MATHEMATICS OF OPERATIONS RESEARCH
Vol. 12, No. 2, May 1987
Printed in U.S.A.

STOCHASTIC ORDERINGS FOR MARKOV PROCESSES ON PARTIALLY ORDERED SPACES*

WILLIAM A. MASSEY

AT & T Bell Laboratories

The purpose of this paper is to develop a unified theory of stochastic ordering for Markov processes on countable partially ordered state spaces. When such a space is not totally ordered, it can induce a wide range of stochastic orderings, none of which are equivalent to sample path comparisons. Similar comparison theorems are also developed for non-Markov processes that are functions of Markov processes and for time-inhomogeneous Markov processes. Such alternative orderings can be quite useful when analyzing multi-dimensional stochastic models such as queueing networks.

1. Introduction. For most areas of applied probability, the goal of explicitly calculating a probability distribution is rarely attainable. This makes alternative methods of analysis attractive. One approach is the notion of stochastic dominance. This is a partial ordering on probability distributions that allows one to define when one distribution is larger than another. So given an unknown distribution, it may be possible to construct a known upper or lower bound for it. For a broad perspective of the theory and utility of stochastic ordering, we cite Bawa [1] and Stoyan [22] as references.

The most widely studied type of stochastic dominance can be illustrated by the following simple example. Let X and Y be two real-valued random variables. We say that X *dominates* Y if the cumulative distribution function of X at each point is less than the same function for Y. Now this distribution of X can be duplicated by applying the inverse of the distribution function to a random variable U having the uniform distribution on $[0, 1]$. Call this random variable $\tilde{X}$. In a similar fashion, we can construct $\tilde{Y}$. Due to the manner of the construction, we have $\tilde{X} \geqslant \tilde{Y}$ a.e. if we use the same U. This says that we have a very strong stochastic ordering. We need only order the distributions of two random variables, and get a sample path ordering between their duplicates. Moreover, this characterization is necessary and sufficient. This concept achieved its fullest generalization in Kamae, Krengel, and O'Brien [8] for comparing stochastic processes on partially ordered Polish spaces. In order to do this, it was necessary to dispense with the notion of comparing distribution functions and instead compare measures of increasing sets. In practice, two processes are compared in this way by establishing the sample path comparisons. Roughly, one would take a common object like U, modify it to get the two different processes, and show that the result of modifying one is consistently less than the modification of the other.

When we restrict ourselves to time homogeneous Markov processes, a related problem arises. In practice, such a process is usually defined only in terms of its infinitesimal generator and initial distribution. Its transient distribution and even steady state distribution (when it exists) may be unknowable. So useful comparison

*Received April 3, 1985; revised April 14, 1986.
AMS 1980 subject classification. Primary: 60J99.
IAOR 1973 subject classification. Main: Stochastic processes.
OR/MS Index 1978 subject classification. Primary: 568 Probability/Markov processes.
Key words. Markov processes, stochastic ordering, stochastic dominance, queuing theory.

0364-765X/87/1202/0350$01.25

theorems are ones that only use the generators and initial distributions of the given Markov processes. Results of this nature were first developed in Kalmykov [6]. Daley [4] identified the critical role that monotone Markov processes play in such theorems. He also gave necessary and sufficient conditions for monotonicity in terms of the transition functions on totally ordered spaces. Kirsten [11] and Keilson and Kester [9] gave many examples of monotone Markov processes and derived other properties for them such as whether a given process is stochastically increasing in time. This approach is summarized in Chapter 4 of Stoyan [22]. Similar results are derived for semi-Markov processes in Sonderman [21].

The purpose of this paper is first to show that when a space is partially ordered, it admits a wide variety of stochastic orderings obtained by merely restricting the type of increasing sets used. Of these, only the *strongest* one may be equivalent to sample path orderings. For all these orderings, the second goal is to develop a unified comparison theory for Markov processes on countable partially ordered spaces in the spirit of Daley, Kalmykov, Keilson, Kester and many others. Such a theory would make it possible to establish a stochastic ordering between two Markov processes where no sample path comparison exists. This is indeed the case of the Jackson network as pointed out in [12] and [13].

In §2, we define a stochastic order and identify three natural candidates for orderings, the strong, weak, and weak*. §3 introduces the notion of a monotone Markov process and the main comparison theorems. Methods for constructing these monotone processes will be developed in §4. In §5 we use Strassen's theorem to strengthen the previous results for strong orderings. §6 gives deeper results for weak orderings. In §7, we extend all of these comparison results to time-inhomogeneous Markov processes. We illustrate in §8, the utility of these many orderings by interpreting the results of [12], [13], and Whitt [26], as stochastic comparisons of various Jackson networks.

2. Stochastic orderings. Given a countable, partially ordered space E with the discrete topology, let $\mathscr{I}(E)$ be a family of subsets of E that includes E itself and the empty set. We can then induce a transitive relation for probability measures on E. If P and Q are two such measures, we say that $P \leqslant_{\mathscr{I}} Q$ whenever $P(\Gamma) \leqslant Q(\Gamma)$ for all Γ in $\mathscr{I}(E)$.

DEFINITION 2.1. We say that $\leqslant_{\mathscr{I}}$ is a *stochastic ordering* on E if

(i) The relation $\leqslant_{\mathscr{I}}$ is a partial order on the space of probability measures.

(ii) Let δ_x be the point mass measure on E for some x in E. For all x and y in E, $x \leqslant y$ if and only if $\delta_x \leqslant_{\mathscr{I}} \delta_y$.

As we stated before, $\leqslant_{\mathscr{I}}$ is already a transitive relation. Condition (i) is then equivalent to requiring that a probability distribution be uniquely determined by its measure on the sets in $\mathscr{I}(E)$. If X and Y are two E-valued random variables, we will say that $X \leqslant_{\mathscr{I}} Y$ whenever their induced measures can be so ordered by $\leqslant_{\mathscr{I}}$. Condition (ii) is then seen to be a compatibility condition where $x \leqslant y$ holds if and only if $x \leqslant_{\mathscr{I}} y$.

For any subset Γ of E, we borrow the following notation from Kamae and Krengel [7].

$$\Gamma^{\uparrow} = \{y \mid y \geqslant x \text{ for some } x \text{ in } \Gamma\}, \quad \Gamma^{\downarrow} = \{y \mid y \leqslant x \text{ for some } x \text{ in } \Gamma\}.$$

DEFINITION 2.2. A subset Γ of E is an *increasing set* if $\Gamma = \Gamma^{\uparrow}$. A family of increasing sets is said to be *strongly separating* if for all $x \nleqslant y$, the family contains a set Γ such that $x \in \Gamma$ and $y \notin \Gamma$. The concept of a *determining class* was used in O'Brien [17] to denote a set of functions where we can recover a probability measure by its integral against these functions. We adapt this notion to a family of sets by using their indicator functions.

PROPOSITION 2.3. *$\mathcal{I}(E)$ induces a stochastic order if and only if $\mathcal{I}(E)$ is a strongly separating family of increasing sets that form a determining class.*

PROOF. Let $\leqslant_{\mathcal{I}}$ be a stochastic ordering on E. For any subset Γ of E, $\delta_x(\Gamma) = 1$ if and only if $x \in \Gamma$, otherwise $\delta_x(\Gamma) = 0$. By condition (ii) of Definition 2.1, we have $x \leqslant y$ and $x \in \Gamma$ implying that $y \in \Gamma$ whenever Γ belongs to $\mathcal{I}(E)$. This holds because $\delta_x(\Gamma) \leqslant \delta_y(\Gamma)$ in this case. Since $x \leqslant y$ is equivalent to $\delta_x \leqslant_{\mathcal{I}} \delta_y$, then $x \nleqslant y$ holds if and only if for some Γ in $\mathcal{I}(E)$, we have $\delta_x(\Gamma) > \delta_y(\Gamma)$. It then follows that $x \in \Gamma$ and $y \notin \Gamma$. So $\mathcal{I}(E)$ is a strongly separating family of sets.

Conversely, if $\mathcal{I}(E)$ is a determining class then any probability distribution is uniquely determined by the sets of $\mathcal{I}(E)$. This makes $\leqslant_{\mathcal{I}}$ a partial order on the space of probability measures on E. Now let $x \in E$. Since $\delta_x(\Gamma)$ equaling 0 or 1 is equivalent to $x \notin \Gamma$ or $x \in \Gamma$ respectively, we have $\delta_x \leqslant_{\mathcal{I}} \delta_y$ whenever $x \leqslant y$. Now if $\delta_x \leqslant_{\mathcal{I}} \delta_y$, we must have $x \leqslant y$. Otherwise $x \nleqslant y$, and by the strong separation condition, there exists $\Gamma \in \mathcal{I}(E)$ such that $x \in \Gamma$ and $y \notin \Gamma$. This however gives us $1 = \delta_x(\Gamma) \leqslant \delta_y(\Gamma) = 0$ which is a contradiction. ∎

We now define some distinguished increasing sets. For any x in E we have $\langle x \rangle = \{x\}^{\uparrow}$, $\langle x \rangle_* = E \setminus \{x\}^{\downarrow}$. This allows us to specify three candidates for stochastic orderings.

DEFINITION 2.4. Let $\mathcal{I}_{st}(E)$, $\mathcal{I}_{wk}(E)$, and $\mathcal{I}_{wk^*}(E)$, denote respectively, the *strong weak, and weak* orderings* where

$$\mathcal{I}_{st}(E) = \{\text{all increasing sets in } E\}, \quad \mathcal{I}_{wk}(E) = \{\langle x \rangle | x \in E\} \cup \{E, \varnothing\},$$

$$\mathcal{I}_{wk^*}(E) = \{\langle x \rangle_* | x \in E\} \cup \{E, \varnothing\}.$$

We will denote these orderings respectively as $\leqslant_{st}$, $\leqslant_{wk}$, and $\leqslant_{wk^*}$. The strong ordering is the one that is equivalent to a sample path comparison of the random variables. Weak orderings are equivalent to comparing tail distribution functions and weak* orderings serve the same role for cumulative distribution functions. Examples of their usefulness can be found in Tong [24] and Stoyan [22]. The similarity of the nomenclature above to the various types of convergence on linear topological spaces is intentional. Just as weaker topologies are defined by restricting the family of open sets used, we can define stochastic orderings weaker than the strong one by restricting the family of increasing sets used. Continuing the analogy, recall that for finite dimensional spaces, all topologies are equivalent. Totally ordered spaces fill the corresponding role for stochastic orderings.

