MATHEMATICS OF OPERATIONS RESEARCH
Vol. 5, No. 1, February 1980
Printed in U.S.A.

SOME USEFUL FUNCTIONS FOR FUNCTIONAL LIMIT THEOREMS*†

WARD WHITT

Bell Laboratories

Many useful descriptions of stochastic models can be obtained from functional limit theorems (invariance principles or weak convergence theorems for probability measures on function spaces). These descriptions typically come from standard functional limit theorems via the continuous mapping theorem. This paper facilitates applications of the continuous mapping theorem by determining when several important functions and sequences of functions preserve convergence. The functions considered are composition, addition, composition plus addition, multiplication, supremum, reflecting barrier, first passage time and time reversal. These functions provide means for proving new functional limit theorems from previous ones. These functions are useful, for example, to establish the stability or continuity of queues and other stochastic models.

1. Introduction. Stochastic processes of interest in operations research models such as queue length processes can often be represented as functions of more basic stochastic processes such as random walks and renewal processes. Consequently, limit theorems for sequences of stochastic processes in operations research models can often be obtained from existing limit theorems for the more basic processes by showing that the connecting functions preserve convergence. This method for proving limit theorems is described in Billingsley (1968) and is well known. The purpose of this paper is to investigate several functions which frequently arise in operations research models. The functions considered are composition, addition, composition plus addition, multiplication, supremum, reflecting barrier, first passage time and time reversal. We find general conditions under which these functions preserve convergence. For example, suppose $X_n \equiv \{X_n(t), t \geq 0\}$ and $Y_n \equiv \{Y_n(t), t \geq 0\}$ are stochastic processes which converge jointly in distribution as $n \to \infty$. Under what conditions does their sum $(X_n + Y_n) \equiv \{X_n(t) + Y_n(t), t \geq 0\}$ also converge as $n \to \infty$? If by "convergence" we mean weak convergence of random elements of the function space $D[0, \infty)$ with Skorohod's (1956) J_1 topology as described in Billingsley (1968) and Lindvall (1973), then the appropriate connecting function is addition mapping $D[0, \infty) \times D[0, \infty)$ into $D[0, \infty)$. Addition is known to preserve convergence when both limit processes have continuous paths w.p.1., but not in general; see Billingsley (1968, Problem 3, p. 123). We determine sufficient conditions for addition to preserve convergence. For example, it suffices for the two limit processes to be independent with one being continuous in probability; see §4.

While our interest here is in stochastic limit theorems, we rarely mention probability measures or stochastic processes in this paper. This is because we can apply the continuous mapping theorem to translate the question of preserving stochastic con-

* Received April 14, 1978; revised January 19, 1979.
AMS 1970 subject classification. Primary 60B10. Secondary 60F05, 60F15.
IAOR 1973 subject classification. Main: Stochastic processes.
OR/MS Index 1978 subject classification. Primary 561 Probability.
Key words. Continuous mapping theorems; preservation of convergence under mappings; functional limit theorems; invariance principles; weak convergence; the function space D; Skorohod topologies; stability of stochastic models; composition; random time transformations; subordinated stochastic processes; maxima; first passage times; time reversal.
† This research was done in the School of Organization and Management, Yale University, supported by National Science Foundation Grants GK-27866 and GK-38149.

0364-765X/80/0501/0067$01.25

vergence to the question of preserving deterministic convergence on the underlying sample space. Suppose X_n, $n \geq 1$, and X are random variables with values in a separable metric space and f_n, $n \geq 1$, and f are Borel measurable functions mapping this separable metric space into another.

CONTINUOUS MAPPING THEOREM (CMT).

(i) *If* $X_n \to X$ *and* f *is continuous almost surely with respect to the distribution of* X, *then* $f(X_n) \to f(X)$.

(ii) *If* $X_n \to X$ *and* $f_n(x_n) \to f(x)$ *for all* $x \in A$ *and* $\{x_n\}$ *with* $x_n \to x$ *for some* A *with* $P(X \in A) = 1$, *then* $f_n(X_n) \to f(X)$.

The mode of convergence above has not been specified because the CMT holds for convergence w.p.l., convergence in probability, and relative compactness as well as the more familiar convergence in distribution (weak convergence) covered by Theorems 5.1 and 5.5 in Billingsley (1968). For relative compactness, with a slight abuse of notation, $X_n \to X$ means that w.p.l. every subsequence of $\{X_n\}$ has a convergent subsequence and its set of limit points is the set X. These four modes of convergence are used in the classical and functional strong laws of large numbers, weak laws of large numbers, laws of the iterated logarithm and central limit laws, respectively.

Since the CMT is so important, we briefly discuss its proof and interpretation. There are several ways to prove the CMT associated with each mode of convergence, but there is one way which we believe helps the understanding. The revealing proof in each case is to represent the mode of convergence in terms of w.p.l. convergence and then apply the CMT associated with w.p.l. convergence. Since the CMT associated with w.p.l. convergence is obvious, the issue for each mode is the representation in terms of w.p.l. convergence. For convergence in distribution, the vehicle is Skorohod's (1956, §3) representation theorem; p. 7 of Billingsley (1971), Dudley (1968), Wichura (1970) and references there. Let $\Rightarrow$ denote both weak convergence of probability measures and weak convergence (convergence in distribution) of random variables. Let $\sim$ denote equality in distribution. As before, let the random variables take values in a separable metric space.

SKOROHOD REPRESENTATION THEOREM. *If* $X_n \Rightarrow X$, *then there exists a probability space supporting random variables* Y_n, $n \geq 1$, *and* Y *such that* $Y_n \sim X_n$ *for all* n, $Y \sim X$ *and* $Y_n \to Y$ *w.p.l.*

It is easy to apply the Skorohod Representation Theorem to prove the CMT for weak convergence. Consider form (ii) of the CMT. If $X_n \Rightarrow X$, the representation theorem gives $Y_n \to Y$ w.p.l. with $Y_n \sim X_n$, $n \geq 1$, and $Y \sim X$. The obvious w.p.l. CMT gives $f_n(Y_n) \to f(Y)$ w.p.l. under the specified conditions. As a consequence, $f_n(Y_n) \Rightarrow f(Y)$. Since $Y_n \sim X_n$ and $Y \sim X$, $f_n(Y_n) \sim f_n(X_n)$ and $f(Y) \sim f(X)$. Hence, $f_n(X_n) \Rightarrow f(X)$ too. A similar argument applies to convergence in probability because a sequence $\{X_n\}$ converges in probability to X if and only if every subsequence of $\{X_n\}$ has a further subsequence converging w.p.l. to X; Theorem 4.4 of Tucker (1967). The theorem in Tucker (1967) is for real-valued random variables, but it extends easily to separable metric spaces because $X_n \to X$ w.p.l. (in probability) if and only if $d(X_n, X) \to 0$ w.p.l. (in probability), where d is the metric. This means that these modes of convergence are characterized by the convergence of associated real-valued random variables. Finally, relative compactness by definition involves w.p.l. convergence.

The CMT obviously operates with greater force when the converging random variables X_n and the limit X have values in a general space such as a function space. Then X_n and X are stochastic processes and many random quantities of interest can be represented as a measurable functions which are continuous almost surely with respect to typical limit processes. It is for precisely this reason that much attention in recent years has been devoted to proving functional limit theorems, invariance principles or stochastic limit theorems in function space settings. One of the function spaces most frequently used is D, the space of all right-continuous functions on a subinterval of the real line which have limits from the left, endowed with Skorohod's (1956) J_1 topology, chapter 3 of Billingsley (1968) and Lindvall (1973). There are now many results of the form $X_n \to X$ in D. With these results in hand and many more forthcoming, it is natural to focus on the other hypotheses in the CMT. For useful functions f_n, $n \geq 1$, and f on D or $D \times D$, it is natural to ask when f is continuous and when $f_n(x_n) \to f(x)$ for all $x_n \to x$. This paper addresses this question. The results are of the form: if x is restricted to a particular subset of D or $D \times D$, then $f_n(x_n) \to f(x)$ whenever $x_n \to x$. In stochastic applications it often requires a little work to identify the appropriate connecting function. The technique is illustrated at the end of §5. Since the range of the functions we consider is also D, functions are mapped into functions, stochastic processes are mapped into stochastic processes, and functional limit theorems are mapped into functional limit theorems. Limits for real-valued random variables, which are often desired in applications, can be obtained later from the CMT using projections or other real-valued functions.

All the functions here have been used to prove limit theorems for queues; see Iglehart and Whitt (1970), Kennedy (1972) and Whitt (1974, 1974a). The earlier results for queues plus various extensions follow easily from the present paper. The results here are especially useful for establishing continuity or stability of stochastic models because then the limiting stochastic processes often do not have continuous sample paths, cf. Kennedy (1972, 1978), Whitt (1974a) and Zolotarev (1978). In fact, many of the functions here are treated in Billingsley (1968, §§5, 11, 17) in conjunction with limit processes such as Brownian motion which have continuous sample paths w.p.l. For other applications in which limit processes do not necessarily have continuous paths and the generality here is important, see Bingham (1973), de Haan and Resnick (1978), Goldie (1977), Lindberger (1978), Serfozo (1973) and Wichura (1974).

The function space D here is slightly more general than in chapter 3 of Billingsley (1968) because the domain of the functions is allowed to be an arbitrary subinterval of the real line instead of a compact subinterval and the range is allowed to be an arbitrary complete separable metric space instead of the real line. The minor gap between this setting and chapter 3 of Billingsley (1968) is filled in §2. The approach here is different from that of Stone (1963) or Lindvall (1973) so should be of independent interest.

For the most part, the topology on D is Skorohod's (1956) J_1 topology, as in chapter 3 of Billingsley (1968), but his M_1 topology is also used in the study of suprema and first passage times in §§5 and 6. It turns out that the essential properties of the J_1 topology carry over to the M_1 topology, but discussion of the M_1 topology is minimized in this paper. The M_1 topology is introduced only when results unavailable (J_1) are available (M_1). However, the investigation of the functions here has been extended to all the other Skorohod (1956) topologies by Pomarede (1976–1976a).

Previous versions. This paper, a revision of several shorter papers written in 1970–1971, has been distributed since 1973 as a Yale University Technical Report entitled, "Continuity of Several Functions on the Function Space D." It is referred to this way in several of the references.

2. The function space D with Skorohod's J_1 topology. The purpose of this section is to generalize the function space $D[0, 1]$ discussed in chapter 3 of Billingsley (1968) by allowing the domain of functions to be an arbitrary subinterval of the real line and the range to be an arbitrary complete separable metric space. The results are natural extensions of the results obtained for $D[0, \infty)$ by Stone (1963) and Lindvall (1973), but the methods here are different. Of particular interest is the simple proof of the theorem characterizing weak convergence of probability measures on the function space in terms of weak convergence of image measures associated with restrictions to compact subintervals (Theorem 2.8). We begin by providing preliminary facts about the topology and the Borel σ-field.

Let T be a subinterval of the real line. The endpoints of T can be finite or infinite and, if finite, open or closed. Let S be a CSMS (complete separable metric space) with metric m. Let $D \equiv D(T) \equiv D(T, S)$ be the set of all right-continuous S-valued functions on T with limits from the left. Let D have Skorohod's (1956) J_1 topology or its natural extension to noncompact intervals: a net $\{x_\alpha\}$ converges to x in $D(T)$ if the restrictions of x_α converge to the restriction of x in $D([a, b])$ for each compact interval $[a, b] \subseteq T$ such that a and b are continuity points of x or endpoints of T. This mode of convergence agrees with previous extensions of the J_1 topology to $T = [0, \infty)$ by Stone (1963) and Lindvall (1973). However, this is by no means the only mode of convergence worth considering. It is often desirable to require more at open boundary points, but we do not here. (Continuity issues associated with a stronger topology have recently been investigated by Bauer (1978).)

