The Annals of Probability
1979, Vol. 7, No. 1, 85–95

LAWS OF LARGE NUMBERS FOR $D[0, 1]$

BY PETER Z. DAFFER AND ROBERT L. TAYLOR

University of South Carolina

Laws of large numbers are obtained for random variables taking their values in $D[0, 1]$ where $D[0, 1]$ is equipped with the Skorokhod topology. The strong law of large numbers is obtained for independent, convex tight random elements $\{X_n\}$ satisfying $\sup_n E\|X_n\|_\infty^r < \infty$ for some $r > 1$ where $\|X\|_\infty = \sup_{0 \leqslant t \leqslant 1}|x(t)|$. A strong law of large numbers is also obtained for almost surely monotone random elements in $D[0, 1]$ for which the hypothesis of convex tightness is not needed. A discussion of the condition of convex tightness is also included.

1. Introduction.. Let D denote the space of functions x: $[0, 1] \to R$ which are right-continuous and possess left-hand limits at each $t \in [0, 1]$. Suppose D is equipped with the topology generated by the Skorokhod metric d, and let $\mathfrak{D}$ denote the Borel field of this topology. Let (X_n) denote a sequence of random elements (r.e.'s) in $(D, \mathfrak{D})$ and put $\bar{X}_n = n^{-1}\Sigma_{k=1}^n X_k$. Some laws of large numbers (LLN's) are proved for D, i.e., conditions under which

$$\lim_{n\to\infty} d(\bar{X}_n, \bar{\mu}_n) = 0 \tag{1.1}$$

for some mode of convergence, where $\bar{\mu}_n = n^{-1}\Sigma_{k=1}^n EX_k$.

The fact that the Skorokhod topology on D is not locally convex motivates the definition of convex tightness of a sequence (X_n) of r.e.'s in D (or of their probability measures). The main result is Theorem 1 which states that the strong law of large numbers (SLLN) holds (i.e., (1.1) holds almost surely (a.s.)) for independent convex tight r.e.'s in D satisfying $\sup_n E\|X_n\|_\infty^r < \infty$ for some $r > 1$, where $\|x\|_\infty = \sup_{0 \leqslant t \leqslant 1}|x(t)|$. Theorem 2 provides a SLLN for almost surely monotone r.e.'s in D, for which the hypothesis of convex tightness is not needed.

When the r.e.'s (X_n) take values in a complete subspace E of D which is separable with respect to the uniform topology on D, convergence can be treated in the uniform topology relativized to the subspace E, and the whole arsenal of results for r.e.'s taking values in a Banach space can be applied. This situation is also discussed in Section 3.

In general, the use of the uniform topology does not suffice, and the convergence of the r.e.'s must be investigated using the Skorokhod topology. The Skorokhod topology on D is separable and has proved extremely fruitful for the study of convergence of probability measures on D, and many central limit results and invariance principles exist (see Billingsley [1]). The Skorokhod topology, however,

Received August 4, 1977; revised January 3, 1978.

AMS 1970 *subject classifications.* Primary 60B05, 60F15; Secondary 60G99.

Key words and phrases. Skorokhod, laws of large numbers, convex tightness, compact, convergence almost surely and in probability.

has many troublesome properties which emerge in studying almost sure convergence or even convergence in probability for D. The metric d is not translation invariant, addition is not a continuous operation, and the topology is not locally convex.

The strong law of large numbers for independent, identically distributed r.e.'s in D was obtained by Rao ([8]). The central limit theorem in D has been investigated in detail by Hahn ([3], [4]), and a law of the iterated logarithm and several consequences thereof have been obtained by Kuelbs ([7]). The convergence of random series in D has been investigated by Kallenberg ([6]).

2. Preliminaries. Let $d(x, y)$ denote the familiar Skorokhod metric on D and let $\|x\|_\infty = \sup_{t\in[0,1]}|x(t)|$. Without express mention to the contrary the space D will be taken to be equipped with the Skorokhod topology. The Borel field $\mathfrak{D}$ on D is generated by the finite-dimensional cylinder sets, and a random element in D is a measurable map from a probability space $(\Omega, \mathfrak{F}, P)$ into $(D, \mathfrak{D})$. A random element X in D is characterized by the property that $X(t)$ is a random variable for each $t \in [0, 1]$ ([1], page 128).

With the Skorokhod topology, D is not a topological vector space (addition is not continuous, for example) and the use of the Pettis integral for expected value does not follow naturally. However, the expectation EX can be defined pointwise by $(EX)(t) = E[X(t)]$, $t \in [0, 1]$, provided that $EX \in D$. The "simplest" sufficient condition yielding $EX \in D$ is $E\|X\|_\infty < \infty$, and will be included in the hypothesis of existence of EX. For $J \subset [0, 1]$ let

$$w_x(J) = \sup_{s,t\in J}|x(s) - x(t)|,$$

and define

$$w'_x(\delta) = \inf \max_{i=1,\cdots,m} w_x([t_{i-1}, t_i)),$$

where $\delta > 0$, and the infimum is taken over all partitions $0 = t_0 < t_1 < \cdots < t_{m-1} < t_m = 1$ of $[0, 1]$ such that $\min_i\{t_i - t_{i-1}\} > \delta$. Then $x \in D$ if and only if $\lim_{\delta\to 0} w'_x(\delta) = 0$ ([1], page 110).

