The Annals of Probability
1977, Vol. 5, No. 6, 899-912

STOCHASTIC INEQUALITIES ON PARTIALLY ORDERED SPACES

BY T. KAMAE, U. KRENGEL AND G. L. O'BRIEN[1]

Osaka City University, University of Gottingen and York University

In this paper we discuss characterizations, basic properties and applications of a partial ordering, in the set of probabilities on a partially ordered Polish space E, defined by $P_1 \prec P_2$ iff $\int f\,dP_1 \leqq \int f\,dP_2$ for all real bounded increasing f. A result of Strassen implies that $P_1 \prec P_2$ is equivalent to the existence of E-valued random variables $X_1 \leqq X_2$ with distributions P_1 and P_2. After treating similar characterizations we relate the convergence properties of $P_1 \prec P_2 \prec \cdots$ to those of the associated $X_1 \leqq X_2 \leqq \cdots$. The principal purpose of the paper is to apply the basic characterization to the problem of comparison of stochastic processes and to the question of the computation of the $\bar{d}$-distance (defined by Ornstein) of stationary processes. In particular we get a generalization of the comparison theorem of O'Brien to vector-valued processes. The method also allows us to treat processes with continuous time parameter and with paths in $D[0, 1]$.

1. Characterizations of the stochastic partial ordering for probabilities. We shall consider the class $\mathscr{M}(E)$ of probability measures on a partially ordered Polish space E, i.e., a complete separable metric space E which is assumed throughout to be endowed with a closed partial ordering ($\leqq$) and the σ-algebra $\mathscr{F}$ generated by the open sets. All subsets of E and functions on E are taken without explicit mention to be $\mathscr{F}$-measurable. Let $\mathscr{I}^*(E)$ denote the set of bounded increasing (i.e., $x \leqq y \to f(x) \leqq f(y)$) real-valued functions and $\mathscr{I}(E)$ the family of "increasing" sets $A \subset E$, i.e., sets A for which the indicator function is increasing. Equivalently $A \in \mathscr{I}(E)$ iff $x \in A$ and $x \leqq y$ together imply $y \in A$. See Nachbin (1965) for general information on ordered topological spaces.

We say that $P_1 \in \mathscr{M}(E)$ is stochastically smaller than $P_2 \in \mathscr{M}(E)$, and denote this by $P_1 \prec P_2$, iff $\int f\,dP_1 \leqq \int f\,dP_2$ for all $f \in \mathscr{I}^*(E)$. A simple approximation argument shows that this is equivalent to the requirement that $P_1(A) \leqq P_2(A)$ for all $A \in \mathscr{I}(E)$. Clearly in this case the inequality $\int f\,dP_1 \leqq \int f\,dP_2$ holds even for all increasing f for which the integrals are well defined.

If E_1 and E_2 are Polish spaces with σ-algebras $\mathscr{F}_1$ and $\mathscr{F}_2$ respectively, a stochastic kernel in $E_1 \times E_2$ is a function k: $E_1 \times \mathscr{F}_2 \to [0, 1]$ such that $k(\cdot, A)$ is measurable for each $A \in \mathscr{F}_2$ and $k(x, \cdot) \in \mathscr{M}(E_2)$ for each $x \in E_1$. If k is such a kernel and $P_1 \in \mathscr{M}(E_1)$, then $P_1 * k$ is the element of $\mathscr{M}(E_1 \times E_2)$ determined by

$$(P_1 * k)(A_1 \times A_2) = \int_{A_1} k(x, A_2)P_1(dx) .$$

Received May 10, 1976; revised November 1, 1976.

[1] Research of this author supported in part by the National Research Council of Canada.

AMS 1970 *subject classifications*. Primary 60B99; Secondary 60G99, 60G10.

Key words and phrases. Stochastic comparison, measures on partially ordered spaces, stationary processes, $\bar{d}$-distance.

We denote the second marginal distribution of $P_1 * k$ by P_1^k. A stochastic kernel k in $E_1 \times E_2$ is said to be stochastically monotonic if $k(x, \cdot) \prec k(y, \cdot)$ for all $x \leqq y$. A stochastic kernel in $E \times E$ is called "upward" if for all x, $k(x, \cdot)$ is a measure with support in $\{y \in E : y \geqq x\}$.

The following theorem, much of which is known, will be basic to all that follows:

THEOREM 1. *The following conditions are equivalent for* $P_1, P_2 \in \mathscr{M}(E)$:

(i) $P_1 \prec P_2$;

(ii) *there exists a* $\lambda \in \mathscr{M}(E \times E)$ *with support in* $K = \{(x, y) \in E \times E : x \leqq y\}$, *with first marginal* P_1 *and with second marginal* P_2;

(iii) *there exists a real-valued random variable* Z *and two measurable functions* f *and* g: $\mathfrak{R}^1 \to E$ *with* $f \leqq g$ *such that the distribution of* $f(Z)$ *is* P_1 *and that of* $g(Z)$ *is* P_2;

(iv) *there exist two E-valued random variables* X_1, X_2 *such that* $X_1 \leqq X_2$ a.s. *and the distribution of* X_i *is* P_i $(i = 1, 2)$.

(v) *there exists an upward kernel* k *on* $E \times E$ *such that* $P_2 = P_1^k$;

(vi) $P_1(B) \leqq P_2(B)$ *for all closed* $B \in \mathscr{F}(E)$.

PROOF. The key implication is (i) $\Rightarrow$ (ii). It is a special case of Theorem 11 of Strassen (1965), obtained by taking his $\varepsilon = 0$ and his $\omega = K$.

Now assume (ii). The probability space (K, λ) is isomorphic mod 0 to $(B, \mathfrak{B}, P)$ where B is a Borel subset of $\mathfrak{R}^1$, $\mathfrak{B}$ is the collection of Borel subsets of B, and P is a probability measure on $(B, \mathfrak{B})$. This was shown by Rohlin (1952). See also page 14 of Parthasarathy (1967). Let $Z: K \to B$ be the isomorphism and let $f = p_1(Z^{-1})$ and $g = p_2(Z^{-1})$, where p_1 and p_2 are the projections from $E \times E$ onto the two factors. Then (iii) holds.

Clearly (iv) follows from (iii) (and of course also from (ii), if one prefers not to use Rohlin's result).

To get (v) from (iv) note that E is Polish and hence there exists a regular conditional probability distribution $k(x, B) = P\{X_2 \in B \mid X_1 = x\}$. As the distribution of (X_1, X_2) is concentrated on K we can find a version of k which is upward by modifying the original k on a P_1-null set.

Now assume (v). Let $B \in \mathscr{F}(E)$ be closed. If $x \in B$, then $k(x, B) = 1$. Thus

$$P_2(B) = \int_E k(x, B) P_1(dx) \geqq \int_B k(x, B) P_1(dx) = P_1(B) ,$$

so that (vi) holds.