For $T = [a, b]$, let ρ be the uniform metric on $D([a, b])$, e the identity map on T, Λ the set of increasing homeomorphisms of T, $\circ$ the composition map, and $a \vee b = \max\{a, b\}$. If

$$d(x, y) = \inf_{\lambda \in \Lambda} \{\rho(\lambda, e) \vee \rho(x, y \circ \lambda)\}, \tag{2.1}$$

where ρ is regarded as the uniform metric on both $D(T, T)$ and $D(T, S)$, then d is the usual incomplete metric inducing the J_1 topology on $D([a, b])$; p. 111 of Billingsley (1968). It is sometimes convenient to have another representation for d which shows that there is considerable freedom in the choice of λ. For $n \geq 1$, let $\mathcal{C}_n$ be the collections of all sets $A = \{t_j \in T : a = t_0 < \cdots < t_n = b\}$. For $A_1, A_2 \in \mathcal{C}_n$ and $x_1, x_2 \in D$, let $w(A_1, A_2) = \max\{|t_{1j} - t_{2j}|, j = 0, 1, \ldots, n\}$ and

$$w(x_1, x_2; A_1, A_2) = \max_{1 \leq j \leq n} \sup_{\substack{t_{1(j-1)} \leq s_1 < t_{1j} \\ t_{2(j-1)} \leq s_2 < t_{2j}}} \{m[x_1(s_1), x_2(s_2)]\}.$$

LEMMA 2.1. *For $x, y \in D([a, b])$,*

$$d(x, y) = \inf_n \inf_{A_1, A_2 \in \mathcal{C}_n} \{w(A_1, A_2) \vee w(x, y; A_1, A_2)\}.$$

PROOF. ($\leq$) For $A_1, A_2 \in \mathcal{C}_n$, let λ be defined by $\lambda(t_{1j}) = t_{2j}$, $0 \leq j \leq n$, and by linear interpolation elsewhere. Then $\rho(\lambda, e) = w(A_1, A_2)$ and $\rho(x, y \circ \lambda) \leq w(x, y; A_1, A_2)$.

($\geq$) Using Lemma 1 on p. 110 of Billingsley (1968), choose n and $A_1 \in \mathcal{C}_n$ for $\epsilon > 0$ given so that

$$\max_{1 \leq j \leq n} w(x; [t_{1(j-1)}, t_{1j})) \equiv \max_{1 \leq j \leq n} \sup_{t_{1(j-1)} \leq s_1, s_2 < t_{1j}} \{m[x(s_1), x(s_2)]\} < \epsilon.$$

Also choose λ so that $\rho(\lambda, e) \leq d(x, y) + \epsilon$ and $\rho(x, y \circ \lambda) \leq d(x, y) + \epsilon$. Let $A_2 = \lambda(A_1)$. Then $A_2 \in \mathcal{C}_n$, $w(A_1, A_2) \leq \rho(\lambda, e) \leq d(x, y) + \epsilon$, $w(x, y; A_1, A_2) \leq \rho(x, y \circ \lambda) + \max_{1 \leq j \leq n} w(x, [t_{1(j-1)}, t_{1j})) \leq d(x, y) + 2\epsilon$. ∎

REMARKS. The metric d as defined by Lemma 2.1 was introduced and shown to induce the J_1 topology by Kolmogorov (1956). Lemma 2.1 itself is due to Pomarede (1976). Lemma 2.1 shows that it suffices to work with the subset of piecewise linear functions in Λ with only finitely many changes of slope. For further discussion about metrics inducing the J_1 topology on D, see §3 of Straf (1970).

The following lemma shows that the specification of convergence in terms of the restrictions is consistent with the direct definition of convergence for compact domains.

LEMMA 2.2. *If b is a continuity point of x with $a < b < c$, $x_n \to x$ in $D([a, c])$ if and only if the restrictions of x_n convergence to restrictions of x in $d([a, b])$ and $D([b, c])$.*

PROOF. (if) If the restrictions converge, use the homeomorphisms λ_n of $[a, b]$ and $[b, c]$ to construct the homeomorphisms λ_n of $[a, c]$. Since b is a fixed point for the homeomorphisms associated with the restrictions, this construction is always possible.

(only if) Since b is a continuity point of x, the homeomorphisms λ_n of $[a, c]$ can be altered to make b a fixed point and still have $\rho(\lambda_n, e) \to 0$ and $\rho(x_n, x \circ \lambda_n) \to 0$. The restrictions of λ_n can now be used to get convergence for the restrictions of x_n. ∎

Let all topological spaces be endowed with Borel σ-fields (generated by the open subsets). Let $\pi_t : D(T, S) \to S$ be the one-dimensional projection defined for any $x \in D$ and $t \in T$ by $\pi_t(x) = x(t)$, and let $r_{ab} : D(T) \to D([a, b])$ be the restriction to $[a, b]$ defined for any $a < b$ in T by $r_{ab}(x)(t) = x(t)$, $a \leqslant t \leqslant b$.

LEMMA 2.3. *The projections π_t and the restrictions r_{ab} are measurable for each t and $a < b$ in T.*

PROOF. If t is an endpoint of T, then π_t is continuous and thus measurable. If t is not an endpoint, then π_t is continuous at x for almost all t by Lemma 2.2 because continuity points can be made endpoints. For $S = R$, follow p. 121 of Billingsley (1968) and let

$$h_{tn}(x) = n \int_0^{n^{-1}} \pi_{t+u}(x)\, du, \quad n \geqslant 1.$$

Then h_{tn} is continuous and $\pi_t = \lim_{n\to\infty} h_{tn}$ by the right-continuity of functions in D, so π_t is measurable. For S an arbitrary CSMS, recall that S is homeomorphic to a G_δ subset (countable intersection of open subsets) of R^∞, p. 308 of Dugundji (1968), so we can make the identification and enlarge the range to R^∞. (Note that $D(T, S)$ is homeomorphic to $D(T, S')$ if S and S' are homeomorphic.) Let h_{tn} map D into R^∞ by letting $h_{tn}(x)_i = h_{tn}(x_i)$, where $x_i(t)$ is the ith coordinate of $x(t) \in R^\infty$. Then the argument above for $S = R$ carries over to $S = R^\infty$. Thus, π_t is measurable with respect to the Borel σ-field of $D(T, R^\infty)$. Note that the J_1 topology on $D(T, S)$ coincides with the relative topology induced on $D(T, S)$ by the J_1 topology on $D(T, R^\infty)$. (To see this, recall that convergence is characterized by the metric convergence of the restrictions to compact subintervals of T and recall that a subspace of a metric space with that same metric is a metric space with the relative topology; Theorem 5.1, p. 186, of Dugundji (1968).) Thus the Borel σ-field on $D(T, S)$ coincides with the trace on $D(T, S)$ of the Borel σ-field on $D(T, R^\infty)$, i.e., $\mathfrak{B}(D(T, S)) = \mathfrak{B}(D(T, R^\infty)) \cap D(T, S)$; cf. Theorem 1.9, p. 5, of Parthasarathy (1967). Hence, the restriction of π_t to $D(T, S)$ is measurable. A similar argument applies to r_{ab}. For example, if b is not a right endpoint of T, work with $h_{abn}(x)(t) = h_{tn}(x)$, $a \leqslant t \leqslant b$, where S is regarded as a subset of R^∞ as above. Note that h_{abn} maps D into $C[a, b]$ for each n. Since $C[a, b]$ is a closed subset of $D[a, b]$ with the J_1 topology, $h_{abn}(x)$ cannot converge to $r_{ab}(x)$ as $n \to \infty$ in the J_1 topology if x has jumps in (a, b). However, it is not difficult to see that $h_{abn}(x)$ converges to $r_{ab}(x)$ in the M_1 topology. Moreover, it is not difficult to show that $h_{abn} : D \to C[a, b]$ is continuous for each n. Hence, r_{ab} is measurable as a mapping from D to $D[a, b]$ if the Borel σ-field associated with the J_1 topology is used on the domain and the Borel σ-field associated with the M_1 topology is used on the range. However, the M_1 topology is known to be metrizable as a complete separable metric space, Whitt (1973) and Pomarede (1976). Hence, the M_1 and J_1 topologies are comparable Souslin topologies, so their Borel σ-fields coincide, p. 124 of Schwartz (1973). ∎

REMARK. An alternate proof that π_t is measurable can be obtained from a simple modification of p. 249 of Parthasarathy (1967).

We now want to define a metric on D which induces the J_1 topology. To do this, we treat the different possible endpoints of T as separate cases. First, assume that T contains no finite open endpoints. This assumption entails no loss of generality because it is easy to construct a homeomorphic space in which finite open endpoints of T are replaced by $\pm\infty$. For example, it is easy to see that $D([a, b), S)$ is homemorphic to $D([0, \infty), S)$ under the mapping $\phi : D([0, \infty), S) \to D([a, b), S)$ defined by $\phi(x)(t) = x(\psi(t))$, $a \leqslant t < b$, where $\psi(t) = t(b-t)^{-1} - a(b-a)^{-1}$, $a \leqslant t < b$.

This leaves $[a, \infty)$, $(-\infty, a]$ and $(-\infty, \infty)$ as the relevant possibilities for T. Since $(-\infty, a]$ is similar to $[a, \infty)$, we do not discuss $(-\infty, a]$ further. For any $x, y \in D([a, \infty))$, let d be defined by

$$d(x, y) = \int_a^\infty dt\, e^{-(t-a)} d_{at}[r_{at}(x), r_{at}(y)] \wedge 1, \tag{2.2}$$

where $a \wedge b = \min\{a, b\}$ and d_{at} is the metric in (2.1) on $D([s, t])$. For any x, $y \in D((-\infty, \infty))$, let d be defined by

$$d(x, y) = \int_{-\infty}^0 ds \int_0^\infty dt\, e^{s-t} d_{st}[r_{st}(x), r_{st}(y)] \wedge 1. \tag{2.3}$$

The idea in (2.2) and (2.3) is to weight the tails relatively less and to have d-convergence determined by d_{st}-convergence of the restrictions for almost all s and t.

