# INVARIANTS UNDER MIXING WHICH GENERALIZE DE FINETTI'S THEOREM: CONTINUOUS TIME PARAMETER[1]

By David A. Freedman

*University of California, Berkeley*

**0. Introduction.** In a previous paper [4], the problem of characterizing mixtures of certain families of discrete-time stochastic processes was solved. The analysis will now be extended to continuous time. At least under a suitable continuity condition, the present discussion gives necessary and sufficient conditions for a process to be a mixture of stationary Markoff chains, or of processes with stationary, independent increments. Roughly, the law of a process is a weighted average of laws of stationary Markoff chains if and only if the probability that the process passes through a given finite sequence of states at given times depends only on the initial state, the number of transitions between each pair of states, and the length of time these transitions take (Theorem 2). The law of a process is a weighted average of laws of processes with stationary, independent increments if and only if the process has exchangeable increments (Theorem 3). This theorem was obtained, for real-valued processes, by a different method in Section 4.3 of Bühlmann (1960). The law of a process is a weighted average of laws of Brownian motions if and only if the process is $d$-isotropic (Definition 6, Theorem 4). The law $P$ of a process $\{X_t\}$ is a weighted average of laws of Poisson processes if and only if, for $i_j$ non-negative integers and $-\infty < t_j < t_{j+1} < \infty$, $P(X_{t_j} \geqq X_{t_{j-1}}) = 1$ and $P[X_{t_j} - X_{t_{j-1}} = i_j, 2 \leqq j \leqq n] \prod_{j=2}^{n} i_j!(t_j - t_{j-1})^{-i_j}$ is a function of $n$, $t_n - t_1$ and $i_2 + \cdots + i_n$ alone (Theorem 5). Theorem 6 makes precise the idea that the law of a process is a weighted average of laws of Poisson processes if and only if the process has exchangeable increments and its sample functions are counting-functions (Definition 7), and Theorem 7 gives equivalent conditions in terms of holding times.

These results follow easily from Theorems 1 and D.1 of Section 2, which constitute a long and technical discussion of the Kriloff-Bogoliouboff (1937) theory (see also Oxtoby (1952)) in the appropriate probability space.

Here are two results for discrete-time processes. The first does not follow from the present study formally, but can be proved by obvious modifications of the argument. Let $P$ be a probability on (the Borel subsets of) the space of bilateral sequences of real numbers. Then $P$ can be represented as a weighted average of probabilities on this space under each of which the coordinates are independent and Gaussian with mean 0 and common variance if and only if under $P$ any finite number of coordinates have a spherically symmetric joint distribution. The second result is proved as Lemma 10. The probability $P$ can be represented as a weighted average of probabilities under each of which the coordinates are independent random variables with common exponential distribution if and

Received February 5, 1962; revised May 30, 1963.

[1] Prepared with the partial support of the National Science Foundation, Grant G-14648.



Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.*
www.jstor.org

only if the coordinates are non-negative with $P$-probability 1, and for any natural number $n$, any non-negative numbers $x_\nu$, $-n \leqq \nu \leqq n$, the $P$-probability that the $\nu$th coordinate exceeds $x_\nu$, $-n \leqq \nu \leqq n$ depends only on $n$ and $\sum_{\nu=-n}^{n} x_\nu$.

The results and proofs are stated here for bilateral sequences, and for processes with time parameter ranging from $(-\infty, \infty)$. With obvious and minor changes, they hold for one-sided sequences and time parameter set $[0, \infty)$. On the other hand, as Doob (1953) points out on pp. 456–458, the law of a stationary process on $[0, 1, \cdots)$ or on $[0, \infty)$ can be uniquely extended, so as to preserve stationarity, to $(\cdots, -1, 0, 1, \cdots)$ or to $(-\infty, \infty)$. The properties of interest for the present study, such as stationarity, metric transitivity, transition exchangeability, and so on are preserved by this extension. Similar remarks apply to the stationary increments case.

**1. Notation.** Some of the notational devices used throughout the paper will be listed here. The letters $j$, $k$, $m$, $n$, $\nu$, refer to integers; $i$ to an integer or an element of a discrete space $I$ (in Theorem 2); $t$ to a real number; $r$ to a binary rational. When written in bold-face, these letters refer to vectors of the same type of quantity: thus $\mathbf{r} = (r_1 \cdots r_n)$ is a vector of binary rationals. The symbols $\Omega$, $\Omega^*$, $W$, $W^*$ denote sets, with elements $\omega$, $\omega^*$, $w$, $w^*$. Upper case script letters, $\mathfrak{F}$, $\mathfrak{I}$, $\mathfrak{G}$ refer to $\sigma$-fields. If $\mathfrak{F}$ is a $\sigma$-field of subsets of $\Omega$, and $A \subset \Omega$, then $A\mathfrak{F}$ denotes the $\sigma$-field of subsets of $A$ of the form $AB$, $B \in \mathfrak{F}$. The symbols $P$ and $Q$ are used for probabilities; bold-face $\mathbf{P}$ and $\mathbf{Q}$ for collections of probabilities. The letters $X$, $Y$, $f$, $g$, $h$ mean functions; $E_P$ is used for expectation under the probability $P$; $L$ and $U$ with subscripts are functionals. If $f$ and $g$ are real functions on the same space, $fg$ means their pointwise product. If $f : B \to C$ while $g : A \to B$, then $f \circ g : A \to C$ by sending $a$ to $f[g(a)]$.

In the formal exposition, *stochastic process* means an indexed collection of measurable functions from one space endowed with a $\sigma$-field to a second space similarly endowed. Only in Section 0, and the first paragraphs of Section 3, does *stochastic process* also connote that a specified probability has been placed on the first $\sigma$-field.

**2. The Kriloff-Bogoliouboff Theory.** As in [4], the decomposition of the sample space into ergodic sets is the fundamental technical device. In the continuous time situation, however, the classical theory does not apply, for two reasons: the relevant sample space is not second countable, or even first countable; and the stochastic processes do not satisfy the requisite measurability conditions. Decomposing the sample space is actually equivalent to finding (see p. 353 of Loève (1960)) a regular conditional probability, on one inseparable $\sigma$-field, given another inseparable $\sigma$-field, simultaneously for a large class of probabilities. The discussion in this section is therefore quite technical. Granted the results of this section, however, it is hoped that the proofs in Section 3 will be found simple.

If $(S, \tau)$ is a Hausdorff space, $C(S)$ denotes the space of bounded $\tau$-continuous real functions on $S$, in the uniform norm, $\|f\| = \sup_{x \in S} |f(x)|$. If $S$ is compact,

$C(S)$ is separable if and only if $\tau$ is second-countable. The $\sigma$-field $\mathfrak{F}(S)$ is the smallest $\sigma$-field of subsets of $S$ over which all the functions of $C(S)$ are measurable. This constitutes the Baire $\sigma$-field, as opposed to the Borel $\sigma$-field; see Chapter 10 of [5]. If $(S, \tau)$ is locally compact and second-countable, this distinction is vacuous.

Let $R^* = (-\infty, \infty)$, in the usual topology, and let $R$ be the binary rationals in $R^*$. The customary representation space for a stochastic process with values in $S$ and time running through $R^*$ is $\Omega^* = \prod_{t \varepsilon R^*} S$, in the product topology $\tau^*$ and product $\sigma$-field $\mathfrak{F}^*$. It is convenient to introduce the space $\Omega = \prod_{r \varepsilon R} S$, in the product topology $\pi$ and Baire $\sigma$-field $\mathfrak{F} = \mathfrak{F}(\Omega)$. This space approximates $\Omega^*$ in a sense to be explained later.

There is a group of homeomorphisms $T = \{T^r : r \varepsilon R\}$ on $\Omega$, defined by $(T^r\omega)(s) = \omega(r + s)$, $\omega \varepsilon \Omega$; $r, s \varepsilon R$. By $\mathbf{P}$ is meant the family of all probabilities $P$ on $\mathfrak{F}$ such that $PT^r = P$, all $r \varepsilon R$. There is a corresponding group $T^* = \{T^{*t} : t \varepsilon R^*\}$ on $\Omega^*$, which defines the family $\mathbf{P}^*$ of probabilities $P$ on $\mathfrak{F}^*$ for which $PT^{*t} = P$, all $t \varepsilon R^*$. If $f$ is a function on $\Omega$, $r \varepsilon R$, define $(T^r f)(\omega) = f(T^r\omega)$. Notice that this maps $C(\Omega)$ homeomorphically onto $C(\Omega)$. There is a corresponding statement for $T^{*t}$ and $C(\Omega^*)$.

There is a natural stochastic process, the coordinate process $\{\xi_t^* : t \varepsilon R^*\}$, on $(\Omega^*, \mathfrak{F}^*)$. It is defined by the relation $\xi_t^*(\omega^*) = \omega^*(t)$. The coordinate process $\{\xi_r : r \varepsilon R\}$ on $(\Omega, \mathfrak{F})$ is defined similarly. The product $\sigma$-field $\mathfrak{F}^*$ is the smallest $\sigma$-field of subsets of $\Omega^*$ containing the $\sigma$-fields $\xi_t^{*-1}\mathfrak{F}(S)$.

Since $\Omega^* = \Omega \times \Omega_u$, $\Omega_u = \prod_{t \varepsilon R^* - R} S$, there is a continuous projection map $\alpha$ from $\Omega^*$ onto $\Omega : \alpha(\omega^*)$ is $\omega^*$ restricted to $R$. If $P^*$ is a probability on $\mathfrak{F}^*$, $\alpha P^*$ is the probability on $\mathfrak{F}$ defined by $(\alpha P^*)(B) = P^*(\alpha^{-1}B)$.

For future reference, let $J_n = \{\mathbf{r} \mid \mathbf{r} \varepsilon R^n \text{ and } r_1 < \cdots < r_n\}$ and $J_n^* = \{\mathbf{t} \mid \mathbf{t} \varepsilon R^{*n}, t_1 < \cdots < t_n\}$. For $\mathbf{r} \varepsilon J_n$ let $\pi(\mathbf{r})$ project $\Omega$ onto $S^n$ according to the relation $\pi(\mathbf{r})(\omega) = [\omega(r_1), \cdots, \omega(r_n)]$. The projection $\pi^*(\mathbf{t})$ of $\Omega^*$ onto $S^n$ is defined in a similaı way for $\mathbf{t} \varepsilon J_n^*$.

From now to the end of Theorem 1, unless noted otherwise, *suppose that* $(S, \rho)$ *is compact metric*. Then $(\Omega^*, \tau^*)$ is compact Hausdorff, but not first countable. The $\sigma$-fields $\mathfrak{F}^*$ and $\mathfrak{F}(\Omega^*)$ coincide. Indeed if

$$\mathbf{A} = \{f \circ \pi^*(\mathbf{t}) \mid f \varepsilon C(S^n), \mathbf{t} \varepsilon J_n^*, n = 1, 2 \cdots\}$$

then $\mathbf{A}$ is dense in $C(\Omega^*)$ by the Stone-Weierstrass theorem (Loomis (1953), Section 4). The space $(\Omega, \pi)$ is compact metrizable; its Baire $\sigma$-field, Borel $\sigma$-field, and product $\sigma$-field all coincide.

It is convenient to begin the analysis with $\Omega$. Lemmas 1–6 provide a decomposition of $\Omega$ under the group $T$, completely analogous to the Kriloff-Bogoliouboff decomposition. To rephrase this, let $\mathfrak{I}_m = \{A \mid A \varepsilon \mathfrak{F} \text{ and } T^{2^{-m}}A = A\}$, $m = 0, 1, \cdots$ and $\mathfrak{I} = \{A \mid A \varepsilon \mathfrak{F} \text{ and } T^r A = A, \text{ all } r \varepsilon R\}$ so that, as $m \uparrow \infty$ $\mathfrak{I}_m \downarrow \mathfrak{I}$. Then Lemmas 1–6 construct one regular conditional probability on $\mathfrak{F}$ given $\mathfrak{I}$, simultaneously for all $P \varepsilon \mathbf{P}$.

Let $\Omega_0 = \{\omega \mid \lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1} f(T^j\omega) = L_0(\omega, f)$ exists simultaneously for all $f \varepsilon C(\Omega)\}$. According to Section 2 of [11], $\Omega_0 \varepsilon \mathfrak{F}$ and $P(\Omega_0) = 1$, all $P \varepsilon \mathbf{P}$.

Moreover, $L_0(\cdot, f)$ is clearly $\Omega_0\mathfrak{F}$-measurable, and $L_0(T^1\omega, f) = L_0(\omega, T^1 f) = L_0(\omega, f)$.

LEMMA 1. *If* $\omega \in \Omega_0$, $f \in C(\Omega)$ and $r \in R$, *then* $T^r\omega \in \Omega_0$, *and in fact* $L_0(T^r\omega, f) = L_0(\omega, T^r f)$.

PROOF. Since $T^r f \in C(\Omega)$,

$$\lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1} f(T^jT^r\omega) = \lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1} (T^r f)(T^j\omega) = L_0(\omega, T^r f),$$

as required.

For $\omega \in \Omega_0$, $f \in C(\Omega)$, $m = 0, 1, \cdots$ define

$$L_m(\omega, f) = 2^{-m}\sum_{j=0}^{2^m-1} L_0(\omega, T^{j2^{-m}} f). \tag{1}$$

It is easy to verify that

$$L_m(\omega, f) = \lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1} f(T^{j2^{-m}}\omega). \tag{2}$$

DEFINITION 1. A probability $P$ on $\mathfrak{F}$ will be called $(m)$-ergodic provided $PT^{2^{-m}} = P$ and $A \in \mathfrak{I}_m$ implies $P(A) = 0$ or 1. An equivalent condition is that for a.a. $\omega$ $[P]$, for all $f \in C(\Omega)$, $L_m(\omega, f) = \int_\Omega f\, dP$. The probability $P \in \mathbf{P}$ will be called $(T)$-ergodic provided $A \in \mathfrak{I}$ implies $P(A) = 0$ or 1.

