WEAK CONVERGENCE OF CONDITIONED SUMS
OF INDEPENDENT RANDOM VECTORS

BY

THOMAS M. LIGGETT

Abstract. Conditions are given for the weak convergence of processes of the form $(X_n(t) \mid X_n(1) \in E^n)$ to tied-down stable processes, where $X_n(t)$ is constructed from normalized partial sums of independent and identically distributed random vectors which are in the domain of attraction of a multidimensional stable law. The conditioning events are defined in terms of subsets E^n of R^d which converge in an appropriate sense to a set of measure zero. Assumptions which the sets E^n must satisfy include that they can be expressed as disjoint unions of "asymptotically convex" sets. The assumptions are seen to hold automatically in the special case in which E^n is taken to be a "natural" neighborhood of a smooth compact hypersurface in R^d.

1. **Introduction and notation.** Empirical distribution functions have been widely studied in probability theory and statistics. Often, the fact that they may be represented as conditioned sums of independent random variables has played an important role in these investigations. This has led to an interest in the behavior of these sums under various forms of conditioning. It is the purpose of this paper to present conditions under which certain stochastic processes, obtained from the partial sums $\{S_k, k \le n\}$ of independent identically distributed random vectors in the domain of attraction of a stable law by conditioning on information concerning S_n, converge to a limiting process.

The following, for example, is a consequence of the main theorems of this paper if the basic random vectors are nonlattice and require no centering constants: When properly normalized and conditioned on the event $\{b_n - \gamma_n \le \|S_n\| \le b_n + \gamma_n\}$, the partial sums $(S_k, k \le n)$ converge to a stable process which is tied down at time one to the surface of the unit sphere in R^d. Here $\{b_n\}$ is the usual sequence of normalizing constants and $\{\gamma_n\}$ is any sequence of positive numbers which is bounded away from zero and is $o(b_n)$. If, instead, we condition on the event $\{\|S_n\| \le \gamma_n\}$, the limiting process is tied down to the origin at time one.

The limiting process will be a multidimensional stable process $X(t)$ which is conditioned, in a certain sense, on the event that $X(1)$ lies in a set E. In all cases of interest, E will be a set of Lebesgue measure zero in R^d, so there is no unique

---

natural way of defining the conditioned process. Roughly speaking, the non-uniqueness in this definition is a consequence of the fact that given that $X(1)$ is in E, one could assign different weights to the event that $X(1)$ lies in one part of E as opposed to another part of E. It is for this reason that for a fixed stable process X, the processes which occur as limits of the conditioned approximating processes are parametrized not by the set E on which conditioning takes place in the limit, but by measures μ on R^d. In a sense, $\mu(d\mathbf{x})$ is the relative importance one gives to the possibility that $X(1)$ is in $d\mathbf{x}$ when given that $X(1)$ is in E, the support of μ.

If the approximating processes constructed from the partial sums are denoted by $X_n(t)$, the conditioning will usually be on the event $\{X_n(1) \in E^n\}$ for an appropriate choice of sets E^n in R^d which decrease to E. Most of the assumptions to be made in the main theorem (Theorem 4) concern these sets E^n. These conditions are of a somewhat involved and seemingly technical nature. It is seen in §4, however, that they are satisfied in many cases, and that they can often be verified quite easily. In fact, the choice of this particular form of the conditions on the sets E^n was motivated by the situation considered in §4, where E is taken to be a smooth hypersurface in R^d and E^n is taken to be a sequence of "natural" neighborhoods of E.

In an earlier paper [8], the author investigated the question of weak convergence of these conditioned processes for the case in which $d=1$, E is a singleton in R^1, and $\{E^n\}$ is a decreasing sequence of intervals about E. The present emphasis is on weak convergence to stable processes which are tied down at time one to more general sets in R^d, such as lower dimensional hypersurfaces. Wichura [13] has obtained some related results. His treatment, insofar as it relates to sums of independent random vectors, deals with the case in which the random vectors are in the domain of attraction of a Gaussian law, and are lattice distributed or have a continuous density. Furthermore, the sets E^n and E are singletons, in which case there is a unique natural way of defining the conditioned processes.

The main results concern a sequence $\{\xi_n\}$ of independent, identically distributed, d-dimensional random vectors with common characteristic function $g(\mathbf{s})$. The distribution of the vectors ξ_n is assumed to be nondegenerate and normalized in the sense of Stone [10]. That is, the distribution of ξ_n does not have its support in any $(d-1)$-dimensional hyperplane, and "there is an integer d_1, $0 \leq d_1 \leq d$, and there are real numbers $\alpha^{(1)}, \ldots, \alpha^{(d_1)}$ such that (i) if (for some k) $s^{(k)} \neq 0$ and either $d_1 < k \leq d$ or $|s^{(k)}| < 2\pi$, then $|g(\mathbf{s})| < 1$; and (ii) if (for some k) $1 \leq k \leq d_1$, $s^{(k)} = 2\pi$, and $s^{(j)} = 0$ for $j \neq k$, then $g(\mathbf{s}) = \exp(2\pi i \alpha^{(k)})$." In other words, the coordinate system of R^d is chosen so that ξ_n has a lattice distribution of span one in the first d_1 directions, and a nonlattice distribution in the remaining directions. Letting S_n denote the partial sum $\xi_1 + \cdots + \xi_n$, we define the support lattice of S_n by

$$D_n = \{\mathbf{x} \in R^d \mid x^{(k)} - n\alpha^{(k)} \text{ is an integer for } 1 \leq k \leq d_1\}.$$

We assume that the distribution of ξ_n is in the domain of attraction of a non-

degenerate stable law. Specifically, there are normalizing constants $b_n > 0$ and $\mathbf{a}_n$ so that $(\mathbf{S}_n - \mathbf{a}_n)/b_n$ converges in distribution to a random vector which will be called $\mathbf{X}(1)$. The support lattice of $(\mathbf{S}_n - \mathbf{a}_n)/b_n$ is $D_n^* = (D_n - \mathbf{a}_n)/b_n$. It is known [7] that the characteristic function of the nondegenerate d-dimensional stable random vector $\mathbf{X}(1)$ has the form

$$f(\mathbf{s}) = \exp \{i\mathbf{b} \cdot \mathbf{s} - c\omega_1(\mathbf{s}, \alpha) - ic\omega_2(\mathbf{s}, \alpha)\},$$

where

$$\omega_1(\mathbf{s}, \alpha) = \int |\boldsymbol{\theta} \cdot \mathbf{s}|^\alpha \, dH(\boldsymbol{\theta}),$$

$$\omega_2(\mathbf{s}, \alpha) = -\tan \frac{\pi \alpha}{2} \int \text{sign} (\boldsymbol{\theta} \cdot \mathbf{s})|\boldsymbol{\theta} \cdot \mathbf{s}|^\alpha \, dH(\boldsymbol{\theta}) \quad \text{if } \alpha \neq 1,$$

$$= \frac{2}{\pi} \int (\boldsymbol{\theta} \cdot \mathbf{s}) \log |\boldsymbol{\theta} \cdot \mathbf{s}| \, dH(\boldsymbol{\theta}) \qquad \text{if } \alpha = 1,$$

$\mathbf{b}$ is a constant vector in R^d, c is a positive number, α is in $(0, 2]$, and $dH(\boldsymbol{\theta})$ is a probability measure on the surface of the unit sphere in R^d which does not concentrate on any lower dimensional subspace of R^d. Since $dH(\boldsymbol{\theta})$ does not concentrate on a lower dimensional subspace, $|f(\mathbf{s})| \leq \exp \{-\varepsilon \|\mathbf{s}\|^\alpha\}$ for some $\varepsilon > 0$. So, $f \in L_1(R^d)$, and it follows that $\mathbf{X}(1)$ has a density $p(\mathbf{x})$ which is continuous and tends to zero as $\|\mathbf{x}\| \to \infty$. Here $\|\cdot\|$ refers to the ordinary norm in R^d. Let

$$Q = \{\mathbf{x} \in R^d \mid p(\mathbf{x}) > 0\}.$$

Weak convergence will be considered in the space $D^d[0, 1]$, which is the topological product of d copies of the space $D[0, 1]$. $D[0, 1]$ is the space of real-valued functions on $[0, 1]$ which are right continuous and have no discontinuities of the second kind. There is a metric in $D[0, 1]$ which makes it into a complete separable metric space whose relative topology on the subspace of continuous functions is the same as the topology of uniform convergence. A description of this metric and of various results concerning weak convergence in $D[0, 1]$ may be found in Chapter 3 of Billingsley [1].