LEMMA 2.4. *The integrals in* (2.2) *and* (2.3) *are well defined.*

PROOF. For each x, y, and s, $d_{st}(r_{st}(x), r_{st}(y)]$ is continuous in t at each point of continuity of both x and y. Thus the integrands as functions of t are Riemann (and thus Lebesgue) integrable. For each x, y, and t, $d_{st}[r_{st}(x), r_{st}(y)]$ is also continuous in s at each point of continuity of both x and y. By the Lebesgue Dominated Convergence Theorem, the integral over t as a function of s is continuous at all points of continuity of both x and y. Thus the integral over t is Riemann integrable in s. ∎

THEOREM 2.5. *The functions d in* (2.2) *and* (2.3) *are metrics which induce the extended J_1 convergence on $D(T)$.*

PROOF. It is well known that $\rho \wedge 1$ is a bounded metric equivalent to ρ for any metric ρ, from which it is easy to deduce that d in (2.2) and (2.3) are metrics. If $x_n \to x$, then $d_{st}[r_{st}(x_n), r_{st}(x)] \to 0$ for almost all s and t, including $s = a$ in the setting of (2.2). By the Lebesgue Dominated Convergence Theorem, $d(x_n, x) \to 0$. On the other hand, if $d(x_n, x) \to 0$, then $d_{st}(r_{st}(x_n), r_{st}(x)) \to 0$ for all pairs (s, t) such that both s and t are continuity points of x. To see this, first note that $\{d_{st}[r_{st}(x_n), r_{st}(x)] \wedge 1\}$ has a convergent subsequence for each s and t. Let s and t be continuity points of x and suppose $d_{st}[r_{st}(x_{n'}), r_{st}(x)] \wedge 1 \to \epsilon > 0$ for a subsequence indexed by n'. Then it is not difficult to see that there is a δ depending on x and ϵ such that

$$\liminf_{n' \to \infty} d_{s_0 t_0}[r_{s_0 t_0}(x_{n'}), r_{s_0 t_0}(x)] > \delta$$

for all pairs (s_0, t_0) with $s - \delta < s_0 \leqslant s$ and $t \leqslant t_0 < t + \delta$. This implies that $\liminf_{n \to \infty} d(x_{n'}, x) > 0$, which is a contradiction. Hence, $d_{st}[r_{st}(x_{n'}), r_{st}(x)] \to 0$ whenever s and t are continuity points of x. Since every subsequence converges to 0, the entire sequence converges to 0. Hence, $x_n \to x$. ∎

THEOREM 2.6. *The space (D, J_1) is metrizable as a complete separable metric space.*

PROOF. The metrics d in (2.1) if $T=[a, b]$, in (2.2) if $T=[a, \infty)$ and in (2.3) if $T=(-\infty, \infty)$ are not complete, p. 112 of Billingsley (1968), but it is well known that there is a metric inducing the same topology as d in (2.1) which is complete, namely, d_0 on p. 113 of Billingsley (1968). If this complete metric is used for d_{st}, then d in (2.2) and (2.3) also become complete and induce the same topology. All the previous results in this section extend to this setting. The main fact to check is that the new $d_{st}[r_{st}(x), r_{st}(y)]$ is continuous in s and t at points which are continuity points of both x and y, which is needed in the new version of Lemma 2.4. For separability, a countable dense set in (D, d) consists of those x with values in a countable dense set of S which are constant over each interval $[(i-1)/k, i/k) \cap T$, $-nk \leqslant i \leqslant nk$, and each outside interval $(-\infty, (nk-1)/k) \cap T$ and $[n, \infty) \cap T$. ∎

LEMMA 2.7. *The Borel σ-field on D coincides with the Kolmogorov σ-field (generated by π_t, $t \in T$).*

The following proof is a variant of one communicated by David Pollard, who in turn attributes it to Michael Wichura.

PROOF. Lemma 2.3 implies that the Kolmogorov σ-field is contained in the Borel σ-field. To go the other way, it is possible to follow the proof of Theorem 14.5 of Billingsley (1968), which involves introducing a new basis for the topology. More directly, it suffices to show that each continuous real-valued function f on D is measurable with respect to the Kolmogorov σ-field, p. 4 of Parthasarathy (1967). Suppose $T=[0, \infty)$, the other cases being treated similarly. For each $n \geqslant 1$, let $\phi_n : D \to D$ be defined by $\phi_n(x)(t) = x(k/n)$, $k/n \leqslant t < (k+1)/n$, $0 \leqslant k < n^2$, and $\phi_n(x)(t) = x(n)$, $t \geqslant n$. It is easy to see that $\phi_n(x) \to x$ as $n \to \infty$ for each $x \in D$. Since f is continuous, $f(\phi_n(x)) \to f(x)$ as $n \to \infty$. Hence, it suffices to show that $f \circ \phi_n$ is measurable. However, it is possible to represent $f \circ \phi_n$ as the composition of three measurable functions. Let $\pi_n : D \to S^{n^2+1}$ be the projection map defined by $\pi_n(x) = [x(0), x(1/n), x(2/n), \ldots, x(n)]$. Let $\psi_n : S^{n^2+1} \to D$ be defined so that $\psi_n(\pi_n(x)) = \phi_n(x)$ for each $x \in D$. Hence, $f \circ \phi_n = f \circ \psi_n \circ \pi_n$. Let the domain of π_n be endowed with the Kolmogorov σ-field; let S^{n^2+1} be endowed with the product topology, under which the Borel σ-field coincides with the product σ-field; and let the domain of f be endowed with the J_1 topology. Since ψ_n and f are continuous, they are Borel-measurable. By Lemma 2.3, π_n is measurable. Since the composition of measurable functions is measurable, $f \circ \psi_n \circ \pi_n$ is measurable. ∎

As on p. 124 of Billingsley (1968), let $T_P = \{t \in T : P(\{x \in D : x(t) \neq x(t-)\}) = 0\}$ for any probability measure P on $D(T)$. The set $T - T_P$ is at most countable. As before, let $\Rightarrow$ denote weak convergence of probability measure.

THEOREM 2.8. *Let P_n, $n \geqslant 1$, and P be probability measures on $D(T)$. Then $P_n \Rightarrow P$ if and only if $P_n r_{s_k t_k}^{-1} \Rightarrow P r_{s_k t_k}^{-1}$ on $D([s_k, t_k])$ for all k and some sequence $\{[s_k, t_k], k \geqslant 1\}$ with $\bigcup_{k=1}^{\infty} [s_k, t_k] = T$.*

PROOF. (only if) By Lemmas 2.2 and 2.3, r_{st} is measurable and continuous almost surely with respect to P for $s, t \in T_P$. Thus the CMT can be applied.

(if) Let $\{s_k\}$ be a nonincreasing sequence of points in T_P and let $\{t_k\}$ be a nondecreasing sequence of points in T_P with $t_k > s_k$ for $k \geqslant 1$ and $T = \bigcup_{k=1}^{\infty} [s_k, t_k]$. Let F be an arbitrary closed subset of $D(T, S)$ and let $H_k = r_{s_k t_k}^{-1}(r_{s_k t_k}(F))$. It is easy to see that $F \subseteq H_k$ for all k. Moreover, we shall show that $F = \bigcap_{k=1}^{\infty} H_k$. Assuming this for the moment, let $\epsilon > 0$ be given and choose k_0 so that $P(H_{k_0}) \leqslant P(F) + \epsilon$. Then, for $k \geqslant k_0$,

$$\begin{aligned}\limsup_{n\to\infty} P_n(F) &\leqslant \limsup_{n\to\infty} P_n(H_k) = \limsup_{n\to\infty} \left(P_n r_{s_k t_k}^{-1}\right)\left(\overline{r_{s_k t_k}(F)}\right)\\ &\leqslant \left(P r_{s_k t_k}^{-1}\right)\left(\overline{r_{s_k t_k}(F)}\right) = P(H_k)\\ &\leqslant P(F) + \epsilon,\end{aligned}$$

by virtue of the convergence $P_n r_{s_k t_k}^{-1} \Rightarrow P r_{s_k t_k}^{-1}$ using the characterization in Theorem 2.1 (iii) of Billingsley (1968). Since ϵ was arbitrary, this implies $P_n \Rightarrow P$.

It remains to show that $\bigcap_{k=1}^{\infty} H_k \subseteq F$. Suppose $x \in \bigcap_{k=1}^{\infty} H_k$. We shall show that, for any $\epsilon > 0$, $d(x, F) < \epsilon$. Since F is closed, this implies that $x \in F$. To be definite, suppose $T = [a, \infty)$ so that d is defined in (2.2). For $\epsilon > 0$ given, choose k_0 so that $s_{k_0} = a$ and $e^{-(t_{k_0} - a)} < \epsilon/2$. This implies that $d(x, y) < d_k(x, y) + \epsilon/2$ for $k \geqslant k_0$, where

$$d_k(x, y) = \int_a^{t_{k_0}} \circ\, dt\, e^{-(t-a)} d_{at}\left[r_{at}(x), r_{at}(y)\right] \wedge 1.$$

Since $x \in H_{k_0}$, $x \in r_{at_{k_0}}^{-1}(\overline{r_{at_{k_0}}(F)})$. This means that there is a sequence $\{y_n\}$ in F such that $d_{at_{k_0}}[r_{at_{k_0}}(x), r_{at_{k_0}}(y_n)] \to 0$ as $n \to \infty$. By Lemma 2.2, $d_{at}[r_{at}(x), r_{at}(y_n)] \to 0$ as $n \to \infty$ for all t which are continuity points of x in $[a, t_{k_0}]$, which is almost all t. Hence, by the Lebesgue Dominated Convergence Theorem, $d_k(x, y_n) \to 0$ as $n \to \infty$. Choose n_0 so that $d_k(x, y_n) < \epsilon/2$ for $n \geqslant n_0$. Then, for $n \geqslant n_0$, $d(x, y_n) < d_k(x, y_n) + \epsilon/2 < \epsilon$, so that $d(x, F) < \epsilon$. ∎

REMARK. The proof of Theorem 2.8 is closely related to the general theory of weak convergence induced by mappings recently developed by Pollard (1977). While the proof here was obtained independently of Pollard (1977), it is a fairly recent addition, reaching this form only in 1978 with the aid of a suggestion by Richard Serfozo.

3. Composition. This section is devoted to the composition function, which is often used in random time transformations (subordination), cf. §17 of Billingsley (1968). Recent related results about weak convergence with random time transformations are contained in Aldous (1978) and Durrett and Resnick (1977).

We begin by defining some subsets of $D \equiv D(T, S)$. Throughout this paper, we assume subsets of D and other spaces are topologized with the relevant relative topology and Cartesian products are topologized with the relevant product topology. Let $C \equiv C(T, S)$ be the subset of continuous funcfions in D; let $C_0 = C_0(T_1, T_2)$ be the subset of strictly-increasing T_2-valued functions in C; let $D_0 \equiv D_0(T_1, T_2)$ be the subset of nondecreasing T_2-valued functions in D. It is well known that the J_1 topology on C coincides with the topology of uniform convergence on compact intervals. Obviously C and D_0 are closed subsets of D, while C_0 is neither open nor closed. However, C_0 is a G_δ because

$$C_0 = \bigcap_{p \in Q} \bigcap_{\substack{q \in Q\\ q > p}} \{x \in C : x(q) - x(p) > 0\},$$

where Q is the set of rationals in T_1.

Let composition be defined for any $(x, y) \in D(T_3, S) \times D_0(T_1, T_2)$ with $T_2 \subseteq T_3$ by $(x \circ y)(t) = x(y(t))$, $t \in T_1$. It is easy to see that $x \circ y \in D(T_1, S)$, which need not be the case if $y \notin D_0$. For example, let $T_1 = T_2 = T_3 = S = [0, 1]$; if $x = I_{[2^{-1}, 1]}$ and $y = 2^{-1} + \sum_{n=1}^{\infty} (-2)^{-n} I_{[2^{-1} - 2^{-n}, 2^{-1} - 2^{-(n+1)})}$, where I_A is the indicator function of the set A, then $x, y \in D$ but $x \circ y$ has no limit from the left at $t = 2^{-1}$.

Billingsley (1968, pp. 145, 232) has shown that composition on $D([0, 1], R) \times D_0([0, 1], [0, 1])$ is measurable and continuous at $(x, y) \in C \times (C \cap D_0)$. With Lemma 2.7, Billingsley's arguments also apply when the domains of x and y are more general intervals and the range of x is a CSMS. To see that composition is not continuous on $D \times D_0$ in any of Skorohod's (1956) topologies, let $T_1 = T_2 = T_3 = S = [0, 1]$, $x_n = x = I_{[1/2, 1]}$, $y(t) = 2^{-1}$ and $y_n(t) = (2^{-1} - n^{-1})$, $0 \leqslant t \leqslant 1$. Then $(x_n \circ y_n)(t) = 0$ while $(x \circ y)(t) = 1$, $0 \leqslant t \leqslant 1$.

