*The Annals of Probability*
1981, Vol. 9, No. 5. 831–851

# LIPSCHITZ SMOOTHNESS AND CONVERGENCE WITH APPLICATONS TO THE CENTRAL LIMIT THEOREM FOR SUMMATION PROCESSES

ROY V. ERICKSON

*Michigan State University*

We prove that certain jump summation processes converge in distribution for the uniform topology to the Brownian sheet, while smoothed summation processes converge for various Lipschitz topologies. These results follow after a careful study of abstract, generalized Lipschitz spaces. Along the way we affirm a conjecture about smoothness and continuity of processes defined on $[0, 1]^d$.

**1. Introduction, motivation and summary.** This paper generalizes Donsker's invariance principle in two respects. First, it considers summation processes whose summands are indexed by $d$-tuples of integers, so that convergence is to the so-called Brownian sheet process. Secondly, the metric spaces on which these processes reside are given the finest (largest) topology compatible with our proofs of convergence in distribution. In order to accomplish these generalizations, we have looked closely at conditions which imply smoothness of stochastic processes with "time" set $\{T, d\}$ a totally bounded semimetric space. Our results in this direction extend the classical techniques of Kolmogorov, Loève (1963) and Delporte (1964) in a manner suggested by Dudley (1973). They do not include all of those found in Fernique (1974, 1978), Garsia and Rodemich (1974) and Nanopoulos and Nobelis (1978); but see our discussion following theorems (5.2) and (6.1). While doing all of this, we also confirm a conjecture of Garsia and Rodemich (1974) concerning processes on $T = [0, 1]^d$; this is done in the spirit of Hahn and Klass (1977).

To motivate this study, let us look at Donsker's theorem and previous extensions. Let $\xi_1, \xi_2, \cdots$ be iid $\xi$ with $E\xi = 0$, $E\xi^2 = 1$ and define the *summation processes*, for $0 \le t \le 1$,

(jump) $$Z_n(t) := \textstyle\sum_1^n \xi_k I(k/n \le t)/n^{1/2}$$

(smooth) $$X_n(t) := \textstyle\sum_1^n \xi_k \mid [(k-1)/n, k/n] \cap [0, t] \mid n^{1/2},$$

where $|A|$ denotes the (one dimensional) Lebesgue measure of measurable $A \subset [0, 1]$. Also, let $W$ denote the Wiener process.

These processes each reside in different, natural function spaces, and each such space can be given many different topologies. Let us tabulate these:

| (space in decreasing size) | topologies | processes |
|---|---|---|
| $D[0, 1]$ | Skorohod, uniform $=: \\| \ \\|$ | $Z_n$ |
| $C[0, 1]$ | $\\| \ \\|$ | |
| $\mathrm{Lip}_\alpha$, $\alpha < ½$ | Lip norm $=: \\| \ \\|_\beta$, $\beta \le \alpha$ | $W$ |
| $\mathrm{Lip}_1$ | $\\| \ \\|_\beta$, $\beta \le 1$ | $X_n$ |

It is usual to prove $Z_n \to_{\text{dist'n}} W$ for the Skorohod topology on $D[0, 1]$ and $X_n \to_{\text{dist'n}} W$ for the uniform topology on $C[0, 1]$. But more can be proven sometimes. Billingsley (1968, page 152) notes that $Z_n$ converges on $D[0, 1]$ with $\| \ \|$ topology in general. On the other

Received July 25, 1979; revised July 14, 1980.

*AMS* 1980 *subject classifications.* Primary 60B10, 60G17; secondary 60G50.

*Key words and phrases.* Lipschitz spaces, weak convergence, central limit theorem, summation processes.

hand, Lamperti (1962) proves that, under certain moment conditions, $X_n$ converges on $\mathrm{Lip}_\alpha$, $\| \|_\beta$, $\beta < \alpha$. Precisely, for $p > 2$, Lamperti shows

(a) $E|\xi|^p < \infty$ implies $X_n \to_{\text{dist'n}} W$ in $\mathrm{Lip}_\alpha$, $\| \|_\beta$ for all $\beta < \alpha := (p-2)/2p$, while
(b) if $P(|\xi| > x) = x^{-p}$, $x \ge 1$, the convergence in (a) cannot hold for $(p-2)/2p < \beta < \alpha < ½$, even though all the $X_n$ are in $\mathrm{Lip}_1$ and $W$ is in all $\mathrm{Lip}_\gamma$, $\gamma < ½$.

Because of Levy's modulus result (see Dudley (1973, page 75)), which states

$$\limsup_{\delta := |s-t| \downarrow 0} |W(s) - W(t)|/(2\delta\,|\log \delta|)^{1/2} = 1 \qquad \text{a.s.,}$$

we see that $W$ lives on a generalized Lipschitz space $L$, $\mathrm{Lip}_{1/2} \overset{\subset}{\neq} L \overset{\subset}{\neq} \bigcap_{\alpha<1/2} \mathrm{Lip}$. A natural question that arises from Lamperti's work is this: if $\xi$ is bounded (or Gaussian), is there a topology $\tau$ on $L$, finer than all of those given by $\| \|_\beta$, $\beta < ½$, so that $X_n \to_{\text{dist'n}} W$ as processes on $L$, $\tau$?

It was this question that stimulated our work. As we answered the above problem, it became apparent that our methods were applicable more generally. And because of the results of Wichura (1969) for certain multiparameter processes, we began to investigate general summation processes and their convergence to a Brownian sheet.

To further motivate our study let us define fairly general summation processes and mention some of the results we obtain. To do so necessitates some preliminary definitions which we give as well.

DEFINITIONS 1.1. Fix dimension $d$ and let $T_{\mathscr{B}} :=$ Borel subsets of $\mathbf{R}^d$ for which $|A| < \infty$, where $|A| :=$ Lebesgue $d$-dimensional measure of $A \in T_{\mathscr{B}}$. We give $T_{\mathscr{B}}$ one of two semimetrics: (Lebesgue)

$$d_L(A, B) := \|I_A - I_B\|_1 = |A \,\Delta\, B|$$

or (Hausdorff)

$$d_H(A, B) := \inf\{\varepsilon > 0; A \subset B^\varepsilon, B \subset A^\varepsilon\}.$$

For the above, $I_A = I(A)$ is the indicator function for $A$, $A \Delta B := A \backslash B \cup B \backslash A$ and

$$A^\varepsilon := \bigcup \{B(x, \varepsilon); x \in A\}$$

where $B(x, \varepsilon) := \{y \in \mathbf{R}^d, \|y - x\| < \varepsilon\}$. ($\| \|$ denotes the supremum norm.)

DEFINITION 1.2. A process $W = \{W(A), A \in T_{\mathscr{B}}\}$ is called a Brownian sheet process if each $W(A)$ is Gaussian and $EW(A) = 0$, $EW(A)W(B) = |A \cap B|$.

One would expect that the Brownian sheet arises naturally as the limit of certain *summation* processes.

DEFINITIONS 1.3. Given any random variable $\xi$ with $E\xi = 0$, $E\xi^2 = 1$, build $\{\xi_j, j \in \mathbf{N}^d\}$ iid $\xi$, where $\mathbf{N}^d := \{j = (j_1, \cdots, j_d); j_i = 1, 2, \cdots,\}$. For $j, n \in \mathbf{N}^d$ let $j/n := (j_1/n_1, \cdots, j_d/n_d)$, write $j \le n$ iff $j_i \le n_i$ for $i = 1, \cdots d$, and set

$$R_{n,j} := \{x \in \mathbf{R}^d; (j - \mathbf{1})/n \le x \le j/n\}$$

where $\mathbf{1} = (1, \cdots, 1) \in \mathbf{N}^d$. As usual, let $\|x\| := \max\{|x_i|, i = 1, \cdots, d\}$. Now define the summation processes: for $A \in T_{\mathscr{B}}$

(jump) $$Z_n(A) := \textstyle\sum_{j \le n} |R_{n,j}|^{1/2} I(j/n \in A)\xi_j,$$

(smoothed) $$X_n(A) := \textstyle\sum_{j \le n} |R_{n,j}|^{-1/2} |R_{n,j} \cap A| \xi_j.$$

Obviously each $X_n$ is continuous as a process on $T_1 := T_{\mathscr{B}} \cap [0, 1]^d$, $d_L$. But $Z_n$ and $X_n$ can be very far apart and need not converge to $W$: consider their values for $A$ consisting of only rational or only irrational points. On the other hand, if we restrict attention to

$$T_0 := \{A \in T_1, |\operatorname{bdry} A| = 0\}$$

and write $X_n | T_0 := \{X_n(A), A \in T_0\}$, we show

THEOREM 7.4. *Both $X_n | T_0$ and $Z_n | T_0$ converge in finite dimensional distribution to $W | T_0$ as $\| 1/n \| \to 0$.*

Our goal is to extend finite dimensional convergence to convergence in distribution for both jump processes $Z_n | T$ under the uniform topology and smooth processes $X_n | T$ under Lipschitz topologies, where $T \subset T_0$ is suitably chosen. This we do in Sections 7 and 8. While Wichura deals only with rectangles $[0, t] \subset [0, 1]^d$, we treat general polytopes and even some classes of sets with smooth boundaries.

Roughly, the technique used is as follows. We extend classical conditions, which imply Lipschitz smooth paths, to such processes as $X_n | T$; then we note that these conditions imply tightness for the induced measures on certain generalized Lipschitz spaces with quite specific topologies. What makes this most interesting is the interplay between tail conditions on $\xi$ and the size of $T \subset T_0$ as measured by a certain type of metric entropy. To treat the jump processes we find conditions which imply $\| X_n | T - Z_n | T \| \to 0$ in probability.

To carry out the above plan, it seems best to study general Lipschitz spaces, over arbitrary totally bound sets $T$, and only later to deal specifically with summation processes.

Here is an outline of what we do.

In Section 2 we define generalized Lipschitz spaces and identify certain compact sets. Section 3 specifies certain subsets of these Lipschitz spaces, membership in which can be checked by observing a function (or process) on a countable dense set. These two subsets arise naturally from the two separate techniques used to show smoothness of processes, one due to Kolmogorov-Loève and another due to Delporte. These techniques, generalized in the spirit of Dudley (1973), are presented in Section 4. The next section gives specific examples and applications of the preceding general theorem. Section 6 deals with processes defined on $T = [0, 1]^d$ and shows that the above results on smoothness cannot be much improved. Certain trigonometric examples are constructed to verify a conjecture of Garsia and Rodemich (1974); the methods used follow those of Hahn (1977) and Hahn and Klass (1977). Finally, Sections 7 and 8 treat, respectively, the Lipschitz (uniform) convergence in distribution of the smooth (jump) summation processes to the Brownian sheet for fairly general classes $T \subset T_{\mathscr{B}}$.

**2. Generalized Lipschitz Spaces.** Let $T$, $d$ be a semimetric space and let $\mathbf{C} = \mathbf{C}(T, \mathbf{R})$ be the set of real valued continuous functions from $T$ to $\mathbf{R}$. The usual Lipschitz spaces are subsets of $\mathbf{C}$ which have certain oscillation properties. To define generalized Lipschitz spaces, we first define a family of comparison functions

$$\mathbf{G} := \{g\colon (0, 1] \to (0, 1];\ \text{increasing}\}.$$

For $x \in \mathbf{C}$, $g \in \mathbf{G}$ and $0 < r \le 1$ set

$$g\text{-osc}(x, r) := \sup\{|x(s) - x(t)|/g \circ d(s, t);\ 0 < d(s, t) \le r,\ s, t \in T\},$$

$$_g\| x \| := \| x \| + g\text{-osc}(x, 1), \qquad \| x \| := \sup\{|x(t)|;\ t \in T\}.$$

Now define

$$g\text{-Lip} := g\text{-Lip}(T) := \{x \in \mathbf{C};\ _g\| x \| < \infty\}.$$

To compare various Lip spaces, notice that for $f, g \in \mathbf{G}$, $0 < r \le 1$ and $x \in \mathbf{C}$, we have

$$\begin{aligned} _f\| x \| &\le \| x \| (1 + 2/f(r)) + \| (g/f) I_{[0,r]} \| g\text{-osc}(x, r) \\ &\le (1 + \| (g/f) I_{[0,r]} \| + 2/f(r))\, _g\| x \|. \end{aligned} \tag{2.1}$$

Note first, by taking $g = f$, that admission to $g$-Lip depends only on $g$ near zero. For further comparison make the usual definitions

$$g = 0(f) \quad \text{at} \quad 0 \quad \text{or} \quad g = o(f) \quad \text{at} \quad 0$$

if $\|(g/f)I_{[0,r]}\| \downarrow a < \infty$ or 0, respectively, as $r \downarrow 0$. We omit "at 0" when no confusion can arise. Finally, define the ${}_g\| \ \|$ discs

$${}_gB_a := \{x \in \mathbf{C}, {}_g\|x\| \leq a\}, \qquad 0 \leq a < \infty.$$

THEOREM 2.2. *Let* $f, g \in \mathbf{G}$.