Since for $\omega \in \Omega_0$, $L_m(\omega, \cdot)$ is a non-negative linear functional on $C(\Omega)$ such that $L_m(\omega, 1) = 1$, by the Riesz representation theorem (Loomis (1953), Section 16) it determines a probability $P_{m,\omega}$ on $\mathfrak{F}$. By a recurrent argument (which from now on will be called "The usual extension argument"), if $f$ is any bounded, $\mathfrak{F}$-measurable function, $\omega \in \Omega_0$,

$$\int_\Omega f\, dP_{m,\omega} = 2^{-m}\sum_{j=0}^{2^m-1}\int_\Omega T^{j2^{-m}} f\, dP_{0,\omega}. \tag{3}$$

Indeed, let $F_+$ be the set of non-negative $\mathfrak{F}$-measurable functions, $C_+ = F_+ \cap C(\Omega)$, and $k$ a natural number. Define $F_k$ to be the set of all $g \in F_+$ such that (3) holds with $f = \max [g, k]$. Then (Loomis (1953), 12$H$) $F_k$ is monotone, and contains $C_+$. Since $F_+$ is the smallest monotone family containing $C_+$, in fact $F_k = F_+$. Allowing $k \uparrow \infty$, by monotone convergence (3) holds for $f \in F_+$, and by linearity for any bounded $\mathfrak{F}$-measurable function.

Define $\Omega_1 = \{\omega \mid \omega \in \Omega_0 \text{ and } P_{0,\omega} \text{ is } (0)\text{-ergodic}\}$. By [11], $\Omega_1 \in \mathfrak{F}$ and $P(\Omega_1) = 1$, $P \in \mathbf{P}$. Moreover,

LEMMA 2. *If* $\omega \in \Omega_1$, $r \in R$, *then* $T^r\omega \in \Omega_1$ *and* $P_{m,\omega}$ *is* $(m)$-*ergodic.*

REMARK. For some $\omega$, the probabilities $P_{m,\omega}$ are not invariant under $T^{2^{-n}}$, $n \geqq m + 1$.

PROOF. The necessary and sufficient condition for $\omega$ to be in $\Omega_1$, given in (2.4) of [11], is

$$\lim_{k\to\infty}\lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1}\left[k^{-1}\sum_{j=0}^{k-1} f(T^{j+i}\omega) - L_0(\omega, f)\right]^2 = 0,$$

for all $f \varepsilon C(\Omega)$. But $f(T^{j+i}T^r\omega) = (T^r f)(T^{j+i}\omega)$ and $L_0(T^r\omega, f) = L_0(\omega, T^r f)$, proving that $T^r\Omega_1 = \Omega_1$.

Next, if $A \varepsilon \mathcal{I}_m$, apply (3) with $f$ the indicator of $A$, to see that $P_{m,\omega}(A) = P_{0,\omega}(A)$. But $\mathcal{I}_m \subset \mathcal{I}_0$, so that $\omega \varepsilon \Omega_1$ implies $P_{0,\omega}(A) = 0$ or 1, completing the proof.

Let $\Omega_2 = \{\omega \mid \omega \varepsilon \Omega_1$, and $\lim_{m\to\infty} L_m(\omega, f) = L(\omega, f)$ exists simultaneously for all $f \varepsilon C(\Omega)\}$. Then

LEMMA 3. *The set* $\Omega_2 \varepsilon \mathfrak{F}$, *and* $P(\Omega_2) = 1$, *all* $P \varepsilon \mathbf{P}$.

PROOF. Let $D$ be a countable dense subset of $C(\Omega)$. Let $\Omega_f = \{\omega \mid \omega \varepsilon \Omega_1$, and $\lim_{m\to\infty} L_m(\omega, f)$ exists$\}$. Then $\Omega_2 = \bigcap_{f\varepsilon D} \Omega_f$, so that $\Omega_2 \varepsilon \mathfrak{F}$ (and $\omega \to L(\omega, f)$ is $\Omega_2\mathfrak{F}$-measurable). It remains to prove that $P(\Omega_f) = 1$ for any fixed $P \varepsilon \mathbf{P}$. But it is well-known (p. 465, Doob (1953)) that $L_m(\cdot, f) = E_p(f \mid \mathcal{I}_m)$ a.s. $[P]$. Here $E_P(\cdot \mid \mathcal{I}_m)$ is the usual conditional expectation (see Section I.7 of [3] or 24.2 of [8]) computed for the probability $P$. Since the $\mathcal{I}_m \downarrow \mathcal{I}$, by a martingale theorem ([3], Theorem 4.2 on p. 328),

$$\lim_{m\to\infty} L_m(\cdot, f) = L(\cdot, f) = E_P(f \mid \mathcal{I}) \quad \text{a.s. } [P]. \tag{4}$$

This completes the proof. The lemma, and in particular Equation (4), will be used constantly in Section 3.

LEMMA 4. *If* $r \varepsilon R$, $\omega \varepsilon \Omega_2$, $f \varepsilon C(\Omega)$, *then* $T^r\omega \varepsilon \Omega_2$, *and in fact* $L(T^r\omega, f) = L(\omega, T^r f) = L(\omega, f)$.

PROOF. If $r = 2^{-b}a$, $b$ is a non-negative integer, $a$ is an integer, and $m \geqq b$, $L_m(T^r\omega, f) = L_m(\omega, T^r f) = L_m(\omega, f)$. Let $m \uparrow \infty$ to give the result.

Since $P(\Omega_2) = 1$ for $P \varepsilon \mathbf{P}$, it follows from e.g., the dominated convergence theorem that for such $P$ and $f \varepsilon C(\Omega)$,

$$\int_\Omega f\,dP = \int_{\Omega_2} L_m(\omega, f)\,dP \tag{5}$$

and by a second application of this theorem to the right side of (5)

$$\int_\Omega f\,dP = \int_{\Omega_2} L(\omega, f)\,dP. \tag{6}$$

But for $\omega \varepsilon \Omega_2$, $L(\omega, \cdot)$ is a non-negative linear functional on $C(\Omega)$ with $L(\omega, 1) = 1$. Thus $L(\omega, \cdot)$ determines a probability $P_\omega$ on $\mathfrak{F}$, such that: (i) for $f \varepsilon C(\Omega)$, $\int f\,dP_\omega = L(\omega, f)$; (ii) $P_\omega \varepsilon \mathbf{P}$; (iii) if $f$ is bounded and $\mathfrak{F}$-measurable, $\omega \to \int f\,dP_\omega$ is $\Omega_2\mathcal{I}$-measurable, and if $P \varepsilon \mathbf{P}$,

$$\int_\Omega f\,dP = \int_{\Omega_2} \int_\Omega f\,dP_\omega\,dP. \tag{7}$$

Indeed, (i) follows from the Riesz theorem, (ii) from Lemma 4, and (iii) from (6) and Lemma 4 by the usual extension argument. Equation (7) is of considerable importance in Section 3. It follows trivially from (iii) that $(\omega, A) \to P_\omega(A)$ is a regular conditional probability, in the usual sense (see p. 353 of [8]), on $\mathfrak{F}$, given $\mathcal{I}$, simultaneously for all $P \varepsilon \mathbf{P}$. Thus $\mathcal{I}$ is sufficient for $\mathbf{P}$ (in the sense of Lehmann (1959), p. 18).

The next lemma is perhaps of independent interest; it will not be used in Section 3. Let $\Omega_3 = \{\omega \mid \omega \varepsilon \Omega_2$ and $P_\omega$ is $(T)$-ergodic in the sense of Definition 1$\}$. Then by Lemma 4, $T^r\Omega_3 = \Omega_3$, $r \varepsilon R$ and

LEMMA 5. *The set* $\Omega_3 \varepsilon \mathfrak{F}$, *and* $P(\Omega_3) = 1$, *all* $P \varepsilon \mathbf{P}$.

PROOF. As usual, a probability $P \varepsilon \mathbf{P}$ is $(T)$-ergodic if and only if for a.a. $\omega[P]$, for all $f \varepsilon C(\Omega)$

$$L(\omega, f) = \int_\Omega f\, dP.$$

Following (2.4) of [11], it is therefore enough to show that for any one $f \varepsilon C(\Omega)$, the set $\Omega_f = \{\omega \mid \omega \varepsilon \Omega_2$, and for a.a. $v \varepsilon \Omega_2$, $[P_\omega]$, $L(v, f) = L(\omega, f)\}$ is in $\mathfrak{F}$ and has $P$-measure 1, $P \varepsilon \mathbf{P}$. This is equivalent to

$$\int_{\Omega_2} [L(v,f) - L(\omega,f)]^2 P_\omega\,(dv) = 0. \tag{8}$$

But by successive applications of dominated convergence, together with Lemma 3, and Equation (2), the left side of (8) is given by

$$\begin{aligned}
&\lim_{M\to\infty} \int_{\Omega_2} [L_M(v,f) - L(\omega,f)]^2 P_\omega\,(dv) \\
&= \lim_{M\to\infty} \lim_{N\to\infty} \int_{\Omega_2} \left[N^{-1} \sum_{j=0}^{N-1} f(T^{j2^{-M}}v) - L(\omega,f)\right]^2 P_\omega\,(dv) \\
&= \lim_{M\to\infty} \lim_{N\to\infty} \lim_{m\to\infty} L_m \left\{\omega, \left[N^{-1} \sum_{j=0}^{N-1} f(T^{j2^{-m}}v) - L(\omega,f)\right]^2\right\} \\
&= \lim_{M\to\infty} \lim_{N\to\infty} \lim_{m\to\infty} \lim_{n\to\infty} h(M, N, m, n, \omega)
\end{aligned} \tag{9}$$

where

$$h(M, N, m, n, \omega) = n^{-1} \sum_{k=0}^{n-1} \left[N^{-1} \sum_{j=0}^{N-1} f(T^{j2^{-M}+k2^{-m}}\omega) - L(\omega,f)\right]^2.$$

The last iterated limit exists everywhere on $\Omega_2$, and vanishes on a Borel subset of $\Omega_2$. Hence $\Omega_f \varepsilon \mathfrak{F}$. To show that $P(\Omega_f) = 1$ for a $P \varepsilon \mathbf{P}$, it is enough, since $h \geqq 0$, to prove that the $P$-integral of the last iterated limit is 0. But denoting $\int_\Omega X\, dP$ by $E_P X$, the pointwise ergodic theorem and Lemma 4 imply

$$E_P[\lim_{n\to\infty} h(M, N, m, n, \omega)] = E_P \left\{\left[N^{-1} \sum_{j=0}^{N-1} f(T^{j2^{-M}}\omega) - L(\omega,f)\right]^2\right\}$$

so that

$$\lim_{m\to\infty} E_P[\lim_{n\to\infty} h(M, N, m, n, \omega)] = E_P \left\{\left[N^{-1} \sum_{j=0}^{N-1} f(T^{j2^{-M}}\omega) - L(\omega,f)\right]^2\right\}$$

and by (2) and dominated convergence,

$$\lim_{N\to\infty} \lim_{m\to\infty} E_P[\lim_{n\to\infty} h(M, N, m, n, \omega)] = E_P\{[L_M(\omega, f) - L(\omega, f)]^2\}$$

so that by Lemma 3 and dominated convergence,

$$\lim_{M\to\infty}\lim_{N\to\infty}\lim_{m\to\infty}E_P[\lim_{n\to\infty}h(M,N,m,\omega)]=0$$

which proves (8) via three more applications of dominated convergence.

The next problem is to connect the spaces $\Omega$ and $\Omega^*$ (cf. the beginning of this section.) To do this it seems necessary to introduce a continuity condition. Lemmas 6–8 give the discussion. For Definition 2, suppose only that $(S, r)$ is metric.

DEFINITION 2. A probability $P$ on $\mathfrak{F}$ has a fixed point of discontinuity at $r \varepsilon R$ if and only if there is a sequence $s_n \to r$, $s_n \varepsilon R$, such that $P[\omega \mid \lim_{n\to\infty} \omega(s_n) = \omega(r)] < 1$. Write $\mathbf{P}_0 = \{P \mid P \varepsilon \mathbf{P}$, and $P$ has no fixed point of discontinuity$\}$. Equally $P^*$ on $\mathfrak{F}^*$ has a fixed point of discontinuity at $t \varepsilon R^*$ if and only if there is a sequence $s_n \to t$, $s_n \varepsilon R^*$, such that $P^*[\omega^* \mid \lim_{n\to\infty} \omega^*(s_n) = \omega^*(t)] < 1$. Write $\mathbf{P}_0^* = \{P^* \mid P^* \varepsilon \mathbf{P}^*$, and $P^*$ has no fixed point of discontinuity.$\}$

If $P \varepsilon \mathbf{P}$, or $P^* \varepsilon \mathbf{P}$, either no point is a fixed point of discontinuity, or all are. Further, if $P^* \varepsilon \mathbf{P}_0^*$, then $\alpha P^* \varepsilon \mathbf{P}_0$.