THEOREM 3.1. *The composition mapping on* $D(T_3, S) \times D_0(T_1, T_2)$ *is continuous at each* $(x, y) \in (C \times D_0) \cup (D \times C_0)$.

REMARK. H. Bauer has noted [2, p. 28] that an additional assumption is needed when $(x, y) \in (D \times C_0)$. It suffices for T_1 to be open on the right. If T_1 has the right endpoint b, then it suffices for x to be continuous at $y(b)$.

PROOF. (i) Suppose $(x_n, y_n) \to (x, y)$ in $D \times D_0$ with $(x, y) \in C \times D_0$. It suffices to look at compact domains. Choose $[a, b] \subseteq T_1$ so that a and b are continuity points of y or endpoints of T_1. Since $y_n \to y$ for the restrictions to $[a, b]$, the range of the restrictions of y_n, $n \geqslant 1$, and y to $[a, b]$ is contained in a compact subinterval $[c, d] \subseteq T_2$. Since $x \in C$ and $x_n \to x$, $\rho(x_n, x) \to 0$ for the restrictions to $[c, d]$. Let λ_n be homeomorphisms of $[a, b]$ such that $\rho(\lambda_n, e) \to 0$ and $\rho(y_n, y \circ \lambda_n) \to 0$. Working with the restrictions to compact domains, we have by the triangle inequality $\rho(x_n \circ y_n, x \circ y \circ \lambda_n) \leqslant \rho(x_n \circ y_n, x \circ y_n) + \rho(x \circ y_n, x \circ y \circ \lambda_n)$. The first term converges to 0 because $\rho(x_n, x) \to 0$ and the second term converges to 0 because $x \in C$ and $\rho(y_n, y \circ \lambda_n) \to 0$.

(ii) Suppose $(x_n, y_n) \to (x, y)$ in $D \times D_0$ with $(x, y) \in D \times C_0$. Again it suffices to look at compact domains. Choose $[a, b]$ so that a and b are continuity points of $x \circ y$ or endpoints of T_1. Choose $[c, d] \subseteq T_2$ so that the restrictions of y_n, $n \geqslant 1$, and y to $[a, b]$ have values in $[c, d]$ and so that c and d are continuity points of x or endpoints of T_3. We now work with the restrictions to the compact domains. Let λ_n, $n \geqslant 1$, be homeomorphisms of $[c, d]$ such that $\rho(x_n, x \circ \lambda_n) \to 0$ and $\rho(\lambda_n, e) \to 0$. Let μ_n, $n \geqslant 1$, be elements of $\Lambda([a, b])$. By the triangle inequality, $\rho(x_n \circ y_n, x \circ y \circ \mu_n) \leqslant \rho(x_n \circ y_n, x \circ \lambda_n \circ y_n) + \rho(x \circ \lambda_n \circ y_n, x \circ y \circ \mu_n)$. Since the first term on the right converges to 0 as $n \to \infty$, it suffices to construct μ_n for each $\epsilon > 0$ so that $\rho(\mu_n, e) < \epsilon$ and $\rho(x \circ \lambda_n \circ y_n, x \circ y \circ \mu_n) < \epsilon$ for sufficiently large n. The idea is to let μ_n be approximately $y^{-1} \circ \lambda_n \circ y_n$. This cannot actually be the solution because y_n need not be continuous or strictly increasing, but it suffices to define μ_n on a finite subset this way and use linear interpolation elsewhere. This works because $y_n \to y$ where $y \in C_0$.

By Lemma 1 on p. 110 of Billingsley (1968) there exists a finite set of points $\{t_j\}$ such that $c = t_0 < \cdots < t_n = d$ and $w(x; [t_{j-1}, t_j)) = \sup_{t_{j-1} \leqslant s_1, s_2 < t_j} \{m[x(s_1), x(s_2)]\} < \epsilon$, $1 \leqslant j \leqslant n$. Since a and b are continuity points of $x \circ y$, the finite set can be chosen so that either $y(a) = c$ or $y(a)$ is not included in the set. Similarly, $y(b) = d$ or it is not included. Furthermore, the finite set $\{t_j\}$ can be chosen, by adding points if necessary, so that $y^{-1}(t_{j+1}) - y^{-1}(t_j) < \epsilon/2$ for $y(a) < t_j < t_{j+1} < y(b)$, $y^{-1}(t_j) - a < \epsilon/2$ for t_j the smallest point greater than $y(a)$, and $b - y^{-1}(t_j) < \epsilon/2$ for t_j the largest point less than $y(b)$. Let n_0 be such that $\rho(y^{-1} \circ \lambda_n \circ y_n, e) < \epsilon/2$ for $n \geqslant n_0$, which exists by part (i). Let $l^* = \max\{i : t_i \leqslant y(a)\}$, $m^* = \min\{i : t_i \geqslant y(b)\}$ and $u_i = t_{l^*+i}$, $0 \leqslant i \leqslant m^* - l^*$. For sufficiently large n, $\{u_i\}$ is the relevant subset of $\{t_i\}$. Let $s_{ni} = \inf\{s : \lambda_n \circ y_n(s) \geqslant u_i\}$, $0 \leqslant i \leqslant m^* - l^*$ and $n \geqslant 1$. For all n sufficiently large, $s_{ni} < s_{n(i+1)}$, $0 \leqslant i \leqslant m^* - l^* - 1$, and $|s_{ni} - y^{-1}(u_i)| < \epsilon/2$, $1 \leqslant i \leqslant m^* - l^* - 1$. Hence, for sufficiently large n, we can define μ_n by $\mu_n(s_{ni}) = y^{-1}(u_i)$, $1 \leqslant i \leqslant m^* - l^* - 1$, and by linear interpolation elsewhere. Since $(\lambda_n \circ y_n)(t) \in [u_i, u_{i+1})$, if and only if $(y \circ \mu_n)(t) \in [u_i, u_{i+1})$,

$$\rho(x \circ \lambda_n \circ y_n, x \circ y \circ \mu_n) \leqslant \max_{1 \leqslant j \leqslant n} w(x, [t_{j-1}, t_j)) < \epsilon.$$

Also, by the triangle inequality,

$$\rho(\mu_n, e) \leqslant \rho(\mu_n, y^{-1} \circ \lambda_n \circ y_n) + \rho(y^{-1} \circ \lambda_n \circ y_n, e).$$

By construction,

$$\rho(\mu_n, y^{-1} \circ \lambda_n \circ y_n) \leqslant \max\{y^{-1}(u_{i+1}) - y^{-1}(u_i)\} < \epsilon/2$$

and, for $n \geqslant n_0$, $\rho(y^{-1} \circ \lambda_n \circ y_n, e) < \epsilon/2$. ∎

REMARK. A different proof of (ii) above can be obtained from Theorem 2.6.1 of Skorohod (1956). The idea is to show that $(x_n \circ y_n)$ converges pointwise on a dense set, which is easy, and then control Skorohod's J_1 modulus for $x_n \circ y_n$ in terms of moduli for x_n and y_n. In fact, for any $(x, y) \in D \times D_0$, $w''_{x \circ y}(\delta) \leqslant w''_x(\delta)$ for w'' in (14.44) of Billingsley (1968). This inequality does not hold for w' in (14.6) of Billingsley (1968): let $T_1 = T_2 = T_3 = S = [0, 1]$, $x = I_{[1/2, 1]}$, $y = (1/3)I_{[0, 1/2)} + (2/3)I_{[1/2, 1]}$, and $\delta = 2/3$ (example due to Michael Wichura).

Composition is also continuous in the following more special situation; Lemma 3.1 of Kennedy (1972). Let $F \equiv F(T_3, S)$ be the subset of functions in D with discontinuities only at integer points in T_3. Let $G \equiv G(T_1, T_2)$ be the subset of functions in D_0 with integer values. Obviously F and G are closed subsets of D.

THEOREM 3.2. *Composition is continuous on $F \times G$.*

PROOF. Suppose $(x_n, y_n) \to (x, y)$ in $F \times G$. Working with the restrictions to compact domains, we have via the triangle inequality that $\rho(x_n \circ y_n, x \circ y \circ \lambda_n) \leqslant \rho(x_n \circ y_n, x \circ y_n) + \rho(x \circ y_n, x \circ y \circ \lambda_n)$. If λ_n are homeomorphsims such that $\rho(y_n, y \circ \lambda_n) \to 0$ and $\rho(\lambda_n, e) \to 0$, then $y_n = y \circ \lambda_n$ for sufficiently large n because y_n and y are integer-valued. Hence, the second term on the right is 0 for sufficiently large n. The first term converges to 0 because d-convergence coincides with ρ-convergence on F. ∎

It is also of interest to have a "converse" to continuity for composition, which provides convergence for stochastic processes based on convergence of embedded processes.

THEOREM 3.3. *Suppose $x_n \circ y_n \to z$ in $D([a, b])$ and $y_n \to y$ in $C \cap D_0([a, b])$ with $y \in C_0$. Then*

(i) *$x_n \to z \circ y^{-1}$ in $D((y(a), y(b)))$; and*

(ii) *if $y_n(a) = y(a)$ and $y_n(b) = y(b)$ for all n sufficiently large, then $x_n \to z \circ y^{-1}$ in $D([y(a), y(b)])$.*

PROOF. (i) Since $y_n \to y$, $y_n(a) < c$ and $y_n(b) > d$ for any compact $[c, d] \subseteq (y(a), y(b))$ for sufficiently large n. Since $y_n \to y$, $y_n^{-1} \to y^{-1}$ on $[c, d]$, where $y_n^{-1}(t) = \inf\{u \geqslant a : y_n(u) > t\}$, $c \leqslant t \leqslant d$, cf. Theorem 7.2. Choose $[c, d]$ so that $y^{-1}(c)$ and $y^{-1}(d)$ are continuity points of z. By Theorem 3.1, $x_n \circ y_n \circ y_n^{-1} \to z \circ y^{-1}$ on $[c, d]$. When $y_n(a) < c$ and $y_n(b) > d$, $y_n \circ y_n^{-1} = e$, where e is the identity map on $[c, d]$ since $y_n \in C$, so $x_n = x_n \circ y_n \circ y_n^{-1}$ on $[c, d]$ for sufficiently large n. Note that $y_n \circ y_n^{-1} = e$ while $y_n^{-1} \circ y_n \neq e$ in general.

(ii) If $y_n(a) = y(a)$ and $y_n(b) = y(b)$, $[y(a), y(b)]$ can be used instead of $[c, d]$ in (i).

REMARK. If y_n is not continuous in Theorem 3.3, an extra condition on the fluctuations of x_n is needed. For example, let all domains be $[0, 1]$, let $y = e$, $y_n = (e + n^{-1}I_{[1/2, 1]}) \wedge 1$, $x(t) = 0$, $0 \leqslant t \leqslant 1$, $x_n(2^{-1} + (2n)^{-1}) = 1$, $x_n(t) = 0$, $t \notin [2^{-1}, 2^{-1} + n^{-1}]$, and x_n defined by linear interpolation elsewhere. Then $x_n \circ y_n = x$, $y_n \to y$, $y \in C_0$ and all functions are continuous except y_n, but $\{x_n\}$ does not converge.