(a) $g$-Lip *is a Banach space with norm* ${}_g\| \ \|$; ${}_gB_a$ *is closed in* $g$-Lip *and complete for each of the* $\| \ \|$ *and* ${}_g\| \ \|$ *topologies.*

(b) $g = 0(f)$ *at* 0 *implies* $g$-Lip $\subset$ $f$-Lip; *and* $g = o(f)$ *at* 0 *implies that* ${}_f\| \ \|$ *and* $\| \ \|$ *give the same topology on each* ${}_gB_a$.

(c) $g = 0(f)$ *and* $f = 0(g)$ *imply that* $g$-Lip = $f$-Lip *and that* ${}_g\| \ \|$ *and* ${}_f\| \ \|$ *are equivalent norms.*

(d) *If* $T$ *is totally bounded and* $g = o(f)$, *then each* ${}_gB_a$ *is compact in the* ${}_f\| \ \|$ *topology; and* $g$-Lip *is a* $\sigma$-*compact, separable linear space with norm* ${}_f\| \ \|$.

REMARK. We leave it to the reader to construct examples in case $T = [0, 1]$, with the usual metric, and $g(t) := |t| =: f^2(t)$, $t \geq 0$ which show that

(e) *g-Lip is not separable in the* ${}_g\| \ \|$ *topology,*

(f) $\| \ \|$ *and* ${}_f\| \ \|$ *induce different topologies on g-Lip and g-Lip is complete in neither.*

PROOF. (a) If $\{x_n\} \subset g$-Lip is ${}_g\| \ \|$ Cauchy, then $\{x_n\} \subset {}_gB_a$, for some $a < \infty$, and $\{x_n\}$ is $\| \ \|$ Cauchy. Under only these last two conditions, for each $t \in T$, $\{x_n(t)\}$ is clearly Cauchy and $x := \lim x_n$ is in ${}_gB_a$: for all $t$ and $s$ with $0 < d(s, t) \leq 1$,

$$|x(s) - x(t)\|/g \circ d(s, t) \leq \limsup g\text{-osc}(x_n, 1), \quad \text{and}$$

$$|\|x\| - \|x_k\|| \leq \limsup_n \|x_n - x_k\| \to 0 \quad \text{as } k \to \infty.$$

This gives ${}_g\|x\| \leq \limsup_n {}_g\|x_n\|$, and (a) follows.

(b) and (c) follow easily from (2.1).

(d) When $T$ is totally bounded, we can show that the compact open and $\| \ \|$ topologies coincide on ${}_gB_a$, and the result follows from Ascoli's theorem (Wilansky (1970, Theorem 13.3.4)). It is just as easy to argue directly. Let $\varepsilon_i \downarrow 0$ and pick finite $\varepsilon_i$-nets $S_i$, $i = 1, 2, \cdots$, $S := \cup_1^\infty S_i$. (See the second paragraph of Section 3 if a definition is needed.) Given $\{x_n\} \subset {}_gB_a$, by a diagonal argument we may assume $x_n(s)$ converges for each $s \in S$. But then, for every $t \in T$ and $i = 1, 2, \cdots$, we see

$$|x_n(t) - x_m(t)| < 2a\|(g/f)I_{[0,\varepsilon_i]}\|\|f\| + \|(x_n - x_m)I_S\|,$$

by approximating $t$ with an element from $S_i$. This shows that $\{x_n\}$ is $\| \ \|$-Cauchy. Hence, by parts (a) and (b), every sequence in ${}_gB_a$ has a $\| \ \|$, so also ${}_f\| \ \|$, convergent subsequence. Thus ${}_gB_a$ is sequentially compact, and therefore compact (see Wilansky (1970, Theorem 7.2.1)). The final result is clear, because a compact metric space is complete and separable (Wilansky (1970, Theorem 11.3.7)). □

**3. Two subsets of $g$-Lip.** In this section we will give two techniques for showing that a function is in $g$-Lip: one due to Kolmogorov and Loève, the other to Delporte, and both extended from the case $T = [0, 1]$ to totally bounded $T$ by a method suggested by Dudley.

We will begin by recalling that a set $S$ contained in a semimetric space $T$, $d$ is an $\varepsilon$-*net* if for each $t \in T$ there is an $s \in S$ such that $d(s, t) < \varepsilon$. Further, $T$ is *totally bounded* if $T$ has a finite $\varepsilon$-net for every $\varepsilon > 0$.

We do not need to consider $\varepsilon$-nets for every $\varepsilon > 0$. Instead we assume given and fixed a sequence $\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots)$ with $1 = \varepsilon_1 > \varepsilon_2 \cdots \downarrow 0$ and a totally bounded $T$. Now choose and fix the sets and maps

$$\begin{array}{lll} & \text{finite } \varepsilon_n\text{-net:} & S_n \\ (3.1) & \text{selector maps:} & \sigma_n: T \to S_n \quad \text{such that} \quad d(\sigma_n(t), t) \le \varepsilon_n \\ & \text{neighboring pairs:} & T_n := \{(s, t) \in S_n \times S_{n+1}; d(s, t) \le 3\varepsilon_n\}. \end{array}$$

Let us take a second to see how these concepts might help. Note that $x \in \mathbf{C}$ implies

$$x(t) - x \circ \sigma_n(t) = \textstyle\sum_{k \ge n} [x \circ \sigma_{k+1}(t) - x \circ \sigma_k(t)].$$

Because $d(s, t) \le \varepsilon_n$ implies $(\sigma_n(s), \sigma_{n+1}(t)) \in T_n$, it follows for such $s$, $t$ and for $x \in \mathbf{C}$:

$$|x(s) - x(t)| \le 2 \textstyle\sum_{k \ge n} M_k(x) \tag{3.2}$$

where

$$M_k(x) := \max\{|x(s) - x(t)|; (s, t) \in T_k\}.$$

As simple as (3.2) is, it is the key to both the Kolmogorov-Loève and Delporte techniques. Specifically, choose and fix some $g \in \mathbf{G}$ and define

$$\begin{array}{lll} & \text{nodes} & g_n := g(\varepsilon_{n+1}+), \\ (3.3) & \text{jumps} & \delta g_n := g_n - g_{n+1} \\ & \text{minorant} & g^* = \sum_1^\infty I(\varepsilon_{n+1}, \varepsilon_n] g_n. \end{array}$$

Note that $g^*$ is a right continuous step function just majorized by $g$, and that ${}_{g^*}\|\ \| \ge {}_g\|\ \|$, $g^*$-Lip $\subset g$-Lip.

Loève (1963) and Kolmogorov (see Delporte (1964, page 178, footnote 6) use a Borel-Cantelli argument to bound $\sum_{k \ge n} M_k(x)$ by the introduction of

$$C_n := \{x \in \mathbf{C}; M_k(x) \le \delta g_k, k \ge n\}. \tag{3.4.a}$$

But Delporte (1964) uses a different method and defines

$$A_a := \{x \in \mathbf{C}; G(x) \le a\} \uparrow A_\infty \tag{3.4.b}$$

where

$$G(x) := \textstyle\sum_1^\infty M_k(x)/g_k.$$

The above sets are useful in bounding the oscillation of a function in $\mathbf{C}$; to get control of its maximal absolute value, introduce

$$\begin{array}{ll} (3.4.c) & B_{n,b} := \{x \in C; |x(s)| \le b, \text{ all } s \in \cup_1^n S_k\} \uparrow B_{n,\infty}, \\ & B_b := \cap_1^\infty B_{n,b} \uparrow B_\infty. \end{array}$$

From (3.2) and (3.4) we see that

$$\begin{array}{ll} & g^*\text{-osc}(x, \varepsilon_n) \le 2 \quad \text{if} \quad x \in C_n, \\ (3.5) & g^*\text{-osc}(x, 1) \le 2G(x) \le 2a \quad \text{if} \quad x \in A_a \\ & \|x\| \le \sum_{k \ge n} M_k(x) + b \quad \text{if} \quad x \in B_{n,b}. \end{array}$$

Combining these we obtain

$$B_{n,b} \cap C_n \subset B_{1+b} \cap C_n \subset {}_{g^*}B_\beta \subset {}_gB_\beta \tag{3.6.a}$$

for $\beta := (1 + b)(1 + 2/g_n) + 2$, while

$$A_a \cap B_{n,b} \subset A_a \cap B_{2a+b} \subset {}_{g^*}B_\alpha \subset {}_gB_\alpha \tag{3.6.b}$$

for $\alpha := 3a + b$.

We now have the compact sets which are the key for our proof of convergence in distribution.

THEOREM 3.7. *Fix $\varepsilon$, totally bounded $T$, the sets and maps of* (3.1), *a $g \in \mathbf{G}$ and the notation of* (3.4). *Given any $f \in \mathbf{G}$ with $g = o(f)$ at* 0, *we have $B_{n,b} \cap C_n$ and $A_a \cap B_{n,b}$ compact in $g$-Lip with ${}_f\|\ \|$ topology.*

PROOF. $A_a$, $B_{n,b}$ and $C_n$ are closed in the $\|\ \|$ topology. □

**4. Basic smoothness and tightness results.** Throughout this section assume given and fixed $\varepsilon$, with $1 = \varepsilon_1 > \varepsilon_2 > \cdots \downarrow 0$, totally bounded $T$, $d$ and $g \in \mathbf{G}$. Associated with these and also fixed we need for $T$, $d$:

finite $\varepsilon_n$-nets $S_n$, selector maps $\sigma_n$ and neighboring pairs $T_n$ (3.1),

and for $g$:

nodes $g_n$, jumps $\delta g_n$ and minorant $g^*$ (3.3).

Since we will be constructing measures on $g$-Lip $= g$-Lip$(T, d)$, we need certain $\sigma$-algebras. For each $f \in \mathbf{G}$, with $g = 0(f)$ at 0, $g$-Lip can be given norm ${}_f\|\ \|$; and each such norm leads to the $\sigma$-algebras generated by the balls and opens, which we denote by ${}_f\underline{\text{Ball}}$ and ${}_f\underline{\text{Borel}}$, respectively. In addition, let ${}_g$Cyl denote the $\sigma$-algebra on $g$-Lip generated by cylinders. Then it is easy to check, noting Theorem (2.2.d):

$$\text{if} \quad g = o(f) \quad \text{at} \quad 0, \qquad \text{then on} \quad g\text{-Lip}$$

$${}_g\text{Cyl} = {}_f\text{Ball} = {}_f\text{Borel} = {}_g\text{Ball} \subset {}_g\text{Borel}.$$

Given a real valued process $X = \{X(t), t \in T\} = \{X(t), t \in T, \Omega, \mathscr{F}, P\}$, our goal is to find conditions which imply $X$ has a version in $g$-Lip, ${}_g$Ball. As our starting point we always assume that $X$ is *continuous in probability*;

$$X \circ \sigma_n(t) \to X(t) \text{ in probability, each } t \in T.$$

We then introduce conditions which relate the size of $T$, as measured by

$$\nu_n := \text{card } T_n,$$

and the variation of $X$, as measured by either

$$p_n \geq p_n(X) := \max\{P(|X(s) - X(t)| > \delta g_n);\ (s, t) \in T_n\}$$

or

$$m_n \geq m_n(X) := \max\{(E|X(s) - X(t)|^p/g_n^p)^{\gamma;}\ (s, t) \in T_n\},$$

where $0 < p < \infty$ and $\gamma := 1 \wedge (1/p)$, $\wedge :=$ minimum.