Suppose again that $(S, \rho)$ is compact metric. Define, for $r \varepsilon R$, $\omega \varepsilon \Omega$, $n = 1, 2, \cdots$

$$f(n, \omega) = \sup \{\rho[\omega(r), \omega(s)] \mid s \varepsilon R, |r - s| < n^{-1}\}. \tag{10}$$

Then $f(n, \cdot)$ is bounded, $\mathfrak{F}$-measurable and

LEMMA 6. *The probability $P$ on $\mathfrak{F}$ has a fixed point of discontinuity at $r$ if and only if* $P[\omega \mid \lim_{n\to\infty} f(n, \omega) = 0] < 1$.

PROOF. The "only if" part is clear. The other implication is a very special case of some considerations in Chapter II.2 of [3], but it is easier to prove directly than by reference. As $n$ increases, $f(n, \omega)$ decreases. Suppose that $M \varepsilon \mathfrak{F}$, $P(M) = \delta > 0$, and for $\omega \varepsilon M$, $f(n, \omega) \geqq \delta' > 0$, all $n$. Then there exist numbers $s_j(n) \varepsilon R$, $1 \leqq j \leqq j_n$, such that $r - n^{-1} < s_1(n) < \cdots < s_{j_n}(n) < r + n^{-1}$ and $P\{\omega \mid \max_{1\leqq j\leqq j_n} \rho\{\omega(s_j(n)), \omega(r)\} \geqq \delta'/2\} \geqq \delta/2$. Define the sequence $\{s_j\}$ as $(s_1, s_2, \cdots, s_{j_1}, s_{j_1+1}, s_{j_1+2}, \cdots) = (s_1(1), s_2(1), \cdots, s_{j_1}(1), s_1(2), s_2(2), \cdots)$. By Egoroff's theorem $P\{\omega \mid \lim_{n\to\infty}\omega(s_n) = \omega(r)\} < 1$, completing the proof.

There is no analogue for Lemma 6 within the $\sigma$-field $\mathfrak{F}^*$.

If $A$ and $B$ are sets, $A \triangle B = (A - B) \cup (B - A)$ is their symmetric difference. If $P^*$ is a probability on $\mathfrak{F}^*$ then

DEFINITION 3. The set $A \varepsilon \mathfrak{F}^*$ is $P^*$-almost invariant if and only if $P^*(A \triangle T^{*t}A) = 0$, all $t \varepsilon R^*$. The probability $P^* \varepsilon \mathbf{P}^*$ is ergodic if and only if $P^*(A)$ is 0 or 1 for each $P^*$-almost invariant $A \varepsilon \mathfrak{F}^*$.

If $S$ has more than one element, then $A \varepsilon \mathfrak{F}^*$ and $T^{*t}A = A$ for all $t \varepsilon R^*$ implies $A$ is empty or $\Omega^*$, making this definition necessary. A simple example of a non-ergodic probability may be obtained by letting $S$ consist of two points, 0 and 1. Define $\zeta^*$ as the function on $R^*$ which is identically 0, and let $\eta^*$ be identically 1. Write $\zeta = \alpha\zeta^*$, $\eta = \alpha\eta^*$. If $1_A$ denotes the indicator function of the set $A$, then the probability $P^*$ defined on $(\Omega^*, \mathfrak{F}^*)$ by $P^*(A) = \frac{1}{2}1_A(\zeta^*) + \frac{1}{2}1_A(\eta^*)$ is in $\mathbf{P}_0^*$ but is not ergodic. Indeed, $\{\omega^* \mid \omega^* \varepsilon \Omega^*$ and $\omega^*(0) = 0\}$ is in

$\mathfrak{F}^*$, is $P^*$-almost invariant, and has $P^*$-probability $\frac{1}{2}$. In the terminology of Theorem 1, $\alpha P^*$ assigns mass $\frac{1}{2}$ each to $\zeta$ and $\eta$; and $P^*_\zeta(A) = 1_A(\zeta^*)$, $P^*_\eta(A) = 1_A(\eta^*)$, $A \varepsilon \mathfrak{F}^*$.

LEMMA 7. *If $P \varepsilon \mathbf{P}_0$, there is a unique $P^*$ on $\mathfrak{F}^*$ such that $P^* \varepsilon \mathbf{P}^*_0$ and $P = \alpha P^*$. Moreover, if $P$ is ergodic, so is $P^*$.*

PROOF. Let $d_i$ operate on vectors of length $n \geqq i$ by deleting the $i$th coordinate. Thus, $d_2(a, b, c) = (a, c)$. If $n \geqq i$ and $n \geqq 2$, then $d_i$ projects $S^n$ onto $S^{n-1}$. If $f \varepsilon C(S^{n-1})$, then $d_i f \varepsilon C(S^n)$, where $(d_i f)(\mathbf{x}) = f(d_i \mathbf{x})$. For each $n$, each $\mathbf{t} \varepsilon J^*_n$, let $L^*(\cdot, \mathbf{t})$ be a non-negative linear functional on $C(S^n)$ which is 1 at the function 1, and which satisfies the *consistency* condition: if $n \geqq 2$, $i \leqq n$, $\mathbf{t} \varepsilon J^*_n$ and $f \varepsilon C(S^{n-1})$, then $L^*(d_i f, \mathbf{t}) = L^*(f, d_i \mathbf{t})$. The collection $\{L^*(\cdot, \mathbf{t}) \mid \mathbf{t} \varepsilon \bigcup_n J^*_n\}$ will be called a *law*. Since the algebra $\mathbf{A}$ is dense in $C(\Omega^*)$, by the Riesz theorem there is a unique probability $P^*$ on $\mathfrak{F}^*$ such that for each $n$, each $\mathbf{t} \varepsilon J^*_n$, each $f \varepsilon C(S^n)$,

$$L^*(f, \mathbf{t}) = \int_{\Omega^*} f \circ \pi^*(\mathbf{t})\, dP^*. \tag{L}$$

This is a version of the Kolmogoroff consistency theorem ([8], p. 93).

Starting with a $P \varepsilon \mathbf{P}_0$, define for each $n$ and each $\mathbf{r} \varepsilon J_n$ a functional $L(\cdot, \mathbf{r})$ on $C(S^n)$ by $L(f, \mathbf{r}) = \int_\Omega f \circ \pi(\mathbf{r})\, dP$. Now define $L^*(\cdot, \mathbf{t})$ on $C(S^n)$ for $\mathbf{t} \varepsilon J^*_n$ thus. If $f \varepsilon C(S^n)$, then $L(f, \cdot)$ is uniformly continuous on $J_n$. Indeed, for all $\delta_1 > 0$, $\delta_2 > 0$ there is a $\delta > 0$ such that $r \varepsilon R$ and $|r| < \delta$ implies $P\{\omega \mid \rho[\omega(r), \omega(0)] > \delta_1\} < \delta_2$. By stationarity, if $\mathbf{r}$ and $\mathbf{s}$ are points of $J_n$ with $|r_i - s_i| < \delta$, $1 \leqq i \leqq n$ then $P\{\omega \mid \max_{1 \leqq i \leqq n} \rho[\omega(r_i), \omega(s_i)] > \delta_1\} \leqq \sum_{i=1}^n P\{\omega \mid \rho[\omega(r_i - s_i), \omega(0)] > \delta_1\} < n\delta_2$. Write $(\theta, f) = \max \{|f(\mathbf{x}) - f(\mathbf{y})| : \mathbf{x} \varepsilon S^n, \mathbf{y} \varepsilon S^n, \rho(x_i, y_i) \leqq \theta, 1 \leqq i \leqq n\}$, for $\theta > 0$. Then $|L(f, \mathbf{r}) - L(f, \mathbf{s})| < n\delta_2\|f\| + (\delta_1, f)$, establishing the uniform continuity. Consequently, $L(f, \cdot)$ has a unique continuous extension $L^*(f, \cdot)$ to $J^*_n$. That is, $L^*(f, \mathbf{t}) = \lim_{\mathbf{r} \to \mathbf{t}, \mathbf{r} \varepsilon J_n} L^*(f, \mathbf{r})$. Clearly, $L^*(\cdot, \mathbf{t})$ is a non-negative linear functional on $C(S^n)$ which is 1 at the function 1, and $\{L^*(\cdot, \mathbf{t}) \mid \mathbf{t} \varepsilon \bigcup_n J^*_n\}$ is consistent. Let $P^*$ be the unique probability on $\mathfrak{F}^*$ in relation (L) to $L^*$. It is clear that $\alpha P^* = P$. For each $n$, each $\mathbf{r} \varepsilon J_n$, each $r \varepsilon R$ and $f \varepsilon C(S^n)$ the equation $L(f, \mathbf{r} + r) = L(f, \mathbf{r})$ holds; this implies that for each $n$, each $\mathbf{t} \varepsilon J^*_n$, each $t \varepsilon R^*$ and $f \varepsilon C(S^n)$, the equation $L(f, \mathbf{t} + t) = L(f, \mathbf{t})$ holds, and $P^* \varepsilon \mathbf{P}$.

The next step is to prove that $P^* \varepsilon \mathbf{P}^*_0$, i.e., that 0 is not a fixed point of discontinuity. Now $\rho$ is bounded continuous on $S \times S$, so that $P^* \varepsilon \mathbf{P}^*$ implies

$$\int_{\Omega^*} \rho[\omega^*(t + h), \omega^*(t)]\, dP^* = \int_{\Omega^*} \rho[\omega^*(h), \omega^*(0)]\, dP^*$$

$$= \lim_{r \to h, r \varepsilon R} \int_\Omega \rho[\omega(r), \omega(0)]\, dP$$

from the construction of $P^*$. This equality, and $P \varepsilon \mathbf{P}_0$, together imply

$$\lim_{h \to 0} P^*\{\omega^* \mid \rho[\omega^*(t + h), \omega^*(t)] \geqq \delta > 0\} = 0 \tag{11}$$

uniformly in $t$. Now suppose $s_n \to 0$, and by way of contradiction, that for some $\delta > 0$, $P^*\{\omega^* \mid \limsup_{n\to\infty} \rho[\omega^*(s_n), \omega^*(0)] \geqq \delta\} > 0$. For each $n$, find $r_n \varepsilon R$ such that $P^*\{\omega^* \mid \rho[\omega^*(r_n), \omega^*(s_n)] \geqq \delta/2\} \leqq n^{-2}$, and $|s_n - r_n| < 1/n$. This is always possible by (11). By the Borel-Cantelli Lemmas, a.s. $[P^*]$, $\rho[\omega^*(r_n), \omega^*(s_n)] < \delta/2$ eventually, so that because $P = \alpha P^*$,

$$0 < P^*\{\omega^* \mid \limsup_{n\to\infty} \rho[\omega^*(r_n), \omega^*(0)] \geqq \delta/2\}$$
$$= P\{\omega \mid \limsup_{n\to\infty} \rho[\omega(r_n), \omega(0)] \geqq \delta/2\},$$

the promised contradiction.

The uniqueness is clear.

If $A \varepsilon \mathfrak{F}^*$, there is a set $A_0 \varepsilon \alpha^{-1}\mathfrak{F}$, with $P^*(A \triangle A_0) = 0$. If $P$ is $(T)$-ergodic and $A \varepsilon \mathfrak{F}^*$ is $P^*$-almost invariant, then $P^*(A) = P^*(A_0) = P(\alpha A_0)$ since $P = \alpha P^*$, and to prove that $P^*$ is ergodic it suffices to prove $P(\alpha A_0) = 0$ or 1. But if $r \varepsilon R$,

$$P(\alpha A_0 \triangle T^r \alpha A_0) = P^*(A_0 \triangle T^{*r} A_0) = 0.$$

Let $A_1 = \bigcup_{r\varepsilon R} T^r \alpha A_0$. Then $P(A_1 \triangle \alpha A_0) = 0$, and $A_1 \varepsilon \mathcal{I}$, i.e., $A_1 \varepsilon \mathfrak{F}$ and $T^r A_1 = A_1$, $r \varepsilon R$, so that $P(A_1) = 0$ or 1, which completes the proof.

REMARK. Given any $P \varepsilon \mathbf{P}$ which is not carried by a single point of $\Omega$, there are precisely $c$ elements $P^* \varepsilon \mathbf{P}^*$ (but not $\varepsilon \mathbf{P}_0^*$) such that $\alpha P^* = P$. This is true even when, e.g., under $P$ the process $(\xi_r, r \varepsilon R)$ is a two-state Markoff chain with standard transition matrix ([2], Section II.2); $c$ of the probabilities constructed before will not be Markoff.

LEMMA 8. *Let* $\Omega_4 = \{\omega \mid \omega \varepsilon \Omega_3$ *and* $P_\omega \varepsilon \mathbf{P}_0\}$. *Then* $\Omega_4 \varepsilon \mathfrak{F}$, *and* $P(\Omega_4) = 1$, $P \varepsilon \mathbf{P}_0$.

PROOF. Define $f(n, \omega)$ by (10), with e.g., $r = 0$, and let $A = \{\omega \mid \lim_{n\to\infty} f(n, \omega) = 0\}$. Then $\Omega_4 = \{\omega \mid P_\omega(A) = 1\} \varepsilon \mathfrak{F}$, and since from (7), if $P \varepsilon \mathbf{P}_0$, then $1 = P(A) = \int_{\Omega_3} P_\omega(A)\, dP$, so that $P(\Omega_4) = 1$, as required.

Analogues for Lemma 7 may be obtained for weaker continuity assumptions than Definition 3. However, I have not been able to extend Lemma 8. Thus, if $\mathbf{P}_1 = \{P \mid P \varepsilon \mathbf{P}$ and $\lim_{r\to 0, r\varepsilon R} \int_\Omega \rho[\omega(r), \omega(0)]\, dP = 0\}$, and $\Omega_5 = \{\omega \mid \omega \varepsilon \Omega_3$ and $P_\omega \varepsilon \mathbf{P}_1\}$, then $\Omega_5 \varepsilon \mathfrak{F}$. I conjecture but cannot prove that $P(\Omega_5) = 1$, for all $P \varepsilon \mathbf{P}_1$.