We conclude this section with an extension of Theorem 3.3 in the weak convergence setting, motivated by Lemma 2 of Iglehart and Whitt (1971). For a different approach, see Serfozo (1975). As in the remark following Theorem 3.1, we need T_1 to be open on the right in the following theorem and corollary.

THEOREM 3.4. *Let (X_n, Y_n) be random elements of $D(T_3, S) \times D_0(T_1, T_2)$ with $T_2 \subseteq T_3$ and let y be a nonrandom element of $C_0(T_1, T_2)$. If*
(i) *$\{X_n\}$ is relatively compact in $D(y(T_1), S)$,*
(ii) *$Y_n \Rightarrow y$ in D_0, and*
(iii) *the finite-dimensional distributions of $X_n \circ Y_n$ converge at all points in a dense subset of T_1, then $X_n \circ Y_n \Rightarrow Z$ in $D(T_1 S)$ for some random element Z and $X_n \Rightarrow Z \circ y^{-1}$ in $D(y(T_1), S)$.*

PROOF. By (i) and Theorem 2.3 of Billingsley (1968), it suffices to show that every weakly convergent subsequence has the same limit. Suppose $X_{n'} \Rightarrow X'$ and $X_{n''} \Rightarrow X''$. By (ii) and Theorem 4.4 of Billingsley (1968), $(X_{n'}, Y_{n'}) \Rightarrow (X', y)$ and $(X_{n''}, Y_{n''}) \Rightarrow (X'', y)$. By Theorem 3.1, $X_{n'} \circ Y_{n'} \Rightarrow X' \circ y$ and $X_{n''} \circ Y_{n''} \Rightarrow X'' \circ y$. However, (iii) implies that $X' \circ y \sim X'' \circ y$, where $\sim$ means equality in distribution, because the finite-dimensional distributions uniquely determine a measure on D. Since $y \in C_0$, $X' \sim X''$ on $[y(a), y(b)]$. Hence, $X_n \Rightarrow X$ for some random element X. Let $Z = X \circ y$. By Theorem 3.1, $X_n \circ Y_n \Rightarrow Z$ and $X_n \circ Y_n \circ y^{-1} \Rightarrow X \circ y \circ y^{-1} = Z \circ y^{-1} = X$. ∎

REMARK. Conditions (ii) and (iii) above do not imply convergence of the finite-dimensional distributions of $\{X_n\}$ in Theorem 3.4. For example, let $T_1 = T_2 = T_3 = S = [0, 1]$, $y(t) = t$, $y_n(t) = [nt]/n$, $x(t) = 0$, $0 \leqslant t \leqslant 1$, and $x_n = \sum_{k=0}^{n-1} I_{[kn^{-1}+(2n)^{-1}, (k+1)n^{-1})}$, where $[a]$ is the integer part of a. Then $(x, y) \in C \times C_0$, $y_n \to y$, $x_n \circ y_n(t) = x \circ y(t) = 0$, $0 \leqslant t \leqslant 1$, but $x_n(t)$ fails to converge to $x(t)$ at any $t \in (0, 1)$.

Let $F(x_n, y_n; a, b)$ denote the maximum fluctuation of x_n in $D([y(a), y(b)])$ over any jump of y_n in $D_0([a; b])$; that is, let

$$F(x_n, y_n; a, b) = \sup\{ m[x(u), x(v)] : y(a) \leqslant y_n(t-) \leqslant u, v \leqslant y_n(t) \leqslant y(b), a \leqslant t \leqslant b\}, \tag{3.1}$$

where $y_n(a-)$ and $y_n(b)$ are redefined to be $y(a)$ and $y(b)$ if $y_n(b-) > y(a)$ or $y_n(b) < y(b)$. Let T_Z be the subset of points in T_1 which are either continuity points of Z w.p.l. or endpoints of T_1.

COROLLARY 1. *Let (X_n, Y_n) be random elements of $D(T_3, S) \times D_0(T_1, T_2)$ with $T_2 \subseteq T_3$, $S = R$, and y a nonrandom element of C_0. If*
(i) *$Y_n \Rightarrow y$ in D_0;*
(ii) *$X_n \circ Y_n \Rightarrow Z$ in D;*
(iii) *$F(X_n, Y_n; a, b) \Rightarrow 0$ for all a and b in T_Z, then*

$$X_n \Rightarrow Z \circ y^{-1} \quad in\ D(y(T_1), R).$$

If $Z \in C$ w.p.l., then (iii) *is a necessary condition.*

PROOF. Obviously (i) and (ii) here imply (ii) and (iii) in Theorem 3.4, so it suffices to show that $\{X_n\}$ is tight. Let $a, b \in T_Z$. Obviously

$$\cdot\sup_{y(a) \leqslant t \leqslant y(b)} |X_n(t)| \leqslant \sup_{a \leqslant t \leqslant b} |X_n \circ Y_n(t)| + F(X_n, Y_n; a, b),$$

so (i) of Theorem 15.2 of Billingsley (1968) follows from (ii) and (iii). Since $y \in C_0$, for any $\delta > 0$ there is a $\gamma > 0$ such that $y(t) - y(s) > 2\gamma$ for all s and t such that $a \leqslant s < t \leqslant b$ and $t - s > \delta$. (The continuous function $f(t) = y(t) - y(t - \delta)$ attains its infimum on $[a + \delta, b]$.) Since $Y_n \Rightarrow y$, for this δ and any $\eta > 0$, $P(Y_n(t_j) - Y_n(t_{j-1}) > \gamma$, all $j) > 1 - \eta$ for any set $\{t_j\}$ in $[y(a), y(b)]$ with $t_j - t_{j-1} > \delta$ for sufficiently large n, where n is independent of $\{t_j\}$. Hence, for each $\epsilon, \eta > 0$, there exist $\delta, \gamma > 0$ and n_0 so that

$$P(w'_{X_n}(\gamma) > 2\epsilon) \leq P(w'_{X_n \circ Y_n}(\delta) > \epsilon) + P(2F(X_n, Y_n; a, b) > \epsilon) + \eta \leq 3\eta$$

for $n \geq n_0$, where w' is the modulus in (14.6) of Billingsley (1968), so that (ii) of Theorem 15.2 of Billingsley (1968) is satisfied. Thus $\{X_n\}$ is tight.

To see that (iii) is necessary when $P(Z \in C) = 1$, note that

$$P(F(X_n, Y_n; a, b) > \epsilon) \leq P(w_{X_n}(\delta) > \epsilon) + P\Big(\sup_{a \leq t \leq b} |Y_n(t) - Y_n(t-)| > \delta\Big),$$

where w is the modulus in (14.2) of Billingsley (1968) on $[y(a), y(b)]$. If $P(Z \in C) = 1$, then $P(Z \circ y^{-1} \in C) = 1$ so that $\{X_n\}$ is C-tight. Lemmas 3 and 4 of Iglehart and Whitt (1971) show that the right side converges to 0 as $n \to \infty$ for any $\epsilon > 0$. ∎

Prior to Theorem 3.4, our results have been expressed in a deterministic setting, from which applications to stochastic settings are immediate. However, it is also possible to go the other way, which shows that Theorem 3.4 applies to all modes of stochastic convergence. Since $x_n \to x$ if and only if $P_n \Rightarrow P$, where P_n and P are measures attaching unit mass to x_n and x, p. 12 of Billingsley (1968), we have the following.

COROLLARY 2. *Let (x_n, y_n) be elements of $D(T_3, S) \times D_0(T_1, T_2)$ with $T_2 \subseteq T_3$, $S = R$ and $y \in C_0$. If $y_n \to y$ in D_0, $x_n \circ y_n \to z$ in D, and $F(x_n, y_n; a, b) \to 0$ for all a and b which are continuity points of z, then*

$$x_n \to z \circ y^{-1} \text{ in } D(y(T_1), R).$$

If $z \in C$, then $F(x_n, y_n; a, b) \to 0$ for all a and b in T_1 is a necessary condition.

REMARK. Obviously (i) in Theorem 3.4 is a necessary condition and we have just seen that (iii) in Corollaries 1 and 2 is necessary when $P(Z \in C) = 1$, but (iii) in Corollaries 1 and 2 is not necessary in general. For example, let all domains be [0, 1], let $x = I_{[2^{-1}, 1]}$, $x_n = I_{[2^{-1}+n^{-1}, 1]}$, $y = e$, $y_n(t) = tI_{[0, 2^{-1}] \cup [2^{-1}+2n^{-1}, 1]}(t) + (2^{-1} + 2n^{-1}) I_{[2^{-1}, 2^{-1}+2n^{-1}]}(t)$, $0 \leq t \leq 1$. Then $F(x_n, y_n; 0, 1) = 1$ while $x_n \to x$, $y_n \to y$, and $x_n \circ y_n \to x \circ y$.

4. Addition and multiplication. Let a binary operation + called addition be defined on the CSMS (S, m) such that $m(s_1 + s_2, s_3 + s_4) \leq m(s_1, s_3) + m(s_2, s_4)$ for all $s_1, s_2, s_3\ s_4 \in S$. It suffices for S to be a topological group with m generated by a norm: $m(s_1, s_2) = \|s_1 - s_2\|$ and $\|s_1 + s_2\| \leq \|s_1\| + \|s_2\|$. For example, a Banach space will do. Define addition in $D(T, S)$ by $(x + y)(t) = x(t) + y(t)$, $t \in T$. Addition is obviously continuous on $D \times D$ at $(x, y) \in C \times C$ because for compact intervals $\rho(x_n + y_n, x + y) \leq \rho(x_n, x) + \rho(y_n, y)$ by virture of the initial assumption. However, addition is not continuous in general (Problem 3 on p. 123 of Billingsley (1968)): let all domains be [0, 1] and let $S = R$; if $y = -x = -x_n = I_{[1/2, 1]}$ and $y_n = I_{[2^{-1}+n^{-1}, 1]}$, then $x_n = x$ and $y_n \to y$, but $\{x_n + y_n\}$ does not converge. (The minus sign in this example shows that addition is not continuous in any of Skorohod's (1956) topologies.) Consequently, D is not a topological group and much of classical functional analysis does not apply. The example above also illustrates that the product topology on $D(T, S)^2$ is weaker than the J_1 topology on $D(T, S^2)$ with the product topology on S^2: $(x_n, y_n) \to (x, y)$ in the first topology, but not the second. It is easy to see that addition is continuous in the J_1 topology on $D(T, S^2)$, so an alternate way to show that addition preserves convergence is to demonstrate convergence in $D(T, S^2)$ with the J_1 topology. This approach has recently been pursued by Pakshirajan and Mohan (1978).

Let $\text{Disc}(x)$ be the set of discontinuity points of x in T. Since $\{(x,y) : \text{Disc}(x) \cap \text{Disc}(y) = \emptyset\}$ is the intersection over n of the (open) sets of (x,y) which have no common discontinuities of size at least n^{-1}, this set is a G_δ.

THEOREM 4.1. *Addition on $D \times D$ is measurable and continuous at those (x,y) for which* $\text{Disc}(x) \cap \text{Disc}(y) = \emptyset$.

PROOF. (measurability) By Lemma 2.7, it suffices to show that the map $(x,y) \to x(t) + y(t)$ from $D \times D$ to S is measurable for each $t \in T$, but this follows since this map is the composition of the measurable maps $(x,y) \to (x(t), y(t))$ and $(s,s') \to s + s'$.