Introduce the conditions on $X$:

$$(K\text{–}L) \qquad \sum \nu_k p_k < \infty \qquad \text{or}$$

$$(D_p) \qquad \sum \nu_k^\gamma m_k < \infty, \qquad 0 < p < \infty \quad \gamma := 1 \wedge (1/p).$$

Of course, when restricted to $T = [0, 1]$, the $(KL)$ condition is of the type applied by Kolmogorov and Loève, while $(\mathbf{D}_p)$ is used by Delporte.

THEOREM 4.1. *Given $\varepsilon$, $T$, $f$, $g \in \mathbf{G}$ with $g = o(f)$ at* 0, *a process $X$ continuous in probability and the notation above.*

(*a*) *If $X$ has $(K - L)$, then $X$ has a version on $g$-Lip, ${}_g$Ball.*

(b) *If X has* $(D_p)$, *then X has a version in g*-Lip, ${}_g$Ball; *and, in addition, there is* a r.v. $G(X)$, *function G defined below, so that* $E\,|\,G(X)\,|^p < \infty$ *and*

$$|\,X(s) - X(t)\,| \le 2G(X)g \circ d(s, t), \qquad all \quad s,t \in T \quad \text{a.s.}$$

COMMENTS: (a) the conditions $(D_p)$ and $(K - L)$ seem not to be comparable. Specifically, under the conditions of Theorem 5.1.b, $(K - L)$ seems to give $X$ a version in a slightly smaller space than does $(D_p)$. On the other hand, specializing (5.1.a) to $T = [0, 1]$, we see that $(D_p)$ gives $X$ a version in a "Log-Lip" space, but $(K - L)$ may not even imply continuity (c.f. Loève (1963, page 519, corollary)).

(b) In the Gaussian case, Dudley (1973, Theorem 2.1) is able to replace the neighboring pairs $T_n$ by the sets $S_n \times S_{n+1}$. If we were to make this replacement in the non-Gaussian case, we would not even get the classical results on [0, 1].

Proof of (4.1). Introduce the random variables

$$M_n := \max\{\,|\,X(s) - X(t)\,|;\ (s, t) \in T_n\}$$

and the events

$$C_n := \{M_k \le \delta g_k;\ k \ge n\} \uparrow C_\infty, \qquad A_a := [G \le a] \uparrow A_\infty,$$

where $G := G(X) := \sum_1^\infty M_k/g_k$.

(a) Now assume $\sum \nu_k p_k < \infty$. Since $P(M_k > \delta g_k) \le \nu_k p_k$, we have

$$P(C_\infty) \ge P(C_n) \ge 1 - \textstyle\sum_{k\ge n} \nu_k p_k \uparrow 1.$$

On $C_n$, $|\,X \circ \sigma_n(t) - X \circ \sigma_{n+j}(t)\,| \le \sum_{k\ge n} M_k \le \sum_{k\ge n} \delta g_k$, forcing $I(C_\infty)X \circ \sigma_n(t)$ to be Cauchy everywhere and thus convergent to $Y(t)$, say, for each $t \in T$. By assumption $X \circ \sigma_n(t) \to X(t)$ in probability, and we see that $X(t) = Y(t)$ a.s., each $t$. Set $S := \cup_1^\infty S_n$. Then $\Omega^* := C_\infty \cap \bigcap_{s\in S}\{X(s) = Y(s)\}$ has $P\Omega^* = 1$, and $\Omega^*$ is in the $\sigma$-algebra $\sigma\{X(s), s \in S\}$. In addition

$$Y(t) - Y \circ \sigma_n(t) = \textstyle\sum_{k\ge n}[X \circ \sigma_{k+1}(t) - X \circ \sigma_k(t)] \qquad \text{on} \quad \Omega^*,$$

and the arguments for (3.5) and (3.6.a) imply

$$g\text{-osc}(Y, \varepsilon_n) \le 2 \quad \text{on} \quad \Omega^* \cap C_n \qquad \text{and} \qquad {}_g\|\,Y\,\| \le (1 + b)(1 + 2/g_n) + 2 \quad \text{on} \quad \Omega^*_{n,b},$$

where $\Omega^*_{n,b} := \Omega^* \cap C_n \cap B_{n,b}$, $B_{n,b} := \{\,|\,X(s)\,| \le b;\ s \in \cup_1^n S_k\}$. As in (3.4), let $B_b := \cap_1^\infty B_{n,b} \uparrow B_\infty$. For $\Omega^{**} := \Omega^* \cap B_\infty$ we have, as in (3.6.a),

$$P(\Omega^{**}) \ge P(\Omega^* \cap C_n \cap B_{1+b}) \ge P(\Omega^*_{n,b}) \ge P(C_n) + P(B_{n,b}) - 1 \to 1$$

if $b \to \infty$ and then $n \to \infty$. Now let $\lambda$ be the mapping which identifies each $\omega \in \Omega^{**}$ with the function $\omega : t \to Y(t, \omega)$. The version of $X$ that we seek is given by the measure space ($g$-Lip, ${}_g$Ball, $P_X := P \circ \lambda^{-1}$).

(b) For $1 \le p < \infty$, the inequality $M_n^p \le \sum_{T_n} |\,X(s) - X(t)\,|^p$ implies $\|\,M_n\,\|_p \le \nu_n^{1/p} m_n g_n$, so that $\|\,G\,\|_p \le \sum_1^\infty \nu_k^{1/p} m_k < \infty$ and $P(A_\infty) \ge P(A_a) \ge 1 - \|\,G\,\|_p^p/a^p \uparrow 1$ if (b) holds. For $0 < p < 1$, $E\,|\,G\,|^p \le \sum_1^\infty \nu_k m_k$. The rest follows as above and from (3.5). □

Let us now turn our attention to conditions for tightness. Suppose given a family $\chi = \{X\}$ of processes on $T$, each continuous in probability, and now choose

$$p_n \ge \sup\{p_n(X), X \in \chi\}, \qquad m_n \ge \sup\{m_n(X), X \in \chi\}.$$

If $\sum \nu_n p_n < \infty$, then each $X$ can be realized as ($g$-Lip, ${}_g$Ball, $P_X$), and we think of the $C_n$ and $B_{n,b}$ of (4.1) as subsets of $g$-Lip. For tightness of the family $\mathscr{P}_\chi = \{P_X, X \in \chi\}$ as measures on $g$-Lip with any $f$ for which $g = o(f)$ at 0, by (3.6.a) it suffices to show that for all $\eta > 0$

$$\inf_\chi P_X(C_n \cap B_{n,b}) \ge 1 - \eta$$

if $n$ and $b$ are large. We already have one uniform bound: $P_X(C_n) \ge 1 - \sum_{k\ge n} \nu_k p_k$. The other is obtained by assuming $\chi$ is *ε-net bounded*: for all $\eta > 0$ there is a $b = b(n, \eta)$ such that $\inf_\chi P_X(B_{n,b}) = \inf_\chi P(\,|\,X(s)\,| \le b;\ s \in \cup_1^n S_k) \ge 1 - \eta$.

For applications it is important to notice that any finite $\chi$ is $\varepsilon$-net bounded, and also that if $\chi = \{X_0, X_1, \cdots\}$ and $X_n \to X_0$ in finite dimensional distribution, then $\chi$ is $\varepsilon$-net bounded.

THEOREM 4.2. *Given $\varepsilon$, $T$, $f$, $g \in G$ with $g = o(f)$ at 0, the notation above, and a family $\chi = \{X\}$ of processes such that each $X$ is continuous in probability and $\chi$ is $\varepsilon$-net bounded. Either of the conditions*

$$(K - L)_\chi : \sum \nu_k p_k < \infty \quad \text{or} \quad (D_p)_\chi : \sum \nu_k^\gamma m_k < \infty, \qquad \gamma := 1 \wedge (1/p),$$

*implies $\mathscr{P}_\chi$ is tight on $g$-Lip with ${}_f\| \ \|$ topology and $\sigma$-algebra ${}_f$Borels.*

**5. Applications of the smoothness theorem.** In this section, until the end of Theorem 5.2, the $\varepsilon$ sequence is fixed as $\varepsilon_n := 2^{-n+1}$, and the size of the totally bounded space $T$ is specified by certain exponents of metric entropy, $r_k(T)$.

To be specific, let

$$L_0(s) = |s|, \qquad L_k = |\log L_{k-1}|, \qquad k = 1, 2.$$

For $k = 1, 2$, write

$$r_k(T) \leq r$$

if for some constant $K$, sets $T_n$ can be found, as in (3.1), such that

$$L_k(\nu_n) \leq K + rL_1(\varepsilon_n) \quad \text{eventually},$$

where again $\nu_n :=$ card $T_n$. To avoid trivialities, assume $\nu_n \to \infty$.

For comparison, recall that Dudley (1973, page 70) defined the *exponent of metric entropy of $T$* as

$$r(T) := \limsup_{\varepsilon \to 0} L_2(\text{card } S(\varepsilon))/L_1(\varepsilon),$$

where each $S(\varepsilon)$ is an $\varepsilon$-net for $T$ of minimal cardinality. By choosing $S_n = S(\varepsilon_n)$ in (3.1) and observing that card $S_n \leq \nu_n \leq (\text{card } S_{n+1})^2$, it is clear that

$$r_2(T) \leq r \quad \text{implies} \quad r(T) \leq r \qquad \text{and} \qquad r_2(T) \leq r(T) + \delta, \quad \text{all } \delta > 0.$$

In addition, it should be verified that

$$r_1(T) \leq r < \infty \quad \text{implies} \quad r_2(T) \leq \delta, \quad \text{all } \delta > 0, \qquad \text{and} \qquad r_1([0, 1]^d) \leq d.$$

One convention is needed before the theorem is stated. If $h : A \subset (0, \infty] \to (0, \infty]$ is increasing on some interval $(0, a) \subset A$, $a$ unspecified, let

$$g := \mathbf{G}(h) := h(\cdot \wedge b) \wedge 1,$$

where $0 < b \leq a$ is chosen so that $g \in \mathbf{G}$. Because $g$-Lip depends on $g$ only near zero, any such choice of $b$ will give the same Lipschitz space.

THEOREM 5.1. *Let $X$ be a process on a totally bounded space $T$ with $r_1(T) \leq r$. Below, $K$ is a finite constant, $\delta := d(s, t)$ and the hypothesized inequalities are assumed valid for all $s, t \in T$.*

(*a*) *$E|X(s) - X(t)|^p \leq K\delta^r/L_1^q(\delta)L_2^v(\delta)$ implies $X$ has a version in $g$-*Lip *for $g^{\gamma p} := \mathbf{G}(L_1^\alpha L_2^\beta)$ if $p > r$, $\alpha := 1 - q\gamma \leq 0$ and if $\beta$ can be chosen so that $\beta > 1 - v\gamma$ with $\beta < 0$ when $\alpha = 0$. ($\gamma := 1 \wedge (1/p)$). In addition, for this version we have*

$$|X(s) - X(t)| \leq 2G(X)g \circ d(s, t) \qquad \text{for all} \quad s, t \in T \quad \text{a.s.},$$

*where $E|G(X)|^p < \infty$.*

(*b*) *$E|X(s) - X(t)|^p \leq K(\delta)^{r+q}$, $p, q > 0$ implies $X$ has a version in $g$-*Lip $\subset$ Lip$_\alpha$ *and in $f$-*Lip $\subset$ Lip$_\alpha$ *for each $\alpha < q/p$, where $g := \mathbf{G}(L_0^{q/p}[L_1 L_2^\beta]^{1/p})$ and $f := \mathbf{G}(L_0^{q/p}[L_1 L_2^\beta]^{1 \vee 1/p})$,* *any* $\beta > 1$. ($\vee := \max$.) *In addition, for this version we have*

$$|X(s) - X(t)| \le 2G(X) f \circ d(s, t) \qquad all \quad s, t \in T, \quad \text{a.s.}$$

*where* $E|G(X)|^p < \infty$.