The following operator will play a fundamental role. Define $J_{m,i} = [(i-1)2^{-m}, i2^{-m})$, $i = 1, \cdots, 2^m - 1$; and $J_{m,2^m} = [(2^m - 1)2^{-m}, 1]$; $m = 1, 2, \cdots$. For $x \in D$ define

$$T_m x = \sum_{i=1}^{2^m} x\left(\frac{i-1}{2^m}\right) I_{J_{m,i}} \tag{2.1}$$

where I_J is the indicator function of the set J. Then for each positive integer m, $T_m : D \to D_m \subset D$ and T_m is linear (however, T_m is not continuous). We have $\dim(D_m) = 2^m$, $D_m \subset D_{m+1}$ and $\cup D_m$ is dense in D.

The following facts about T_m are easily established:

(1) $\lim_{m\to\infty} d(x, T_m x) = 0$ for each $x \in D$;
(2) $d(x, T_m x) < w'_x(2^{-m}) + 2^{-m}$, for each $x \in D$ and each m.

The following lemma thus follows directly from (2).

LEMMA 1. *If K is a compact subset of D, then*

$$\lim_{m\to\infty}\sup_{x\in K} d(x, T_m x) = 0. \tag{2.2}$$

The operation of building the closed convex hull of a set does not preserve compactness in D. This motivates the following definition.

DEFINITION. A family (X_n) of random elements taking values in a linear space E with a topology τ is said to be *convex tight* if to each $\varepsilon > 0$ there is a convex τ-compact subset K of E such that

$$\sup_n P[X_n \notin K] \leqslant \varepsilon.$$

If E is a Fréchet space, then the closed convex hull of a compact subset of E is again compact ([9], page 72, Theorem 3.25), and the notion of convex tightness coincides with the classical notion of tightness ([1], page 37).

The need for convexity arises from the desired condition that a convex combination of elements $\{x_i\}$ of a set $K \subset D$, in particular, $\Sigma_{i=1}^{n}(1/n)x_i$, again belong to K.

The following lemma easily follows and will be used in obtaining the laws of large numbers for $D[0, 1]$.

LEMMA 2. *If $x, y, u, v \in D$, then*

$$d(x + u, y + v) \leqslant d(x, y) + \|u\|_\infty + \|v\|_\infty.$$

3. Laws of large numbers for random elements in D. We first investigate to what extent Banach space results can be applied to obtain laws of large numbers for r.e.'s in D.

Consider a linear subspace E of D and suppose that E is a metric space. Denote the metric topology on E by τ and suppose that (E, τ) is an analytic space. Then, if τ' is any other Hausdorff topology on E such that $\tau' \subset \tau$, the Borel sets of (E, τ') are exactly the Borel sets of (E, τ): for this see Hoffmann-Jørgensen ([5], page 112, Corollary 6). From this it follows that if E is a complete separable subspace of D in the uniform topology, then a map $X : \Omega \to D$ such that $X \in E$ a.s. is a random element in D with the uniform topology if and only if X is a random element in D with the Skorokhod topology. This leads to a generic law of large numbers for a sequence (X_n) of random elements in D taking values almost surely in a closed linear subspace E of D which is separable with respect to the uniform topology on D:

$$(H) \Rightarrow \text{SLLN or WLLN in } \boldsymbol{D} \text{ with the uniform topology}$$

implies that

$$(H) \Rightarrow \text{SLLN or WLLN in } \boldsymbol{D} \text{ with the Skorokhod topology,}$$

where (H) denotes an hypothesis on (X_n) and/or E.

Of course, the hypothesis (H) will involve conditions yielding the law of large numbers (strong or weak) in the Banach space D with the uniform topology, and the question can still be asked: which of these conditions can be weakened to involve only the Skorokhod metric and still yield laws of large numbers in D with the Skorokhod topology?

Obviously there are many random elements whose ranges do not have separable support with respect to the uniform topology. However, the concept of convex tight random elements will yield several laws of large numbers.

THEOREM 1. *Let* (X_n) *be a sequence of independent convex tight* r.e.'s *in* D *satisfying* $\sup_n E\|X_n\|_\infty^r \leq \Gamma$, *where* $r > 1$ *and* Γ *is a constant. Then*

$$\lim_{n\to\infty} d\left(n^{-1}\Sigma_{k=1}^n X_k,\ n^{-1}\Sigma_{k=1}^n EX_k\right) = 0,$$

with probability one.

