The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 938–946

FUNCTIONS OF PROCESSES WITH MARKOVIAN STATES[1]

By Martin Fox[2] and Herman Rubin[3]

Michigan State University

1. Summary. Let $\{Y_k\}$ be a stochastic process where either $k = 1, 2, \cdots$, or $k = 0, \pm 1, \cdots$. Let S and T be measurable sets of sequences of states. Let ϵ be a state. Assume $P(Y_n = \epsilon) > 0$. Let $p_n(S\epsilon T) = P((\cdots, Y_{n-2}, Y_{n-1}) \varepsilon S, Y_n = \epsilon, (Y_{n+1}, Y_{n+2}, \cdots) \varepsilon T)$. We define the *rank of* ϵ *at time* n to be maximal rank of matrices $(p_n(S_i\epsilon T_j); i, j = 1, \cdots, m)$ as m, the S_i and the T_j vary. In the stationary case, since the rank does not depend on n, we will refer to the rank of ϵ. In this case, with finite state space, Gilbert [6] denoted the rank of ϵ by $n(\epsilon)$.

A state which has rank 1 at time n is a *Markovian state at time n*. A stochastic process all of whose states are Markovian at all times is a Markov process.

Let $\{X_k\}$ be a second stochastic process indexed as $\{Y_k\}$. Gilbert proved (but stated in far less generality) that if $\nu_n(\epsilon)$ and $\mu_n(\delta)$ are the ranks at time n of the states ϵ in $\{Y_k\}$ and δ in $\{X_k\}$, respectively, and if $Y_n = f(X_n)$, then

$$\nu_r(\epsilon) \leqq \sum_{f(\delta)=\epsilon} \mu_n(\delta). \tag{1.1}$$

In the stationary finite state space case, Dharmadhikari [2] gave conditions under which if $\{Y_k\}$ has all states of finite rank it is possible to express $Y_k = f(X_k)$ where $\{X_k\}$ is stationary and Markovian. An example was given to show equality need not hold in (1.1). A similar proof can be used to show in the more general case that if ϵ is a state of finite rank in $\{Y_k\}$ and his conditions hold for ϵ, then it is possible to write $Y_k = f(X_k)$ where $\epsilon = f(\epsilon_i)$ for a finite number of states ϵ_i of $\{X_k\}$ which are Markovian and $\delta = f(\delta)$ for $\delta \neq \epsilon_i$. See [4].

In Section 2 of the present paper we present an example showing that finite rank alone does not guarantee Dharmadhikari's result. In this case we have stationarity for $\{X_k\}$ and $\{Y_k\}$ and $\epsilon = f(\delta)$ for a countable set of states δ of $\{X_k\}$ which are Markovian. This is an example referred to by Dharmadhikari [2] and disproves a conjecture due to Gilbert.

In Section 3 it is proved that if $\{Y_k\}$ has a state ϵ of finite rank at time n, then there exists a stochastic process $\{X_k\}$ such that $Y_k = f(X_k)$ where $\delta = f(\delta)$ if $\delta \neq \epsilon_i$ and $\epsilon = f(\epsilon_i)$ for a countable family $\{\epsilon_i\}$ of Markovian states of $\{X_k\}$ at time n. However, we are unable to prove that the ranks of states at times other than n are undisturbed. Trivially, no rank may be decreased. Alternative con-

Received September 26, 1967.

[1] Prepared under Contract No. Nonr-2587(02) Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government. Distribution of this Document is Unlimited.

[2] Work partially done while this author was a Fulbright lecturer at Tel-Aviv University.

[3] Work by both authors partially done under USAF contract No. AF19 (628)-355.

structions are given to show Markovianess and rank 2 of states may be preserved in $\{X_k\}$. The latter is deferred to Section 4.

Our construction is valid both for $n = 0, \pm 1, \cdots$ and for $n = 1, 2, \cdots$ without any assumption of finiteness of the state space of $\{Y_k\}$. Trivially, if this state space is countable and $n = 1, 2, \cdots$ it is possible to represent $Y_k = f(X_k)$ where $\{X_k\}$ is a countable state Markov process. In this case Carlyle [1] gave a particular construction of $\{X_k\}$ which, in the case that $\{X_k\}$ has a finite state space, yields the minimal state space.