(c) $P(|X(s) - X(t)| > a) \le K \exp(-a^q/K\delta^w)$ *for all* $a > 0$ *and some* $q, w > 0$ *implies* $X$ *has a version in* $g$-Lip *for* $g^q := \mathrm{G}(bL_0^w L_1)$, $b$ *sufficiently large.*

PROOF. Clearly, each hypothesis implies $X$ is continuous in probability. Note that $r_1(T) \le r$ implies $\nu_k \le K_1 2^{kr}$ eventually. In the statements below, all $n$ and $k$ are sufficiently large, and $b_k \backsim c_k$ means $b_k/c_k \to 1$ as $k \to \infty$.

In part (a) we have $h_k^{\gamma p} := g^{\gamma p}(2^{-k}) \backsim K_2 k^\alpha(\log k)^\beta$, and the hypothesis implies that $m_k$ can be chosen to be $K_3(2^{-kr}/h_k^p k^q \log^v k)^\gamma$. Thus $\sum_{k\ge n} \nu_k^\gamma m_k \le K_4 \sum_{k \ge n} k^{-1}(\log k)^{-\beta - v\gamma} < \infty$, and $(D_p)$ holds. Now apply (4.1.b).

In part (b) we have $h_k^p \backsim K_2 2^{-kq} k(\log k)^\beta$. Thus $h_{k+1}/h_k \to 2^{-q/p}$, and we can find $0 < c < 1$ so that $a_k := h_k - h_{k+1} > ch_k$. The Markov inequality gives

$$p_k(X) \le \max_{T_k} E|X(s) - X(t)|^p/(ch_k)^p \le K_3 2^{-kr}/k \log^\beta k =: p_k,$$

and $\sum p_k \nu_k < \infty$. This gives the $g$-Lip result. For the $f$-Lip case, we note that

$$\nu_k^\gamma m_k(X) \le K_4[2^{kr} \cdot 2^{-k(r+q)}/f_n^p]^\gamma \backsim K/k(\log k)^\beta$$

since $f_n^p := [f(\varepsilon_n)]^p \backsim K_6 2^{-kq}[k(\log k)^\beta]^{1/\gamma}$.

In part (c) we have $h_k^{q'} = K_2 b 2^{-kw} k$. Thus $h_{k+1}/h_k \to 2^{-w/q}$ and again there is $0 < c < 1$ such that $a_k := h_k - h_{k+1} > ch_k$. Hence

$$p_k(X) \le K \exp\{-bK_3 k \log 2\} =: p_k(b)$$

and $\sum \nu_k p_k(b) < \infty$ if $b > r/K_3$. □

Actually, the inequality in (5.1.c) suffices to imply the existence of a smooth version of $X$ on a much larger domain $T$. This will be useful in the study of Brownian processes.

THEOREM 5.2. *Let $X$ be a process on $T$ and make all the assumptions of* (5.1.c) *except assume* $r_2(T) \le r < w$. *If* $q, v > 0$ *then* $X$ *has a version in* $g$-Lip *for* $g := \mathrm{G}(L_0^\alpha L_1^\beta)$ *with* $\alpha := (w - r)/q$ *and any* $\beta > 0$.

PROOF. Notice that $r_2(T) \le r$ implies $\nu_k \le \exp(K_1 s^{kr})$ eventually. Again for large $k$, as above we have

$$p_k(X) \le K \exp\{-K_2 2^{kr} k^{\beta q}\} =: p_k. \ □$$

Finally, let us apply our results to a centered, real valued Gaussian process $X = [X(t), t \in T\}$ with semimetric on $T$ given by

$$d(s, t) := \|X(s) - X(t)\|_2.$$

There are two techniques commonly used to show that $X$ has a version with continuous paths: one is to introduce an entropy condition, called (EN) below, and the other is to show the existence of a *mesure majorante*, the (MM) condition below.

Following Dudley (1973), given $\varepsilon > 0$ let $N(\varepsilon) = \exp H(\varepsilon)$ be the smallest $n$ such that there exist sets $A_1, \cdots, A_n \subset T$ with both $T \subset \cup_1^n A_k$ and diam $A_k \le 2\varepsilon$ for all $k$. Introduce the condition on $X$

$$\text{(EN)} \qquad f(t) := \int_0^t (H(s))^{1/2}\, ds < \infty, \qquad \text{some} \quad t > 0.$$

Following Fernique (1974), introduce the condition on $X$

(MM) there exist a probability measure $\lambda$ on $T$, $d$ with Borels such that

$$\lim_{\varepsilon\downarrow 0} \sup_{t\in T} \int_0^{\varepsilon} |\log \lambda(B(t, r))|^{1/2}\, dr = 0,$$

where $B(t, r) := \{s \in T, d(s, t) < r\}$.

If $X$ has either (EN) or (MM), it has a continuous version. If $X$ is stationary and $T = [0, 1]$, then (EN) and (MM) are equivalent with $\lambda :=$ Lebesgue, and necessary for $X$ to have a continuous version (Fernique, 1974). Also, Fernique (1974) shows that (EN) implies (MM), and Heinkel (1978) constructs an example of an $X$ on $T = [0, 1]$ with (MM) and not (EN). In addition, roughly stated, even for non-Gaussian processes, Fernique (1978, Theorems 1.3.3, and 2.1.1) shows that (MM) is necessary and sufficient for a continuous version.

Our results are patterned after those of Dudley, so it is not surprising that they will provide a continuous version of $X$ when (EN) holds. On the other hand, our methods do not seem to handle the case (MM). But in the next section we will compare the smoothness provided by our Theorem 4.1 with that given by the (MM) method as used by Nanopoulos and Nobelis (1978).

THEOREM 5.3. (Dudley, 1973, Theorem 2.1). *Let $X$ be a Gaussian process with property* (EN) *above. Then $X$ has a version in $g$-*Lip *for $g := \mathbf{G}(f)$.*

PROOF. Pick $\varepsilon_n$, $\delta_n$ as in Dudley (1973, Theorem 2.1) so that $(\varepsilon_n - \varepsilon_{n+1}) > 2\varepsilon_n/3$, $\varepsilon_{n+2} \le \delta_n/3$ and $\delta_n \le 2\varepsilon_n$, all $n$. Let $S_n$ be a minimal $3\delta_n/2$ net. Then $\nu_n \le \operatorname{card}(S_n \times S_{n+1}) \le \exp 2H(2\delta_{n+1}) \le \exp 4H(\varepsilon_{n+1})$, the last by definition of $\delta_{n+1} := 2 \inf\{\varepsilon; H(\varepsilon) \le 2H(\varepsilon_{n+1})\}$. Choosing $a_n = b[f(\varepsilon_{n+1}) - f(\varepsilon_{n+2})] \ge 2b\varepsilon_{n+1}H(\varepsilon_{n+1})^{1/2}/3$ and $b = 9^3/2$ we get, for large $n$,

$$q_n := \nu_n p_n(X) \le \exp -H(\varepsilon_{n+1}) =: r_n$$

and hence $q_{n+2} \le r_{n+2} \le r_n^2$. If $T$ is infinite then $H(\varepsilon) \to \infty$ as $\varepsilon \downarrow 0$, forcing $r_n < 1$ eventually and $\sum q_n < \infty$. If $T$ is finite there is no problem. □

## 6. Smoothness of processes on $T = [0, 1]^d$.

Here we consider Section 5 when $T = [0, 1]^d$ and $r_1(T) \le d$, and we construct examples to show that (5.1.a, b) cannot be improved much in these cases. These results are in answer to a conjecture by Garsia and Rodemich (1974) and follow techniques used by Hahn (1977) and Hahn and Klass (1977).

On $T = [0, 1]^d$ we use the metric induced by the supremum norm $\|t\| := \|(t_1, \cdots, t_d)\| = \max\{|t_1|, \cdots, |t_d|\}$ on $\mathbf{R}^d$.

THEOREM 6.1. *Let $X$ be a process on $T = [0, 1]^d$. Then $r_1(T) \le d$ and the results of Theorem* 5.1 *hold with $r = d$.* □

Let us compare these results with those of Nanopoulos and Nobelis (1978, page 641) who use the *mesure majorante* (MM) method mentioned at the end of Section 5 and Garsia and Rodemich (1974, page 105) whose method is somewhat similar. Assuming $T = [0, 1]$, so that $r_1(T) \le 1 = r$, they show that the inequality of (5.1.a) with $1 < p < q$, $v = 0$ implies $X$ has a version such that

$$|X(s) - X(t)| \le CYL_1^{\alpha}(|s - t|) \quad \text{all} \quad s, t \in [0, 1] \quad \text{a.s.},$$

where $C$ is a constant and $Y$ is r.v. with $E|Y|^p < \infty$ and $\alpha := 1 - (q/p) < 0$. Our (5.1.a) gives a slightly weaker result: $L_1^{\alpha}$ is replaced with $L_1^{\alpha}L_2^{\beta}$ with $\beta > 1$. On the other hand, in the case of (5.1.b) Nanopoulos and Nobelis only prove that $p > 1 = r = d$ implies $X$ has a version such that

$$|X(s) - X(t)| \le CY|s - t|^{q(p-1)/p^2} \quad \text{all} \quad s, t \in [0, 1] \quad \text{a.s.}$$

with $E|Y|^p < \infty$, while our result replaces $|s-t|^{q(p-1)/p^2}$ with the much smaller bound $|s-t|^{q/p}L_1(|s-t|)L_2^\beta(s-t)$, any $\beta > 1$.

We now expand the statement of (6.1) and specify in what sense (5.1.a, b) are nearly best possible.

THEOREM 6.2. *Let $X$ be a process on $T = [0, 1]$.*

(a) *Suppose inequality* (5.1.a) *holds with $r = d, p > 0$ and $\gamma := 1 \wedge (1/p)$.*

(i) *Suppose $p \le d$ so that $d\gamma \ge 1$. If $d\gamma > 1$, or $d\gamma = 1$ and $q\gamma > 0$, or $d\gamma = 1$ and $q\gamma = 0$ and $v\gamma > 0$, then $X$ has a constant version: $E|X(t) - X(0)|^p = 0$ for all $t$.*

(ii) *$p > d$ and $p < q$ or $p = q < v$ imply $X$ has a continuous and even Lipschitz version, as in* (5.1.a).

(iii) *$X$ need have no continuous version if $p \ge 2, p > d, p = q$ and $v = 0$.*

(b) *Suppose inequality* (5.1.b) *holds with $r = d, p, q > 0$.*

(i) *$p < d + q$ implies $X$ has a constant version.*

(ii) *$p \ge d + q$ implies $X$ has a version in $\mathrm{Lip}_\alpha$ for all $\alpha < q/p$.*

(iii) *If $p > d + q$ and $p \ge 2$, $X$ need have no version in $\mathrm{Lip}_\alpha$ for $\alpha > q/p$.*

PROOF. Part (ii) of both (a) and (b) is a restatement of (5.1.a, b). Part (i) of both is easy: for $1 \le p \le d$ the inequality of (5.1.a) implies

$$\|X(t) - X(0)\|_p \le \textstyle\sum_{k=1}^n \|X((k-1)t/n) - X(kt/n)\|_p$$
$$\le K_1 n/(n^d L_1^q(n) L_2^v(n))^{1/p} \to 0.$$

under the hypothesis of (a)(i). A similar argument works in case $p < 1$ and in case (b). Part (iii) will be proven by the examples constructed in the rest of this section.