PROOF. Let $\varepsilon > 0$ be given. By convex tightness let $K \subset D$ be convex and compact such that $P[X_k \notin K] < \varepsilon^{r/(r-1)}$, for all k, and w.l.o.g., it can be assumed that $0 \in K$. Then

$$E\|X_k I_{[X_k \notin K]}\|_\infty \leq \left(E\|X_k\|^r\right)^{1/r} \cdot P\left[X_k \notin K\right]^{(r-1)/r} \leq \Gamma^{1/r}\varepsilon. \tag{3.1}$$

Note that EX_k exists for every k since $E\|X_k\|_\infty \leq \Gamma + 1$.

$$\begin{aligned} d\left(n^{-1}\Sigma_{k=1}^n X_k, n^{-1}\Sigma_{k=1}^n EX_k\right) &= d\Big(n^{-1}\Sigma_{k=1}^n\left(X_k I_{[X_k\in K]} + X_k I_{[X_k\notin K]}\right),\\ &\qquad n^{-1}\Sigma_{k=1}^n\left(EX_k I_{[X_k\in K]} + EX_k I_{[X_k\notin K]}\right)\Big)\\ \text{(I)}\quad &\leq d\left(n^{-1}\Sigma_{k=1}^n X_k I_{[X_k\in K]}, n^{-1}\Sigma_{k=1}^n T_m\left(X_k I_{[X_k\in K]}\right)\right)\\ \text{(II)}\quad &\quad + d\left(n^{-1}\Sigma_{k=1}^n T_m\left(X_k I_{[X_k\in K]}\right), n^{-1}\Sigma_{k=1}^n T_m\left(EX_k I_{[X_k\in K]}\right)\right)\\ \text{(III)}\quad &\quad + d\left(n^{-1}\Sigma_{k=1}^n T_m\left(EX_k I_{[X_k\in K]}\right), n^{-1}\Sigma_{k=1}^n EX_k I_{[X_k\in K]}\right)\\ \text{(IV)}\quad &\quad + \|n^{-1}\Sigma_{k=1}^n X_k I_{[X_k\notin K]}\|_\infty\\ \text{(V)}\quad &\quad + \|n^{-1}\Sigma_{k=1}^n EX_k I_{[X_k\notin K]}\|_\infty. \end{aligned}$$

The above inequality is a consequence of Lemma 2.

This sum can be made $< \varepsilon$ almost surely.

For (I), we have

$$d\left(n^{-1}\Sigma_{k=1}^n X_k I_{[X_k\in K]}, T_m\left(n^{-1}\Sigma_{k=1}^n X_k I_{[X_k\in K]}\right)\right) \leq \sup_{x\in K} d(x, T_m(x)),$$

pointwise in Ω, since K is convex and $0 \in K$. Since $\lim_m \sup_{x\in K} d(x, T_m(x)) = 0$ from Lemma 1, there is m_0 such that (I) $< \varepsilon$ for all $m \geq m_0$, and every sample point $\omega \in \Omega$.

As for (II), using the linearity of T_m:

$$\begin{aligned} \text{(II)} &\leq \|T_m\left(n^{-1}\Sigma_{k=1}^n\left(X_k I_{[X_k\in K]} - E\left[X_k I_{[X_k\in K]}\right]\right)\right)\|_\infty\\ &\leq \Sigma_{i=1}^{2^m}\left|n^{-1}\Sigma_{k=1}^n\left(X_k\left(\frac{i}{2^m}\right)I_{[X_k\in K]} - EX_k\left(\frac{i}{2^m}\right)I_{[X_k\in K]}\right)\right|. \end{aligned}$$

Since K is compact, by Billingsley [1], Theorem 14.3, page 116, there is a constant C such that $\sup_{x\in K}\sup_{t\in I}|x(t)| \leq C$. For each $i = 1, \cdots, 2^m$, $|X_k(i/2^m)I_{[X_k\in K]}|$

$\leqslant C$, and hence, by the strong law of large numbers for random variables

$$\lim_{n\to\infty} n^{-1}\Sigma_{k=1}^{n}\left(X_k\left(\frac{i}{2^m}\right)I_{[X_k\in K]} - E\left[X_k\left(\frac{i}{2^m}\right)I_{[X_k\in K]}\right]\right) = 0 \qquad \text{a.s.}$$

Thus, for each given m, almost surely for sufficiently large n (II) $< \varepsilon$.