Combining these results:

THEOREM 1. *There is a subset* $\Omega_4 \varepsilon \mathfrak{F}$, *and corresponding to each* $\omega \varepsilon \Omega_4$ *a probability* $P_\omega^*$ *on* $\mathfrak{F}^*$ *such that*:

(i) $\tau^r \Omega_4 = \Omega_4$ *and* $P_{\tau^r\omega}^* = P_\omega^*$, *for all* $r \varepsilon R$;

(ii) $P(\Omega_4) = 1$, $P \varepsilon \mathbf{P}_0$;

(iii) $P_\omega^* \varepsilon \mathbf{P}_0^*$, *and* $P_\omega^*$ *is ergodic*;

(iv) *if* $f$ *is bounded and* $\mathfrak{F}^*$*-measurable, then* $\omega \to \int_{\Omega^*} f\, dP_\omega^*$ *is* $\Omega_4\mathcal{I}$*-measurable*; *and if* $P^* \varepsilon \mathbf{P}_0^*$, *then* $\int_{\Omega^*} f\, dP^* = \int_{\Omega_4} \int_{\Omega^*} f\, dP_\omega^*\, d\alpha P^*$.

PROOF. The set $\Omega_4$ was constructed in Lemma 8. Corresponding to each $\omega \varepsilon \Omega_4$, $P_\omega \varepsilon \mathbf{P}_0$, and so has a unique extension $P_\omega^*$ on $\mathfrak{F}^*$, such that $P_\omega^* \varepsilon \mathbf{P}_0^*$, and $P_\omega =$

$\alpha P_\omega^*$, by Lemma 7. Part (i) then follows by Lemma 4, part (ii) by Lemma 8, (iii) by Lemmas 5 and 7. Part (iv) follows from an extension argument. Indeed, part (iv) holds for $f = g \cdot \pi^*(\mathbf{r})$, where $\mathbf{r} \varepsilon J_n$ and $g \varepsilon C(S^n)$. This follows easily from equation (7). Now let $\mathbf{r} \to \mathbf{t} \varepsilon J_n^*$. By the construction of $P_\omega^*$ from $P_\omega$ in Lemma 7, $\int g \circ \pi^*(\mathbf{r})\, dP_\omega^* \to \int g \circ \pi^*(\mathbf{t})\, dP_\omega^*$, and by dominated convergence part (iv) holds for $f \varepsilon \mathbf{A}$ and so for $f \varepsilon C(\Omega^*)$ by uniform passage to the limit, and the usual extension argument completes the proof. Notice that $\Omega_4 \subset \Omega$, while $P_\omega^*$ acts on subsets of $\Omega^*$.

REMARK. As a consequence of this theorem, the $\sigma$-field $\alpha^{-1}\mathfrak{G}$ is sufficient for $\mathbf{P}_0^*$.

The result of Theorem 1 may be clarified by an example. Let $(S, \rho)$ be the interval $[-1, 1]$ in the usual metric, and $\varphi^*$ be a continuous function of period 1 from the real line $R^*$ to $S$; a suitable $\varphi^*$ maps $t$ into $\sin 2\pi t$. The map $M^*$ of $[0, 1]$ into $\Omega^*$ defined by $M^*(\theta)(t) = \varphi^*(\theta + t)$, $\theta \varepsilon [0, 1]$, $t \varepsilon R^*$ is continuous, so measurable. The image $P^*$ of Lebesgue measure by $M^*$ is ergodic and has no fixed points of discontinuity. Let $M$ map $[0, 1]$ into $\Omega$ by the relation $M(\theta) = \alpha M^*(\theta)$. Then $M$ is continuous and has a compact, so measurable, range $\Omega_M$. The image $P$ of Lebesgue measure by $M$ is clearly $\alpha P^*$ and is concentrated on $\Omega_M$. For $\omega \varepsilon \Omega_M$, $f \varepsilon C(\Omega)$, it is easy to check that $L_0(\omega, f) = f(\omega)$. Therefore $P_{0,\omega}$ is 0-ergodic but invariant under $T^{2^{-m}}$ for no $m \geqq 1$;

$$L(\omega, f) = \int_0^1 f \circ M(\theta)\, d\theta,$$

thus $P_\omega = P$ and $P$ is $(T)$-ergodic. Clearly, $P$ is $m$-ergodic for no $m$.

This theorem gives a decomposition more than adequate for the purposes of Section 3, when the underlying process is (strictly) stationary. Some further analysis is needed for processes which have only stationary increments. The discussion is entirely parallel to that of the stationary case. Lemmas are numbered correspondingly, with the prefix $D$. Most of the proofs are omitted as obvious modifications of earlier arguments.

The relevant sample space is obtained as follows. Let $(G, \tau)$ be a topological group which is $T_1$, so Hausdorff (see Chapter VI of [9]). Suppose that $(G, \tau)$ is locally compact and second countable. In Section 3, the additive real line or the additive integers will be taken for $G$, but a concrete discussion is harder than an abstract one. If $G$ is compact, put $Ge = G$; if not, put $Ge$ for the one-point compactification of $G$. Thus $Ge$ is compact metrizable, but may have no algebraic structure. This leads to certain complications.

The group operations are written as addition and subtraction, and the identity is written 0; it is not assumed that $G$ is commutative. The product spaces $\Omega$ and $\Omega^*$ (see the beginning of Section 2) with $(G, \tau)$ for $(S, \tau)$ will be written $W$ and $W^*$. The product space $\Omega$ with $Ge$ for $S$ will be written $W_\infty$. Two $\sigma$-fields are needed on $W^*$; the product $\sigma$-field $\mathfrak{F}^*$ already defined, and the difference $\sigma$-field $\mathfrak{F}_d^* \subset \mathfrak{F}^*$, which is the smallest $\sigma$-field of subsets of $W^*$ containing the $\sigma$-fields $(\xi_t^* - \xi_s^*)^{-1}\mathfrak{F}(G)$, $s$, $t \varepsilon R^*$.

Notice that $W \subset W_\infty$ in an obvious way, and then $W \varepsilon \mathfrak{F}(W_\infty)$ and $\mathfrak{F} = W\mathfrak{F}(W_\infty)$ coincides with the product $\sigma$-field of $W$, the Baire $\sigma$-field of $W$, and the Borel $\sigma$-field of $W$. The difference $\sigma$-field $\mathfrak{F}_d \subset \mathfrak{F}$ is defined similarly to $\mathfrak{F}_d^*$. Let $K$ be the class of functions $f$ on $W$ for which there exist a natural number $n$, a vector $\mathbf{r} \varepsilon J_{n+1}$, a function $g \varepsilon C(Ge^n)$, such that

$$f(w) = g[w(r_2) - w(r_1), \cdots, w(r_{n+1}) - w(r_n)] = \langle g, \mathbf{r}\rangle(w).$$

Notice that $f \varepsilon K$ is bounded and continuous; the smallest $\sigma$-field over which all functions in $K$ are measurable is precisely $\mathfrak{F}_d$; and $K$ is separable in the uniform norm $\|f\| = \sup_{w\varepsilon W} |f(w)|$, because each $C(Ge^n)$ is.

Let $\mathbf{Q}$ be the set of all probabilities $P$ on $\mathfrak{F}$ such that for all $D \varepsilon \mathfrak{F}_d$, $r \varepsilon R$, $PT^rD = PD$. Define $\mathbf{Q}^*$ for $W^*$ in a similar way.

Suppose $f \varepsilon K$, so that $f = \langle g, \mathbf{r}\rangle$. The process $\langle g, \mathbf{r} + j\rangle$; $j = 0, 1, \cdots$ is strictly stationary under $P \varepsilon \mathbf{Q}$. Hence by the pointwise ergodic theorem ([3], Theorem 2.1, p. 465),

$$\lim_{N\to\infty} N^{-1}\sum_{j=0}^{N-1} \langle g, \mathbf{r} + j\rangle(w) = U_0(w, f)$$

exists on a set of $P$-measure 1 for all $P \varepsilon \mathbf{Q}$. Now define $W_0 = \{w \mid w \varepsilon W$ and $U_0(w, f)$ exists simultaneously for all $f \varepsilon K\}$. Then the reasoning of [11], together with the separability of $K$, implies.

LEMMA D.0. *The set* $W_0 \varepsilon \mathfrak{F}_d$, *and* $P(W_0) = 1$, *all* $P \varepsilon \mathbf{Q}$.

Notice that $T^rK = K$, for $r \varepsilon R$, since

$$\langle T^rg, \mathbf{r}\rangle(w) = \langle g, \mathbf{r} + r\rangle(w) = \langle g, \mathbf{r}\rangle(T^rw).$$

Hence

LEMMA D.1. *If* $f \varepsilon K$, $r \varepsilon R$, *then* $T^rW_0 = W_0$ *and* $U_0(T^rw, f) = U_0(w, T^rf)$. *Also* $U_0(w, T^1f) = U_0(w, f)$.

For $w \varepsilon W_0$, $f \varepsilon K$, $m = 0, 1, \cdots$ write

$$U_m(w, f) = 2^{-m} \sum_{j=0}^{2^m-1} U_0(w, T^{j2^{-m}}f). \tag{D.1}$$

Then

$$U_m(w, f) = \lim_{N\to\infty} N^{-1}\sum_{j=0}^{N-1} f(T^{j2^{-m}}w). \tag{D.2}$$

DEFINITION D.1. Let $\mathcal{G}_m = \{A \mid A \varepsilon \mathfrak{F}_d$ and $T^{2^{-m}}A = A\}$; $\mathcal{G} = \{A \mid A \varepsilon \mathfrak{F}_d$ and $T^rA = A$, all $r \varepsilon R\}$. A probability $P$ on $\mathfrak{F}$ is called $(m)$ $d$-ergodic provided $PT^{2^{-m}} = P$ on $\mathfrak{F}_d$, and $A \varepsilon \mathcal{G}_m$ implies $P(A) = 0$ or 1. The probability $P$ is called $d$-ergodic provided $P \varepsilon \mathbf{Q}$ and $A \varepsilon \mathcal{G}$ implies $P(A) = 0$ or 1.

Notice that $\mathcal{G}_m$ is strictly smaller than the sub-$\sigma$-field of $\mathfrak{F}$ consisting of the sets invariant under $T^{2^{-m}}$. Hence there exist $d$-ergodic but not ergodic probabilities. As before, $m \uparrow \infty$ implies $\mathcal{G}_m \downarrow \mathcal{G}$. For most $w$, the functional $U_m(w, \cdot)$ on $K$, defined by (D.2), induces a unique (under a convention stated below) probability $Q_{m,w}$ on $\mathfrak{F}$. The technique, being somewhat different from the one used to construct $P_{m,\omega}$, will be explained in detail. For a fixed $\mathbf{r} \varepsilon J_{n+1}$, $U_m(w, \langle \cdot, \mathbf{r} \rangle)$ is a non-negative linear functional on $C(Ge^n)$, and determines a probability $P_m(w, \mathbf{r}, \cdot)$ on $\mathfrak{F}(Ge^n)$. If $A \varepsilon \mathfrak{F}(Ge^n)$,

$$A_0 = \{w \mid w \varepsilon W, [w(r_2) - w(r_1), \cdots, w(r_{n+1}) - w(r_n)] \varepsilon A\},$$

then $A_0 \varepsilon \mathfrak{F}_d$, and $w \to P_m(w, \mathbf{r}, A)$ is $W_0\mathfrak{F}_d$-measurable. If $P \varepsilon \mathbf{Q}$, then as usual $P(A_0) = \int_{W_0} P_m(w, \mathbf{r}, A)\, dP$. Hence, if $W_{01} = \{w \mid w \varepsilon W_0$, and for all $m$, all $n$, all $\mathbf{r} \varepsilon J_{n+1}$, $P_m(w, \mathbf{r}, G^n) = 1\}$, then $W_{01} \varepsilon \mathfrak{F}_d$, $T^r W_{01} = W_{01}$, and $P(W_{01}) = 1$ for all $r \varepsilon R$ and all $P \varepsilon \mathbf{Q}$.

Now fix $w \varepsilon W_{01}$. There is a unique probability $Q_{m,w}$ on $\mathfrak{F}$ in $W$ such that

(*i*) $Q_{m,w}[w' \mid w'(0) = 0] = 1$

(ii) $\int_W f\, dQ_{m,w} = U_m(w, f)$ for all $f \varepsilon K$.

Moreover, it follows from (ii) that

(iii) $Q_{m,w} T^{2^{-m}} = Q_{m,w}$ on $\mathfrak{F}_d$.

