The Annals of Mathematical Statistics
1968, Vol. 38, No. 3, 1069–1077

SPLITTING A SINGLE STATE OF A STATIONARY PROCESS INTO MARKOVIAN STATES[1]

BY S. W. DHARMADHIKARI

Michigan State University and Indian Statistical Institute

1. Introduction and summary. Let $\{Y_n, n \geqq 1\}$ be a stationary process with a finite state-space J. Let δ denote a state of J and let s, t denote finite sequences of states of J. If $s = (\delta_1, \cdots, \delta_n)$, let $p(s) = P[(Y_1, \cdots, Y_n) = s]$. The *rank* $n(\delta)$ of a state δ is defined to be the largest integer n such that we can find $2n$ sequences $s_1, \cdots, s_n, t_1, \cdots, t_n$ such that the $n \times n$ matrix $\|p(s_i\delta t_j)\|$ is non-singular. The number $n(\delta)$ was first defined by Gilbert [5] and the term rank was first used by Fox and Rubin [4]. A state δ is called *Markovian* if $n(\delta) = 1$. It is easy to check that δ is Markovian if, and only if, $p(s\delta t) = p(s\delta)p(\delta t)/p(\delta)$ for all s and t.

Suppose that μ is a fixed state of J. Let $J' = J - \{\mu\}$. Assume that $n(\mu) < \infty$. Fox and Rubin have shown that there exists a stationary process $\{X_n\}$ with a countable state-space $I = J' \cup J''$ and a function f on I onto J such that (a) $f(i) = \mu$ if $i \varepsilon J''$ and $f(\delta) = \delta$ if $\delta \varepsilon J'$; (b) states of J'' are Markovian states of $\{X_n\}$; and (c) $\{Y_n\}$ and $\{f(X_n)\}$ have the same distribution. Gilbert [5] has shown that J'' must have at least $n(\mu)$ elements whereas Fox and Rubin [4] have given an example to show that J'' cannot always be chosen to be finite. For $\delta \varepsilon J'$ let $\nu(\delta)$ denote the rank of δ in $\{X_n\}$. In general $\nu(\delta) \geqq n(\delta)$. But Fox and Rubin have shown that $\{X_n\}$ can be constructed in such a way that $\nu(\delta) = 1$ whenever $n(\delta) = 1$. Finally they have shown that, if $n(\mu) = 2$, then $\{X_n\}$ can be chosen in such a way that J'' has 2 elements and $\nu(\delta) = n(\delta)$ for all $\delta \varepsilon J'$.

In this paper we give some conditions under which J'' can be chosen to be finite. These conditions are similar to those imposed in [2]. It is shown that $\{X_n\}$ can be constructed in such a way that, for $\delta \varepsilon J'$, $\nu(\delta) = 1$ whenever $n(\delta) = 1$. Finally it is proved that if $N(\mu) = n(\mu)$, then $\nu(\delta) = n(\delta)$ for all $\delta \varepsilon J'$. This generalizes the result proved by Fox and Rubin for the case $n(\mu) = 2$. However, they have given results for the non-stationary case also. The results of this paper were partially reported in [3].

2. The main result. We recall that μ is a fixed state of J of finite rank. The finiteness of $n(\mu)$ can be used (see [1] and [2]) to find $2n(\mu)$ sequences $s_{\mu i}$, $t_{\mu i}$, $i = 1, \cdots, n(\mu)$, such that the matrix $\|p(s_{\mu i}\mu t_{\mu j})\|$ is non-singular. Let $\pi_\mu(t)$ denote the row vector whose ith element is $p(s_{\mu i}\mu t)$. Then, for every s, there is a unique row vector $\alpha_\mu(s)$ such that, for all t,

$$p(s\mu t) = \alpha_\mu(s)\pi_\mu'(t). \tag{1}$$

Received 30 October 1967.

[1] Work partially done while the author was at the University of Arizona.

Let $\mathcal{C}(\alpha_\mu)$ denote the closed convex cone generated by the vectors $\alpha_\mu(s)$ where s varies over all finite sequences of states of J. Define $\mathcal{C}(\pi_\mu)$ similarly. If $\mathcal{C}^+$ denotes the dual cone of a cone $\mathcal{C}$, then (1) shows that $\mathcal{C}(\alpha_\mu) \subset [\mathcal{C}(\pi_\mu)]^+$.

Let H_m denote the set of all sequences of length m of states of J. We interpret H_0 as the set consisting of the empty sequence $\varnothing$. For conventions regarding $\varnothing$, see [1]. Let $H = \bigcup_{m=0}^{\infty} H_m$. Define H_m' and H' from J' similarly.