By Lemma 1, $\lim_{m\to\infty}\sup_n(\text{III}) = 0$. Then, we have

$$\begin{aligned}(\text{IV}) &\leqslant n^{-1}\Sigma_{k=1}^{n}\|X_k\|_\infty I_{[X_k\notin K]}\\ &= n^{-1}\Sigma_{k=1}^{n}\left[\|X_k\|_\infty I_{[X_k\notin K]} - E\|X_k\|_\infty I_{[X_k\notin K]}\right]\\ &\quad + n^{-1}\Sigma_{k=1}^{n}E\|X_k\|_\infty I_{[X_k\notin K]}.\end{aligned}$$

For the first of these terms we have

$$\begin{aligned}&\Sigma_{k=1}^{\infty}\frac{1}{k^r}E\left|\|X_k\|_\infty I_{[X_k\notin K]} - E\|X_k\|_\infty I_{[X_k\notin K]}\right|^r\\ &\leqslant \Sigma_{k=1}^{\infty}\frac{2^{r-1}}{k^r}\left\{E\|X_k I_{[X_k\notin K]}\|^r + \left(E\|X_k\|_\infty I_{[X_k\notin K]}\right)^r\right\}\\ &\leqslant 2^r\Sigma_{k=1}^{\infty}\frac{1}{k^r}E\|X_k\|_\infty^r I_{[X_k\notin K]} \leqslant 2^r\Sigma_{k=1}^{\infty}\frac{\Gamma}{k^r} < \infty,\end{aligned}$$

since $r > 1$. Hence by Chung's strong law of large numbers, the first term tends to zero almost surely as $n \to \infty$. For the second term, (3.1) yields

$$n^{-1}\Sigma_{k=1}^{n}E\|X_k\|_\infty I_{[X_k\notin K]} < \varepsilon\Gamma^{1/r} \qquad \text{for every} \quad n.$$

Finally, by (3.1) we have $(V) \leqslant \Gamma^{1/r}\cdot\varepsilon$.

Thus, a null set can be excluded for each m, and the countable union Ω_0 is obtained. For $\varepsilon > 0$ and $\omega \notin \Omega_0$, m is chosen large enough so that (I) and (III) are $< \varepsilon$ and then $N(\varepsilon, \omega)$ is chosen large enough so that (II) and (IV) are each $< \varepsilon$. □

The case of identical distributions for Theorem 1 was obtained by Rao [8] without using convex tightness, but a similar use of compact sets was needed.

Unfortunately, not all random elements are convex tight. An example is provided in the next section of a random element in D which is not convex tight. However, a proof similar to the proof of the Glivenko-Cantelli theorem will yield laws of large numbers for a class of random elements in D which may not be convex tight.

Let $D\uparrow$ denote the cone of nondecreasing elements of D.

THEOREM 2. *Let (X_n) be a sequence of independent random elements in D satisfying*

(i) $X_n \in D\uparrow$ *almost surely, for each* n;
(ii) $\Sigma\dfrac{E\|X_n\|_\infty^r}{n^r} < \infty$ *for some* $1 \leqslant r \leqslant 2$;
(iii) $EX_n = EX_1$, *for all* n.

Then

$$\lim_{n\to\infty}\|n^{-1}\Sigma_{k=1}^{n}X_k - EX_1\|_\infty = 0, \qquad \textit{with probability one.}$$

PROOF. Put $\bar{X}_n = n^{-1}\Sigma_{k=1}^{n}X_k$ and $\mu = EX_1$. Note that $E\|X_1\|_\infty^r < \infty$ implies the existence of μ. By Billingsley [1], page 110, Lemma 1, given $m \in N$ there is a partition of [0, 1], $0 = t_0 < t_1 < \cdots < t_{k(m)} = 1$ such that $\sup_{t_i \leqslant t, s < t_{i+1}}|\mu(s) - \mu(t)| < 1/m$, for $i = 0, 1, \cdots, k(m) - 1$. Since $\mu \in D\uparrow$ this means $\mu(t_{i+1} - 0) - \mu(t_i) < 1/m$, for $i = 0, 1, \cdots, k(m) - 1$. Let $t \in [0, 1]$, then $t \in [t_{i-1}, t_i)$ for some $i = 1, 2, \cdots, k(m)$ or $t = 1$. In any case,

$$\begin{aligned}\bar{X}_n(t) - \mu(t) &\leqslant \bar{X}_n(t_i - 0) - \mu(t_{i-1})\\ &\leqslant \bar{X}_n(t_i - 0) - \mu(t_i - 0) + 1/m\end{aligned}$$

and

$$\begin{aligned}\bar{X}_n(t) - \mu(t) &\geqslant \bar{X}_n(t_{i-1}) - \mu(t_i - 0)\\ &\geqslant \bar{X}_n(t_{i-1}) - \mu(t_{i-1}) - 1/m.\end{aligned}$$

Thus,

$$\begin{aligned}&|\bar{X}_n(t) - \mu(t)|\\ &\quad\leqslant \max\{|\bar{X}_n(t_i - 0) - \mu(t_i - 0)|, |\bar{X}_n(t_{i-1}) - \mu(t_{i-1})|\} + 1/m\\ &\quad\leqslant \max_{1\leqslant i\leqslant k(m)}\max\{|\bar{X}_n(t_i - 0) - \mu(t_i - 0)|, |\bar{X}_n(t_{i-1}) - \mu(t_{i-1})|\} + 1/m\end{aligned}$$