To give the examples, which will complete the proofs of part (iii) of (6.2.a, b), we look at certain trigonometric series. In particular, for dimension $d = 1$ we consider

$$\textstyle\sum_{k\ge1} k^{-\beta} \cos kt, \quad \beta > 1 \quad \text{and} \quad \sum_{k\ge3} (k \log k)^{-1} \cos kt.$$

The first series clearly converges and will be shown to be in certain Lipschitz spaces, while the second will be shown to be integrable and continuous on $[-\pi, \pi]\backslash\{0\}$, yet large near 0. The ideas and inequalities we use have their source in Hahn and Klass (1977), although for $d \ge 2$ we are unable to attain their generality.

To set forth these examples let $\Omega := T := [-\pi, \pi]^d$ have probability measure $P :=$ $d$-dimensional Lebesgue$/(2\pi)^d$. For $s = (s_1, \cdots, s_d)$, $t \in \mathbf{R}^d$ and real $a < b$, let $s \cdot t := \sum_1^d s_i t_i$, $|s| := (|s_1|, \cdots, |s_d|)$, $\|s\| := \max\{|s_i|, 1 \le i \le d\}$, $s \le t$ means $s_i \le t_i$, $i = 1, \cdots, d$ and $a \le s$ $(\le b)$ means $a\mathbf{1} \le s$ $(\le b\mathbf{1})$ for $\mathbf{1} = (1, \cdots, 1)$ in $\mathbf{R}^d$. Finally

$$T_0 := \{t \in T, t_i = 0 \text{ some } i = 1, \cdots, d\}$$
$$I_j := \{k \in \mathbf{R}^d, 1 \le k_1 \le k_2 \le \cdots \le k_d < j + 1, k_i \text{ integral}\},$$
$$I := I_\infty, \quad I_j^c := I\backslash I_j, \quad j = 1, 2, \cdots$$

PROPOSITION 6.3. *Fix $p, q > 0, p > d + q, p \ge 2, r := p/(p-1), \beta := d + (q/p)$ and let*

$$f(t) := \textstyle\sum_I \|k\|^{-\beta} \cos k \cdot t, \quad t \in \mathbf{R}^d,$$
$$X(t, \omega) := f(t - \omega), \quad t \in T, \quad \omega \in \Omega.$$

(a) *$f$ is in $\mathrm{Lip}_{q/p}(T)$ but in no $\mathrm{Lip}_\alpha$ for $\alpha > q/p$.*

(b) *$E|X(s+t) - X(t)|^p \le K\|s\|^{d+q}$ for some $K < \infty$ and all $t, s + t \in T$, yet $X$ has no version in $\mathrm{Lip}_\alpha$ if $\alpha > q/p$.*

PROPOSITION 6.4. *Let*

$$v(k) := (k \vee 3)^{-d}\log^{-1}(k \vee 3),$$

*and*

$$g(t) := I(t \notin T_0) \textstyle\sum_{k\geq 1} v(\|k\|)\cos k\cdot t$$

*and*

$$X(t, \omega) = g(t - \omega).$$

(a) *$g$ is in $L_1(T, P)$ and is continuous on $T\backslash T_0$.*
*In addition there are constants $K_1$, $K_2$ and $K_p$ so that*
(b) *if $\pi/4d(n + 1) \leq |t|$ then*

$$|g(t)| \leq K_2 \log\log n,\ n \geq 3, \quad \textit{while}$$

(c) *if $\pi/4d(n + 1) \leq |t| \leq \pi/4dn$ then*

$$|g(t)| \geq K_1 \log\log n,\ n \geq 3, \quad \textit{and}$$

*(d) $X$ has no continuous version even though*

$$E|X(s + t) - X(t)|^p \leq K_p \|s\|^d/L_1^p(\|s\|)$$

$$\textit{for all} \quad t, s + t \in T, p > d, p \geq 2.$$

In many of the calculations of these proofs it is necessary to know the asymptotic ($\sim$) behavior of $f_{\alpha,\beta}(n) := \sum k^{-\alpha}\text{Log}^{-\beta}k$, where the summation is over 4 to $n$ if $\alpha < 1$ or $\alpha = 1 \geq \beta$ while over $n$ to $\infty$ if $\alpha > 1$ or $\alpha = 1 < \beta$. It is easy to show, directly for $\alpha \neq 1$ and using the theory of slowly varying functions as in Feller (1971, VIII-9) for $\alpha \neq 1$:

$$\begin{aligned} f_{\alpha,\beta}(n) &\sim n^{1-\alpha}(\log n)^{-\beta}/|1 - \alpha| && \text{if } \alpha \neq 1,\\ &\sim (\log n)^{1-\beta} && \text{if } \alpha = 1 \neq \beta\\ &\sim \log\log n && \text{if } \alpha = \beta = 1. \end{aligned}$$

PROOF OF PROPOSITION (6.3.). (a) Notice that

$$f(s + t) - f(t) = \textstyle\sum_{\pm I} a(k, s)e^{ik\cdot t}$$

where $a(k, s) := \|k\|^{-\beta}(e^{ik\cdot s} - 1)/2$, so that $|a(k, s)| \leq \|k\|^{-\beta}(d\|k\|\|s\| \wedge 2)/2$. Let $\sum_j(\sum_j^c)$ denote summation over $I_j(I_j^c)$. Observe that, for each $d$ and $j$, $\sum_j 1 = \binom{k+d-1}{d} \leq j^d$. Applying this for the first $d - 1$ variables in $I$, when $1/(j + 1) \leq \|s\| < 1/j$ we have

$$\begin{aligned} |f(s + t) - f(t)| &\leq dj^{-1}\textstyle\sum_j \|k\|^{1-\beta} + 2\sum_j^c \|k\|^{-\beta}\\ &= 0(j^{-1}\textstyle\sum_1^j k^{d-\beta} + \sum_{j+1}^{\infty} k^{d-1-\beta})\\ &= 0(\|s\|^{q/p}). \end{aligned}$$

For the second statement use the Dini-Lipschitz theorem (Igari (1968, page 12)): If $f$ is in $\text{Lip}_\alpha$, then

$$0(j^{-q/p}) = (\textstyle\sum_j^c \|k\|^{-\beta}) \leq \sup_T |f(t) - \sum_j \|k\|^{-\beta}\cos k\cdot t| = 0(j^{-\alpha}\log^d j),$$

and $\alpha > q/p$ is impossible.

(b) For the first part use the Hausdorff-Young inequality (extend Katznelson (1968, page 99)): for $1/(j + 1) \leq \|s\| < 1/j$ and $p \geq 2$,

$$\begin{aligned} \|X(t + s) - X(t)\|_p = \|f(\cdot + s) - f(\cdot)\|_p &\leq (\textstyle\sum_{\pm I} |a(k, s)|^r)^{1/r}\\ &= 0(j^{-r}\textstyle\sum_j \|k\|^{r(1-\beta)} + \sum_j^c \|k\|^{-r\beta})^{1/r}\\ &= 0(j^{d-\beta r})^{1/r} = 0(\|s\|^{d+q})^{1/p}, \end{aligned}$$

since $-r < -(d+q)r/p = d - \beta r < 0$. The fact that $X$ has no version in $\mathrm{Lip}_\alpha$ for $\alpha > q/p$ is clear since $f \notin \mathrm{Lip}_\alpha$. □

Before we can prove (6.4) we need two well known facts.

LEMMA 6.5 (SUMMATION BY PARTS): *Given $\{u(k)\}$ and $\{v(\|k\|)\}$ for $k \in \mathbf{Z}^d$, define*

$$s(n) := \textstyle\sum_{1\le k\le n} u(k)v(\|k\|).$$

*For fixed $n \ge 1$ with $m := \|n\|$ and $j = 1, 2, \cdots$, let*

$$U_n(j) := \textstyle\sum_{1\le k\le j1} [u(k)I(k \le n)],$$

$$V(j) := v(j) - v(j+1).$$

*Then, by an appropriate induction,*

$$s(n) = \textstyle\sum_{j=1}^{m-1} U_n(j)V(j) + U_n(m)v(m).$$

*Hence, if $|\sum_{1\le k\le n} u(k)| \le B < \infty$ for all $n$ and $v(j) \downarrow 0$ as $j_0 \le j \to \infty$, then* $\min(n_1, \cdots, n_d)$ *$\to \infty$ implies*

$$s(n) \to s := \textstyle\sum_1^\infty U(j)V(j), \qquad |s| \le Bv(1)$$

*where $U(j) := U_{j1}(j)$.*

To apply this with $u(k) := \cos k\cdot t$, fixed $t$, use induction and

$$\sin(s/2) \textstyle\sum_{j=1}^{m} \cos(\alpha + js) = \sin(ms/2)\cos\{(m+1)s/2 + \alpha\}$$

*To prove*

LEMMA 6.6. Define $C(n, t) := \sum_{1\le k\le n} \cos k\cdot t$. Then

(a) $C(n, t) = \cos(n+1)\cdot t/2 \,\Pi_1^d \sin(nt_i/2)/\sin(t_i/2)$,
(b) $|C(n, t)| \le (\Pi_1^d n_i) \wedge (\Pi_1^d \pi/|t_i|)$ *and*
(c) $|C(n, t)| \ge (\Pi_1^d n_i)/2$ *if* $|k\cdot t| \le \pi/3$ *for* $1 \le k \le n$.

PROOF OF PROPOSITION 6.4. Define $u(k) := \cos k\cdot t$, $t \notin T_0$, $v(\|k\|)$ as in (6.4) and check that $v(j) \downarrow 0$ for large $j$ and

$$V(j) := v(j) - v(j+1) \sim d/j^{d+1}\log j$$

as $j \to \infty$ ($\sim$ := "asymptotic to"). Then (6.4.a, b, c) follow from (6.5) and (6.6). Because of (6.4.c), $X$ has no continuous version. To prove the moment inequality of (6.4.d), use the Hausdorff-Young inequality as before. □

**7. Brownian sheet and summation processes.** We now turn to the processes which motivated this study. Recall that we defined the Brownian sheet in (1.2) and guessed it would be a limit of a smooth or jump summation process in (1.3). To see that this is so we begin by proving two general inequalities.

PROPOSITION 7.1. *Given $\zeta_1, \zeta_2, \cdots$* iid *$\zeta$ with $E\zeta = 0$, $E\zeta^2 = 1$. For $2 \le p < \infty$ there is a constant $K_p$, depending only on $p$, such that for all real $a$;*

$$\tfrac{1}{2}(\textstyle\sum a_j^2)^{p/2} \le E|\sum a_j\zeta_j|^p \le K_p E|\zeta|^p(\sum a_j^2)^{p/2}.$$

PROOF. Apply Theorem 3 of H. P. Rosenthal (1970) and notice that $\sum |a_j|^p = \sum(a_j^2)^{p/2} \le (\sum a_j^2)^{p/2}$. □

Better results are possible when $\zeta$ is *generalized Gaussian*, i.e. when there is a constant $K = K(\zeta)$ such that for all $t > 0$

$$E\exp(t\zeta) \le \exp Kt^2/2.$$

Such $\zeta$ are also called sub Gaussian if $K > 0$. Note that bounded or Gaussian variables have the above property.

PROPOSITION 7.2. *If $\zeta_1, \zeta_2, \cdots$ iid $\zeta$ and $\zeta$ is generalized Gaussian, then for all real $a_j$ and $t > 0$*

$$P(|\sum a_j\zeta_j| > t) \le 2\exp\{-t^2/2K(\sum a_j^2)\}.$$

PROOF. Since $E \exp t(\sum a_j\zeta_j) \le \exp K \sum a_j^2t^2/2$, the result follows from Chow (1966, equation 4).

Recall the notation given in (1.1). We also need the set $A^{-\varepsilon} := R^d \backslash (R^d \backslash A)^{\varepsilon}$ and the fact, proven in Bhattacharya and Rao (1976, page 14),

$$|A^{\varepsilon}\backslash A^{-\varepsilon}| = |(\text{bdry } A)^{\varepsilon}|, \qquad A \in T_{\mathscr{B}}.$$

For our next result write

$$\beta_n := \beta_n(A) = |(\text{bdry } A)^{\varepsilon_n}| = |A^{\varepsilon_n}\backslash A^{-\varepsilon_n}|,$$

for $A \in T$ and $\varepsilon_n := \| 1/n \|$.