In Section 4 we consider a state ϵ of rank 2 at time n. It is proved that in this case the construction of the process $\{X_k\}$ can be carried out so that there are only two states, ϵ_1 and ϵ_2, for which $\epsilon = f(\epsilon_i)$. Furthermore, it is proved that ranks of all states $\delta \neq \epsilon$ are undisturbed. If, in addition, $\{Y_k\}$ is stationary, it is proved in Section 5 that a stationary $\{X_k\}$ can be constructed with all these properties.

The definition of rank can be readily extended to the case of densities. All results of this paper as well as Gilbert's and Dharmadhikari's results can be obtained in this case. These considerations are deferred to a later paper.

The results of Section 4 have found applications to a stochastic process describing the temporal behavior of cloud cover [5]. The observable states of cloud cover are clear (under 5% of the sky covered by clouds), partly cloudy (between 5% and 50% covered), cloudy (between 50% and 95% covered) and overcast (over 95% covered). A nonstationary Markov chain does not fit the data. Certain matrices of observed frequencies are, approximately, of rank 2 for data from Boston. Assume the observable process is a function of a nonstationary Markov chain with eight states, each state of the observable process being the image of two Markovian states. This model fits the Boston data very well.

2. Example. Let the transition probabilities for $\{X_k\}$ be given by:

Initial State	Final State	Transition Probability
δ	$\epsilon_{m,1}$	$(\frac{1}{2})^m \sin^2 m$
δ	δ	$1 - \sum_{k=1}^{\infty} (\frac{1}{2})^k \sin^2 k$
$\epsilon_{m,i}$	$\epsilon_{m,i+1}$	1 if $i = 1, \ldots, m-1$
$\epsilon_{m,m}$	δ	1

Let $Y_k = f(X_k)$ where $f(\delta) = \delta$ and $f(\epsilon_{m,i}) = \epsilon$ for all m and i. Then δ is Markovian in both processes.

In this case it is possible to consider finite sequences s_i and t_j of states in place of arbitrary sets S_i and T_j of sequences. In the matrices of the form $(p(s_i \epsilon t_j))$ the sequences s_i can be truncated by omitting any portions to the left of Markovian states and the sequences t_j can be truncated by omitting any portions to the right of Markovian states. Thus, let s_i consist of a δ followed by i ϵ's and t_j consist of a δ proceeded by j ϵ's. Then, the infinite matrix $(p(s_i \epsilon t_j)$: $i, j = 0, 1, \cdots)$ is, except for constant factors in each row (or column), the matrix

$$P = (\sin^2 (i + j + 1) : i, j = 0, 1, \cdots).$$

But P is of rank 3 so the rank of ϵ is 3.

Given any $\tau > 0$, the matrix P can be transformed by rearranging rows and columns to form a matrix $P^*(\tau)$ with main diagonal elements less than τ. But a non-negative hollow matrix of rank greater than 1 cannot be expressed as a sum of a finite number of non-negative matrices of rank 1 so that, by continuity, the same property holds for P, which proves that there is no stochastic process $\{Z_k\}$ with Markovian states $\beta_1, \cdots, \beta_r$ and function g with $Y_k = g(Z_k)$, $\epsilon = g(\beta_i)$ $(i = 1, \cdots, r)$ and $g(\alpha) \neq \epsilon$ for all other states α of $\{Z_k\}$.

3. The general case of finite rank. Since we do not assume stationarity, we may assume the state spaces at different times are disjoint.

THEOREM 1. *Let $\{Y_k\}$ have state space U_k at each time k. Let $\epsilon \varepsilon U_n$ have finite rank at time n. Then, there exists a process $\{X_k\}$ such that*

(i) *$\{X_k\}$ has state space U_k at time $k \neq n$ and $(U_n - \{\epsilon\}) \cup \{\epsilon_1, \epsilon_2, \cdots\}$ at time n;*

(ii) *The states $\epsilon_1, \epsilon_2, \cdots$ are Markovian;*

(iii) *$Y_k = f(X_k)$ where $f(\delta) = \delta$ if $\delta \neq \epsilon_i$ and $f(\epsilon_i) = \epsilon$ for $i = 1, 2, \cdots$.*

PROOF. Let the rank of ϵ at time n be $\nu_n(\epsilon)$. Let $\mathcal{S}$ and $\mathfrak{T}$ be the classes of all measurable sets of sequences of states prior to and after time n, respectively. Then, there exist functions Φ_i and Ψ_i $(i = 1, \cdots, \nu_n(\epsilon))$ on $\mathcal{S}$ and $\mathfrak{T}$ respectively, such that

$$p_n(S\epsilon T) = \sum_{i=1}^{\nu_n(\epsilon)} \Phi_i(S)\Psi_i(T) \tag{3.1}$$

for all $S \varepsilon \mathcal{S}$, $T \varepsilon \mathfrak{T}$.

