The Annals of Probability
1985, Vol. 13, No. 3, 860–884

THE SPACE $D(A)$ AND WEAK CONVERGENCE FOR SET-INDEXED PROCESSES[1]

BY RICHARD F. BASS AND RONALD PYKE

University of Washington

In this paper we consider weak convergence of processes indexed by a collection $\mathscr{A}$ of subsets of I^d. As a suitable sample space for such processes, we introduce the space $\mathscr{D}(\mathscr{A})$ of set functions that are outer continuous with inner limits. A metric is defined for $\mathscr{D}(\mathscr{A})$ in terms of the graphs of its elements and then we give a sufficient condition for a subset of $\mathscr{D}(\mathscr{A})$ to be compact in this topology. This framework is then used to provide a criterion for probability measures on $\mathscr{D}(\mathscr{A})$ to be tight. As an application, we prove a central limit theorem for partial-sum processes indexed by a family of sets, $\mathscr{A}$, when the underlying random variables are in the domain of normal attraction of a stable law. If $\alpha \in (1, 2)$ denotes the exponent of the limiting stable law, if r denotes the coefficient of metric entropy of $\mathscr{A}$, and if $\mathscr{A}$ satisfies mild regularity conditions, we show that the partial-sum processes converge in law to a stable Lévy process provided $r < (\alpha - 1)^{-1}$.

1. Introduction. The main purpose of this paper is to provide a useful topology for a space of set functions that are "outer continuous with inner limits." The space is denoted by $\mathscr{D}(\mathscr{A})$ where $\mathscr{A}$, the domain of the functions, is a family of Borel subsets in the d-dimensional unit cube $I^d = [0, 1]^d$ (see Definition 3.4 below). This space of set functions was introduced in Bass and Pyke (1984b) as a range space for set-indexed Lévy processes. It should be viewed as a natural generalization of the space $D[0, 1]$ of real functions on $[0, 1]$ having left limits and right continuity, functions which were originally referred to as having "discontinuities of the first kind."

Whereas $D[0, 1]$ and its extensions are suitable range spaces for many discontinuous processes indexed by points, the space $\mathscr{D}(\mathscr{A})$ studied in this paper is a natural range space of sample paths for processes indexed by a family of sets. Examples of such processes include set-indexed empirical processes, partial-sum processes, Brownian processes, Lévy (infinitely divisible) processes and general point processes. Topologies on range spaces of processes are necessary for the measurability and distribution theory of functionals of processes and especially for the study of weak convergence of image laws. In the examples to date of weak convergence results for set-indexed processes, the limiting processes have had continuous sample paths, enabling one to make use of the uniform topology. In such situations the nonseparability and resulting nonmeasurability problems can be circumvented in various ways. However, a much smaller topology is essential when the limiting process itself does not have continuous paths, as is the case

---

Received December 1983; revised February 1985.

[1] Prepared under National Science Foundation Grants MCS-83-00581 and MCS-82-02861.

AMS 1980 *subject classifications.* Primary 60B10; secondary 60F05, 60F17, 60B05.

Key words and phrases. D-space, weak convergence, central limit theorem, domains of attraction, stable laws, tightness, partial-sum processes, empirical processes.

for the central limit problem studied below in Section 5 in which the limit processes are stable Lévy processes with index $\alpha < 2$.

There are many ways in which set functions can have discontinuities. In this paper we focus upon discontinuities that are due to the existence of point masses (atoms) at fixed points. The motivating principle for the construction of a suitable topology is the same as for $D[0, 1]$; namely, two functions should be close if the large atoms (the "jumps" in the real case) of one function and their locations are approximately equal to those of the other function, while the two functions minus their large atoms are uniformly close. The challenge is to make this precise in such a way as to enable one to be able to provide a usable characterization of the compact sets. This is the purpose of Section 3.

The first topologies for $D[0, 1]$ were provided by Prohorov (1953, 1956), Skorokhod (1955, 1956a, 1956b, 1957) and Kolmogorov (1956). See Billingsley (1968) for a presentation of the main topology on $D[0, 1]$, commonly referred to now as the Skorokhod topology. This is a topology determined by a metric which makes $D[0, 1]$ into a complete and separable metric space.

Many topologies are possible for $D[0, 1]$. Skorokhod (1956) introduced four topologies, known as the J_1, J_2, M_1 and M_2 topologies. Convergence in each of these topologies consists of convergence in the weakest M_2 topology plus perhaps additional conditions. Each thereby postulates the convergence of graphs with respect to the Hausdorff metric, and this is true also of our topology for $\mathscr{D}(\mathscr{A})$.

The d-dimensional generalization of $D[0, 1]$, $D(I^d)$ say, and the provision of suitable topologies for it has been given by Bickel and Wichura (1971), Neuhaus (1971), and Straf (1972), with the latter reference focusing on more general index sets than I^d. Other extensions have been made to the cases $D[0, \infty)$ and $D[0, \infty)^d$ of noncompact index sets; see Lindvall (1973). Also of interest, is a recent approach by Vervaat (1981), where a different sample space is introduced in which functions are equated with pairs of upper and lower semicontinuous fits, and then a corresponding topology is introduced.

The topology for $\mathscr{D}(\mathscr{A})$ that we introduce is similar in spirit to the M_2 topology for functions on $\mathbb{R}$. (J_2 and M_1 are clearly not suitable.) The idea is to define the distance between two functions by the Hausdorff distance between their graphs. The set of right continuous, left limit functions are not closed under this topology, but this is a lesser difficulty. Most of the work comes in developing a criterion for when a set is compact (see Theorem 3.4).

The main application and motivation of the topology of Section 3 is the central limit theorem of Section 5. For this, the necessary characterization of weak convergence is provided in Section 4. The central limit theorem then states that the partial-sum processes obtained by suitably smoothing and normalizing the partial-sums formed from an array of independent random variables (r.v.) with common distribution in the domain of attraction of nonnormal stable distributions converge to a stable process indexed by sets.

Central limit theorems for partial-sum processes (random walks) on $[0, 1]$ date back to Donsker (1951). For $\mathscr{A}$ a class of sets much larger than the orthants, a uniform central limit theorem for smoothed partial-sum processes was given by Pyke (1983). A different method of proof, as well as a law of the iterated logarithm, was given in Bass and Pyke (1984a). These results have now been shown to hold under a finite variance condition only; cf. Alexander and Pyke (1985) and Bass (1985). All of these results, however, are concerned with convergence to a normal limiting process. For a central limit theorem for partial-sum processes in $D[0, 1]$ converging to a nonnormal Lévy process on $[0, 1]$, see the book of Gikhman and Skorokhod (1969). A martingale approach and further references may be found in Jacod, Kłopotowski, and Memin (1982). For more general index families $\mathscr{A}$, the structure and existence of suitable limit laws has been studied by Adler and Feigin (1984) and Bass and Pyke (1984b).

Before introducing a topology for $\mathscr{D}(\mathscr{A})$ and deriving these limit theorems, we first introduce in the next section the assumptions we impose on the index families $\mathscr{A}$. Finally, in Section 6, some remarks and open problems are given.

2. The assumptions on A. Before beginning our study of the topology on $\mathscr{D}(\mathscr{A})$, and from there to the study of weak convergence and central limit theorems, some conditions on our family $\mathscr{A}$ must be imposed. For convenience, we collect the assumptions here in one place.

(A1) (i) $\mathscr{A}$ is a collection of closed subsets of I^d;
(ii) $\mathscr{A}$ itself is closed with respect to the Hausdorff metric d_H;
(iii) for each δ, there is a finite subset $\mathscr{A}_\delta$ of $\mathscr{A}$ such that if $A \in \mathscr{A}$, there exists $A_\delta^+ \in \mathscr{A}_\delta$ such that $A \subseteq (A_\delta^+)^0 \subseteq A^\delta$.

Here $(A_\delta^+)^0$ is the interior of A_δ^+ with respect to the relative Euclidean topology on I^d, and A^δ is the set of all points of I^d that are less than δ from some point of A. Recall $d_H(A, B) = \inf\{\varepsilon: A \subseteq B^\varepsilon \text{ and } B \subseteq A^\varepsilon\}$. We use $|\cdot|$ to denote Lebesgue measure.

(A2) (i) for each δ, there is a finite subset $\mathscr{A}_\delta$ of $\mathscr{A}$ satisfying (A1) (iii) such that if $A \in \mathscr{A}$, there exists $A_\delta, A_\delta^+ \in \mathscr{A}_\delta$ with $A_\delta \subseteq A \subseteq A_\delta^+$ and $|A_\delta^+ \setminus A_\delta| \leq \delta$;
(ii) there are constants K and $r > 0$ such that if $\#\mathscr{A}_\delta$ is the cardinality of $\mathscr{A}_\delta$, and $H(\delta) = \ln(\#\mathscr{A}_\delta)$, then $H(\delta) \leq K\delta^{-r}$ for δ sufficiently small.

Examples of index families covered by these assumptions when $d > 1$ include $\mathscr{C}^d$, the set of closed convex sets in I^d, for which $r = (d-1)/2$ (cf. Dudley, 1974) and $\mathscr{I}(d, q, M)$, the family of closed sets with "smooth" boundaries determined by q-differentiable functions whose Lipschitz norm of order q is bounded by M. For the latter, $r = (d-1)/q$ (cf. Dudley, 1974). A related family with the same r has been introduced by Révész (1976). In connection with the central limit theorem of Section 5, where α is the exponent of the limiting stable distribution, we see from the above that $\mathscr{C}^d$ is a possible index family provided $(d-1)/2 < (\alpha-1)^{-1}$, or equivalently $\alpha < (d+1)/(d-1)$. For example, if $\alpha = 5/4$, the convex sets form a possible index family if the dimension does not exceed 8.

The classical case of processes indexed by the points in I^d can of course be viewed as set-indexed processes by the identification of $\mathbf{t}$ with $[\mathbf{0}, \mathbf{t}]$. If $\mathscr{I}^d := \{[\mathbf{0}, \mathbf{t}]: \mathbf{t} \in I^d\}$ denotes the resulting index family of intervals, then note that any coefficient $r > 0$ satisfies (A2)(ii) for $\mathscr{A} = \mathscr{I}^d$.

3. The $\mathscr{D}_0(\mathscr{A})$ space. Throughout this section we assume that $\mathscr{A}$ satisfies (A1). We begin by defining the subfamily $\mathscr{D}_0 = \mathscr{D}_0(\mathscr{A})$ of $\mathscr{D}(\mathscr{A})$ which will contain the sample paths of our processes. We denote elements of $\mathscr{D}_0$ by x, y, z.

If $x: \mathscr{A} \to R$, set

$$\| x \|_{\mathscr{A}} = \sup_{A \in \mathscr{A}} | x(A) |. \tag{3.1}$$

DEFINITION 3.1. A function $x: \mathscr{A} \to \mathbb{R}$ is purely atomic if there exist finitely many locations $\mathbf{t}_1, \cdots, \mathbf{t}_n \in I^d$ and masses $a_1, \cdots, a_n$ such that for all $B \in \mathscr{A}$,

$$x(B) = \sum_{\mathbf{t}_j \in B} a_j. \tag{3.2}$$

For such an x, let

$$\text{Variation}(x) = \sum_j | a_j |. \tag{3.3}$$

Let $\mathscr{C}(\mathscr{A})$ be the class of real-valued functions that are uniformly continuous with respect to d_H.

DEFINITION 3.2. Let $\mathscr{D}_0 = \{x: \mathscr{A} \to \mathbb{R}$, such that there exist uniformly continuous functions $C_m \in \mathscr{C}(\mathscr{A})$ and purely atomic functions J_m such that $\| x - C_m - J_m \|_{\mathscr{A}} \to 0$ as $m \to \infty\}$.

Our set $\mathscr{D}_0$ is closely related to the set $\mathscr{D}$ in Dudley (1978).