PROPOSITION 2.5. *If E is a totally ordered space, then all of its stochastic orderings are equivalent.*

PROOF. Let $\leqslant_{\mathcal{I}}$ be a stochastic ordering. We will show that for all x, there exists a sequence of increasing sets Γ_n in $\mathcal{I}(E)$ such that

$$P(\langle x \rangle) = \lim_{n \to \infty} P(\Gamma_n). \tag{2.1}$$

Conversely, given any Γ in $\mathcal{I}(E)$, there exists a sequence $\{y_n\}$ in E such that

$$P(\Gamma) = \lim_{n \to \infty} P(\langle y_n \rangle). \tag{2.2}$$

From (2.1) and (2.2) we will have $P \leqslant_{\mathcal{I}} Q$ if and only if $P \leqslant_{wk} Q$.

E is a countably infinite set, so for any x in E, there exists an increasing sequence $\{x_n\}$ such that $x_n \leqslant x$ for all n, and $y \leqslant x$ implies that $y \leqslant x_k$ for some integer k. Merely take $\{z_1, z_2, \dots\}$ to be an enumeration of all the elements strictly less than x. Now let $x_n = \max(z_1, \dots, z_n)$. By recursion, we can construct the following sequence of sets $\{\Gamma_n\}$ in $\mathscr{I}(E)$ where $\langle x \rangle = \bigcap_{n=1}^{\infty} \Gamma_n$ as follows:

$[n = 1]$. Let Γ_1 be an increasing set in $\mathscr{I}(E)$ such that $x_1 \notin \Gamma_1$ but $x \in \Gamma_1$.

$[n \to n+1]$. Let $\Gamma_{n+1} = \Gamma_n$ if $x_{n+1} \notin \Gamma_n$. Otherwise, choose a new set in $\mathscr{I}(E)$ such that $x_{n+1} \notin \Gamma_{n+1}$ and $x \in \Gamma_{n+1}$.

By Proposition 2.3, we can always construct such a sequence. Since E is a totally ordered space, then $\{\Gamma_n\}$ is a decreasing chain of sets in $\mathscr{I}(E)$ with $\langle x \rangle = \bigcap_{n=1}^{\infty} \Gamma_n$, so $P(\langle x \rangle) = \lim_{n \to \infty} P(\Gamma_n)$.

If Γ belongs to $\mathscr{I}(E)$, then let $\{x_n\}$ be an enumeration of all the elements in Γ. Since Γ is an increasing set, then $\Gamma = \bigcup_{n=1}^{\infty} \langle x_n \rangle$ and so $P(\Gamma) = \lim_{n \to \infty} P(\bigcup_{k=1}^{n} \langle x_k \rangle)$. But E is totally ordered, so there exists a unique y_n for $x_1, \dots, x_n$ such that $y_n = \min(x_1, \dots, x_n)$ and $\langle y_n \rangle = \bigcup_{k=1}^{n} \langle x_k \rangle$, hence $P(\Gamma) = \lim_{n \to \infty} P(\langle y_n \rangle)$ and this finishes the proof. ∎

We now give a familiar example of a space where the various types of stochastic orderings can be quite distinct.

EXAMPLE 2.6. Let $E = Z_+^2$, the space of ordered pairs of nonnegative integers. We say that $(m_1, m_2) \leqslant (n_1, n_2)$ if $m_1 \leqslant n_1$ and $m_2 \leqslant n_2$. Define P and Q as $P(0,1) = P(1,0) = \frac{1}{2}$ and $Q(1,1) = Q(0,0) = \frac{1}{2}$.

We then have $P \leqslant_{wk} Q$ but $Q \leqslant_{wk^*} P$ so there can be no strong ordering between P and Q. Moreover, increasing functions on E can behave badly with respect to the wrong stochastic ordering.

EXAMPLE 2.7. Let $E = Z_+^2$, with the same partial ordering as in Example 2.6. Let f map Z_+^2 into Z_+^2 where

$$f(m,n) = \begin{cases} (m-1, n+1) & \text{if } m > 0, \\ (m,n) & \text{if } m = 0. \end{cases}$$

Let X and Y be two Z_+^2-valued random variables whose induced measures on Z_+^2 are P and Q as given in Example 2.6. We then have $X \leqslant_{wk} Y$ but $f(X) \nleqslant_{wk} f(Y)$, despite the fact that f is an increasing function. We can remedy this situation in the following way.

DEFINITION 2.8. Let E and F be two partially ordered spaces, with $\mathscr{I}(E)$ and $\mathscr{I}(F)$ defining their respective stochastic orderings. We say that a function f from E to F is an *isotone mapping* from $\mathscr{I}(E)$ to $\mathscr{I}(F)$ if $f^{-1}(\mathscr{I}(F)) \subseteq \mathscr{I}(E)$. If $E = F$ and $f^{-1}(\mathscr{I}(E)) \subseteq \mathscr{I}(E)$, we say that f is *$\mathscr{I}(E)$-isotone*.

PROPOSITION 2.9. *If f is an isotone mapping from $\mathscr{I}(E)$ to $\mathscr{I}(F)$, then for any two E-valued random variables X and Y we have*

$$X \leqslant_{\mathscr{I}} Y \Rightarrow f(X) \leqslant_{\mathscr{I}} f(Y).$$

Any such isotone mapping is then an increasing function.

PROOF. Let Γ belong to $\mathscr{I}(F)$. Now $f^{-1}(\Gamma)$ belongs to $\mathscr{I}(E)$, so

$$\Pr\{f(X) \in \Gamma\} = \Pr\{X \in f^{-1}(\Gamma)\} \leqslant \Pr\{Y \in f^{-1}(\Gamma)\} = \Pr\{f(Y) \in \Gamma\}$$

and this completes the proof. ∎

3. Comparison theorems. Any probability measure P on E is uniquely characterized by its value on all of the singleton sets $\{x\}$ for $x \in E$ which we denote as $P(x)$. We can then alternatively view P as a function that maps E into the real line or more precisely, the closed interval $[0,1]$. Let $l_1(E)$ denote the Banach space of all absolutely summable real valued functions on E. Now define $\mathbf{e}_x$ for each x in E to be the indicator function for the singleton set $\{x\}$. All the $\mathbf{e}_x$'s comprise a basis for $l_1(E)$. We can then encode every measure P on E as a vector $\mathbf{p}$ in $l_1(E)$ where

$$\mathbf{p} = \sum_{x \in E} P(x)\mathbf{e}_x. \tag{3.1}$$

Notice that the l_1-norm of $\mathbf{p}$ is the sum of the absolute values of the "e_x-based" coefficients, which equals 1 in this case.

Let $l_\infty(E)$ equal the Banach space of bounded real-valued functions on E. If Γ is a subset of E, let $\mathbf{1}_\Gamma$ be its indicator function. All the $\mathbf{1}_\Gamma$'s comprise an *uncountable* basis for $l_\infty(E)$. We can then define the natural inner product between elements in $l_1(E)$ and $l_\infty(E)$ by setting

$$\mathbf{e}_x \cdot \mathbf{1}_\Gamma = \begin{cases} 1, & x \in \Gamma, \\ 0, & x \notin \Gamma. \end{cases}$$

If we apply $\mathbf{1}_\Gamma$ to $\mathbf{p}$ in (3.1), we get $\mathbf{p} \cdot \mathbf{1}_\Gamma = P(\Gamma)$. So now the statement $P \leqslant_{\mathscr{I}} Q$ is equivalent to

$$\mathbf{p} \cdot \mathbf{1}_\Gamma \leqslant \mathbf{q} \cdot \mathbf{1}_\Gamma \text{ for all } \Gamma \text{ in } \mathscr{I}(E)$$

where $\mathbf{p}$ and $\mathbf{q}$ are the $l_1(E)$ representations of P and Q.

The spaces $l_1(E)$ and bounded linear operators on $l_1(E)$ admit natural partial orderings through the definition of positive vectors (see [15, p. 179]). We will use $\leqslant$ to denote such orderings. We can also endow these spaces with a partial ordering induced by $\mathscr{I}(E)$.

PROPOSITION 3.1. *Let $\mathscr{I}(E)$ be a stochastic ordering on E, then the following relations are partial orders*:

(i) *For $\mathbf{f}$ and $\mathbf{g}$ in $l_1(E)$ we say that $\mathbf{f} \leqslant_{\mathscr{I}} \mathbf{g}$ whenever*

$$\mathbf{f} \cdot \mathbf{1}_\Gamma \leqslant \mathbf{g} \cdot \mathbf{1}_\Gamma \quad \textit{for all } \Gamma \textit{ in } \mathscr{I}(E).$$

(ii) *If $\mathbf{A}$ and $\mathbf{B}$ are two bounded linear operators on $l_1(E)$, we say that $\mathbf{A} \leqslant_{\mathscr{I}} \mathbf{B}$ provided*

$$\mathbf{A}\mathbf{1}_\Gamma \leqslant \mathbf{B}\mathbf{1}_\Gamma \quad \textit{for all } \Gamma \textit{ in } \mathscr{I}(E).$$

PROOF. For $\mathbf{f}$ and $\mathbf{g}$ in $l_1(E)$, $\mathbf{f} \leqslant_{\mathscr{I}} \mathbf{g}$ is a transitive relation. It then remains to show that $\mathbf{f} \leqslant_{\mathscr{I}} \mathbf{g}$ and $\mathbf{g} \leqslant_{\mathscr{I}} \mathbf{f}$ imply $\mathbf{f} = \mathbf{g}$. Given $\mathbf{f} \leqslant_{\mathscr{I}} \mathbf{g}$ and $\mathbf{g} \leqslant_{\mathscr{I}} \mathbf{f}$, we have

$$(\mathbf{f} - \mathbf{g}) \cdot \mathbf{1}_\Gamma = 0 \quad \text{for all } \Gamma \text{ in } \mathscr{I}(E).$$

The case $\Gamma = E$ says that all of the components of $\mathbf{f} - \mathbf{g}$ sum to zero. This means that there are two positive $l_1(E)$-vectors $\mathbf{h}_+$ and $\mathbf{h}_-$, where $\mathbf{f} - \mathbf{g} = \mathbf{h}_+ - \mathbf{h}_-$. We then have

$$\mathbf{h}_+ \cdot \mathbf{1}_\Gamma = \mathbf{h}_- \cdot \mathbf{1}_\Gamma \quad \text{for all } \Gamma \text{ in } \mathscr{I}(E).$$

Since $\mathbf{h}_+ \cdot \mathbf{1}_E = \mathbf{h}_- \cdot \mathbf{1}_E$, we can scale $\mathbf{h}_+$ and $\mathbf{h}_-$ so that they both represent probability distributions. If P_+ is the measure corresponding to $\mathbf{h}_+$, and P_- to $\mathbf{h}_-$, then we have $P_+ \leqslant_{\mathscr{I}} P_-$ and $P_- \leqslant_{\mathscr{I}} P_+$. But $\mathscr{I}(E)$ is a stochastic order so $P_+ = P_-$, then $\mathbf{h}_+ = \mathbf{h}_-$, and finally $\mathbf{f} = \mathbf{g}$.

