhedral relative to $\mathscr{C} = \operatorname{cone}\{\mathbf{c}_i \mid i \in \mathbf{I}\}$, that card $\mathbf{I} = n < \infty$, that $\mathbf{c}_i\mathbf{b} = 1$ for $i \in \mathbf{I}$, and that $\mathbf{c}_i I_\alpha$ is either 0 or $\mathbf{c}_i$ (since cone $\{\mathbf{c}_i I_\alpha \mid i \in \mathbf{I}, \alpha \in \mathbf{J}\} = \mathscr{C}$). Let C be the matrix with rows $\mathbf{c}_i$, so that $C\mathbf{b} = \mathbf{1}$. Since $\mathbf{a} \in \mathscr{C}$, there is a nonnegative row vector $\mathbf{p}$ such that $\mathbf{a} = \mathbf{p}C$. Similarly, there is a nonnegative matrix P such that $CR(1) = PC$ if $\mathbf{T}$ is discrete or $C(I+t_0R) = PC$ if $\mathbf{T}$ is continuous (in the notation of the lemma of Section 3). Define $P(n) = P^n$ if $\mathbf{T}$ is discrete and $P(t) = \exp[(P-I)t/t_0]$ if $\mathbf{T}$ is continuous. Then $P(t)$ is a nonnegative semigroup such that $P(t)C = CR(t)$, $t \in \mathbf{T}$. Since $R(t)\mathbf{b} = \mathbf{b}$, $P(t)\mathbf{1} = P(t)C\mathbf{b} = CR(t)\mathbf{b} = C\mathbf{b} = \mathbf{1}$, and $\mathbf{p}\mathbf{1} = \mathbf{a}\mathbf{b} = \sum_{\alpha \in J} p_Y(0, \alpha) = 1$. Thus $\mathbf{p}$ and $P(t)$ determine a finite M.c. $X = \{x_t, t \in \mathbf{T}, \mathbf{I}\}$. If $f\colon \mathbf{I} \to \mathbf{J}$ is defined by $f(i) = \alpha$ iff $\mathbf{c}_i I_\alpha = \mathbf{c}_i$, then (f, X) is a function of a finite M.c. equal to Y in joint distribution. Notice that (f, X) has a representation $\mathscr{P} = \{\mathbf{p}, P(t), I_\alpha^{n \times n} = \sum_{f(i)=\alpha} I_i^{n \times n}, \mathbf{1}\}$, where $n = \text{size } \mathscr{P}$, and that $C \downarrow \mathscr{P} = \mathscr{R}$. Since a representation of the type $\mathscr{P}$ is polyhedral relative to $\mathscr{D} \cap$ span row $\mathscr{P}$, where $\mathscr{D}$ is the positive orthant in $\mathbf{R}^n$, the converse is obvious. In case each state of Y has rank less than or equal to two, each state is polyhedral, which implies that Y is polyhedral. This completes the proof of Theorem 3.

The proof of the corollary is now obvious and is omitted.

Note that if there exists another M.c. X', with initial distribution $\mathbf{p}'$ and transition matrix $P'(t)$, and a function f' such that $Y = (f', X')$ then Y has a representation $\mathscr{P}' = \{\mathbf{p}', P'(t), I_\alpha' = \sum_{f(i)=\alpha} I_i', \mathbf{1}\}$ and $\mathscr{R} = C' \downarrow \mathscr{P}'$, where C' is constructed as in Section 3. Hence, given Y with polyhedral representation, every X' and f' such that $Y = (f', X')$ can be constructed as above by choosing various generators for various cones relative to which Y is polyhedral.

It should also be pointed out that Dharmadhikari [4] has constructed a five state M.c. X and a function f such that (f, X) has two states, rank four, and is a function of no four state M.c. The number of states in any pre-image M.c. of a polyhedral process is equal to the number of generators chosen to produce the cone relative to which the process is polyhedral. This number is only bounded below by rank Y, and it may never attain rank Y.

5. Proof of Theorem 2. The proof of Theorem 2 parallels that of Theorem 3. Suppose, then, that Y has discrete time set $\mathbf{T}$ and countable state space $\mathbf{J}$ and that each state has finite rank. Let $\mathscr{R} = \{\mathbf{a}, R(t), I_\alpha, \mathbf{b}\}$ be the representation for Y constructed in Theorem 1. Let $\mathscr{C}_\alpha = \text{cone}\{\mathbf{c}_{\alpha i} \mid \mathbf{c}_{\alpha i}\mathbf{b} = 1, i \in J_\alpha\}$ denote a cone relative to which α is polyhedral, if it is; and let $\mathscr{C}_\alpha = \{\mathbf{c}_{\alpha i} \mid i \in J_\alpha\}$ be an enumeration of the distinct vectors in $(\text{row } \mathscr{R})I_\alpha \backslash \{0\}$ scaled so that $\mathbf{c}_{\alpha i}\mathbf{b} = 1$, if α is not polyhedral.

