The Annals of Probability
1986, Vol. 14, No. 4, 1329–1338

EMPIRICAL PROCESSES INDEXED BY LIPSCHITZ FUNCTIONS[1]

BY EVARIST GINÉ AND JOEL ZINN

Texas A & M University

Necessary and sufficient conditions on P for the unit balls of $BL(\mathbf{R})$ and $\mathrm{Lip}(\mathbf{R})$ to be functional P-Donsker classes are obtained.

1. Introduction and results. Given an interval $A \subseteq \mathbf{R}$, we let

$$BL(A) = \{f: A \to \mathbf{R}, \|f\|_\infty < \infty, \|f\|_L < \infty\}$$

where

$$\|f\|_L := \sup_{s \neq t,\, s,\, t \in A} \frac{|f(t) - f(s)|}{|t - s|}$$

is the Lipschitz pseudonorm of f and $\|f\|_\infty = \sup_{s \in A} |f(s)|$;

$$BL_1(A) = \{f \in BL(A): \|f\|_\infty \leq 1, \|f\|_L \leq 1\};$$

$$\mathrm{Lip}(A) = \{f: A \to \mathbf{R}, \|f\|_L < \infty\} \text{ and } \mathrm{Lip}_1(A) = \{f \in \mathrm{Lip}(A): \|f\|_L \leq 1\}.$$

Let P be a Borel probability measure on $\mathbf{R}$ and let X_i be i.i.d. random variables with law P (which we assume to be the coordinates of the infinite product probability space $(\mathbf{R}^N, B^N, P^N)$). The empirical measure corresponding to $\{X_i\}$ is

$$P_n = \sum_{i=1}^{n} \frac{\delta_{X_i}}{n}.$$

A class of measurable functions $F \subset B$ is a functional P-Donsker class if and only if $n^{1/2}(P_n - P)$ converges weakly as a sequence of random variables with values in $l^\infty(F)$. For the exact definitions as well as proofs we refer to [1], [2], and [5]. (Actually, in the unbounded case we consider some classes such that $\sup_{f \in F} |f(s)| = \infty$; in this case we say that F is functional P-Donsker if $F \subseteq \{f + c: f \in G, c \in \mathbf{R}\}$ and G is a functional P-Donsker class according to the usual definition.)

The purpose of this note is to prove the following

THEOREM 1. *A necessary and sufficient condition for $BL_1(\mathbf{R})$ to be a functional P-Donsker class is*

$$(1) \qquad \sum_{j=1}^{\infty} [\mathrm{pr}\{j - 1 < |X| \leq j\}]^{1/2} < \infty,$$

where $L(X) = P$.

Received December 1984; revised May 1985.

[1]Research partially supported by National Science Foundation grants DMS-83-18610 and MCS-83-01367.

AMS 1980 *subject classification*. Primary 60F17.

Key words and phrases. Empirical processes, functional Donsker classes, Lipschitz and bounded Lipschitz functions.

In the process of studying this problem we also noticed that

THEOREM 2. *A necessary and sufficient condition for* $\mathrm{Lip}_1(\mathbf{R})$ *to be a functional P-Donsker class is*

$$\sum_{j=1}^{\infty} [\mathrm{pr}\{|X| > j\}]^{1/2} < \infty, \tag{2}$$

where $L(X) = P$.

It turns out that almost any method works for Theorem 1. For example, one can

1. show that $BL(\mathbf{R})$ is the dual of a cotype 2 space, in fact of an L_1 space, thus reducing the problem to the well known central limit theorem in L_1 (see Proposition 1 below), or
2. give an essentially direct random entropy argument, or
3. obtain Theorem 1 as a corollary of Theorems 3.2 (5.1, 5.4) of [8], or
4. give a direct proof using summation by parts and the usual approximation ([12], [6]).

We became intrigued by this problem because of Stute's remark [16] to the effect that the techniques of [8] are not useful in the case of smooth functions. Nonetheless we are able to prove Theorem 1 by method 3 above. We also give a direct proof because it is reasonably elementary. We use method 1 in the proof of Theorem 2, and for this we must examine the relationship between Donsker classes and the central limit theorem in Banach spaces. Methods 2 and 4 also work for Theorem 2, which is in fact easier to prove than Theorem 1.