For notational compactness we adopt the conventions $t\varnothing = t$ and $\varnothing t = t$. For $u \varepsilon H$, let $A_\mu(u)$ denote the $n(\mu) \times n(\mu)$ matrix whose ith row is $\alpha_\mu(s_{\mu i}\mu u)$. Then equation (1) and the uniqueness of $\alpha_\mu(s)$ can be used to show that for all $s \varepsilon H$, $t \varepsilon H$ and $u \varepsilon H$,

$$(2) \qquad \alpha_\mu(s)A_\mu(u) = \alpha_\mu(s\mu u) \quad \text{and} \quad A_\mu(u)\pi_\mu'(t) = \pi_\mu'(u\mu t).$$

The state μ of finite rank will be split into a finite number of Markovian states under the following condition.

Condition C_μ. There is a convex polyhedral cone $\mathcal{C}_\mu$ generated by $N(\mu)$ non-zero vectors $\beta_{\mu i}$, $i = 1, \cdots, N(\mu)$, such that

$$(3) \qquad \mathcal{C}(\alpha_\mu) \subset \mathcal{C}_\mu \subset [\mathcal{C}(\pi_\mu)]^+;$$

$$(4) \qquad \beta_{\mu i}A_\mu(u) \varepsilon \mathcal{C}_\mu \quad \text{for all} \quad i \quad \text{and all} \quad u \varepsilon H'.$$

It is a straightforward consequence of (2) that if either $\mathcal{C}(\alpha_\mu)$ or $\mathcal{C}(\pi_\mu)$ is polyhedral then condition C_μ holds with $\mathcal{C}_\mu = \mathcal{C}(\alpha_\mu)$ or $\mathcal{C}_\mu = [\mathcal{C}(\pi_\mu)]^+$.

We now assume that condition C_μ holds. Let B_μ be the $N(\mu) \times n(\mu)$ matrix whose ith row is $\beta_{\mu i}$. It follows from (3) that for every $u \varepsilon H'$ there is a non-negative vector $q_\mu(u)$ such that $q_\mu(u)B_\mu = \alpha_\mu(u)$. Further (4) shows that, for every $u \varepsilon H'$, we can choose a non-negative matrix $M_\mu(u)$ such that $B_\mu A_\mu(u) = M_\mu(u)B_\mu$.

Observe that $q_\mu(\varnothing)$ has been defined. For sequences $s \varepsilon (H' - H)$, define $q_\mu(s)$ by induction as follows.

$$(5) \qquad q_\mu(s\mu u) = q_\mu(s)M_\mu(u), \qquad u \varepsilon H'.$$

Lemma 1. *For all* $s \varepsilon H$, $\alpha_\mu(s) = q_\mu(s)B_\mu$.

Proof. The lemma holds for all $s \varepsilon H'$ and hence for sequences of length zero in H. Suppose it holds for all sequences in H of length $\leqq n$. Let s have length $(n + 1)$ and belong to $H - H'$. Then $s = s'\mu u$ where s' has length $\leqq n$ and $u \varepsilon H'$. Therefore

$$q_\mu(s)B_\mu = q_\mu(s')M_\mu(u)B_\mu = q_\mu(s')B_\mu A_\mu(u) = \alpha_\mu(s')A_\mu(u) = \alpha_\mu(s'\mu u) = \alpha_\mu(s).$$

The lemma thus follows by induction.

The Markov-state $\{X_n\}$ that will be constructed will have state-space $I = J' \cup J''$ where $J'' = \{\mu_i, i = 1, \cdots, N(\mu)\}$. If $q_{\mu i}(s)$ denotes the ith entry of $q_\mu(s)$ then, for a sequence $s \varepsilon H_n$, we want to have

$$q_{\mu i}(s) = P[(Y_1, \cdots, Y_n) = s, X_{n+1} = \mu_i].$$

But we also want $\{X_n\}$ to be stationary. This means that $q_\mu(s)$ must satisfy certain stationarity conditions. We proceed to show that a choice satisfying these conditions can be made.