Let $\mathbf{I} = \bigcup_{\beta \in J} J_\beta$ and let $f: \mathbf{I} \to \mathbf{J}$ be defined so that $f^{-1}(\beta) = J_\beta$. Arguing as in the proof of Theorem 3, using $\mathscr{R}$ and the matrix C with rows $\mathbf{c}_{\alpha i}$, $\alpha \in J$, $i \in J_\alpha$, we construct an M.c. X such that Y and (f, X) are equal in joint distribution. Clearly, the polyhedral states of Y have finite pre-image. This completes the proof of Theorem 2.

6. Splitting states of processes. The following results generalize those of Dharmadhikari [5] and those of Fox and Rubin [7] to the continuous time parameter case.

Let $Y = \{y_t, t \in \mathbf{T}, \mathbf{J}\}$ be a process with countable state space $\mathbf{J} = \mathbf{F} + \mathbf{G}$ ($+$ = disjoint union). A state $\beta \in \mathbf{J}$ is termed a *Markovian state* if rank $\beta = 1$, and for such a state it is easy to show that $p_Y(\mathbf{r}s\mathbf{t}, \boldsymbol{\alpha}\beta\boldsymbol{\gamma})p_Y(s, \beta) = p_Y(\mathbf{r}s, \boldsymbol{\alpha}\beta)p_Y(s\mathbf{t}, \beta\boldsymbol{\gamma})$, so that a stochastic process all of whose states are Markovian is a (not necessarily standard) Markov chain with stationary transition probabilities.

A countable state process $X = \{x_t, t \in \mathbf{T}, \mathbf{H} + \mathbf{G}\}$ is termed a *splitting of* Y *on* $\mathbf{F}$ if each state $\eta \in \mathbf{H}$ is Markovian and if there exists a function $f: \mathbf{H} + \mathbf{G} \to \mathbf{F} + \mathbf{G}$ which is a surjection of $\mathbf{H}$ to $\mathbf{F}$ and the identity on $\mathbf{G}$ and is such that (f, X) and Y are equal in joint distribution.

In order to phrase Theorem 1, Theorem 2 and Theorem 3 in the above context, we need to extend the notion of a matricial process: say that $Y = \{y_t, t \in \mathbf{T}, \mathbf{F}+\mathbf{G}\}$ is *matricial on* $\mathbf{F}$ if there is a countable set K and functions

$\mathbf{a}: \mathscr{T} \times \mathscr{G} \times \mathbf{T} \to \mathbf{R}^K$ (real row vectors)

$R: \mathscr{T} \times \mathscr{G} \times \mathbf{T} \to \mathbf{R}^{K \times K}$ (real matrices indexed on $K \times K$),

$A: \mathbf{F} \to \mathbf{R}^{K \times K}$, and

$\mathbf{b}: \mathscr{T} \times \mathscr{G} \to \mathbf{R}^K$ (real column vectors) such that

$$p_Y(\mathbf{t}_1 s_1 \cdots \mathbf{t}_n s_n \mathbf{t}_{n+1}, \gamma_1 \varphi_1 \cdots \gamma_n \varphi_n \gamma_{n+1}) = \mathbf{a}(\mathbf{t}_1, \gamma_1, s_1) A(\varphi_1) \cdots R(\mathbf{t}_n, \gamma_n, s_n) A(\varphi_n) \mathbf{b}(\mathbf{t}_{n+1}, \gamma_{n+1})$$

for all $(\mathbf{t}_i, \gamma_i) \in \mathscr{T} \times \mathscr{G} = \{\phi, \phi\} \cup \bigcup_{n=1}^{\infty} \mathbf{T}^n \times \mathbf{G}^n$, $(s_i, \varphi_i) \in \mathbf{T} \times \mathbf{F}$, $i = 1, \cdots, n+1$, $n = 1, 2, \cdots$. (Multiplication is from the left if K is infinite.) The set $\{\mathbf{a}, R, A, \mathbf{b}\} = \mathscr{R}(\mathbf{F})$ is termed a *representation of* Y *on* $\mathbf{F}$. We set row $\mathscr{R}(\mathbf{F}) = \{\mathbf{a}(\mathbf{t}_1, \gamma_1 s_1) A(\varphi_1) \cdots R(\mathbf{t}_n, \gamma_n, s_n) \hat{A}\}$, where $\hat{A} = I$ or $A(\varphi)$, $(\mathbf{t}_i, \gamma_i) \in \mathscr{T} \times \mathscr{G}$, $(s_i, \varphi_i) \in \mathbf{T} \times \mathbf{F}$, $i = 1, \cdots, n$, $n \geqq 1$, and define col $\mathscr{R}(\mathbf{F})$ similarly. The *polyhedral* property of Y on $\mathbf{F}$ is defined analogously except that (iv) is weakened to read (iv′) $\mathbf{cb}(\mathbf{t}, \gamma) \geqq 0$ for $\mathbf{c} \in \mathscr{C}$ and $(\mathbf{t}, \gamma) \in \mathscr{T} \times \mathscr{G}$; the definition of a *polyhedral state* is changed similarly.

With these extensions we can now rephrase Theorem 1, Theorem 2 and Theorem 3 for a countable state process $Y = \{y_t, t \in \mathbf{T}, \mathbf{F}+\mathbf{G}\}$.