The basic processes to be considered are defined as follows. For each n, let $\mathbf{X}_n$ be the random element of $D^d[0, 1]$ given by

(1)
$$\mathbf{X}_n(t) = \frac{\mathbf{S}_{[(n+1)t]} - ([(n+1)t]/n)\mathbf{a}_n}{b_n} \quad \text{for } 0 \leq t < 1,$$

$$= \frac{\mathbf{S}_n - \mathbf{a}_n}{b_n} \qquad \qquad \text{for } t = 1,$$

where $[\cdot]$ is the greatest integer function.

2. **Preliminary results.** As in the one-dimensional case, the main theorems for the conditioned processes follow from two basic results. The first of these is the weak convergence of the processes $\mathbf{X}_n$, which is stated as Theorem 1 below, while

the second is the local limit theorem for the partial sums S_n, which appears as Theorem 2. We will need the function $\eta(t)$ which is defined by

$$\eta(t) = \mathbf{b}(t - t^{1/\alpha}) \qquad\qquad \text{if } \alpha \neq 1,$$

$$= \frac{2c}{\pi} t \log t \int \boldsymbol{\theta} \, dH(\boldsymbol{\theta}) \quad \text{if } \alpha = 1,$$

where the quantities $\mathbf{b}$, c, and $dH(\boldsymbol{\theta})$ are those appearing in the expression for the characteristic function of $\mathbf{X}(1)$ in §1.

THEOREM 1. $\mathbf{X}_n$ *converges weakly to* $\mathbf{X}$ *in* $D^d[0, 1]$, *where* $\mathbf{X}(t)$ *is a temporally homogeneous process with independent increments which has no fixed times of discontinuity and whose one-dimensional distributions are the same as those of* $\eta(t) + t^{1/\alpha}\mathbf{X}(1)$.

The proof of this theorem follows that of Theorem 6 of [5], which is exactly the above result for the case $\alpha = 2$. Hence it will not be reproduced here. The only change in the argument is that Theorem 1 of [8] is used in place of Donsker's theorem. For the details of the proof, see Lemma 3.6 of [9, p. 23].

To obtain the general results of §3, we will need versions of the local limit theorem which deal with the probability that $\mathbf{X}_n(1)$ is in a small convex set and, in fact, that it is in a small set which is nearly convex in a sense which will be made precise later. Since the local limit theorem which is available (Theorem 1 of [10]) applies only to $(d - d_1)$-dimensional cubes, it is necessary to prove the following lemma, which shows that uniformly bounded convex sets can be approximated uniformly by unions of cubes. Without the uniformity statement, the proof of the lemma would be much simpler. However, this uniformity will be necessary in the next section.

In order to obtain the required uniformity, we will introduce the standard metric topology on the class of nonempty compact convex subsets of R^k. For this purpose, if $\delta > 0$ and A is any subset of R^k, define the δ-neighborhood of A by

$$(2) \qquad\qquad N(A; \delta) = \left\{ \mathbf{x} \in R^k \,\middle|\, \inf_{y \in A} \|\mathbf{x} - \mathbf{y}\| < \delta \right\}.$$

Then, if C_1 and C_2 are nonempty compact convex subsets of R^k, the distance $\Delta(C_1, C_2)$ between C_1 and C_2 is defined to be

$$\inf \{\delta > 0 \mid C_1 \subset N(C_2; \delta)\} + \inf \{\delta > 0 \mid C_2 \subset N(C_1; \delta)\}.$$

It is known that this is, in fact, a metric [4, p. 60]. Furthermore, by the Blaschke Selection Theorem [4, Theorem 32], the class of uniformly bounded compact convex subsets of R^k is compact in this metric topology. We may now proceed to the lemma.

LEMMA 1. *For* $\varepsilon > 0$ *and integers* $n_1, \ldots, n_k$, *define*

$$I^\varepsilon(\mathbf{n}) = I^\varepsilon(n_1, \ldots, n_k) = \{\mathbf{x} \in R^k \mid (n_i - 1)\varepsilon < x^{(i)} \leqq n_i \varepsilon \text{ for } i = 1, \ldots, k\}.$$

For a bounded convex subset C of R^k, let $V_\varepsilon(C)$ be the union of all $I^\varepsilon(\mathbf{n})$ which have the property that $I^\varepsilon(\mathbf{n}) \cap \partial C \neq \varnothing$. Then, for each constant M, $m_k(V_\varepsilon(C)) \to 0$ as $\varepsilon \to 0$ uniformly for all convex sets C such that the diameter of C is $\leq M$. Here m_k denotes k-dimensional Lebesgue measure.

Proof. Since the closure of a convex set is convex, we may without loss of generality, consider only compact convex sets. Also, since the statement of the lemma is invariant under translations in R^k, we may assume that all convex sets to be considered are contained in a fixed sphere in R^k. By the Blaschke Selection Theorem, it is now sufficient to prove that if C_m and C are nonempty compact convex sets such that

$$\Delta(C_m, C) \to 0 \quad \text{as } m \to \infty,$$

and if $\varepsilon_m \to 0$, then $m_k(V_{\varepsilon_m}(C_m)) \to 0$.

Now, if $\mathbf{x} \in V_{\varepsilon_m}(C_m)$ then $\mathbf{x} \in I^{\varepsilon_m}(\mathbf{n})$ for some $\mathbf{n}$ such that $I^{\varepsilon_m}(\mathbf{n}) \cap \partial C_m \neq \varnothing$. So, there is a $\mathbf{y} \in \partial C_m$ so that

$$\|\mathbf{x} - \mathbf{y}\| \leq \varepsilon_m k^{1/2}.$$

Using the definition of the metric Δ, we see that for this $\mathbf{y}$, there is a $\mathbf{z} \in \partial C$ so that

$$\|\mathbf{y} - \mathbf{z}\| \leq \Delta(C_m, C).$$

Hence, given $\eta > 0$, if m is so large that $\varepsilon_m k^{1/2} + \Delta(C_m, C) < \eta$ we have $V_{\varepsilon_m}(C_m)$ contained in $N(\partial C; \eta)$. Since ∂C is compact and $m_k(\partial C) = 0$, we see that

$$m_k[N(\partial C; \eta)] \to 0$$

as $\eta \to 0$. This completes the proof of the lemma.

The next step is to use this approximation lemma to extend Stone's local limit theorem to convex sets. In order to state the following result, it is necessary to define a sequence ν_n of measures on R^d which are essentially the Haar measures of the support lattices D_n^*. If A is a Borel measurable subset of R^d, $A \cap D_n^*$ can be regarded as a union of countably many $(d-d_1)$-dimensional Borel measurable sets. So, we may define

$$\nu_n(A) = m_{d-d_1}(A \cap D_n^*)/b_n^{d_1}.$$

Here it is understood that $m_0(A \cap D_n^*)$ is the number of points in the set $A \cap D_n^*$. The set $\mathbf{y} + b_n^{-1} C$ denotes, as usual, the set $\{\mathbf{y} + b_n^{-1} \mathbf{x} \mid \mathbf{x} \in C\}$.

THEOREM 2.

$$P(\mathbf{X}_n(1) \in \mathbf{y} + b_n^{-1} C) = \nu_n(\mathbf{y} + b_n^{-1} C) p(\mathbf{y}) + o(b_n^{-d})$$

where, for each finite M, $b_n^d o(b_n^{-d}) \to 0$ uniformly for all $\mathbf{y} \in R^d$ and for all convex subsets C of R^d such that $\sup_{\mathbf{x} \in C} \|\mathbf{x}\| \leq M$.

Proof. For each n, D_n^* is a countable union of disjoint $(d-d_1)$-dimensional

hyperplanes, which we may denote by $\{H_k^n, k \geq 1\}$. Let $\pi \colon R^d \to R^{d-d_1}$ be the projection defined by

$$\pi(x^{(1)}, \ldots, x^{(d)}) = (x^{(d_1+1)}, \ldots, x^{(d)}).$$

If we let $Q_k^n = \pi[(b_n H_k^n) \cap (b_n \mathbf{y} + C)]$, Q_k^n is a convex subset of R^{d-d_1}, and has a smaller diameter than C does in R^d. In what follows, $\mathbf{m} = (m_1, \ldots, m_{d-d_1})$ will range through the integer lattice of R^{d-d_1}.