For ordering defined in (ii), we note that $\mathbf{A} \cdot \mathbf{1}_\Gamma \leqslant \mathbf{B} \cdot \mathbf{1}_\Gamma$ holds in $l_\infty(E)$ if and only if $\mathbf{fA} \cdot \mathbf{1}_\Gamma \leqslant \mathbf{fB} \cdot \mathbf{1}_\Gamma$ for all positive $\mathbf{f}$ in $l_1(E)$. But now we have the same relation between $\mathbf{fA}$ and $\mathbf{fB}$ in $l_1(E)$ that was defined in (i). By the previous arguments, $\mathbf{fA} = \mathbf{fB}$ for all positive $\mathbf{f}$ in $l_1(E)$. Since the $\mathbf{e}_x$'s are positive and are a basis for $l_1(E)$, we have $\mathbf{A} = \mathbf{B}$. ∎

Given a stochastic process $\{X(t)|t \geqslant 0\}$ on E, let $\mathbf{p}(t)$ be the $l_1(E)$-vector representing the probability distribution of $X(t)$. Whenever $\{X(t)|t \geqslant 0\}$ is a conservative Markov process that is pure jump (right continuous sample paths), $\mathbf{p}(t)$ solves a simple differential equation in $l_1(E)$,

$$\frac{d}{dt}\mathbf{p}(t) = \mathbf{p}(t)\mathbf{A}$$

where $\mathbf{A}$ is a linear operator that acts on at least some dense subspace of $l_1(E)$. This is called the *forward equation* for $\{X(t)|t \geqslant 0\}$ and $\mathbf{A}$ is the *infinitesimal generator* of the process. Since any contraction semigroup is the strong operator limit of similar semigroups with bounded generators, we will assume $\mathbf{A}$ to be a bounded operator acting on $l_1(E)$. Such an $\mathbf{A}$ is easily characterized by the following:

1. $\mathbf{A}\mathbf{1}_E = 0$.
2. The off-diagonal entries of $\mathbf{A}$ are positive.

Such generators are also called *uniformizable* (see Keilson and Kester [9]). If $\lambda > 0$ is greater than the absolute value of any diagonal term (guaranteed if $\lambda \geqslant |\mathbf{A}|_1$), then $\mathbf{P}_\lambda(\mathbf{A}) = \mathbf{I} + \mathbf{A}/\lambda$ is a stochastic matrix and we can expand $\exp(t\mathbf{A})$ as

$$\exp(t\mathbf{A}) = \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!}\mathbf{P}_\lambda(\mathbf{A})^n.$$

DEFINITION 3.2. $\mathbf{A}$ is the generator for an *$\mathscr{I}(E)$-monotone Markov process* if for all probability vectors $\mathbf{p}$ and $\mathbf{q}$ in $l_1(E)$, we have

$$\mathbf{p} \leqslant_{\mathscr{I}} \mathbf{q} \quad \text{implies that} \quad \mathbf{p}\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{A}).$$

LEMMA 3.3. *If $\mathbf{A}$ is the generator for an $\mathscr{I}(E)$-monotone Markov process, then for $\mathbf{f}$ and $\mathbf{g}$ in $l_1(E)$ with $\mathbf{f} \cdot \mathbf{1}_E = \mathbf{g} \cdot \mathbf{1}_E$ we have*

$$\mathbf{f} \leqslant_{\mathscr{I}} \mathbf{g} \quad \text{implies that} \quad \mathbf{f}\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \mathbf{g}\exp(t\mathbf{A}).$$

Similarly, if $\mathbf{B}$ and $\mathbf{C}$ are bounded operators on $l_1(E)$ with $\mathbf{B}\mathbf{1}_E = \mathbf{C}\mathbf{1}_E$, then

$$\mathbf{B} \leqslant_{\mathscr{I}} \mathbf{C} \quad \textit{implies that} \quad \mathbf{B}\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \mathbf{C}\exp(t\mathbf{A}).$$

PROOF. The argument is very much like the one for Proposition 3.1.

THEOREM 3.4. *Let $X(t)$ and $Y(t)$ be two Markov processes with state space E, where $X(t)$ or $Y(t)$ is $\mathscr{I}(E)$-monotone. If $\mathbf{A}$ and $\mathbf{B}$ are their respective generators, $X(0) \leqslant_{\mathscr{I}} Y(0)$, and $\mathbf{A} \leqslant_{\mathscr{I}} \mathbf{B}$, then $X(t) \leqslant_{\mathscr{I}} Y(t)$ for all $t \geqslant 0$.*

PROOF. Let $\mathbf{p}$ be the $l_1(E)$ representation for the distribution of $X(0)$, and $\mathbf{q}$ be the same for $Y(0)$. In operator notation, we want to prove that $\mathbf{p} \leqslant_{\mathscr{I}} \mathbf{q}$ implies that

$$\mathbf{p}\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{B}).$$

Assume that $\mathbf{B}$ is the generator for an $\mathscr{I}(E)$-monotone process, then it suffices to show that $\mathbf{p}\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \mathbf{p}\exp(t\mathbf{B})$ for all probability vectors $\mathbf{p}$. This is equivalent to asserting that $\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \exp(t\mathbf{B})$. This, in turn, is equivalent to $0 \leqslant_{\mathscr{I}} \exp(t\mathbf{B}) - \exp(t\mathbf{A})$, but $\exp(t\mathbf{B}) - \exp(t\mathbf{A})$, can be rewritten as follows

$$\exp(t\mathbf{B}) - \exp(t\mathbf{A}) = \int_0^t \frac{d}{ds}\exp((t-s)\mathbf{A})\exp(s\mathbf{B})\,ds$$

$$= \int_0^t \exp((t-s)\mathbf{A})(\mathbf{B}-\mathbf{A})\exp(s\mathbf{B})\,ds.$$

Now our problem reduces to showing that

$$0 \leqslant_{\mathscr{I}} \exp((t-s)\mathbf{A})(\mathbf{B}-\mathbf{A})\exp(s\mathbf{B})$$

for all s in $[0, t]$.

Since $\exp((t-s)\mathbf{A})$ maps positive vectors into positive vectors if and only if the off-diagonal terms of $\mathbf{A}$ are positive (see Keilson and Kester [9]), then we need only show that for all $t \geqslant 0$, $0 \leqslant_{\mathscr{I}} (\mathbf{B}-\mathbf{A})\exp(t\mathbf{B})$. By Lemma 3.3, we are done. ∎

Let E and F be two partially ordered spaces. If f is a function mapping E into F, then let $\Phi(f)$ be the following bounded operator that maps $l_1(E)$ into $l_1(F)$:

$$\mathbf{e}_x\Phi(f) = \mathbf{e}_{f(x)}.$$

We will use this operator in the next theorem that gives us a means to compare a Markov process with the functional of another Markov process. To this extent, coupled with Proposition 2.9, we acquire a comparison technique for non-Markov processes also.

THEOREM 3.5. *Let $X(t)$ and $Y(t)$ be Markov processes with state spaces E' and E respectively, having corresponding generators $\mathbf{A}$ and $\mathbf{B}$. If $Y(t)$ is $\mathscr{I}(E)$-monotone, f maps E' into E, $f(X(0)) \leqslant_{\mathscr{I}} Y(0)$, and $\mathbf{A}\Phi(f) \leqslant_{\mathscr{I}} \Phi(f)\mathbf{B}$, then $f(X(t)) \leqslant_{\mathscr{I}} Y(t)$ for all $t \geqslant 0$. The same result holds when the direction of the ordering is reversed.*

PROOF. The argument runs exactly the same as in the proof of Theorem 3.1, except we are proving that

$$\mathbf{p}\exp(t\mathbf{A})\Phi(f) \leqslant_{\mathscr{I}} \mathbf{q}\Phi(f)\exp(t\mathbf{B})$$

and the key identity here is

$$\Phi(f)\exp(t\mathbf{B}) - \exp(t\mathbf{A})\Phi(f) = \int_0^t \frac{d}{ds}\exp((t-s)\mathbf{A})\Phi(f)\exp(s\mathbf{B})\,ds$$

$$= \int_0^t \exp((t-s)\mathbf{A})(\Phi(f)\mathbf{B} - \mathbf{A}\Phi(f))\exp(s\mathbf{B})\,ds.$$

Now $\Phi(f)\mathbf{1}_E = \mathbf{1}_{E'}$, so $(\Phi(f)\mathbf{B} - \mathbf{A}\Phi(f))\mathbf{1}_E = 0$. Given $\mathbf{A}\Phi(f) \leqslant_{\mathscr{I}} \Phi(f)\mathbf{B}$ the rest follows. ∎

Special cases of this theorem were used by the author in the second half of Theorem 3.1 in [13], and Theorems 5.4 and 6.4 in [15]. Such a comparison result was derived independently by Whitt [27] for general state spaces by omitting supplementary variables, and he applied it to blocking networks in [25].

We now generalize a property introduced in Keilson and Kester [9].

DEFINITION 3.6. A Markov process $\{X(t)\}_{t\geq 0}$ is $\mathscr{I}(E)$*-time increasing* if $\Pr\{X(t) \in \Gamma\}$ is an increasing function of time for all Γ in $\mathscr{I}(E)$. If such quantities are decreasing functions of time, then $\{X(t)\}_{t\geq 0}$ is $\mathscr{I}(E)$*-time decreasing.*

THEOREM 3.7. *Let $\{X(t)\}_{t\geq 0}$ be a monotone Markov process with generator* $\mathbf{A}$ *and initial distribution vector* $\mathbf{p}$. *The following statements hold*:

(i) *$\{X(t)\}_{t\geq 0}$ is $\mathscr{I}(E)$-time increasing if and only if* $\mathbf{pA} \geq_{\mathscr{I}} 0$.

(ii) *$\{X(t)\}_{t\geq 0}$ is $\mathscr{I}(E)$-time decreasing if and only if* $\mathbf{pA} \leq_{\mathscr{I}} 0$.