THEOREM 7.3. *Let $X_n$ and $Z_n$ denote the smooth and jump summation processes as defined in* (1.3). *Let $A \in T_{\mathscr{B}}$.*

(a) *Given $p \ge 2$ there is a constant $K_p$ such that*

$$E|X_n(A) - Z_n(A)|^p \le K_pE|\xi|^p[\beta_n(A)]^{p/2}.$$

*Thus $A \in T_0$ implies*

$$E|X_n(A) - Z_n(A)|^2 \le K_2\beta_n(A) \to 0 \quad \text{as} \quad \varepsilon_n \to 0.$$

(b) *If $\xi$ is generalized Gaussian then*

$$P(|Z_n(A) - X_n(A)| > t) \le \exp\{-t^2/2K\beta_n(A)\}.$$

PROOF. $X_n(A) - Z_n(A) = \sum a_j\xi_j$ where

$$a_j := |R_{n,j}|^{-1/2}\{|R_{nj} \cap A| - |R_{n,j}| I(j/n \in A)\}.$$

But $a_j^2 = 0$ if either $R_{n,j} \subset A$ or $R_{n,j} \subset \mathbf{R}^d\backslash A$, while on the other hand $R_{n,j} \subset A^{\varepsilon_n}\backslash A^{-\varepsilon_n}$ and in general $a_j^2 \le |R_{n,j}|$. This implies that $\sum a_j^2 \le |A^{\varepsilon_n}\backslash A^{-\varepsilon_n}| = |(\text{bdry}A)^{\varepsilon_n}| = \beta_n(A)$. □

We will use (7.3) in Section 8 to show that $Z_n \mid T$ converges in distribution to $W \mid T$ for the uniform topology for certain $T \subset T_0$. For now let us prove finite dimensional convergence. Recall that

$$T_0 := \{A \subset [0, 1]^d, A \text{ Borel}, |\text{bdry } A| = 0\}.$$

THEOREM 7.4. *Both $X_n \mid T_0$ and $Z_n \mid T_0$ converge in finite dimensional distribution to $W \mid T_0$ as $\| 1/n \| \to 0$.*

PROOF. Treat $Z_n$ first. Fix $A_1, \cdots, A_r \in T_0$, $v_1, \cdots, v_r$ real and set $v := \sum |v_p|$,

$$V_n := \textstyle\sum_{p=1}^r v_pZ_n(A_p) = \sum_{j\le n} \alpha_{nj}\xi_j,$$

where $\alpha_{nj} := \sum_p v_p|R_{n,j}|^{1/2}I(j/n \in A_p)$, $|\alpha_{nj}| \le v\| 1/n \|^{d/2}$. If $U_n$ denotes the random variable uniformly distributed on the points $j/n$, $1 \le j \le n$, then

$$\begin{aligned} b_n := EV_n^2 &= \textstyle\sum_j \alpha_{nj}^2 = \sum_{j\le n} |R_{n,j}| \sum_{p,q} v_pv_qI(j/n \in A_p \cap A_q) \\ &= \textstyle\sum_{p,q} v_pv_qP(U_n \in A_p \cap A_q) \\ &\to b := \textstyle\sum_{p,q} v_pv_q \mid A_p \cap A_q| = E(\sum v_pW(A_p))^2. \end{aligned}$$

Now apply the Lindeberg theorem: for $b > 0$, $\| 1/n \|$ so small that $b_n > b/2$, and any $t > 0$,

$$\textstyle\sum_{j \le n} E(\alpha_{nj}^2 \xi_j^2 / b_n;\ \alpha_{nj}^2 \xi_j^2 > b_n t) \le E(\xi^2, \xi^2 > bt/2v^2 \| 1/n \|^d) \to 0$$

as $\| 1/n \| \to 0$. The result for $X_n \mid T_0$ follows from (7.3.a). □

The next inequality allows us to apply the basic smoothness and tightness theorems and their applications, Sections 4 and 5, to obtain that $X_n$ converges to $W$ in distribution for certain domains $T \subset T_0$.

THEOREM 7.5. (a) *Given $p \le 2$, there is a constant $K_p$ such that for all $A, B \in T_{\mathscr{B}}$*

$$E \mid X_n(A) - X_n(B) \mid^p \le K_p E \mid \xi \mid^p [d_L(A, B)]^{p/2}.$$

(b) *If $\xi$ is generalized Gaussian, then for all $t > 0$, all $A, B \in T_{\mathscr{B}}$ and $K = K(\xi)$ we have*

$$P(| X_n(A) - X_n(B) | > t) \le 2 \exp\{-t^2/2K\, d_L(A, B)\}.$$

PROOF. $X_n(A) - X_n(B) = \sum a_j \xi_j$ with

$$a_j := | R_{nj} |^{-1/2} (| R_{nj} \cap A \backslash B | - | R_{nj} \cap B \backslash A |),$$

so that $a_j^2 \le | R_{nj} \cap (A \Delta B) |$. Now apply (7.1). □

We are now ready to prove Lipschitz convergence of $X_n \mid T$ to $W \mid T$. Before doing this we should check that the Brownian sheet $W$ is smooth. Because $W(A) - W(B) \sim \mathscr{N}(0, d_L(A, B))$, for all $a > 0$

$$P(| W(A) - W(B) | > a) \le 2 \exp\{-a^2/2\, d_L(A, B)\}.$$

Thus, if $T \subset T_{\mathscr{B}}$ has $r_2(T) \le r < 1$ then $W \mid T := \{W(A), A \in T\}$ has a version in $g$-Lip for $g := \mathbf{G}(L_0^\alpha L_1^\beta)$, $\alpha := (1 - r)/2$ and any $\beta > 0$, by Theorem 5.2. (Remember that $L_0 :=$ identity and $L_{k+1} := | \log \mid L_k | |$.)

The next theorem is due to Lamperti (1962) in case $d = 1$ and $T = \{[0, t], 0 \le t \le 1\}$, so that $r_1(T) = 1$. This is our main theorem for convergence of summation processes.

THEOREM 7.6. *Fix totally bounded $T \subset T_0$, $d_L$.*

(a) *Suppose $E \mid \xi \mid^p < \infty$ for some $p > 2$. If $r_1(T) \le r$ and $q := (p/2) - r > 0$, then $W \mid T$ and each $X_n \mid T$ has a version in $g$-Lip $\subset$ Lip$_\alpha$ for $g^p := \mathbf{G}(L_0^q L_1 L_2^\beta)$, for each $\beta > 1$ and $\alpha < q/p$. Also, $X_n \mid T$ converges in distribution to $W \mid T$ as processes on $g$-Lip, ${}_f\| \ \|$ for each $f$ such that $g = o(f)$ at 0.*

(b) *Suppose $\xi$ is generalized Gaussian. If $r_1(T) \le r$ ($r(T) < r < 1$), then $W \mid T$ and each $X_n \mid T$ has a version in $g$-Lip $\subset$ Lip$_\gamma$ for $g := \mathbf{G}(bL_0^{1/2} L_1^{1/2})$ and $b$ large ($g := \mathbf{G}(L_0^\lambda L_1^\beta)$ with $\lambda := (1 - r)/2$ and any $\beta > 0$) and $0 < \gamma < ½$ ($0 < \gamma < \lambda$). Also, $X_n \mid T$ converges in distribution to $W \mid T$ as processes on $g$-Lip, ${}_f\| \ \|$ for each $f$ such that $g = o(f)$ at 0.*

PROOF. There is really nothing left to prove, but let us point to the relevant theorems as they are used. For the first statements in (a) and (b) use (7.5.a), (5.1.b) and (7.5.b), (5.1.c), (5.2) respectively. The second statements follow from (7.4), (4.2) and the remark preceding it, (7.5) and the calculations of (5.1) and (5.2). □

Let us now turn to the Hausdorff metric $d_H$. For this we introduce, for each $K > 0$,

$$T_{\partial, K} := \{A \in T_0;\ | A^\varepsilon \backslash A | \le K\varepsilon/2, \quad \text{all} \quad 0 < \varepsilon \le 1\}.$$

Notice that $A \subset B^\varepsilon$ implies $A \backslash B \subset B^\varepsilon \backslash B$. Hence $\varepsilon > d_H(A, B)$ implies $A \Delta B \subset A^\varepsilon \backslash A \cup B^\varepsilon \backslash B$, and we have

$$(7.7) \qquad A, B \in T_{\partial, K} \quad \text{implies} \quad d_L(A, B) \le K\, d_H(A, B).$$

It is well known that there is a $K$ such that $T_{cvx} := \{A \subset [0, 1]^d, A \textit{ convex}\} \subset T_{\partial, K}$. The example of two distinct points shows that the converse of (7.7) fails.

THEOREM 7.8. *For $T \subset T_{\partial,K}$, (7.6) remains valid with $r_1$ and $r_2$ calculated using $d_H$.* □

To apply (7.7) and (7.8) we exhibit some appropriate subsets $T \subset T_{\partial,K}$:

$$T_d := \{[0, t] = [0, t_i] \times \cdots \times [0, t_d]; \mathbf{0} \le t = (t_1, \cdots, t_d) \le \mathbf{1}\},$$

$$T_{d,c} := \{\text{convex hull } (v_1, \cdots, v_c); v_i \in [0, 1]^d, i = 1, \cdots, c\}$$

$$T_{d,\alpha,M} := \text{Dudley's (1978, page 917) class of sets } J(d, \alpha, M).$$

It is easy to show that both $T_d$, $T_{d,c} \subset T_{\partial,K}$, some $K = K(d, c)$, and

$$r_1(T_d) \le d, r_1(T_{d,c}) \le dc, \qquad \text{for both} \quad d_L \quad \text{and} \quad d_H.$$

Let us describe the sets in $T_{d,\alpha,M}$ and some of Dudley's (1974) results about them. Let $S^{d-1} := \{x \in \mathbf{R}^d; \sum x_i^2 = 1\}$ and let $G(d, \alpha, M)$ be the set of all $f: S^{d-1} \to [0, 1]^d$ which, with all their partial derivatives of orders $\le \alpha$, are bounded in norm by $M$. (For our convenience we have chosen range $(f)$ in $[0, 1]^d$ rather than in $\mathbf{R}^d$ as in Dudley.) Now let $I(f)$ be the open set of all $y \in \mathbf{R}^d \backslash$range $(f)$ such that $f$ is not homotopic to a constant map in $\mathbf{R}^d\{y\}$. Finally, set $J(f) := I(f) \cup$ range $(f)$ and $T_{d,\alpha,M} := \{J(f); f \in G(d, \alpha, M)\}$. Then Dudley (1974) proves that

$$T_{d,\alpha,M} \subset T_{\partial,K} \qquad \text{some} \quad K = K(d, \alpha, M)$$

and for $\alpha \ge 1$

$$r(T_{d,\alpha,M}) = (d-1)/\alpha \qquad \text{for both} \quad d_L \quad \text{and} \quad d_H.$$

Recall (Section 5) that $r_2(T) \le r(T) + \delta$ for all $\delta > 0$.

THEOREM 7.9. (a) (7.6.a) *holds for $T = T_d(T = T_{d,c})$ and either metric $d_L$ or $d_H$ if $E \mid \xi \mid^p < \infty$ for some $p > 2d$ ($p > 2dc$).*

(b) (7.6.b) *holds for $T = T_{d,\alpha,M}$ and either metric $d_L$ or $d_H$ if $\alpha > d - 1$* (iff $r(T) < 1$) *and $d \ge 2$.*

**8. Summation processes; the jump version.** Let $Z_n \mid T$ again denote the jump summation process. Under suitable conditions on the moments of $\xi$ and the size of $T$, we prove that

$$Z_n \mid T \to W \mid T \qquad (\mathscr{D}, \mathscr{U}),$$

which is to say that there is convergence in distribution for the uniform topology.