To complete the proof we need only show that (vi) implies (i). Assume (vi), let $\varepsilon > 0$ and let $A \in \mathscr{F}(E)$. There exists a compact subset K_ε of A such that $P_i(A) - P_i(K_\varepsilon) < \varepsilon$ $(i = 1, 2)$. Let $H = \{y \in E : x \leqq y$ for some $x \in K_\varepsilon\}$. By Proposition 4 on page 44 of Nachbin (1965), H is a closed set in $\mathscr{F}(E)$. This may also be checked directly. Therefore $P_1(H) \leqq P_2(H)$. Since $K_\varepsilon \subseteqq H \subseteqq A$,

$$P_1(A) \leqq P_1(H) + \varepsilon \leqq P_2(H) + \varepsilon \leqq P_2(A) + \varepsilon .$$

Since ε is arbitrary $P_1(A) \leqq P_2(A)$, proving (i). □

REMARKS. The main import of the theorem is that (i) implies any of (ii)—(v). In the case $E = \Re^1$ this is very easy to see using distribution functions.

Nawrotski (1962) treats $E = \mathbb{Z}^2$ (with $(x_1, x_2) \leqq (y_1, y_2) \Leftrightarrow x_1 \leqq y_1$ and $x_2 \leqq y_2$) and his proof can be adapted to arbitrary countable E. His proof is still of interest since it is constructive in nature. If E is finite, λ on $E \times E$ is found with finitely many operations. Other proofs of (i) $\Rightarrow$ (ii) with various degrees of generality were given by Meyer (1966, page 246) and Preston (1974) and in unpublished notes of Snijders (1976) and Major (1975), who uses a continuous form of the marriage lemma, the König–Egerváry theorem, to get a result containing that of Strassen and a theorem of Dudley (1968) on the Prokhorov metric; for this method see also Lovász and Major (1973).

We thank P. Major, L. Lovász, H. G. Kellerer, R. Pyke and W. Vervaat for pointing out these references.

Nachbin (1951) has proved that any finite signed Radon measure μ on a compact partially ordered space with the property that $\int f\,d\mu \geqq 0$ for all continuous increasing f has the form

$$\int_E f\,d\mu = \int_K \{f(y) - f(x)\}\nu(dx, dy)$$

where ν is a nonnegative measure, and Hommel (1973) has a generalization of this to locally compact spaces. In Polish spaces their results follow easily from Theorem 1. Nachbin also showed that ν satisfies $||\nu|| = \frac{1}{2}||\mu||$. Nachbin's results can be used to obtain Theorem 1 (i) $\Rightarrow$ (ii) for the case of compact ordered spaces, and hence also of uniformizable ordered spaces (cf. page 104 of Nachbin (1965)).

We remark that $\mathscr{M}(E)$ is not in general a lattice under $\prec$ even if E is a lattice. For example, let $E = \{a = (0, 0), b = (0, 1), c = (1, 0), d = (1, 1)\}$ with $a \leqq b \leqq d$, $a \leqq c \leqq d$. Let ε_a denote the measure with unit mass at a, etc. Then $0.5(\varepsilon_a + \varepsilon_b)$ and $0.5(\varepsilon_a + \varepsilon_c)$ are both maximal among those measures which are stochastically smaller than both $0.5(\varepsilon_a + \varepsilon_d)$ and $0.5(\varepsilon_c + \varepsilon_b)$.

Since in $\Re^1$ the stochastic inequality $P \prec Q$ between two measures is equivalent to the statement that the distribution function of P is $\geqq$ the distribution function of Q, it is tempting to expect this also in $\Re^2$. However, the partial order defined by inequality between the distribution functions is in $\Re^n$ $(n \geqq 2)$ not equivalent to the concept considered in this paper.

2. Stochastically ordered probabilities on product spaces. We now consider products $E^n = E_1 \times E_2 \times \cdots \times E_n$ and $E^\infty = \prod_{i=1}^\infty E_i$ of partially ordered Polish spaces E_i with the product topology and the coordinate-wise partial ordering. Their elements will be denoted by $x^n = (x_1, \cdots, x_n)$ and $x^\infty = (x_1, x_2, \cdots)$ respectively, i.e., we have $x^n \leqq y^n$ iff $x_i \leqq y_i$ for all i. E^n and E^∞ are again partially ordered Polish spaces (see Billingsley (1968), page 218).

PROPOSITION 1. *Let* $E_1, \cdots, E_n$ *be partially ordered Polish spaces,* $P_1, Q_1 \in \mathscr{M}(E_1)$, $P_1 \prec Q_1$, *and let* p_i, q_i $(i = 2, \cdots, n)$ *be stochastic kernels on* $E^{i-1} \times E_i$

such that

$$p_i(x^{i-1}, \cdot) \prec q_i(y^{i-1}, \cdot) \tag{1}$$

whenever $x^{i-1} \leqq y^{i-1}$. *Then*

$$P_1 * p_2 * \cdots * p_n \prec Q_1 * q_2 * \cdots * q_n . \tag{2}$$

PROOF. Let $n = 2$. By Theorem 1(v) there exists an upward kernel k on $E_1 \times E_1$ with $Q_1 = P_1^k$. Let $A \in \mathscr{I}(E^2)$ and let

$$f(x_1) = \int_{E_2} 1_A(x_1, x_2) p_2(x_1, dx_2) ; \qquad g(y_1) = \int_{E_2} 1_A(y_1, y_2) q_2(y_1, dy_2) .$$

If $x_1 \leqq y_1$ then $f(x_1) \leqq g(y_1)$ since $A \in \mathscr{I}(E^2)$ and (1) holds. Hence

$$\begin{aligned} \int_{E_1} f(x_1) P_1(dx_1) &= \int_{E_1} \int_{\{y_1 \geqq x_1\}} f(x_1) k(x_1, dy_1) P_1(dx_1) \\ &= \int_{E_1} \int_{\{y_1 \geqq x_1\}} g(y_1) k(x_1, dy_1) P_1(dx_1) \\ &= \int_{E_1} \int_{E_1} g(y_1) k(x_1, dy_1) P_1(dx_1) \\ &= \int_{E_1} g(y_1) Q_1(dy_1) \end{aligned}$$

and this means that $P_1 * p_1 \prec Q_1 * q_1$. Now apply induction. □

REMARKS. Proposition 1 is similar to the independently obtained Satz 4.1 of Franken and Kirstein (1977), who give the result in the case $E_1 = E_2 = \cdots = E_n = \mathfrak{R}^1$. Note that (2) implies $P_k \prec Q_k$, where P_k and Q_k are the kth marginals of the measures in (2). Thus, Proposition 1 generalizes the theorem of Kalmykov (1962). If p_i is stochastically monotonic, then (1) need only be assumed for $x^{i-1} = y^{i-1}$, since then, if $x^{i-1} \leqq y^{i-1}$, $p_i(x^{i-1}, \cdot) \prec p_i(y^{i-1}, \cdot) \prec q_i(y^{i-1}, \cdot)$, so that (1) still holds. Similarly we need only assume (1) for $x^{i-1} = y^{i-1}$ if q_i is stochastically monotonic.