Let $p_n(S\epsilon Y) = Q(S)$ and $p_n(X\epsilon T) = R(T)$ where $X = \cdots \times U_{n-2} \times U_{n-1}$ and $Y = U_{n+1} \times U_{n+2} \times \cdots$. We wish to show $\Phi_i \ll Q$ and $\Psi_i \ll R$ for $i = 1, \cdots, \nu_n(\epsilon)$. Suppose the former is false. Let E be such that $Q(E) = 0$ and $\Phi_i(E) \neq 0$ for some i. Then, $p_n(E\epsilon T) = 0$ so that

$$\sum_{i=1}^{\nu_n(\epsilon)} \Phi_i(E)\Psi_i(T) = 0$$

for all $T \varepsilon \mathfrak{T}$. But then the Ψ_i are not linearly independent, so that the rank of ϵ at time n is less than $\nu_n(\epsilon)$, a contradiction. The proof that the $\Psi_i \ll R$ is the same. Thus, (3.1) becomes

$$p_n(S\epsilon T) = \sum_{i=1}^{\nu_n(\epsilon)} \int_{S\times T} \varphi_i(x)\psi_i(y)\, dQ(x)\, dR(y) \tag{3.2}$$

where

$$\varphi_i = d\Phi_i/dQ \quad \text{and} \quad \psi_i = d\Psi_i/dR.$$

Then, each φ_i is a function on X and each ψ_i is a function on Y.

Let $\varphi = (\varphi_1, \cdots, \varphi_{\nu_n(\epsilon)})$ and $\psi = (\psi_1, \cdots, \psi_{\nu_n(\epsilon)})$. Let $\varphi^*(x) = \varphi(x)/\|\varphi(x)\|$ and $\psi^*(y) = \psi(y)/\|\psi(y)\|$. Then, $(\varphi^*(x), \psi^*(y)) \geqq 0$ for almost all $x \varepsilon X$, $y \varepsilon Y$ since $p_n(S\epsilon T) \geqq 0$ for all $S \varepsilon \mathcal{S}$, $T \varepsilon \mathfrak{T}$. By changing φ^* and ψ^* on appropriate sets of measure 0, we obtain this inequality for all x, y.

Let $\mathfrak{U}$ and $\mathfrak{V}$ be the range spaces of φ^* and ψ^*, respectively. Then, $\overline{\mathfrak{U}}$ and $\overline{\mathfrak{V}}$ are compact and inner product is a continuous function on $\mathfrak{U} \times \mathfrak{V}$. Hence by a theorem due to Rubin [7],

$$(u, v) = \sum_{j=1}^{\infty} \alpha_j(u)\beta_j(v) \tag{3.3}$$

for all $u \varepsilon \mathfrak{U}$, $v \varepsilon \mathfrak{V}$ where the $\alpha_j \geqq 0$ and $\beta_j \geqq 0$.

Inserting (3.3) in (3.2) we obtain

$$p_n(S\epsilon T) = \sum_{j=1}^{\infty} \int_{S\times T} \|\varphi(x)\| \, \alpha_j(\varphi^*(x)) \|\psi(y)\| \, \beta_j(\psi^*(y)) \, dQ(x) \, dR(y). \tag{3.4}$$

The α_j and β_j can be normalized by multiplication of α_j by a constant $c_j > 0$ and dividing β_j by c_j. We normalize for $j = 1, 2, \cdots$ so that

$$\int_Y \|\psi(y)\| \, \beta_j(\psi^*(y)) \, dR(y) = 1. \tag{3.5}$$

Then from (3.4),

$$p_n(S\epsilon Y) = \sum_{j=1}^{\infty} \int_S \|\varphi(x)\| \, \alpha_j(\varphi^*(x)) \, dQ(x). \tag{3.6}$$