(continuity) Let a and b be continuity points of both x and y in T and fix $\epsilon > 0$. Apply Lemma 1 on p. 110 of Billingsley (1968) to construct finite subsets $A_1 = \{t_j\}$ and $A_2 = \{s_j\}$ of $[a,b]$ such that $a = t_0 < \cdots < t_n = b$, $a = s_0 < \cdots < s_m = b$, $w(x; [t_{j-1}, t_j)) < \epsilon$ and $w(y; [s_{j-1}, s_j)) < \epsilon$ for all j. Since $\text{Disc}(x) \cap \text{Disc}(y) = \emptyset$, the two sets A_1 and A_2 can be chosen so that $A_1 \cap A_2 = \{a, b\}$. Note that $w(x; [t_{j-1}, t_j)) < \epsilon$ and $w(y; [t_{j-1}, t_j)) < \epsilon$ for $\{t_j\} = A_1 \cup A_2$. Let 2δ be the distance between the closest two points in $A_1 \cup A_2$. Choose n_0 and homeomorphisms λ_n and μ_n so that $\rho(x_n, x \circ \lambda_n) < (\delta \wedge \epsilon)$, $\rho(\lambda_n, e) < (\delta \wedge \epsilon)$, $\rho(y_n, y \circ \mu_n) < (\delta \wedge \epsilon)$, and $\rho(\mu_n, e) < (\delta \wedge \epsilon)$ for $n \geq n_0$. Thus, for $n \geq n_0$, $\lambda_n^{-1}(A_1) \cap \mu_n^{-1}(A_2) = \{a, b\}$ and $\lambda_n^{-1}(A_1) \cup \mu_n^{-1}(A_2)$ has corresponding points in the same order as $A_1 \cup A_2$. Let γ_n be homeomorphsims of $[a,b]$ defined by $\gamma_n(t_j) = t_j'$ for corresponding points $t_j \in \lambda_n^{-1}(A_1) \cup \mu_n^{-1}(A_2)$ and $t_j' \in A_1 \cup A_2$ and by linear interpolation elsewhere. Then $\rho(\gamma_n, e) < \epsilon$ and $\rho(x_n + y_n, (x+y) \circ \gamma_n) \leq \rho(x_n, x \circ \gamma_n) + \rho(y_n, y \circ \gamma_n) < 2\epsilon$, with the first inequality holding because of the initial assumption about m and $+$. ∎

Now let $S = R$ and consider pointwise multiplication on $D \times D$, defined by $(xy)(t) = x(t)y(t)$, $t \in T$. The example for addition with 1 added to y_n, y and 1 subtracted from x_n, x shows that multiplication is not continuous in general.

THEOREM 4.2. *Multiplication is measurable on $D(T,R) \times D(T,R)$ and continuous at those x, y for which* $\text{Disc}(x) \cap \text{Disc}(y) = \emptyset$.

PROOF. The argument for Theorem 4.1 applies with only minor modification. Only the last step of the continuity argument must be changed: $\rho(x_n y_n, (xy) \circ \gamma_n) \leq \rho(x_n y_n, y_n(x \circ \gamma_n)) + \rho(y_n(x \circ \gamma_n), (xy) \circ \gamma_n) \leq \|y_n\| \rho(x_n, x \circ \gamma_n) + \|x\| \rho(y_n, y \circ \gamma_n)$, where $\|x\| = \sup_{a \leq t \leq b} |x(t)| < \infty$. Since $y_n \to y$, $\sup_{n \geq 1} \|y_n\| < \infty$. ∎

For stochastic applications of Theorems 4.1 and 4.2, we have

LEMMA 4.3. *If X and Y are independent random elements of D with $P(X(t) = X(t-)) = 1$ for all $t \in T$, then*

$$P(\text{Disc}(X) \cap \text{Disc}(Y) \neq \emptyset) = 0.$$

PROOF. Since X and Y are independent, we can write

$$P(\text{Disc}(X) \cap \text{Disc}(Y) \neq \phi) = \int P(X \in A(y)) P(Y \in dy),$$

where $A(y) = \{x : \text{Disc}(x) \cap \text{Disc}(y) \neq \emptyset\}$. For any $y \in D$, $P(X \in A(y)) = \sum_{t \in \text{Disc}(y)} P(X(t) \neq X(t-)) = 0$. ∎

5. Composition with translation. Theorem 4.1 can be combined with §3 to treat composition with translation. For the following theorem, let $x_n, x, z \in D(T,R)$, $y_n \in D_0(T,T)$, $y_0 \in C_0(T,T)$ and $y_n \to y$. Let b_n and c_n be real constants, and let $F(x_n, y_n; a, b)$ be as in (3.1).

THEOREM 5.1.

(i) *If* $(x_n - c_n e) \to x$, $c_n(y_n - b_n e) \to z$, *and* $\mathrm{Disc}(x \circ y) \cap \mathrm{Disc}(z) = \emptyset$, *then*

$$(x_n \circ y_n - b_n c_n e) \to x \circ y + z \quad \text{in } D(T, R).$$

(ii) *If* $(x_n \circ y_n - b_n c_n e) \to x$, $c_n(y_n - b_n e) \to z$, $y_n \in C \cap D_0$ *and* $\mathrm{Disc}(x) \cap \mathrm{Disc}(z) = \emptyset$, *then*

$$(x_n - c_n e) \to x \circ y^{-1} - z \circ y^{-1} \quad \text{in } D(T, R).$$

(iii) *If* $(x_n \circ y_n - b_n c_n e) \to x$, $c_n(y_n - b_n e) \to z$, $F(x_n, y_n;\ a, b) \to 0$ *for all a and b which are continuity points of z or endpoints of T, and* $\mathrm{Disc}(x) \cap \mathrm{Disc}(z) = \emptyset$ *then*

$$(x_n - c_n e) \to x \circ y^{-1} - z \circ y^{-1} \quad \text{in } D(T, R).$$

PROOF. (i) A direct application of Theorems 3.1 and 4.1 does the job because $(x_n \circ y_n - b_n c_n e) = (x_n - c_n e) \circ y_n + c_n(y_n - b_n e)$.

(ii) By Theorem 4.1, $(x_n - c_n e) \circ y_n = (x_n \circ y_n - b_n c_n e) - c_n(y_n - b_n e) \to x - z$. Now apply Theorem 3.3.

(iii) Apply Theorem 4.1 as in (ii), then use Corollary 2 to Theorem 3.4. ∎

REMARK. This relatively trivial extension of previous results is included because there are many stochastic applications. As a simple example, suppose (X, Y) is a random element of $D([0, \infty), R) \times D_0([0, \infty), R)$ such that $(U_n, V_n) \Rightarrow (U, V)$, where $U_n(t) = n^{-\alpha}[X(nt) - n\lambda_n t]$ and $V_n(t) = n^{-\alpha}[Y(nt) - n\mu_n t]$, $t \geq 0$, with λ_n, μ_n, and α being constants such that $\alpha \leq 1$, $\mu_n \to \mu > 0$ and $\lambda_n \to \lambda$. To put this in our setting, let $x_n(t) = n^{-\alpha}X(nt)$, $y_n(t) = n^{-1}Y(nt)$, $c_n = n^{1-\alpha}\lambda_n$, and $b_n = \mu_n$. Then $y_n \to \mu e$ by the CMT because $V_n \to V$, and $[(x_n - c_n e), y_n, c_n(y_n - b_n e)] \Rightarrow [U, \mu e, V]$ by Theorem 4.4 of Billingsley (1968). If $P(\mathrm{Disc}(U \circ \mu e) \cap \mathrm{Disc}(V) \neq \emptyset) = 0$, then $W_n \Rightarrow U \circ \mu e + V$ by the CMT and Theorem 5.1(i), where $W_n(t) = n^{-\alpha}[X(Y(nt)) - \lambda_n \mu_n nt]$, $t \geq 0$. In this way, Corollary 1 and Theorem 1 of Iglehart and Kennedy (1970) are extended. See Serfozo (1973), (1975) and references there for more applications of §§3–5.

6. Supremum. Throughout §§6 and 7 let $S = R$ and let T have 0 as a closed left endpoint. For any $x \in D$, let $x^{\uparrow}(t) = \sup_{0 \leq s \leq t} x(s)$, $t \in T$. (We could look at $x^{\uparrow}(s, t) = \sup_{s \leq u \leq t} x(u)$, $u \in T$, which would put $x^{\uparrow}$ in the space D with a two-dimensioanl parameter space, cf. Straf (1970), but we have chosen not to consider multidimensional parameter spaces in this paper.) The supremum function $\uparrow$ is easily seen to be continuous in each of Skorohod's (1956) topologies. In fact, with the metric d in (2.1) or (2.2), we have a Lipschitz property which does not hold for the previous functions.

THEOREM 6.1. *For all* $x, y \in D$, $d(x^{\uparrow}, y^{\uparrow}) \leq d(x, y)$.

PROOF. If $x^{\uparrow}(t) > y^{\uparrow}(\lambda(t)) + \epsilon$ for some $t \in T$ and $\lambda \in \Lambda$, then there is an $s \leq t$ such that $x(s) > y^{\uparrow}(\lambda(t)) + \epsilon > y(\lambda(s)) + \epsilon$. Similarly, if $y^{\uparrow}(t) > x^{\uparrow}(\lambda^{-1}(t)) + \epsilon$, then $y(s) > x(\lambda^{-1}(s)) + \epsilon$, so $x(s) < y(\lambda(s)) - \epsilon$. ∎

Below we will have occasion to use Skorohod's (1956) M_1 topology. Virtually everything up to this point, including Theorem 2.6, carries over. The M_1 topology is weaker than the J_1 topology and coincides with the topology of pointwise convergence on the subset D_0 of nondecreasing real-valued functions in D. We say that $x_n \to x(A)$ if the A topology is used on D, where $A = J_1$, M_1 or U (uniform convergence on compact intervals). Let c_n, $n > 1$, be real numbers.

THEOREM 6.2. *Suppose that* $(x_n - c_n e) \to x(J_1)$.

(i) *If* $c_n \to c$, *then* $x_n^{\uparrow} \to (x + ce)^{\uparrow}(J_1)$.

(ii) *If* $c_n \to +\infty$ *and x has no negative jump, then* $(x_n^{\uparrow} - c_n e) \to x(J_1)$.

(iii) *If* $c_n \to -\infty$, *then* $x_n^{\uparrow} \to y(U)$, *where* $y(t) = x(0)$, $t \in T$.

PROOF. (i) By Theorem 4.1, $x_n \to x + ce$. Then apply Theorem 6.1.

(ii) It suffices to use the compact domain $[0, b]$ where b is a continuity point of x or the right endpoint of T. Let $\lambda_n \in \Lambda[0, b]$ be such that $\rho(x_n - c_n e, x \circ \lambda_n) \to 0$ and $\rho(\lambda_n, e) \to 0$. Since $x_n \leqslant x_n^\uparrow$, it suffices to show for any $\epsilon > 0$ that there exists an n_0 such that $x_n^\uparrow(t) - c_n t - x(\lambda_n(t)) \leqslant \epsilon$, $0 \leqslant t \leqslant b$, $n \geqslant n_0$. Since $x_n - c_n e \to x$, there is an n_1 such that $x_n(s) - c_n s - x(\lambda_n(s)) \leqslant \epsilon/2$ for all $s \leqslant t$, $n \geqslant n_1$. Since x has no negative jumps and $c_n \to \infty$, there exists an $n_0 \geqslant n_1$ such that $x(\lambda_n(s)) - x(\lambda_n(t)) \leqslant c_n(t - s) + \epsilon/2$ for all $s \leqslant t$, $n > n_0$. Adding completes the proof.