PROPOSITION 2. *Let* $E_1, E_2, \cdots$ *be Polish spaces and let* $F = E^\infty = \prod_{i=1}^\infty E_i$. *Let* $P, Q \in \mathscr{M}(F)$ *and let* $P^{(i)}(Q^{(i)})$ *be the marginal distribution of the first* i *coordinates of* P *(respectively* Q*), so that* $P^{(i)}, Q^{(i)} \in \mathscr{M}(E^i)$. *If*

$$P^{(i)} \prec Q^{(i)} , \qquad i = 1, 2, \cdots , \tag{3}$$

then $P \prec Q$.

PROOF. Let $(z_1, z_2, \cdots)$ be a fixed element of F. By Theorem 1 (iv), there are, for each m, random m-vectors $(X_m^1, \cdots, X_m^m)$ and $(Y_m^1, \cdots, Y_m^m)$ in E^m with distributions $P^{(m)}$ and $Q^{(m)}$ respectively and such that $X_m^j \leqq Y_m^j$ for $j = 1, 2, \cdots, m$. We obtain from these vectors random elements X_m and $Y_m \in F$ by defining $X_m^j = Y_m^j = z_j$, $j = m+1, m+2, \cdots$. Let P_m and Q_m be the distributions of X_m and Y_m respectively. Since $X_m \leqq Y_m$ we have $P_m \prec Q_m$. For fixed j, the jth one-dimensional marginal of P_m is independent of m for $m \geqq j$; hence the sequence of such marginals is tight by Theorem 1.4 of Billingsley (1968). By Tychonov's theorem, it follows that the sequences $\{P_m\}$ and $\{Q_m\}$ are tight. Let U_m be the joint distribution of (X_m, Y_m) in $F \times F$. The sequence $\{U_m\}$ is also tight and hence has a subsequence $\{U_{m_i}, i = 1, 2, \cdots; m_1 < m_2 < \cdots\}$ which converges weakly to a probability measure $U \in \mathscr{M}(F \times F)$. The marginals (chosen appropriately) of U are $P^{(i)}$ and $Q^{(i)}$ and (hence) P and Q. Also, each U_m has support in the closed set $H \equiv \{(x, y) \in F \times F : x \leqq y\}$. By Theorem 2.1 of Billingsley (1968),

$$U(H) \geqq \limsup_{i\to\infty} U_{m_i}(H) = 1 ,$$

so that U has support in H. Thus $P \prec Q$. □

REMARK. Suppose that for $i = 1, 2, \cdots$ probability measures $P^{(i)}$ and $Q^{(i)}$ are given on $E_1 \times \cdots \times E_i$ in such a way that Kolmogorov's consistency conditions are met and such that (3) holds. Then there exist probability measures P and Q on E with $P \prec Q$ and marginals $P^{(i)}$ and $Q^{(i)}$ respectively.

PROPOSITION 3. *Let $\{P_m\}$ and $\{Q_m\}$, $m = 1, 2, \cdots$ be sequences of probability measures on a Polish space E and converging weakly to probability measures P and Q (respectively). If $P_m \prec Q_m$ for all m, then $P \prec Q$.*

PROOF. For each m, there is a probability measure λ_m on $E \times E$ with marginals P_m and Q_m and with $\lambda_m(\{(x, y) : x \leqq y\}) = 1$, by Theorem 1 (ii). The sequences $\{P_m\}$ and $\{Q_m\}$ are tight; hence $\{\lambda_m\}$ is also tight and has a subsequence which converges to some probability measure λ. The marginals of λ must be P and Q and $\lambda(\{(x, y) : x \leqq y\}) = 1$ as in Proposition 2. By Theorem 1, $P \prec Q$. □

REMARKS. The relation $\prec$ on $\mathscr{M}(E)$ is a closed partial ordering with the weak convergence topology. Sections 6, 7 and 8 (except for some remarks) can be read independently of Sections 3, 4 and 5.

3. The discrete-time comparison theorem. The following theorem is an extension of the main case (that is, the case related to Kalmykov's (1962) theorem) of the comparison theorem of O'Brien (1975) to random sequences in Polish spaces.

THEOREM 2. *Let $E_1, E_2, \cdots$ be Polish spaces, let $P_1, Q_1 \in \mathscr{M}(E_1)$, and let p_n, q_n be stochastic kernels on $E^{n-1} \times E_n$ for $n = 2, 3, \cdots$. Assume $P_1 \prec Q_1$ and*

$$p_n(x^{n-1}, \cdot) \prec q_n(y^{n-1}, \cdot) \tag{4}$$

whenever $(x_1, \cdots, x_{n-1}) \leqq (y_1, \cdots, y_{n-1})$. Then there exist random variables X_i and Y_i with values in E_i, $i = 1, 2, \cdots$, such that the distribution of $X_1(Y_1)$ is $P_1(Q_1)$, the conditional distribution of X_n given $X^{n-1} = x^{n-1}$ (Y_n given y^{n-1}) is $p_n(x^{n-1}, \cdot)(q_n(y^{n-1}, \cdot))$, and

$$P(X_i \leqq Y_i, i = 1, 2, \cdots) = 1 . \tag{5}$$

PROOF. Construct $P^{(n)} = P_1 * p_2 * \cdots * p_n$ and $Q^{(n)} = Q_1 * q_2 * \cdots * q_n$ as in Proposition 1, with $P^{(n)} \prec Q^{(n)}$. By Kolmogorov's consistency conditions, there exist probability measures P and Q on E^∞, where P and Q are related to $P^{(i)}$ and $Q^{(i)}$ as in Proposition 2. By that proposition $P \prec Q$. The result now follows from Theorem 1 (iv). □

Let $\{X_n, n = 1, 2, \cdots\}$ and $\{Y_n, n = 1, 2, \cdots\}$ be *any* random sequences in E^∞ with initial and conditional distributions as in Theorem 2. Then there exist random sequences $\{\tilde{X}_n\}$ and $\{\tilde{Y}_n\}$ which are defined on a common space and have the same distributions as $\{X_n\}$ and $\{Y_n\}$ respectively, such that

$$P(\tilde{X}_n \leqq \tilde{Y}_n, n = 1, 2, \cdots) = 1 . \tag{6}$$

This leads to a variety of inequalities between probabilities which are determined (separately) by the distributions of $\{X_n\}$ and $\{Y_n\}$. Some examples are given in