We now define $\{X_k\}$. Let $\{X_k\}$ have state space as defined in (i) of this theorem. For $A \subset U_n - \{\epsilon\}$, let

$$P((\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n \varepsilon A, (X_{n+1}, X_{n+2}, \cdots) \varepsilon T)$$
$$= P((\cdots, Y_{n-2}, Y_{n-1}) \varepsilon S, Y_n \varepsilon A, (Y_{n+1}, Y_{n+2}, \cdots) \varepsilon T).$$

Furthermore, let

$$P((\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n = \epsilon_j) = \int_S \|\varphi(x)\| \, \alpha_j(\varphi^*(x)) \, dQ(x) \tag{3.7}$$

and

$$P((X_{n+1}, X_{n+2}, \cdots) \varepsilon T \,|\, (\cdots X_n, X_{n-1}) \varepsilon S, X_n = \epsilon_j)$$
$$= \int_T \|\psi(y)\| \, \beta_j(\psi^*(y)) \, dR(y).$$

By (3.5) and (3.6) these probabilities are consistent and $Y_k = f(X_k)$ which completes the proof.

COROLLARY. *Under the conditions of Theorem* 1, *there exists* $\{X_k\}$ *satisfying the conclusions for which every Markovian state* $\delta \neq \epsilon$ *in* $\{Y_k\}$ *is Markovian in* $\{X_k\}$.

PROOF. From (3.4) we have

$$p_n(S\epsilon T) = \sum_1^{\infty} \Lambda_i(S)\Xi_i(T) \tag{3.8}$$

for all $S \varepsilon \mathcal{S}$, $T \varepsilon \mathfrak{I}$ where $\Lambda_i \geqq 0$, $\Xi_i \geqq 0$.

We say T is a right M-set if whenever $u\delta v \varepsilon T$ and δ is Markovian, then also $u\delta w \varepsilon T$ for all sequences w. We define left M-sets similarly. Let $\mathcal{S}_M$ and $\mathfrak{I}_M$ be, respectively, the classes of all measurable left M-sets of states prior to time n and right M-sets of states after time n.

Let

$$g(T, t) = \chi_T(t) \quad \text{if} \quad t \quad \text{has no Markovian states} \tag{3.9}$$
$$= P((Y_{n+1}, Y_{n+2}, \cdots) \varepsilon T \,|\, (Y_{n+1}, \cdots, Y_{n+m}) = u, Y_{n+m+1} = \delta)$$

if $t = u\delta w$ where u is a sequence of length m without Markovian states and δ is Markovian. Define

$$\Xi_i^*(T) = \int g(T, t)\, d\Xi_i(t). \tag{3.10}$$

Similarly define $f(S, s)$ and

$$\Lambda_i^*(S) = \int f(S, s)\, d\Lambda_i(s). \tag{3.11}$$

Then $\Lambda_i^*(S) = \Lambda_i(S)$ for $S \varepsilon \mathfrak{S}_M$ and $\Xi_i^*(T) = \Xi_i(T)$ for $T \varepsilon \mathfrak{T}_M$. Furthermore

$$p_n(S\epsilon T) = \sum_1^\infty \Lambda_i^*(S)\Xi_i^*(T)$$

from (3.8) and (3.9).

We now define $\{X_k\}$. For $A \subset U_n - \{\epsilon\}$ use (3.7). Let

$$P((\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n = \epsilon_j) = \Lambda_j^*(S)$$

and

$$P((X_{n+1}, X_{n+2}, \cdots) \varepsilon T \mid (\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n = \epsilon_j) = \Xi_j^*(T).$$

It suffices to show that a state $\delta \neq \epsilon$ which is Markovian in $\{Y_k\}$ is Markovian in $\{X_k\}$. Examining (3.9) and (3.10) we see that $\{X_k\}$ has the transition probabilities from ϵ_j at time n given by Ξ_j until a Markovian state of $\{Y_k\}$ is reached. After a Markovian state is reached, the transition probabilities for $\{X_k\}$ as given by the Ξ_j^* become those of $\{Y_k\}$. Thus any state of $\{Y_k\}$ at time $k > n$ which is Markovian in $\{Y_k\}$ is Markovian in $\{X_k\}$. Similarly, examining (3.11) and the analogue of (3.9) which defines f we see that a state of $\{Y_k\}$ at time $k < n$ which is Markovian in $\{Y_k\}$ is Markovian in $\{X_k\}$. Trivially a state of $\{Y_k\}$ Markovian at time n is Markovian in $\{X_k\}$. This completes the proof.