(iii) Since $x_n^\uparrow(s) \geqslant x_n^\uparrow(0) = x_n(0) \to x(0)$ as $n \to \infty$, it suffices to show for each $t > 0$ and $\epsilon > 0$ that there exists an n_0 such that $x_n(s) \leqslant x(0) + \epsilon$, $0 \leqslant s \leqslant t$, $n \geqslant n_0$. Using the right-continuity of x, choose $t_0 > 0$ so that $x^\uparrow(t_0) < x(0) + \epsilon/2$. Choose n_0 so that $c_n < 0 \wedge (-2(x^\uparrow(t)/t_0))$, $d_t(x_n - c_n e, x) < \epsilon/2$, and $\rho_t(\lambda_n, e) < t_0/2$ for $n \geqslant n_0$. Then $x_n(s) - c_n s \leqslant x(\lambda_n(s)) + \epsilon/2$, $0 \leqslant s \leqslant t$, so that $x_n(s) \leqslant c_n s + x^\uparrow(t_0) + \epsilon/2 < x(0) + \epsilon$ if $s < t_0/2$ and $x_n(s) \leqslant c_n s + x^\uparrow(t) + \epsilon/2 < -x^\uparrow(t) + x^\uparrow(t) + \epsilon/2 \leqslant \epsilon/2$ if $t_0/2 \leqslant s < t$. ∎

THEOREM 6.3. *Suppose that* $(x_n - c_n e) \to x(M_1)$.

(i) *If* $c_n \to c$, *then* $x_n^\uparrow \to (x + ce)^\uparrow(M_1)$.

(ii) *If* $c_n \to +\infty$, *then* $(x_n^\uparrow - c_n e) \to x(M_1)$.

(iii) *If* $c_n \to -\infty$, *then* $x_n^\uparrow \to y(U)$ *where* $y(t) = x(0)$, $t \in T$.

PROOF. We only prove (ii). The argument in the proof of Theorem 6.2(ii) applies again, but since the parametric representations move continuously along the completed graphs, it is not necessary to prohibit negative jumps in x. Again, let b be a continuity point of x. If $(x_n - c_n e) \to x$, there exist parametric representations $[y_n(s), t_n(s)]$ and $[y(s), t(s)]$ of the completed graphs of $(x_n - c_n e)$ and x on $[0, b]$ such that $\rho(y_n, y) + \rho(t_n, t) \to 0$, where y_n, y, t_n, and t are continuous functions on $[0, 1]$ say. Since e is the identity, $y_n = z_n - c_n t_n$ on $[0, 1]$, where $[z_n, t_n]$ is a parametric representation of the completed graph of x_n. As before, since $x_n^\uparrow \geqslant x_n$, it suffices to show that $z_n(r) - c_n t_n(s) \leqslant y(s) + \epsilon$, $0 \leqslant r \leqslant s \leqslant 1$, for sufficiently large n. Since $c_n \to \infty$ and y is bounded and continuous on $[0, 1]$, $y(r) \leqslant y(s) + c_n[t_n(s) - t_n(r)] + \epsilon/2$, $0 \leqslant r \leqslant s \leqslant 1$. Since $x_n - c_n e \to x(M_1)$, $z_n(r) - c_n t_n(r) \leqslant y(r) + \epsilon/2$, $0 \leqslant r \leqslant 1$. Adding completes the proof. ∎

Needless to say, corresponding results hold for the infimum function $x^\downarrow$ because $x^\downarrow = -(-x)^\uparrow$. Also closely related is the function $f : D \to D$ defined for any $x \in D$ by $f(x) = x - x^\downarrow$. The function f, which corresponds to the addition of an impenetrable barrier at the origin, frequently arises in the study of queues. The following is a generalization of Theorem 1 of Iglehart and Whitt (1970). Let $\theta(t) = 0$, $t \in T$.

THEOREM 6.4. *Suppose* $(x_n - c_n e) \to x(J_1)$, *where* $x(0) = 0$.

(i) *If* $c_n \to c$, *then* $f(x_n) \to f(x + ce)(J_1)$.

(ii) *If* $c_n \to +\infty$, *then* $f(x_n) - c_n e \to x(J_1)$.

(iii) *If* $c_n \to -\infty$ *and* x *has no positive jumps, then* $f(x_n) \to \theta(U)$.

PROOF. (i) By Theorem 4.1, $x_n \to x + ce$, but then f is continuous in each of Skorohod's (1956) topologies. In fact, an analog of Theorem 6.1 is easy to prove.

(ii) By Theorem 4.1 and 6.2(iii), $f(x_n) - c_n e = x_n - c_n e - x_n^\downarrow = x_n - c_n e + (-x_n)^\uparrow \to x$.

(iii) Note that $f(x_n) = x_n - c_n e - (x_n^\downarrow - c_n e)$. By Theorem 6.2(ii), $(x_n^\downarrow - c_n e) \to x$. Theorem 4.1 cannot be applied for the subtraction unless $x \in C$, but the same homeomorphsims λ_n can be used for both $x_n - c_n e \to x$ and $x_n^\downarrow - c_n e \to x$. Hence, $f(x_n) \to \theta$.

7. **Inverse or first passage time.** It is convenient to discuss first passage times in the subset E of functions in $D([0, \infty), R)$ which are unbounded above and have

$x(0) \geqslant 0$. Since $E = \bigcap_{n=1}^{\infty} B_n^c$, where B_n is the closed subset of functions in D which are bounded above by n and have $x(0) \geqslant 0$, E is a G_δ subset. For $x \in E$, let the first passage time function be defined by $x^{-1}(t) = \inf\{s \geqslant 0 : x(s) > t\}$, $t \geqslant 0$. It is easy to see that $x^{-1} \in E \cap D_0([0, \infty), [0, \infty))$. As in §6, we use the J_1, M_1 and U topologies on D. To start with, the first passage time function is not continuous in the J_1 topology. For example, let $x = 2I_{[0,2)} + eI_{[2,\infty)}$ and $x_n = (2 - n^{-1})I_{[0,1)} + (2 + n^{-1}) I_{[1,2+n^{-1})} + eI_{[2+n^{-1},\infty)}$; then $x_n^{-1} \to x^{-1}(M_1)$ but not (J_1) since $x_n^{-1}(1) = 1$ for all $n \geqslant 1$.

Theorem 7.1. *The first passage time function mapping $E(M_1)$ into $E \cap D_0(M_1)$ is continuous.*

Proof. Whitt (1971). It suffices to look at nondecreasing functions in E because the supremum function is continuous (M_1). Then each parametric representation $[y, t]$ of the completed graph of x can serve as a parametric representation of the completed graph of x^{-1} when the roles of y and t are switched.

Remark. Continuity of the supremum and first passage time functions in the M_1 topology is not as useful as it might appear because M_1 convergence on D_0 is equivalent to pointwise convergence at all continuity points of the limit functior plus all closed bounded endpoints of T. This follows from 2.4.1 of Skorohod (1956) because $\Delta_{M_1}(c, x) = 0$ for all $x \in D_0$. Furthermore, weak convergence of random functions in (D_0, M) is characterized by weak convergence of the corresponding finite-dimensional distributions for $t \in T_P$.

Theorem 7.2. *The first passage time function mapping $E(J_1)$ into $E \cap D_0(J_1)$ is measurable and continuous at each strictly increasing x.*

Proof. (measurability) By Theorem 7.1, continuity implies measurability if the M_1 topology is used. However, the σ-fields generated by the M_1 and J_1 topologies coincide. This can be seen from Lemma 2.7 plus a corresponding result for the M_1 topology. More generally, the topological σ-fields associated with two comparable Souslin topologies on the same space coincide by virtue of the separation lemma; p. 124 of Schwartz (1973). Alternatively, by an analog of Lemma 2.7 for M_1, it suffices to show that the mapping $x \to x^{-1}(t)$ is measurable for each t, but this follows because $x^{-1}(t) \geqslant a$ is equivalent to $x(a) \leqslant t$, cf. Lemma 4 of Vervaat (1972).

(continuity) Since $x_n \to x(J_1)$, $x_n \to x(M_1)$ because the M_1 topology is weaker than the J_1. Then $x_n^{-1} \to x^{-1}(M_1)$ by Theorem 7.1. Since $x^{-1} \in C$, $x_n^{-1} \to x^{-1}(U)$ and thus also (J_1). ∎

As a dual to Theorem 6.2(ii), we have the following result (due to W. Vervaat, who in turn was inspired by (2.9) of Gut (1975)).

Theorem 7.3. *If $c_n \to \infty$, $x_n \in E$, x has no positive jumps, $x(0) = 0$, and $c_n(x_n - e) \to x(J_1)$, then $c_n(x_n^{-1} - e) \to -x(J_1)$.*

Remark. The need for the condition $x(0) = 0$ was noted by Bauer [2, p. 49].

Proof. Since $c_n(x_n - e) \to x(J_1)$, $x_n \to e(U)$. By Theorem 7.2, $x_n^{-1} \to e(U)$. By Theorem 3.1, $c_n(x_n \circ x_n^{-1} - x_n^{-1}) \to x$. By Theorem 4.1, it suffices to show that $c_n(x_n \circ x_n^{-1} - e) \to 0$. Let $J_t^+(x)$ be the supremum of the positive jumps of x in $[0, t]$, where $x(0)$ is interpreted as the jump at $t = 0$. Since x has no positive jumps and $c_n(x_n - e) \to x$, $c_n J_t^+(x_n) \to 0$. However, $x_n \circ x_n^{-1}(t) - t \leqslant J^+_{x_n^{-1}(t)}(x_n)$. Hence, $c_n(x_n \circ x_n^{-1} - e)(t) \leqslant c_n J_{2b}^+(x_n) < \epsilon$ for all $t \leqslant b$ and all n sufficiently large. ∎

Remark. The special case of Theorem 7.3 in which $x \in C$ is covered by Theorem 6.2 here plus Theorem 1 of Iglehart and Whitt (1971) or Lemmas 1 and 2 of Vervaat (1972).

The following result shows that $x \in C$ is necessary if we want convergence for both the supremum and first passage times.

THEOREM 7.4. *If* $c_n \to \infty$, $x_n \in E$, $c_n(x_n^{\uparrow} - e) \to x(J_1)$ *and* $c_n(x_n^{-1} - e) \to -x(J_1)$, *then* $x \in C$.

PROOF. Since the J_1 topology is in effect and $c_n(x_n^{\uparrow} - e)$ has no negative jumps, neither does x. Since $c_n(x_n^{-1} - e)$ has no negative jumps, neither does $-x$. ∎

Corresponding to Theorem 6.3(ii), we have

THEOREM 7.5. *Let* $x_n \in E$ *and* $c_n \to \infty$. *If* $c_n(x_n - e) \to x(M_1)$ *then* $c_n(x_n^{-1} - e) \to -x(M_1)$.