The norm induced by $BL_1(\mathbf{R})$ on measures of finite total variation (the dual bounded Lipschitz norm d_{BL^*}) metrizes weak (i.e., weak-star) convergence of probability measures and because of this it has some interest in statistics (see, e.g., [9]). In particular, it is interesting to obtain speeds of convergence in probability for the limit $d_{BL^*}(P_n, P) \to 0$. For instance, by Theorem 1, if P satisfies (1) then the sequence $\{n^{1/2} d_{BL^*}(P_n, P)\}_{n=1}^{\infty}$ converges in distribution and is therefore bounded in probability.

As far as we know, the first result on the problem considered here is found in Strassen and Dudley [14] where, as an application of their central limit theorem, they show (in present terminology) that $BL_1(\mathbf{R})$ is functional P-Donsker if P has bounded support. In [15] Stute considers the case of P with unbounded support. However, only boundedness in probability of $\{n^{1/2} d_{BL^*}(P_n, P)\}_{n=1}^{\infty}$ is obtained there, and assuming both a condition stronger than the existence of second moment for P, and a smoothness condition (on P). Notice that the condition in Theorem 1 is "between" $E|X| < \infty$ and $E|X|^{1+\delta} < \infty$ for some $\delta > 0$.

The problem in two or more dimensions seems to be of a different nature. For one thing, the identification $\mathrm{Lip}(\mathbf{R})/\mathbf{R} \leftrightarrow L_\infty(\mathbf{R})$ given by $f \to f'$ does not extend to $\mathbf{R}^d$, $d > 1$. Also, Strassen and Dudley show in [14] that $BL_1(\mathbf{R}^2)$ is not P-pregaussian if P is Lebesgue measure on the unit square.

2. Proofs. It can be easily proved (see, e.g., [10]; also [14]) that if Q is any finite measure on a *finite* interval A, then the $L_2(Q)$ covering number of $BL_1(A)$ satisfies

$$N\big(\varepsilon, BL_1(A), \|\cdot\|_{L_2(Q)}\big) \le N\big(\varepsilon/Q^{1/2}(A), BL_1(A), \|\cdot\|_\infty\big) \tag{3}$$
$$\le \exp\{c|A|Q^{1/2}(A)/\varepsilon\}$$

for all $\varepsilon > 0$, and for some constant c independent of A, Q, and ε. Using this estimate in the Strassen–Dudley central limit theorem or in Theorem 5.4 of [8], one immediately obtains

LEMMA 1 (Strassen and Dudley). *Let A be a finite interval. Then the class of functions $F = \{fI_A\colon f \in BL_1(\mathbf{R})\}$ is a functional P-Donsker class for any probability measure P on $\mathbf{R}$.*

PROOF OF THE SUFFICIENCY PART OF THEOREM 1, USING METHODS FROM [8]. Since $BL_1(\mathbf{R})$ is separable for the usual metric of uniform convergence on compact sets, it is l_2-deviation measurable for any P (in the notation of [8], Definition 2.2). Next we prove that if (1) holds, then $BL_1(\mathbf{R})$ is P-pregaussian. Define for $j \in \mathbf{Z}$ processes $\tilde{G}_j\colon f \in BL_1(\mathbf{R}) \to$ centered normal variables with covariances $E\tilde{G}_j(f)\tilde{G}_j(g) = Ef(X)g(X)I(j-1 < X \le j)$. Then

$$E|\tilde{G}_j(f) - \tilde{G}_j(g)|^2 \le \|f-g\|_\infty \mathrm{pr}(j-1 < X \le j)$$

and therefore, by [8], Theorem 2.15, and by (3), it follows that $\tilde{G}_j$ has a version G_j with bounded uniformly continuous paths on $(BL_1(\mathbf{R}), \|\cdot\|_{L_2(P)})$ and moreover,

$$E \sup_{f \in BL_1(\mathbf{R})} |G_j(f)| \le K\,[\mathrm{pr}(j-1 < X \le j)]^{1/2} \tag{4}$$

for some fixed constant K. Then if

$$G(f) = \sum_{j=-\infty}^{\infty} G_j(f), \tag{5}$$

where G_j are independent and have bounded uniformly continuous paths, so does G by (1) and (4) [(1) and (4) imply that (5) converges uniformly a.s. by Itô–Nisio's theorem]. Since the covariance of $G(f)$ is that of $f(X)$ [which dominates that of $f(X) - Ef(X)$] it follows that $BL_1(\mathbf{R})$ is P-pregaussian.