PROOF. Since $\Pr\{X(t) \in \Gamma\} = \mathbf{p} \cdot \exp(t\mathbf{A})\mathbf{1}_\Gamma$, then $\Pr\{X(t) \in \Gamma\}$ is an increasing function of time if and only if

$$\frac{d}{dt}\Pr\{X(t) \in \Gamma\} = \mathbf{pA}\exp(t\mathbf{A}) \cdot \mathbf{1}_\Gamma \geq 0.$$

Therefore $\{X(t)\}_{t\geq 0}$ is $\mathscr{I}(E)$-increasing if and only if $\mathbf{pA}\exp(t\mathbf{A}) \geq_{\mathscr{I}} 0$. If $\mathbf{pA} \geq_{\mathscr{I}} 0$ and $\mathbf{A}$ is monotone, then $\mathbf{pA}\exp(t\mathbf{A}) \geq_{\mathscr{I}} 0$ by Lemma 3.3. If $\mathbf{pA}\exp(t\mathbf{A}) \geq_{\mathscr{I}} 0$ for all $t \geq 0$, we merely set $t = 0$ and get $\mathbf{pA} \geq_{\mathscr{I}} 0$. The proof for statement (ii) is similar. ∎

4. Constructing monotone Markov processes.

In §3, we demonstrated the importance of monotone Markov processes to deriving comparison theorems. Now, we will introduce methods for constructing them. For all functions f that map E into itself, and all $\boldsymbol{\alpha}$ in $l_\infty(E)$, we can define two bounded operators $\boldsymbol{\Phi}(\mathbf{f})$ and $\boldsymbol{\Delta}(\boldsymbol{\alpha})$ where

$$\mathbf{e}_x\boldsymbol{\Phi}(\mathbf{f}) = \mathbf{e}_{f(x)}, \quad \mathbf{e}_x\boldsymbol{\Delta}(\boldsymbol{\alpha}) = \alpha(x)\mathbf{e}_x.$$

Notice that $\boldsymbol{\Phi}(\mathbf{f})\mathbf{1}_\Gamma = \mathbf{1}_{f^{-1}(\Gamma)}$ and so $\boldsymbol{\Phi}(\mathbf{f})\mathbf{1}_E = \mathbf{1}_E$. Before we proceed, we define $\Gamma\#\Gamma'$ to be the symmetric difference for two subsets of E or $(\Gamma \setminus \Gamma') \cup (\Gamma' \setminus \Gamma)$.

THEOREM 4.1. $\mathbf{A}$ *is an $\mathscr{I}(E)$-monotone generator for a Markov process if*:

(i) $\mathbf{A}$ *is the strong operator limit of other $\mathscr{I}(E)$-monotone generators.*

(ii) $\mathbf{A}$ *is the sum of two $\mathscr{I}(E)$-monotone generators.*

(iii) $\mathbf{A} = \boldsymbol{\Delta}(\boldsymbol{\alpha})(\boldsymbol{\Phi}(\mathbf{f}) - \mathbf{I})$, *where f is an $\mathscr{I}(E)$-isotone function and $\boldsymbol{\alpha}$ is a positive bounded function on E that is constant on all sets of the form $\Gamma\#f^{-1}(\Gamma)$ for all Γ in $\mathscr{I}(E)$.*

PROOF. Let $\mathbf{p} \leq_{\mathscr{I}} \mathbf{q}$. By the hypothesis of (i), there is a sequence $\{\mathbf{A}_n\}$ of $\mathscr{I}(E)$-monotone generators that converge to $\mathbf{A}$ in the strong operator topology. It then follows that the sequence $\{\exp(t\mathbf{A}_n)\}$ converges similarly to $\exp(t\mathbf{A})$. We have $\mathbf{p}\exp(t\mathbf{A}_n) \leq_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{A}_n)$ for all n since each $\mathbf{A}_n$ is $\mathscr{I}(E)$-monotone. Taking the limit gives us $\mathbf{p}\exp(t\mathbf{A}) \leq_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{A})$, and so we are done.

Now suppose that $\mathbf{A} = \mathbf{B} + \mathbf{C}$ where $\mathbf{B}$ and $\mathbf{C}$ are $\mathscr{I}(E)$-monotone generators. If $\mathbf{p} \leq_{\mathscr{I}} \mathbf{q}$, we then have $\mathbf{p}\exp(t\mathbf{B}) \leq_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{B})$. From this we get

$$\mathbf{p}\exp(t\mathbf{B})\exp(t\mathbf{C}) \leq_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{B})\exp(t\mathbf{C}).$$

By induction, we have for all n

$$\mathbf{p}\left(\exp\left(\frac{t}{n}\mathbf{B}\right)\exp\left(\frac{t}{n}\mathbf{C}\right)\right)^n \leq_{\mathscr{I}} \mathbf{q}\left(\exp\left(\frac{t}{n}\mathbf{B}\right)\exp\left(\frac{t}{n}\mathbf{C}\right)\right)^n.$$

Using the Trotter product formula (see Reed and Simon [19, p. 295]), we have in the norm operator topology,

$$\lim_{n\to\infty}\left(\exp\left(\frac{t}{n}\mathbf{B}\right)\exp\left(\frac{t}{n}\mathbf{C}\right)\right)^n = \exp(t\mathbf{A}).$$

From this it follows that $\mathbf{p}\exp(t\mathbf{A}) \leqslant_{\mathscr{I}} \mathbf{q}\exp(t\mathbf{A})$ and so (ii) holds.

Now let $\mathbf{A} = \Delta(\alpha)(\Phi(\mathbf{f}) - \mathbf{I})$ as given by (iii). We are given that $\mathbf{p}\cdot\mathbf{1}_\Gamma \leqslant \mathbf{q}\cdot\mathbf{1}_\Gamma$ for all Γ in $\mathscr{I}(E)$. To show that $\mathbf{p}\cdot\exp(t\mathbf{A})\mathbf{1}_\Gamma \leqslant \mathbf{q}\cdot\exp(t\mathbf{A})\mathbf{1}_\Gamma$, it is sufficient to prove that $\exp(t\mathbf{A})$ maps the l_∞-closure of the positive linear cone generated by $\{\mathbf{1}_\Gamma\}_{\Gamma\in\mathscr{I}(E)}$ into itself. To show this, it is sufficient to find some positive constant β such that $\mathbf{A} + \beta\mathbf{I}$ has this property since

$$\exp(t\mathbf{A}) = \sum_{n=0}^{\infty}\frac{e^{-\beta t}t^n}{n!}(\mathbf{A}+\beta\mathbf{I})^n.$$

Setting $\beta = |\alpha|_\infty$ gives us

$$\begin{aligned}(\mathbf{A}+\beta\mathbf{I})\mathbf{1}_\Gamma &= (\beta\mathbf{I} + \Delta(\alpha)(\Phi(f) - \mathbf{I}))\mathbf{1}_\Gamma\\ &= \beta\mathbf{1}_\Gamma + \Delta(\alpha)(\mathbf{1}_{f^{-1}(\Gamma)} - \mathbf{1}_\Gamma).\end{aligned}$$

As indicator functions for all subsets Δ and Γ of E we have

$$|\mathbf{1}_\Delta(x) - \mathbf{1}_\Gamma(x)| = \mathbf{1}_{\Delta\#\Gamma}(x)$$

for all x in E. Since α is a bounded function on E that is constant on $f^{-1}(\Gamma)\#\Gamma$, we can replace $\Delta(\alpha)(\mathbf{1}_{f^{-1}(\Gamma)} - \mathbf{1}_\Gamma)$ by $\alpha(\Gamma)(\mathbf{1}_{f^{-1}(\Gamma)} - \mathbf{1}_\Gamma)$ where $\alpha(\Gamma)$ equals α evaluated at some element in $f^{-1}(\Gamma)\#\Gamma$. We now have

$$(\mathbf{A}+\beta\mathbf{I})\mathbf{1}_\Gamma = (\beta - \alpha(\Gamma))\mathbf{1}_\Gamma + \alpha(\Gamma)\mathbf{1}_{f^{-1}(\Gamma)}.$$

By definition of β, $\beta - \alpha(\Gamma) > 0$ and so $\mathbf{A} + \beta\mathbf{I}$ preserves the positive cone generated by $\{\mathbf{1}_\Gamma\}_{\Gamma\in\mathscr{I}(E)}$ if f^{-1} preserves $\mathscr{I}(E)$, so we are done. ∎

Theorem 4.1 gives us a way to construct two stochastically ordered processes.

THEOREM 4.2. *Let $\mathbf{A} = \sum_{i\in I}\Delta(\alpha_i)(\mathbf{E}(\mathbf{f}_i) - \mathbf{I})$ be a bounded operator where the α_i and f_i satisfy hypothesis* (iii) *of Theorem* 4.1, *and I is a countable index set. Now let $\mathbf{B} = \sum_{i\in I}\Delta(\alpha_i)(\mathbf{E}(\mathbf{g}_i) - \mathbf{I})$ where the g_i are functions that map E into itself such that $f_i(x) \leqslant g_i(x)$ for all x in E. If $X(t)$ is the Markov process with generator* $\mathbf{A}$, *$Y(t)$ is similarly defined by* $\mathbf{B}$, *and $X(0) \leqslant_{\mathscr{I}} Y(0)$, then $X(t) \leqslant_{\mathscr{I}} Y(t)$. The same result holds when the directions of all the orderings are reversed.*

PROOF. By Theorem 4.1, we know that $\mathbf{A}$ is $\mathscr{I}(E)$-monotone. By Theorem 3.4, we need only show that $\mathbf{A} \leqslant_{\mathscr{I}} \mathbf{B}$. By the similar decompositions of $\mathbf{A}$ and $\mathbf{B}$, it is sufficient to show that $\mathbf{E}(\mathbf{f}_i) \leqslant_{\mathscr{I}} \mathbf{E}(\mathbf{g}_i)$ for all $i \in I$. By the definitions of $\leqslant_{\mathscr{I}}$ and $\mathbf{E}(\mathbf{f}_i)$, this means that we want to prove for all $i \in I$ and $\Gamma \in \mathscr{I}(E)$ that $\mathbf{1}_{f_i^{-1}(\Gamma)} \leqslant \mathbf{1}_{g_i^{-1}(\Gamma)}$. This holds if and only if $f_i^{-1}(\Gamma) \subseteq g_i^{-1}(\Gamma)$. But if $y \in f_i^{-1}(\Gamma)$, then $f_i(y) \in \Gamma$ and by hypothesis $g_i(y) \geqslant f_i(y)$. So we have $g_i(y) \in \Gamma$ since Γ is an increasing set, or $y \in g_i^{-1}(\Gamma)$. This finishes the proof. ∎

5. Strong orderings. The theorems derived in §4 for a general comparison of Markov processes can be sharpened considerably for the strong ordering. This follows from the key result below.

THEOREM 5.1 (Strassen). *Let* $\mathbf{p}$ *and* $\mathbf{q}$ *be two probability vectors in* $l_1(E)$. *We have* $\mathbf{p} \leqslant_{st} \mathbf{q}$ *if and only if there exists a probability vector* $\mathbf{r}$ *in* $l_1(E \times E)$ *such that* $r(x, y) = 0$ *if* $x \nleqslant y$ *and*

$$\mathbf{p} = \sum_{x \in E} \sum_{y \geqslant x} r(x, y)\mathbf{e}_x, \quad \mathbf{q} = \sum_{x \in E} \sum_{y \geqslant x} r(x, y)\mathbf{e}_y.$$

Strassen's result [23] is actually more general than the statement given above. For the case of E being a finite set, we refer the reader to an elegant proof of Theorem 5.1 in Preston [18] that uses the min-cut, max-flow theorem. The following two theorems are generalizations of results proved in Kester [10].