An extension of Theorem 1 to allow the simultaneous construction of Markovian states at a fixed finite set of time points can be carried out. Stronger results are possible if we assume Markovian states occur infinitely often (in both directions in the case $k = 0, \pm 1, \cdots$). These considerations are deferred to a later paper.

4. The case of rank 2. In this section we do not, as yet, assume stationarity.

THEOREM 2. *Let $\{Y_k\}$ have state space U_k at each time k. Let $\epsilon \varepsilon U_n$ have rank 2 at time n. Then, there exists a process $\{X_k\}$ such that*

(i) *$\{X_k\}$ has state space U_k at time $k \neq n$ and $(U_n - \{\epsilon\}) \cup \{\epsilon_1, \epsilon_2\}$ at time n;*
(ii) *The states ϵ_1, ϵ_2 are Markovian;*
(iii) *$Y_k = f(X_k)$ where $f(\delta) = \delta$ if $\delta \neq \epsilon_i$ and $f(\epsilon_i) = \epsilon$ for $i = 1, 2$;*
(iv) *If $\delta \neq \epsilon$, then the rank at any time of δ is the same in $\{X_k\}$ as in $\{Y_k\}$.*

PROOF. Let X, Y, $\mathfrak{S}$, $\mathfrak{T}$ be as in the proof of Theorem 1. We first prove that there exist measurable $S_1 \varepsilon \mathfrak{S}$ and $T_1 \varepsilon \mathfrak{T}$ such that

$$\begin{vmatrix} p_n(S_1\epsilon T_1) & p_n(S_1\epsilon Y) \\ p_n(X\epsilon T_1) & p_n(X\epsilon Y) \end{vmatrix} \neq 0. \tag{4.1}$$

Suppose false. Then, for all choices of S_1 and T_1

$$\begin{aligned} p_n(S_1 \epsilon T_1) &= p_n(S_1 \epsilon Y) p_n(X \epsilon T_1)/p_n(X \epsilon Y) \\ &= \alpha(S_1)\beta(T_1) \end{aligned}$$

so that the rank of ϵ is 1, a contradiction.

Choose S_1 and T_1 to satisfy (4.1). Then, for all measurable $S \varepsilon \mathcal{S}$, $T \varepsilon \mathfrak{T}$,

$$\begin{vmatrix} p_n(S\epsilon T) & p_n(S\epsilon T_1) & p_n(S\epsilon Y) \\ p_n(S_1\epsilon T) & p_n(S_1\epsilon T_1) & p_n(S_1\epsilon Y) \\ p_n(X\epsilon T) & p_n(X\epsilon T_1) & p_n(X\epsilon Y) \end{vmatrix} = 0$$

so that

$$p_n(S\epsilon T) = p_n(S\epsilon T_1)\psi_1(T) + p_n(S\epsilon Y)\psi_2(T).$$

Now $p_n(S\epsilon T) \leqq p_n(S\epsilon Y)$ so that we may assume $p_n(S\epsilon Y) \neq 0$. Let

$$\lambda(S) = p_n(S\epsilon T_1)/p_n(S\epsilon Y).$$

Then $0 \leqq \lambda(S) \leqq 1$. Let $A = \sup_{S\varepsilon\mathcal{S}} \lambda(S)$ and $B = \inf_{S\varepsilon\mathcal{S}} \lambda(S)$. Now,

$$\lambda(S)\psi_1(T) + \psi_2(T) = p_n(S\epsilon T)/p_n(S\epsilon Y) \geqq 0. \tag{4.2}$$