Since the $Z_n$ processes live on a space like $D[0, 1]$, there is no well established definition for $(\mathscr{D}, \mathscr{U})$ convergence. A recent paper of Dudley (1978) presents one possible definition (page 900) and mentions two other proposals (page 902). Our meaning will be a little stronger than Dudley's. Let us sketch the distinction briefly.

Begin with metric space $(S, \rho)$ having Borels $\mathscr{B}$ and ball $\sigma$-algebra $\mathscr{B}_b \subset \mathscr{B}$. Given probability measures $\{P_n\}$, $P$ on $S$, $\mathscr{B}_b$, Dudley (1978) defines $P_n \to P$ to mean $\int f\, dP_n \to \int f P$ for every bounded, continuous $f: S \to \mathbf{R}$ which is $\mathscr{B}_b$ measurable. Our usage removes the condition that $f$ be $\mathscr{B}_b$ measurable and then replaces the integrals by upper and lower integrals, $\int^*$ and $\int_*$. Let us return to the $Z_n \mid T$ processes to see how this change comes about.

We already have that $Z_n \mid T \to W \mid T$ in finite dimensional distribution (Theorem 7.4), and by analogy with Dudley (1978, Theorem 1.2) we should have $Z_n \mid T \to W \mid T$ $(\mathscr{D}, \mathscr{U})$ if $\{Z_n \mid T\}$ has *small oscillation*, i.e. if for all $\varepsilon > 0$ there is a $\delta > 0$ such that

$$P(1 - \operatorname{osc}(Z_n \mid T, \delta) < \varepsilon) > 1 - \varepsilon \qquad \text{eventually}.$$

As is seen in Theorem 8.6 below, small oscillations for $\{Z_n \mid T\}$ implies not only Dudley's type of convergence (8.6.c) but ours as well (8.6.b).

Let us summarize the contents of this section. First, we give an example to show that when $T$ is too large it is impossible for $Z_n \mid T$ to converge to $W \mid T$ for the uniform topology, even when the random summands have only two values and the sets in $T$ have very smooth boundaries. After this discouraging example we show, on the other hand, that there is such convergence in case $T = T_{d,c}$ consists of polytopes with a bounded number, $c$, of vertices in $[0, 1]^d$ and if the summands have enough moments (Theorem 8.1). Finally, we look at the implications of finite dimensional convergence combined with small oscillations of $Z_n \mid T$, and we characterize small oscillations in terms of a certain type of convergence (Theorem 8.6).

Let us give an *example*, communicated to us orally by R. M. Dudley, to show that $Z_n \mid T$ cannot converge in distribution to $W \mid T$ for the uniform topology if $T$ contains too many sets, even if they have smooth boundaries. To do this let $d = 2$ and assume $\xi = \pm 1$ with probability ½, ½. For any positive integer $m$ let $C_\infty := \{f\colon [0, 1] \to [0, 1]$, $k$th derivative $D^k f$ exists for all $k\}$, $C_{\infty,m} := \{f \in C_\infty;\ \|D^k f\| \le 1,\ k = 0, \cdots, m\}$, and for $f \in C_{\infty,m}$ define $A(f) := \{(x, y);\ 0 \le x \le 1,\ 0 \le y \le f(x)\}$.

For even $n = 2k$, when $f \in C_{\infty,m}$ is such that $\|f - ½\| \le ½n$, we have

$$Z_n(A(f)) - Z_n(A(1 - f)) = \frac{1}{n} \textstyle\sum_{i=1}^n \operatorname{sgn}\left[ f(i/n) - \frac{1}{2} \right] \xi_{i,k},$$

and

$$d(A(f), A(1 - f)) \le 1/n \text{ for } d := d_L \text{ and } d := d_H.$$

If we take $T := \{A = A(f);\ f \in C_{\infty,m},\ \|f - ½\| \le ½\}$ it is clear that $1 - \operatorname{osc}\{Z_n \mid T, 1/n\} \ge 1$: for each $\omega$ choose $f$ so that $\operatorname{sgn}[f(i/n) - ½]\xi_{i,k}(\omega) = 1$, $i = 1, \cdots, n$.

Let us now turn to positive results. We will find conditions on $T$ and the summands so that $Z_n \mid T$ has small oscillations. Note that

$$1\text{-osc}(Z_n \mid T, \delta) \le 2\, \| Z_n - X_n \|_T + 1 - \operatorname{osc}(X_n \mid T, \delta)$$

where $\| x \|_T := \sup\{ | x(t) |,\ t \in T\}$. Hence, $Z_n \mid T$ has small oscillations if

$$\| Z_n - X_n \|_T \to 0 \text{ probability}$$

and if $X_n \mid T$ has small oscillations. The latter is easy to deduce for suitable $\xi$ and $T$: for nice $\xi$ and $T$, (a) if $g \in \mathrm{G}$ and $g(r) \downarrow 0$ as $r \downarrow 0$, then

$$1 - \operatorname{osc}(X_n \mid T, \delta) \le {}_g\| X_n \mid T \|\, g(\delta) \text{ and}$$

(b) for all $\varepsilon > 0$ there is a $b > 0$ such that

$$P\{{}_g\| X_n \mid T \| \le b\} > 1 - \varepsilon \text{ eventually}$$

(see Theorems 4.1 and 7.6).

For totally bounded $T \subset T_0$ we let $\varepsilon_n$ − net $S_n$ and map $\sigma_n$ be as in (3.1). To show that $\| Z_n - \mathrm{X}_n \|_T \to 0$ in probability, notice that

$$\| Z_n - X_n \|_T \le \| Z_n - Z_n \circ \sigma_n \|_T + \| Z_n \circ \sigma_n - X_n \circ \sigma_n \|_T + 1 - \operatorname{osc}(X_n \| T, \varepsilon_n).$$

The last term is easily handled, as above. For the second, we have

$$\begin{aligned} P(\| Z_n \circ \sigma_n - X_n \circ \sigma_n \|_T > \varepsilon) &\le [\operatorname{card} \sigma_n(T)] \max\{P(| Z_n(A) - X_n(A) | > \varepsilon),\ A \in \sigma_n(T)\} \\ &\le [\operatorname{card} \sigma_n(T)] K_p E | \xi |^p \varepsilon^{-p} \max\{\beta_n(A);\ A \in \sigma_n(T)\}^{p/2} \end{aligned}$$

by (7.3.a). Finally

$$\begin{aligned} P(\| Z_n - Z_n \circ \sigma_n \|_T > \varepsilon) \le{}& \operatorname{card} \sigma_n(T) \\ &\times \max\{P(\sup_{B \in \sigma_n^{-1}(A)} | Z_n(B) - Z_n(A) | > \varepsilon);\ A \in \sigma_n(T)\}. \end{aligned}$$

The key in handling this last term is to notice that the "sup" is really a "max", for $Z_n(B) - Z_n(A) = Z_n(B\backslash A) - Z_n(A\backslash B)$ depends on $A$ and $B$ only through the subsets of $J_n := \{k/n, \mathbf{1} \le k \le n\}$ that can be formed by $B\backslash A$ and $A\backslash B$.

Let us now apply the above technique to the case $T = T_{d,c} := \{A \subset [0, 1]^d; A$ is a convex polytope with at most $c$ vertices$\}$.

To keep matters simple from now on we assume vector $n = (n_1, \cdots, n_d)$ has $n_1 = n_2 = \cdots = n_d = \|n\|$ and also take $\varepsilon_n := \|\mathbf{1}/n\| = 1/\|n\|$.

THEOREM 8.1. *Let $T = T_{d,c}$ with Hausdorff metric $d_H$ and form the summation process $Z_n$ for summands* iid. *$\xi$, where $E|\xi|2^p < \infty$ for some $p > 2\left[dc + (d-1)\left\{d\binom{c}{d} + 1\right\}\right]$. Then $Z_n | T_{d,c}$ converges in distribution to $W | T_{d,c}$ for the uniform topology.*

PROOF. Theorem (7.9.a) and the arguments above show that $X_n | T_{d,c}$ has small oscillations if $E|\xi|^p < \infty$ for some $p > 2dc$.

To treat the term $\|Z_n \circ \sigma_n - X_n \circ \sigma_n\|_T$ introduce some notation: For $0 \le x \le 1$ let $[x]_n = k/\|n\|$ if $k - 1 < \|n\| x \le k$, $= 1/\|n\|$ if $x = 0$. For $w = (w_1, \cdots, w_d) \in [0, 1]^d$ let $[w]_n := ([w_1]_n, \cdots, [w_d]_n)$. Finally, if $A \in T_{d,c}$ has vertices $v_1, \cdots, v_b$, $b \le c$, let

$$\sigma_n(A) := \text{convex hull } ([v_1]_n \cdots, [v_b]_n).$$

Clearly, $S_n := \sigma_n(T_{d,c})$ is an $\varepsilon_n$-net for $T_{d,c}$ with metric $d_H$ and

$$c_n := \text{card } S_n \le n^{dc}.$$

In addition, from Bhattacharya and Rao (1976, Theorem 3.1)

$$\beta_n(A) := |A^{\varepsilon_n}\backslash A^{-\varepsilon_n}| \le K_d \varepsilon_n$$

for all convex $A \in [0, 1]^d$. This gives

$$P(\|Z_n \circ \sigma_n - X_n \circ \sigma_n\|_T > \varepsilon) \le K_{d,p} \varepsilon^{-p} E|\xi|^p n^{dc-(p/2)} \to 0$$

if $E|\xi|^p < \infty$ for some $p > 2dc$.

The bounding of the $\|Z_n - Z_n \circ \sigma_n\|_T$ term is much more complicated. We need a little more notation and two lemmas: add a parenthetical (2) to denote objects defined using the $l_2$ norm of $v = (v_1, \cdots, v_d) \in \mathbf{R}^d$:

$$\|v\| \le \|v\|_{(2)} := (\textstyle\sum v_j^2)^{1/2} \le d^{1/2}\|v\|.$$

LEMMA 8.2. *$A \in \sigma_n(T_{d,c})$ and $B \in \sigma_n^{-1}(A)$ implies*

$$A^{-\varepsilon d^{1/2}} \subset B \subset A^{\varepsilon}, \ \varepsilon > 1/\|n\|.$$

PROOF. The second inclusion is clear and the first can be derived as follows: since $A^{-\eta} \subset A^{-\eta,(2)}$ and $A \subset B^{\varepsilon} \subset B^{\eta,(2)}$ where $\eta := d^{1/2}\varepsilon$, it suffices to show that

$$B^{(2)}(x, \eta) := \{y; \|x - y\|_{(2)} < \eta\} \subset B^{\eta,(2)} := \cup\{B^{(2)}(b, \eta); b \in B\}$$

implies $x \in B$ if $B$ closed and convex.

If $x \notin B$, there exists a unique $b \in B$ such that $\|x - b\|_2 = d_{(2)}(x, B) =: \delta > 0$; then $y := x + \rho(x - b)/\delta \in B^{(2)}(x, \eta) \subset B^{\eta,(2)}$ while $d_{(2)}(y, B) = \delta + \rho$ for all $\rho < \eta$, a contradiction. □

Define $\eta_n := 2d^{1/2}/\|n\|$. Then for $A \in \sigma_n(T_{d,c}) =: S_n$ and $B \in \sigma_n^{-1}(A)$, we have $A^{-\eta_n} \subset B \subset A^{\eta_n}$. Further, let $\mathscr{F}(A) := \{J_n \cap B\backslash A^{-\eta_n}, B \in \sigma_n^{-1}(A)\}$ for $A \in S_n$, where $J_n := \{k/n, \mathbf{1} \le k \le n\} \subset [0, 1]^d$. Because

$$Z_n^*(A) := \sup\{|Z_n(B) - Z_n(A)|; B \in \sigma_n^{-1}(A)\}$$
$$\le 2 \max\{|Z_n(C)|; C \in \mathscr{F}(A)\},$$

we see that

$$P(\|Z_n - Z_n \circ \sigma_n\|_T > \varepsilon)$$
$$\le \|n\|^{dc} \max_{A \in \sigma_n(T)} \{\text{card } \mathscr{F}(A) \times 2^p/\varepsilon^p \times \max_{C \in \mathscr{F}(A)} E|Z_n(C)|^p\}.$$

LEMMA 8.3. *There is a constant $K = K_d$ such that for each $A \in \sigma_n(T)$ and each $n$*

$$\text{card } \mathscr{F}(A) \le K_d \|n\|^\rho$$

*where* $\rho = \rho_{d,c} := (d-1)\left[d\binom{c}{d} + 1\right]$.