$\to_{\text{a.s.}} 0 + 1/m$ by the SLLN using Chung's condition of (ii). Thus $\lim_{n\to\infty}\|\bar{X}_n - \mu\|_\infty \leqslant 1/m$ a.s. and since m is arbitrary, $\lim_{n\to\infty}\|\bar{X}_n - \mu\|_\infty = 0$ a.s. □

REMARK. The conclusion immediately implies convergence in the Skorokhod metric. A corresponding theorem holds mutatis mutandis for almost surely nonincreasing random elements.

Weak laws of large numbers can be obtained in a similar manner. However, the hypotheses for the weak laws of large numbers involve less restrictive conditions (often pointwise conditions will suffice). For comparison to the strong laws of large numbers in this section, the following three results from Taylor and Daffer [10] are listed.

THEOREM 3. *Let (X_n) be a sequence of convex tight* r.e.'s *in D such that $E\|X_n\|_\infty^r \leqslant \Gamma < \infty$, for all n, with $r > 1$. If*

$$\lim_{n\to\infty}n^{-1}\Sigma_{k=1}^{n}(X_k(t) - EX_k(t)) = 0$$

in probability, for each dyadic rational $t \in [0, 1]$, then

$$\lim_{n\to\infty}d\big(n^{-1}\Sigma_{k=1}^{n}X_k, n^{-1}\Sigma_{k=1}^{n}EX_k\big) = 0 \qquad \textit{in probability.}$$

COROLLARY 4. *Let (X_n) be a sequence of convex tight* r.e.'s *in D such that $E\|X_n\|_\infty^r \leqslant \Gamma$ for all n, with $r > 1$. If*

(i) $\mathrm{Cov}\{X_l(t), X_k(t)\} = 0$, *for each* $k \neq l$ *and each* $t \in [0, 1]$
and

(ii) $\Sigma_{k=1}^{n}\mathrm{Var}\{X_k(t)\} = o(n^2)$, *for each* $t \in [0, 1]$,
then

$$\lim d\left(n^{-1}\Sigma_{k=1}^{n}X_k, n^{-1}\Sigma_{k=1}^{n}EX_k\right) = 0 \quad \textit{in probability.}$$

THEOREM 5. *Let* (X_n) *be a sequence of identically distributed, convex tight* r.e.'s *in* D *such that* $E\|X_1\|_\infty < \infty$. *Then*

$$\lim_{n\to\infty} n^{-1}\Sigma_{k=1}^{n}X_k(t) = EX_1 \quad \textit{in probability,}$$

for each dyadic rational $t \in [0, 1]$, *if and only if*

$$\lim_{n\to\infty} d\left(n^{-1}\Sigma_{k=1}^{n}X_k, EX_1\right) = 0 \quad \textit{in probability.}$$

If $n^{-1}\Sigma_{k=1}^{n}EX_k$ converges to a constant, then Theorem 3 and Corollary 4 are if and only if since convergence in probability implies weak convergence (in distribution), and hence pointwise convergence in distribution to a constant, which in turn implies pointwise convergence in probability.

4. A discussion of convex tightness. Let $K = \{x_s : 0 < a \leqslant s \leqslant 1 - a < 1\}$, $0 < a < \frac{1}{2}$, where $x_s = I_{[s, 1]}$. That K is compact follows easily by applying the criteria of Theorem 14.3 of [1]. That the closed convex hull $\overline{co(K)}$ of K is not compact is seen as follows. Let $\delta > 0$ be given and choose $s, t \in [a, 1 - a]$ such that $s < t$ and $t - s < \delta$. Then $x = \frac{1}{2}x_t + \frac{1}{2}x_s = \frac{1}{2}I_{[s, t)} + I_{[t, 1]} \in \overline{co(K)}$, but for any partition $\{t_0, t_1, \cdots, t_m\}$ of $[0, 1]$ satisfying $\min\{t_i - t_{i-1}\} > \delta$, $\max_i w_x([t_{i-1}, t_i)) \geqslant \frac{1}{2}$. Thus, $\lim_{\delta\downarrow 0}\sup_{x\in\overline{co(K)}} w'_x(\delta) \geqslant \frac{1}{2}$, and $\overline{co(K)}$ is not compact by Theorem 14.3 of [1]. This property is related to the fact that the Skorokhod topology on D is not locally convex. Some points of D possess a convex local base but most do not.

This example can be strengthened by making the set K countably infinite. Let $K = \{x_{s_i} : i \in N, 0 < a \leqslant s_i \leqslant 1 - a < 1\}$, $0 < a < \frac{1}{2}$. Then, if the sequence (s_i) contains infinitely many different points of $[a, 1 - a]$, it is shown as above that the convex hull of K is not relatively compact.