For each pair of positive integers k and n, let

$$(3) \qquad U_k^n = \sum_{I^\varepsilon(\mathbf{m}) \subset Q_k^n} P(\mathbf{X}_n(1) \in H_k^n, \pi(b_n \mathbf{X}_n(1)) \in I^\varepsilon(\mathbf{m})),$$

and

$$(4) \qquad T_k^n = \sum_{I^\varepsilon(\mathbf{m}) \cap \partial Q_k^n \neq \varnothing} P(\mathbf{X}_n(1) \in H_k^n, \pi(b_n \mathbf{X}_n(1)) \in I^\varepsilon(\mathbf{m})).$$

We will show that T_k^n is negligible compared with b_n^{-d}, but that $b_n^d U_k^n$ is approximately equal to $p(\mathbf{y}) m_{d-d_1}(Q_k^n)$. Let $\mathbf{z}_k^n(\mathbf{m})$ be the unique point in H_k^n with the property that $\pi(\mathbf{z}_k^n(\mathbf{m}))$ is the midpoint of the cube $b_n^{-1} I^\varepsilon(\mathbf{m})$. An application of Theorem 1 of [10] then shows that, as $n \to \infty$,

$$P(\mathbf{X}_n(1) \in H_k^n, \pi(b_n \mathbf{X}_n(1)) \in I^\varepsilon(\mathbf{m})) = b_n^{-d} \varepsilon^{d-d_1} p(\mathbf{z}_k^n(\mathbf{m})) + o(b_n^{-d})$$

uniformly for all k, all $\mathbf{m}$, and all positive ε which are bounded above.

If $\mathbf{m}$ is such that $I^\varepsilon(\mathbf{m}) \cap \bar{Q}_k^n \neq \varnothing$, there is a $\mathbf{w} \in b_n H_k^n$ which satisfies $\pi \mathbf{w} \in I^\varepsilon(\mathbf{m}) \cap \bar{Q}_k^n$. Then $\mathbf{w} \in b_n \mathbf{y} + \bar{C}$, so $\|b_n \mathbf{y} - \mathbf{w}\| \leq M$. Furthermore,

$$\|\mathbf{w} - b_n \mathbf{z}_k^n(\mathbf{m})\| = \|\pi \mathbf{w} - \pi(b_n \mathbf{z}_k^n(\mathbf{m}))\| \leq \varepsilon (d - d_1)^{1/2}/2.$$

So, $\|\mathbf{y} - \mathbf{z}_k^n(\mathbf{m})\| \leq (M + \varepsilon(d-d_1)^{1/2})/b_n$. Using this inequality, the uniform continuity of $p(\cdot)$, and the fact that the number of terms in each of the sums (3) and (4) is bounded by $[2M + 2\varepsilon]^{d-d_1} \varepsilon^{d_1 - d}$, we have

$$U_k^n = p(\mathbf{y}) b_n^{-d} m_{d-d_1} \left\{ \bigcup_{I^\varepsilon(\mathbf{m}) \subset Q_k^n} I^\varepsilon(\mathbf{m}) \right\} + [2M + 2\varepsilon]^{d-d_1} \varepsilon^{d_1 - d} o(b_n^{-d}),$$

and

$$T_k^n = p(\mathbf{y}) b_n^{-d} m_{d-d_1} \left\{ \bigcup_{I^\varepsilon(\mathbf{m}) \cap \partial Q_k^n \neq \varnothing} I^\varepsilon(\mathbf{m}) \right\} + [2M + 2\varepsilon]^{d-d_1} \varepsilon^{d_1 - d} o(b_n^{-d}),$$

where $b_n^d o(b_n^{-d}) \to 0$ uniformly for all k and all bounded ε.

Now, let K_n be the set of integers k such that $\bar{Q}_k^n \neq \varnothing$. The number of elements in K_n is bounded in n and in C as long as $\sup_{\mathbf{x} \in C} \|\mathbf{x}\| \leq M$. By Lemma 1,

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \sup_k b_n^d T_k^n = 0.$$

Furthermore, since we can write

$$b_n^d U_k^n \leq p(\mathbf{y}) m_{d-d_1}(Q_k^n) + [2M + 2\varepsilon]^{d-d_1} \varepsilon^{d_1 - d} o(1)$$
$$\leq b_n^d(U_k^n + T_k^n) + 2[2M + 2\varepsilon]^{d-d_1} \varepsilon^{d_1 - d} o(1),$$

it follows that

$$(5) \qquad \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \sup_k \left| p(\mathbf{y}) m_{d-d_1}(Q_k^n) - b_n^d U_k^n \right| = 0.$$

To conclude the proof of the theorem, note that the first two lines below are true by definition, and that they, together with (5), imply the third:

$$b_n^d \sum_{k \in K_n} U_k^n \leqq b_n^d P(\mathbf{X}_n(1) \in \mathbf{y} + b_n^{-1}C) \leqq b_n^d \sum_{k \in K_n} (U_k^n + T_k^n),$$

$$b_n^d \nu_n(\mathbf{y} + b_n^{-1}C) = b_n^{d-d_1} \sum_{k \in K_n} m_{d-d_1}[(\mathbf{y} + b_n^{-1}C) \cap H_k^n] = \sum_{k \in K_n} m_{d-d_1}(Q_k^n),$$

$$\lim_{n \to \infty} b_n^d \left| P(\mathbf{X}_n(1) \in \mathbf{y} + b_n^{-1}C) - \nu_n(\mathbf{y} + b_n^{-1}C) p(\mathbf{y}) \right| = 0.$$

3. **General conditioned convergence theorems.** The two principal theorems in this section give conditions on a sequence of Borel measurable sets $E^n \subset R^d$ which guarantee that the processes $(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$ converge weakly to a limiting process. The first of these considers the case in which the sets E^n converge to a point of R^d in a sense to be defined, while the second uses this result to obtain a similar conclusion for sets E^n which converge to a more general set in R^d.

It will always be the case in what follows that the event $\{\mathbf{X}_n(1) \in E^n\}$ has positive probability, at least for sufficiently large n. So, the conditioned processes

$$(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$$

are defined simply as measures on $D^d[0, 1]$ which take the value

$$P(\mathbf{X}_n \in A \mid \mathbf{X}_n(1) \in E^n)$$

at each measurable subset A of $D^d[0, 1]$. With this definition, we are in a position to state the following theorem.

THEOREM 3. *Let* $\mathbf{y} \in R^d$ *be such that* $p(\mathbf{y}) > 0$, *and let* $\{E^n\}$ *be a sequence of Borel measurable subsets of* R^d *which satisfy the following conditions*:

$$(6) \qquad \lim_{n \to \infty} \sup_{\mathbf{z} \in E^n} \|\mathbf{z} - \mathbf{y}\| = 0,$$

$$(7) \qquad \sup_n [b_n \text{ diameter } (E^n)] < \infty,$$

and

$$(8) \qquad \inf_n [b_n^d \nu_n(E^n)] > 0.$$

Suppose there are two sequences of Borel measurable convex sets $\{C^n\}$ *and* $\{F^n\}$ *which satisfy* $C^n \subset E^n \subset F^n$ *and*

$$(9) \qquad \lim_{n \to \infty} b_n^d \nu_n(F^n \backslash C^n) = 0.$$

Then the sequence of processes $(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$ *converges weakly in* $D^d[0, 1]$ *to a process* $\mathbf{Y}_\mathbf{y}$ *with finite-dimensional distributions given by*

$$P(\mathbf{Y}_\mathbf{y}(t_1) \in A_1, \ldots, \mathbf{Y}_\mathbf{y}(t_k) \in A_k)$$

$$= \int_{A_1} \cdots \int_{A_k} \frac{p_{1-t_k}(\mathbf{y} - \mathbf{z}_k)}{p(\mathbf{y})} P(\mathbf{X}(t_1) \in d\mathbf{z}_1, \ldots, \mathbf{X}(t_k) \in d\mathbf{z}_k)$$

for $0 < t_1 < \cdots < t_k < 1$; $\mathbf{Y}_\mathbf{y}(0) \equiv 0$ *and* $\mathbf{Y}_\mathbf{y}(1) \equiv \mathbf{y}$. *Here* p_t *is the density of* $\mathbf{X}(t)$:

$$p_t(\mathbf{x}) = \frac{1}{t^{d/\alpha}} p\left(\frac{\mathbf{x} - \boldsymbol{\eta}(t)}{t^{1/\alpha}}\right).$$

REMARKS. It should be noted that if $d_1 = 0$ and if E^n is convex, the above conditions simplify considerably. In this case ν_n is d-dimensional Lebesgue measure for each n and condition (9) is automatically satisfied. The proof of Theorem 3 is a straightforward extension of the proof of Theorem 4 of [8], so we will only outline the main ideas of the proof here. The details may be found in [9, p. 25].