PROOF. By Theorem 6.3, we can assume x_n is nondecreasing for each n. Let $[y_n, t_n]$ and $[y, t]$ be parametric representations of the graphs of $c_n(x_n - e)$ and x associated with $c_n(x_n - e) \to x$. If $z_n = c_n^{-1}y_n + t_n$, then $[z_n, t_n]$, $[t_n, z_n]$, and $[c_n(t_n - z_n), z_n]$ are parametric representations for x_n, x_n^{-1}, and $c_n(x_n^{-1} - e)$, respectively. Since $y_n \to y$ and $t_n \to t$ uniformly on compact sets, $z_n = c_n^{-1}y_n + t_n \to t$ and $c_n(t_n - z_n) = -y_n \to -y$ uniformly on compact sets. Hence, $c_n(x_n^{-1} - e) \to -x(M_1)$. ∎

Applications are facilitated by the easily proved

LEMMA 7.6. *If* $x \in E$ *and* λ *is a homeomorphism of* $[0, \infty)$, *then* $(x^{-1}) \circ \lambda = (\lambda^{-1} \circ x)^{-1}$ *and* $(x \circ \lambda)^{-1} = \lambda^{-1} \circ x^{-1}$.

COROLLARY. *Let* $x_n \in E$, $c_n \to \infty$, *and* $a > 0$. *If* $c_n(x_n - ae) \to x(M_1)$ *or* (J_1) *if* x *has no positive jumps, then* $c_n(x_n^{-1} - a^{-1}e) \to -a^{-1}x \circ a^{-1}e$ *in the same topology*.

PROOF. Since $c_n(x_n - ae) = ac_n(a^{-1}x_n - e)$, $ac_n([a^{-1}x_n]^{-1} - e) \to -x$ by Theorems 7.3 and 7.5. Since $[a^{-1}x_n]^{-1} = x_n^{-1} \circ ae$ by Lemma 7.6, $ac_n(x_n^{-1} \circ ae - e) \to -x$. Finally, $ac_n(x_n^{-1} - a^{-1}e) \to -x \circ a^{-1}e$ by Theorem 3.1 and $c_n(x_n^{-1} - a^{-1}e) \to -a^{-1}x \circ a^{-1}e$. ∎

REMARK. Distributional complements to the results in §§6 and 7 plus an indication of possible applications are contained in Bingham (1973) and references there. First passage times with more general boundary functions are treated by Chow and Hsiung (1976), Gut (1973), (1975), Lindberger (1978), Mohan (1976) and Pakshirajan and Mohan (1978). Recent related results are contained in Bauer (1978), Goldie (1977) and de Haan and Resnick (1978).

8. Time reversal. Let D be $D([0, 1], S)$ where (S, m) is a complete separable metric group under $+$ and m is translation invariant ($S = R^k$ is one possibility). Use the metric d in (2.1). Assume that $x(1) = x(1-)$ for all $x \in D$, which amounts to looking at a closed subset. Let D_θ be the closed subset of D in which $x(0) = \theta$, where θ is the identity in $(S, +)$. Let $R : D \to D$ be the *reverse-time function* defined for any $x \in D$ by

$$R(x)(t) = \begin{cases} \lim_{s \downarrow t+} x(1 - s), & 0 \leq t < 1, \\ x(0), & t = 1. \end{cases}$$

Let $r : D \to D_\theta$ be the *reverse-time function starting at* θ, defined for any $x \in D$ by $r(x)(t) = R(x)(t) - x(1)$, $0 \leq t \leq 1$. To see how the time reversal functions can be applied, note that $f(x) = r(x)^{\uparrow}$ for $x \in D_\theta([0, 1], R)$ and f the barrier function in Theorem 6.4.

THEOREM 8.1. $d[R(x), R(y)] = d(x, y)$ *on* $D \times D$ *and* $d[r(x), r(y)] \leq 2d(x, y)$ *on* $D \times D$ *and* $D_\theta \times D_\theta$.

PROOF. Obviously the maps are one-to-one and onto. Note that $(-r)$ maps Λ into itself and $(-r)(e) = e$, where e is the identity in Λ. Note that $\rho(Rx, Ry) = \rho(x, y)$ and $\rho(rx, ry) \leqslant \rho(x, y) + m[x(1), y(1)] \leqslant 2\rho(x, y)$. Then note that $R(x) \circ (-r)(\lambda) = R(x \circ \lambda)$ for $(x, \lambda) \in D \times \Lambda$ and $r(x) \circ (-r)(\lambda) = r(x \circ \lambda)$ for $(x, \lambda) \in D_0 \times \Lambda$. For any $(x, y) \in D \times D$, let $d_\lambda(x, y)$ be the minimum of $\rho(x, y \circ \lambda)$ and $\rho(\lambda, e)$. Then

$$\begin{aligned} d_{(-r)(\lambda)}(Rx, Ry) &= \rho(Rx, Ry \circ (-r)\lambda) \vee \rho((-r)\lambda, e) \\ &= \rho(Rx, R(y \circ \lambda)) \vee \rho((-r)\lambda, (-r)e) \\ &= \rho(x, y \circ \lambda) \vee \rho(\lambda, e) \\ &= d_\lambda(x, y). \end{aligned}$$

Consequently, $d(Rx, Ry) = d(x, y)$. The same argument applies to r. ∎

Acknowledgments. I am grateful to Herbert Bauer, David Pollard, Jean-Michel Pomarede, Richard Serfozo, Wim Vervaat, Michael J. Wichura and the referees for their assistance.

References

[1] Aldous, D. J. (1978). Weak Convergence of Randomly Indexed Sequences of Random Variables. *Math. Proc. Cambridge Philos. Soc.* **83** 117–126.

[2] Bauer, H. (1978). *Invarianzprinzipien für Erneuerungsprozesse und für ihre Verallgemeinerungen*. Ph.D. dissertation, Mathematics Institute, University of Stuttgart.

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

[4] ———. (1971). *Weak Convergence of Measures: Applications in Probability*. SIAM Regional Conference Series in Applied Mathematics, Philadelphia.

[5] Bingham, N. H. (1973). Maxima of Sums of Random Variables and Suprema of Stable Processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **26** 273–296.

[6] Chow, Y. S. and Hsiung, A. C. (1976). Limiting Behavior of $\max_{j \leqslant n} S_j / j^{-\alpha}$ and First Passage Time in a Random Walk with Positive Drift. *Bull. Inst. Math. Acad. Sinica.* **4** 35–44.

[7] Dudley, R. M. (1968). Distances of Probability Measures and Random Variables. *Ann. Math. Statist.* **39** 1563–1572.

[8] Dugundji, J. (1968). *Topology*. Allyn and Bacon, Boston.

[9] Durrett, R. T. and Resnick, S. I. (1977). Weak Convergence with Random Indices. *Stochastic Processes Appl.* **5** 213–220.

[10] Goldie, C. M. (1977). Convergence Theorems for Empirical Lorenz Curves and Their Inverses. *Advances in Appl. Probability.* **9** 765–791.

[11] Gut, A. (1973). A Functional Central Limit Theorem Connected with Extended Renewal Theory. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **27** 123–129.

[12] ———. (1975). Weak Convergence and First Passage Times. *J. Appl. Probability.* **12** 324–332.

[13] Haan, L. de and Resnick, S. I. (1978). Conjugate π-variation and Process Inversion. Report 7801/S, Erasmus University, Rotterdam.

[14] Iglehart, D. L. and Kennedy, D. P. (1970). Weak Convergence for the Average of Flag Processes. *J. Appl. Probability.* **7** 747–753.

[15] ——— and Whitt, W. (1970). Multiple Channel Queues in Heavy Traffic, II: Sequences, Networks, and Batches. *Advances in Appl. Probability.* **2** 355–369.

[16] ——— and ———. (1971). The Equivalence of Functional Central Limit Theorems for Counting Processes and Associated Partial Sums. *Ann. Math. Statist.* **42** 1372–1378.

[17] Kennedy, D. P. (1972). The Continuity of the Single Server Queue. *J. Appl. Probability.* **9** 370–381.

[18] ———. (1978). The Stability of Queueing Systems. *Proceedings of the 41st Session of the International Statistical Institute. New Delhi, December 1977.*

[19] Kolmogorov, A. N. (1956). On Skorohod Convergence. *Theor. Probability Appl.* **1** 215–222.

[20] Lindberger, K. (1978). Functional Limit Theorems for Cumulative Processes and Stopping Times. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **44** 47–56.

[21] Lindvall, T. (1973). Weak Convergence of Probability Measures and Random Functions in the Function Space $D[0, \infty)$. *J. Appl. Probability.* **10** 109–121.

[22] Mohan, N. R. (1976). A Note on J_1-Convergence of a First-passage Process. *Stochastic Processes Appl.* **4** 157–165.

[23] Pakshirajan, R. P. and Mohan, N. R. (1978). Some Functional Limit Theorems for Multidimensional Random Walks. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **41** 263–271.

[24] Parthasarathy, K. R. (1967). *Probability Measures on Metric Spaces*. Academic Press, New York.
[25] Pollard, D. (1977). Induced Weak Convergence and Random Measures. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **37** 321–328.
[26] Pomarede, J. L. (1976). *A Unified Approach via Graphs to Skorohod's Topologies on the Function Space D*. Ph.D. dissertation, Department of Statistics, Yale University.
[27] ———. (1976a). On the Continuity of Functions on *D*. Department of Mathematics, Boston College.
[28] Schwartz, L. (1973). *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. Oxford University Press, London.
[29] Serfozo, R. F. (1973). Weak Convergence of Superpositions of Randomly Selected Partial Sums. *Ann. Probability*. **1** 1044–1056.
[30] ———. (1975). Functional Limit Theorems for Stochastic Processes Based on Embedded Processes. *Advances in Appl. Probability*. **7** 123–139.
[31] Silvestrov, D. (1972). The Convergence of Composite Random Functions in the *J*-Topology. *Soviet Math. Dokl.* **13** 152–154.
[32] Skorohod, A. V. (1956). Limit Theorems for Stochastic Processes. *Theor. Probability Appl.* **1** 261–290.
[33] Stone, C. (1963). Weak Convergence of Stochastic Processes Defined on Semi-Infinite Time Intervals. *Proc. Amer. Math. Soc.* **14** 694–696.
[34] Straf, M. L. (1970). Weak Convergence of Stochastic Processes with Several Parameters. *Proc. Sixth Berk. Symp. Math. Stat. Prob.* II. 187–221.
[35] Tucker, H. G. (1967). *A Graduate Course in Probability*. Academic Press, New York.
[36] Vervaat, W. (1972). Functional Central Limit Theorems for Processes with Positive Drift and Their Inverses. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*. **23** 245–253.
[37] Whitt, W. (1971). Weak Convergence of First Passage Time Processes. *J. Appl. Probability*. **8** 417–422.
[38] ———. (1973). Representation and Convergence of Point Processes on the Line. Department of Administrative Sciences, Yale University. Unpublished paper.
[39] ———. (1974). Heavy Traffic Limit Theorems for Queues: A Survey. *Mathematical Methods in Queueing Theory*, A. B. Clarke, ed. Lecture Notes in Economics and Mathematical Systems No. 98, Springer-Verlag, New York, 307–350.
[40] ———. (1974a). The Continuity of Queues. *Advances in Appl. Probability* **6** 175–183.
[41] Wichura, M. J. (1970). On the Construction of Almost Uniformly Convergent Random Variables with Given Weakly Convergent Image Laws. *Ann. Math. Statist.* **41** 284–291.
[42] ———. (1974). Functional Laws of the Interated Logarithm for the Partial Sums of i.i.d. Random Variables in the Domain of Attraction of a Completely Asymmetric Stable Law. *Ann. Probability*. **2** 1108–1138.
[43] Zolotarev, V. M. (1978). General Problems of the Stability of Mathematical Models. *Proceedings of the 41st Session of the International Statistical Institute, New Delhi, December 1977.*

BELL LABORATORIES, HOLMDEL, NEW JERSEY 07733