COROLLARY 1. *Let $\{X_n\}$ and $\{Y_n\}$ be as above and assume $E_1 = E_2 = \cdots$. Let $A \in \mathscr{I}(E_1)$.*

(i) *$P(X_n \in A) \leqq P(Y_n \in A)$.*

(ii) *Let $N_x = \min\{n \geqq 1 : X_n \in A\}$ ($= \infty$ if $X_n \notin A$ for all n) and define N_y similarly. Then $P(N_x < n) \leqq P(N_y < n)$, $n = 1, 2, \cdots$, and $n = \infty$.*

(iii) *Let $f\colon E_1 \to \mathfrak{R}^1$ be nondecreasing and measurable. Then $E(f(X_n)) \leqq E(f(Y_n))$ for all n for which the expectations exist.*

If the stochastic kernels p_n, $n = 2, 3, \cdots$ (or q_n, $n = 2, 3, \cdots$) of Theorem 2 are stochastically monotonic, it is enough to assume in the theorem that (4) holds for $x^{n-1} = y^{n-1}$ (see Proposition 1 and the subsequent remark). If $\{X_n\}$ is related to p_n as in the theorem, we also say $\{X_n\}$ is stochastically monotonic. This extends the concept introduced by Daley (1968).

Various other comparison-type theorems may be obtained from Theorem 2 by making suitable changes of variables. The following theorem on the comparison of increments is related to Example 7.2 in O'Brien (1975).

THEOREM 3. *Let E be a Polish space which also has a compatible vector space structure (that is, addition and scalar multiplication are continuous and, if $A \in \mathscr{I}(E)$ and $x \in E$, then $A + x = \{y + x : y \in A\} \in \mathscr{I}(E)$). Let P_1 and Q_1 be probability measures on E with $P_1 \prec Q_1$. Let p_n and q_n, $n = 2, 3, \cdots$, be stochastic kernels in $E^{n-1} \times E$ such that*

$$p_{n+1}(s_1, \cdots, s_n; s_n + A) \leqq q_{n+1}(t_1, \cdots, t_n; t_n + A) \tag{7}$$

for all $s^n, t^n \in E^n$ with $s_1 \leqq t_1$ and with (for $n > 1$) $s_{i+1} - s_i \leqq t_{i+1} - t_i$, $i = 1, 2, \cdots, n-1$, and all $A \in \mathscr{I}(E)$. Then there exist processes $\{S_n, n \geqq 1\}$ and $\{T_n, n \geqq 1\}$ on a common space $(\Omega, \mathscr{F}, P)$ such that the distributions of S_1 and T_1 are P_1 and Q_1 respectively, $P(S_{n+1} - S_n \in A \mid S_1, \cdots, S_n) = p_{n+1}(S_1, \cdots, S_n; A + S_n)$ and $P(T_{n+1} - T_n \in A \mid T_1, \cdots, T_n) = q_{n+1}(T_1, \cdots, T_n; A + S_n)$, $n = 1, 2, \cdots$, and $P[S_{n+1} - S_n \leqq T_{n+1} - T_n, n = 1, 2, \cdots] = 1$ and $P(S_1 \leqq T_1) = 1$.

PROOF. Define a sequence $\hat{p}_n$, $n = 2, 3, \cdots$ of stochastic kernels on $E^{n-1} \times E$ (respectively) by

$$\hat{p}_n(x^{n-1}, A) = p_n(x_1, x_1 + x_2, \cdots, x_1 + \cdots + x_{n-1}, x_1 + x_2 + \cdots + x_{n-1} + A) .$$

Define $\hat{q}_n$ similarly in terms of q_n. By the hypotheses, $\hat{p}_n(x^{n-1}, \cdot) \prec \hat{q}_n(y^{n-1}, \cdot)$

if $x^{n-1} \leqq y^{n-1}$. By Theorem 2, there are random sequences $\{X_n\}$ and $\{Y_n\}$, $n = 1, 2, \cdots$ with initial and conditional distributions P_1, $\hat{p}_n$, Q_1 and $\hat{q}_n$ respectively such that $P(X_n \leqq Y_n, n = 1, 2, \cdots) = 1$. Let $S_n = X_1 + \cdots + X_n$ and $T_n = Y_1 + \cdots + Y_n$. It is easily checked that the sequences $\{S_n\}$ and $\{T_n\}$ have the required properties. □

4. The continuous-time comparison theorem. We shall now derive an analogue of the above comparison theorem for processes with continuous time parameter. This result will be new even for real-valued processes. Let $[a, b] = I$ be a compact interval in $\mathfrak{R}^1$. If E is a partially ordered Polish space, so is the space $D_E(I)$ of functions from I to E which are right continuous and have left limits at all $t \in I$. Here the Skorohod metric is used and $x \leqq y$ if $x(t) \leqq y(t)$ for all $t \in I$. If I is open on the right (and possibly $b = \infty$), $D_E(I)$ is still Polish, where we use the Stone (1963) modification of Skorohod's metric. In either case the subclass $C_E(I)$ of continuous functions from I to E is also a Polish space with the induced topology and partial ordering. We refer occasionally to the book of Billingsley even though he treats only the case $D_{\mathfrak{R}^1}(I)$. However, the considerations for more general E are quite analogous.

We consider processes with paths in $D_E(I)$. We require some additional notation. Let $J = \{t \in I: t \text{ is rational}, t = a \text{ or } t = b\}$. In writing $t^n = (t_1, t_2, \cdots, t_n) \in I^n$, we assume throughout that $t_1 < t_2 < \cdots < t_n$. A collection $\{p_{t^n}, n = 1, 2, \cdots; t^n \in I^n : E^{n-1} \times \mathscr{F} \to [0, 1]\}$ of stochastic kernels is called a $D_E(I)$-family if there exists a stochastic process $\{X_t, t \in I\}$ with paths in $D_E(I)$ such that

$$p_{t^n}(x^{n-1}, B) = P(X_{t_n} \in B \mid X_{t^{n-1}} = x^{n-1}) \tag{8}$$

for all $n = 2, 3, \cdots$, $x^{n-1} \in E^{n-1}$ and $B \subseteqq E$.

The following lemma has points in common with Billingsley's discussion of separable processes (pages 134–136).