DEFINITION 3.3. $x: \mathscr{A} \to \mathbb{R}$ is outer continuous at $A \in \mathscr{A}$ if $A \subseteq A_n$, $A_n \to_{d_H} A$, and $A_n \in \mathscr{A}$ implies $x(A_n) \to x(A)$. x has inner limits if $A_n \subseteq A^0$, $A_n \to_{d_H} A^0$, and $A_n, A \in \mathscr{A}$ implies $\lim_n x(A_n)$ exists, where A^0 is the interior of A.

As pointed out in Bass and Pyke (1984b), one does not want to require $A_n \to A$ monotonically in the definition, even in the case where $\mathscr{A} = \mathscr{I}^d = \{[\mathbf{0}, \mathbf{t}]: \mathbf{t} \in I^d\}$, the family of lower orthants.

DEFINITION 3.4. Let $\mathscr{D}(\mathscr{A}) = \{x: \mathscr{A} \to \mathbb{R}$, such that x is outer continuous with inner limits at each $A \in \mathscr{A}\}$.

It is clear that both purely atomic functions and continuous functions are outer continuous with inner limits. Since outer continuity is preserved under sums and uniform limits, every element of $\mathscr{D}_0(\mathscr{A})$ is outer continuous with inner limits. Thus, $\mathscr{D}_0(\mathscr{A}) \subseteq \mathscr{D}(\mathscr{A})$.

Next we proceed to define a metric on $\mathscr{D}(\mathscr{A})$. Let the distance between two elements of $\mathscr{A} \times \mathbb{R}$ be defined by

$$\rho((A_1, r_1), (A_2, r_2)) = d_H(A_1, A_2) + | r_1 - r_2 |. \tag{3.4}$$

Let $\mathscr{G}$ be the collection of closed subsets of $\mathscr{A} \times \mathbb{R}$, and let d_G be the Hausdorff metric on $\mathscr{G}$ induced by ρ.

For $x \in \mathscr{D}(\mathscr{A})$ define the graph of x, denoted by $G(x)$, to be the closure in $\mathscr{A} \times \mathbb{R}$ of $\{(A, x(A)): A \in \mathscr{A}\}$. Note that we therefore do not require $G(x)$ to contain the "vertical" pieces (A, r), $r_1 < r < r_2$, whenever (A, r_1) and (A, r_2) are both in $G(x)$, as is done in Skorokhod's M_2 topology for $D[0, 1]$.

Now define

$$d_D(x, y) = d_G(G(x), G(y)). \tag{3.5}$$

Since d_G is a metric, d_D is a pseudo-metric on $\mathscr{D}$. In the usual way, let us identify functions x and y if $d_D(x, y) = 0$. (Note, e.g., the functions $x_a(t) = \sin(1/t)$, $t \in (0, 1]$, $x_a(0) = a$ for $a \in [-1, 1]$ are all identified by d_D.)

It is known that if $\mathscr{K}$ is the class of compact subsets of a metric space $\mathscr{S}$, then $\mathscr{K}$ under the Hausdorff metric is complete (separable) (compact) if $\mathscr{S}$ is complete (separable) (compact). See, for example, Debreu (1967). In view of assumption (A1) and the compactness of I^d, this implies that $\mathscr{A}$ is compact, separable and complete. In particular, $\mathscr{A}$ is totally bounded under the Hausdorff metric d_H.

It is easy to characterize the compact subsets of $\mathscr{G}$.

PROPOSITION 3.1. *Suppose $\mathscr{F} \subseteq \mathscr{G}$ is closed and contains at least one set G_0 that is bounded in $\mathscr{A} \times \mathbb{R}$. Then $\mathscr{F}$ is compact if and only if $R = \sup\{|r|: (a, r) \in G$ for some $A \in \mathscr{A}$ and some $G \in \mathscr{F}\} < \infty$.*

PROOF. First suppose $\mathscr{F}$ is compact. Let $K_1 = \sup\{|s|: (A, s) \in G_0\}$, $K_2 = \sup\{d_G(G_0, H): H \in \mathscr{F}\}$, and $K = K_1 + K_2$. K is finite since G_0 is bounded and $\mathscr{F}$ is compact, hence bounded. If $(A, r) \in H \in \mathscr{F}$, then $|r| \le K_1 + K_2$.

Now suppose $R < \infty$. If we define $\mathscr{G}_R$ by

$$\mathscr{G}_R = \{H \in \mathscr{G}: H \subseteq \mathscr{A} \times [-R, R]\}, \tag{3.6}$$

then $\mathscr{G}_R$ is totally bounded since $\mathscr{A} \times [-R, R]$ is a bounded set. Since $\mathscr{G}$ is complete, $\mathscr{G}_R$ is compact. If $H \in \mathscr{F}$, then $H \subseteq \mathscr{A} \times [-R, R]$. Thus $\mathscr{F}$ is a closed subset of $\mathscr{G}_R$, hence compact. □

Unfortunately, characterizing the compact subsets of $\mathscr{G}$ is not that useful because neither $\mathscr{D}_0(\mathscr{A})$ nor $\mathscr{D}(\mathscr{A})$ is closed in $\mathscr{G}$. For example, let $\mathscr{A} = \mathscr{I}^1 = \{[0, t], 0 \le t \le 1\}$, and let $x_n([0, t]) = \sin(nt)$. Each $x_n \in \mathscr{C}(\mathscr{A})$, and $G(x_n)$ converges as $n \to \infty$ to $G_0 = [0, 1] \times [-1, 1]$, but x_n does not converge to an element of $\mathscr{D}(\mathscr{A})$. What is needed is to characterize those subsets of $\mathscr{D}_0(\mathscr{A})$ which are compact with respect to d_D. We do not give a complete characterization of such sets, but we do prove compactness for a class of sets sufficiently large for the purposes of our central limit theorem.

If x is purely atomic with atoms at locations $\mathbf{t}_1, \cdots, \mathbf{t}_n$, define

$$\operatorname{gap}(x) = \inf_{i \ne j} |\mathbf{t}_i - \mathbf{t}_j|. \tag{3.7}$$

DEFINITION 3.5. If h is an increasing real function with $h(\delta) \le \delta$, N is a finite positive integer-valued function, and η and R positive real numbers, define $\mathscr{F}_{PA}(h, N, \eta, R)$ to be the set of all purely atomic x such that

(3.8) (i) $\operatorname{gap}(x) \ge \eta$,

(ii) Variation $(x) \le R$, and

(iii) for each δ, there exist sets $A_1, A_2, \cdots, A_{N(\delta)} \in \mathscr{A}$, possibly depending on x, such that
  (a) each point of $G(x)$ is within a distance $\leq \delta$ from some $(A_i, x(A_i))$, $i = 1, \cdots N(\delta)$,
  (b) $A_i^{h(\delta)} \backslash A_i$ contains no atoms of x.

The crucial proposition is

PROPOSITION 3.2. *$\mathscr{F}_{PA}(h, N, \eta, R)$ is a compact subset of $\mathscr{D}_0(\mathscr{A})$.*

PROOF. For the duration of this proof, abbreviate $\mathscr{F}_{PA}(h, N, \eta, R)$ by $\mathscr{F}_0$. If $x \in \mathscr{F}_0$, then $\| x \|_{\mathscr{A}} \leq$ Variation $(x) \leq R$. Since the image of $\mathscr{F}_0$ under G is thus a subset of $\mathscr{G}_R$ (defined by (3.6)), which is compact in $\mathscr{G}$, it suffices to show that the image of $\mathscr{F}_0$ under G is closed. To do this, it suffices to show that if $x_n \in \mathscr{F}_0$ and $d_G(G(x_n), G_0) \to 0$ as $n \to \infty$ for some G_0, then there exists an $x \in \mathscr{F}_0$ with $G(x) = G_0$.

Since $\inf_n \operatorname{gap}(x_n) \geq \eta$ by (3.8)(i), the number of atoms in x_n is uniformly bounded. Replacing x_n by a subsequence if necessary, we may assume that there is a positive integer M and each x_n has precisely M atoms. Let us denote the locations by $\mathbf{t}_{in}$, $i = 1, \cdots, M$, and the masses by a_{in}, $i = 1, \cdots, M$. Replacing x_n by a further subsequence if necessary, we may assume, since I^d is compact, that the $\mathbf{t}_{in}$ converge as $n \to \infty$, say to $\mathbf{t}_i$, $i = 1, \cdots, M$, and that the a_{in} converge as $n \to \infty$, say to a_i, $i = 1, \cdots, M$. Let x be the purely atomic function which has an atom of size a_i at location $\mathbf{t}_i$, $i = 1, \cdots, M$.

First we show $G(x) \subseteq G_0$. For $A \in \mathscr{A}$, choose ε small enough and then choose $B \in \mathscr{A}$ such that $A \subseteq B^0 \subseteq A^\varepsilon$, $x(B) = x(A)$, and no atom of x lies on the boundary of B. This is possible since x has only finitely many atoms. Since the locations and sizes of the atoms of x_n converge to those of x, and every atom of x lies either in B^0 or $(B^c)^0$, $x_n(B) \to x(B)$. Thus $(B, x(B))$ is in $\lim G(x_n) = G_0$. The fact that $x(B) = x(A)$ and ε is arbitrary shows that $(A, x(A)) \in G_0$, or $G(x) \subseteq G_0$.

Now we show $G_0 \subseteq G(x)$. Suppose $(B, r) \in G_0$. Because $G(x_n) \to_{d_G} G_0$, there exists sets B_n such that $B_n \to_{d_H} B$ and $x_n(B_n) \to r$. Let $\delta > 0$, and use (3.8)(iii) to find A_n such that $d_H(A_n, B_n) < 2\delta$, $| x_n(A_n) - x_n(B_n) | < 2\delta$, and $A_n^{h(\delta)} \backslash A_n$ contains no atom of x_n. Replacing x_n by a subsequence if necessary, we may assume $A_n \to_{d_G}$, say to A. Finally, choose ε and C such that $\varepsilon < h(\delta)/2$, $A \subseteq C^0 \subseteq A^\varepsilon$, $x(C) = x(A)$, and no atom of x lies on the boundary of C.

Now $d_H(C, B) \leq \varepsilon + 2\delta \leq 5\delta/2$. Since no atom of x lies on the boundary of C, $x_n(C) \to x(C)$. And since $A_n \to_{d_H} A$, for n large enough, $A_n \subseteq C^0 \subseteq A_n^{h(\delta)}$. But then $x_n(C) = x_n(A_n)$ by (3.8)(iii)(b), and so $| x(C) - r | \leq 2\delta$. This shows that $\rho((B, r), (C, x(C))) \leq 9\delta/2$; since δ is arbitrary, $(B, r) \in G(x)$, and hence $G_0 \subseteq G(x)$.

It remains to show $x \in \mathscr{F}_0$. (3.8)(i) and (ii) are obvious. Fix δ, and for each n, choose A_{in}, $i = 1, \cdots, N(\delta)$, to satisfy (3.8)(iii) for x_n. Replacing x_n by a subsequence if need be and using the fact that $\mathscr{A}$ is compact, we may suppose $A_{in} \to_{d_H}$, say to A_i, as $n \to \infty$, $i = 1, \cdots, N(\delta)$.

Let $\gamma > 0$ and let $A \in \mathscr{A}$. As above, choose B such that $d_H(A, B) < \gamma$,

$x(B) = x(A)$, and x has no atoms on the boundary of B. If n is sufficiently large, $|x_n(B) - x(B)| \le \gamma$. Choose $A_{i_n,n}$ such that $\rho((A_{i_n,n}, x_n(A_{i_n,n})), (B, x_n(B))) \le \delta + \gamma$. Replacing x_n by a subsequence if need be, we may assume that $A_{i_n,n}$ converges to one of the A_i, $i = 1, \cdots, N(\delta)$; for the sake of definiteness, suppose it is A_1. By (3.8)(iii)(b), x has no atoms in $A_{i_n,n}^{h(\delta)} \backslash A_{i_n,n}$. Since the atoms of x_n converge to those of x and A_1 is closed, $x_n(A_{i_n,n}) \to x(A_1)$. For n sufficiently large, $\rho((A_{i_n,n}, x_n(A_{i_n,n})), (A_1, x(A_1))) < \gamma$. Hence $\rho((A, x(A)), (A_1, x(A_1))) \le \delta + 4\gamma$, which, since γ was arbitrary, shows that the A_i satisfy (3.8)(iii)(a).