Hence, by Theorem 3.2 in [8], the result will follow if we prove that

$$\limsup_n \mathrm{pr}\left\{ \sup_{\substack{f, g \in BL_1(\mathbf{R}),\\ E(f(X)-g(X))^2 \le \varepsilon/n^{1/2}}} \left| \frac{\sum_{i=1}^n \varepsilon_i(f(X_i) - g(X_i))}{n^{1/2}} \right| > \tau\varepsilon \right\} = 0 \tag{6}$$

for some $\tau > 0$ and all $\varepsilon > 0$ (where $\{\varepsilon_i\}_{i=1}^\infty$ is a Rademacher sequence independent of $\{X_i\}$). By Lemma 1, for all $0 < r < \infty$ we have

$$\limsup_n \operatorname{pr}\left\{ \sup_{\substack{f,g\in BL_1(\mathbf{R}),\\ E(f(X)-g(X))^2\le \varepsilon/n^{1/2}}} \left| \frac{\sum_{i=1}^n \varepsilon_i (f(X_i)-g(X_i)) I_{[-r,r]}(X_i)}{n^{1/2}} \right| > \tau\varepsilon \right\} = 0. \tag{7}$$

On the other hand, by the proof of Theorem 5.1 of [8] [more concretely, by (5.13) there],

$$\begin{aligned} &\operatorname{pr}\left\{ \sup_{\substack{f,g\in BL_1(\mathbf{R}),\\ E(f(X)-g(X))^2\le \varepsilon/n^{1/2}}} \left| \frac{\sum_{i=1}^n \varepsilon_i (f(X_i)-g(X_i)) I_{[-r,r]^c}(X_i)}{n^{1/2}} \right| > \tau\varepsilon \right\} \\ &\qquad \le 5\operatorname{pr}\left\{\ln N_{n,1}\left(\varepsilon/n^{1/2}, BL_1([-r,r]^c)\right) > \sigma\varepsilon n^{1/2}\right\} + o(1), \end{aligned} \tag{8}$$

where $o(1)$ tends to zero as $n \to \infty$ independently of r, and σ is related to $\tau > 1$, but is independent of n and r. Here $N_{n,1}(\varepsilon/n^{1/2}, BL_1([-r,r]^c))$ denotes the $\varepsilon/n^{1/2}$-covering number with respect to the distance $d_{n,1}(f,g) = \sum_{i=1}^n |f(X_i) - g(X_i)|/n$ of the set $\{fI_{[-r,r]^c}\colon f \in BL_1(\mathbf{R})\}$, which we identify with $BL_1([-r,r]^c)$.

To compute the random entropy $\ln N_{n,1}$, set

$$C_{r,n} = \sum_{j=r+1}^{\infty} P_n^{1/2}(j-1 < |X| \le j)$$

and $I_j = \{x\colon j-1 < |x| \le j\}$. Then

$$\begin{aligned} d_{n,1}(f,g) &= \sum_{j=r+1}^{\infty} \sum_{k=1}^{n} |f(X_k) - g(X_k)| I_{I_j}(X_k)/n \\ &\le \sum_{j=r+1}^{\infty} \|(f-g)I_{I_j}\|_\infty P_n(I_j) \\ &= \sum_{j=r+1}^{\infty} \left(\|(f-g)I_{I_j}\|_\infty C_{r,n} P_n^{1/2}(I_j)\right) \frac{P_n^{1/2}(I_j)}{C_{r,n}}. \end{aligned}$$