THEOREM 5.2. *Let* $X(t)$ *be a Markov process with generator* $\mathbf{A}$. *Then* $X(t)$ *is strongly monotone if and only if for all* $x \leqslant y$ *in* E, *and all increasing sets* Γ *such that either* $x \in \Gamma$ *or* $y \notin \Gamma$ *holds, we have*

$$\mathbf{e}_x\mathbf{A}\mathbf{1}_\Gamma \leqslant \mathbf{e}_y\mathbf{A}\mathbf{1}_\Gamma. \tag{5.1}$$

The proof of this follows immediately from the theorem below.

THEOREM 5.3. *Let* $X(t)$ *and* $Y(t)$ *be two Markov processes with generators* $\mathbf{A}$ *and* $\mathbf{B}$ *respectively. Then* $X(t) \leqslant_{st} Y(t)$ *given that* $X(0) \leqslant_{st} Y(0)$, *if and only if for all* $x \leqslant y$ *in* E, *and all increasing sets* Γ *such that* $x \in \Gamma$ *or* $y \notin \Gamma$, *we have*

$$\mathbf{e}_x\mathbf{A}\mathbf{1}_\Gamma \leqslant \mathbf{e}_y\mathbf{B}\mathbf{1}_\Gamma. \tag{5.2}$$

PROOF. Given that $X(0) \leqslant_{st} Y(0)$ implies $X(t) \leqslant_{st} Y(t)$, we then have for all $x \leqslant y$ in E and all increasing sets Γ

$$\mathbf{e}_x\exp(t\mathbf{A})\mathbf{1}_\Gamma \leqslant \mathbf{e}_y\exp(t\mathbf{A})\mathbf{1}_\Gamma. \tag{5.3}$$

Now if $x \in \Gamma$ or $y \notin \Gamma$, we still have $\mathbf{e}_x \cdot \mathbf{1}_\Gamma = \mathbf{e}_y \cdot \mathbf{1}_\Gamma$. Subtracting this from (5.3), dividing by t, and letting $t \to 0$, gives us (5.2).

To prove the converse, let $\mathbf{P}_\lambda(\mathbf{A}) = \mathbf{I} + \mathbf{A}/\lambda$ and $\mathbf{P}_\lambda(\mathbf{B}) = \mathbf{I} + \mathbf{B}/\lambda$, where $\lambda > 0$ is sufficiently large enough to make $\mathbf{P}_\lambda(\mathbf{A})$ and $\mathbf{P}_\lambda(\mathbf{B})$ stochastic matrices. It is sufficient to show that for all probability vectors $\mathbf{p} \leqslant_{st} \mathbf{q}$, that $\mathbf{p} \cdot \mathbf{P}_\lambda(\mathbf{A}) \leqslant_{st} \mathbf{q} \cdot \mathbf{P}_\lambda(\mathbf{B})$ holds. By induction, we then have $\mathbf{p} \cdot \mathbf{P}_\lambda(\mathbf{A})^n \leqslant_{st} \mathbf{q} \cdot \mathbf{P}_\lambda(\mathbf{B})^n$ for all n, and so

$$\begin{aligned}\mathbf{p} \cdot \exp(t\mathbf{A}) &= \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!}\mathbf{p} \cdot \mathbf{P}_\lambda(\mathbf{A})^n \\ &\leqslant_{st} \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!}\mathbf{q} \cdot \mathbf{P}_\lambda(\mathbf{B})^n \\ &\leqslant_{st} \mathbf{q} \cdot \exp(t\mathbf{B}),\end{aligned}$$

which proves that $X(0) \leqslant_{st} Y(0)$ implies $X(t) \leqslant_{st} Y(t)$.

Showing that $\mathbf{p} \leqslant_{st} \mathbf{q}$ implies $\mathbf{p} \cdot \mathbf{P}_\lambda(\mathbf{A}) \leqslant_{st} \mathbf{q} \cdot \mathbf{P}_\lambda(\mathbf{B})$, follows if for all $x \leqslant y$ in E, we have

$$\mathbf{e}_x\mathbf{P}_\lambda(\mathbf{A}) \leqslant_{st} \mathbf{e}_y\mathbf{P}_\lambda(\mathbf{B}). \tag{5.4}$$

This is due to Theorem 5.1 (Strassen) which allows $\mathbf{p}$ to be written as a convex combination of $\mathbf{e}_x$'s such that if each one is replaced by $\mathbf{e}_y$ for some $y \geqslant x$, we get $\mathbf{q}$. To show (5.4), we need to verify that $\mathbf{e}_x \cdot \mathbf{P}_\lambda(\mathbf{A})\mathbf{1}_\Gamma \leqslant \mathbf{e}_y \cdot \mathbf{P}_\lambda(\mathbf{B})\mathbf{1}_\Gamma$ holds only for the following three cases:

(1) $x \in \Gamma$ and $y \in \Gamma$,
(2) $x \notin \Gamma$ and $y \notin \Gamma$,
(3) $x \notin \Gamma$ and $y \in \Gamma$.

Since Γ is an increasing set, it is impossible to have $x \in \Gamma$ but $y \notin \Gamma$, so cases (1)–(3) exhaust all of the possibilities. For cases (1) and (2), $\mathbf{e}_x \cdot \mathbf{1}_\Gamma = \mathbf{e}_y \cdot \mathbf{1}_\Gamma$ so (5.4) is equivalent to (5.2), our hypothesis. For case (3) $\mathbf{e}_x \cdot \mathbf{1}_\Gamma = 0$ and $\mathbf{e}_y \cdot \mathbf{1}_\Gamma = 1$. If we require that $\lambda \geqslant |\mathbf{A}|_1 + |\mathbf{B}|_1$, then

$$\begin{aligned}\mathbf{e}_x\mathbf{P}_\lambda(\mathbf{A})\mathbf{1}_\Gamma - \mathbf{e}_y\mathbf{P}_\lambda(\mathbf{B})\mathbf{1}_\Gamma &= -1 + \frac{1}{\lambda}(\mathbf{e}_x\mathbf{A}\mathbf{1}_\Gamma - \mathbf{e}_y\mathbf{B}\mathbf{1}_\Gamma)\\ &\leqslant -1 + \frac{1}{\lambda}(|\mathbf{A}|_1 + |\mathbf{B}|_1)\\ &\leqslant 0.\end{aligned}$$

So (5.4) holds in general for $\lambda \geqslant |\mathbf{A}|_1 + |\mathbf{B}|_1$, and this completes the proof. ■

Theorems 5.2 and 5.3 allow us to strengthen condition (iii) of Theorem 4.1, for the strong ordering case.

THEOREM 5.4. $\mathbf{A} = \Delta(\boldsymbol{\alpha})(\boldsymbol{\Phi}(\mathbf{f}) - \mathbf{I})$ *is a strongly monotone Markov generator if and only if the following three conditions hold for all* $x \leqslant y$ *in* E:

(i) *If both* $\alpha(x)$ *and* $\alpha(y)$ *are nonzero, then either* $f(x) \leqslant f(y)$ *holds, or the conjunction of* $x \leqslant f(y)$ *and* $f(x) \leqslant y$.
(ii) *If* $\alpha(x) < \alpha(y)$, *then* $x \leqslant f(y)$.
(iii) *If* $\alpha(x) > \alpha(y)$, *then* $f(x) \leqslant y$.

PROOF. By Theorem 5.2, we know that $\mathbf{A}$ is strongly monotone if and only if for all $x \leqslant y$ in E we have

$$\mathbf{e}_x\mathbf{A}\mathbf{1}_\Gamma \leqslant \mathbf{e}_y\mathbf{A}\mathbf{1}_\Gamma \tag{5.5}$$

for all increasing sets Γ with $x \in \Gamma$ or $y \notin \Gamma$. Since $\mathbf{A} = \Delta(\boldsymbol{\alpha})(\boldsymbol{\Phi}(\mathbf{f}) - \mathbf{I})$, (5.5) becomes

$$\alpha(x)\mathbf{1}_\Gamma(f(x)) \leqslant \alpha(y)\mathbf{1}_\Gamma(f(y)) \tag{5.6}$$

for all increasing sets Γ such that $y \notin \Gamma$, and

$$\alpha(x)\mathbf{1}_{\Gamma^c}(f(x)) \geqslant \alpha(y)\mathbf{1}_{\Gamma^c}(f(y)) \tag{5.7}$$

where Γ^c is the complement of all increasing sets Γ with $x \in \Gamma$.

We first show that conditions (i), (ii), and (iii) imply (5.6) and (5.7). If $\alpha(x) = \alpha(y) \neq 0$, then (5.6) and (5.7) are both equivalent to showing that

$$f(x) \in \Gamma \quad \text{implies} \quad f(y) \in \Gamma \tag{5.8}$$

for all increasing sets such that $x \in \Gamma$ or $y \notin \Gamma$. Given (i), we know that (5.8) holds if $f(x) \leqslant f(y)$. Otherwise, $x \leqslant f(y)$ and $f(x) \leqslant y$. Given $x \in \Gamma$, we have $f(y) \in \Gamma$, so (5.8) is still true.

If $\alpha(x) < \alpha(y)$, then by (ii) we have $x \leqslant f(y)$. By (i), we must have $f(x) \leqslant f(y)$ or $f(x) \leqslant y$. If the former, then $\mathbf{1}_\Gamma(f(x)) \leqslant \mathbf{1}_\Gamma(f(y))$, and so (5.8) holds when $y \in \Gamma$. If the latter, then $y \notin \Gamma$ means that $f(x) \notin \Gamma$, so $1_\Gamma(f(x)) = 0$, which proves (5.6). Condition (5.7) follows from (i) since $x \leq f(y)$ and $x \in \Gamma$ mean that $f(y) \in \Gamma$, so $1_{\Gamma^c}(f(y)) = 0$. For the case $\alpha(x) > \alpha(y)$, the proof is similar.

Now we prove the converse. Once again, we consider the three cases of $\alpha(x) = \alpha(y)$, $\alpha(x) < \alpha(y)$, and $\alpha(x) > \alpha(y)$. If $\alpha(x) = \alpha(y) \neq 0$, then (5.6) and (5.7) both imply (5.8) whenever $x \in \Gamma$ or $y \notin \Gamma$ for all increasing sets Γ. Suppose that $f(x) \nleq f(y)$, then $\Gamma = \langle f(x) \rangle$ is an increasing set with $f(x) \in \Gamma$ and $f(y) \notin \Gamma$. Given condition (5.8) however, we must have $x \notin \Gamma$ and $y \in \Gamma$. From $y \in \Gamma$ follows $f(x) \leq y$. Repeating this argument with $\Gamma = \langle f(y) \rangle_*$ gives us $x \leq f(y)$.