To see this, consider the class $H$ of all $f \varepsilon C(W_\infty)$ of the form $f(w) = g[w(r_i), 1 \leqq i \leqq n; w(0); w(s_i), 1 \leqq i \leqq n]$, where $\mathbf{r} \varepsilon J_n$, $r_i < 0$ and $\mathbf{s} \varepsilon J_n$, $s_i > 0$, and $g \varepsilon C(Ge^{2n+1})$ vanishes off a compact subset of $G^{2n+1}$. Let $(Df)(w) = g[w(r_i) - w(0), 1 \leqq i \leqq n; 0; w(s_i) - w(0), 1 \leqq i \leqq n], w \varepsilon W; = 0, w \varepsilon W_\infty - W$. Since e.g., $w(s_i) - w(0) = [w(s_i) - w(s_{i-1})] + \cdots + [w(s_1) - w(0)]$, therefore $Df$ when restricted to $W$ is in $K$. If $f \varepsilon K \cap H$, $Df = f$. Define the functional $V_m(w, \cdot)$ on $H$ as

$$V_m(w, f) = U_m(w, Df).$$

Then $V_m(w, \cdot)$ is non-negative, linear, and assigns the value 1 to the function 1. Moreover, by the Stone-Weierstrass theorem the algebra $H$ is dense in $C(W_\infty)$, so by the Riesz theorem, $V_m(w, \cdot)$ determines a probability on $\mathfrak{F}(W_\infty)$, which when restricted to $\mathfrak{F} = W\mathfrak{F}(W_\infty)$ is $Q_{m,w}$. The fact $P_m(w, \mathbf{r}, G^n) \equiv 1$ ensures $Q_{m,w}(W) = 1$. Assertions (i), (ii), (iii) are easy to verify, and this finishes the demonstration. This is a variant of the discussion of p. 97 of [3].

Define $W_1 = \{w \mid w \varepsilon W_{01}$ and $Q_{0,w}$ is (0) $d$-ergodic$\}$. By the technique of [11], $W_1 \varepsilon \mathfrak{F}_d$, and $P(W_1) = 1$, all $P \varepsilon \mathbf{Q}$.

LEMMA D.2. *For $r \varepsilon R$, $T^r W_1 = W_1$, and $w \varepsilon W_1$ implies $Q_{m,w}$ is $(m)$ $d$-ergodic.*

Define $W_2 = \{w \mid w \varepsilon W_1$, and $\lim_{m \to \infty} U_m(w, f) = U(w, f)$ exists simultaneously for all $f \varepsilon K\}$.

LEMMA D.3. *The set $W_2 \varepsilon \mathfrak{F}_d$, and $P(W_2) = 1$, all $P \varepsilon \mathbf{Q}$.*

PROOF. Since $U_m(\cdot, f)$ is a version of $E_P(f \mid \mathcal{G}_m)$ for any $P \varepsilon \mathbf{Q}$, and $\mathcal{G}_m \downarrow \mathcal{G}$, convergence follows from the previous martingale argument.

LEMMA D.4. *If $r \varepsilon R$, $f \varepsilon K$, $w \varepsilon W_2$, then $T^r W_2 = W_2$, and $U(T^r w, f) = U(w, T^r f) = U(w, f)$.*

Following the argument for $U_m(w, f)$, it is possible to find an invariant $\mathfrak{F}_d$ set, $W_{21} \subset W_2$, such that $P(W_{21}) = 1$ for $P \varepsilon \mathbf{Q}$, and such that for $w \varepsilon W_{21}$ there is a unique probability $Q_w$ on $\mathfrak{F}$ satisfying:

(i) $Q_w[w' \mid w'(0) = 0] = 1$;

(ii) $\int_W f\, dQ_w = U(w, f)$, all $f \varepsilon K$;

(iii) $Q_w \varepsilon \mathbf{Q}$.

Moreover, if $f$ is bounded and $\mathfrak{F}_d$-measurable, $w \to \int_W f d\, Q_w$ is $W_{21}\mathfrak{F}_d$-measurable, and if $P \, \varepsilon \, \mathbf{Q}$.

$$\text{(D.7)} \qquad \int_W f\, dP = \int_{W_{21}} \int_W f\, dQ_w\, dP.$$

Let $W_3 = \{w \mid w \, \varepsilon \, W_{21}$ and $Q_w$ is $d$-ergodic$\}$. Then $T^r W_3 = W_3$, $r \, \varepsilon \, R$ and

LEMMA D.5. *The set* $W_3 \, \varepsilon \, \mathfrak{F}_d$, *and* $P(W_3) = 1$, *all* $P \, \varepsilon \, \mathbf{Q}$.

Definition 2 applies with $\Omega$ and $\Omega^*$ replaced by $W$ and $W^*$.

DEFINITION D.2. Write $\mathbf{Q}_0 = \{P \mid P \, \varepsilon \, \mathbf{Q}$, and $P$ has no fixed point of discontinuity$\}$. Write $\mathbf{Q}_0^* = \{P \mid P \, \varepsilon \, \mathbf{Q}^*$ and $P$ has no fixed point of discontinuity$\}$.

If $P \, \varepsilon \, \mathbf{Q}$ or $P^* \, \varepsilon \, \mathbf{Q}^*$, either no point is a fixed point of discontinuity or all are. If $P^* \, \varepsilon \, \mathbf{Q}_0^*$ then $\alpha P^* \, \varepsilon \, \mathbf{Q}_0$ (where $\alpha$ projects $W^*$ onto $W$). Recall that $Ge$ is compact metrizable; let $\rho$ be a metric on $Ge$ inducing its topology. Fix $r \, \varepsilon \, R$, and define

$$g(n, w) = \sup \{\rho[w(r) - w(s), 0] \mid w \, \varepsilon \, W, s \, \varepsilon \, R, |r - s| < 1/n\}.$$

Since $w \, \varepsilon \, W$, $t \, \varepsilon \, R$ imply $w(t) \, \varepsilon \, G$, $w(r) - w(s)$ is defined. Thus $g(n, \cdot)$ is bounded, $\mathfrak{F}_d$-measurable, and

LEMMA D.6. *The probability* $P$ *on* $\mathfrak{F}$ *in* $W$ *has a fixed point of discontinuity at* $r$ *if and only if* $P\{w \mid w \, \varepsilon \, W, \lim_{n\to\infty} g(n, w) = 0\} = 1$.

DEFINITION D.3. The probability $P^* \, \varepsilon \, \mathbf{Q}^*$ will be called $d$-ergodic provided $A \, \varepsilon \, \mathfrak{F}^*$ and $P^*(A \bigtriangleup T^{*t}A) = 0$ for all $t \, \varepsilon \, R^*$ imply $P^*(A) = 0$ or 1.

LEMMA D.7. *If* $P \, \varepsilon \, \mathbf{Q}_0$, $P[w \mid w(0) = 0] = 1$, *there is a unique* $P^* \, \varepsilon \, \mathbf{Q}_0^*$ *such that* $\alpha P^* = P$. *Moreover, if* $P$ *is d-ergodic, so is* $P^*$.

PROOF. The functions $\mathbf{r} \to U(P, \langle f, \mathbf{r}\rangle) = \int_W \langle f, \mathbf{r}\rangle \, dP$ on $J_{n+1}$, for fixed $f \, \varepsilon \, C(Ge^n)$, are uniformly continuous, and so may be extended to functions $\mathbf{t} \to U(\langle f, \mathbf{t}\rangle)$ on $J_{n+1}^*$. These determine $P^*$ just as $U_m(w, \cdot)$ determined $Q_{m,w}$, using e.g. Theorem 1.6 on p. 604 of [3]. Notice that $\alpha P^* = P$ entails

$$P^*[w^* \mid w^*(0) = 0] = 1.$$

LEMMA D.8. *Define* $W_4 = \{w \mid w \, \varepsilon \, W_3$ *and* $Q_w \, \varepsilon \, \mathbf{Q}_0\}$. *Then* $W_4 \, \varepsilon \, \mathfrak{F}_d$, *and* $P(W_4) = 1$, *all* $P \, \varepsilon \, \mathbf{Q}$.

These lemmas together imply

THEOREM D.1. *There is a set* $W_4 \, \varepsilon \, \mathfrak{F}_d$, *and corresponding to each* $w \, \varepsilon \, W_4$ *a probability* $Q_w^*$ *on* $\mathfrak{F}^*$ *such that:*

(i) $T^r W_4 = W_4$, *and* $Q_{T^r w}^* = Q_w^*$, *for all* $r \, \varepsilon \, R$;

(ii) $P(W_4) = 1$, *all* $P \, \varepsilon \, \mathbf{Q}_0$;

(iii) $Q_w^* \, \varepsilon \, \mathbf{Q}_0^*$ *is d-ergodic, and* $Q_w^*\{w^* \mid w^*(0) = 0\} = 1$;

(iv) *if* $f$ *is bounded and* $\mathfrak{F}_d^*$*-measurable, then* $w \to \int_{W^*} f\, dQ_w^*$ *is* $W_4\mathfrak{F}_d$*-measurable, and* $P^* \, \varepsilon \, \mathbf{Q}_0^*$ *implies*

$$\int_{W^*} f\, dP^* = \int_{W_4} \int_{W^*} f\, dQ_w^*\, d\alpha P^*.$$

**3. Applications.** The results of Section 2 will now be used to characterize mixtures of the laws of the following families of stochastic processes:

(i) stationary Markoff chains, with standard transition matrix ([2], II.2) and no instantaneous states, ([2], p. 149)—a stationary Markoff chain has no fixed points of discontinuity if and only if it has a standard transition matrix and no instantaneous states; see Theorem 3 on p. 141 and Theorem 2 on p. 154 of [2];

(ii) processes with stationary, independent increments and no fixed points of discontinuity;

(iii) the subfamily of (ii) consisting of Brownian motions;

(iv) the subfamily of (ii) consisting of Poisson processes.

To explain e.g. (i) more precisely, let $I$ be a countable set with the discrete topology. Let $(I', \rho)$ be (a metric space homeomorphic to) the one point compactification of $I$. Let $L = \{L(\cdot, \mathbf{t}); \mathbf{t} \in \bigcup_n J_n^*\}$ be a law in the sense of Lemma 7, with $I'$ for $S$. Suppose that $(\mathfrak{X}, \mathfrak{A}, P)$ is a probability space and for each $t \in R^*$, $X_t$ is a measurable function from $(\mathfrak{X}, \mathfrak{A})$ to $(I', \mathfrak{F}(I'))$. Then $L$ is called the *law of the (stochastic) process* $\{X_t : t \in R^*\}$ *(under* $P$*, if any ambiguity is possible)* if for each $n$, each $\mathbf{t} \in J_n^*$, each $f \in C(I'^n)$,

$$\text{(S)} \qquad L(f, \mathbf{t}) = \int_{\mathfrak{X}} f[X_{t_1}, \cdots, X_{t_n}]\, dP.$$

It is well known that many properties of a process are determined by its law—for example: the property of being stationary; the property of having a fixed point of discontinuity at 0; or the property of being a stationary Markoff chain with state space a subset of $I$, standard transition matrix, and no instantaneous states.

The problem faced in part (i) of Section 3 is how to decide whether there exists a probability space $(\Lambda, \mathfrak{B}, \mu)$ with three properties:

(a) to each $\lambda \in \Lambda$ there corresponds a law $L_\lambda$ and any process in relation (S) to $L_\lambda$ is a stationary Markoff chain with state space a subset of $I$ and

(b) standard transition matrix with no instantaneous states;

(c) for each $n$, each $\mathbf{t} \in J_n^*$, each $f \in C(I'^n)$, the function $\lambda \to L_\lambda(f, \mathbf{t})$ is $\mathfrak{B}$-measurable and

$$L(f, \mathbf{t}) = \int_\Lambda L_\lambda(f, \mathbf{t})\mu\, (d\lambda).$$

Three conditions are obviously necessary for the existence of such a $(\Lambda, \mathfrak{B}, \mu)$. Any process in relation (S) to $L$, i.e., of which $L$ is the law, must (i) be $I$-valued a.s. at time 0; and (ii) have no fixed point of discontinuity. Moreover (iii) $L$ must be *transition-exchangeable*. This concept is most easily defined using the Kronecker symbol, $\delta(a, b) = 1$ when $a = b$; $= 0$ when $a \neq b$. Let $\sigma \in \Sigma$ if and only if for some $n$ (called the length of $\sigma$) $\sigma = [\mathbf{\sigma}(1), \mathbf{\sigma}(2)]$, where $\mathbf{\sigma}(1) \in J_n^*$ and $\sigma_\nu(2) \in I$, $1 \leqq \nu \leqq n$. Write $\sigma \sim \tau$ when both are in $\Sigma$ if and only if both have the same length (say $n$), $\sigma_1(2) = \tau_1(2)$; and if $n \geqq 2$ for each $t > 0, j \in I$, $k \in I$,

$$\sum_{\nu=2}^{n} \delta[t, \sigma_\nu(1) - \sigma_{\nu-1}(1)]\delta[\sigma_{\nu-1}(2), j]\delta[\sigma_\nu(2), k]$$
$$= \sum_{\nu=2}^{n} \delta[t, \tau_\nu(1) - \tau_{\nu-1}(1)]\delta[\tau_{\nu-1}(2), j]\delta[\tau_\nu(2), k].$$

Thus $\sigma \sim \tau$ if and only if both have the same length, begin with same state of $I$, and exhibit the same transitions although possibly in different order. Then $L$ is *transition-exchangeable* if and only if for any $\sigma \varepsilon \Sigma$, $\tau \varepsilon \Sigma$, $\sigma \sim \tau$ implies

$$(*) \qquad L\{\delta[\boldsymbol{\sigma}(2), \cdot], \boldsymbol{\sigma}(1)\} = L\{\delta[\boldsymbol{\tau}(2), \cdot], \boldsymbol{\tau}(1)\}.$$

To paraphrase, $L$ is transition-exchangeable means that a process having $L$ for law is in states $\sigma_\nu(2)$ at times $\sigma_\nu(1)$ respectively with the same probability that it is in states $\tau_\nu(2)$ at times $\tau_\nu(1)$ respectively whenever $\sigma \sim \tau$. This implies, among other things, that the process is stationary. If $\mathbf{u} \varepsilon I^n$, then $\mathbf{v} \to \delta(\mathbf{u}, \mathbf{v})$ is a continuous function on $I'^n$, since $I$ has the discrete topology. Hence both sides of $(*)$ are defined.