Therefore, by (3), for all $\varepsilon > 0$

$$\begin{aligned} N_{n,1}(\varepsilon, BL([-r,r]^c)) &\le \prod_{j=r+1}^{\infty} N\left(\varepsilon, BL_1(I_j), \|\cdot\|_\infty P_n^{1/2}(I_j) C_{r,n}\right) \\ &\le \exp\left(c\varepsilon^{-1} \left(\sum_{j=r+1}^{\infty} P_n^{1/2}(I_j) \right)^2 \right) \end{aligned}$$

and

$$\operatorname{pr}\left\{\ln N_{n,1}\left(\frac{\varepsilon}{n^{1/2}}, BL_1([-r,r]^c)\right) > \sigma\varepsilon n^{1/2}\right\}$$
$$\leq \operatorname{pr}\left\{\sum_{j=r+1}^{\infty} P_n^{1/2}(I_j) > \left(\frac{\sigma\varepsilon^2}{c}\right)^{1/2}\right\}$$
$$\leq \left(\frac{c}{\sigma\varepsilon^2}\right)^{1/2} \sum_{j=r+1}^{\infty} EP_n^{1/2}(I_j)$$
$$\leq \left(\frac{c}{\sigma\varepsilon^2}\right)^{1/2} \sum_{j=r+1}^{\infty} \operatorname{pr}^{1/2}\{j-1<|X|\leq j\},$$

which can be made arbitrarily small by choosing r sufficiently large. Plug this into (8) and obtain $\limsup_r \limsup_n((7)+(8)) = 0$, thus proving (6). This gives the sufficiency part of the theorem. □

DIRECT PROOF OF THEOREM 1. Although randomization is not essential, it does shorten some expressions, so we will use Theorem 2.14 $(a) \Leftrightarrow (b)$ from [8]. By (3), $BL_1(\mathbf{R})$ is totally bounded in $L_2(P)$ for any P. So, for the sufficiency part it suffices to prove

$$\lim_{\delta\downarrow 0}\limsup_n \operatorname{pr}\left\{\sup_{\substack{f,g\in BL_1(\mathbf{R}),\\ E(f-g)^2(X))\leq\delta}}\left|\frac{\sum_{k=1}^n \varepsilon_k(f(X_k)-g(X_k))}{n^{1/2}}\right| > \varepsilon\right\} = 0 \tag{9}$$

for all $0<r<\infty$. The corresponding limit for $fI_{[-r,r]}$ instead of f and $gI_{[-r,r]}$ instead of g is zero by Lemma 1, for all $0<r<\infty$. Therefore, by the triangle inequality, it will be enough, for (9) to hold, that

$$\lim_{r\to\infty}\sup_n E\sup_{f\in BL_1(\mathbf{R})}\left|\frac{\sum_{k=1}^n \varepsilon_k f(X_k)I(|X_k|>r)}{n^{1/2}}\right| = 0. \tag{10}$$

Note that

$$\begin{aligned}
&\sum_{k=1}^n \varepsilon_k f(X_k)I(|X_k|>r)/n^{1/2}\\
&\quad= \sum_{k=1}^n\sum_{j=r+1}^{\infty} \varepsilon_k(f(X_k)-f(j-1))I(j-1<X_k\leq j)/n^{1/2}\\
&\qquad+ \sum_{k=1}^n\sum_{j=r+1}^{\infty} f(j-1)\varepsilon_k I(j-1<X_k\leq j)/n^{1/2}\\
&\qquad+ \sum_{k=1}^n\sum_{j=-r}^{\infty} \varepsilon_k(f(X_k)-f(j-1))I(j-1\leq X_k<j)/n^{1/2}\\
&\qquad+ \sum_{k=1}^n\sum_{j=-r}^{\infty} f(j-1)\varepsilon_k I(j-1\leq X_k<j)/n^{1/2}\\
&\quad:= (\mathrm{I})+(\mathrm{II})+(\mathrm{III})+(\mathrm{IV}).
\end{aligned} \tag{11}$$