Finally, since the set $A_i^{h(\delta)} \backslash A_i$ is open, x has an atom there only if x has an atom in $A_i^{h(\delta)-2\varepsilon} \backslash A_i^{2\varepsilon}$ for some ε. But then for n large enough, x_n would have an atom in $A_i^{h(\delta)-\varepsilon} \backslash A_i^{\varepsilon} \subseteq A_{in}^{h(\delta)} \backslash A_{in}$, a contradiction. □

LEMMA 3.3. *Suppose $y_n \in \mathscr{D}(\mathscr{A})$, $y \in \mathscr{C}(\mathscr{A})$ and $\| y_n - y \|_{\mathscr{A}} \to 0$. Suppose z_n, $z \in \mathscr{D}(\mathscr{A})$ and $z_n \to_{d_D} z$. Then $y_n + z_n \to_{d_D} y + z$.*

PROOF. Let $\varepsilon > 0$. Recalling that $\mathscr{A}$ is compact, choose $\delta < \varepsilon$ small enough so that $|y(A) - y(B)| < \varepsilon$ whenever $d_H(A, B) < \delta$ and $A, B \in \mathscr{A}$. Suppose n is large enough so that $\| y_n - y \|_{\mathscr{A}} < \varepsilon$ and $d_D(z_n, z) < \delta$.

Suppose $A \in \mathscr{A}$. By the definition of d_D, there exists $B_n \in \mathscr{A}$ such that $d_H(A, B_n) < \delta$ *and* $|z(A) - z_n(B_n)| < \delta$. Then

$$\begin{aligned} |(y + z)(A) - (y_n + z_n)(B_n)| &\le \delta + |y(A) - y(B_n)| + |y(B_n) - y_n(B_n)| \\ &\le \delta + 2\varepsilon \le 3\varepsilon. \end{aligned}$$

Hence $G(y + z) \subseteq (G(y_n + z_n))^{3\varepsilon}$ if n is large. By the same argument with the roles reversed, we can conclude $d_G(G(y + z), G(y_n + z_n)) \le 3\varepsilon$, from which the lemma follows. □

Our criterion for compactness is based on the following definition:

DEFINITION 3.6. Suppose for each $m \ge 1$ that η_m, R_m, and M_m are real numbers, h_m and N_m are functions as in (3.8), and ω_m an increasing function with $\omega_m(r) \to 0$ as $r \to 0$. Suppose Δ_m is a positive sequence tending to 0. Define $\mathscr{F}$ to be the set of functions x in $\mathscr{D}_0(\mathscr{A})$ such that for each m there is a purely atomic function $J_m(x)$ and a function $C_m(x) \in \mathscr{C}(\mathscr{A})$ with

(3.9) (i) $J_m(x) \in \mathscr{F}_{PA}(h_m, N_m, \eta_m, R_m)$,
(ii) (a) $\| C_m(x) \|_{\mathscr{A}} \le M_m$,
(b) $\sup\{ |C_m(x)(A) - C_m(x)(B)| : A, B \in \mathscr{A}, d_H(A, B) \le r\} \le \omega_m(r)$ for all r, and
(iii) $\| x - (J_m(x) + C_m(x)) \|_{\mathscr{A}} \le \Delta_m$.

Note that we allow the bounds and moduli of continuity of $C_m(x)$ to depend on m. Even for Lévy processes on the line, this is necessary. There, for example, we might take the $J_m(x)$ to be $\sum_{s \le t} \Delta X_s(\omega) 1_{(|\Delta X_s(\omega)| \ge 1/m)}$, $t \le 1$, and $C_m(x) \equiv c_m t$, where, in general, $c_m \to \infty$ as $m \to \infty$.

In Section 5 we will show that our partial-sum processes have, with high probability, paths in sets such as $\mathscr{F}$. There, $C_m(x)$ will be identically 0 and the $J_m(x)$ will be the purely atomic functions whose atoms are the points of the partial-sum process that are larger in absolute value than γ_m, where γ_m is a sequence of positive reals tending to 0.

The case where one has purely atomic functions converging uniformly to a continuous function (e.g., as in most empirical processes) can also be fitted into the above framework. For example, if x_n are purely atomic, converging uniformly to x continuous, then $\{x_n, x\}$ form a subset of a set $\mathscr{F}$ of the above type if we let $C_m(x_n) = C_m(x) = x$ for each n and $J_m(x_n) = J_m(x) \equiv 0$ for each n. Alternatively, we can let $C_m(x_n)$ be a "smoothed" version of x_n.

Our theorem is

THEOREM 3.4. *Suppose $\mathscr{F}$ is as in (3.9). Then $\mathscr{F}$ is compact relative to d_D.*

PROOF. Note that

$$\sup_{x\in\mathscr{F}} \| x \|_{\mathscr{A}} \le R_m + \sup_{x\in\mathscr{F}} \| C_m(x) \|_{\mathscr{A}} + \sup_{x\in\mathscr{F}} \| x - (J_m(x) + C_m(x)) \|_{\mathscr{A}} < \infty.$$

Thus, as in the proof of Proposition 3.2, it suffices to show that if $x_n \in \mathscr{F}$, $G(x_n) \to G_0$, then there exists $x \in \mathscr{F}$ such that $G(x) = G_0$.

Using (3.9)(i) and Proposition 3.2, (3.9)(ii) and Ascoli-Arzela, and a diagonalization argument, we may replace x_n by a subsequence (also denoted x_n) such that $J_m(x_n) \to_{d_D}$ to an element of $\mathscr{F}_{PA}(h_m, N_m, \eta_m, R_m)$, call it j_m, and $C_m(x_n) \to_{\|\cdot\|_{\mathscr{A}}}$ to an element of $\mathscr{C}(\mathscr{A})$, call it c_m, for each $m = 1, 2, \cdots$.

The main step is to show $\{j_m + c_m\}$ is a Cauchy sequence with respect to $\|\cdot\|_{\mathscr{A}}$. Let $\varepsilon > 0$. Suppose M is large enough so that whenever $m \ge M$,

$$\sup_{x\in\mathscr{F}} \| x - (J_m(x) + C_m(x)) \|_{\mathscr{A}} \le \varepsilon. \tag{3.10}$$

Suppose $k, m \ge M$ and for some $A \in \mathscr{A}$, $|(j_m(A) + c_m(A)) - (j_k(A) + c_k(A))| \ge 5\varepsilon$. Since c_m and c_k are continuous and j_m and j_k are purely atomic, there exists a set B with $A \subseteq B^0 \subseteq A^\gamma$ for some γ such that $j_m(B) = j_m(A)$, $j_k(B) = j_k(A)$, neither j_m nor j_k have any atoms on the boundary of B, and $|c_k(A) - c_k(B)|$, $|c_m(A) - c_m(B)| \le \varepsilon$.

Then

$$|(j_m(B) + c_m(B)) - (j_k(B) + c_k(B))| \ge 3\varepsilon.$$

Since neither j_k nor j_m have any atoms on the boundary of B, as we argued in Proposition 3.2, $J_k(x_n)(B) \to j_k(B)$ and $J_m(x_n)(B) \to j_m(B)$ as $n \to \infty$. Thus, for n sufficiently large,

$$|(J_m(x_n)(B) + C_m(x_n)(B)) - (J_k(x_n)(B) + C_k(x_n)(B))| \ge 2\varepsilon,$$

contradicting (3.10).

Therefore, $\|(j_m + c_m) - (j_k + c_k)\|_{\mathscr{A}} < 5\varepsilon$, and $\{j_m + c_m\}$ is a Cauchy sequence. Since $\|\cdot\|_{\mathscr{A}}$ is complete, $j_n + c_m \to_{\|\cdot\|_{\mathscr{A}}}$, say to y. Let $J_m(y) = j_m$, $C_m(y) = c_m$ for each m. For each $\varepsilon > 0$,

$$\begin{aligned}\| y - (j_m + c_m) \|_{\mathscr{A}} &\leq \limsup_{k\to\infty} \| (j_k + c_k) - (j_m + c_m) \|_{\mathscr{A}} \\ &\leq \limsup_{k\to\infty} \limsup_{n\to\infty} \| (J_k(x_n) + C_k(x_n)) \\ &\qquad - (J_m(x_n) + C_m(x_n)) \|_{\mathscr{A}} + 2\varepsilon \\ &\leq \limsup_{k\to\infty} \Delta_k + \Delta_m + 2\varepsilon,\end{aligned}$$

the argument for the second equality being similar to the proof that $\{j_m + c_m\}$ is Cauchy. It is now easy to see that $y \in \mathscr{F}$.

Finally, let $\varepsilon > 0$. By Lemma 3.3,

$$d_G(G(j_m + c_m), G(J_m(x_n) + C_m(x_n))) < \varepsilon$$

if n is large enough. Since clearly $d_G(G(j_m + c_m), G(y))$ and $d_G(G(J_m(x_n) + C_m(x_n)), G(x_n)) < \Delta_m$, and also $d_G(G(x_n), G) < \varepsilon$ if n is large enough, then $d_G(G(y), G_0) < 2\varepsilon + 2\Delta_m$. Hence $G(y) = G_0$, and the theorem is proved. □

The next theorem concerns d_D convergence when the limit is a continuous function.

THEOREM 3.5. *Suppose $x_n \to x$ relative to d_D and x is continuous. Then $x_n \to x$ relative to the uniform norm $\| \cdot \|_{\mathscr{A}}$.*

PROOF. Let $\varepsilon > 0$. Choose $\delta < \varepsilon$ so that if $d_H(A, B) < \delta$, then $| x(A) - x(B) | < \varepsilon$. Suppose $d_D(x_n, x) < \delta$. If $A \in \mathscr{A}$, there exists $B \in \mathscr{A}$ such that $\rho((A, x_n(A)), (B, x(B))) < \delta$. Then $d_H(A, B) < \delta$, and so

$$| x_n(A) - x(A) | \leq | x_n(A) - x(B) | + | x(B) - x(A) | < \delta + \varepsilon \leq 2\varepsilon.$$

Hence $\| x_n - x \|_{\mathscr{A}} \leq 2\varepsilon$. □

4. Weak convergence. In this section we study weak convergence with respect to the topology introduced in Section 3. We suppose throughout that (A1) holds.

First, we rather easily dispose of the question of measurability for processes with sample paths in $\mathscr{D}(\mathscr{A})$. Recall that the graph function G has domain $\mathscr{D}(\mathscr{A})$. The σ-field $\mathscr{B}$ induced on $\mathscr{D}(\mathscr{A})$ by G and the Borel σ-field of $\mathscr{G}$, $\mathscr{B}_{\mathscr{G}}$, are related by

$$\mathscr{B} = G^{-1}(\mathscr{B}_{\mathscr{G}}). \tag{4.1}$$

Let $\mathscr{A}^*$ be a subset of $\mathscr{A}$ that is dense with respect to d_H. Let $\pi_A: \mathscr{D}(\mathscr{A}) \to \mathbb{R}$ denote the one-dimensional projection. Thus $\pi_A(x) = x(A)$. Analogously, let $\pi_{A_1,\cdots,A_k}(x) = (x(A_1), \cdots, x(A_k))$.

Now let $\mathscr{B}_0 = \sigma(\{\pi_A: A \in \mathscr{A}^*\})$, the smallest σ-field with respect to which each of these one-dimensional projections is measurable.

PROPOSITION 4.1. $\mathscr{B}_0 = \mathscr{B}$.