If $\alpha(x) < \alpha(y)$, then for (5.6) to hold when $\alpha(x) \neq 0$, we must *not* have $1_\Gamma(f(x)) = 1$ and $1_\Gamma(f(y)) = 0$. This gives us condition (5.8) again, with $y \notin \Gamma$. So either $f(x) \leq f(y)$ holds or $f(x) \leq y$. Using (5.7), we must have $1_{\Gamma^c}(f(y)) = 0$ or $f(y) \in \Gamma$, whenever $x \in \Gamma$. Setting $\Gamma = \langle x \rangle$, we get $x \leq f(y)$. This proves that $\alpha(x) < \alpha(y)$ implies (i) and (ii). By a similar argument, $\alpha(x) > \alpha(y)$ implies (i) and (iii). ∎

6. Weak orderings. Whereas $\mathscr{I}_{st}(E)$ always includes a stochastic order, the same is not true for $\mathscr{I}_{wt}(E)$ or $\mathscr{I}_{wk*}(E)$. In [16], we give an example of a partially ordered space where $\mathscr{I}_{wk}(E)$ does *not* induce a stochastic order. On the other hand, we demonstrate there that $\mathscr{I}_{wk}(E)$ induces a stochastic order for many partially ordered spaces such as finite sets and lattices (even upper-semilattices). We will henceforth assume that E is a partially ordered space that admits $\mathscr{I}_{wk}(E)$ as a stochastic order.

PROPOSITION 6.1. *Let $\mathscr{I}(E)$ be a strongly separating family of increasing sets. If $\mathscr{I}(E)$ is closed under intersections, then $\mathscr{I}_{wk}(E) \subseteq \mathscr{I}(E)$. If $\mathscr{I}(E)$ is closed under unions, then $\mathscr{I}_{wk*}(E) \subseteq \mathscr{I}(E)$.*

PROOF. For any $x \in E$, let $\mathscr{I}_x = \{\Gamma | \Gamma \in \mathscr{I}(E) \text{ and } x \in \Gamma\}$ and set $\Delta = \bigcap_{\Gamma \in \mathscr{I}_x} \Gamma$. If $\mathscr{I}(E)$ is closed under intersections, then Δ belongs to $\mathscr{I}(E)$ and $x \in \Delta$. If $y \ngeq x$, then there exists some Γ in $\mathscr{I}_x$ such that $y \notin \Gamma$. From this it follows that $\Delta = \langle x \rangle$. A similar proof holds when $\mathscr{I}(E)$ is closed under unions. ∎

We now provide a characterization theorem for all weakly isotone functions.

THEOREM 6.2. *A function f is weakly isotone if and only if f is increasing on E and there is an increasing function g defined on* Dom(g), *a subset of E such that*

$$f \circ g(x) \geq x \quad \text{for } x \in \mathrm{Dom}(g), \quad g \circ f(x) \leq x \quad \text{for } x \in f^{-1}(\mathrm{Dom}(g)) \quad \text{where}$$

$$\mathrm{Dom}(g) = \begin{cases} f(E)^{\downarrow} & \text{if } E = \langle x_0 \rangle \text{ for some } x_0 \in E, \\ f(E)^{\downarrow} - \bigcap_{x \in E} \{f(x)\}^{\downarrow} & \text{otherwise.} \end{cases}$$

PROOF. We will only consider the case $E \neq \langle x_0 \rangle$ for all x_0 in E. The proof for the other case is similar. If f is weakly isotone, then for all x in E we have $f^{-1}\langle x \rangle$ equalling E, $\varnothing$, or $\langle y \rangle$ for some y in E. We also have

$$\begin{aligned} f^{-1}\langle x \rangle \text{ equals } E \text{ or } \varnothing &\Leftrightarrow E \text{ equals } \{z | f(z) \ngeq x\} \text{ or } \{x | f(z) \geq x\} \\ &\Leftrightarrow x \notin f(E)^{\downarrow} \text{ or } x \in \bigcap_{z \in E} \{f(z)\}^{\downarrow} \\ &\Leftrightarrow x \notin \mathrm{Dom}(g). \end{aligned}$$

So for all x in Dom(g), there exists some $g(x)$ in E such that $f^{-1}\langle x\rangle = \langle g(x)\rangle$. Since $g(x)$ belongs to $f^{-1}\langle x\rangle$, then $f \bigcirc g(x) \geqslant x$. If x is an element of $f^{-1}(\mathrm{Dom}(g))$, then $f^{-1}\langle f(x)\rangle = \langle g \bigcirc f(x)\rangle$ which gives us $x \geqslant g \bigcirc f(x)$ since x belongs to $f^{-1}\langle f(x)\rangle$. Finally, if $x_1 \leqslant x_2$ in Dom(g), then $f^{-1}\langle x_2\rangle \subseteq f^{-1}\langle x_1\rangle$. We then have $\langle g(x_2)\rangle \subseteq \langle g(x_1)\rangle$ and so $g(x_1) \leqslant g(x_2)$. Therefore g is an increasing function on Dom(g).

We now prove the converse. Given f, g, and Dom(g) with the properties described in the hypothesis, we want to show that $f^{-1}\langle x\rangle$ equals E, $\varnothing$, or $\langle g(x)\rangle$. As before, $f^{-1}\langle x\rangle$ equals $\varnothing$ or E if and only if $x \notin \mathrm{Dom}(g)$. If $x \in \mathrm{Dom}(g)$, then $f \bigcirc g(x) \geqslant x$ hence $g(x) \in f^{-1}\langle x\rangle$. But f is increasing so $f^{-1}\langle x\rangle$ is an increasing set and $\langle g(x)\rangle \subseteq f^{-1}\langle x\rangle$.

If $y \in f^{-1}\langle x\rangle$, then $f(y) \geqslant x$. We claim that $f(y)$ belongs to Dom(g). Since $x \in \mathrm{Dom}(g)$, then $x \notin \bigcap_{z \in E}\{f(z)\}^{\downarrow}$. The intersection of decreasing sets is a decreasing set also and $f(y) \geqslant x$ means that $f(y) \notin \bigcap_{z \in E}\{f(z)\}^{\downarrow}$. Clearly $f(y) \in f(E)^{\downarrow}$, so $f(y) \in \mathrm{Dom}(g)$. Since g is increasing, then $f(y) \geqslant x$ means that $g \bigcirc f(y) \geqslant g(x)$ and $y \geqslant g \bigcirc f(y)$ implies $y \geqslant g(x)$. Therefore $f^{-1}\langle x\rangle \subseteq \langle g(x)\rangle$, and so $f^{-1}\langle x\rangle = \langle g(x)\rangle$. ∎

Definition 6.3. Let $\mathbf{Z}(E)$ be a bounded linear operator that maps $l_1(E)$ into $l_\infty(E)$ where, for all $x \in E$, $\mathbf{e}_x\mathbf{Z}(E) = \sum_{y \leqslant x} \mathbf{e}_y$. $\mathbf{Z}(E)$ is called the *zeta function* of E. $\mathbf{A}$ is the generator for a *Möbius monotone* Markov process if for some $\lambda > 0$, there exists a bounded, positive operator $\mathbf{M}$ that maps $l_\infty(E)$ into itself such that

$$\mathbf{P}_\lambda(\mathbf{A}) \cdot \mathbf{Z}(E) = \mathbf{Z}(E) \cdot \mathbf{M}.$$

Kester [10] discusses Möbius monotonicity for the case of E being a finite set. In Rota [19], we see that when E is a finite set, then the matrix inverse of $Z(E)$ is referred to as the *Möbius function*, hence the name for this type of monotonicity. We now show that it is a special case of weak monotonicity.

Theorem 6.4. *Möbius monotonicity implies weak monotonicity, but the converse does not hold.*

Proof. Expanding $\exp(t\mathbf{A})$ in terms of $\mathbf{P}_\lambda(\mathbf{A})$ gives us

$$\exp(t\mathbf{A}) \cdot \mathbf{Z}(E) = \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \mathbf{P}_\lambda(\mathbf{A})^n \cdot \mathbf{Z}(E)$$

$$= \mathbf{Z}(E) \cdot \sum_{n=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^n}{n!} \mathbf{M}^n.$$

$\mathbf{M}$ is positive, so the last sum is positive. We then have that $\mathbf{p}\mathbf{Z}(E) \leqslant \mathbf{q}\mathbf{Z}(E)$ implies $\mathbf{p}\exp(t\mathbf{A})\mathbf{Z}(E) \leqslant \mathbf{q}\exp(t\mathbf{A})\mathbf{Z}(E)$. Since $\mathbf{p} \leqslant_{wk} \mathbf{q}$ if and only if $\mathbf{p}\mathbf{Z}(E) \leqslant \mathbf{q}\mathbf{Z}(E)$, we have shown that $\mathbf{A}$ is weakly monotone.

We now show by counterexample that weak monotonicity does not imply Möbius monotonicity. Let $E = \{a, b, c\}$ where $a \leqslant b$, $c \leqslant b$, but a and c are incomparable. Define a Markov process on E with generator $\mathbf{A} = \alpha(\mathbf{\Phi}(f) - \mathbf{I})$ where $\alpha > 0$ and $f(a) = f(b) = f(c) = a$. Since $f^{-1}(\Gamma)$ equals E or $\varnothing$ for any subset Γ of E, then f is weakly monotone. By Theorem 4.1, $\mathbf{A}$ is then a weakly monotone generator. Let $\mathbf{e}_1 = \mathbf{e}_a$, $\mathbf{e}_2 = \mathbf{e}_b$, $\mathbf{e}_3 = \mathbf{e}_c$ be the unit basis vectors of $l_1(E)$. When then have

$$A = \begin{bmatrix} 0 & 0 & 0 \\ \alpha & -\alpha & 0 \\ \alpha & 0 & -\alpha \end{bmatrix}, \quad \mathbf{Z}(E) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Z}(E)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Computing with these matrices, $\mathbf{M}$ must have the form

$$\mathbf{M} = \mathbf{Z}(E)^{-1}\mathbf{P}_\lambda(\mathbf{A})\mathbf{Z}(E)$$

$$= \frac{1}{\lambda}\mathbf{Z}(E)^{-1}\mathbf{A}\mathbf{Z}(E) + \mathbf{I}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ \frac{-\alpha}{\lambda} & 1 - \frac{\alpha}{\lambda} & 0 \\ \frac{\alpha}{\lambda} & 0 & 1 - \frac{\alpha}{\lambda} \end{bmatrix}.$$

With both α and λ positive, it is impossible for $\mathbf{M}$ to be a positive matrix. Therefore $\mathbf{A}$ is weakly monotone, but not Möbius monotone. ■

In conclusion, we remark that all of the results in this section can be dualized for the weak* ordering. Merely let E^* be the partially ordered space dual to E (see Birkhoff [3, p. 3]). The map that sends a subset of E to its compliment, induces a bijection between $\mathscr{I}_{wk^*}(E)$ and $\mathscr{I}_{wk}(E^*)$.