LEMMA 1. *Let E be a Polish space with metric m and let X be the set E^J of functions from J to E, endowed with the product topology. The set $A = \{x \in X$: there exists an $f \in D_E(I)$ such that $x(t) = f(t)$ for all $t \in J\}$ is a Borel subset of X.*

PROOF. The interval I is a countable union of compact intervals; it therefore suffices to consider the case $I = [a, b]$. Let d denote the Skorohod metric on $D_E(I)$. Let Λ be the class of increasing bijections of $[a, b]$ onto $[a, b]$. Then

$$d(f, g) = \inf\{\varepsilon > 0 : \exists\, \lambda \in \Lambda \text{ such that (9) and (10) hold}\} :$$

$$\sup_t |\lambda(t) - t| \leqq \varepsilon\,, \tag{9}$$

$$\sup_t m(f(t), g(\lambda(t))) \leqq \varepsilon\,. \tag{10}$$

Then $D_E(I)$ is separable and has a countable dense subset $\{f_1, f_2, \cdots\}$. Let $T_1 \subseteqq T_2 \subseteqq \cdots$ be finite sets whose union is J. The set

$$B = \{x \in X : x \text{ is right-continuous at all } t \in J\}$$

is a Borel subset of X. For each $\varepsilon > 0$, n, j, define the open set $U(\varepsilon, j, n) = \{x \in X : \exists \lambda$ satisfying (9) such that $m(x(t), f_j(\lambda(t))) < \varepsilon$ for all $t \in T_n\}$. Hence $B(\varepsilon, j) \equiv \bigcap_{n=1}^{\infty} U(\varepsilon, j, n)$ is a Borel set. It is clear that

$$A \subseteqq B \cap [\bigcap_{k=1}^{\infty} \bigcup_{j=1}^{\infty} B(k^{-1}, j)] \tag{11}$$

by the denseness of $\{f_1, f_2, \cdots\}$. We show that the reverse inclusion is also valid. Let x be an element of the right side of (11). Fix $r \in (a, b]$. We show x has a left limit at r (right limits are similar). Let $\varepsilon > 0$ and let $s_1 < s_2 < \cdots$ be elements of J which converge to r. Fix $k > \varepsilon^{-1}$ and fix j such that $x \in B(k^{-1}, j)$. For all n, there is an N_n such that $\{s_1, s_2, \cdots, s_n\} \subseteqq T_{N_n}$ and hence there is a $\lambda_n \in \Lambda$ satisfying (9) and $m(x(s_i), f_j(\lambda_n(s_i))) < \varepsilon$, $i = 1, 2, \cdots, n$. By Lemma 1 (page 110) of Billingsley (1968), $m(f_j(\lambda_n(s_i)), f_j(\lambda_n(s_{i+1}))) < \varepsilon$ at all but a set of l values of i, where l is determined by j and ε. Hence $m(x(s_i), x(s_{i+1})) < 3\varepsilon$ at all but finitely many i. Since ε and the sequence $\{s_1, s_2, \cdots\}$ are arbitrary, $x(t)$ must have a limit as $t \uparrow r$.

THEOREM 4. *Let P_a and $Q_a \in \mathscr{M}(E)$ with $P_a \prec Q_a$. Let $\{p_{t^n}\}$ and $\{q_{t^n}\}$ be $D_E(I)$-families such that*

$$p_{t^n}(x^{n-1}, \cdot) \prec q_{t^n}(y^{n-1}, \cdot) \qquad \textit{whenever} \quad x^{n-1} \leqq y^{n-1} . \tag{12}$$

Then there exist stochastic processes $\{X_t, t \in I\}$ and $\{Y_t, t \in I\}$ on a common probability space, with paths in $D_E(I)$, with distributions at time a given by P_a and Q_a respectively, and with conditional probabilities of $\{X_t\}$ given by (8) *and those of Y_t given by* (8) *with p replaced by q, such that*

$$P(X_t \leqq Y_t \ \textit{for all} \ t \in I) = 1 . \tag{13}$$

PROOF. Let $\{U_t, t \in I\}$ be a stochastic process in $D_E(I)$ such that (8) holds with X replaced by U. Let $P^{(t^n)}$ be the joint distribution of $(U_t, \cdots, U_{t_n})$ and define $Q^{(t^n)}$ similarly. By considering any enumeration $\{t_1 = a, t_2, t_3, \cdots\}$ of the countable set J, we may construct processes $\{X_t, t \in J\}$ and $\{Y_t, t \in J\}$ on the same space $(\Omega, \mathscr{F}, P)$ and with suitable initial and conditional distributions such that $P(X_t \leqq Y_t$ for all $t \in J) = 1$, exactly as in the discrete time case (Theorem 2). Define the set $A = \{\omega \in \Omega : X_t(\omega)$ can be extended to a path in $D_E(I)\}$. By Lemma 1, A is a measurable set. Thus $P(A)$ is defined and is determined by P_a and p_{t^n}. Since almost all paths of U_t are in $D_E(I)$, we conclude that $P(A) = 1$. Performing this extension, we see that (8) remains valid. Extending Y_t almost surely to $D_E(I)$ by the same means, we have (13). □

REMARK. The corollary to Theorem 2 has an obvious analogue in the continuous time case.

5. Markov processes. The theorems of Sections 3 and 4 are more simply expressed in the case of Markov processes with stationary transition probabilities. We restrict attention to the continuous-time case since the discrete-time situation is similar. Let E be a Polish space and let $I = [a, b]$ or $[a, b)$ be an interval in $\mathfrak{R}^1$. A collection $\{p_t: E \times \mathscr{F} \to [0, 1], 0 \leqq t \leqq b - a\}$ of stochastic kernels on $E \times E$ is called a *Markov* $D_E(I)$ *family of transition functions* if there exists a Markov process $\{X_t, t \in I\}$ with paths in $D_E(I)$ such that

$$p_t(x, B) = P(X_{t+s} \in B \mid X_s = x) \tag{14}$$

for all $x \in E$, $B \subseteq E$, and $t, s \in \mathfrak{R}^1$ such that $a \leqq s \leqq t + s \leqq b$. Theorem 4 now reduces to

THEOREM 5. *Let P_a and $Q_a \in \mathscr{M}(E)$ with $P_a \prec Q_a$. Let $\{p_t\}$ and $\{q_t\}$ be Markov $D_E(I)$ families such that*

$$p_t(x, \cdot) \prec q_t(y, \cdot) \qquad \textit{whenever} \quad x \leqq y\,. \tag{15}$$

Then there exist Markov processes $\{X_t\}$ and $\{Y_t\}$ on the same space with paths in $D_E(I)$, with initial distributions P_a and Q_a respectively and with (14) *and the similar statement for Y and q holding, such that*

$$P(X_t \leqq Y_t, \textit{ for all } t \in I) = 1\,. \tag{16}$$

If either p_t or q_t is stochastically monotonic for all t, then we need only assume (15) with $x = y$. As remarked by Keilson and Kester (1977), for the case $E = \mathfrak{R}^1$ the class of continuous-time processes with stochastically monotonic transition functions includes all diffusions processes, even in $\mathfrak{R}^n$. Keilson and Kester (1977), Franken and Kirstein (1977), O'Brien (1977), and Sonderman and Whitt (1977) all show under various circumstances that for pure jump processes, the condition (15) may be expressed equivalently as an inequality between the two corresponding transition rate functions. In the case of diffusions, Anderson (1972) has obtained pathwise inequalities from inequalities between the corresponding infinitesimal generators, in the case when the variance term is the same for both generators and is independent of the location.