Outline of proof. (a) Theorem 2 is used to show that for $t \in (0, 1)$,

$$\frac{P(\mathbf{X}_n(t) \in \mathbf{x} + E^n)}{P(\mathbf{X}_n(1) \in E^n)} \to \frac{p_t(\mathbf{x} + \mathbf{y})}{p(\mathbf{y})}$$

uniformly for all $\mathbf{x}$ such that $-\mathbf{x}$ is in the support of $P(\mathbf{X}_n(1) - \mathbf{X}_n(t) \in d\mathbf{z})$.

(b) Part (a) is used to show that the finite-dimensional distributions of the processes $(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$ converge to those of $\mathbf{Y}_\mathbf{y}$.

(c) Part (a) is used again to show that for $t < 1$ there are constants $M(t)$ and $n(t)$ so that

$$P(X_n^{(i)} \in A \mid \mathbf{X}_n(1) \in E^n) \leqq M(t) P(X_n^{(i)} \in A)$$

for each $i = 1, \ldots, d$, for each $n \geqq n(t)$, and for each Borel set A in $D[0, 1]$ which depends only on times in $[0, t]$.

(d) Part (c) is used to show that the processes $(X_n^{(i)} \mid \mathbf{X}_n(1) \in E^n)$ are tight in $D[0, t]$ for each $t < 1$ and each i.

(e) A time reversal argument is used to show that the sequence $(X_n^{(i)} \mid \mathbf{X}_n(1) \in E^n)$ is tight in $D[0, 1]$ for each i.

(f) Finally, Lemma 1 of [5] is used to show that the processes $(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$ are tight in $D^d[0, 1]$.

Before proceeding to the main theorem of this section, it is necessary, for technical reasons, to prove a lemma concerning the behavior of $P(\mathbf{Y}_\mathbf{y} \in F)$ as a function of $\mathbf{y}$ for a fixed set F. Note that this function is well defined only for $\mathbf{y} \in Q$, the support of $p(\cdot)$. The method of proof is of interest because it is an application of Theorem 3.

LEMMA 2. *For each closed subset* F *of* $D^d[0, 1]$, $P(\mathbf{Y}_\mathbf{y} \in F)$ *is an uppersemi-continuous function of* $\mathbf{y}$ *in* Q. *For every Borel subset* F *of* $D^d[0, 1]$, $P(\mathbf{Y}_\mathbf{y} \in F)$ *is a Borel measurable function of* $\mathbf{y}$ *in* Q.

Proof. Let F be a closed subset of $D^d[0, 1]$, and let $y_n \to y_0$, where y_n and y_0 are in Q. It will suffice to show that the measures Y_{y_n} on $D^d[0, 1]$ converge weakly to the measure Y_{y_0}, since then it will follow from Theorem 2.1 of [1] that

$$P(Y_{y_0} \in F) \geqq \limsup_{n \to \infty} P(Y_{y_n} \in F).$$

Let β be any number greater than $(d_1)^{1/2}$. Since the span of the lattice part of the distribution of ξ_i is equal to one, we have for any $y \in R^d$,

$$m_{d-d_1}\{x \in R^d \mid \|x - y\| \leqq \beta, x \in D_n\} \geqq m_{d-d_1}\{x \in R^{d-d_1} \mid \|x\|^2 \leqq \beta^2 - d_1\} > 0.$$

For each pair of integers n and m, let

$$E_n^m = \{x \in R^d \mid \|x - y_n\| \leqq \beta/b_m\}.$$

Then b_m diameter $(E_n^m) \leqq 2\beta$, and

$$b_m^d \nu_m(E_n^m) = m_{d-d_1}(b_m E_n^m \cap b_m D_m^*) \geqq m_{d-d_1}\{x \in R^{d-d_1} \mid \|x\|^2 \leqq \beta^2 - d_1\}.$$

Furthermore, E_n^m is convex for each n and m. So, by Theorem 3, $(X_m \mid X_m(1) \in E_n^m)$ converges weakly to Y_{y_n} for each n, and $(X_{m_k} \mid X_{m_k}(1) \in E_{n_k}^{m_k})$ converges weakly to Y_{y_0} for each pair of sequences of integers $\{m_k\}$ and $\{n_k\}$ which tend to infinity. Now, let T be a bounded, continuous function on $D^d[0, 1]$, and define

$$a_{n,m} = \int_{D^d[0,1]} T(h) P(X_m \in dh \mid X_m(1) \in E_n^m).$$

By Theorem 2.1 of [1]

$$\lim_{m \to \infty} a_{n,m} = \int_{D^d[0,1]} T(h) P(Y_{y_n} \in dh)$$

and

$$\lim_{k \to \infty} a_{n_k, m_k} = \int_{D^d[0,1]} T(h) P(Y_{y_0} \in dh)$$

whenever $n_k \to \infty$ and $m_k \to \infty$. So,

$$\lim_{n \to \infty} \left[\lim_{m \to \infty} a_{n,m} \right] = \int_{D^d[0,1]} T(h) P(Y_{y_0} \in dh).$$

Since this is true for every such function T, we may conclude that Y_{y_n} converges weakly to Y_{y_0}.

This second part of the lemma follows from the first by applying the Monotone Class Theorem [2, p. 5].

The next theorem gives general conditions under which one has convergence of the processes $(X_n \mid X_n(1) \in E^n)$. Most of these conditions relate to the type of sets E^n which may be used. In spite of their somewhat involved appearance, it is usually relatively easy to verify them, as will be seen in §4. One condition is a probabilistic one, and again it is automatically satisfied in many cases.

THEOREM 4. *Let $\{N_n; n = 1, 2, \ldots\}$ be a sequence of integers such that $0 < N_n \leqq \infty$, and let $\{E_k^n\}$, $\{C_k^n\}$ and $\{F_k^n\}$ be three families of Borel subsets of R^d, where k runs through the integers $\{1, 2, \ldots, N_n\}$ for each $n = 1, 2, \ldots$. Define $E^n = \bigcup_k E_k^n$, and assume that these sets satisfy the following conditions:*

    (i) *C_k^n and F_k^n are convex; $C_k^n \subset E_k^n \subset F_k^n$; $\{E^n\}$ are disjoint for each n.*

    (ii) *$b_n^d \nu_n(F_k^n \backslash C_k^n) \to 0$ as $n \to \infty$ uniformly in k.*

    (iii) *$\sup_{k,n} [b_n \text{ diameter } (E_k^n)] < \infty$ and $\inf_{k,n} b_n^d \nu_n(E_k^n) > 0$.*

    (iv) *If we define $\mu_n(dx) = \nu_n(dx \cap E^n)$, then there are constants β_n and a measure $\mu(dx)$ on R^d satisfying $0 < \int_{R^d} p(x)\mu(dx) < \infty$ so that $\beta_n \mu_n(dx)$ converges to $\mu(dx)$ in the sense that*

$$\beta_n \int h(\mathbf{x})\mu_n(d\mathbf{x}) \to \int h(\mathbf{x})\mu(d\mathbf{x})$$

for every continuous function $h(\mathbf{x})$ which has its support in a compact subset of Q.

    (v) *For each $\varepsilon > 0$, there is a compact subset K of Q so that*

$$\liminf_{n \to \infty} P(\mathbf{X}_n(1) \in K \mid \mathbf{X}_n(1) \in E^n) \geqq 1 - \varepsilon.$$

Then, $(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$ converges weakly to $\mathbf{Y}_\mu$ in $D^d[0, 1]$, where $\mathbf{Y}_\mu$ is a random element satisfying

(10)
$$P(\mathbf{Y}_\mu \in A) = \frac{\int P(\mathbf{Y}_y \in A) p(\mathbf{y}) \mu(d\mathbf{y})}{\int p(\mathbf{y}) \mu(d\mathbf{y})}$$

for each Borel subset A of $D^d[0, 1]$.

REMARK. It will be shown later (see Lemma 3) that condition (v) above could be replaced by the condition $\beta_n \mu_n(R^d) = O(1)$.

Proof. Lemma 2 guarantees that $P(\mathbf{Y}_y \in A)$ is a measurable function of $\mathbf{y}$ for each Borel set A in $D^d[0, 1]$, so the integral in (10) makes sense. Furthermore, the right-hand side of (10) is a probability measure on $D^d[0, 1]$, and may therefore be taken as the definition of $P(\mathbf{Y}_\mu \in A)$. So, by Theorem 2.1 of [1], it will suffice to show that for every closed subset F of $D^d[0, 1]$,

(11)
$$P(\mathbf{Y}_\mu \in F) \geqq \limsup_{n \to \infty} P(\mathbf{X}_n \in F \mid \mathbf{X}_n(1) \in E^n).$$

Let F be a closed set in $D^d[0, 1]$, and let K be a fixed compact subset of Q such that $\mu(\partial K) = 0$. Define $h(\mathbf{y}) = P(\mathbf{Y}_y \in F)$ on Q. Then, $h(\mathbf{y})$ is bounded by one on Q, and is upper semicontinuous on Q by Lemma 2. So, there is a sequence $h_m(\mathbf{y})$ of continuous functions on Q such that $h_m(\mathbf{y}) \leqq 1$ and $h_m(\mathbf{y}) \downarrow h(\mathbf{y})$ on Q [11, Lemma 6-9 IV].