7. **Time-inhomogeneous Markov processes.** We now extend the comparison results of the previous sections to time-inhomogeneous Markov processes. This section was motivated by work that the author did for analyzing the time-dependent $M/M/1$ queue (see Theorem 10.1 of [14]). Let $\{X(t_0, s)|t_0 \leqslant s \leqslant t\}$ denote such a process with state space E that has evolved from time t_0 to time t with $X(t_0) = X(t_0, t_0)$. If $\mathbf{p}(t_0, s)$ represents the probability vector for $X(t_0, s)$ (resp. $\mathbf{p}(t_0)$ for $X(t_0)$), we will consider processes that satisfy the integral version of the Kolmogorov forward equation namely

$$\mathbf{p}(t_0, t) = \mathbf{p}(t_0) + \int_{t_0}^{t} \mathbf{p}(t_0, s)\mathbf{A}(s)\, ds.$$

The process $\{X(t_0, s)|t_0 \leqslant s \leqslant t\}$ is governed by a *family of generators* $\{A(s)|t_0 \leqslant s \leqslant t\}$. Each $\mathbf{A}(s)$ is the generator of a time-homogeneous Markov process in its own right. If the original process was time-homogeneous ($\mathbf{A}(s) = \mathbf{A}$ for all $t_0 \leqslant s \leqslant t$), we would then have

$$\mathbf{p}(t_0, t) = \mathbf{p}(t_0)\exp((t - t_0)\mathbf{A}).$$

We want a similar representation for time-inhomogeneous processes. Unfortunately,

$$\mathbf{p}(t_0, t) \neq \mathbf{p}(t_0)\exp\left(\int_{t_0}^{t} \mathbf{A}(s)\, ds\right)$$

in general because $\mathbf{A}(s)$ and $\mathbf{A}(s')$ may not commute for $s \neq s'$. We can obtain our desired representation if we appeal to the theory of product integration (see Dollard and Friedman [5]). First, let our family of generators $\{\mathbf{A}(s)|t_0 \leqslant s \leqslant t\}$ satisfy the following three conditions:

(P1). For all s in $[t_0, t]$, $\mathbf{A}(s)$ is a bounded operator on $l_1(E)$.

(P2). $\mathbf{A}$ as an operator-valued function on $[t_0, t]$ is strongly measurable.

(P3). The upper integral of $|\mathbf{A}(\cdot)|$ on $[t_0, t]$ is finite.

We then get the following theorem, which is Theorem 5.3 (Chapter 3) in Dollard and Friedman [5].

THEOREM 7.1 (Dollard and Friedman). *When conditions* (P1), (P2), *and* (P3) *are satisfied, the product integral* $E_A(t_0, t)$ *exists and is the unique solution of*

$$\mathbf{E_A}(t_0, t) = \mathbf{I} + \int_{t_0}^{t} \mathbf{E_A}(t_0, s)\mathbf{A}(s)\, ds \quad \text{or} \quad \mathbf{E_A}(t_0, t) = \mathbf{I} + \int_{t_0}^{t} \mathbf{A}(s)\mathbf{E_A}(s, t)\, ds$$

where integration is defined by the strong operator topology.

Now let $X(t_0, t)$ be a time-inhomogeneous Markov process with a family of generators satisfying conditions (P1), (P2), and (P3). We can then write the probability vector for $X(t_0, t)$ as $\mathbf{p}(t_0, t) = \mathbf{p}(t_0)\mathbf{E_A}(t_0, t)$.

DEFINITION 7.2. The family of generators $\{\mathbf{A}(s) | t_0 \leqslant s \leqslant t\}$ is *$\mathscr{I}(E)$-monotone on* $[t_0, t]$ if for all probability vectors $\mathbf{p}$ and $\mathbf{q}$ in $l_1(E)$ we have

$$\mathbf{p} \leqslant_{\mathscr{I}} \mathbf{q} \quad \text{implies that} \quad \mathbf{p}\mathbf{E_A}(s, t) \leqslant_{\mathscr{I}} \mathbf{q}\mathbf{E_A}(s, t)$$

for all s in $[t_0, t]$.

Correspondingly, we will say that $X(t_0, t)$ is $\mathscr{I}(E)$-monotone on $[t_0, t]$ if $\{\mathbf{A}(s) | t_0 \leqslant s \leqslant t\}$ is $\mathscr{I}(E)$-monotone on $[t_0, t]$.

THEOREM 7.3. *Let $X(t_0, t)$ and $Y(t_0, t)$ be two Markov processes with state space E. We are given that either $X(t_0, t)$ or $Y(t_0, t)$ is $\mathscr{I}(E)$-monotone on $[t_0, t]$. If $\{\mathbf{A}(s) | t_0 \leqslant s \leqslant t\}$ and $\{\mathbf{B}(s) | t_0 \leqslant s \leqslant t\}$ are their respective family of generators, $X(t_0) \leqslant_{\mathscr{I}} Y(t_0)$, and $\mathbf{A}(s) \leqslant_{\mathscr{I}} \mathbf{B}(s)$ for almost all s in $[t_0, t]$, then $X(t_0, t) \leqslant_{\mathscr{I}} Y(t_0, t)$.*

PROOF. The argument is identical to Theorem 3.4, once we establish the following identity

$$\mathbf{E_A}(t_0, t) - \mathbf{E_B}(t_0, t) = \int_{t_0}^{t} \frac{d}{ds}\mathbf{E_A}(t_0, s)\mathbf{E_B}(s, t)\, ds$$

$$= \int_{t_0}^{t} \mathbf{E_A}(t_0, s)(\mathbf{A}(s) - \mathbf{B}(s))\mathbf{E_B}(s, t)\, ds$$

and make the observation that $\mathbf{E_A}(t_0, s)$ is a positive operator for all s in $[t_0, t]$. ∎

THEOREM 7.4. *Let $X(t_0, t)$ and $Y(t_0, t)$ be two Markov processes with state spaces E' and E respectively, having corresponding families of generators $\{\mathbf{A}(s) | t_0 \leqslant s \leqslant t\}$ and $\{\mathbf{B}(s) | t_0 \leqslant s \leqslant t\}$. If $Y(t_0, t)$ is $\mathscr{I}(E)$-monotone on $[t_0, t]$, f maps E' into E, $f(X(t_0)) \leqslant_{\mathscr{I}} Y(t_0)$, and $\mathbf{A}(s)\mathbf{\Phi}(\mathbf{f}) \leqslant_{\mathscr{I}} \mathbf{\Phi}(\mathbf{f})\mathbf{B}(s)$ for almost all s in $[t_0, t]$, then $f(X(t_0, t)) \leqslant_{\mathscr{I}} Y(t_0, t)$.*

PROOF. Once again, the proof here is identical to its counterpart, Theorem 3.5 after establishing the following identity

$$\mathbf{E_A}(t_0, t)\mathbf{\Phi}(f) - \mathbf{\Phi}(f)\mathbf{E_B}(t_0, t) = \int_{t_0}^{t} \frac{d}{ds}\mathbf{E_A}(t_0, s)\mathbf{\Phi}(f)\mathbf{E_B}(s, t)\, ds$$

$$= \int_{t_0}^{t} \mathbf{E_A}(t_0, s)(\mathbf{A}(s)\mathbf{\Phi}(f) - \mathbf{\Phi}(f)\mathbf{B}(s))\mathbf{E_B}(s, t)\, ds$$

from this the theorem follows. ∎

Now there is left only the task of constructing families of $\mathscr{I}(E)$-monotone generators. The following theorem enables us to use the results of Theorem 4.1.

THEOREM 7.5. *$\{\mathbf{A}(s) | t_0 \leqslant s \leqslant t\}$ is $\mathscr{I}(E)$-monotone on $[t_0, t]$ if each individual $\mathbf{A}(s)$ is an $\mathscr{I}(E)$-monotone Markov generator for almost all s in $[t_0, t]$.*

PROOF. It is enough to show that $\mathbf{p} \leqslant_{\mathscr{I}} \mathbf{q}$ implies $\mathbf{p}\mathbf{E}_{\mathbf{A}}(t_0, t) \leqslant_{\mathscr{I}} \mathbf{q}\mathbf{E}_{\mathbf{A}}(t_0, t)$. For each n, let $\{t_i(n) | i = 0, 1, \ldots, n\}$ be a partition of $[t_0, t]$ where $t_0 = t_0(n) < t_1(n) < \cdots < t_n(n) = t$ with $\Delta t_i(n) = t_i(n) - t_{i-1}(n)$ for $i = 1, \ldots, n$ and $\lim_{n\to\infty} \max_{1 \leqslant i \leqslant n} \Delta t_i(n) = 0$. Now let $\{\mathbf{A}_n(s) | t_0 \leqslant s \leqslant t\}$ be a family of generators such that

$$\mathbf{A}_n(s) = \mathbf{A}_n^{(i)} = \frac{1}{\Delta t_i(n)} \int_{t_{i-1}(n)}^{t_i(n)} \mathbf{A}(r)\, dr$$

for $t_{i-1}(n) < s \leqslant t_i(n)$. Each $\mathbf{A}_n(\cdot)$ is a step function on $[t_0, t]$, so

$$\mathbf{E}_{\mathbf{A}_n}(t_0, t) = \prod_{i=1}^{n} \exp\left(\Delta t_i(n) \mathbf{A}_n^{(i)}\right).$$

By Theorems 3.2 and 5.1 (Chapter 3) of Dollard and Friedman [5], we have as $n \to \infty$, $\mathbf{E}_{\mathbf{A}_n}(t_0, t) \to \mathbf{E}_{\mathbf{A}}(t_0, t)$ with respect to the strong operator topology.