We note that the vectors $\beta_{\mu i}$ are non-zero. This easily implies that $\beta_{\mu i}\pi_\mu'(\varnothing) > 0$. Therefore the $\beta_{\mu i}$'s can be chosen in such a way that $\beta_\mu \pi_\mu'(\varnothing) = e_\mu$, where e_μ is the column vector all of whose $N(\mu)$ elements equal 1. We assume that this has been done. Then, for all $s \varepsilon H$,

$$(6) \qquad q_\mu(s)e_\mu = q_\mu(s)B_\mu \pi_\mu'(\varnothing) = \alpha_\mu(s)\pi_\mu'(\varnothing) = p(s\mu).$$

For $s \varepsilon H$, define $q_\mu^{m}(s) = \sum_{t\varepsilon H_m} q_\mu(ts)$. Then (6) and the stationarity of $\{Y_n\}$ imply that

$$(7) \qquad q_\mu^{m}(s)e_\mu = p(s\mu)$$

for all $s \varepsilon H$ and for $m = 1, 2, \cdots$. It follows from (7) that $0 \leqq q_\mu^{m}(s) \leqq e_\mu'$. Define

$$\theta_n(s) = n^{-1}\sum_{m=1}^{n} q_\mu^{(m)}(s).$$

Then $0 \leqq \theta_n(s) \leqq e_\mu'$ for all n and s. Since the number of sequences s is countable, there is a single subsequence $\{n_k, k \geqq 1\}$ of positive integers such that $\bar{q}_\mu(s) = \lim_{k\to\infty} \theta_{n_k}(s)$ exists for all $s \varepsilon H$.

LEMMA 2. *For all* $s \varepsilon H$, $\bar{q}_\mu(s)B_\mu = \alpha_\mu(s)$.

PROOF. The uniqueness of $\alpha_\mu(s)$ and the stationarity of $\{Y_n\}$ show that

$$q_\mu^{m}(s)B_\mu = \sum_{t\varepsilon H_m} \alpha_\mu(ts) = \alpha_\mu(s).$$

Therefore $\theta_n(s)B_\mu = \alpha_\mu(s)$. This proves the lemma.

LEMMA 3. *For all* $s \varepsilon H$, $\bar{q}_\mu(s) = \sum_{t\varepsilon H_m} \bar{q}_\mu(ts)$.

PROOF. If the lemma holds for $m = 1$, then

$$\sum_{t\varepsilon H_{m+1}} \bar{q}_\mu(ts) = \sum_{u\varepsilon H_m}\sum_{v\varepsilon H_1} \bar{q}_\mu(vus) = \sum_{u\varepsilon H_m} \bar{q}_\mu(us)$$

and the lemma follows by induction for all m. It is thus enough to prove the lemma for $m = 1$. Observe that

$$q_\mu^{(m+1)}(s) = \sum_{u\varepsilon H_{m+1}} q_\mu(us) = \sum_{t\varepsilon H_1}\sum_{v\varepsilon H_m} q_\mu(vts) = \sum_{t\varepsilon H_1} q_\mu^{(m)}(ts).$$

Summing for $m = 1, \cdots, n$ and dividing by n, we get

$$\theta_n(s) + n^{-1}[q_\mu^{(n+1)}(s) - q_\mu^{(1)}(s)] = \sum_{t\varepsilon H_1} \theta_n(ts).$$

Replacing n by n_k and letting $k \to \infty$ we get the lemma for $m = 1$. This proves the lemma.

LEMMA 4. *For all* $s \varepsilon H$ *and* $u \varepsilon H'$, $\bar{q}_\mu(s\mu u) = \bar{q}_\mu(s)M_\mu(u)$.

PROOF. Straightforward.

The preceding three lemmata show that $\bar{q}_\mu(s)$ has all the properties of $q_\mu(s)$ and also has the required stationarity properties. From now on we will use $\bar{q}_\mu(s)$ without any reference to the original $q(s)$ and will suppress the bar over q.

Recall that $I = J' \cup J''$, where $J'' = \{\mu_i, i = 1, \cdots, N(\mu)\}$. Let G_m be the set of all sequences of length m of states of I. Let $G = \bigcup_{m=0}^{\infty} G_m$. Define F_m and F similarly from $I \cup \{\mu\}$.

For $u \varepsilon H'$, let $r_{\mu i}(u) = \beta_{\mu i}\pi_\mu'(u)$. Recall that $\beta_{\mu i}$'s have been chosen in such a way that $r_{\mu i}(\varnothing) = 1$ for all i. For $t \varepsilon G$, we define $r_{\mu i}(t)$ by induction as follows.

$$r_{\mu i}(u\mu_j t) = [M\mu(u)]_{ij} r_{\mu j}(t), \tag{8}$$

where $u \varepsilon H'$ and $[M_\mu(u)]_{ij}$ denotes the (i, j)th term in $M_\mu(u)$. For $t \varepsilon F$, define $r_{\mu i}(t)$ by induction as follows.

$$r_{\mu i}(u\mu t) = \sum_{j=1}^{N(\mu)} r_{\mu i}(u\mu_j t), \qquad u \varepsilon G.$$

Finally $r_\mu(t)$ will denote the column vector whose ith entry is $r_{\mu i}(t)$.