Fix an $m \geqq 1$. We will show that

(12)
$$\limsup_{n \to \infty} \sup_{E_k^n \cap K \neq \varnothing} \left[P(\mathbf{X}_n \in F \mid \mathbf{X}_n(1) \in E_k^n) - \inf_{\mathbf{y} \in E_k^n} h_m(\mathbf{y}) \right] \leqq 0.$$

Suppose that (12) is in fact positive. Then there is an $\varepsilon > 0$ such that for each integer N, there is an $n \geq N$, an integer k, and a $\mathbf{y} \in E_k^n$ so that

$$P(\mathbf{X}_n \in F \mid \mathbf{X}_n(1) \in E_k^n) - h_m(\mathbf{y}) \geq \varepsilon$$

and $E_k^n \cap K \neq \varnothing$. By condition (iii), all these $\mathbf{y}$'s lie in a compact subset of Q for sufficiently large N. So, there are sequences $\{\mathbf{y}_j\}$ in Q, $\{k_j\}$ and $\{n_j\}$ so that $\mathbf{y}_j \in E_{k_j}^{n_j}$, $n_j \uparrow \infty$, $\mathbf{y}_j \to \mathbf{y}_0 \in Q$, and

$$P(\mathbf{X}_{n_j} \in F \mid X_{n_j}(1) \in E_{k_j}^{n_j}) - h_m(\mathbf{y}_j) \geq \varepsilon.$$

Noting that if the hypotheses of Theorem 3 hold for a subsequence $\{n_j\}$ then so does the conclusion, we may apply that theorem to conclude that $(\mathbf{X}_{n_j} \mid X_{n_j}(1) \in E_{k_j}^{n_j})$ converges weakly to $\mathbf{Y}_{\mathbf{y}_0}$. So, since F is closed,

$$h(\mathbf{y}_0) \geq \limsup_{j \to \infty} P(\mathbf{X}_{n_j} \in F \mid \mathbf{X}_{n_j}(1) \in E_{k_j}^{n_j}).$$

Hence,

$$h_m(\mathbf{y}_0) \geq h(\mathbf{y}_0) \geq \varepsilon + \limsup_{j \to \infty} h_m(\mathbf{y}_j),$$

which contradicts the continuity of h_m. So, we conclude that (12) holds.

Now, let $\tilde{K}$ be a compact subset of the interior of K. Applying (12) and using the fact that E^n is the disjoint union of the E_k^n's, we obtain

$$\limsup_{n \to \infty} P(\mathbf{X}_n \in F \mid \mathbf{X}_n(1) \in E^n)$$

(13)
$$\leq \limsup_{n \to \infty} \left\{ \frac{\sum\limits_{E_k^n \subset K} \left[\inf\limits_{\mathbf{y} \in E_k^n} h_m(\mathbf{y})\right] P(\mathbf{X}_n(1) \in E_k^n)}{\sum\limits_{E_k^n \cap K \neq \varnothing} P(\mathbf{X}_n(1) \in E_k^n)} + \frac{P(\mathbf{X}_n(1) \in E^n \backslash \tilde{K})}{P(\mathbf{X}_n(1) \in E^n)} \right\}.$$

Let $\mathbf{y}_k^n$ be a point in the closure of E_k^n so that $p(\mathbf{y}_k^n) = \inf_{\mathbf{y} \in E_k^n} p(\mathbf{y})$. For each k and n, there is a sphere G_k^n in R^d of diameter at most twice the diameter of E_k^n such that $E_k^n \subset G_k^n$. So, by replacing F_k^n by $F_k^n \cap G_k^n$ if necessary, we see that we may assume that

$$\sup_{k,n} [b_n \text{ diameter } (F_k^n)] < \infty.$$

Then, by Theorem 2,

$$P(\mathbf{X}_n(1) \in F_k^n) = \nu_n(F_k^n) p(\mathbf{y}_k^n) + o(b_n^{-d})$$

and

$$P(\mathbf{X}_n(1) \in C_k^n) = \nu_n(C_k^n) p(\mathbf{y}_k^n) + o(b_n^{-d})$$

where, in each case, $b_n^d o(b_n^{-d}) \to 0$ uniformly in k. By condition (ii),

(14)
$$b_n^d P(\mathbf{X}_n(1) \in E_k^n) = b_n^d \nu_n(E_k^n) p(\mathbf{y}_k^n) + o(1)$$

where $o(1) \to 0$ as $n \to \infty$ uniformly in k. Since $b_n^d \nu_n(E_k^n) p(\mathbf{y}_k^n)$ is bounded away

from zero for k such that $E_k^n \cap K \neq \varnothing$ and for sufficiently large n, we may rewrite (14) as

$$P(\mathbf{X}_n(1) \in E_k^n) = (1 + o(1)) \nu_n(E_k^n) \inf_{y \in E_k^n} p(\mathbf{y})$$

where $o(1) \to 0$ uniformly for k such that $E_k^n \cap K \neq \varnothing$. Similarly, we have

$$P(\mathbf{X}_n(1) \in E_k^n) = (1 + o(1)) \nu_n(E_k^n) \sup_{y \in E_k^n} p(\mathbf{y}) \tag{15}$$

with the same uniformity statement.

We will now apply these estimates to the first term on the right side of (13) to obtain

$$\limsup_{n \to \infty} \beta_n \sum_{E_k^n \subset K} \left[\inf_{y \in E_k^n} h_m(\mathbf{y}) \right] P(\mathbf{X}_n(1) \in E_k^n) \leqq \limsup_{n \to \infty} \beta_n \sum_{E_k^n \subset K} \left[\inf_{y \in E_k^n} h_m(\mathbf{y}) p(\mathbf{y}) \right] \nu_n(E_k^n)$$

$$\leqq \limsup_{n \to \infty} \beta_n \sum_{E_k^n \subset K} \int_{E_k^n} h_m(\mathbf{y}) p(\mathbf{y}) \mu_n(d\mathbf{y})$$

$$\leqq \int_K h_m(\mathbf{y}) p(\mathbf{y}) \mu(d\mathbf{y})$$

since $\beta_n \mu_n(d\mathbf{y})$ converges weakly to $\mu(d\mathbf{y})$ on compact subsets of Q, $h_m(\mathbf{y})$ is continuous, and $\mu(\partial K) = 0$. In the same way, we obtain

$$\liminf_{n \to \infty} \beta_n \sum_{E_k^n \cap K \neq \varnothing} P(\mathbf{X}_n(1) \in E_k^n) \geqq \int_K p(\mathbf{y}) \mu(d\mathbf{y}).$$

Applying the dominated convergence theorem, we now have

$$\limsup_{n \to \infty} P(\mathbf{X}_n \in F \mid \mathbf{X}_n(1) \in E^n) \leqq \frac{\int_K h(\mathbf{y}) p(\mathbf{y}) \mu(d\mathbf{y})}{\int_K p(\mathbf{y}) \mu(d\mathbf{y})} + \limsup_{n \to \infty} \frac{P(\mathbf{X}_n(1) \in E^n \setminus \tilde{K})}{P(\mathbf{X}_n(1) \in E^n)}.$$

This is true for any pair of compact subsets K and $\tilde{K}$ of Q such that $\tilde{K}$ is contained in the interior of K. So, (11) follows from condition (v) of the theorem, and the fact that $0 < \int_Q p(\mathbf{y}) \mu(d\mathbf{y}) < \infty$.

REMARK. A straightforward verification shows that the random elements $\mathbf{Y}_\mu$ are Markov processes. However, they do not have stationary transition probabilities.

One might expect that in obtaining the weak convergence of the conditioned processes $(\mathbf{X}_n \mid \mathbf{X}_n(1) \in E^n)$, it would not be necessary to impose any probabilistic conditions other than those needed to guarantee the weak convergence of $\mathbf{X}_n$ to $\mathbf{X}$. However, we have included an additional such condition in this theorem—condition (v). This condition is not very restrictive, since it is easy to check that it follows from the conclusion of the theorem. Unfortunately, though, the theorem is not true without condition (v) as may be seen by referring to Example 5.1 of [9].