For $x_s = I_{[s, 1]}$ and $0 \leqslant s < 1$, let the sequence (X_n) be i.i.d. where $P[X_1 \in \{x_s : s_1 \leqslant s \leqslant s_2\}] = s_2 - s_1$, $0 \leqslant s_1 < s_2 < 1$. Then each X_n is a r.e. in D, and (X_n) does not have separable support with respect to the uniform topology since the linear span of the set $\{x_s : 0 \leqslant s < 1\}$ is dense in D. Now, if a convex compact set K_1 satisfying $P[X_1 \notin K_1] \leqslant \varepsilon$ existed, then K_1 necessarily would contain a set of the form $K_J = \{x_s : s \in J\}$, where $J \subset [a, 1 - a]$ and $m(J) \geqslant 1 - 2a - \varepsilon$ (m is Lebesgue measure). Being convex and closed, K_1 would contain the closed convex hull of K_J. But from the above example it is clear that the closed convex hull of such a set K_J is not compact. Hence no such set K_1 can exist. While not being convex tight, the range of X_1 is in $D\uparrow$. Hence, the SLLN of Theorem 2 applies for independent copies of X_1.

Roughly speaking, convex compact sets in D are those sets of functions which become arbitrarily small in absolute value at cluster points of jumps in [0, 1]. The following theorem characterizes the convex compact subsets K of D in terms of the jumps of functions in K.

For $A \subset D$ and $\varepsilon > 0$, let

$$S_\varepsilon(A) = \{t \in [0, 1] : \sup_{x \in A} |x(t) - x(t-0)| > \varepsilon\}.$$

THEOREM 6. *If K is a relatively compact subset of D, then $co(K)$ is relatively compact if and only if $S_\varepsilon(K)$ is finite for every $\varepsilon > 0$.*

PROOF ("only if" part). Suppose that for some $\varepsilon > 0$, $S_\varepsilon(K)$ is infinite. Then there is $t_0 \in [0, 1]$ and a sequence $\{t_n\}$ of distinct points in $S_\varepsilon(K)$ converging to t_0 and a sequence $\{x_n\}$ of elements of K such that $|x_n(t_n) - x_n(t_n - 0)| > \varepsilon$, for all $n \in N$. Since K is compact, there is $\delta > 0$ such that

$$\sup_{x \in K} w'_x(\delta) < \varepsilon/2. \tag{4.1}$$

Find n and n', $n \neq n'$, such that $t_n, t_{n'} \in (t_0 - \delta/2, t_0 + \delta/2)$ and such that $|x_n(t_n) - x_n(t_n - 0)| > \varepsilon$ and $|x_{n'}(t_{n'}) - x_{n'}(t_{n'} - 0)| > \varepsilon$.

Since x_n, $x_{n'}$ have jumps at t_n, $t_{n'}$, respectively, of magnitudes $> \varepsilon$, we have from (4.1), for $i = n, n'$:

$$\sup_{t_i \leqslant t,\, s < t_i + \delta} |x_i(t) - x_i(s)| < \varepsilon/2 \tag{4.2}$$

$$\sup_{t_i - \delta \leqslant t,\, s < t_i} |x_i(t) - x_i(s)| < \varepsilon/2.$$

Now let $T = \{t_0, t_1, \cdots, t_m\}$ be any partition of [0, 1] with $\min_{1 \leqslant i \leqslant m}\{t_i - t_{i-1}\} \geqslant \delta$. Since $|t_n - t_{n'}| < \delta$, T can contain t_n or $t_{n'}$ or neither one, but not both. If, for example, $t_n \in T$, then $t_{n'} \notin T$ and

$$\begin{aligned}|x_n(t_{n'}) + x_{n'}(t_{n'}) - (x_n(t_{n'} - 0) + x_{n'}(t_{n'} - 0))| \\ \geqslant |x_{n'}(t_{n'}) - x_{n'}(t_{n'} - 0)| - |x_n(t_{n'}) - x_n(t_{n'} - 0)| \\ > \varepsilon - \varepsilon/2 = \varepsilon/2, \quad \text{using the relations (4.2).}\end{aligned}$$

Define $x_\delta = \frac{1}{2}x_n + \frac{1}{2}x_{n'}$. We then have $|x_\delta(t_{n'}) - x_\delta(t_{n'} - 0)| > \varepsilon/4$. If $t_{n'} \in T$ then $t_n \notin T$ and the same reasoning as above yields $|x_\delta(t_n) - x_\delta(t_n - 0)| > \varepsilon/4$. If $t_n \notin T$ and $t_{n'} \notin T$, then a fortiori $|x_\delta(t_n) - x_\delta(t_n - 0)| > \varepsilon/4$. Thus, for any partition T with $\min_{1 \leqslant i \leqslant m}\{t_i - t_{i-1}\} \geqslant \delta$, we have $\sup_{s,\, t \in [t_{i-1}, t_i)} |x_\delta(s) - x_\delta(t)| > \varepsilon/4$, for some $i = 1, \cdots, m$, and hence $w'_{x_\delta}(\delta) \geqslant \varepsilon/4$. But $x_\delta \in co(K)$ and so $\sup_{x \in co(K)} w'_x(\delta) \geqslant w'_{x_\delta}(\delta) \geqslant \varepsilon/4$.