LEMMA 5. *For all $t \varepsilon H$, $r_\mu(t) = B_\mu \pi_\mu'(t)$.*

PROOF. Straightforward by induction.

LEMMA 6. *For all $u \varepsilon F$ and $v \varepsilon F$,*

$$r_\mu(u\mu v) = \sum_{j=1}^{N(\mu)} r_\mu(u\mu_j v).$$

PROOF. The definitions yield the lemma for $u \varepsilon G$. For $u \varepsilon F - G$, the lemma follows easily by induction.

LEMMA 7. *For all $u \varepsilon F$ and $v \varepsilon F$,*

$$r_{\mu i}(u\mu_j v) = r_{\mu i}(u\mu_j) r_{\mu j}(v).$$

PROOF. For $u \varepsilon H'$ and $v \varepsilon G$, the lemma follows from definitions. For $u \varepsilon F - H'$ and $v \varepsilon F - G$, we can use induction and Lemma 6 to prove the lemma.

LEMMA 8. *For all $t \varepsilon F$*

$$\sum_{u \varepsilon G_m} r_\mu(tu) = r_\mu(t).$$

PROOF. As in the case of Lemma 3 it is sufficient to prove the lemma for $m = 1$. If $t \varepsilon H$, then

$$\begin{aligned}\sum_{u\varepsilon G_1} r_\mu(tu) &= \sum_{j=1}^{N(\mu)} r_\mu(t\mu_j) + \sum_{u\varepsilon H_1'} r_\mu(tu) = r_\mu(t\mu) + \sum_{u\varepsilon H_1'} r_\mu(tu)\\ &= \sum_{u\varepsilon H_1} r_\mu(tu) = \sum_{u\varepsilon H_1} B_\mu\pi_\mu'(tu) = B_\mu \sum_{u\varepsilon H_1} \pi_\mu'(tu) = B_\mu\pi_\mu'(t)\\ &= r_\mu(t).\end{aligned}$$

If $t \varepsilon F - H$ then $t = v\mu_j w$ where $v \varepsilon F$ and $w \varepsilon H$. We then have

$$\begin{aligned}\sum_{u\varepsilon G_1} r_\mu(ru) &= \sum_{u\varepsilon G_1} r_\mu(v\mu_j wu) = \sum_{u\varepsilon G_1} r_\mu(v\mu_j) r_{\mu j}(wu)\\ &= r_\mu(v\mu_j) \sum_{u\varepsilon G_1} r_{\mu j}(wu) = r_\mu(v\mu_j) r_{\mu j}(w) = r_\mu(v\mu_j w) = r_\mu(t).\end{aligned}$$

This proves the lemma.

LEMMA 9. *For all $s \varepsilon H$ and $t \varepsilon H$,*

$$q_\mu(s) r_\mu(t) = p(s\mu t).$$

PROOF. $q_\mu(s) r_\mu(t) = q_\mu(s) B_\mu \pi_\mu'(t) = \alpha_\mu(s)\pi_\mu'(t) = p(s\mu t)$.

We are now ready to define the underlying stochastic process $\{X_n\}$ with state-space I. Define the finite dimensional distributions as follows.

$$(9) \quad P[(X_1, \cdots, X_n) = u] = p(u), \text{ if } u \,\varepsilon\, H_n', \quad \text{and}$$
$$P[(X_1, \cdots, X_n) = u\mu_i t] = q_{\mu i}(u) r_{\mu i}(t), \text{ if } u \,\varepsilon\, H' \text{ and } t \,\varepsilon\, G.$$

THEOREM 1. *The finite dimensional distributions defined by* (9) *are consistent and the resulting process* $\{X_n\}$ *is stationary. Every* μ_i *is a Markovian state of* $\{X_n\}$. *Moreover, if* $f(\mu_i) = \mu$ *for all* i *and* $f(\delta) = \delta$ *for* $\delta \,\varepsilon\, J'$, *then* $\{Y_n\}$ *and* $f(X_n)\}$ *have the same distribution.*