PROOF. Since $\mathscr{A}$ is separable, we can without loss of generality assume $\mathscr{A}^*$ is countable. For B in $\mathscr{G}$, let

$$B_A = B \cap (A \times \mathbb{R}), \quad A \in \mathscr{A}. \tag{4.2}$$

Since elements of $\mathscr{A}$ and $\mathscr{G}$ are closed, it is clear that

$$G^{-1}(B) = \cap_{A \in \mathscr{A}^*} \{x \in \mathscr{D}(\mathscr{A}): (A, x(A)) \in B_A\}. \tag{4.3}$$

This, together with the easily checked fact that $\mathscr{B}_0 \subseteq \mathscr{B}$, proves the proposition. □

The above result shows that to check measurability of a $\mathscr{D}(\mathscr{A})$-valued process X, it suffices to check the measurability of the one-dimensional projections $X(A)$, $A \in \mathscr{A}^*$.

A simple monotone class argument applied to (4.3) shows that the finite-dimensional distributions are a determining class for probabilities on $\mathscr{D}(\mathscr{A})$. More precisely,

PROPOSITION 4.2. *Suppose P and Q are two probabilities supported on $(\mathscr{D}(\mathscr{A}), \mathscr{B})$ such that for all choices $A_1, \cdots, A_k \in \mathscr{A}^*$, $k \geq 1$,*

$$P \circ \pi^{-1}_{A_1,\cdots,A_k} = Q \circ \pi^{-1}_{A_1,\cdots,A_k}.$$

Then $P = Q$.

As usual, we say that probabilities P_n on $\mathscr{D}(\mathscr{A})$ converge weakly to a probability P, denoted $P_n \to_w P$, if $\int f\,dP_n \to \int f\,dP$ for all f continuous and bounded on $\mathscr{G}$. In view of Proposition 4.2, Prohorov's theorem (cf. Billingsley, 1968, page 37), the fact that $\mathscr{G}$ is a complete and separable metric space, the imbedding of $\mathscr{D}(\mathscr{A})$ into $\mathscr{G}$ by means of the mapping $x \to G(x)$, and standard arguments, we have

THEOREM 4.3. *Suppose for all ε, there exists a compact subset $\mathscr{F}_\varepsilon$ of $\mathscr{D}(\mathscr{A})$ such that $\inf_n P_n(\mathscr{F}_\varepsilon) \geq 1 - \varepsilon$. Then there exists a subsequence of the P_n's which converge weakly to a probability P on $\mathscr{D}(\mathscr{A})$. If, moreover,*

$$P_n \circ \pi^{-1}_{A_1,\cdots,A_k} \to_w P \circ \pi^{-1}_{A_1,\cdots,A_k}$$

for all choices of $A_1, \cdots, A_k \in \mathscr{A}^*$, $k \geq 1$, then $P_n \to_w P$.

It is perhaps worth mentioning that Skorokhod's representation theorem still holds. Thus if X_n, X are set-indexed processes with the law of X_n converging weakly to that of X, then we can find another probability space and processes X_n', X' equal in law to X_n, X, respectively, such that $d_D(X_n', X') \to 0$, a.s. Since $\mathscr{G}$ is a complete and separable metric space, this is a special case of the general result of Skorokhod (1956a); see also Billingsley (1971) and Pollard (1979).

By Theorem 4.3, to prove weak convergence, one must show that the finite-dimensional distributions converge and prove tightness. Criteria for a subset of $\mathscr{D}(\mathscr{A})$ to be compact are given above in Section 3, but as they stand, it may be difficult in general cases to show that a given process is in a compact set with high probability. For the applications to our central limit theorem of Section 5, we can achieve a considerable simplification since our partial-sum processes will have sample paths in $\mathscr{D}_0(\mathscr{A}) \subset \mathscr{D}(\mathscr{A})$.

Let $C_{n\mathbf{j}} = n^{-1}(\mathbf{j} - \mathbf{1}, \mathbf{j}]$ for $\mathbf{j} \in J^d$ and $J = \{1, 2, \cdots\}$. Let T_n be a purely atomic set-indexed process on $\mathscr{A}$. Clearly the domain of T_n can be considered to contain all subsets of I^d, and hence in particular, the cubes $C_{n\mathbf{j}}$. Set

$$\xi_n(A) = \#\{atoms\ of\ T_n\ in\ A\}, \quad \xi_{n\mathbf{j}} = \xi_n(C_{n\mathbf{j}}).$$

Assume that T_n satisfies the following condition:

(4.4) (i) For each n, $\{\xi_{n\mathbf{j}}: n^{-1}\mathbf{j} \in I^d\}$ are i.i.d. Bernoulli r.v.'s with parameter $p_n \le c_1 n^{-d}$ for a positive constant c_1. (In particular, this implies that the number of atoms in each cube $C_{n\mathbf{j}}$ is either 0 or 1.)

(ii) If $A \subset C_{n\mathbf{j}}$,

$$P(\xi_n(A) = 1 \mid \xi_{n\mathbf{j}} = 1) = |A|/|C_{n\mathbf{j}}|$$

so that the location of any atom in $C_{n\mathbf{j}}$ is uniformly distributed there.

(iii) As $M_0 \to \infty$,

$$P(|T_n(C_{n\mathbf{j}})| > M_0 \mid \xi_{n\mathbf{j}} = 1) \to 0$$

uniformly in $\mathbf{j}$ and n.

This condition will be satisfied by the "large atoms" of our partial-sum processes.

A key result is

THEOREM 4.4. *If T_n satisfies* (4.4), *then given $\varepsilon > 0$, there exist η, M, N, and h (independent of n) such that $P(T_n \in \mathscr{F}_{PA}(h, N, \eta, M)) \ge 1 - \varepsilon$ for all n.*

Before proving this result, we first need a lemma.

LEMMA 4.5. *Suppose T_n satisfies* (4.4).

(a) *Let $\mathbf{t} \in I^d$ be fixed. Then*

$$r_1 := P(\xi_n(\{\mathbf{s}: |\mathbf{s} - \mathbf{t}| \le \varepsilon\}) > 1) \le c_2 \varepsilon^{2d},$$

with c_2 independent of n and $\mathbf{t}$.

(b) $r_2 := P(\xi_n(A) > 0) \le c_3 |A|$, *$c_3$ independent of n.*

(c) *If $A_1, \cdots, A_m$ are disjoint,*

$$r_3 := P(\xi_n(A_i) > 0,\ 1 \le i \le m \mid \xi_n(I^d) = m) \le c_4 \textstyle\prod_{i=1}^m |A_i|;$$

here c_4 may depend on m but not on n.

PROOF. Measurability is not a difficulty here, since all the relevant events are measurable with respect to the σ-field generated by the events (T_n has an atom in $(\mathbf{s}, \mathbf{t})$ of size $> a$: $\mathbf{s}, \mathbf{t} \in I^d$, $a \in R$).

To prove (a), note that $\{\mathbf{s}: |\mathbf{s} - \mathbf{t}| \le \varepsilon\}$, the ball of radius ε about $\mathbf{t}$, intersects at most $K(n, \varepsilon)$ cubes $C_{n\mathbf{j}}$, where $K(n, \varepsilon) = k_0([n^d \varepsilon^d] \vee 1)$ for an integer k_0 independent of n and ε. We need to bound the probability that this ball contains two or more atoms of T_n. Let $B(k, p)$ denote a Binomial r.v. with parameters k

and p. If $n\,\varepsilon \le 1$, by (4.4)(ii)

$$r_1 \le P(B(k_0,\, c_5\varepsilon^d) > 1) \le c_6\varepsilon^{2d}$$

where c_5 is the volume of the unit ball in $\mathbb{R}^d$.

If $n\,\varepsilon \ge 1$, by (4.4)(i),

$$r_1 \le P(B(K(n,\,\varepsilon),\, p_n) > 1) \le c_7(k_0[n^d\varepsilon^d]p_n)^2 \le c_8\varepsilon^{2d}.$$

This proves (a). Next, consider (b). Consider first the case where $A \subseteq C_{n\mathbf{j}}$ for some $\mathbf{j}$. Then by (4.4)(i, ii),

$$r_2 = P(\xi_n(A) > 0 \mid \xi_{n\mathbf{j}} = 1)P(\xi_{n\mathbf{j}} = 1) \le (\mid A \mid / C_{n\mathbf{j}} \mid)c_1 n^{-d} = c_1 \mid A \mid.$$

Since in the general case we can write $A = \cup_{\mathbf{j}}\, (A \cap C_{n\mathbf{j}})$, the above suffices to prove (b) because of subadditivity and linearity.

Finally, to prove (c), fix n and m. As above, we may assume each A_i is wholly contained in some $C_{n\mathbf{j}}$. Observe that without loss of generality we may assume that each $C_{n\mathbf{j}}$ intersects only one A_i. By (4.6)(ii), it suffices to suppose each A_i equals some $C_{n\mathbf{j}}$. But then

$$r_3 = \binom{n^d}{m}^{-1} \le c_9(n^d)^{-m} = c_9 \prod_{i=1}^{m} \mid A_i \mid. \qquad \square$$

PROOF OF THEOREM 4.4. First of all, by Lemma 4.5(a),

$$\begin{aligned} P(\mathrm{gap}(T_n) \le k^{-1}) &= P(\text{for some } \mathbf{j} \in J^d \text{ with } k^{-1}\mathbf{j} \in I^d, \text{ the ball of radius} \\ &\qquad 3/k \text{ about } k^{-1}\mathbf{j} \text{ contains at least two atoms of } T_n) \\ &\le k^d c_2 (3/k)^{2d} \to 0 \end{aligned}$$

as $k \to \infty$. Thus we may choose η sufficiently small so that $P(\mathrm{gap}(T_n) < \eta) < \varepsilon/4$.

On the set where $\mathrm{gap}(T_n) \ge k^{-1}$, T_n can have at most $c_{10}k^d$ atoms for a suitable constant c_{10}. This fact together with (4.4) (iii) shows that for a given $\varepsilon > 0$ we may choose M_0 sufficiently large so that if $M = c_{10}k^d M_0$,

$$P(\mathrm{Variation}(T_n) \ge M,\ \mathrm{gap}(T_n) \ge k^{-1}) \le \varepsilon/4.$$

Fix δ. We want to show next that we can find N and sets $A_1, \cdots, A_N$ so that $\{(A_i,\, T_n(A_i)) \colon i = 1, \cdots, N\}$ is a δ-net for $G(T_n)$ with high probability. Although the A_i's could be allowed to depend on ω, that is not necessary here. If $\mathscr{A}_\delta$ is given by (A1)(iii), let $\mathscr{B}_q = \cup_{i=1}^{q} \mathscr{A}_{1/i}$. Let $G_q(x) = \{(A,\, x(A)) \colon A \in \mathscr{B}_q\}$ be the graph of the restriction of x to $\mathscr{B}_q$. We now show that if q is sufficiently large,

$$\text{(4.5)} \qquad P(G_q(T_n) \text{ is not a } \delta\text{-net for } G(T_n),\ \mathrm{gap}(T_n) \ge k^{-1}) \le \varepsilon/4.$$

Since T_n has at most $c_{10}k^d$ atoms when $\mathrm{gap}(T_n) \ge k^{-1}$, it will suffice to show that for each fixed m

$$P(G_q(T_n) \text{ is not a } \delta\text{-net for } G(T_n) \mid \xi_n(I^d) = m)$$

can be made arbitrarily small if q is large.

Let $\mathbf{t}_1, \cdots, \mathbf{t}_m$ be given. Let $A \in \mathscr{A}$. By (A1)(iii), there exists a q and a $B \in \mathscr{B}_q$

such that $A \subseteq B$, $B \backslash A$ contains no $\mathbf{t}_i$'s and $d_H(A, B) < \delta/2$. If x is *any* purely atomic function that has exactly m atoms at locations $\mathbf{t}_1, \cdots, \mathbf{t}_m$, then $x(B) = x(A)$ and $\rho((A, x(A)), (B, x(B))) < \delta/2$.