But,

$$\begin{aligned} \lambda(S)\psi_1(T) + \psi_2(T) &= (\lambda(S) - B)(A - B)^{-1}[\psi_2(T) + A\psi_1(T)] \\ &\quad + (A - \lambda(S))(A - B)^{-1}[\psi_2(T) + B\psi_1(T)] \\ &= (\lambda(S) - B)(A - B)^{-1}\Psi_1(T) \\ &\quad + (A - \lambda(S))(A - B)^{-1}\Psi_2(T), \end{aligned} \tag{4.3}$$

say, where $\Psi_1 \geqq 0$, $\Psi_2 \geqq 0$. From (4.2) and (4.3)

$$p_n(S\epsilon T) = \Phi_1(S)\Psi_1(T) + \Phi_2(S)\Psi_2(T) \tag{4.4}$$

where

$$\begin{aligned} \Phi_1(S) &= p_n(S\epsilon Y)[\lambda(S) - B]/(A - B) \\ &= [p_n(S\epsilon T_1) - Bp_n(S\epsilon Y)]/(A - B) \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} \Phi_2(S) &= p_n(S\epsilon Y)[A - \lambda(S)]/(A - B) \\ &= [Ap_n(S\epsilon Y) - p_n(S\epsilon T_1)]/(A - B). \end{aligned} \tag{4.6}$$

Now, $0 \leqq \Phi_i(S) \leqq p_n(S\epsilon Y) \leqq 1$ for $i = 1, 2$, and $\Phi_1(S) + \Phi_2(S) = p_n(S\epsilon Y)$.

We wish to show $\Psi_1(T) \leqq 1, \Psi_2(T) \leqq 1$. The two proofs are similar and we will prove the first. From (4.3) and (4.2) we see that

$$\Psi_1(T) = A\psi_1(T) + \psi_2(T) = \sup_S p_n(S\epsilon T)/p_n(S\epsilon Y)$$

and the last expression is at most 1 since $T \subset Y$. Furthermore, $\Psi_1(Y) = \Psi_2(Y) = 1$ since otherwise,

$$p_n(S\epsilon Y) < \Phi_1(S) + \Phi_2(S) = p_n(S\epsilon Y).$$

The ψ_i, and hence the Ψ_i, are linear combinations of $p_n(S_1\epsilon T)$ and $p_n(S_2\epsilon T)$ so that the Ψ_i are σ-additive. Similarly the Φ_i are σ-additive.

We now define $\{X_k\}$. Let the state space for $\{X_k\}$ be as in (i) of this theorem. For measurable $A \subset U_n - \{\epsilon\}$, let

$$(4.7)\quad P((\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n \varepsilon A, (X_{n+1}, X_{n+2}, \cdots) \varepsilon T) = P((\cdots, Y_{n-2}, Y_{n-1}) \varepsilon S, Y_n \varepsilon A, (Y_{n+1}, Y_{n+2}, \cdots) \varepsilon T).$$

For $i = 1, 2$, let

$$(4.8)\quad P((\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n = \epsilon_i) = \Phi_i(S)$$

and

$$(4.9)\quad P((X_{n+1}, X_{n+2}, \cdots) \varepsilon T \mid (\cdots, X_{n-2}, X_{n-1}) \varepsilon S, X_n = \epsilon_i) = \Psi_i(T).$$

Finally, we must show that if $\delta \neq \epsilon$, then the rank of δ at any time is the same in $\{X_k\}$ as in $\{Y_k\}$. If $\delta \varepsilon U_n$ this is trivial. Let $\nu_k(\delta)$ and $\nu_k^*(\delta)$ be the ranks of δ in $\{Y_k\}$ and $\{X_k\}$, respectively, at time $k \neq n$.

Trivially $\nu_k^*(\delta) \geqq \nu_k(\delta)$. We assume $\delta \varepsilon U_k$, $k < n$. Let

$$p_n(S\delta T\epsilon U) = P[(\cdots, Y_{k-2}, Y_{k-1}) \varepsilon S, \quad Y_k = \delta, \quad (Y_{k+1}, \cdots, Y_{n-1}) \varepsilon T, \quad Y_n = \epsilon, \quad (Y_{n+1}, Y_{n+2}, \cdots) \varepsilon U]$$

for appropriate S, T and U. Then, by (4.5) and (4.8)

$$\begin{aligned} P((\cdots, X_{k-2}, X_{k-1}) \varepsilon S, X_k = \delta, (X_{k+1}, \cdots, X_{n-1}) \varepsilon T, X_n = \epsilon_1) &= [p_n(S\delta T\epsilon T_1) - Bp_n(S\delta T\epsilon Y)]/(A - B) \\ &= \textstyle\sum_{i=1}^{\nu_k(\delta)} [f_i(S)g_i(T\epsilon T_1) - Bf_i(S)g_i(T\epsilon Y)]/(A - B) \\ &= \textstyle\sum_{i=1}^{\nu_k(\delta)} f_i(S)G_i(T) \end{aligned}$$

and similarly if ϵ_1 is replaced by ϵ_2 in the above. Hence $\nu_k^*(\delta) \leqq \nu_k(\delta)$ so $\nu_k^*(\delta) = \nu_k(\delta)$.