THEOREM 1′. *Every state* $\varphi \in \mathbf{F}$ *has finite rank* iff Y *is matricial on* $\mathbf{F}$ *relative to a representation* $\mathscr{R}(\mathbf{F})$ *for which rank* $A(\varphi)$ *is finite for each* $\varphi \in \mathbf{F}$.

THEOREM 2′. *If* Y *has discrete time and each state* $\varphi \in \mathbf{F}$ *has finite rank, then there exists a splitting process for* Y *on* $\mathbf{F}$, *and a state has finite pre-image if it is polyhedral.*

THEOREM 3′. Y, *on the states* $\mathbf{F}$, *has a splitting* $X = \{x_t, t \in \mathbf{T}, \mathbf{H}+\mathbf{G}\}$ *with finite* $\mathbf{H}$ iff $\sum_{\varphi \in F}$ rank φ *is finite and* Y *is polyhedral on* $\mathbf{F}$.

PROOFS. Theorem 1′ may be proved using the techniques of Theorem 1. I $\sum_{\varphi \in F}$ rank φ is finite, then the representation $\mathscr{R}(\mathbf{F})$ constructed in Theorem 1′ has finite size; and if $\mathscr{R}(\mathbf{F})$ is polyhedral relative to $\mathscr{C} = \text{cone}\{c_i \mid i \in H, H \text{ finite}\}$ and the rows of C are $\mathbf{c}_i$, then Y also has a representation $\mathscr{P}(\mathbf{F}) = \{\mathbf{p}, P, I(\varphi), \mathbf{q}\}$ on $\mathbf{F}$ such that $\mathscr{R}(\mathbf{F}) = C \downarrow \mathscr{P}(\mathbf{F})$ and such that $\mathbf{p}$, $\mathbf{P}$ and $\mathbf{q}$ are nonnegative matrix functions and $I(\varphi)$ is zero off the diagonal, zero or one on the diagonal, and $\sum_{\varphi \in F} I(\varphi) = I =$ identity matrix on $\mathbf{R}^{H \times H}$. The process $X = \{x_t, t \in \mathbf{T}, \mathbf{H}+\mathbf{G}\}$ is now defined as a matricial process on $\mathbf{H}$ using $\mathscr{P}(\mathbf{H}) = \{\mathbf{p}, P, I_\eta, \mathbf{q}\}$, where I_η has entry (i, j) equal to one if $i = j = \eta \in \mathbf{H}$ and zero otherwise. $\mathscr{P}(\mathbf{H})$ satisfies consistency requirements since $\mathscr{P}(\mathbf{F})$ does and since $\sum_{\eta \in H} I_\eta = I$. Conversely, if $X = \{x_t, t \in \mathbf{T}, \mathbf{H}+\mathbf{G}\}$ is a splitting of Y on $\mathbf{F}$ and $\mathbf{H}$ is finite, the construction of Theorem 1′ gives a representation $\mathscr{P}(\mathbf{H}) = \{\mathbf{p}, P, I_\eta, \mathbf{q}\}$ of X on $\mathbf{H}$, where the matrices of $\mathscr{P}(\mathbf{H})$ are nonnegative and I_η is as above. Y thus has a polyhedral representation $\mathscr{P}(\mathbf{F}) = \{\mathbf{p}, P, I_\varphi = \sum_{f(\eta) = \varphi} I_\eta, \mathbf{q}\}$. The needed parts of the proposition of Section 3 extend to show that Y is polyhedral on $\mathbf{F}$ with a representation $\mathscr{R}(\mathbf{F}) = V \downarrow \mathscr{P}(\mathbf{F})$ of size equal $\sum_{\varphi \in F}$ rank $\varphi \leqq$ size $\mathscr{P}(\mathbf{H}) < \infty$. Theorem 2′ may now be proved as was Theorem 2 using the above method to define the splitting process.

Acknowledgment. The author wishes to express sincere thanks to Professor A. Bruce Clarke (Western Michigan University) for an introduction to these problems and for encouragement in their study.

REFERENCES

[1] Boudreau, Paul E. (1968). Functions of finite Markov chains and exponential type processes. *Ann. Math. Statist.* **39** 1020–1029.

[2] Chung, Kai Lai. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, New York.

[3] Dharmadhikari, S. W. (1963a). Functions of finite Markov chains. *Ann. Math. Statist.* **34** 1022–1032.

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

[5] Dharmadhikari, S. W. (1968c). Splitting a single state of a stationary process into Markovian states. *Ann. Math. Statist.* **39** 1069–1077.

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

[7] Fox, Martin and Rubin, Herman. (1968). Functions of processes with Markovian states. *Ann. Math. Statist.* **39** 938–946.

[8] Heller, A. (1965). On stochastic processes derived from Markov chains. *Ann. Math. Statist.* **36** 1286–1291.

[9] Hille, E., and Phillips, R. S. (1957). *Functional Analysis and Semigroups*. American Mathematical Society, Providence.

[10] Kemeny, John G., Snell, J. Laurie and Knapp, Anthony W. (1960). *Denumerable Markov Chains*. Van Nostrand, Princeton.

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