PROOF. (a) *Consistency.* First let $u \,\varepsilon\, H_n'$. Then

$$\begin{aligned}
\textstyle\sum_{v \varepsilon G_1} & P[(X_1, \cdots, X_{n+1}) = uv] \\
&= \textstyle\sum_{i=1}^{N(\mu)} P[(X_1, \cdots, X_{n+1}) = u\mu_i] + \sum_{v\varepsilon H_1'} P[(X_1, \cdots, X_{n+1}) = uv] \\
&= \textstyle\sum_{i=1}^{N(\mu)} q_{\mu i}(u) + \sum_{v\varepsilon H_1'} p(uv) = q_\mu(u) r_\mu(\varnothing) + \sum_{v\varepsilon H_1'} p(uv) \\
&= p(u\mu) + \textstyle\sum_{v\varepsilon H_1'} p(uv) = \sum_{v\varepsilon H} p(uv) = p(u) \\
&= P[(X_1, \cdots, X_n) = u].
\end{aligned}$$

Next let $s = u\mu_i v$ where $u \,\varepsilon\, H'$ and $v \,\varepsilon\, G$. Then

$$\begin{aligned}
\textstyle\sum_{w\varepsilon G_1} & P[(X_1, \cdots, X_{n+1}) = sw] \\
&= \textstyle\sum_{w\varepsilon G_1} P[(X_1, \cdots, X_{n+1}) = u\mu_i vw] = \sum_{w\varepsilon G_1} q_{\mu i}(u) r_{\mu i}(vw) \\
&= q_{\mu i}(u) \textstyle\sum_{w\varepsilon G_1} r_{\mu i}(vw) = q_{\mu i}(u) r_{\mu i}(v) = P[(X_1, \cdots, X_n) = u\mu_i v].
\end{aligned}$$

This verifies consistency

(b) *Stationarity.* First let $u \,\varepsilon\, H_n'$. Then

$$\begin{aligned}
P[(X_2, & \cdots, X_{n+1}) = u] \\
&= \textstyle\sum_{v\varepsilon G_1} P[(X_1, \cdots, X_{n+1}) = vu] \\
&= \textstyle\sum_{i=1}^{N(\mu)} P[(X_1, \cdots, X_{n+1}) = \mu_i u] + \sum_{v\varepsilon H_1'} P[(X_1, \cdots, X_{n+1}) = vu] \\
&= \textstyle\sum_{i=1}^{N(\mu)} q_{\mu i}(\varnothing) r_{\mu i}(u) + \sum_{v\varepsilon H_1'} p(vu) = p(\mu u) + \sum_{v\varepsilon H_1'} p(vu) \\
&= \textstyle\sum_{v\varepsilon H_1} p(vu) = p(u).
\end{aligned}$$

Next let $s = u\mu_i v$ where $u \,\varepsilon\, H'$ and $v \,\varepsilon\, G$. Then

$$\begin{aligned}
P[(X_2, & \cdots, X_{n+1}) = s] \\
&= \textstyle\sum_{w\varepsilon G_1} P(X_1, \cdots, X_{n+1}) = wu\mu_i v] \\
&= \textstyle\sum_{j=1}^{N(\mu)} P[(X_1, \cdots, X_{n+1}) = \mu_j u\mu_i v] + \sum_{w\varepsilon H_1'} P[(X_1, \cdots, X_{n+1}) = wu\mu_i v] \\
&= \textstyle\sum_{j=1}^{N(\mu)} q_{\mu j}(\varnothing) r_{\mu j}(u\mu_i v) + \sum_{w\varepsilon H_1'} q_{\mu i}(wu) r_{\mu i}(v) \\
&= \textstyle\sum_{j=1}^{N(\mu)} q_{\mu j}(\varnothing) [M_\mu(u)]_{ji} r_{\mu i}(v) + \sum_{w\varepsilon H_1'} q_{\mu i}(wu) r_{\mu i}(v) \\
&= q_{\mu i}(\mu u) r_{\mu i}(v) + \textstyle\sum_{w\varepsilon H_1'} q_{\mu i}(wu) r_{\mu i}(v) \\
&= [\textstyle\sum_{w\varepsilon H_1} q_{\mu i}(wu)] r_{\mu i}(v) = q_{\mu i}(u) r_{\mu i}(v) = P[(X_1, \cdots, X_n) = u\mu_i v].
\end{aligned}$$

This checks stationarity.

(c) The second statement of the theorem follows easily from (9) and the last statement follows easily from Lemma 9.