Let $U_1, \cdots, U_u$ be a finite collection of balls of d_H diameter $\delta/2$ that cover $\mathscr{A}$. Let $\{\tau_i, i = 1, \cdots, 2^m\}$ be the collection of all subsets of $\{\mathbf{t}_1, \cdots, \mathbf{t}_m\}$. For each $i = 1, \cdots, 2^m$, $l = 1, \cdots, u$, let $A_{il} \in U_l \cap \mathscr{A}$ denote a set satisfying $A_{il} \cap \{\mathbf{t}_1, \cdots, \mathbf{t}_m\} = \tau_i$ whenever such a set exists. For each such A_{il}, select a q_{il} and a B_{il} as in the preceding paragraph. Let $q = \max_{i,l} q_{il}$.

Let x be any purely atomic function whose atoms are at $\mathbf{t}_1, \cdots, \mathbf{t}_m$. We claim $G_q(x)$ is a δ-net for $G(x)$. For if $A \in \mathscr{A}$, then $A \in U_l$ for some l and $A \cap \{\mathbf{t}_1, \cdots, \mathbf{t}_m\} = \tau_i$ for some i. Then there must exist an $A_{il} \in U_l \cap \mathscr{A}$ with $A_{il} \cap \{\mathbf{t}_1, \cdots, \mathbf{t}_m\} = \tau_i$ for which $d_H(A, B_{il}) \leq d_H(A, A_{il}) + d_H(A_{il}, B_{il}) < \delta$ and $x(A) = x(A_{il}) = x(B_{il})$, which proves that $G_q(x)$ is a δ-net for $G(x)$.

Now let us turn this around. Let

$$W_q = \{(\mathbf{t}_1, \cdots, \mathbf{t}_m)\text{: there exists a purely atomic } x \text{ with atoms located at } \mathbf{t}_1, \cdots, \mathbf{t}_m \text{ such that } G_q(x) \text{ is not a } \delta\text{-net for } G(x)\}. \tag{4.6}$$

What we have shown, then, is that $W_q \downarrow \phi$. The W_q are subsets of $(I^d)^m$. It is not hard to see that the W_q are open, hence measurable; hence the Lebesgue measure of the $W_q \to 0$. For each q, select countably many sets V_{qi}, $i = 1, \cdots$, such that $W_q \subseteq \cup_{i=1}^{\infty} V_{qi}$, $|\cup_{i=1}^{\infty} V_{qi}| \leq 2|W_q|$, and each V_{qi} is of the form $A_{qi1} \times \cdots \times A_{qim}$ where the A_{qil}'s are disjoint subsets of I^d.

Then by Lemma 4.5(c),

$$\begin{aligned}
&P(G_q(T_n) \text{ is not a } \delta\text{-net for } G(T_n) \mid \xi_n(I^d) = m) \\
&\quad \leq P(\text{there is a labeling } (\mathbf{t}_1, \cdots, \mathbf{t}_m) \text{ of the atoms of } T_n \\
&\qquad \text{with } (\mathbf{t}_1, \cdots, \mathbf{t}_m) \in W_q \mid \xi_n(I^d) = m) \\
&\quad \leq \textstyle\sum_i P(T_n \text{ has atoms in each of } A_{qi1}, \cdots, A_{qim} \mid \xi_n(I^d) = m) \\
&\quad \leq \textstyle\sum_i c_4 \prod_{l=1}^{m} |A_{qil}| \leq \sum_i c_4 |V_{qi}| \leq 2c_4 |W_q| \to 0
\end{aligned}$$

as $q \to \infty$.

We have thus shown that by taking q sufficiently large and letting $N = N(\delta) = \#(\mathscr{B}_q)$, then $\{(A, T_n(A))\colon A \in \mathscr{B}_q\}$ is a δ-net for $G(T_n)$ with high probability. Finally, if $A \in \mathscr{B}_q$, by Lemma 4.5(b),

$$P(T_n \text{ has an atom in } A^h \backslash A) \leq c_3 |A^h \backslash A| \to 0$$

as $h \to 0$. Since $\mathscr{B}_q$ is a finite set, we can take $h = h(\delta)$ small enough so that

$$P(T_n \text{ has an atom in } A^h \backslash A \text{ for some } A \in \mathscr{B}_q) < \varepsilon/4.$$

Thus $P(T_n \notin \mathscr{F}_{PA}(h, N, \eta, M)) < \varepsilon$. $\square$

5. Central limit theorem for partial-sum processes. In this section we prove a central limit theorem for partial-sum processes in the domain of normal attraction of a stable law of index α. The main result is for the cases of $\alpha \in$

(1, 2). The cases $\alpha \in (0, 1]$ will be discussed at the end of this section, while some remarks concerning more general central limit theorems are given in Section 6.

Consider an array of i.i.d. r.v.'s $X_{\mathbf{j}}$: $\mathbf{j} \in J^d\}$ where each $X_{\mathbf{j}}$ is in the domain of normal attraction of a stable law of exponent $\alpha \in (1, 2)$ and $J = \{1, 2, \cdots\}$. Without loss of generality, assume $EX_{\mathbf{j}} = 0$. For this context, we define the partial-sum process indexed by $\mathscr{A}$, $S_n \equiv \{S_n(A): A \in \mathscr{A}\}$, by

$$S_n(A) = n^{-d/\alpha} \textstyle\sum_{\mathbf{j}\in nA} X_{\mathbf{j}}. \tag{5.1}$$

Just as in the case where $X_{\mathbf{j}}$ is in the domain of normal attraction of the Normal law, there is no central limit theorem for S_n itself; rather it is first necessary to "smooth" S_n (cf. Pyke, 1983; and Bass and Pyke, 1984a). By the "smoothing" of a partial-sum process, we mean a deterministic or random perturbation of the masses $X_{\mathbf{j}}$. In Pyke (1973, 1983) and Bass and Pyke (1984a) the smoothing that was considered spread the mass $X_{\mathbf{j}}$ uniformly over the cube $C_{\mathbf{j}} := (\mathbf{j} - \mathbf{1}, \mathbf{j}] = nC_{n\mathbf{j}}$ in the sense that the purely atomic set function $X_{\mathbf{j}}\delta_{\mathbf{j}}(nA)$, where $\delta_x(B) = 1$ or 0 according as $x \in B$ or $x \notin B$, is replaced by $X_{\mathbf{j}} \,|\, A \cap C_{n\mathbf{j}} \,|$ to give

$$X_n(A) := n^{-d/\alpha} \textstyle\sum_{\mathbf{j}\in J^d} |\, A \cap C_{n\mathbf{j}} \,|\, X_{\mathbf{j}}. \tag{5.2}$$

The main reason that this smoothing could be used for these previous applications is that the limiting processes (Brownian processes, either tied-down or not) were always in $\mathscr{C}(\mathscr{A})$. Thus $X_n(\cdot)$, which is continuous, could be expected to be a suitable perturbation of the S_n-process.

In the present paper, where the natural limiting processes are not Gaussian and whose paths are not continuous but are in $\mathscr{D}_0(\mathscr{A})$, one cannot expect to be able to smooth the partial-sum processes in a manner as in (5.2) that results in continuous paths. Rather, we retain the discontinuous purely atomic nature of S_n by simply perturbing the location of the atoms. Namely, let $\{U_{\mathbf{j}}: \mathbf{j} \in J^d\}$ be independent r.v.'s that are independent of the $X_{\mathbf{j}}$ with $U_{\mathbf{j}}$ being uniform over $C_{\mathbf{j}}$. Then write

$$Y_n(A) = n^{-d/\alpha} \textstyle\sum_{U_{\mathbf{j}}\in nA} X_{\mathbf{j}}. \tag{5.3}$$

This type of smoothing involves a random relocation of the grid points and can be viewed as being determined by a random transformation of the sets in $\mathscr{A}$. The mapping $\mathbf{j} \to U_{\mathbf{j}}$ determines, by linear interpolation for example, a transformation τ_n: $I^d \to I^d$ (note that $U_{\mathbf{j}} \le \mathbf{j}$) that maps $\mathbf{j}/n \to U_{\mathbf{j}}/n$. By interpreting $\tau_n(A)$ in the natural way, we see that

$$Y_n(A) = S_n(\tau_n(A)) \tag{5.4}$$

for all $A \subset I^d$.

Let us now assume $\mathscr{A}$ satisfies (A1) and (A2). Moreover, let us assume that the r defined in (A2)(ii) satisfies

$$1 < r < (\alpha - 1)^{-1}. \tag{5.5}$$

There is no loss of generality in taking $r > 1$ since increasing the size of $\mathscr{A}$ merely results in a stronger theorem. On the other hand, the existence of a limit law with paths in $\mathscr{D}(\mathscr{A})$ requires $r \le (\alpha - 1)^{-1}$ (Bass and Pyke, 1984b); hence (5.5) is optimal, except for the boundary case $r = (\alpha - 1)^{-1}$. See Section 6.

By assumption, the $X_{\mathbf{j}}$ are in the domain of normal attraction of a stable law of index α. Let ν be the Lévy measure of that stable law, and let $Z := \{Z(A): A \in \mathscr{A}\}$ be a Lévy process with Lévy measure ν; cf. Bass and Pyke (1984b). We need to show

PROPOSITION 5.1. *The finite-dimensional distributions of Y_n converge to those of Z.*

PROOF. For notational simplicity, we will show that the two-dimensional distributions converge, the general case being completely analogous. For any Borel set A in I^d, define

$$\mu_n(A) = \textstyle\sum_{\mathbf{j}} 1_{nA}(U_{\mathbf{j}}) = \#\{\mathbf{j}: U_{\mathbf{j}} \in nA\}. \tag{5.6}$$

Note that $\mu_n(A)$ is a sum of independent Bernoulli r.v.'s. Since

$$E\,\mu_n(A) = \textstyle\sum_{\mathbf{j}} |A \cap C_{n\mathbf{j}}| = |nA| = n^d|A|,$$

and

$$\operatorname{var}(n^{-d}\mu_n(A)) = n^{-2d} \textstyle\sum_{\mathbf{j}} |A \cap C_{n\mathbf{j}}|\{1 - |A \cap C_{n\mathbf{j}}|\} \le n^{-d}|A|,$$

it follows by the Borel-Cantelli lemma and Chebyshev's inequality that

$$\mu_n(A)/n^d \to |A|, \quad \text{a.s.} \tag{5.7}$$

Now let us consider the joint characteristic function of $Y_n(A)$, $Y_n(B)$. Let f be the characteristic function of $X_{\mathbf{j}}$, and f_0 the characteristic function of the limiting stable law. Then, using the independence of the $U_{\mathbf{j}}$'s from the $X_{\mathbf{j}}$'s,

$$\begin{aligned}
&E\exp(i\{uY_n(A) + vY_n(B)\}) \\
&\quad = E[E[\exp(i \textstyle\sum_{\mathbf{j}} n^{-d/\alpha}X_{\mathbf{j}}\{u1_{nA}(U_{\mathbf{j}}) + v1_{nB}(U_{\mathbf{j}})\}) \mid U_{\mathbf{j}}, \mathbf{j} \in J^d]] \\
&\quad = E[\textstyle\prod_{\mathbf{j}} f(n^{-d/\alpha}\{u1_{nA}(U_{\mathbf{j}}) + v1_{nB}(U_{\mathbf{j}})\})] \\
&\quad = E[f((u+v)n^{-d/\alpha})^{n^d(\mu_n(A\cap B)n^{-d})} f(un^{-d/\alpha})^{n^d(\mu_n(A\setminus B)n^{-d})} \times f(vn^{-d/\alpha})^{n^d(\mu_n(B\setminus A)n^{-d})}] \\
&\quad \to f_0(u+v)^{|A\cap B|} f_0(u)^{|A\setminus B|} f_0(v)^{|B\setminus A|} \\
&\quad = E\exp(i\{uZ(A) + vZ(B)\}).
\end{aligned} \tag{5.8}$$

The limit follows from the facts that $f(wn^{-d/\alpha})^{n^d} \to f_0(w)$ for all w, $|f| \le 1$, $f_0 \ne 0$ and (5.7) together with the dominated convergence theorem. □

We do not need the convergence of the finite-dimensional distributions of S_n, but if A has a smooth boundary in the sense that $|A(\delta)| \to 0$ as $\delta \to 0$, where $A(\delta) = \{x: \text{distance of } x \text{ to } \partial A < \delta\}$, then $S_n(A) - Y_n(A) \to 0$ in probability. In fact,

$$S_n(A) - Y_n(A) = \textstyle\sum_{\mathbf{j}} n^{-d/\alpha}X_{\mathbf{j}}(1_{nA}(\mathbf{j}) - 1_{nA}(U_{\mathbf{j}})).$$

Under the assumption of a smooth boundary, it is not hard to show that

$$E \sum_{\mathbf{j}} | 1_{nA}(\mathbf{j}) - 1_{nA}(U_{\mathbf{j}}) | / n^d = \sum_{\mathbf{j} \notin nA} | A \cap C_{n\mathbf{j}} | + \sum_{\mathbf{j} \in nA} | A^c \cap C_{n\mathbf{j}} | \to 0$$

and hence that

$$n^{-d} \sum_{\mathbf{j}} | 1_{nA}(\mathbf{j}) - 1_{nA}(U_{\mathbf{j}}) | \to_P 0.$$

An argument using characteristic functions as in Proposition 5.1 then shows that $E \exp(iu(S_n(A) - Y_n(A))) \to 1$.