A useful sufficient condition for condition (v) is given in the following lemma. An application of it will occur in the next section.

LEMMA 3. *Suppose all conditions except* (v) *of Theorem* 4 *are satisfied. Then, if* $\{\beta_n\mu_n(R^d)\}$ *is bounded in n, condition* (v) *holds also.*

Proof. For $\varepsilon > 0$, define $K_\varepsilon = \{\mathbf{x} \in R^d \mid p(\mathbf{x}) \geq \varepsilon\}$, and let K be any compact subset of Q with the property that $\mu(\text{interior of } K) > 0$. There exists such a K since $\int p(\mathbf{x})\mu(d\mathbf{x}) > 0$. Since condition (v) was not used to derive equation (14) we have, for a fixed $\varepsilon > 0$,

$$P(\mathbf{X}_n(1) \in E_k^n) \leq 2\varepsilon\nu_n(E_k^n)$$

for k so that $E_k^n \cap K_\varepsilon^c \neq \varnothing$ and for sufficiently large n. On the other hand, it follows from the same equation that

$$P(\mathbf{X}_n(1) \in E_k^n) \geq \tfrac{1}{2} \inf_{\mathbf{y} \in K} p(\mathbf{y})\nu_n(E_k^n)$$

for k so that $E_k^n \cap K \neq \varnothing$ and for sufficiently large n.

So, for large n, we have

$$P(\mathbf{X}_n(1) \in K_\varepsilon^c \mid \mathbf{X}_n(1) \in E^n) \leq \frac{\displaystyle\sum_{E_k^n \cap K_\varepsilon^c \neq \varnothing} P(\mathbf{X}_n(1) \in E_k^n)}{\displaystyle\sum_{E_k^n \cap K \neq \varnothing} P(\mathbf{X}_n(1) \in E_k^n)}$$

$$\leq \frac{4\varepsilon\nu_n(E^n)}{[\inf_{\mathbf{y} \in K} p(\mathbf{y})]\nu_n(E^n \cap K)} \leq \frac{4\varepsilon}{\inf_{\mathbf{y} \in K} p(\mathbf{y})} \frac{\beta_n\mu_n(R^d)}{\beta_n\mu_n(K)}.$$

From condition (iv), it follows that $\lim \inf \beta_n\mu_n(K) \geq \mu(\text{interior of } K)$. So, ε can be chosen so that K_ε satisfies condition (v).

4. The convergence theorem in the "smooth" case. The purpose of this section is to show that Theorem 4 has a wide range of applicability. In particular, we will show that all the conditions of this theorem are automatically satisfied in case the sets E^n are the "natural" neighborhoods of a smooth compact hypersurface in R^d which has no boundary points, at least when the distribution of ξ_i has no lattice component. When ξ_i has a lattice component, the statement of an analogous result would become more involved, and it would probably be easier to investigate the role of the support lattice for each individual example than to try to state a general result. Hence, in this section, we will restrict ourselves to the case $d_1 = 0$.

The hypersurface to be used in the conditioning will be defined in terms of a mapping $L(\mathbf{x})$ from R^d to R^r ($1 \leq r < d$). Writing $L(\mathbf{x}) = (L_1(\mathbf{x}), \ldots, L_r(\mathbf{x}))$, we will require that each function $L_i(\mathbf{x})$ have continuous mixed partial derivatives of order two in all of its variables $x^{(1)}, \ldots, x^{(d)}$. For each $\mathbf{x} \in R^d$, define

$$S(\mathbf{x}) = \left(\sum_{1 \leq i_1 < \cdots < i_r \leq d} \left[\frac{\partial(L_1, \ldots, L_r)}{\partial(x^{(i_1)}, \ldots, x^{(i_r)})} \right]^2 \right)^{1/2},$$

where we have used the standard notation for the Jacobians. We then have the following theorem.

THEOREM 5. *Suppose that the following conditions hold in addition to those stated above*:

(i) $S(\mathbf{x}) > 0$ *for each* $\mathbf{x}$ *where* $L(\mathbf{x}) = 0$.

(ii) $L(\mathbf{x}) = 0$ *for some* $\mathbf{x}$ *in* Q.

(iii) $\lim \inf_{\|\mathbf{x}\| \to \infty} \|L(\mathbf{x})\| > 0$.

Then the assumptions (and hence the conclusion) of Theorem 4 are satisfied when E^n *and* μ *are taken to be*

$$E^n = \{\mathbf{x} \in R^d \mid |L_i(\mathbf{x})| \leq b_n^{-1} \text{ for } i = 1, \ldots, r\},$$

$$\mu(d\mathbf{x}) = m_{d-r}(d\mathbf{x} \cap E)/S(\mathbf{x}),$$

where E *is the set* $\{\mathbf{x} \mid L(\mathbf{x}) = 0\}$.

REMARK. Before proceeding to the proof of this theorem, it may be in order to comment briefly on the conditions listed above. Condition (i) guarantees that E is genuinely $(d-r)$-dimensional at every point. Condition (iii) forces all the sets E^n to lie within a compact part of R^d, thus making μ a finite measure and leading to the application of Lemma 3. Condition (ii) will make the denominator of expression (10) positive. Finally, the smoothness assumption on L makes it possible to write E^n as a disjoint union of asymptotically convex sets as required in the first two conditions of Theorem 4.

Proof. Let $\inf_{\mathbf{y} \in E} \|\mathbf{x} - \mathbf{y}\|$ be denoted by $\rho(x, E)$, and define, as in (2), an ε-neighborhood of E by

$$N(\varepsilon) = \{\mathbf{x} \in R^d \mid \rho(\mathbf{x}, E) < \varepsilon\}$$

for each positive ε. The first part of the proof will show that there is an n_0 and a constant R so that, for $n \geq n_0$,

(16) $N(R^{-1}b_n^{-1}) \subset E^n \subset N(Rb_n^{-1})$.

The left inclusion is easy to obtain. Choose constants $\eta > 0$ and M, and a compact convex subset K of R^d which contains $N(\eta)$, so that all first and second order partial derivatives of each L_i are bounded by M in K. That this is possible is a consequence of the regularity assumptions made on L and condition (iii) above. By the mean value theorem, if $\mathbf{x} \in E$ and $\mathbf{y}$ satisfies $\|\mathbf{x} - \mathbf{y}\| \leq \eta$, then

$$|L_i(\mathbf{y})| \leq dM \|\mathbf{x} - \mathbf{y}\| \quad \text{for } 1 \leq i \leq r.$$

So, for n so large that $(dMb_n)^{-1} \leq \eta$, we have $N((dMb_n)^{-1}) \subset E^n$.

To show that the right inclusion in (16) holds also, it will suffice to show that the function $\rho(\mathbf{x}, E)/\|L(\mathbf{x})\|$ is uniformly bounded in $K \backslash E$. Let $\mathbf{x}_0$ be any point in $K \backslash E$, and let $\mathbf{y}_0$ be any point in E such that $\|\mathbf{x}_0 - \mathbf{y}_0\| = \rho(\mathbf{x}_0, E)$. Such a point exists since E is compact. Since $S(\mathbf{y}_0) > 0$, there is a set of indices, which we will take to be $\{1, \ldots, r\}$ for simplicity of notation, so that the mapping

$$(z^{(1)}, \ldots, z^{(d)}) \to (L_1(\mathbf{z}), \ldots, L_r(\mathbf{z}), z^{(r+1)}, \ldots, z^{(d)})$$

maps a neighborhood $\mathcal{N}$ of $\mathbf{y}_0$ one-to-one and onto a neighborhood $\mathcal{N}'$ of $(0, \ldots, 0, y_0^{(r+1)}, \ldots, y_0^{(d)})$. As $\mathbf{z}$ ranges through $\mathcal{N} \cap E$, the function $\|\mathbf{z} - \mathbf{x}_0\|^2$ achieves its minimum at $\mathbf{z} = \mathbf{y}_0$. So, for $i > r$,

$$0 = (x_0^{(1)} - y_0^{(1)}) dz^{(1)}/dz^{(i)} + \cdots + (x_0^{(r)} - y_0^{(r)}) dz^{(r)}/dz^{(i)} + (x_0^{(i)} - y_0^{(i)}) \quad \text{at } \mathbf{z} = \mathbf{y}_0.$$