We now handle the sums (I) and (II) ((III) and (IV) being obviously analogous). Here is (II):

$$
\begin{aligned}
(12)\qquad E \sup_{f\in BL_1(\mathbf{R})} |(\mathrm{II})| &\le \sum_{j=r+1}^{\infty} E\left|\frac{\sum_{k=1}^{n}\varepsilon_k I(j-1<X_k\le j)}{n^{1/2}}\right| \\
&\le \sum_{j=r+1}^{\infty} [\mathrm{pr}(j-1<X\le j)]^{1/2},
\end{aligned}
$$

which by (1) tends to zero as $r\to\infty$ uniformly in n. For (I) we will use integration by parts and the fact that if $f\in BL_1(\mathbf{R})$ then $\|f\|_\infty\le 1$ and $\|f'\|_\infty\le 1$, where f' is the derivative of f (f is an absolutely continuous function) and $\|f'\|_\infty$ denotes the ess sup of $|f'|$. We have, for $f\in BL_1(\mathbf{R})$,

$$
\begin{aligned}
&\sum_{k=1}^{n}\sum_{j=r+1}^{\infty}\varepsilon_k(f(X_k)-f(j-1))I(j-1<X_k\le j)/n^{1/2} \\
&\quad=\sum_{j=r+1}^{\infty}\sum_{k=1}^{n}\left[\left(\int_{j-1}^{X_k}f'(t)\,dt\right)\varepsilon_k I(j-1<X_k\le j)\right]\Big/ n^{1/2} \\
&\quad=\sum_{j=r+1}^{\infty}\int_{j-1}^{j}\left(\sum_{k=1}^{n}\varepsilon_k I(t<X_k\le j)\right)f'(t)\,dt/n^{1/2}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
(13)\qquad E\sup_{f\in BL_1(\mathbf{R})}|(I)| &\le \sum_{j=r+1}^{\infty}\int_{j-1}^{j}\sup_{j-1\le t\le j} E\left|\sum_{k=1}^{n}\varepsilon_k I(t<X_k\le j)\right|dt/n^{1/2} \\
&\le \sum_{j=r+1}^{\infty}[\mathrm{pr}\{j-1<X\le j\}]^{1/2},
\end{aligned}
$$

which tends to zero as $r\to\infty$ independently of n. (12) and (13), together with their analogues for (III) and (IV), give (10) and therefore the direct part of the theorem.

For the necessity part, consider the functions

$$
f_{3j}(x)=\begin{cases}0 & \text{for } x<3j-1,\\ 1 & \text{for } x=3j,\\ 1 & \text{for } x=3j+1,\\ 0 & \text{for } x>3j+2,\end{cases}\qquad \text{and linear in between.}
$$

For the class $\{\sum_{j=-\infty}^{\infty}\lambda_j f_{3j}\colon \lambda_j=\pm 1\}$ to be P-pregaussian it is necessary that the series $\sum_{j=-\infty}^{\infty}|g_{f_{3j}}|$ converge, where $g_{f_{3j}}$ are independent centered normal random variables with variance equal to the variance of $f_{3j}(X_1)$. But $\mathrm{Var}\, f_{3j}(X)\ge [1-P\{3j-1\le x\le 3j+2\}]^2 P\{3j\le x\le 3j+1\}$. So, if $\sum_{j=-\infty}^{\infty}|g_{f_{3j}}|$ converges a.s. then $\sum_{j=-\infty}^{\infty}(P\{3j\le x\le 3j+1\})^{1/2}<\infty$. Convergence of the other thirds of the series (1) follows in the same manner if one chooses the analogous classes of functions. (A similar argument can also be found in [14].) □