The following result is needed to prove tightness of the Y_n. Let

$$X_{\mathbf{j},n}^{(\gamma)} = X_{\mathbf{j}} 1_{[|X_{\mathbf{j}}| \le \gamma n^{d/\alpha}]}, \tag{5.9}$$

$$Y_n^{(\gamma)}(A) = n^{-d/\alpha} \sum_{\mathbf{j}} 1_{nA}(U_{\mathbf{j}}) X_{\mathbf{j},n}^{(\gamma)}, \tag{5.10}$$

and

$$y_n^{(\gamma)}(A) = n^{-d/\alpha} \sum_{\mathbf{j}} 1_{nA}(U_{\mathbf{j}}) E X_{\mathbf{j},n}^{(\gamma)}. \tag{5.11}$$

The process $y_n^{(\gamma)}$ has atoms at the same random locations as $Y_n^{(\gamma)}$, but with deterministic masses.

PROPOSITION 5.2. *Let $\varepsilon > 0$ and $\eta > 0$. Then for all $\gamma > 0$ sufficiently small*

$$P[\| Y_n^{(\gamma)} - y_n^{(\gamma)} \|_{\mathscr{A}} \ge \eta] \le \varepsilon \text{ for all } n.$$

PROOF. Since Y_n is purely atomic, $\| Y_n^{(\gamma)} - y_n^{(\gamma)} \|_{\mathscr{A}} \to 0$, a.s. as $\gamma \to 0$ for each n. It thus suffices to find γ and n_0 such that

$$P[\| Y_n^{(\gamma)} - y_n^{(\gamma)} \|_{\mathscr{A}} \ge \eta] \le \varepsilon \quad \text{for all} \quad n \ge n_0. \tag{5.12}$$

For $m \ge 1$, $n \ge 1$, define

$$Y_{n,m}(A) = n^{-d/\alpha} \sum_{\mathbf{j}} 1_{nA}(U_{\mathbf{j}}) X_{\mathbf{j},n,m} \tag{5.13}$$

and

$$y_{n,m}(A) = n^{-d/\alpha} \sum_{\mathbf{j}} 1_{nA}(U_{\mathbf{j}}) E X_{\mathbf{j},n,m}, \tag{5.14}$$

where

$$X_{\mathbf{j},n,m} = X_{\mathbf{j}} 1_{[|X_{\mathbf{j}}| \in (a_{m+1}, a_m) n^{d/\alpha}]} \tag{5.15}$$

and where the sequence of constants a_m will be specified later with $a_m \to 0$. Set

$$f_{n,m} = P[a_{m+1} < n^{-d/\alpha} | X_{\mathbf{j}} | \le a_m]. \tag{5.16}$$

We will need the following bounds for the mean and variances of $X_{\mathbf{j},n,m}$:

$$E | X_{\mathbf{j},n,m} | \le a_m n^{d/\alpha} f_{n,m} \tag{5.17}$$

and

$$E | X_{\mathbf{j},n,m} |^2 \le a_m^2 f_{n,m} n^{2d/\alpha}. \tag{5.18}$$

To prove the proposition, it suffices to prove that for $\varepsilon > 0$ and $\eta > 0$, there exists $m_0 = m_0(\varepsilon, \eta)$ and a sequence $a_m \to 0$ such that

$$P[\| \textstyle\sum_{m \ge m_0} (Y_{n,m} - y_{n,m}) \|_{\mathscr{A}} > \eta] \le \varepsilon. \tag{5.19}$$

For simplicity, assume that the X_j are symmetric so that $EX_{j,n,m} = 0$ and hence $y_{n,m} = 0$. At the end of the proof, we will show that this assumption is unnecessary.

We prove (5.19) by showing the existence of constants η_m and ε_m such that

$$\textstyle\sum_{m=1}^{\infty} 3\eta_m \le \eta, \ \sum_{m=1}^{\infty} \varepsilon_m \le \varepsilon/3,$$

and

$$P[\| Y_{n,m} \|_{\mathscr{A}} > 3\eta_m] \le \varepsilon_m, \ m > m_0.$$

Let $0 < \beta < 1$, let δ_0 be fixed, and let $\delta_m = \delta_0 \beta^m$. Consider first of all (for any $\delta_m > 0$) that

$$\begin{aligned} P[\| Y_{n,m} \|_{\mathscr{A}} > 3\eta_m] & \\ &\le P[\| Y_{n,m} \|_{\mathscr{A}_{\delta_m}} > \eta_m] \\ &\quad + P[\max_{A, A^+ \in \mathscr{A}_{\delta_m}, |A^+ \setminus A| < \delta_m} \sup_{B \in \mathscr{A}, A \subseteq B \subseteq A^+} | Y_{n,m}(B \Delta A) | > 2\eta_m] \\ &:= P_1(n, m) + P_2(n, m). \end{aligned}$$

By the usual nesting argument (cf. Bass and Pyke, 1984b, or Dudley, 1973, the second probability is split again based on the inequality

$$\begin{aligned} \max_{A, A^+ \in \mathscr{A}_{\delta_m}, |A^+ \setminus A| < \delta_m} & \sup_{B \in \mathscr{A}, A \subseteq B \subseteq A^+} | Y_{n,m}(B \Delta A) | \\ &\le \textstyle\sum_{j=m}^{k_m - 1} \max_{A \in \mathscr{A}_{\delta_j}, B \in \mathscr{A}_{\delta_{j+1}}} | Y_{n,m}(A \Delta B) | \\ &\quad + \max_{A, A^+ \in \mathscr{A}_{\delta_{k_m}}, |A^+ \setminus A| < \delta_{k_m}} \sup_{B \in \mathscr{A}, A \subseteq B \subseteq A^+} | Y_{n,m}(A \Delta B) | \\ &:= T_{21} + T_{22}, \end{aligned} \tag{5.20}$$

where k_m will be chosen later.

By Bernstein's inequality

$$P_1(n, m) \le 2 \exp\left\{ H(\delta_m) - \frac{\eta_m^2}{2(n^d a_m^2 f_{n,m} + \eta_m a_m/3)} \right\}.$$

To obtain the necessary convergent series, we will show that we can obtain

$$H(\delta_m) \le \ln \varepsilon_m + \frac{\eta_m^2}{4n^d a_m^2 f_{m,n}} \wedge \frac{3\eta_m}{4a_m}. \tag{5.21}$$

To satisfy this inequality, we use the known characterization of the d.f. of a r.v. in the domain of normal attraction of a stable-α distribution (cf. Gnedenko and Kolmogorov, 1954, page 182) to obtain

$$\begin{aligned} f_{n,m} &= P[a_{m+1} < | X_j | n^{-d/\alpha} \le a_m] \\ &\le P[| X_j | > n^{d/\alpha} a_{m+1}] \\ &= (n^{d/\alpha} a_{m+1})^{-\alpha} (c_1 + g(n^{d/\alpha} a_{m+1})) \end{aligned} \tag{5.22}$$

where $g(x) = o(1)$ as $x \to +\infty$. From this it follows that for a sufficiently large constant c_2 we have $f_{n,m} \leq c_2(n^{d/\alpha}a_{m+1})^{-\alpha}\beta^\alpha$ for all n and m. Therefore

$$n^d a_m^2 f_{n,m} \leq c_2 a_m^2 a_{m+1}^{-\alpha}\beta^\alpha \tag{5.23}$$

for all n and m.

Let us choose $a_m = a_0\beta^{am}$ for some $a > 0$ and $a_0 > 0$. Then $a_m/a_{m+1} = \beta^{-1}$ and so

$$n^d a_m^2 f_{n,m} \leq c_2 a_m^{2-\alpha}. \tag{5.24}$$

The inequality (5.21) will then be satisfied if

$$H(\delta_m) \leq \ln \varepsilon_m + \frac{\eta_m^2}{4c_2 a_m^{2-\alpha}} \wedge \frac{3\eta_m}{4a_m},\ m \geq m_0. \tag{5.25}$$

This inequality becomes (5.41) below.

In a similar way, to study $P_2(n, m)$, obtain from Bernstein's inequality that

$$P[T_{21} > \eta_m] \leq \textstyle\sum_{j=m}^{k_m-1} 2\exp\left\{2H(\delta_{j+1}) - \dfrac{\eta_{m,j}^2}{2(n^d a_m^2 f_{n,m}\delta_j + \eta_{m,j}\alpha_m/3)}\right\} \tag{5.26}$$

where we will choose the $\eta_{m,j}$ so that $\sum_{j=m}^{k_m-1} \eta_{m,j} \leq \eta_m$ and where the following bound on the variance is used:

$$\begin{aligned} \operatorname{var}(Y_{n,m}(A)) &= n^{-2d/\alpha} \textstyle\sum_j \{|C_j \cap nA|\, EX_{j,n,m}^2 - |C_j \cap nA|^2 EX_{j,n,m}^2\} \\ &\leq a_m^2 f_{n,m} |nA| = n^d a_m^2 f_{n,m} |A|. \end{aligned} \tag{5.27}$$

To obtain the desired summability, it will suffice to have

$$2H(\delta_{j+1}) \leq \ln \varepsilon_{m,j} + \frac{\eta_{m,j}^2}{4n^d a_m^2 f_{n,m}\delta_j} \wedge \frac{3\eta_{m,j}}{4\alpha_m} \tag{5.28}$$

for $m \leq j \leq k_m$ and $m \geq m_0$, provided $\sum_{j=m}^{k_m-1} \varepsilon_{m,j} \leq \varepsilon_m$ for each m and n. As was done for (5.21), an application of (5.23) permits the above to be simplified to

$$2H(\delta_{j+1}) \leq \ln \varepsilon_{m,j} + \frac{\eta_{m,j}^2}{4c_2 a_m^{2-\alpha}\delta_j} \wedge \frac{3\eta_{m,j}}{4a_m}. \tag{5.29}$$

To handle the term T_{22} we use the bound

$$|Y_{n,m}(B\backslash A)| \leq n^{-d/\alpha} \textstyle\sum_j 1_{n(A^+\backslash A)}(U_j)\, |X_{j,n,m}| \tag{5.30}$$

when $A \subset B \subset A^+$ and $A, A^+ \in \mathscr{A}_{\delta_{k_m}}$. To apply Bernstein's inequality in these cases, we need to center the sums at their expectations and compute the corresponding variances. But here the mean and variance of the right-hand side of (5.30) are, respectively,

$$n^d |A^+\backslash A|\, n^{-d/\alpha} E|X_{j,n,m}| \leq \delta_{k_m} n^d a_m f_{n,m} \tag{5.31}$$

and

$$\begin{aligned} n^{-2d/\alpha} \textstyle\sum_j\{|C_j \cap n(A^+\backslash A)|\, E(X_{j,n,m})^2 - |C_j \cap n(A^+\backslash A)|^2 (E|X_{j,n,m}|)^2\} \\ \leq \delta_{k_m} n^d a_m^2 f_{n,m}. \end{aligned} \tag{5.32}$$

Thus to apply Bernstein's inequality, we need after centering to have

$$\eta_{m,k_m} - \delta_{k_m} n^d a_m f_{n,m} \tag{5.33}$$

remain positive. Write $\delta_m^* = \delta_{k_m}$ and $\eta_m^* = \eta_{m,k_m}$. We will ask that the term in (5.33) exceed $\eta_m^*/2$; that is

$$\delta_m^* n^d a_m f_{n,m} \le \eta_m^*/2. \tag{5.34}$$

In view of (5.23), (5.34) would be satisfied if

$$c_2 \delta_m^* a_m^{1-\alpha} \le \eta_m^*/2 \tag{5.35}$$

for $m \ge m_0$.