By Theorem 4.1, each $\mathbf{A}_n^{(i)}$ is $\mathscr{I}(E)$-monotone. Given $\mathbf{p} \leqslant_{\mathscr{I}} \mathbf{q}$, we then have by induction $\mathbf{p}\mathbf{E}_{\mathbf{A}_n}(t_0, t) \leqslant_{\mathscr{I}} \mathbf{q}\mathbf{E}_{\mathbf{A}_n}(t_0, t)$. Taking the limit as $n \to \infty$ gives us the desired result. ■

8. Applications to the Jackson network. We will define a Jackson network with single server queues by constructing its generator.

$$E = Z_+^N, \quad I = \{(i, j) | 0 \leqslant i, j \leqslant N\},$$

$$a_{(i,j)} = \begin{cases} 0, & i = j = 0, \\ \lambda_j, & i = 0,\ j \neq 0, \\ \mu_i q_i, & i \neq 0,\ j = 0, \\ \mu_i p_{ij}, & i \neq 0,\ j \neq 0, \end{cases}$$

$$f_{(i,0)}(\ldots, n_i, \ldots) = (\ldots, (n_i - 1)^+, \ldots),$$

$$f_{(0,j)}(\ldots, n_j, \ldots) = (\ldots, n_j + 1, \ldots),$$

$$f_{(i,j)}(\ldots, n_i, \ldots, n_j) = \begin{cases} (\ldots, n_i - 1, \ldots, n_j + 1, \ldots), & n_i > 0, \\ (\ldots, n_i, \ldots, n_j, \ldots), & n_i = 0, \end{cases}$$

and finally

$$\mathbf{A} = \sum_{(i,j)\in I} a_{(i,j)}\left(\Phi(\mathbf{f}_{(i,j)}) - \mathbf{I}\right).$$

We will let $\mathbf{Q}(t)$ denote the associated queue length process. Notice that each $f_{(i,j)}$ is an increasing function with respect to the canonical (componentwise) partial ordering on Z_+^N. For any partition $\mathscr{P}$ of the set $\{1, 2, \ldots, N\}$, we define $\mathbf{A}_{\mathscr{P}}$ to be the generator for an associated Jackson network with queue length vector $\mathbf{Q}_{\mathscr{P}}(t)$, where $\mathbf{A}_{\mathscr{P}} = \sum_{(i,j)\in I} a_{(i,j)}(\Phi_{\mathscr{P}}^{ij}(f) - \mathbf{I})$ using the notation above, with

$$\Phi_{\mathscr{P}}^{ij}(f) = \begin{cases} \Phi(f_{(i,0)}) + \Phi(f_{(0,j)}) - \mathbf{I}, & \mathscr{P} \text{ separates } i \text{ and } j, \\ \Phi(f_{(i,j)}), & \text{otherwise.} \end{cases}$$

If J is a subset of $\{1, 2, \ldots, N\}$, let Z_+^J be the set of nonnegative integer $|J|$-tuples with the same partial ordering. So, for example, $Z_+^N = \times_{J \in \mathscr{P}} Z_+^J$. Now for every partition $\mathscr{P}$, define

$$\mathscr{I}_{\mathscr{P}}(Z_+^N) = \left\{ \underset{J \in \mathscr{P}}{\times} \Gamma_J | \Gamma_J \subseteq Z_+^J \text{ and } \Gamma_J \text{ is increasing} \right\}.$$

Clearly $\mathscr{I}_{wk}(Z_+^N) \subseteq \mathscr{I}_{\mathscr{P}}(Z_+^N)$, so $\mathscr{I}_{\mathscr{P}}(Z_+^N)$ induces a stochastic ordering which we will denote by $\leqslant_{\mathscr{P}}$. Note that if $\mathscr{P}^*$ is a refinement of $\mathscr{P}$, then $X \leqslant_{\mathscr{P}} Y$ implies that $X \leqslant_{\mathscr{P}^*} Y$. The motivation for $\mathbf{Q}_{\mathscr{P}}(t)$ is as follows. The nodes of the network are partitioned by $\mathscr{P}$. For the new process, we have the same activities. The only difference is that the event of customers leaving node i and entering node j is forbidden. Instead such customers upon leaving node i, leave the entire network. As compensation, node j receives an independent Poisson stream of customers at rate $a_{(i,j)} = \mu_i p_{ij}$. Thus the original output from i to j is replaced by a Poisson stream into j that matches the output rate when node i is not idle. It follows from Theorem 4.1 that $\mathbf{Q}(t)$ as defined here is strongly monotone. For $\mathbf{Q}_{\mathscr{P}}(t)$, we can state a deeper result, which is used implicitly in the theorem that follows.

PROPOSITION 8.1 *$\mathbf{Q}_{\mathscr{P}}(t)$ is $\mathscr{P}$-monotone.*

Let $\mathscr{P}^*$ be a refinement of $\mathscr{P}$. Given the construction of $\mathbf{A}_{\mathscr{P}}$, we see that $(\mathbf{A}_{\mathscr{P}})_{\mathscr{P}^*} = \mathbf{A}_{\mathscr{P}^*}$. Intuitively, one would expect the queue length process $\mathbf{Q}_{\mathscr{P}^*}(t)$ to be "larger" than the queue length process $\mathbf{Q}_{\mathscr{P}}(t)$. In [12], the author proved that this is indeed the case *provided* that the proper type of stochastic ordering is employed. We now restate the results of [12] below using notation consistent with this paper.

THEOREM 8.2 (Massey). *If $\mathscr{P}^*$ is a refinement of $\mathscr{P}$, and $\mathbf{Q}_{\mathscr{P}}(0) \leqslant_{\mathscr{P}^*} \mathbf{Q}_{\mathscr{P}^*}(0)$, then*

$$\mathbf{Q}_{\mathscr{P}}(t) \leqslant_{\mathscr{P}^*} \mathbf{Q}_{\mathscr{P}^*}(t) \tag{8.1}$$

for all $t \geqslant 0$. Moreover, if $\mathbf{Q}_{\mathscr{P}}(0) = \mathbf{Q}_{\mathscr{P}^*}(0)$, then $\mathbf{Q}_{\mathscr{P}}(t) \leqslant_{\mathscr{P}} \mathbf{Q}_{\mathscr{P}^*}(t)$ if and only if $\mathbf{Q}_{\mathscr{P}}(t)$ and $\mathbf{Q}_{\mathscr{P}^*}(t)$ are identical in distribution.

For any network that can be nontrivially partitioned, we see that the relation (8.1) *cannot* be extended to a stronger stochastic ordering like $\leqslant_{\mathscr{P}}$. In turn, this means that there is no relation like (8.1) with respect to $\leqslant_{st}$. Or equivalently we see that this is a relation that does not extend to a sample path comparison between the two processes.

These results by no means exhaust the possibilities for stochastic ordering on Jackson networks. There are other partial orderings on Z_+^N that may be appropriate. A special case of Theorem 14 in Whitt [26] for non-Markovian series networks with multiservers and possibly finite waiting rooms, reduces to the following result for Jackson series networks.

THEOREM 8.3 (Whitt). *For m and n in Z_+^N, let $\mathbf{m} \leqslant_* \mathbf{n}$ denote $\sum_{i=1}^j m_i \leqslant \sum_{i=1}^j n_i$ for all $j = 1, \ldots, N$. Let $\mathbf{Q}_1(t)$ and $\mathbf{Q}_2(t)$ be two queue length processes for series Jackson networks with $\lambda_1(1) \leqslant \lambda_2(2)$ but $\mu_i(1) \geqslant \mu_i(2)$ for all $i = 1, \ldots, N$. If $\leqslant_{st}$ denotes the strong ordering induced by $\leqslant_*$, then $\mathbf{Q}_1(0) \leqslant_{st} \mathbf{Q}_2(0)$ implies that $\mathbf{Q}_1(t) \leqslant_{st} \mathbf{Q}_2(t)$.*

We also note that all of the theorems in this section hold for time varying Jackson networks. This follows from using Theorems 7.3 and 7.5.

Acknowledgements. The author is indebted to Ward Whitt of AT & T Bell Laboratories for his helpful comments and especially for providing A. Kester's thesis. Thanks are also due to Tepper Gill of Howard University for giving me his insight into the theory of product integration.

References

[1] Bawa, Vijay S. (1982). Stochastic Dominance: A Research Bibliography. *Management Sci.* **28**, **6**. June.

[2] Billingsley, Patrick (1968). *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York.

[3] Birkhoff, Garrett, (1948). *Lattice Theory*, Amer. Math. Soc. Colloq. Publ. **25**.

[4] Daley, D. J. (1968). Stochastically Monotone Markov Chains. *Z. Wahrsch. Verw. Gebiete*. **10**, 305–317.

[5] Dollard, J. D. and Friedman, C. N. (1984). *Product Integration*. Cambridge University Press, Cambridge.

[6] Kalmykov, G. I. (1962). On the Partial Ordering of One-Dimensional Markov Processes. *Theory Prob. Appl.* **7**, 456–459.

[7] Kamae, T. and Krengel, U. (1978). Stochastic Partial Ordering. *Ann. Probab.* **6**, **6**, 1044–1049.

[8] ______, ______, and O'Brien, G. L. (1977). Stochastic Inequalities on Partially Ordered Spaces. *Ann. Probab.* **5**, 899–912.

[9] Keilson, J. and Kester, A. (1977). Monotone Matrices and Monotone Markov Processes, *Stochastic Process. Appl.* **5**, **3**, 231–241.

[10] Kester, Adrianus J. M. (1977). Preservation of Cone Characterizing Properties in Markov Chains, Ph.D. Thesis, University of Rochester.

[11] Kirstein, B. M. (1976). Monotonicity and Comparability of Time-Homogeneous Markov Processes with Discrete State Space. *Math. Operationsforsch. Statist.* **1**, S. 151–168.

[12] Massey, William A. (1986). A Family of Bounds for the Transient Behaviors of a Jackson Network. *J. Appl. Probab.* **23**, 543–549.

[13] ______. (1984). An Operator Analytic Approach to the Jackson Network. *J. Appl. Probab.* **21**, 379–393.

[14] ______. (1985). Asymptotic Analysis of the Time-Dependent $M/M/1$ Queue. *Math. Oper. Res.* 10, 2, 305–327.

[15] ______. (1984). Open Networks of Queues: Their Algebraic Structure and Estimating Their Transient Behavior. *Adv. Appl. Probab.* **16**, **1**, 176–201.

[16] ______. Tail Distribution Functions on Partially Ordered Spaces. (to appear).

[17] O'Brien, G. (1982). Determining Classes and Independence. *Statist. and Probab. Lett.* **1**, 57–59.

[18] Preston, C. J. (1974). A Generalization of the FKG Inequalities. *Comm. Math. Phys.* **36**, 233–241.

[19] Reed, M. and Simon, B. (1972). *Functional Analysis. Vol.* I. Academic Press, New York and London.

[20] Rota, Gian-Carlo (1964). On the Foundations of Combinatorial Theory. I. Theory of Möbius Functions. *Z. Wahrsch. Verw. Gebiete*. **2**, 340–368.

[21] Sonderman, D. (1980). Comparing Semi-Markov Processes, *Math. Oper. Res.* **5**, 110–119.

[22] Stoyan, Dietrich (1983). *Comparison Methods for Queues and Other Stochastic Models*. John Wiley & Sons, New York.

[23] Strassen, V. (1965). The Existence of Probability Measures with Given Marginals. *Ann. Math. Statist.* **36**, 423–439.

[24] Tong, Y. L. (1980). *Probability Inequalities in Multivariate Distributions*. Academic Press, New York.

[25] Whitt, Ward (1985). Blocking When Service is Required from Several Facilities Simultaneously. *AT & T Bell Lab. Tech. J.* **64**, 1807–1856.

[26] ______. (1981). Comparing Counting Processes and Queues, *Adv. Appl. Probab.* **13**, **1**, 207–220.

[27] ______. (1986). Comparing Non-Markov Processes. *Math. Oper. Res.* **11**, 608–618.

AT & T BELL LABORATORIES, 600 MOUNTAIN AVENUE, MURRAY HILL, NEW JERSEY 07974