REMARK 1. Method 1 and Proposition 5.1 of [13] show that $BL_1(\mathbf{R})$ is a functional P-Donsker class if and only if

$$\left\{ \sup_{f \in BL_1(\mathbf{R})} |n^{1/2}(P_n - P)(f)| \right\}_{n=1}^{\infty} \tag{14}$$

is stochastically bounded (as, e.g., in the comment at the end of the proof of Theorem 2 below). In particular this, with Theorem 1, settles a question from [15]: if $dP(x) = c\,dx/x^{\alpha}$ for $x \geq a$ and for some c, α, and $a > 0$, then (14) is stochastically bounded if and only if $\alpha > 2$. One can also prove directly that if (1) fails then (14) is not stochastically bounded (in a similar way as it has been proved above that $BL_1(\mathbf{R})$ is not P-pregaussian). By uniform integrability (see, e.g., [12], Prop. and Remarque 2.1) (14) is stochastically bounded if and only if

$$\sup_n E \sup_{f \in BL_1(\mathbf{R})} |n^{1/2}(P_n - P)(f)| < \infty,$$

or equivalently (see, e.g., [8], Lemma 2.7),

$$\sup_n E \sup_{f \in BL_1(\mathbf{R})} \left| \sum_{k=1}^{n} \varepsilon_k f(X_k)/n^{1/2} \right| < \infty. \tag{15}$$

For concreteness, suppose $\sum_{j=1}^{\infty}[P(3j < x \leq 3j + 1)]^{1/2} = \infty$, and let $P_{3j} = P\{3j < x \leq 3j + 1\}$. Using Corollary 3.4 from [7], we have

$$\begin{aligned} &\sup_n E \sup_{f \in BL_1(\mathbf{R})} \left| \sum_{k=1}^{n} \varepsilon_k f(X_k) \right| \Big/ n^{1/2} \\ &\geq \sup_n E \sup_{f = \sum \lambda_j f_{3j},\, \lambda_j = \pm 1} \left| \sum_{k=1}^{n} \varepsilon_k f(X_k) \right| \Big/ n^{1/2} \\ &= \sup_n \sum_{j=1}^{\infty} E \left| \sum_{k=1}^{n} \varepsilon_k f_{3j}(X_k) \right| \Big/ n^{1/2} \\ &\geq C_1^{1/2} \sup_n \sum_{\{j:\, P_{3j} \geq 1/72n\}} P_{3j}^{1/2} = \infty, \end{aligned}$$

showing that (15) does not hold.

REMARK 2. Theorems 1 and 2 hold in more generality, e.g., for i.i.d. random measures ν_i instead of just i.i.d. point masses δ_{X_i}. Let us describe the analogue of Theorem 1. We let $S = M(\mathbf{R})$ be the set of measures of finite total variation on $\mathbf{R}$, with a σ-algebra S for which the total variation norm $\|\nu\|$ and the maps $\nu \to \int f\,d\nu$, $f \in BL_1(\mathbf{R})$ are all measurable. Then $BL_1(\mathbf{R})$ can be thought of as a family of functions on S: $f(\nu) = \int f\,d\nu$. Let P be a probability measure on $M(\mathbf{R})$ and let ν be a random measure with $L(\nu) = P$. We then have: If

$$\sum_{j=1}^{\infty} \left(E\|\nu I_{[j-1,j]}\|^2 \right)^{1/2} < \infty$$

then $BL_1(\mathbf{R})$ is a functional P-Donsker class, and if ν is positive this condition is also necessary for $BL_1(\mathbf{R})$ to be a functional P-Donsker class. The proof of this result differs from the direct proof of Theorem 1 only in formal details.

Before proving Theorem 2 we formalize the relationship between the central limit theorem (CLT) for a B-valued random variable Y with law Q (B a Banach space) and the P-Donsker property for the unit ball B_1' of the dual B' of B (and, for completeness, even in the nonseparable case). We use the notation $Y \in \mathrm{CLT}(B)$ to mean that if Y_i, with $L(Y_i) = L(Y)$, are the coordinate functions on an infinite product probability space, there exists a Radon centered gaussian measure γ such that for every $H: B \to \mathbf{R}$ bounded and continuous,

$$\int^* H\left(\sum_{i=1}^n Y_i/n^{1/2}\right) dpr \to \int H(x)\, d\gamma(x).$$

(see, e.g., Andersen [1], Definition 6.1; see [1] or [6] for the definition of $\int^*$; if B is separable this reduces to the usual definition of the CLT for Y). The following proposition, in this generality, is an immediate consequence of the main result in Andersen and Dobrić [21].