Subject to (5.34) being satisfied, we may apply Bernstein's inequality to obtain

$$P[T_{22} > \eta_m^*] \le 2 \exp\left\{2H(\delta_m^*) - \frac{(\eta_m^*/2)^2}{2(\delta_m^* n^d a_m^2 f_{n,m} + \eta_m^* a_m/6)}\right\} \tag{5.36}$$

We will therefore aim to satisfy the inequality

$$2H(\delta_m^*) \le \ln \varepsilon_m + \frac{(\eta_m^*)^2}{16\delta_m^* n^d a_m^2 f_{n,m}} \wedge \frac{3\eta_m^*}{4a_m}. \tag{5.37}$$

Again, (5.23) permits the replacement of this inequality by

$$2H(\delta_m^*) \le \ln \varepsilon_m + \frac{(\eta_m^*)^2}{16\delta_m^* c_2 a_m^{2-\alpha}} \wedge \frac{3\eta_m^*}{8a_m};\ m \ge m_0. \tag{5.38}$$

To show that it is possible to satisfy all of the above inequalities, it suffices to choose

$$\delta_m = \delta_0 \beta^m, \quad a_m = a_0 \beta^{am} \tag{5.39}$$

$$\eta_m = \eta_0 \beta^{bm}, \quad \eta_{m,j} = \eta_j(1 - \beta^b). \tag{5.40}$$

The key inequalities (5.25), (5.29), (5.35) and (5.38) can then be rewritten for $m \ge m_0$ as, respectively.

$$K\delta_0^{-r}\beta^{-rm} \le \ln \varepsilon_m + \left(\frac{\eta_0^2}{4c_2 a_0^{2-\alpha}}\right)\beta^{-((2-\alpha)a-2b)m} \wedge \left(\frac{3\eta_0}{4a_0}\right)\beta^{-(a-b)m}; \tag{5.41}$$

$$2K\delta_0^{-r}\beta^{-r(j+1)} \le \ln \varepsilon_{m,j} + \left\{\frac{\eta_0^2}{4c_2 a_0^{2-\alpha}\delta_0}\right\}\beta^{-(2-\alpha)am-(1-2b)j} \wedge \left(\frac{3\eta_0}{4a_0}\right)\beta^{-am+bj}, \quad m \le j < k_m; \tag{5.42}$$

$$\delta_0 \beta^{-a\alpha m} \le \left\{\frac{\eta_0}{2c_2 a_0^{1-\alpha}}\right\}\beta^{(b-1)k_m}; \tag{5.43}$$

$$2K\delta_0^{-r}\beta^{-rk_m} \le \ln \varepsilon_m + \left\{\frac{\eta_0^2}{16\delta_0 c_2 a_0^{2-\alpha}}\right\}\beta^{-(1-2b)k_m-(2-\alpha)am} \wedge \left(\frac{3\eta_0}{8a_0}\right)\beta^{-am+bk_m}. \tag{5.44}$$

Since $r < (\alpha - 1)^{-1}$, one can choose $b > 0$ sufficiently small so that $\alpha - 1 < (1-b)/(r+b)^{-1}$ or equivalently,

$$r(\alpha - 1) < (1-b)/(1+b/r) < 1.$$

Moreover, one can also choose b sufficiently small to insure

$$r(\alpha - 1) < 1 - \alpha b < 1,$$

which insures that

$$(2 - \alpha)/(r - 1 + 2b) > 1/(r + b).$$

Let $a \geq (r + 2b)/(2 - \alpha)$, which in turn is larger than $r + b$. Then let k_m be the first integer strictly larger than $(\alpha - 1)am/(1 - b)$ for m sufficiently large.

Since $0 < \beta < 1$, straightforward algebra shows that we have

(5.45) $$r \leq (2 - \alpha)a - 2b \quad \text{and} \quad r \leq a - b;$$

(5.46) $$(r - 1 + 2b)j \leq (2 - \alpha)am \quad \text{and} \quad (r + b)j \leq am, \quad \text{for} \quad m \leq j < k_m;$$

(5.47) $$k_m \geq (\alpha - 1)am/(1 - b);$$

(5.48) $$k_m \leq (2 - \alpha)am/(r - 1 + 2b) \quad \text{and} \quad k_m \leq am/(r + b),$$

The choice of $\varepsilon_m = \varepsilon_0 m^{-3}$ and $\varepsilon_{m,j} = \varepsilon_m$ suffices to provide a suitable convergent series for which, by an appropriate choice of ε_0, $\sum_m \varepsilon_m \leq \varepsilon/3$ and $\sum_m \sum_{j=m}^{k_m} \varepsilon_{m,j} \leq \sum_m k_m \varepsilon_m \leq \varepsilon/3$.

Choose η_0 so that $\sum_{m=1}^{\infty} \eta_m \leq \eta$ and choose a_0 and $\delta_0 \leq 1$ arbitrarily. Then, provided m_0 is sufficiently large, (5.41), (5.42), (5.43), and (5.44) will be satisfied by examining the appropriate powers of β and taking into account (5.45), (5.46), (5.47), and (5.48), respectively.

The only place in the above where we use the symmetry of $X_{\mathbf{j}}$ is to insure that $EX_{\mathbf{j},n,m} = 0$, which was needed in the use of Bernstein's inequality. In the general case, we instead apply Bernstein's inequality to the sums $\hat{Y}_{n,m}(A) = Y_{n,m}(A) - y_{n,m}(A)$. Since

$$|Ey_{n,m}(A)| \leq n^{-d/\alpha} \textstyle\sum_{\mathbf{j}} |(\eta A) \cap C_{\mathbf{j}}| \; E\,|X_{\mathbf{j},n,m}| \leq n^d a_m f_{n,m} |A|$$

for $A \in \mathscr{A}$, then

$$\sup_{A \subseteq B \subseteq A^+} |EY_{n,m}(B \backslash A)| \leq n^d a_m f_{n,m} |A^+ \backslash A|.$$

A comparison with (5.30) and (5.31) shows that only minor modifications are required in (5.34) and (5.37). With these modifications, the proof proceeds as above. □

We can now prove

THEOREM 5.3. *With $X_{\mathbf{j}}$, Y_n, Z, $\mathscr{A}$ as above, Y_n converges to Z weakly with respect to d_D.*

PROOF. Since the finite-dimensional distributions converge (Proposition 5.1), by Theorem 4.3 it suffices to show tightness of the Y_n. We do that by means of Theorem 3.4. Given $\Delta_m \to 0$, for each m, let $C_m(Y_n) = 0$. Choose γ_m sufficiently small so that $P[\| Y_n^{(\gamma_m)} - y_n^{(\gamma_m)} \|_{\mathscr{A}} \geq \Delta_m] \leq \varepsilon/2^{m+1}$ for all n. This can be done by Proposition 5.2. Let $J_m(Y_n) = Y_n - (Y_n^{(\gamma_m)} - y_n^{(\gamma_m)})$.

Fix m, and let $T_n = J_m(Y_n)$. Since $|y_n^{(\gamma_m)}(C_{n\mathbf{j}})| \leq \gamma_m$, T_n has an atom in $C_{n\mathbf{j}} = n^{-1}(\mathbf{j} - \mathbf{1}, \mathbf{j}]$ only if $X_\mathbf{j}$ is larger than $\gamma_m(n^{d/\alpha} - 1)$ in absolute value. By (5.22), the probability of this happening is $O(n^{-d})$. Thus (4.4)(i, ii, and iii) follow easily, and so by Theorem 4.4, we can find η_m, M_m, N_m, and h_m such that

$$P(J_m(Y_n) \notin \mathscr{F}_{PA}(h_m, N_m, \eta_m, M_m)) \le \varepsilon/2^{m+1}.$$

We can therefore apply Theorems 3.4 and 4.3, and we are done. □

Let us now discuss the cases $\alpha \in (0, 1]$. If $X_\mathbf{j}$ is in the domain of normal attraction of a stable law of index α, $\alpha = 1$, it is necessary first to center the Y_n processes appropriately (cf. Gnedenko and Kolmogorov, 1954, page 175) and then use techniques similar to those employed in Proposition 5.2. When $\alpha < 1$, the methods needed are considerably easier, since

$$\| Y_n^{(\gamma)} \|_{\mathscr{A}} \le n^{-d/\alpha} \textstyle\sum_\mathbf{j} | X_\mathbf{j} | 1_{[|X_\mathbf{j}| \le \gamma n^{d/\alpha}]},$$

and so

$$\begin{aligned} E \| Y_n^{(\gamma)} \|_{\mathscr{A}} &\le n^{d-d/\alpha} E | X_\mathbf{j} | 1_{[|X_\mathbf{j}| \le \gamma n^{d/\alpha}]} \\ &\le n^{d-d/\alpha} \int_0^{\gamma n^{d/\alpha}} P[|X_\mathbf{j}| > x]\, dx \\ &\le c_2 n^{d-d/\alpha} \int_0^{\gamma n^{d/\alpha}} x^{-\alpha}\, dx \\ &= c_2(1-\alpha)^{-1}\gamma^{1-\alpha} \to 0 \quad \text{as} \quad \gamma \to 0. \end{aligned}$$

6. Remarks. 1. Consider the one-dimensional situation in which $\mathscr{A} = \{[0, t]: 0 \le t \le 1\}$, the family of right closed intervals. Identifying $Z(t)$ with $Z([0, t])$, we are in the well-known case of processes whose sample paths are right continuous with left limits, that is, processes with paths in $D[0, 1]$. The results of Section 5 then provide a central limit theorem for $D[0, 1]$, where, however, weak convergence is stated with respect to Skorokhod's metric M_2 rather than with respect to the more usual metric J_1. To see that, in fact, our approach also yields the stronger result, recall that the way the central limit theorem is proved is to show both the convergence of the finite-dimensional distributions and tightness, and recall that the way tightness is proved is to show that the hypotheses of Theorem 3.4 are satisfied. However, observe that once the hypotheses of Theorem 3.4 are satisfied, we have, with high probability, a bound on how close together the jumps of S_n that exceed a given size can be. By the tightness criteria for J_1 (cf. Billingsley, 1968, page 116), we see then that we also have, in this case, tightness with respect to J_1, and hence a central limit theorem with respect to J_1 as well.

2. Consider a general array $X_{n\mathbf{j}}$: $\mathbf{j} \in J^d$, $\mathbf{j} \le \mathbf{k}_n\}$ of independent infinitesimal r.v.'s. In this case we view $X_{n\mathbf{j}}$ as a random mass placed at $\mathbf{j}_{(n)} := (j_1/k_{n1}, \cdots, j_d/k_{nd})$. Assume $\mathbf{k}_n$ approaches infinity in the sense that $k_{ni} \to +\infty$ for each $1 \le i \le d$. We use our same notation for the resulting unsmoothed and smoothed

partial-sum processes, namely

$$S_n(A) = \sum_{\mathbf{j}_{(n)} \in A} X_{n\mathbf{j}}, \; Y_n(A) = \sum_{U_{n\mathbf{j}} \in \mathbf{A}} X_{n\mathbf{j}} \tag{6.1}$$

where now the normalizing constants are included in the $\{X_{n\mathbf{j}}\}$, and the $\{U_{n\mathbf{j}}\}$ are independent r.v.'s independent of the $\{X_{n\mathbf{j}}\}$ with $U_{n\mathbf{j}}$ being uniform over the interval $C_{n\mathbf{j}} = (\mathbf{j}_{(n)} - \mathbf{1}_{(n)}, \mathbf{j}_{(n)}]$. Now define for any real Borel set B and any $A \in \mathscr{A}$,

$$\nu_n(B, A) = \sum_{\mathbf{j}_{(n)} \in A} P[X_{n\mathbf{j}} \in B]. \tag{6.2}$$

This is then the expected number of atoms located in A that have masses with magnitudes in B. Assume that on $(-\infty, -x] \cup [x, \infty)$ for any $x > 0$, the measures $\nu_n(\cdot A)$ converge weakly to $\nu(\cdot, A)$, the limiting Lévy measure, uniformly over $A \in \mathscr{A}$. For purposes of this discussion, let us assume further that we have a homogeneous case in which $\nu(\cdot, A)$ is of the form $|A|\, \nu(\cdot)$ for some Lévy measure ν.