3. Markovian states of $\{Y_n\}$ can be kept Markovian. In Section 2 the state μ of $\{Y_n\}$ was split into $N(\mu)$ Markovian states of $\{X_n\}$. We will use the same letter p to denote the probability function of the process $\{X_n\}$. For $\delta \varepsilon J'$, let $\nu(\delta)$ be the rank of δ in $\{X_n\}$. For $u \varepsilon H$ and $t \varepsilon H$, the probability $p(u\delta t)$ can be obtained by adding probabilities $p(v\delta w)$ where v and w vary over certain subsets of G. It therefore follows that $\nu(\delta) \geqq n(\delta)$. It is desirable to construct $\{X_n\}$ in such a way that $\nu(\delta) = n(\delta)$ for all $\delta \varepsilon J'$. Whether this can be achieved under the condition C_μ is an open question. In this section we show that if $n(\delta) = 1$ then we can arrange to have $\nu(\delta) = 1$. We will exhibit this only for one Markovian state.

Let ξ be a fixed state of J' and let $n(\xi) = 1$. In this section s will denote a sequence in H' which does not involve ξ. We define $q_\mu(u)$ for $u = s$ and ξs as before. We also define $M_\mu(s)$ as before. For $u \varepsilon H'$ let $q_\mu(u\xi s) = p(u\xi)q_\mu(\xi s)/p(\xi)$. For sequences t in $H - H'$ which do not involve ξ define $q_\mu(t)$ by $q_\mu(u\mu s) = q_\mu(u)M_\mu(s)$. For $t \varepsilon H'$ define $r_\mu(t)$ as before. Complete the definition of $M_\mu(t)$ for $t \varepsilon H'$ as follows:

$$M_\mu(u\xi s) = r_\mu(u\xi)q_\mu(\xi s)/p(\xi), \qquad u \varepsilon H'.$$

We can now define $q_\mu(t)$ for all sequences t in H which involve both μ and ξ by using (5). Finally we can use (8) to define $r_\mu(t)$ for all sequences t in $F - H'$.

It is straightforward to verify that all the lemmata of Section 2 hold for the above choices of q_μ and r_μ. It is also easy to prove that for $t \varepsilon G$ and $u \varepsilon G$,

$$r_\mu(u\xi t) = r_\mu(u\xi)p(\xi t)/p(\xi),$$

and for $v \varepsilon H$ and $w \varepsilon H$,

$$q_\mu(v\xi w) = p(v\xi)q_\mu(\xi w)/p(\xi).$$

THEOREM 2. *The process $\{X_n\}$ given by Theorem 1 through the above choices of q_μ and r_μ has $\nu(\xi) = 1$.*

PROOF. We must show that, for $t \varepsilon G$ and $u \varepsilon G$,

$$p(t\xi u) = p(t\xi)p(\xi u)/p(\xi). \tag{10}$$

(a) If $t \varepsilon H'$ and $u \varepsilon H'$, then (10) follows because $n(\xi) = 1$.

(b) Let $t \varepsilon G - H'$ and $u \varepsilon G$. Then $t = v\mu_i w$ where $v \varepsilon H'$ and $w \varepsilon G$. We have

$$p(t\xi u) = p(v\mu_i w\xi u) = q_{\mu i}(v)r_{\mu i}(w\xi u) = q_{\mu i}(v)r_{\mu i}(w\xi)p(\xi u)/p(\xi)$$
$$= p(v\mu_i w\xi)p(\xi u)/p(\xi) = p(t\xi)p(\xi u)/p(\xi),$$

which is the same as (10).

(c) Let $t \varepsilon H'$ and $u \varepsilon G - H'$. Then $u = v\mu_i w$ where $v \varepsilon H'$ and $w \varepsilon G$. We have

$$p(t\xi u) = p(t\xi v\mu_i w) = q_{\mu i}(t\xi v)r_{\mu i}(w) = p(t\xi)q_{\mu i}(\xi v)r_{\mu i}(w)/p(\xi)$$
$$= p(t\xi)p(\xi v\mu_i w)/p(\xi) = p(t\xi)p(\xi u)/p(\xi).$$

This verifies (10) and completes the proof of the theorem.

4. The regular case. In this section we assume that conditions C_μ hold with $N(\mu) = n(\mu)$. We call this the regular case. In this case the matrix B_μ is non-singular and therefore a vector $q_\mu(s)$, non-negative or not, satisfying $q_\mu(s)B_\mu = \alpha_\mu(s)$ is uniquely determined as $q_\mu(s) = \alpha_\mu(s)B_\mu^{-1}$. Similarly $M_\mu(u)$ is uniquely determined. Non-negativity of $q_\mu(s)$ and $M_\mu(u)$ is guaranteed by condition C_μ and the stationarity properties are guaranteed by Lemma 3. Since $M_\mu(u)$ is unique, so is $r_\mu(t)$ for all $t \varepsilon F$.