PROPOSITION 1. *Let B be a Banach space and let Y be a weakly centered Baire random variable on B with law Q. Then $Y \in CLT(B)$ if and only if B_1' is a functional Q-Donsker class.*

PROOF. If B is separable, this follows directly from Theorem 1 in Philipp [11] and Theorem 0.3 ($a \Leftrightarrow b$) in Dudley [4] (a proof of $a \Rightarrow b$ in this theorem has been kindly supplied to us by Dudley; see also [2], Theorem 3.2). In the general case we have, for Y and Q as in the statement of the proposition:

$$Y \in \mathrm{CLT}(B) \Leftrightarrow Y \in \mathrm{CLT}(l^\infty(B_1'))$$

$$\Leftrightarrow \begin{cases} B_1' \text{ is } Q\text{-pregaussian and } \left\{\sum_{i=1}^n f(Y_i)/n^{1/2}\colon f \in B_1'\right\} \\ \text{is eventually uniformly equicontinuous for the} \\ L_2(Q) \text{ distance in } B_1'. \end{cases}$$

$$\Leftrightarrow B_1' \text{ is a functional } Q\text{-Donsker class.}$$

The first implication follows by Tietze's extension theorem and the third is just Theorem 4.1.1 in [5]. The crucial second implication follows from Theorem 5.5 and 4.1 in [2] (see also their example 5.7). □

PROOF OF THEOREM 2. We must show that condition (2) is equivalent to $\mathrm{Lip}_1^0(\mathbf{R}) = \{f \in \mathrm{Lip}_1(\mathbf{R}), f(0) = 0\}$ being a functional P-Donsker class. Since by, e.g., 2.12 and 2.13 in [8], the condition

$$n\mathrm{pr}\left\{\sup_{f \in \mathrm{Lip}_1^0(\mathbf{R})} |f(X)| > n^{1/2}\right\} = n\mathrm{pr}\{|X| > n^{1/2}\} \to 0$$

is necessary for $\mathrm{Lip}_1^0(\mathbf{R})$ to be P-Donsker, we may assume in particular that $\int_{-\infty}^{\infty}|x|\,dP(x) < \infty$. Note that for $f \in \mathrm{Lip}(\mathbf{R})$,

$$\int_{-\infty}^{\infty} f d(P_n - P) = \int_{-\infty}^{0} (P_n - P)\{x \le t\} f'(t)\,dt + \int_{0}^{\infty} (P_n - P)\{x \ge t\} f'(t)\,dt. \tag{16}$$

Define

$$Y(t) = \begin{cases} (\delta_X - P)\{x \le t\} & \text{for } t < 0, \\ (\delta_X - P)\{x \ge t\} & \text{for } t \ge 0, \end{cases}$$

and $Y_i(t)$ as $Y(t)$ with X_i replacing X. Then $Y, Y_i \in L_1(\mathbf{R})$ a.s. since $\int_{-\infty}^{\infty}|x|\,dP(x) < \infty$. Denote

$$\langle Y, g\rangle = \int_{-\infty}^{\infty} Y(t) g(t)\,dt, \qquad g \in L_\infty(\mathbf{R}).$$

Then (16) can be rewritten as

$$(P_n - P)(f) = \left\langle \sum_{i=1}^{n} Y_i/n^{1/2}, f' \right\rangle, \qquad f \in \mathrm{Lip}(\mathbf{R}). \tag{16'}$$

Since the correspondence $f \leftrightarrow f'$ is one-to-one (and isometric) between $\mathrm{Lip}_1^0(\mathbf{R})$ and the unit ball of $L_\infty(\mathbf{R})$, which we will denote by B_1', (16′) and the Dudley–Philipp $L_2(P)$-equicontinuity criterion for Donsker classes ([5], Theorem 4.1.1) imply that

(17) $\mathrm{Lip}_1^0(\mathbf{R})$ is a functional P-Donsker class if and only if the unit ball of $L_\infty(R)$, B_1', is a functional Q-Donsker class, where $Q = L(Y)$.

Hence, by Proposition 1, $\mathrm{Lip}_1^0(\mathbf{R})$ is a functional P-Donsker class if and only if $Y \in \mathrm{CLT}(L_1(\mathbf{R}))$.