6. Stochastically increasing sequences of probability measures. We begin with a generalization of Theorem 1 to sequences of random variables.

PROPOSITION 4. *Let E be a Polish space and $\{P_1, P_2, \cdots\}$ a sequence of probability measures on E. The following are equivalent.*

(i) $P_1 \prec P_2 \prec P_3 \prec \cdots$;

(ii) *there exist random elements $X_1, X_2, \cdots$ in E defined on a common space in such that $X_1 \leqq X_2 \leqq \cdots$* a.s., *and such that the distribution of X_i is P_i, $i = 1, 2, \cdots$; and*

(iii) *there exists a real-valued random variable Z and a sequence of functions $f_i: \mathfrak{R}^1 \to E$ with $f_1 \leqq f_2 \leqq \cdots$ such that P_i is the distribution of $f_i(Z)$, $i = 1, 2, \cdots$.*

PROOF. It is clear that (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). Now assume (i). By Theorem 1 (v), there exist stochastic kernels $k_1, k_2, \cdots$ on $E \times E$ such that $P_{i+1} = P_i^{k_i}$ where $k_i(x, \cdot)$ has support in $\{y \in E: x \leqq y\}$, $i = 1, 2, \cdots$. Therefore a probability measure P exists on E^∞ such that the ith marginal of P is P_i and P is concentrated on $\{x^\infty \in E^\infty : x_1 \leqq x_2 \leqq \cdots\}$. To get (ii), take $X_i : E^\infty \to E$ to be the ith projection mapping. By Rohlin (1952), there is an isomorphism $\Psi : (E^\infty, P) \to (\mathfrak{R}^1, P^1)$ for some probability measure P^1 on $\mathfrak{R}^1$. To obtain (iii), take $Z = \Psi : E^\infty \to \mathfrak{R}^1$ and $f_i = X_i \circ \Psi^{-1}$. □

REMARKS. The sequence $\{X_i\}$ of condition (ii) is of course a submartingale sequence. Moreover, if any of (i) to (iii) holds, then (ii) holds for a Markov sequence $\{X_i\}$ (with nonstationary transition probabilities), by the proof of (i) → (ii). The question of existence of an increasing process $\{X_t, t \in \mathfrak{R}^1\}$ for a given stochastically increasing family $\{P_t, t \in \mathfrak{R}^1\}$ will be considered in a forthcoming paper of Kamae and Krengel (1977). In the continuous time case, the answer depends on the nature of E. The next theorem shows that various modes of convergence are equivalent under an assumption of monotonicity.

THEOREM 6. *Let E be a Polish space with metric d, let $X_1, X_2, \cdots$ be random variables from $(\Omega, \mathscr{F}, P)$ to E with $X_1 \leqq X_2 \leqq \cdots$, and let $P_1, P_2, \cdots$ be their distributions. Then the following conditions are equivalent.*

(i) *The family $\{P_i, i \geqq 1\}$ is tight;*
(ii) *$\{P_i, i \geqq 1\}$ converges weakly to a probability measure; and*
(iii) *$\{X_i, i \geqq 1\}$ converges in probability.*

The condition

(iv) *$\{X_i, i \geqq 1\}$ converges almost surely*

is also equivalent for all nondecreasing sequences $\{X_i\}$ if and only if E satisfies the following regularity conditions:

(17) *for all sequences (x_n) and (y_n) of elements of E with* $x_n \leqq y_n \leqq x_{n+1} (n \geqq 1)$, $x_n \to x \in E$ *implies* $y_n \to x$.

PROOF. Clearly (iv) ⇒ (iii) ⇒ (ii) ⇒ (i). Now assume (i) and let $\varepsilon > 0$. There is a compact set K such that

$$P(X_i \in K) = P_i(K) \geqq 1 - \varepsilon, \qquad i \geqq 1.$$

Let $A = \bigcap_{n=1}^\infty \bigcup_{i=n}^\infty \{\omega \in \Omega : X_i(\omega) \in K\}$. Then $P(A) \geqq 1 - \varepsilon$. Fix $\omega \in A$ and let $S(\omega)$ be the (infinite) sequence of n's for which $X_n(\omega) \in K$. Every subsequence of $\{X_n(\omega), n \in S(\omega)\}$ contains a convergent subsequence. If x and y are two limit points, then $X_n(\omega) \leqq x$ for all $n \in S(\omega)$ and hence $y \leqq x$ since the partial ordering is closed. Similarly $x \leqq y$, so that $\{X_n(\omega), n \in S(\omega)\}$ converges to a point $X(\omega) \in E$. There is a positive integer $N(\omega)$ such that for $n \geqq N(\omega)$ and $n \in S(\omega)$, $d(X(\omega), X_n(\omega)) < \varepsilon$. There exists a positive integer N, independent of ω, such that $N \geqq N(\omega)$ for all $\omega \in A$ outside a set of probability $< \varepsilon$. For $n \geqq N$,

$$P(d(X, X_n) \geqq \varepsilon) \leqq \varepsilon + 1 - P(X_n \in K) < 2\varepsilon.$$

Since ε is arbitrary, we have proved (iii).

We now assume (iii) and (17). There is an increasing sequence $\{n_i\}$ of positive integers such that $X_{n_i} \to X$ a.s. Let $\{m_i\}$ be another increasing sequence of positive integers. We may assume, by deleting terms of one or both sequences if necessary that $n_i < m_i < n_{i+1}$ for all i; hence, $X_{n_i} \leqq X_{m_i} \leqq X_{n_{i+1}}$ a.s. and hence $X_{m_i} \to X$ a.s. by (17). Thus $X_n \to X$ a.s.