The necessity of condition (iii) is almost trivial, since $(*)$ holds with $L$ replaced by $L_\lambda$, for each $\lambda \varepsilon \Lambda$. Then $(*)$ follows by integrating out $\lambda$ with respect to $\mu$ and applying (c).

Conversely, if $L$ satisfies these three conditions then a $(\Lambda, \mathfrak{B}, \mu)$ with properties (a), (b), (c) exists. This is non-trivial, and is a consequence of Theorem 2.

It is possible that a $(\Lambda, \mathfrak{B}, \mu)$ with properties (a) and (c) exists if and only if $L$ has properties (i) and (iii). I have not been able to settle this.

To avoid clumsy formulations, Theorem 2 is stated in terms of probability measures in function space. Construct $(\Omega^*, \mathfrak{F}^*)$ and $(\Omega, \mathfrak{F})$ as in Section 2, with $(I', \rho)$ for $(S, \rho)$. Recall that $\mathbf{P}_0^*$ is the class of invariant probabilities on $\mathfrak{F}^*$ which have no fixed points of discontinuity (Definition 2), and the coordinate process $\{\xi_t^* : t \varepsilon R^*\}$ on $(\Omega^*, \mathfrak{F}^*)$ is defined by $\xi_t^*(\omega^*) = \omega^*(t^*)$. By the Kolmogoroff consistency theorem (see Lemma 7 for the discussion) there is a unique $P^*$ on $\mathfrak{F}^*$ such that $L$ is the law of $\{\xi_t^* : t \varepsilon R^*\}$ under $P^*$. If $L$ has properties (i), (ii), (iii) clearly $P^* \varepsilon \mathbf{P}_0^*$. Write $M$ for the set of all $\omega \varepsilon \Omega_4 \subset \Omega$ (see Lemma 8) such that under $P_\omega^*$ (see Theorem 1) the process $\{\xi_t^* : t \varepsilon R^*\}$ is Markoff with state space contained in $I$. According to Theorem 2, $\alpha P^*(M) = 1$. (Recall that $\alpha$ projects $\Omega^*$ onto $\Omega$.) This demonstrates the existence of a $(\Lambda, \mathfrak{B}, \mu)$ with properties (a), (b), (c). Indeed, for $\Lambda$ take $M$, for $\mathfrak{B}$ take $M\mathfrak{F}$, for $\mu$ take $\alpha P^*$ restricted to $M\mathfrak{F}$, and for $\lambda \varepsilon \Lambda$ take $L_\lambda$ to be the law of the process $\{\xi_t^* : t \varepsilon R^*\}$ under $P_\lambda^*$. The relation $\alpha P^*(M) = 1$ implies $\mu(\Lambda) = 1$. Condition (a) follows from the definition of $M$; so does (b), using the remark made earlier that a stationary Markoff chain has no fixed points of discontinuity if and only if it has a standard transition matrix and no instantaneous states. Condition (c) follows from Theorem 1 (iv), bearing in mind that $\alpha P^*(M) = 1$.

Similar discussions for Theorems 3–7 are omitted.

THEOREM 2. *The set $M \varepsilon \mathfrak{F}$. Moreover, if $P^* \varepsilon \mathbf{P}_0^*$, then $\alpha P^*(M) = 1$ if and only if $P^*\{\omega^* \mid \omega^*(0) \varepsilon I\} = 1$ and the law of the process $\{\xi_t^* : t \varepsilon R^*\}$ under $P^*$ is transition exchangeable.*

PROOF. Since $\omega \varepsilon \Omega_4$ implies $P_\omega^* \varepsilon \mathbf{P}_0^*$, $\omega \varepsilon M$ if and only if $\omega \varepsilon \Omega_4$,

$$P_\omega\{\omega' \mid \omega' \varepsilon \Omega, \omega'(0) \varepsilon I\} = 1,$$

and for all $n$, all $\mathbf{r} \varepsilon J_{n+1}$, all $\mathbf{i} \varepsilon I^{n+1}$,

$$(12) \quad \begin{aligned} &P_\omega[\omega' \mid \omega'(r_j) = i_j, 1 \leqq j \leqq n+1]P_\omega[\omega' \mid \omega'(r_n) = i_n] \\ &\quad = P_\omega[\omega' \mid \omega'(r_j) = i_j, 1 \leqq j \leqq n]P_\omega[\omega' \mid \omega'(r_n) = i_n, \omega'(r_{n+1}) = i_{n+1}], \end{aligned}$$

where $\omega' \varepsilon \Omega$. This follows because for $\omega \varepsilon \Omega_4$ the probabilities $P_\omega$ and $P_\omega^*$ have no fixed points of discontinuity, so that (12) implies a similar relation with $P_\omega$ replaced by $P_\omega^*$, $\omega'$ replaced by $\omega^* \varepsilon \Omega^*$, and $\mathbf{r}$ by $\mathbf{t} \varepsilon J_{n+1}^*$. Thus $M \varepsilon \mathfrak{F}$.

The "only if" assertion is clear from part (iv) of Theorem 1. In the other direction, suppose $P^*$ satisfies the condition. Since

$$1 = \alpha P^*\{\omega' \mid \omega' \varepsilon \Omega, \omega'(0) \varepsilon I\} = \int_{\Omega_4} P_\omega\{\omega' \mid \omega' \varepsilon \Omega, \omega'(0) \varepsilon I\}\alpha P^*(d\omega),$$

there is an $\alpha P^*$-null set $N \varepsilon \mathfrak{F}$ such that $\omega \varepsilon \Omega_4 - N$ implies

$$P_\omega\{\omega' \mid \omega' \varepsilon \Omega, \omega'(0) \varepsilon I\} = 1.$$

Let $\Omega(m) = \prod_{r \varepsilon R_m} I$, in the product topology, where

$$R_m = \{j2^{-m} : j = 0, \pm 1, \cdots\}.$$

Let $\theta_m$ be the homeomorphism of $\Omega(m)$ onto itself defined by $(\theta_m u)(r) = u(r + 2^{-m})$, $u \varepsilon \Omega(m)$, $r \varepsilon R_m$. Let $\pi_{m,\omega}$ be the projection of $P_{m,\omega}$ into $\Omega(m)$, for $\omega \varepsilon \Omega_4 - N$. (The $P_{m,\omega}$ were defined in the paragraph before Equation (3)). There is an $\alpha P^*$-null set $N_m \varepsilon \mathfrak{F}$, such that $\omega \varepsilon \Omega_4 - N - N_m$ implies $\pi_{m,\omega}$ is summarized (Definition 3 of [4]) by the statistics $\{T_n\}$ (paragraph after Definition 2 of [4]). This follows using the argument for Theorem 1 of [4], appealing to Equation (2) of the present paper, instead of Equation (2) of [4]. But $\omega \varepsilon \Omega_4$ implies $P_{m,\omega}$ is $(m)$-ergodic, so $\omega \varepsilon \Omega_4 - N$ implies $\pi_{m,\omega}$ is ergodic under $\theta_m$. Hence $\omega \varepsilon \Omega_4 - N - N_m$ implies that under $\pi_{m,\omega}$ the coordinate process on $\Omega(m)$ is Markoff, by Theorem 2 of [4]. Consequently, if $\omega \varepsilon \Omega_4 - N - \bigcup_{m=0}^\infty N_m$, and $m$ is so large that all the $r_j$ in (12) are of the form $n_j 2^{-m}$, the $n_j$ integers, then (12) will hold with $P_\omega$ replaced by $P_{m,\omega}$. But ($I$ has the discrete topology) the indicators of the sets in (12) are continuous functions of $\omega'$. Allow $m \uparrow \infty$, and apply Lemma 3, to complete the proof of (12), and with it Theorem 2.

Theorem 3 deals with the spaces $W^*$ and $W$ of Section 2. Let $(\mathfrak{X}, \mathfrak{A}, P)$ be a probability space. For each $t \varepsilon R^*$, let $X_t$ be measurable from $(\mathfrak{X}, \mathfrak{A})$ to $(G, \mathfrak{F}(G))$. Definitions 4 and 5 will be applied primarily when $X_t = \xi_t^*$ and $(\mathfrak{X}, \mathfrak{A}) = (W^*, \mathfrak{F}^*)$, but it is convenient to state them more generally.

Definition 4. The process $\{X_t : t \varepsilon R^*\}$ has independent increments under $P$, if and only if for all $n$, $\mathbf{t} \varepsilon J_n^*$, the random variables

$$\{X_{t_1}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}\}$$

are independent under $P$.

The concept of independent increments for $\{\xi_r : r \varepsilon R\}$ under $P$ on $(W, \mathfrak{F})$ is defined similarly. The next step is to define *exchangeable increments*. Let $\sigma \varepsilon \Sigma$ if and only if for some $n$ (called the length of $\sigma$), $\sigma = [(\sigma_{\nu,1}, \sigma_{\nu,2}) : 1 \leqq \nu \leqq n]$, where $\sigma_{\nu,1} > 0$ and $\sigma_{\nu,2} \varepsilon \mathfrak{F}(G)$ for $1 \leqq \nu \leqq n$. For each $\sigma \varepsilon \Sigma$, let $C(\sigma)$ be the set of all $\mathbf{s} \varepsilon R^{*n}$ (where $n$ is the length of $\sigma$) having the property that the open intervals $(s_\nu, s_\nu + \sigma_{\nu,1}) : 1 \leqq \nu \leqq n$ are disjoint. For $\mathbf{s} \varepsilon C(\sigma)$, define $A(\sigma, \mathbf{s}) = \{x \mid x \varepsilon \mathfrak{X}, X_{s_\nu + \sigma_{\nu,1}}(x) - X_{s_\nu}(x) \varepsilon \sigma_{\nu,2}, 1 \leqq \nu \leqq n\}$.

DEFINITION 5. The process $\{X_t : t \varepsilon R^*\}$ has exchangeable increments under the probability $P$ on $(\mathfrak{X}, \mathfrak{A})$ if and only if $\sigma \varepsilon \Sigma$, $\mathbf{t} \varepsilon C(\sigma)$, $\mathbf{s} \varepsilon C(\sigma)$ imply

$$P[A(\sigma, \mathbf{s})] = P[A(\sigma, \mathbf{t})].$$

If $(\mathfrak{X}, \mathfrak{A}) = (W^*, \mathfrak{F}^*)$ and $X_t = \xi_t^*$, then $P \varepsilon \mathbf{Q}^*$ automatically.

Let $w \varepsilon L$ if and only if $w \varepsilon W_4$ (Lemma D.8) and under $Q_w^*$ (Theorem D.1) the process $\{\xi_t^* : t \varepsilon R^*\}$ has independent increments. Then

THEOREM 3. *The set $L \varepsilon \mathfrak{F}_d$. Moreover, if $P^* \varepsilon \mathbf{Q}_0^*$ then $\alpha P^*(L) = 1$ if and only if $\{\xi_t^* : t \varepsilon R^*\}$ has exchangeable increments under $P^*$.*

PROOF. A point $w$ of $W_4$ is in $L$ if and only if, for all $n$, all $k$, and all $\nu$

$$(13)\qquad \begin{aligned} &\int_W \prod_{j=1}^n f_\nu\{w'[j2^{-k}] - w'[(j-1)2^{-k}]\} Q_w(dw') \\ &\qquad = \prod_{j=1}^n \int_W f_\nu\{w'[j2^{-k}] - w'[(j-1)2^{-k}]\} Q_w(dw'), \end{aligned}$$

where $\{f_\nu\}$ is dense in $C(Ge)$. Hence $L \varepsilon \mathfrak{F}_d$.

The necessity of the condition being clear from (iv) of Theorem D.1, suppose $P^*$ satisfies the condition. Introduce the processes $Y_{m,j}(w) = w[j2^{-m}] - w[(j-1)2^{-m}]$, $j = 0, \pm 1, \pm 2, \cdots$. For $w$ outside an $\alpha P^*$-null set $N_m \varepsilon \mathfrak{F}_d$, under $Q_{m,w}$ the process $\{Y_{m,j} : j = 0, \pm 1, \cdots\}$ is exchangeable; that is, any $n$ distinct variables of the process induce the same measure in $G^n$.

To prove this, let $\{g_\nu\}$ be dense in $C(Ge^n)$. Let $\pi(1 \cdots n) = (i_1 \cdots i_n)$ be a permutation, and for $g \varepsilon C(Ge^n)$ define $(\pi g)(x_1 \cdots x_n) g(x_{i_1} \cdots x_{i_n})$. Write

$$g_\nu[Y_{m,1}(w) \cdots Y_{m,n}(w)] = G_\nu(w), \qquad \pi g_\nu[Y_{m,1}(w) \cdots Y_{m,n}(w)] = \pi G_\nu(w).$$

Then it is enough to prove that, for a fixed $n$, $\nu$ and $\pi$, a.s. $[\alpha P^*]$,

$$(14)\qquad \int_W G_\nu(w') Q_{m,w}(dw') = \int_W \pi G_\nu(w') Q_{m,w}(dw').$$

For then on discarding a countable union of $\alpha P^*$—null sets of $\mathfrak{F}_d$, (14) would hold for all $\nu$ and all $\pi$, and hence for all $f \varepsilon C(Ge^n)$ and all $\pi$. It would then hold for all $n$ by a similar ritual, and that gives exchangeability.