Now, since $L_1(\mathbf{R})$ is of cotype 2,

$$Y \in \mathrm{CLT}(L_1(\mathbf{R})) \Leftrightarrow Y \text{ is pregaussian} \Leftrightarrow \int_{-\infty}^{\infty} \left(EY^2(t)\right)^{1/2} dt < \infty. \tag{18}$$

(See, e.g., exercises 3.8.13 and 3.8.14 in [3].) But this last integral condition is precisely (2) (more exactly, the equivalent one $\int_{-\infty}^{0}(\mathrm{pr}\{X \le t\} - \mathrm{pr}^2\{X \le t\})^{1/2}\,dt + \int_0^\infty(\mathrm{pr}\{X \ge t\} - \mathrm{pr}^2\{X \ge t\})^{1/2}\,dt < \infty$). □

It is worthwhile to note that since $L_1(\mathbf{R})$ is of cotype 2, the random vector Y in the above proof satisfies the CLT if and only if the sequence $\{\|\sum_{i=1}^n Y_i/n^{1/2}\|_{L_1(\mathbf{R})}\}_{n=1}^\infty$ is stochastically bounded (Proposition 5.1 in [13]). This shows by (16′) that condition (2) is also necessary for the sequence

$$\left\{\sup_{f \in \mathrm{Lip}_1(\mathbf{R})} |n^{1/2}(P_n - P)(f)|\right\}_{n=1}^{\infty}$$

to be stochastically bounded.

Acknowledgments. We thank R. Dudley and G. Pisier for interesting conversations on this and on related topics.

REFERENCES

[1] ANDERSEN, N. T. (1984). A central limit theorem for non-separable valued functions. Research Report No. 119, Math. Inst., Aarhus Univ.

[2] ANDERSEN, N. T. and DOBRIĆ, V. (1984). The central limit theorem for stochastic processes. Research Report No. 124, Math. Inst., Aarhus Univ.

[3] ARAUJO, A. and GINÉ, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, New York.

[4] DUDLEY, R. M. (1973). Sample functions of the Gaussian process. *Ann. Probab.* **1** 66–103. Complemented by lecture notes for a course at MIT.

[5] DUDLEY, R. M. (1984). A course on empirical processes. *Lecture Notes in Math.* **1097** 2–142. Springer, New York.

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

[7] GINÉ, E. and ZINN, J. (1983). Central limit theorems and weak laws of large numbers in certain Banach spaces. *Z. Wahrsch. verw. Gebiete* **62** 323–352.

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

[9] HUBER, P. (1981). *Robust Statistics*. Wiley, New York.

[10] KOLMOGOROV, A. N. and TIKHOMIROV, V. M. (1961). The ε-entropy and ε-capacity of sets in functional spaces. *Amer. Math. Soc. Translations* (2) **17** 277–364. (From *Uspekhi Mat. Nauk* **14** (1959) 3–86.)

[11] PHILIPP, W. (1980). Weak and L^p-invariance principles for sums of B-valued random variables. *Ann. Probab.* **8** 68–82; Correction. *Ann. Probab.* **14** (1986) 1095–1101.

[12] PISIER, G. (1975). Le théorème limite central et la loi du logarithme iteré dans les espaces de Banach. Sém. Maurey-Schwartz 75-76, Exp. III et IV. Ecole Polytechnique, Paris.

[13] PISIER, G. and ZINN, J. (1978). On the limit theorems for random variables with values in L_p, $p \leq 2$. *Z. Wahrsch. verw. Gebiete* **41** 289–304.

[14] STRASSEN, V. and DUDLEY, R. M. (1969). The central limit theorem and ε-entropy. *Lecture Notes in Math.* **89** 224–231. Springer, New York.

[15] STUTE, W. (1983). Empirical processes indexed by smooth functions. *Stochastic Process. Appl.* **14** 55–66.

[16] STUTE, W. (1984). Discussion of the paper of Giné and Zinn. *Ann. Probab.* **12** 997–998.

DEPARTMENT OF MATHEMATICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843