PROOF. This is based on the following result of Vapnik and Chervonenkis (1971) which we state only for our sets: if $X := J_n \backslash A^{-\eta_n}$ where $\eta_n := 2d^{1/2}/\|n\|$, $\Delta(F) := \text{card}\{F \cap B, B \in \sigma_n^{-1}(A)\}$, $m(k) := \max\{\Delta(F), \text{card } F = k, F \subset X\}$ and $r := \min\{k; m(k) < 2^k\}$ ($\min \phi := \infty$), then $m(k) \le k^r + 1$ for $k = 0, 1, \cdots$ In our case we claim that $r \le d\binom{c}{d} + 1$: each $B \in \sigma_n^{-1}(A)$ has at most $\binom{c}{d}$ faces (of dimension $d-1$); and each face and its half space containing $B$ can produce, as $B$ varies over $\sigma_n^{-1}(A)$, all subsets of at most $d$ points. Hence, $m(j) < 2^j$ for $j = d\binom{c}{d} + 1$. Since

$$A^{-\eta_n} \subset B \subset A^{\eta_n} \text{ for } B \in \sigma_n^{-1}(A),$$

the result follows if we take $F = J_n \cap A^{\eta_n} \backslash A^{-\eta_n}$ and $k = \text{card } F$. But $j/n \in F$ implies $R_{n,j} := [j - 1/n, j/n] \subset (A^{\eta_n} \backslash A^{-\eta_n})^{\eta_n} = (\partial A^{\eta_n})^{\eta_n} \subset A^{2\eta_n} \backslash A^{-2\eta_n} =: A_n$. Thus

$$\text{card } F = \|n\|^d \textstyle\sum_{j/n \in F} |R_{n,j}|$$
$$= \|n\|^d \,|\textstyle\bigcup_{j/n \in F} R_{n,j}|$$
$$\le \|n\|^d |A_n| \le K_d' \|n\|^{d-1}. \quad \square$$

PROOF OF THEOREM 8.1 COMPLETED: It remains to bound $E|Z_n(C)|^p$ for $C \in \mathscr{F}(A)$, $A \in \sigma_n(T)$. By (7.1) we have

$$E|Z_n(C)|^p \le K_p E|\xi|^p (\text{card } C)^{p/2} \|n\|^{-pd/2}.$$

Now $C \subset J_n \cap A^{\eta_n} \backslash A^{-\eta_n} =: F$, and in the previous lemma we saw that card $F \le K_d' n^{d-1}$.
All together this gives

$$P(\|Z_n - Z_n \circ \sigma_n\|_T > \varepsilon) \le K_{d,p} \varepsilon^{-p} E|\xi|^p \|n\|^{\alpha(p)},$$

where $\alpha(p) = dc + \rho + (d-1)p/2 - dp/2 = dc + \rho - p/2$. $\square$

We should compare Theorem 8.1 with the results of Wichura (1969): he assumes only $E|\xi|^2 < \infty$, but his result holds only for $T = \{[0, t] = [0, t_1] \times \cdots \times [0, t_d], t = (t_1, \ldots, t_d) \in [0, 1]^d\}$. Our methods do not give his result because we do not have a maximal inequality in case $T = T_{d,c}$.

Finally, let us describe in more detail the implications of

(a) $Z_n | T \to W | T$ in finite dimensional distribution

and

(b) $\{Z_n | T\}$ has small oscillations.

To do this we think of $Z_n | T$ and $W | T$ as all realized on a path space like $D[0, 1]$: assume that $r_2(T) \le r < 1$ so that $W | T$ has continuous paths; notice that each $Z_n | T(\omega)$ has a *simple path* $s \in S(T, \mathbf{R})$ i.e. a path of the form

$$s(A) := \textstyle\sum_1^r a_k I(b_k \in A),\ a \in T$$

for some $r = 1, 2, \ldots$, and $a_k \in \mathbf{R}$, $b_k \in [0, 1]^d$. Now define

$$D_0 := D_0(T) := \{x + s;\ x \in C(T, \mathbf{R}),\ s \in S(T, \mathbf{R})\}$$

and give $D_0$ the uniform topology. For $\sigma$-algebras introduce the Borels $\mathscr{B}$, the ball-Borels

$$\mathscr{B}_b := \sigma\{B(y, \varepsilon);\ y \in D_0;\ \varepsilon > 0\},\ B(y, \varepsilon) := \{z \in D_0,\ \|z - y\| < \varepsilon\}$$

and the cylinder $\sigma$-algebra $\mathscr{C}$. Because $C(T, \mathbf{R})$ is separable iff $T$ is compact (metric) (Dunford and Schwartz (1958), page 437, problem 17), $D_0$ will usually be nonseparable and $\mathscr{B}_b$ will likely be strictly contained in $\mathscr{B}$.

Now let $Z_n$ be realized on $D_0$, $\mathscr{C}$ under measure $P_n$, say. We claim

**PROPOSITION 8.4.** *Sets in $\mathscr{B}_b$ are $P_n$ measurable if $T$ is totally bounded.*

**PROOF.** Given $\varepsilon > 0$ and $y = x + s \in D$ with $x \in C(T, \mathbf{R})$ and $s = \sum_1^r a_k I(b_k \in \cdot) \in S(T, \mathbf{R})$. Let $J_n(s) := \{b_k, j/n;\ k = 1, \ldots, r,\ 1 \le j \le n \in [0, 1]^d\}$. For each $K \subset J_n(s)$ set $T(K) := \{A \in T;\ A \cap J_n(s) = K\}$, and note that $T(K)$ is totally bounded since $T$ is. Finally

$$\{\omega;\ \|Z_n(\omega, \cdot) - y\|_T \le \varepsilon\} = \cap_{K \subset J_n(s)} \cap_{A \in T^*(K)} \{\omega;\ |n^{-d/2} \textstyle\sum_{j/n \in K} \xi_j(\omega) - a(K) - x(A)| \le \varepsilon\}$$

where $a(K) := s(A)$ is constant for $A \in T(K)$ and $T^*(K) := \cup_{i=1}^\infty \{\text{finite } 1/i\text{-net of } T(K)\}$. □

If we build $W | T$ on $C := C(T, \mathbf{R})$ with distribution $Q$ then $\mathscr{B}_b$ (for $C$) $\subset \mathscr{C}$ (for $C$) when $T$ is totally bounded, and the extension of $Q$ to $P$ on $D_0$ via $P(A) := Q(A \cap C)$ gives

**PROPOSITION 8.5.** *Sets in $\mathscr{B}_b$ and $\mathscr{C}$ are $P$ measurable if $r_2(T) \le r < 1$.* □

To state our main result we need some standard notation: given a measure $\mu$ on some measure space $X$, $\mathscr{A}$, let $\mu^*$ and $\mu_*$ denote the outer and inner measures induced by $\mu$ and let

$$\int^* f d\mu := \inf\left\{\int h\, d\mu;\ h \ge f,\ \int h\, d\mu \text{ defined}\right\}.$$

Define $\int_* f\, d\mu$ analogously.

**THEOREM 8.6.** *Suppose $T \subset T_0$ has $r_2(T) \le r < 1$. Let $P_n$ and $P$ denote the distribution of $Z_n | T$ and $W | T$ on $D_0$. Then*

(a) *$\{Z_n | T\}$ has small oscillations* iff

(b) $\operatorname{diam}\{\int^* fdP_n, \int_* fdP_n, \int^* fdP, \int_* fdP\} \to 0$

*for every $f: D_0 \to \mathbf{R}$ which is bounded and continuous when $D_0$ is given the uniform topology.*

*In particular,* (a) *implies*

(c) $\int fdP_n \to \int fdP$

*for every $f: D_0 \to \mathbf{R}$ which is bounded, continuous and $\mathscr{B}_b$ measurable.*

**PROOF.** This follows directly from Erickson and Fabian (1975, equations 2.8, 2.10 and 2.3). Without elaborating on definitions, let us check the conditions imposed in (2.10): convergence on $\pi_T$ means that $Z_n | T \to W | T$ in finite dimensional distribution; the domain of $P$ contains both $\mathscr{B}_b$ and $\mathscr{C}$, so also $\pi_T$. That $P$ and $\{P_n\}$ are almost simple follows from their small oscillations not only when $T$ is compact but also when $T$ is totally bounded: given $\varepsilon$, $\delta > 0$, a finite $\delta$-net $S_\delta$ for $T$ and $B_\delta(s) := \{t \in T,\ d(s, t) < \delta\}$, we have

$$\{x \in D_0;\ 1 - \operatorname{osc}(x, \delta) < \varepsilon\} \subset \{x \in D_0;\ \operatorname{diam} x(B_{\delta/2}(s)) < \varepsilon,\ s \in S_\delta\}. \quad □$$

## REFERENCES


BHATTACHARYA, R. N. and RAO, R. R. (1976). *Normal Approximation and Asymptotic Expansion.* Wiley, New York.

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

CHOW, Y. S. (1966). Some convergence theorems for independent random variables. *Ann. Math. Statist.* **37** 1482–1493.

DELPORTE, J. (1964). Fonctions aléatoires presque sûrement continues sur un intervalle fermé. *Ann. Inst. Henri-Poincare,* B. **1** 111–215.

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

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

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

DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators.* Interscience Publishers, New York.

ERICKSON, R. V. and FABIAN, V. (1975). On vague convergence of stochastic processes. *Ann. Probability.* **3** 1014–1022.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications.* Vol. II, 2nd ed. Wiley, New York.

FERNIQUE, X. (1974). Regularité des trajectoires des fonctions aléatoires gaussiennes. *Lecture Notes in Mathematics* **480** 1–96.

FERNIQUE, X. (1978). Caractérisation de processus a trajectoires majorées ou continues. *Lecture Notes in Mathematics* **649** 691–706.

GARSIA, A. M. and E. RODEMICH (1974). Monotonicity of certain functionals under rearrangement. *Ann. Inst. Fourier (Grenoble)* **24** 67–116.

HAHN, M. (1977). Conditions for sample continuity and the central limit theorem. *Ann. Probability.* **5** 351–360.

HAHN, M. and M. KLASS (1977). Sample-continuity of square-integrable processes. *Ann. Probability* **5** 361–370.

HEINKEL, B. (1978) Sur la relation entre théorème central-limite et loi du logarithme itéré dans les espaces de Banach. preprint.

IGARI, S. (1968). *Lectures on Fourier Series of Several Variables.* (unpublished).

KATZNELSON, Y. (1968). *An Introduction to Harmonic Analysis.* Wiley, New York.

LAMPERTI, J. (1962). On convergence of stochastic processes. *Trans. Am. Math. Soc.* **104** 430–435.

LOÈVE, M. (1963). *Probability Theory,* 3rd ed. Van Nostrand, Princeton, N.J.

NANOPOULOS, C. and NOBELIS, PH. (1978). Régularité et propriétés limites des fonctions aléatoires. *Lecture Notes in Mathematics* **649** 567–690.

ROSENTHAL, H. P. (1970). On the subspaces of $L^p(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8** 273–303.

VAPNIK, V. N. and A. YA. CHERVONENKIS (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theor. Probability Appl.* **16** 264–280.

WICHURA, M. (1969). Inequalities with applications to the weak convergence of random processes with multidimensional time parameters. *Ann. Math. Statist.* **40** 681–687.

WILANSKY, A. (1970). *Topology for Analysis.* Ginn, Waltham, Massachusetts.



MICHIGAN STATE UNIVERSITY
DEPARTMENT OF STATISTICS AND PROBABILITY
WELLS HALL
EAST LANSING, MICHIGAN 48824