Since $\varepsilon > 0$ is fixed and $\delta > 0$ is arbitrary, this yields $\liminf_{\delta \to 0} \sup_{x \in co(K)} w'_x(\delta) > 0$, and thus $co(K)$ is not relatively compact in D. □

For the "if" part we first prove the following lemma.

LEMMA 7. *If K is a relatively compact subset of D such that $S_\varepsilon(K)$ is finite for each $\varepsilon > 0$, then to each $t \in [0, 1]$ the following holds: for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\sup_{x \in K} w_x([t, t + \delta)) < \varepsilon$ and $\sup_{x \in K} w_x((t - \delta, t)) < \varepsilon$.*

PROOF. Let $\varepsilon > 0$ be given and fix $t_0 \in [0, 1]$. By relative compactness of K, find $\delta_0 > 0$ such that $\sup_{x\in K} w'_x(\delta_0) < \varepsilon/3$. Now find δ_1, $0 < \delta_1 \leqslant \delta_0$, such that $t \in (t_0, t_0 + \delta_1)$ implies

$$\sup_{x\in K}|x(t) - x(t-0)| < \varepsilon/3. \tag{4.3}$$

Take $x \in K$ and let $T = \{t_i\}$ be any finite partition of [0, 1] such that $\max_i \sup_{x\in K} w_x([t_{i-1}, t_i)) < \varepsilon/3$ and $\min_i\{t_i - t_{i-1}\} \geqslant \delta_0$. If no point of T falls in $[t_0, t_0 + \delta_1)$, then $w_x([t_0, t_0 + \delta_1)) < \varepsilon/3$. If a point $t_i \in T$ is such that $t_i \in (t_0, t_0 + \delta_1)$, then $\max_i \sup_{x\in K} w_x([t_{i-1}, t_i)) < \varepsilon/3$ yields $w_x([t_0, t_i)) > \varepsilon/3$ and $w_x([t_i, t_0 + \delta_1)) < \varepsilon/3$. By inequality (4.3), if x makes a jump at t_i, its magnitude is necessarily $< \varepsilon/3$. Thus, by the triangle inequality, $w_x([t_0, t_0 + \delta_1)) < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$, for any $x \in K$. Hence $\sup_{x\in K} w_x([t_0, t_0 + \delta_1)) \leqslant \varepsilon$. In a similar manner, a $\delta_2 > 0$ can be found such that $\sup_{x\in K} w_x((t_0 - \delta_2, t_0)) \leqslant \varepsilon$. Now take $\delta = \min\{\delta_1, \delta_2\}$ and the proof is complete.

PROOF ("if" part). Now, let $\varepsilon > 0$ be given and $S_\varepsilon(K)$ be finite. Let $S_\varepsilon(K) = \{t_1, \cdots, t_N\}$.

Let $I_K = [t_k, t_{k+1}]$ for $k = 0, 1, \cdots, N$, where $t_0 = 0$ and $t_{N+1} = 1$. Put $t_{k1} = t_k$ and corresponding to t_k find, by Lemma 7, a $\delta(t_{k1}) > 0$ such that

$$\sup_{x\in K} w_x\big([t_k, t_k + \delta(t_{k1}))\big) < \varepsilon.$$

For $j \geqslant 1$, inductively define $t_{k,j+1} = t_{kj} + \delta(t_{kj})$ and set $I_{k1} = [t_{k1}, t_{k1} + \delta(t_{k1}))$ and $I_{kj} = (t_{kj}, t_{kj} + \delta(t_{kj}))$ if $j \neq 1$. Let t_{kj} and $\delta(t_{kj})$ be determined alternately in the following manner: given $t_{k1}, \cdots, t_{kj}$, find by Lemma 7, $\delta(t_{kj}) > 0$, such that

$$\sup_{x\in K} w_x\big((t_{k,j+1} - \delta(t_{k,j+1}), t_{k,j+1})\big) < \varepsilon$$

and

$$\sup_{x\in K} w_x\big([t_{k,j+1}, t_{k,j+1} + \delta(t_{k,j+1}))\big) < \varepsilon.$$

Each $t_{kj} \leqslant t_{k+1}$ since some $x \in K$ makes a jump at t_{k+1}.