But the $L^2(W, \mathfrak{F}, \alpha P^*)$ norm of the difference of the two sides of (14) (which are real functions of $w$) is, in virtue of (D.7), dominated convergence, and the continuity of $G_\nu$ and $\pi G_\nu$,

$$
\begin{aligned}
&\lim_{N\to\infty} \int_W \left[ N^{-1} \sum_{j=0}^{N-1} (G_\nu - \pi G_\nu)(T^{j2^{-m}} w') \right]^2 \alpha P^*(dw') \\
&\quad = \lim_{N\to\infty} N^{-2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left\{ \int_W G_\nu(T^{j2^{-m}} w') G_\nu(T^{k2^{-m}} w') \alpha P^*(dw') \right. \\
(15)\qquad &\quad + \int_W \pi G_\nu(T^{j2^{-m}} w') \pi G_\nu(T^{k2^{-m}} w') \alpha P^*(dw') \\
&\quad - \int_W G_\nu(T^{j2^{-m}} w') \pi G_\nu(T^{k2^{-m}} w') \alpha P^*(dw') \\
&\quad \left. - \int_W \pi G_\nu(T^{j2^{-m}} w') G_\nu(T^{k2^{-m}} w') \alpha P^*(dw') \right\}.
\end{aligned}
$$

For $|j - k| \geqq n$, each summand vanishes, in view of the exchangeability of $\{Y_{m,j} : j = 0, \pm, \cdots\}$ under $P^*$. The relative frequency of terms with $|j - k| < n$ goes to 0, and each lies in $[-2\, \|g_\nu\|, +2\, \|g_\nu\|]$, proving that both sides of (15) are 0, so that (14) holds, a.s. $[\alpha P^*]$.

But for $w \,\varepsilon\, W_4$, $Q_{m,w}$ is $(m)$ $d$-ergodic, which implies that under it

$$\{Y_{m,j} : j = 0, \pm 1, \cdots\}$$

is metrically transitive. For $w \,\varepsilon\, W_4 - \bigcup_{m=0}^{\infty} N_m$ it follows that under $Q_{m,w}$ the process $\{Y_{m,j} : j = 0, \pm 1 \cdots\}$ consists of independent, identically distributed random variables. To see this clearly, notice that if $A_j \,\varepsilon\, \mathfrak{F}(G)$, $1 \leqq j \leqq n + 1$, and $\nu \geqq n + 1$

$$
\begin{aligned}
&Q_{m,w}\{w' \mid w' \,\varepsilon\, W,\, Y_{m,j}(w') \,\varepsilon\, A_j,\, 1 \leqq j \leqq n + 1\} \\
&\qquad = Q_{m,w}\{w' \mid w' \,\varepsilon\, W,\, Y_{m,j}(w') \,\varepsilon\, A_j,\, 1 \leqq j \leqq n,\, Y_{m,\nu} \,\varepsilon\, A_{n+1}\}
\end{aligned}
$$

in virtue of exchangeability. But as $\nu \to \infty$, the right side goes in first Cesaro mean to

$$
\begin{aligned}
&Q_{m,w}\{w' \mid w' \,\varepsilon\, W,\, Y_{m,j}(w') \,\varepsilon\, A_j,\, 1 \leqq j \leqq n\} \\
&\qquad Q_{m,w}\{w' \mid w' \,\varepsilon\, W,\, Y_{m,n+1}(w') \,\varepsilon\, A_{n+1}\}
\end{aligned}
$$

in virtue of metric transitivity (see p. 435 of [8]). (This argument evidently recapitulates the proof of De Finetti's theorem given in [4]). Hence (13) holds provided $m \geqq k$, and $Q_w$ is replaced by $Q_{m,w}$. Allow $m \uparrow \infty$ and apply Lemma D.3 to complete the proof.

The substantial assertion in Theorem 3 is: if (i) $P^* \,\varepsilon\, \mathbf{Q}_0^*$ and (ii) $\{\xi_t^* : t \,\varepsilon\, R^*\}$ has exchangeable increments under $P^*$ then $P^*$ is a weighted average of probabilities under each of which $\{\xi_t^* : t \,\varepsilon\, R^*\}$ has stationary, independent increments and no fixed points of discontinuity. Now (i) may be weakened when $G$ is abelian and has a connected character group. For then $P \,\varepsilon\, \mathbf{Q}$ and $\{\xi_r : r \,\varepsilon\, R\}$ having exchangeable increments under $P$ imply that $P \,\varepsilon\, \mathbf{Q}_0$ (the argument is the same as on the real line). Hence, if $P^*$ satisfies (ii), it follows that $\alpha P^* \,\varepsilon\, \mathbf{Q}_0$.

Suppose also that $P^*$ is *continuous in distribution*, which means that for any $n$, if $\mathbf{t}(\nu) \varepsilon J_n^*$ converges to $\mathbf{t} \varepsilon J_n^*$ as $\nu \to \infty$, and if $f \varepsilon C(Ge^n)$,

$$\int_{W^*} f \circ \pi^*[\mathbf{t}(\nu)] \, dP^* \to \int_{W^*} f \circ \pi^*(\mathbf{t}) \, dP^*.$$

Then $P^*$ is determined by $\alpha P^*$. For simplicity, suppose $P^*\{w^* \mid w^*(0) = 0\} = 1$, and define $Q^*$ on $\mathfrak{F}^*$ by $Q^*(A) = \int_L Q_w^*(A) \, d\alpha P^*$. It follows that $\alpha Q^* = \alpha P^*$, so $P^* = Q^* \varepsilon \mathbf{Q}_0^*$.

Even when $G$ is the additive real line, (i) may not be dropped entirely. For if $\varphi$ is any non-measurable function with $\varphi(t + s) = \varphi(t) + \varphi(s)$, there is a unique probability on $\mathfrak{F}^*$, assigning outer measure 1 to $\{\varphi\}$; points of $W^*$ are not in $\mathfrak{F}^*$, and have inner measure 0 for any measure on $\mathfrak{F}^*$. Of course, this probability satisfies (ii) (under it the coordinate process has stationary, independent increments) but not (i).

DEFINITION 6. The real random variables $\{Z_1 \cdots Z_n\}$ on $(W^*, \mathfrak{F}^*, P^*)$ will be called isotropic (under $P^*$) if their joint distribution (under $P^*$) has spherical symmetry (in $n$-dimensions). The bilateral sequence $\{ \cdots Z_{-1}, Z_0, Z_1, \cdots \}$ will be called isotropic under $P^*$ if, for every $n$, $\{Z_{-n}, \cdots, Z_n\}$ is isotropic under $P^*$. The process $\{Z_t : -\infty < t < \infty\}$ will be called $d$-isotropic under $P^*$ if, for every $h > 0$, the bilateral sequence $Z_n' = Z_{nh} - Z_{(n-1)h}$ is isotropic under $P^*$.

LEMMA 9. *Let $X$ and $Y$ be random variables on $(W^*, \mathfrak{F}^*)$ which are isotropic under $P^*$ and independent. Then, under $P^*$, they are normal with mean* 0.

PROOF. This is an immediate consequence of known results, but it is easy to give a direct proof. It is clear that $X$ and $Y$ have a common characteristic function $z \to \varphi(z^2)$ under $P^*$, where $\varphi$ is a real continuous function on $[0, \infty)$, $\varphi(0) = 1$, $|\varphi| \leqq 1$. Their joint characteristic function under $P^*$ is $(z_1, z_2) \to \psi(z_1^2 + z_2^2)$, where $\psi$ is real and continuous on $[0, \infty)$, and $\psi(0) = 1$. Hence $\varphi(z_1^2)\varphi(z_2^2) = \psi(z_1^2 + z_2^2)$, and the only continuous solution is $\varphi(u) = e^{\lambda u}$. Since $\varphi$ is real and bounded, $\lambda \leqq 0$, completing the proof.

For Theorem 4, replace $G$ in $W^*$ and $W$ by the additive group of real numbers. Let $B = \{w \mid w \varepsilon L$, and under $Q_w^*$ the process $[\xi_t^* : t \varepsilon R^*]$ is a Brownian motion$\}$. Then

THEOREM 4. *The set $B \varepsilon \mathfrak{F}_d$. Moreover, if $P^* \varepsilon \mathbf{Q}_0^*$, the necessary and sufficient condition that $\alpha P^*(B) = 1$ is: under $P^*$ the process $\{\xi_t^* : t \varepsilon R^*\}$ is $d$-isotropic.*

PROOF. Let $\{f_s\}$ be a dense sequence in the space $C_0(R^n)$ of continuous functions on $R^n$ vanishing off compact sets, and let $\{\rho_\nu\}$ be a sequence of rotations of $R^n$ dense in the set of all rotations of $R^n$. If $f \varepsilon C_0(R^n)$, define $(\rho_\nu f)(\mathbf{x}) = f(\rho_\nu \mathbf{x})$ for $\mathbf{x} \varepsilon R^n$, so $\rho_\nu f \varepsilon C_0(R^n)$.

Then by Lemma 9, $w \varepsilon B$ if and only if $w \varepsilon L$ and for all $k$, $\nu$, $s$, and $n = 2$,

$$\begin{aligned} &\int_W f_s\{w'[j2^{-k}] - w'[(j-1)2^{-k}], 1 \leqq j \leqq n\} Q_w(dw') \\ (16) \qquad &= \int_W \rho_\nu f_s\{w'[j2^{-k}] - w'[(j-1)2^{-k}], 1 \leqq j \leqq n\} Q_w(dw'). \end{aligned}$$

Thus $B \varepsilon \mathfrak{F}_d$. Necessity is again clear by Theorem D.1 (iv). For sufficiency, the argument used in Theorem 3 shows that if $P^*$ satisfies the given condition, there is a $\alpha P^*$-null set $N \varepsilon \mathfrak{F}_d$, such that $w \varepsilon W_4 - N$ implies (16) will be satisfied with $Q_w$ replaced by $Q_{m,w}$, provided $m \geqq k$. Allowing $m \uparrow \infty$ completes the proof.

For Theorems 5, 6, and 7 it is convenient to work in the spaces $W^*$ and $W$ of Section 2, with the additive group of integers for $G$. Let $S = \{w \mid w \varepsilon L$, and under $Q_w^*$, $[\xi_t^* : t \varepsilon R^*]$ is a Poisson process$\}$. Then

THEOREM 5. *The set $S \varepsilon \mathfrak{F}_d$, and if $P^* \varepsilon \mathbf{Q}_0^*$, the necessary and sufficient condition that $\alpha P^*(S) = 1$ is: for each $n \geqq 2$, each $\mathbf{t} \varepsilon J_n^*$, each vector $[i_j : 2 \leqq j \leqq n]$ of non-negative integers, under $P^*$, $w^*(t_j) \geqq w^*(t_{j-1})$* a.s., *and*

$$P^*\{w^*(t_j) - w^*(t_{j-1}) = i_j, 2 \leqq j \leqq n\}\prod_{j=2}^{n} i_j!(t_j - t_{j-1})^{-i_j}$$
$$= f(t_n - t_1, n, \sum_{j=2}^{n} i_j),$$

*where $f(\cdot, \cdot, \cdot)$ is some function of triplets of positive reals, positive integers, non-negative integers.*

PROOF. If $w \varepsilon L$, under $Q_w$ the $\{Y_{m,j} : j = 0, \pm 1 \cdots\}$ process is a sequence of independent, identically distributed random variables; and $w \varepsilon S$ if and only if they are Poisson. Now ([4], Theorem 4) gives a necessary and sufficient condition for this: under $Q_w$, the random variable $Y_{m,j}$ is a.s. non-negative, and

$$Q_w[Y_{m,j} = i_j, 1 \leqq j \leqq n]\prod_{j=1}^{n} (i_j)! = Q_w[Y_{m,j} = k_j, 1 \leqq j \leqq n]\prod_{j=1}^{n} (k_j)! \tag{17}$$

whenever $i_j$ and $k_j$ are non-negative integers with $\sum_{j=1}^{n} i_j = \sum_{j=1}^{n} k_j$. Since there are only a finite number of $n$-tuples of non-negative integers with a fixed sum, (17) implies $S \varepsilon \mathfrak{F}_d$. As before, necessity is clear; conversely if $P^*$ satisfies the condition, then (17) holds with $Q_w$ replaced by $\alpha P^*$. To see this, replace $n$ in the condition of the theorem by $n + 1$, $t_j$ by $(j - 1)2^{-m}$ and $i_j$ by $i_{j-1}$, $1 \leqq j \leqq n + 1$, so that $t_{n+1} - t_1 = n2^{-m}$ depends only on $n$ and $m$, and

$$\prod_{j=2}^{n+1} (t_j - t_{j-1})^{-i_{j-1}} = 2^{-m\Sigma_{j=2}^{n} i_j}$$

depends only on $m$ and $\sum_{j=1}^{n} i_j$. By previous arguments there is an $\alpha P^*$-null $N \varepsilon \mathfrak{F}_d$ such that if $w \varepsilon W_4 - N$, (17) holds with $Q_w$ replaced by $Q_{m,w}$, for sufficiently large $m$. Since the indicator function of the set inside braces in (17) is continuous ($G$ having the discrete topology) an application of Lemma D.3 completes the proof.

The problem of characterizing mixtures of Poisson processes goes back at least to Lundberg (1940). Some variations on Theorem 5 might therefore be of interest.

DEFINITION 7. A point $w^* \varepsilon W^*$ is a counting function if and only if there is a bilateral sequence $\{ \cdots \sigma_{-1}, \sigma_0, \sigma_1, \cdots \}$ of positive extended real numbers, with $\sum_{i=1}^{\infty} \sigma_i = \sum_{i=-1}^{\infty} \sigma_i = \infty$ such that

$$w^*(t) - w^*(0) = \min \{n \mid n \geqq 0, \sum_{i=0}^{n} \sigma_i \geqq t\}$$

for $t \geqq 0$; and $w^*(0) - w^*(t) = \min \{n \mid n \geqq 0, \sum_{i=-(n+1)}^{-1} (-\sigma_i) \leqq t\}$ for $t < 0$. A point $w \varepsilon W$ is a counting function if it is the restriction to $R$ of a counting function $w^* \varepsilon W^*$. Call $W_\sigma$ the set of counting functions in $W$.