Now assume that (17) fails. Take any probability space $(\Omega, \mathscr{F}, P)$ and let $C_i \in \mathscr{F}$ be such that $P(C_i) \to 0$ and $P(\{\omega \in \Omega : \omega \in \text{infinitely many } C_i\}) = 1$. Let x_n, y_n be points in E for which (17) fails and assume without loss of generality that the distance from y_n to x exceeds $\varepsilon > 0$ for all n. Let $X_n(\omega) = y_n$ if $\omega \in C_n$, $= x_n$ otherwise. Then $X_n \to x$ in probability but not a.s. □

REMARKS. A Polish space for which (17) fails is given by the following pathological example. Let

$$E = \left\{x_n = \left(1 - \frac{1}{n}, 0\right); n = 1, 2, \cdots\right\} \cup \{z = (1, 0)\}$$
$$\cup \{y_n = (n, 1); n = 1, 2, \cdots\}$$

with the metric induced by that of $\mathfrak{R}^2$ and the linear ordering $x_n < y_n < x_{n+1} < z$ for all n. Condition (17) is clearly implied by local convexity as defined by Nachbin (1965, page 100). The following example shows that the reverse implication is not valid for Polish spaces. Let E and the metric be as above and define the partial ordering by $a \leqq b$ if $a = x_n$ and $b = y_n$ for the same n or if $b = z$. An application of Theorem 6 for the proof of a new pointwise ergodic theorem will be given in a forthcoming paper of Krengel (1977).

7. Stationary sequences. Given two stationary random sequences, we can obtain pathwise inequalities between them, in the sense of Theorem 2, while at the same time obtaining stationarity of the joint process.

THEOREM 7. *Let E be a Polish space and let P, Q be shift-invariant probability measures on $\Omega = E^{\mathbb{Z}}$ with $P \prec Q$. Then there exists a shift-invariant probability measure ν on $\Omega \times \Omega$ with marginals P and Q and with support in $H = \{(\omega_1, \omega_2) \in \Omega \times \Omega : \omega_1 \leqq \omega_2\}$.*

PROOF. By Theorem 1 (ii), there exists a measure μ on $\Omega \times \Omega$ with marginals P and Q and with support in H. Let T be the shift operator on Ω and let S be the shift operator on $\Omega \times \Omega$; that is, $S(\omega_1, \omega_2) = (T\omega_1, T\omega_2)$. Define a sequence $\{\mu_n\}$ of measures on $\Omega \times \Omega$ by

$$\mu_n = n^{-1} \sum_{k=0}^{n-1} \mu \circ S^{-k} .$$

Since H is invariant under S, each μ_n is concentrated on H. Moreover, each μ_n has P and Q as its marginals in Ω and consequently the sequence $\{\mu_n\}$ is tight. There is a subsequence with a weak limit ν, which is a probability measure also with marginals P and Q. Since each μ_n is concentrated on the closed set H, so is ν. Finally ν is invariant under S since

$$||\mu_n - \mu_n \circ S|| \leqq 2n^{-1} . \qquad \square$$

8. An application to Ornstein's $\bar{d}$-distance. We shall now apply Theorem 7 to derive a formula for the $\bar{d}$-distance as defined by Ornstein (1973) of two real-valued stationary processes. Let again T be the shift in $\Omega = \mathfrak{R}^{\mathbb{Z}}$, S the shift in $\Omega \times \Omega$. Let $\mathscr{J}_T$ be the set of T-invariant probability measures on Ω and $\mathscr{J}_S$ the set of S-invariant probability measures on $\Omega \times \Omega$. For $\nu \in \mathscr{J}_S$ let $\nu_i \in \mathscr{J}_T$ $(i = 1, 2)$ be the marginals in the first, resp. second factor of $\Omega \times \Omega$. We write the elements of $\Omega \times \Omega$ in the form $\omega = (\omega_1, \omega_2)$ where $\omega_i \in \Omega$ $(i = 1, 2)$ has coordinates $\omega_i^n \in \mathfrak{R}$ $(n \in \mathbb{Z})$.

Consider $P, Q \in \mathscr{J}_T$ such that $\int_\Omega \omega_1^0 P(d\omega_1)$ and $\int_\Omega \omega_1^0 Q(d\omega_1)$ are defined and finite. The $\bar{d}$-distance of two processes with P and Q as their distributions can be defined by the formula

$$\bar{d}(P, Q) = \inf\{\int |\omega_1^0 - \omega_2^0|\nu(d\omega) : \nu \in \mathscr{J}_S, \nu_1 = P, \nu_2 = Q\}.$$

(This is formally slightly different from Ornstein's definition, since we do not want to introduce some of the concepts which play a role in Ornstein's work, but it is easily seen to be equivalent.)

LEMMA 2. $\bar{d}(P, Q) \geqq |\int \omega_1^0 P(d\omega_1) - \int \omega_1^0 Q(d\omega_1)|$.

PROOF. Obvious.

LEMMA 3. *If $P, Q \in \mathscr{J}_T$ and $P \prec Q$ then*

$$\bar{d}(P, Q) = \int \omega_1^0 Q(d\omega_1) - \int \omega_1^0 P(d\omega_1).$$

PROOF. Take $\nu \in \mathscr{J}_S$ with $\nu_1 = P$, $\nu_2 = Q$ and support in $H = \{(\omega_1 \omega_2) : \omega_1 \leqq \omega_2\}$. Then

$$\begin{aligned}\int_\Omega \omega_1^0\, dQ - \int_\Omega \omega_1^0\, dP &= \int_{\Omega\times\Omega} (\omega_2^0 - \omega_1^0)\nu(d\omega) = \int_H (\omega_2^0 - \omega_1^0)\nu(d\omega) \\ &= \int_H |\omega_2^0 - \omega_1^0|\nu(d\omega) = \int_{\Omega\times\Omega} |\omega_2^0 - \omega_1^0|\nu(d\omega).\end{aligned}$$

Now apply Lemma 2. □

THEOREM 8. *For any $P, Q \in \mathscr{J}_T$ we have*

$$\begin{aligned}\bar{d}(P, Q) = &\int \omega_1^0 P(d\omega_1) + \int \omega_1^0 Q(d\omega_1) \\ &- 2 \sup\{\int \omega_1^0 R(d\omega_1) : R \in \mathscr{J}_T, R \prec P, R \prec Q\}.\end{aligned} \tag{18}$$

PROOF. For any $R \in \mathscr{J}_T$ such that $R \prec P$ and $R \prec Q$ we have

$$\begin{aligned}\bar{d}(P, Q) &\leqq \bar{d}(P, R) + \bar{d}(R, Q) \quad \text{(since } \bar{d} \text{ is a distance)} \\ &= \int \omega_1^0 P(d\omega_1) + \int \omega_1^0 Q(dw_1) - 2 \int \omega_1^0 R(d\omega_1) \quad \text{(by Lemma 3)},\end{aligned}$$

so that $\leqq$ in (18) follows. To prove the opposite inequality we introduce a function $\varphi : \Omega \times \Omega \to \Omega$ by defining the nth coordinate of $\varphi(\omega_1, \omega_2)$ to be $\varphi(\omega_1, \omega_2)^n = \omega_1^n \wedge \omega_2^n$ $(n \in \mathbb{Z})$.