Suppose now $\delta \varepsilon J'$ and let $n(\delta) < \infty$. For $k = 1, \cdots, n(\delta)$, choose $s_{\delta k}$, $t_{\delta k}$ and, for $t \varepsilon H$, vectors $\pi_\delta(t)$ and $\alpha_\delta(t)$ as in the first paragraph of Section 2. We note that we may choose the $s_{\delta k}$'s and the $t_{\delta k}$'s in such a way that they belong to H'. This is because, for $s \varepsilon H$, $p(s)$ can be obtained by linear combinations of $p(u)$ where u varies over some subset of H'. For $s \varepsilon H$, $A_{\mu\delta}(s)$ will denote the $n(\mu) \times n(\delta)$ matrix whose ith row is $\alpha_\delta(s_{\mu i}\mu s)$. The matrices $A_{\delta\mu}(s)$ are defined similarly. It can be shown from the uniqueness of α that for all $s \varepsilon H$, $t \varepsilon H$, $u \varepsilon H$ and $v \varepsilon H$

$$\alpha_\mu(s)A_{\mu\delta}(u) = \alpha_\delta(s\mu u),$$
$$A_{\mu\delta}(u)\pi_\delta'(t) = \pi_\mu'(u\delta t),$$
$$A_{\mu\delta}(u)A_{\delta\mu}(v) = A_\mu(u\delta v).$$

In the above results μ and δ can be interchanged.

Suppose $a_{\delta k}(s)$ denotes the kth element of $\alpha_\delta(s)$. We need two lemmata.

LEMMA 10. *Let $s \varepsilon H$ and $u \varepsilon H$. Then*

$$\sum_{k=1}^{n(\delta)} a_{\delta k}(s)q_\mu(s_{\delta k}\delta u) = q_\mu(s\delta u). \tag{11}$$

PROOF. The left side of (11) $= \sum_{k=1}^{n(\delta)} a_{\delta k}(s)\alpha_\mu(s_{\delta k}\delta u)B_\mu^{-1} = \alpha_\delta(s)A_{\delta\mu}(u)B_\mu^{-1}$
$= \alpha_\mu(s\delta u)B_\mu^{-1} = q_\mu(s\delta u).$

To state the next lemma we need to define $\alpha_\delta(s)$ for all $s \varepsilon F$ as follows. For $i = 1, \cdots, n(\mu)$ and $s \varepsilon H$, we define

$$\alpha_\delta(\mu_i s) = q_{\mu i}(\varnothing)\beta_{\mu i}A_{\mu\delta}(s).$$

For the remaining sequences in F, we define

$$\alpha_\delta(u\mu_i v) = p(u\mu_i)[q_{\mu i}(\varnothing)]^{-1}\alpha_\delta(\mu_i v), \qquad \text{where } v \varepsilon H.$$

LEMMA 11. *For all $s \varepsilon H$, $t \varepsilon H$ and $i, j = 1, \cdots, n(\mu)$,*

$$[M_\mu(s\delta t)]_{ij} = [q_{\mu i}(\varnothing)]^{-1}\sum_{k=1}^{n(\delta)} a_{\delta k}(\mu_i s)q_{\mu j}(s_{\delta k}\delta t).$$

PROOF.

$$\begin{aligned}
\sum_{j=1}^{n(\mu)} &[M_\mu(s\delta t)]_{ij}\beta_{\mu j} \\
&= \beta_{\mu i}A_\mu(s\delta t) = \beta_{\mu i}A_{\mu\delta}(s)A_{\delta\mu}(t) = [q_{\mu i}(\varnothing)]^{-1}\alpha_\delta(\mu_i s)A_{\delta\mu}(t) \\
&= [q_{\mu i}(\varnothing)]^{-1}\sum_{k=1}^{n(\delta)} a_{\delta k}(\mu_i s)\alpha_\mu(s_{\delta k}\delta t) \\
&= [q_{\mu i}(\varnothing)]^{-1}\sum_{k=1}^{n(\delta)} a_{\delta k}(\mu_i s)\sum_{j=1}^{n(\mu)} q_{\mu j}(s_{\delta k}\delta t)\beta_{\mu j} \\
&= \sum_{j=1}^{n(\delta)} [(q_{\mu i}(\varnothing))^{-1}\sum_{k=1}^{n(\delta)} a_{\delta k}(\mu_i s)q_{\mu j}(s_{\delta k}\delta t)]\beta_{\mu j}.
\end{aligned}$$

The result now follows from the linear independence of $\beta_{\mu j}$'s.

For $t \varepsilon G$ we now define $\pi_\delta(t)$ as the column vector whose kth entry is $p(s_{\delta k}\delta t)$, where this function p now refers to $\{X_n\}$.