The point of this proof is the linearity of the probabilities in $p_n(S\delta T\epsilon T_1)$ and $p_n(S\delta T\epsilon Y)$. But if $\delta \varepsilon U_k$ for $k > n$ we have only to consider the $\Psi_i(T\delta U)$ which are linear in $p_n(S_1\epsilon T\delta U)$ and $p_n(S_2\epsilon T\delta U)$ so the proof is similar.

From Theorem 2 and the corollary to Theorem 1 we obtain the

COROLLARY. *Under the conditions of Theorem 1 it is possible to construct $\{X_k\}$ satisfying the conclusions for which every state $\delta \neq \epsilon$ in $\{Y_k\}$ which is of rank at most 2 has its rank preserved in $\{X_k\}$.*

PROOF. Since only states of positive probability have ranks, we know that at most a countable number of states are of rank 2. We carry out the construction of Theorem 2 to obtain a process $\{Z_k\}$ such that $Y_k = h(Z_k)$ and each rank 2 state of $\{Y_k\}$ is the image under h of two Markovian states of $\{Z_k\}$. Furthermore we recall that the ranks of other states are preserved.

Secondly, carry out the construction of the corollary to Theorem 1 to obtain a process $\{W_k\}$ such that $Z_k = g(W_k)$. Recall that this preserves Markovianness in $\{W_k\}$ of states in $\{Z_k\}$. Furthermore, the only states of $\{W_k\}$ for which g is not the identity are those mapping into ϵ.

Let $m(\delta) = h(\delta)$ if $g(\delta) \neq \epsilon$ and $m(\delta) = g(\delta)$ if $g(\delta) = \epsilon$. Set $X_k = m(W_k)$. Then the obvious f yields $Y_k = f(X_k)$ and $\{X_k\}$ and f have all the properties claimed.

5. The stationary case. With the assumption of stationarity we drop the subscripts n on p_n and U_n.

THEOREM 3. *Let $\{Y_k\}$ be a stationary process satisfying the conditions of Theorem 2. Then, there exists a stationary process $\{X_k\}$ satisfying the conclusions of Theorem 2.*

PROOF. The proof of Theorem 2 is valid for $k = 0, \pm 1, \cdots$ and for $k = 1, 2, \cdots$. In considering the stationary case, we require that it be possible to choose $S_1 \varepsilon \mathcal{S}$, $S_2 \varepsilon \mathcal{S}$, $T_1 \varepsilon \mathfrak{J}$ for (4.1) independent of the time at which ϵ occurs. For this reason we prove the theorem for $k = 0, \pm 1, \cdots$. If $k = 1, 2, \cdots$, extend $\{Y_k\}$ to $k = 0, \pm 1, \cdots$ and after constructing $\{X_k\}$, restrict back to $k = 1, 2, \cdots$.

Let $\{X_k\}$ have state space $(U - \{\epsilon\}) \cup \{\epsilon_1, \epsilon_2\}$. Let s be a sequence of states of $\{X_k\}$. Fix n and assume $s_n = \epsilon_i$ ($i = 1$ or 2). Let $s_k^{(n)} = s_k$ if $k \neq n$ and $s_n^{(n)} = \epsilon_j$ ($j \neq i$). We will say that the set of sequences S is ϵ-free if $s \varepsilon S$ implies $s^{(n)} \varepsilon S$ for all n such that $s_n = \epsilon_i$ for some $i = 1, 2$.

Assume $S \varepsilon \mathcal{S}$ and $T \varepsilon \mathfrak{J}$ are ϵ-free. Define probabilities as in (4.7), (4.8) and (4.9). It suffices to define $P((X_{k+1}, \cdots, X_{n-1}) \varepsilon T, X_n = \epsilon_i)$. $(\cdots, X_{k-2}, X_{k-1}) \varepsilon S, X_k = \epsilon_j)$ for $i, j = 1, 2$. We will do this for $i = 1$. The proof is similar for $i = 2$.