If $X_{n\mathbf{j}}^{(\tau)}$ denotes the truncation of $X_{n\mathbf{j}}$ at τ and if $S_n^{(\tau)}$ and $Y_n^{(\tau)}$ denote, respectively, the unsmoothed and smoothed partial-sum processes formed from the truncated array $\{X_{n\mathbf{j}}^{(\tau)}\}$, then the necessary and sufficient conditions for the classical case (cf. Gnedenko and Kolmogorov, 1954, page 124) could be used to obtain conditions for the weak convergence of the finite-dimensional distributions of the S_n-processes. These conditions would be expressed in terms of $\{\nu_n\}$ and the means and variances for the truncated arrays, namely

$$\mu_n(\tau, A) := ES_n^{(\tau)}(A), \quad \sigma_n^2(\tau, A) := \operatorname{var} S_n^{(\tau)}(A). \tag{6.3}$$

Such conditions could be described as requiring that all subarrays cut out by $A \in \mathscr{A}$ are in the domain of attraction of the appropriate infinitely divisible distributions. By contrast, to obtain conditions for the weak convergence of the finite-dimensional distributions of the smoothed Y_n-processes, it may well be necessary to introduce an assumption about the smoothness of the boundaries of the sets in $\mathscr{A}$.

To complete a central limit theorem for these general arrays one needs to verify the required tightness. A key assumption that may be needed for our arguments to carry through in this case is the uniform domination of the ν_n-measures by the Lévy measures; e.g.,

$$\nu_n(B, A) \leq c_x \nu(B, A) \tag{6.4}$$

for some constant c_x and all $A \in \mathscr{A}$ and $B \subset (-\infty, -x] \cup [x, \infty)$. For the case studied in this paper (that is, i.i.d. arrays in the domain of normal attraction of a stable law,) this property is known to hold; cf. (5.22) above.

It would be very interesting to obtain a central limit theorem for general arrays. We expect one must assume (6.4), sufficient smoothness of the boundaries of $A \in \mathscr{A}$, and bounds on $H(\delta)$, the log-entropy of $\mathscr{A}$. The last condition is necessary to ensure that the limiting Lévy process exists in $\mathscr{D}(\mathscr{A})$ (cf. Bass and Pyke, 1984b).

3. Theorem 3.4 characterizes many of the compact subsets of $\mathscr{D}_0(\mathscr{A})$. It would

be interesting and useful to have a complete characterization of the compact subsets of $\mathscr{D}(\mathscr{A})$. For example, consider a set-indexed process Z where, for each ω, $Z(\omega)$ is the set function that is generated by surface measure of a ball whose center and radius depends on ω. Such a process belongs to $\mathscr{D}(\mathscr{A})$ but not $\mathscr{D}_0(\mathscr{A})$. Of course, much more complicated and interesting examples might arise. To study the weak convergence of such processes, one would first like a suitable criterion for tightness, hence a characterization of compact sets.

4. Recall that our metric is related to Skorokhod's M_2 topology. Of the topologies for $D[0, 1]$, most work since 1956 has concentrated instead on Skorokhod's J_1 topology. There are inherent difficulties in extending the latter to $\mathscr{D}(\mathscr{A})$. The definition of J_1 requires the existence of a group Λ of homeomorphisms on the index set $[0, 1]$, and in this aspect has been extended to very general index sets by Straf (1972). However, a simple example shows why straightforward extensions of J_1 are not suitable for general spaces of set-indexed functions. Let $\mathscr{A}$ be the family of closed convex subsets of I^d, and define functions $x_n\colon \mathscr{A} \to \mathbb{R}$ by $x_n(A) = 1$ if $\mathbf{t}_n \in A$ and 0 otherwise; $n = 0, 1, 2, \cdots$, where $\mathbf{t}_n \in I_d$ is a fixed sequence which converges to a point $\mathbf{t}_0$ as $n \to \infty$. Clearly one requires a topology under which x_n converges to x_0; but can one construct homeomorphisms λ_n on $\mathscr{A}$ so that $x_n \circ \lambda_n = x_0$? Perhaps such λ_n exist, but their construction would not be simple. Of course, $\mathscr{A}$ and x_n can be considerably more complicated.

A more appealing approach might be to consider homeomorphisms λ on I^d itself. But, in general, λ will not map $\mathscr{A}$ to $\mathscr{A}$ and so $x \circ \lambda(A)$ may not make sense. Enlarging $\mathscr{A}$ to $\mathscr{A}^* = \{\lambda A\colon \lambda \in \Lambda, A \in \mathscr{A}\}$ generally results in a class of sets that is too large to support the processes under consideration. It would be extremely interesting to see if a suitable extension of J_1 to $\mathscr{D}(\mathscr{A})$ exists.

5. A further application of our results is to empirical processes. Although the limit processes of empirical processes are usually continuous, the empirical processes themselves are not, and there are considerable difficulties involved with the nonmeasurability of certain necessary random quantities. One possible way of approaching the questions of nonmeasurability is to observe that empirical processes have paths in $\mathscr{D}_0(\mathscr{A})$ and to use the topology we introduce here. Measurability follows from Proposition 4.1. The relation to uniform convergence is given by Theorem 3.5.

It would be worthwhile to compare this approach with those of others, such as in Dudley (1978) and Dudley and Philipp (1983); for other references, see Giné and Zinn (1984). In particular, is the class of central limit problems the same for each approach?

6. Our central limit theorem of Section 5 requires the log-entropy H to satisfy $H(\delta) \le K\delta^{-r}$. $1 < r < (\alpha - 1)^{-1}$. As discussed there, there is no loss of generality in taking $r > 1$. By an example of Adler and Feigin (1984), no limit process can exist if $r > (\alpha - 1)^{-1}$.

This leaves the case $r = (\alpha - 1)^{-1}$. By Bass and Pyke (1984b), there exists a set-indexed stable process whose paths are outer continuous with inner limits if

H satisfies

$$\int_0^1 (H(x)/x)^{1-1/\alpha}\, dx < \infty. \tag{6.5}$$

We would expect that a more refined truncation procedure, such as the one in Bass (1985), would allow one to prove a central limit theorem for H satisfying (6.5).

7. In this paper we have proved a uniform central limit theorem for smoothed versions of partial-sum processes for which the random atoms are at fixed locations, namely at the points of a regular lattice. A central limit theorem may also be obtained for sums of i.i.d. $\mathscr{D}_0(\mathscr{A})$-valued processes and this is to be included in a forthcoming paper by the authors. This provides a central limit theorem for the cases of random masses at random locations, generalizing results for empirical processes to the case of nonnormal limits. By contrast, recall that an empirical process involves nonrandom constant masses at random locations.

Acknowledgment. The authors greatly appreciate the thoughtful reading and helpful comments of the referee and the Associate Editor.

REFERENCES

ADLER, R. and FEIGIN, P. (1984). On the cadlaguity of random measures. *Ann. Probab.* **12** 615–630.

ALEXANDER, K. S. and PYKE, R. (1985). A uniform central limit theorem for set-indexed partial-sum processes with finite variance. *Ann. Probab.* To appear.

BASS, R. F. (1985). Law of the iterated logarithm for set-indexed partial sum processes with finite variance. *Z. Wahrsch. verw. Gebiete.* To appear.

BASS, R. F. and PYKE, R. (1984a). Functional law of the iterated logarithm and uniform central limit theorem for partial-sum processes indexed by sets. *Ann. Probab.* **12** 13–34.

BASS, R. F. and PYKE, R. (1984b). The existence of set-indexed Lévy processes. *Z. Wahrsch. verw. Gebiete* **66** 157–172.

BICKEL, P. J. and WICHURA, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42** 1656–1670.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

BILLINGSLEY, P. (1971). Weak Convergence of Measures: Applications in Probability. *Conference Board for the Mathematical Sciences Regional Conference Series in Applied Mathematics* **5**. Soc. Ind. Appl. Math., Philadelphia.

DEBREU, G. (1967). Integration of correspondences. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **2** 351–372. Univ. of California Press.

DONSKER, M. (1951). An invariance principle for certain probability limit theorems. *Four papers on Probability. Mem. Amer. Math. Soc.* **6** 1–12.

DUDLEY, R. M. (1973). Sample functions of the Gaussian process. *Ann. Probab.* **1** 66–103.

DUDLEY, R. M. (1974). Metric entropy of some classes of sets with differentiable boundaries. *J. Approx. Theory* **10** 227–236.

DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929.

DUDLEY, R. M. and PHILIPP, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. verw Gebiete* **62** 509–552.

GIKHMAN, I. I. and SKOROKHOD, A. V. (1969). *Introduction to the Theory of Random Processes.* Saunders, Philadelphia.

GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–989.

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley, Cambridge, Mass.

JACOD, J., KŁOPOTOWSKI, A., and MEMIN, J. (1982). Théorème de la limite centrale et convergence fonctionelle vers un processus a accroissements indépendants: la methode des martingales. *Ann. Inst. H. Poincare* **18** 1–46.

KOLMOGOROV, A. N. (1956). On Skorokhod convergence. *Theory Probab. Appl.* **1** 215–222.

LINDVALL, T. (1973). Weak convergence of probability measures and random functions on the function space $D(0, \infty)$. *J. Appl. Probab.* **10** 109–121.

NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

POLLARD, D. (1979). Weak convergence on non-separable metric spaces. *J. Austral. Math. Soc. Ser. A* **28** 197–204.

PROHOROV, YU. V. (1953). Probability distributions in functional spaces. *Uspehi Mat. Nauk (N.S.)* **8** 165–167.

PROHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 177–238.

PYKE, R. (1973). Partial sums of matrix arrays, and Brownian sheets. In *Stochastic Analysis.* (D. G. Kendall and E. F. Harding, eds.) 331–348. Wiley, London.

PYKE, R. (1983). A uniform central limit theorem for partial-sum processes indexed by sets. In *Probability, Statistics, and Analysis.* (J.F.C. Kingman and G.E.H. Reuter, eds.) 219–240. Cambridge Univ. Press.

RÉVÉSZ, P. (1976). Three theorems of multivariate empirical process. *Lecture Notes in Math.* **566** 106–126, Springer-Verlag, Berlin.

SKOROKHOD, A. V. (1955). On the limiting transition from a sequence of sums of independent random quantities to a homogeneous random process with independent increments. *Dokl. Akad. Nauk SSSR (N.S.)* **104** 364–367.

SKOROKHOD, A. V. (1956a). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 289–319.

SKOROKHOD, A. V. (1956b). On a class of limit theorems for Markoff chains. *Dokl. Akad. Nauk SSSR (N.S.)* **106** 781–784.

SKOROKHOD, A. V. (1957). Limit theorems for stochastic processes with independent increments. *Theory Probab. Appl.* **2** 145–177.

STRAF, M. L. (1972). Weak convergence of stochastic processes with several parameters. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 187–221. Univ. of California Press.

VERVAAT, W. (1981). Une compactification des espaces fonctionnels C et D; une alternative pour la démonstration de théorèmes limites fonctionnels. *C. R. Acad. Sci. Paris* **292,** 441–444.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195