THEOREM 3. *In the regular case, the process $\{X_n\}$ given by Theorem 1 is such that $\nu(\delta) = n(\delta)$ for all $\delta \varepsilon J'$.*

PROOF. If $n(\delta) = \infty$ then $\nu(\delta) = \infty$. So let $n(\delta) < \infty$. To show that $\nu(\delta) = n(\delta)$ we must verify that, for all $s \varepsilon G$ and $t \varepsilon G$,

$$p(s\delta t) = \alpha_\delta(s)\pi_\delta(t). \tag{12}$$

(a) If $s \varepsilon H'$ and $t \varepsilon H'$, there is nothing to prove.

(b) Let $s \varepsilon H'$ and $t \varepsilon G - H'$. Then $t = u\mu_i v$ where $v \varepsilon G$ and $u \varepsilon H'$. We have

$$\begin{aligned}
p(s\delta t) &= p(s\delta u\mu_i v) = q_{\mu i}(s\delta u)r_{\mu i}(v) = \sum_{k=1}^{n(\delta)} a_{\delta k}(s)q_{\mu i}(s_{\delta k}\delta u)r_{\mu i}(v) \\
&= \sum_{k=1}^{n(\delta)} a_{\delta k}(s)p(s_{\delta k}\delta u\mu_i v) = \alpha_\delta(s)\pi_\delta'(u\mu_i v) = \alpha_\delta(s)\pi_\delta'(t).
\end{aligned}$$

(c) Let $s \varepsilon G - H'$ and $t \varepsilon H'$. Write $s = u\mu_i v$ where $u \varepsilon G$ and $v \varepsilon H'$. Then

$$\begin{aligned}
p(s\delta t) &= p(u\mu_i v\delta t) = p(u\mu_i)r_{\mu i}(v\delta t) = p(u\mu_i)\beta_{\mu i}\,\pi_\mu'(v\delta t) = p(u\mu_i)\beta_{\mu i}A_{\mu\delta}(v)\pi_\delta'(t) \\
&= p(u\mu_i)[q_{\mu i}(\varnothing)]^{-1}\alpha_\delta(\mu_i v)\pi_\delta'(t) = \alpha_\delta(u\mu_i v)\pi_\delta'(t) = \alpha_\delta(s)\pi_\delta'(t).
\end{aligned}$$

(d) Let $s \varepsilon G - H'$ and $t \varepsilon G - H'$. Write $s = u\mu_i v$ and $t = w\mu_j y$ where $u \varepsilon G$, $v \varepsilon H'$, $w \varepsilon H'$ and $y \varepsilon G$. Then

$$\begin{aligned}
p(s\delta t) &= p(u\mu_i v\delta w\mu_j y) = p(u\mu_i)[M_\mu(v\delta w)]_{ij}r_{\mu j}(y) \\
&= p(u\mu_i)[q_{\mu i}\ \varnothing)]^{-1}\sum_{k=1}^{n(\delta)} a_{\delta k}(\mu_i v)q_{\mu j}(s_{\delta k}\delta w)r_{\mu j}(y) \\
&= p(u\mu_i)[q_{\mu i}(\varnothing)]^{-1}\sum_{k=1}^{n(\delta)} a_{\delta k}(\mu_i v)p(s_{\delta k}\delta w\mu_j y) \\
&= p(u\mu_i)[q_{\mu i}(\varnothing)]^{-1}\alpha_\delta(\mu_i v)\pi_\delta'(w\mu_j y) = \alpha_\delta(u\mu_i v)\pi_\delta'(w\mu_j y) = \alpha_\delta(s)\pi_\delta'(t).
\end{aligned}$$

This verifies (12) and completes the proof of the theorem.

COROLLARY. *If $n(\mu) = 2$, then we can split μ into two Markovian states in such a way that $\nu(\delta) = n(\delta)$ for all $\delta \varepsilon J'$.*

PROOF. It was shown on page 1037 of [2] that if $n(\mu) = 2$ then we are in the regular case. Hence the preceding theorem applies.

The result stated in the above corollary has been proved by Fox and Rubin [4]. However, they have considered the non-stationary case also whereas the present paper is restricted to the stationary case.

REFERENCES

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

[2] 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.

[3] DHARMADHIKARI, S. W. (1967). Markovianization of a single state of a stationary process. (Abstract). *Ann. Math. Statist.* **38** 1311.

[4] FOX, MARTIN and RUBIN, HERMAN. (1967). Functions of processes with Markovian states. *Ann. Math. Statist.* **39** 938–947.

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