With the obvious extension of the notation of Section 4, we have

$$(5.1)\quad \begin{aligned} P((\cdots, X_{k-2}, X_{k-1}) \varepsilon S, Y_k = \epsilon, (X_{k+1}, \cdots, X_{n-1}) \varepsilon T, X_n = \epsilon_1) \\ = [p(S\epsilon T\epsilon T_1) - Bp(S\epsilon T\epsilon Y)]/(A - B) \end{aligned}$$

by (4.4) and (4.7). Denote the left side of (5.1) by $\Gamma(S\epsilon T\epsilon_1)$. By (4.4),

$$\begin{aligned} \Gamma(S\epsilon T\epsilon_1) = [\Phi_1(S)\Psi_1(T\epsilon T_1) + \Phi_2(S)\Psi_2(T\epsilon T_1)]/(A - B) \\ - B[\Phi_1(S)\Psi_1(T\epsilon Y) + \Phi_2(S)\Psi_2(T\epsilon Y)]/(A - B) \end{aligned}$$

so that

$$\begin{aligned} \Gamma(S\epsilon T\epsilon_1) &= \Phi_1(S)[\Psi_1(T\epsilon T_1) - B\Psi_1(T\epsilon Y)]/(A - B) \\ &\quad + \Phi_2(S)[\Psi_2(T\epsilon T_1) - B\Psi_2(T\epsilon Y)]/(A - B) \\ &= \Phi_1(S)\Psi_1^*(T) + \Phi_2(S)\Psi_2^*(T), \end{aligned}$$

say. Additivity of Ψ_1^* and Ψ_2^* follow from their linearity in Ψ_1 and Ψ_2. Let

$\lambda^*(S) = \Phi_1(S)/\Phi_2(S) = [\lambda(S) - B]/[A - \lambda(S)]$. Then, $\sup_{S \varepsilon \mathcal{S}} \lambda^*(S) = \infty$ and $\inf_{S \varepsilon \mathcal{S}} \lambda^*(S) = 0$. Thus, if $\Psi_1^*(T) < 0$, there exists $S \varepsilon \mathcal{S}$ such that $\Gamma(S\epsilon T\epsilon_1) < 0$ and similarly if $\Psi_2^*(T) < 0$. Hence, $\Psi_1^* \geqq 0$ and $\Psi_2^* \geqq 0$. Also, $\Gamma(S\epsilon T\epsilon_1) \leqq \Phi_i(S) (i = 1, 2)$ so that, by the argument used in Section 4 for the Ψ_i we have the $\Psi_i^* \leqq 1$. We now set

$$P((X_{k+1}, \cdots, X_{n-1}) \varepsilon T, X_n = \epsilon_1 \mid (\cdots, X_{k-2}, X_{k-1}) \varepsilon S, X_k = \epsilon_j) = \Psi_j^*(T) \quad \text{for} \quad j = 1, 2.$$

Extending the notation in the obvious way we obtain

$$\Gamma(S\epsilon T\epsilon_2) = \Phi_1(S)\Psi_1^{**}(T) + \Phi_2(S)\Psi_2^{**}(T)$$

where the Ψ_i^{**} enjoy the same properties as the Ψ_i^* which are proved in the preceding paragraph. Then,

$$P((X_{k+1}, \cdots, X_{n-1}) \varepsilon T, X_n = \epsilon_2 \mid (\cdots, X_{k-2}, X_{k-1}) \varepsilon S, X_k = \epsilon_j) = \Psi_j^{**}(T)$$

for $j = 1, 2$.

REFERENCES

[1] CARLYLE, J. W. (1967). Identification of state calculable functions of finite Markov chains. *Ann. Math. Statist.* **38** 201–205.

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

[3] 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.

[4] DHARMADHIKARI, S. W. (1967). Splitting a single state of a stationary process into Markovian states. *Ann. Math. Statist.* **39** 1069–1077.

[5] FOX, MARTIN, and RUBIN, HERMAN (1965). A stochastic model for the evolution of cloud cover-estimation of parameters and goodness of fit based on Boston data. Michigan State University RM-132.

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

[7] RUBIN, HERMAN (1967). Representations of non-negative continuous functions on product spaces. *Bull. Amer. Math. Soc.* **74** 332–333.