In this way we get a sequence $\{t_{k1}, t_{k2}, \cdots\}$ of points in $[t_k, t_{k+1}]$ and a sequence of intervals $I_{k1}, I_{k2}, \cdots$ which are all open sets in $[t_k, t_{k+1}]$. Another application of Lemma 7 yields a $\delta(t_{k+1}) > 0$ such that

$$\sup_{x\in K} w_x\big((t_{k+1} - \delta(t_{k+1}), t_{k+1})\big) < \varepsilon.$$

Let $I'_k = (t_{k+1} - \delta(t_{k+1}), t_{k+1}]$.

The collection of relatively open subintervals $\{I'_k, I_{k1}, I_{k2}, \cdots\}$ is an open cover of $[t_k, t_{k+1}]$ which is compact. Consequently there exists an open subcover $\{J_{k1}, \cdots, J_{kN_k}\}$. Let $\{s_{k1}, \cdots, s_{kN_k}\}$ denote the respective centers of these intervals.

Now this can be done for every k, $k = 0, 1, \cdots, N$. The collection of points $\bigcup_{k=0}^{N} \bigcup_{j=1}^{N_k}\{s_{kj}\} \cup \bigcup_{j=1}^{N}\{t_j\}$ forms a partition of [0, 1]; call it T and denote the points of it in ascending order, by

$$0 = s_0 < s_1 < \cdots < s_{m-1} < s_m = 1.$$

The claim is now that

$$\max_{i=1,\dots,m-1} \sup_{x \in co(K)} w_x([s_{i-1}, s_i)) < \varepsilon \tag{4.4}$$

and

$$\sup_{x \in co(K)} w_x([s_{m-1}, 1]) < \varepsilon.$$

Let $x = \sum_{j=1}^n \alpha_j x_j$, $x_j \in K$, $\alpha_j \geqslant 0$ and $\sum_1^n \alpha_j = 1$.
Then

$$\begin{aligned} w_x([s_{i-1}, s_i)) &= \sup_{s,t \in [s_{i-1}, s_i)} \left|\sum_{j=1}^n \alpha_j (x_j(s) - x_j(t))\right| \\ &\leqslant \sum_{j=1}^n \alpha_j \sup_{s,t \in [s_{i-1}, s_i)} |x_j(s) - x_j(t)| \\ &= \sum_{j=1}^n \alpha_j w_{x_j}([s_{i-1}, s_i)) \\ &\leqslant \sum_{j=1}^n \alpha_j \sup_{x \in K} w_x([s_{i-1}, s_i)) \\ &\leqslant \sum_{j=1}^n \alpha_j \cdot \varepsilon \\ &= \varepsilon. \end{aligned}$$

Hence, (4.4) holds and thus

$$\sup_{x \in co(K)} w'_x(\delta) < \varepsilon$$

by taking $\delta < \min_{1 \leqslant i \leqslant m} \{s_i - s_{i-1}\}$.

Thus to each $\varepsilon > 0$ there is $\delta > 0$ such that $\sup_{x \in co(K)} w'_x(\delta) < \varepsilon$ and so

$$\lim_{\delta \to 0} \sup_{x \in \overline{co(K)}} w'_x(\delta) = 0$$

and $co(K)$ is relatively compact. □

Acknowledgments. The authors wish to thank both Ms. Marjorie G. Hahn for her helpful comments on the properties of $D[0, 1]$ and the referee for valuable suggestions.

REFERENCES

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

[2] Daffer, P. and Taylor, R. L. (1977). Laws of large numbers for $D[0, 1]$. Univ. of South Carolina Mathematics Technical Report no. 60B05-10.

[3] Hahn, M. G. (1975). Central limit theorems for $D[0, 1]$-valued random variables. Thesis, M.I.T.

[4] Hahn, M. G. (1977). Conditions for sample-continuity and the central limit theorem. *Ann. Probability* **5** 351–360.

[5] Hoffmann-Jørgensen, J. (1970). The theory of analytic spaces. Aarhus Universitet. Various publication series no. 10.

[6] Kallenberg, O. (1974). Series of random processes without discontinuities of the second kind. *Ann. Probability* **2** 729–737.

[7] Kuelbs, J. (1976). A strong convergence theorem for Banach space valued random variables. *Ann. Probability* **4** 744–771.

[8] RAO, R. RANGA (1963). The law of large numbers for $D[0, 1]$-valued random variables. *Theor. Probability Appl.* **8** 70–74.

[9] RUDIN, W. (1973). *Functional Analysis*. McGraw-Hill, New York.

[10] TAYLOR, R. L. and DAFFER, P. (1978). Some weak laws of large numbers for probability measures on vector spaces. *Proceedings of the Conference on Vector Space Measures and Applications*. Springer-Verlag Lecture Notes **644** 411–417.

DEPARTMENT OF MATHEMATICS
COMPUTER SCIENCE AND STATISTICS
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, SOUTH CAROLINA 29208