Now, let $\mathcal{L}$ and $\tilde{\mathcal{L}}$ be the matrices

$$\mathcal{L} = (\partial L_i / \partial z^{(j)})_{i=1,\ldots,r; j=1,\ldots,r} \quad \text{and} \quad \tilde{\mathcal{L}} = (\partial L_i / \partial z^{(j)})_{i=1,\ldots,r; j=1,\ldots,d},$$

where the partial derivatives are evaluated at $\mathbf{y}_0$. Then, as a consequence of the fact that the inverse of the Jacobian matrix of a transformation is the Jacobian matrix of the inverse of that transformation, we have

$$\begin{pmatrix} \partial z^{(1)}/\partial z^{(i)} \\ \vdots \\ \partial z^{(r)}/\partial z^{(i)} \end{pmatrix} = -\mathcal{L}^{-1} \begin{pmatrix} \partial L_1/\partial z^{(i)} \\ \vdots \\ \partial L_r/\partial z^{(i)} \end{pmatrix} \quad \text{for } i > r.$$

So, $\mathbf{x}_0 - \mathbf{y}_0 = (x_0^{(1)} - y_0^{(1)}, \ldots, x_0^{(r)} - y_0^{(r)}) \mathcal{L}^{-1} \tilde{\mathcal{L}}$. Let $\boldsymbol{\alpha}$ be the r-vector

$$(x_0^{(1)} - y_0^{(1)}, \ldots, x_0^{(r)} - y_0^{(r)}) \mathcal{L}^{-1},$$

and let $\lambda_1, \ldots, \lambda_r$ and $\mathbf{x}_1, \ldots, \mathbf{x}_r$ be the complete set of eigenvalues of $\tilde{\mathcal{L}} \tilde{\mathcal{L}}^*$ and the corresponding orthonormalized eigenvectors. (The asterisk denotes the transpose matrix.) Then

$$\|\boldsymbol{\alpha} \tilde{\mathcal{L}} \tilde{\mathcal{L}}^*\|^2 = \sum_{i=1}^{r} \lambda_i^2 (\boldsymbol{\alpha}, \mathbf{x}_i)^2,$$

while

$$\|\boldsymbol{\alpha} \tilde{\mathcal{L}}\|^2 = \sum_{i=1}^{r} \lambda_i (\boldsymbol{\alpha}, \mathbf{x}_i)^2.$$

So, since $\lambda_i \geqq 0$ for each i,

$$\|\boldsymbol{\alpha} \tilde{\mathcal{L}} \tilde{\mathcal{L}}^*\|^2 \geqq \left(\min_{1 \leqq i \leqq r} \lambda_i \right) \|\boldsymbol{\alpha} \tilde{\mathcal{L}}\|^2.$$

By the Cauchy-Binet formula [6, p. 1], $S^2(\mathbf{y}_0) = \det(\tilde{\mathcal{L}} \tilde{\mathcal{L}}^*)$, which is equal to the product of the eigenvalues λ_i, $i = 1, \ldots, r$. For each i,

$$\lambda_i = \|\mathbf{x}_i \tilde{\mathcal{L}}\|^2 \leqq d^2 M^2,$$

so $\min_{1 \leqq i \leqq r} \lambda_i$ is bounded below by a constant multiple of $S^2(\mathbf{y}_0)$. Since $S(\cdot)$ is bounded away from zero on E,

$$\|\boldsymbol{\alpha} \tilde{\mathcal{L}}\|^2 = \|\mathbf{x}_0 - \mathbf{y}_0\|^2 \quad \text{and} \quad \|\boldsymbol{\alpha} \tilde{\mathcal{L}} \tilde{\mathcal{L}}^*\|^2 = \sum_{i=1}^{r} (\nabla L_i, (\mathbf{x}_0 - \mathbf{y}_0))^2,$$

where ∇L_i is the gradient of L_i, we see that there is a $\delta > 0$ which is independent of $\mathbf{x}_0 \in K \backslash E$ so that

$$\sum_{i=1}^{r} (\nabla L_i, (\mathbf{x}_0 - \mathbf{y}_0))^2 \geqq \delta \|\mathbf{x}_0 - \mathbf{y}_0\|^2 = \delta \rho^2(\mathbf{x}_0, E).$$

By Taylor's Theorem,

$$\|L(\mathbf{x}_0)\|^2 = \sum_{i=1}^{r} (\nabla L_i, (\mathbf{x}_0 - \mathbf{y}_0))^2 + o(\|\mathbf{x}_0 - \mathbf{y}_0\|^2),$$

where $\|\mathbf{x}_0 - \mathbf{y}_0\|^{-2} o(\|\mathbf{x}_0 - \mathbf{y}_0\|^2) \to 0$ as $\|\mathbf{x}_0 - \mathbf{y}_0\| \to 0$ uniformly for $\mathbf{x}_0 \in K \backslash E$. So, the right inclusion in (16) can also be achieved for appropriate choices of n_0 and R.

Fix an R and an n so that (16) holds, and let $\mathscr{S}(\mathbf{x}, \varepsilon)$ denote the d-dimensional open sphere with center at $\mathbf{x}$ and radius ε. Since E is compact, it may be covered with finitely many spheres with center in E and of radius b_n^{-1}. If we increase the radii of these spheres to $(1 + R^{-1})b_n^{-1}$ and eliminate "unnecessary" spheres, we can assume that, if $\{\mathbf{x}_j\}$ is the set of centers of the remaining spheres, $\|\mathbf{x}_i - \mathbf{x}_j\| \geq (Rb_n)^{-1}$ for $i \neq j$. Then,

$$(17) \qquad\qquad E^n \subset N(Rb_n^{-1}) \subset \bigcup_j \mathscr{S}\left(\mathbf{x}_j, \frac{R+1+R^{-1}}{b_n}\right).$$

The next step is to partition the union on the right side of (17) into disjoint convex sets. This is accomplished by setting

$$I_i^n = \left\{\mathbf{z} \in R^d \mid \|\mathbf{z} - \mathbf{x}_i\| = \min_j \|\mathbf{z} - \mathbf{x}_j\| < \frac{R+1+R^{-1}}{b_n},\right.$$
$$\left. \|\mathbf{z} - \mathbf{x}_k\| > \min_j \|\mathbf{z} - \mathbf{x}_j\| \text{ for } k \leq i-1 \right\}.$$

Then, by (17) $E^n \subset \bigcup_i I_i^n$; $\{I_i^n\}$ are disjoint for each n; and each I_i^n is convex, since it is the intersection of a sphere and finitely many half spaces. Also, the diameter of I_i^n is less than or equal to $2(R+1+R^{-1})b_n^{-1}$ and $m_d(I_i^n \cap E^n) \geq B(Rb_n)^{-d}$, where B is a constant depending only on d, since $I_i^n \cap E^n$ contains a sphere of radius $(2Rb_n)^{-1}$. Now, let $B_1 = d^2 M(R+1+R^{-1})^2$ and define

$$E_i^n = I_i^n \cap E^n,$$
$$C_i^n = I_i^n \cap \{\mathbf{z} \in R^d \mid |(\nabla L_j(\mathbf{x}_i), (\mathbf{z} - \mathbf{x}_i))| \leq b_n^{-1} - B_1 b_n^{-2} \text{ for } j = 1, \ldots, r\},$$
$$F_i^n = I_i^n \cap \{\mathbf{z} \in R^d \mid |(\nabla L_j(\mathbf{x}_i), (\mathbf{z} - \mathbf{x}_i))| \leq b_n^{-1} + B_1 b_n^{-2} \text{ for } j = 1, \ldots, r\}.$$

Note that C_i^n and F_i^n are convex sets. Furthermore, $C_i^n \subset E_i^n \subset F_i^n$ since

$$|L_j(\mathbf{z}) - (\nabla L_j(\mathbf{x}_i), (\mathbf{z} - \mathbf{x}_i))| \leq B_1 b_n^{-2}$$

for each $\mathbf{z} \in I_i^n$ by Taylor's Theorem. This completes the verification of conditions (i) and (iii) of Theorem 4.