Notice that $w \varepsilon W_\sigma$ determines $\sigma_i = \sigma_i(w)$ uniquely. Moreover, it is not hard to see that $W_\sigma \varepsilon \mathfrak{F}_d$ and $\sigma_i(\cdot)$ is $\mathfrak{F}_d$-measurable. However, the set of counting functions in $W^*$ is not in $\mathfrak{F}^*$. For $t \varepsilon R^*$, put $Y_t(w) = 0, w \not\varepsilon W_\sigma$; $= \lim_{r \downarrow t, r \varepsilon R} w(r)$, $w \varepsilon W_\sigma$. Then $(t, w) \rightarrow Y_t(w) - Y_0(w)$ is $(-\infty, \infty) \times \mathfrak{F}_d$-measurable.

THEOREM 6. *Let $P^* \varepsilon \mathbf{Q}^*$ and be continuous in distribution; suppose that $\alpha P^*(W_\sigma) = 1$ and the process $\{\xi_t^* : t \varepsilon R^*\}$ has exchangeable increments under $P^*$. Then $P^* \varepsilon \mathbf{Q}_0^*$ and $\alpha P^*(S) = 1$.*

REMARK. The converse is omitted as trivial.

PROOF. From the proof of Theorem 3, there exists an $\alpha P^*$-null set $N \varepsilon \mathfrak{F}_d$ such that for $w \varepsilon W_3 - N$ (Lemma D.5) the process $\{\xi_r : r \varepsilon R\}$ has stationary, independent increments under $Q_w$. From Equation (D.7), there is a further $\alpha P^*$-null set $M \varepsilon \mathfrak{F}_d$ such that $w \varepsilon W_3 - N - M$ implies $Q_w(W_\sigma) = 1$. But then by continuity considerations, under $Q_w$ the process $\{Y_t : t \varepsilon R^*\}$ has stationary, independent increments and all its sample functions are counting functions. For such $w$ ([3], pp. 398–404), $Q_w \varepsilon \mathbf{Q}_0$ and under $Q_w$ the process $\{Y_t : t \varepsilon R^*\}$ is Poisson. Further, $Q_w \varepsilon \mathbf{Q}_0$ implies that the probability measure induced on $(W^*, \mathfrak{F}^*)$ by $\{Y_t\}$ is precisely $Q_w^*$, so that $w \varepsilon S$ and $\alpha P^*(S) = 1$. Since $P^*$ is continuous in distribution, it is uniquely determined by $\alpha P^*$. If $Q^*$ is defined on $\mathfrak{F}^*$ by $Q^*(A) = \int_S Q_w^*(A) \alpha P^*(dw)$, then $\alpha Q^* = \alpha P^*$, so that $P^* = Q^* \varepsilon \mathbf{Q}_0^*$, completing the proof.

Let $(\mathfrak{X}, \mathfrak{A}, P)$ be a probability space, and $(\cdots \sigma_{-1}, \sigma_0, \sigma_1 \cdots)$ a bilateral sequence of real $\mathfrak{A}$-measurable functions, such that (i) $0 < \sigma_i < \infty$; (ii) $\sum_{i=1}^{\infty} \sigma_i = \sum_{i=1}^{\infty} \sigma_{-i} = \infty$. Define $Y_t(x) = \min \{n \mid n \geqq 0, \sum_{i=0}^{n} \sigma_i(x) \geqq t\}$ for $t \varepsilon [0, \infty]$, $x \varepsilon \mathfrak{X}$; $= \min \{n \mid n \geqq 0, \sum_{i=-(n+1)}^{-1} [-\sigma_i(x)] \leqq t\}$ for $t \varepsilon (-\infty, 0)$, $x \varepsilon \mathfrak{X}$. Then $(t, x) \rightarrow Y_t(x)$ is $\mathfrak{F}(-\infty, \infty) \times \mathfrak{A}$-measurable.

THEOREM 7. *The following four conditions are equivalent*:

(i) *for any $n$ and $t_i > 0$, $-n \leqq i \leqq n$: $P\{x \mid x \varepsilon \mathfrak{X}, \sigma_i(x) \geqq t_i, -n \leqq i \leqq n\} = f(n, \sum_{i=-n}^{n} t_i)$, where $f(\cdot, \cdot)$ is some function of pairs of non-negative integers, positive real numbers*;

(ii) *there is a probability $m$ on $\mathfrak{F}(0, \infty)$ for which*

$$P\{x \mid x \varepsilon \mathfrak{X}, \sigma_i(x) \geqq t_i, -n \leqq i \leqq n\} = \int_0^\infty \exp\left(-\lambda \sum_{i=-n}^{n} t_i\right) m(d\lambda);$$

(iii) *$\{Y_t : t \varepsilon R^*\}$ has exchangeable increments under $P$*;

(iv) *for each $n \geqq 2$ and $\mathbf{t} \varepsilon J_n^*$, and $i_j$ non-negative integers, $2 \leqq j \leqq n$,*

$$P\{x \mid x \varepsilon \mathfrak{X}, Y_{t_j}(x) - Y_{t_{j-1}}(x) = i_j, 2 \leqq j \leqq n\}$$
$$= \left[\prod_{j=2}^{n} (t_j - t_{j-1})^{i_j}(i_j!)^{-1}\right] \int_0^\infty \exp\left\{-\lambda(t_n - t_1) + (\log \lambda) \sum_{j=2}^{n} i_j\right\} m(d\lambda).$$

*The* $m$ *in* (ii) *and* (iv) *is the same, and is unique.*

PROOF. The equivalence of (i) and (ii) follows easily from Lemma 10, and the equivalence of (iii) and (iv) follows almost as easily from Theorem 6. Indeed, under (iii) the process $\{Y_r : r \varepsilon R\}$ has exchangeable increments and hence no fixed points of discontinuity. Since $Y.(x)$ is continuous from the right for each $x \varepsilon \mathfrak{X}$, by a routine argument $\{Y_t : t \varepsilon R^*\}$ has no fixed points of discontinuity and Theorem 6 applies to its law. Clearly, (iii) and (iv) imply (i) and (ii) (with the same $m$) because the holding times of a Poisson process are independent, identically distributed exponential random variables. Conversely, (i) and (ii) imply that given its invariant $\sigma$-field, the process $\{\cdots \sigma_{-1}, \sigma_0, \sigma_1 \cdots\}$ is distributed conditionally as a bilateral sequence of independent random variables with common exponential distribution (Lemma 10). Hence, given that $\sigma$-field, the process $\{Y_t\}$ has conditionally a Poisson law. This implies (iii) and (iv). The uniqueness of $m$ is an elementary fact of Laplace transform theory, and this completes the proof.

Lemma 10 is stated and proved in the notation of [11], with the following identifications: $\Omega$ is the space of bilateral sequences of non-negative extended real numbers, in the product topology; $T$ is the shift. In addition, call $\mathfrak{J}$ the $\sigma$-field of all Borel subsets of $\Omega$ invariant under $T$. Let $\{\xi_i : i = 0, \pm 1, \cdots\}$, $\xi_i(p) = p(i)$ be the coordinate process on $\Omega$. A probability $\mu$ on $\Omega$ will be called *planar* if $\mu\{q \mid q \varepsilon \Omega, 0 < \xi_0(q) < \infty\} = 1$ and for all $n$, all positive reals $x_\nu$, $y_\nu$, $-n \leqq \nu \leqq n$, $\sum_{\nu=-n}^{n} x_\nu = \sum_{\nu=-n}^{n} y_\nu$ implies

$$\mu\{q \mid q \varepsilon \Omega, \xi_\nu(q) \geqq x_\nu, -n \leqq \nu \leqq n\} = \mu\{q \mid q \varepsilon \Omega, \xi_\nu(q) \geqq y_\nu, -n \leqq \nu \leqq n\}. \tag{18}$$

A planar $\mu$ is evidently invariant under $T$, for under it the coordinate process is even exchangeable. Write $\Pi$ for the set of all $p \varepsilon Q_T$ such that $\mu_p$ is planar.

LEMMA 10.

(i) *The set* $\Pi \varepsilon \mathfrak{J}$, *and if* $\mu$ *is planar,* $\mu(\Pi) = 1$.

(ii) *If* $\nu$ *is planar and ergodic, under* $\nu$ *the coordinate process is a bilateral sequence of independent random variables with common exponential distribution.*

(iii) *if* $\mu$ *is planar, the* $\mu$*-conditional distribution of the* $\{\xi_i\}$ *given* $\mathfrak{J}$ *is that of a bilateral sequence of independent random variables with common exponential distribution.*

PROOF.

(i) Let $A_n$ be the set of pairs $\mathbf{x}$ and $\mathbf{y}$ of vectors of positive rationals of length $(2n + 1)$ with $\sum_{\nu=-n}^{n} x_\nu = \sum_{\nu=-n}^{n} y_\nu$. Let $B_n(\mathbf{x}, \mathbf{y})$ denote the set of $p \varepsilon Q_T$ for which (18) holds with $\mu_p$ for $\mu$. Then

$$\Pi = \{p \mid \mu_p[q \mid 0 < \xi_0(q) < \infty] = 1\} \cap \{B_n(\mathbf{x}, \mathbf{y}) \mid (\mathbf{x}, \mathbf{y}) \varepsilon A_n, n = 1, 2, \cdots\}$$

because if (18) holds for a given $\mu$ and all $(\mathbf{x}, \mathbf{y}) \varepsilon A_n$, it holds for all real positive $x_\nu$ and $y_\nu$ satisfying $\sum_\nu x_\nu = \sum_\nu y_\nu$, both sides of (18) being continuous from the left. Since for any Borel $B \subset \Omega$, $p \rightarrow \mu_p(B)$ is $Q_T\mathcal{I}$-measurable, it follows that $\Pi \varepsilon \mathcal{I}$.

For any fixed bounded Borel measurable $f$ on $\Omega$, $\{p \mid p \varepsilon Q_T, M(f, p) = \int f\, d\mu_p\}$ is invariant and has invariant probability 1. Indeed, if $\theta$ is an invariant probability, $p \rightarrow M(f, p)$ and $p \rightarrow \int f\, d\mu_p$ are both versions of $E_\theta(f \mid \mathcal{I})$. Let $Q_0$ be the set of all $p \varepsilon Q_T$ for which

$$M(f, p) = \int_\Omega f\, d\mu_p, \quad \text{all } f \varepsilon D \tag{19}$$

where $D$ is the (countable) set of all indicator functions of subsets of $\Omega$ of the type $\{q \mid \xi_\nu(q) \geqq x_\nu, -n \leqq \nu \leqq n\}$ where $n = 1, 2, \cdots$ and the $x_\nu$ are positive rationals. Then $Q_0 \varepsilon \mathcal{I}$ and has invariant probability 1. By familiar reasoning, if $\mu$ is planar then $\mu\{p \mid p \varepsilon Q_0, \mu_p[q \mid 0 < \xi_0(q) < \infty]\} = 1$. Using the even more familiar argument of Theorem 1 of [4], but appealing this time to (19), for each $n$ and $(\mathbf{x}, \mathbf{y}) \varepsilon A_n$, $\mu[B_n(\mathbf{x}, \mathbf{y})] = 1$. It follows that $\mu(\Pi) = 1$, proving (i).

(ii) Under $\nu$ the $\{\xi_i\}$ are exchangeable, so by the argument of Theorem 3 independent and identically distributed. Let $G(x) = \nu\{q \mid \xi_0(q) \geqq x\}$. Since $G(0) = 1$, (18) implies $G(x)G(y) = G(x + y)$; since $G$ is monotone decreasing to 0, and $G(0-) = 1$, $G(x) = e^{-\lambda x}$ for some positive finite $\lambda$, which proves (ii).

(iii) Since $(p, A) \rightarrow \mu_p(A)$ is a regular conditional probability given $\mathcal{I}$ under any invariant $\mu$, (iii) follows from (i) and (ii).

## REFERENCES

[1] Bühlmann, Hans (1960). Austauschbare Stochastische Variabeln und ihre Grenzwertsaetze. *Univ. California Publ. Statist.* **3** 1–36.
[2] Chung, Kai Lai (1960). *Markov Chains With Stationary Transition Probabilities.* Springer, Berlin.
[3] Doob, J. L. (1953). *Stochastic Processes.* Wiley, New York.
[4] Freedman, David A. (1962). Invariants under mixing which generalize De Finetti's theorem. *Ann. Math. Statist.* **33** 916–923.
[5] Halmos, Paul R. (1958). *Measure Theory.* Van Nostrand, Princeton.
[6] Kriloff, N. and Bogoliouboff, N. (1937). La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non-linéaire, *Ann. Math. Second Series* **38** 65–113.
[7] Lehmann, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.
[8] Loève, Michel (1960). *Probability Theory,* 2nd ed. Van Nostrand, Princeton.
[9] Loomis, Lynn H. (1953). *An Introduction to Abstract Harmonic Analysis.* Van Nostrand, New York.
[10] Lundberg, O. (1940). On random processes and their application to sickness and accident statistics. Thesis, Stockholm.
[11] Oxtoby, J. C. (1952). Ergodic Sets. *Bull. Amer. Math. Soc.* **58** 116–136.