The mapping $\nu \to \int |\omega_1^0 - \omega_2^0|\nu(d\omega)$ from $\mathscr{M}(\Omega \times \Omega)$ into $[0, \infty]$ is lower semi-continuous, as shown by Billingsley (1968, Theorem 5.3) so that the infimum in the definition of $\bar{d}(P, Q)$ is assumed; that is, there exists a ν in the class

$\{\nu \in \mathscr{J}_S : \nu_1 = P, \nu_2 = Q\}$ with

$$\bar{d}(P, Q) = \int |\omega_1^0 - \omega_2^0| \nu(d\omega) .$$

Let $R' = \nu \circ \varphi^{-1}$, then $R' \in \mathscr{J}_T$. If $A \in \mathscr{I}(\Omega)$, then $\varphi^{-1}(A) \subseteqq A \times A$. Hence $R'(A) = \nu(\varphi^{-1}(A)) \leqq \nu(A \times \Omega) = P(A)$. Thus $R' \prec P$. Similarly $R' \prec Q$. On the other hand

$$\begin{aligned} \bar{d}(P, Q) &= \int |\omega_1^0 - \omega_2^0| \nu(d\omega) = \int (\omega_1^0 + \omega_2^0 - 2\varphi(\omega_1, \omega_2)^0)\nu(d\omega) \\ &= \int \omega_1^0 P(d\omega_1) + \int \omega_1^0 Q(d\omega_1) - 2 \int \omega_1^0 R'(d\omega_1) . \end{aligned}$$

Thus we have proved $\geqq$ as well. □

The point of the theorem is that the computation of an infimum over a class of distributions in $\Omega \times \Omega$ is replaced by a supremum involving only Ω. If $P \prec Q$, the theorem (or Lemma 3) gives an exact formula for $\bar{d}(P, Q)$.

A large class of examples with $P \prec Q$ can be obtained by considering pairs of irreducible positive recurrent Markov chains with one-dimensional distributions taken to be their stationary distributions. By Theorem 2, we need only verify for such chains that the stationary distributions satisfy $P(\omega_1^0 > z) \leqq Q(\omega_1^0 > z)$ for all z and the transition functions satisfy

$$P(\omega_1^1 > z \mid \omega_1^0 = x) \leqq Q(\omega_1^1 > z \mid \omega_1^0 = y)$$

for all z whenever $x \leqq y$. It is easy to construct pairs of processes with these properties (from the class of birth-death processes, for example). The assumption $P \prec Q$ in these examples is important in that a formula for the $\bar{d}$-distance has not yet been written down for the case of arbitrary stationary Markov measures P and Q, even on $\{0, 1\}^Z$. Ellis (1975) showed that even in this case, the infimum in the definition of $\bar{d}$-distance is not necessarily attained by such ν's that are stationary Markov measures on $\{0, 1\}^Z \times \{0, 1\}^Z$.

Acknowledgment. The authors wish to thank the referee for a number of useful comments, including the comment that Theorem 1 (i) $\Rightarrow$ (ii) can be deduced from the results of Nachbin in the compact case.

REFERENCES

[1] Anderson, W. J. (1972). Local behavior of solutions of stochastic differential equations. *Trans. Amer. Math. Soc.* **164** 309–321.

[2] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[3] Daley, D. J. (1968). Stochastically monotone Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **10** 307–317.

[4] Dudley, R. M. (1968). Distances of probability measures and random variables. *Ann. Math. Statist.* **39** 1563–1572.

[5] Ellis, M. (1975). Achieving the $\bar{d}$-distance between two one-step two-state mixing Markov processes. Ph.D. dissertation, Stanford Univ.

[6] Franken, P. and Kirstein, B. M. (1977). Zur Vergleichbarkeit zufälliger Prozesses. *Math. Nachr.* To appear.

[7] Hommel, G. (1973). Isotone Radonmasse auf lokalkompakten geordneten Räumen. Thesis. Univ. of Erlangen.

[8] KALMYKOV, G. I. (1962). On the partial ordering of one-dimensional Markov processes. *Theor. Probability Appl.* **7** 456–459.

[9] KAMAE, T. and U. KRENGEL (1977). Stochastic partial ordering. *Ann. Probability.* To appear.

[10] KEILSON, J. and KESTER, A. (1977). Monotone matrices and monotone Markov processes. Unpublished manuscript.

[11] KRENGEL, U. (1977). Un theorème ergodique pour les processus sur-stationnaires. Unpublished manuscript.

[12] LOVÁSZ, L. and MAJOR, P. (1973). A note on a theorem of Dudley. *Studia Sci. Math. Hungar.* **8** 151–152.

[13] MAJOR, P. (1975). Private communication.

[14] MEYER, P. A. (1966). *Probability and Potentials.* Blaisdell, Waltham, Massachusetts.

[15] NACHBIN, L. (1951). Linear continuous functionals positive on the increasing continuous functions. *Summa Brasil. Math.* 2 135–150 (reprinted in [16]).

[16] NACHBIN, L. (1965). *Topology and Order.* Van Nostrand, New York.

[17] NAWROTSKI, K. (1962). Eine Monotonieeigenschaft zufälliger Punktfolgen. *Math. Nachr.* **24** 193–200.

[18] O'BRIEN, G. L. (1975). The comparison method for stochastic processes. *Ann. Probability* **3** 80–88.

[19] O'BRIEN, G. L. (1977). Stochastic inequalities for pure jump Markov processes. Unpublished manuscript.

[20] ORNSTEIN, D. S. (1973). An application of ergodic theory to probability theory. *Ann. Probability* **1** 43–65.

[21] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

[22] PRESTON, C. J. (1974). A generalization of the FKG inequalities. *Comm. Math. Phys.* **36** 233–241.

[23] ROHLIN, V. A. (1952). On the fundamental ideas of measure theory. *Amer. Math. Soc. Transl.* **71**.

[24] SNIJDERS, T. (1976). An ordering of probability distributions on a partially ordered outcome space. Preliminary report.

[25] SONDERMAN, D. and WHITT, W. (1977). Comparing continuous time Markov chains. Unpublished manuscript.

[26] STONE, C. (1963). Weak convergence of stochastic processes defined on semi-infinite time intervals. *Proc. Amer. Math. Soc.* **14** 694–696.

[27] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.

T. KAMAE
UER DE MATHEMATIQUES ET D'INFORMATIQUE
UNIVERSITÉ DE BORDEAUX I
357 COURS DE LA LIBÉRATION
33405 TALENCE, FRANCE

U. KRENGEL
INSTITUTE FÜR MATHEMATISCHE STATISTIK
UND WIRTSCHAFTSMATHEMATIK
34 GOTTINGEN
LOTZESTRASSE 13
W. GERMANY

G. L. O'BRIEN
DEPARTMENT OF MATHEMATICS
YORK UNIVERSITY
DOWNSVIEW, ONTARIO, M3J 1P3
CANADA