Turning to condition (ii), note that for each i,

$$F_i^n \backslash C_i^n \subset \bigcup_{j=1}^{r} \left\{\mathbf{z} \in \mathscr{S}\left(\mathbf{x}_i, \frac{R+1+R^{-1}}{b_n}\right) \mid \left| |(\nabla L_j(\mathbf{x}_i), (\mathbf{z} - \mathbf{x}_i))| - b_n^{-1}\right| \leq B_1 b_n^{-2}\right\}.$$

Each set in this union consists of two parts, each of which lies between two $(d-1)$-dimensional hyperplanes which are at a distance $2B_1 b_n^{-2}\|\nabla L_j(\mathbf{x}_i)\|^{-1}$ from each other. So,

$$(18) \qquad m_d(F_i^n \backslash C_i^n) \leqq \sum_{j=1}^{r} \frac{4B_1}{b_n^2} \frac{1}{\|\nabla L_j(\mathbf{x}_i)\|} B_2 \left(\frac{R+1+R^{-1}}{b_n}\right)^{d-1},$$

where B_2 is the volume of the unit $(d-1)$-dimensional sphere. Now, $\mathscr{L}\mathscr{L}^*(\mathbf{x}_i)$ is an $r \times r$ positive definite matrix whose diagonal elements are $\|\nabla L_j(\mathbf{x}_i)\|^2$. So,

$$S^2(\mathbf{x}_i) = \det(\mathscr{L}\mathscr{L}^*(\mathbf{x}_i)) \leqq \prod_{j=1}^{r} \|\nabla L_j(\mathbf{x}_i)\|^2$$

(see [6, p. 9]). Also, $\|\nabla L_j(\mathbf{x}_i)\|^2 \leqq dM^2$, and hence

$$\sum_{j=1}^{r} \|\nabla L_j(\mathbf{x}_i)\|^{-1} \leqq \frac{r(dM)^{r-1}}{S(\mathbf{x}_i)}.$$

This, together with (18), yields

$$m_d(F_i^n \backslash C_i^n) \leqq B_3 [S(\mathbf{x}_i) b_n^{d+1}]^{-1}$$

where B_3 is a constant which is independent of i and of n for $n \geqq n_0$. So, condition (ii) of Theorem 4 holds.

Since E is compact, and since $S(\cdot) > 0$ on E, there is a finite set of points $\{\mathbf{y}_i\}$ in E and a corresponding set of positive numbers $\{\varepsilon_i\}$ so that $E \subset \bigcup_i \mathscr{S}(\mathbf{y}_i, \varepsilon_i)$, and so that for each i, there is a set of indices $\{j_1, \ldots, j_{d-r}\}$ with the property that the mapping

$$(x^{(1)}, \ldots, x^{(d)}) \to (L_1(\mathbf{x}), \ldots, L_r(\mathbf{x}), x^{(j_1)}, \ldots, x^{(j_{d-r})})$$

maps $\mathscr{S}(\mathbf{y}_i, 2\varepsilon_i)$ onto a neighborhood of $(0, \ldots, 0, y_i^{(j_1)}, \ldots, y_i^{(j_{d-k})})$ in a one-to-one fashion. Fix an i, and let $g(\mathbf{x})$ be a continuous function on R^d which is supported by $\mathscr{S}(\mathbf{y}_i, 2\varepsilon_i)$. To simplify the notation, we will assume that $j_1 = r+1, \ldots, j_{d-r} = d$ for this choice of i. Letting $z^{(j)} = L_j(\mathbf{x})$ for $j = 1, \ldots, r$ and transforming the variables of integration, we obtain

$$\left(\frac{b_n}{2}\right)^r \int_{E^n} g(\mathbf{x}) m_d(d\mathbf{x}) \to \int_{-\infty}^{\infty} dx^{(r+1)} \cdots \int_{-\infty}^{\infty} dx^{(d)} g(\mathbf{x}) \left| \frac{\partial(x^{(1)}, \ldots, x^{(r)})}{\partial(z^{(1)}, \ldots, z^{(r)})} \right|$$

where $x^{(j)}$ is evaluated at $(0, \ldots, 0, x^{(r+1)}, \ldots, x^{(d)})$ for $j = 1, \ldots, r$. Now, in $\mathscr{S}(\mathbf{y}_i, 2\varepsilon_i)$, we have the following expression for the surface area of E:

$$m_{d-r}(d\mathbf{x} \cap E) = \left(\sum_{1 \leqq l_1 < \cdots < l_{d-r} \leqq d} \left[\frac{\partial(x^{(l_1)}, \ldots, x^{(l_{d-r})})}{\partial(x^{(r+1)}, \ldots, x^{(d)})} \right]^2 \right)^{1/2} dx^{(r+1)} \cdots dx^{(d)}.$$

Using this, together with the expression for subdeterminants of the inverse of a matrix in terms of the subdeterminants of the matrix (see [6, p. 5]), we obtain

$$(19) \qquad m_{d-r}(d\mathbf{x} \cap E) = S(\mathbf{x}) \left| \frac{\partial(x^{(1)}, \ldots, x^{(r)})}{\partial(z^{(1)}, \ldots, z^{(r)})} \right| dx^{(r+1)} \cdots dx^{(d)}.$$

So, as $n \to \infty$,

$$(20) \qquad \left(\frac{b_n}{2}\right)^r \int_{E^n} g(\mathbf{x}) m_d(d\mathbf{x}) \to \int_E \frac{g(\mathbf{x})}{S(\mathbf{x})} m_{d-r}(d\mathbf{x}).$$

By (16) $E^n \subset \bigcup_i \mathscr{S}(\mathbf{x}_i, \varepsilon_i)$ for large enough n, so it now follows that (20) holds for any bounded continuous function $g(\cdot)$ on R^d. From (19) and condition (ii) of this theorem, we see that $0 < m_{d-r}(E \cap Q) \leqq m_{d-r}(E) < \infty$. So, condition (iv) of Theorem 4 is verified.

To verify condition (v) of Theorem 4, note that applying (20) to the function $g \equiv 1$ yields the fact that $(b_n/2)^r m_d(E^n)$ is bounded in n, since $S(\mathbf{x})$ is bounded away from zero on E. We may now apply Lemma 3. This completes the verification of all the conditions of Theorem 4.

This section has concentrated on the case in which E is very smooth. To show that Theorem 4 can be applied in many other types of cases, we will present the following example, in which E is the ordinary Cantor set.

EXAMPLE. Assume $d_1 = 0$, and define E^n by

$$E^n = \{\mathbf{x} \in R^d \mid x^{(1)} \in [0, 1], x^{(1)} \text{ has a ternary expansion whose first}$$
$$m(n) \text{ terms are 0 or 2, and } |x^{(i)}| \leqq b_n^{-1} \text{ for } i = 2, \ldots, d\},$$

where $m(n)$ is the integer part of $\log_3 b_n$. Suppose $E \cap Q \neq \varnothing$, where

$$E = \bigcap_n E^n = \{\mathbf{x} \in R^d \mid x^{(1)} \text{ is in the Cantor set and } x^{(i)} = 0 \text{ for } i = 2, \ldots, d\}.$$

Then it is easy to see that all the conditions of Theorem 4 are satisfied where $\mu(d\mathbf{x})$ is the $(\log_3 2)$-dimensional Hausdorff measure on E. To check this, it suffices to let $C_k^n = E_k^n = F_k^n$, for $k = 1, \ldots, 2^{m(n)}$, be the $2^{m(n)}$ components of E^n, and use the fact that $1 \leqq b_n 3^{-m(n)} \leqq 3$. Condition (v) will again follow from Lemma 3.

REFERENCES

1. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR **38** #1718.

2. R. M. Blumenthal and R. K. Getoor, *Markov processes and potential theory*, Academic Press, New York, 1968.

3. M. Dwass and S. Karlin, *Conditioned limit theorems*, Ann. Math. Statist. **34** (1963), 1147–1167. MR **29** #4087.

4. H. G. Eggleston, *Convexity*, Cambridge Tracts in Math. and Math. Phys., no. 47, Cambridge Univ. Press, New York, 1958. MR **23** #A2123.

5. D. L. Iglehart, *Weak convergence of probability measures on product spaces with applications to sums of random vectors*, Tech. Rep. No. 109, Operations Research Department, Stanford Univ., Stanford, Calif., 1968.

6. S. Karlin, *Total positivity*. Vol. I, Stanford Univ. Press, Stanford, Calif., 1968. MR **37** #5667.

7. P. Levy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.

8. T. M. Liggett, *An invariance principle for conditioned sums of independent random variables*, J. Math. Mech. **18** (1968), 559–570. MR **38** #6649.

9. ———, *Weak convergence of conditioned sums of independent random vectors*, Thesis, Stanford Univ., Stanford, Calif., 1969.

10. C. Stone, *On local and ratio limit theorems*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), vol. II: Contributions to Probability Theory, part 2, Univ. of California Press, Berkeley, 1967, pp. 217–224. MR **36** #5989.

11. A. E. Taylor, *General theory of functions and integration*, Blaisdell, Waltham, Mass., 1965. MR **31** #2358.

12. B. Trumbo, *Sufficient conditions for the weak convergence of conditional probability distributions in a metric space*, Thesis, Univ. of Chicago, Chicago, 1965.

13. M. Wichura, *On the weak convergence of non-Borel probabilities on a metric space*, Thesis, Columbia Univ., New York, 1968.

UNIVERSITY OF CALIFORNIA,
    LOS ANGELES, CALIFORNIA 90024