[12] — *A Kronecker theorem in functional analysis*, Colloq. Math. (to appear).
[13] A. N. Kolmogorov and V. M. Tihomirov, *ε-entropy and ε-capacity of sets in function spaces*, Uspehi 14 (86) (1959), pp. 3–86. Also in AMS Translations 17 (2) (1961), pp. 277–364.
[14] T. W. Körner, *Some results on Kronecker, Dirichlet, and Helson sets*, Annales Fourier (Grenoble) 20 (1970), pp. 219–324.
[15] G. G. Lorentz, *Lower bounds for the degree of approximation*, Trans. Amer. Math. Soc. 97 (1960), pp. 25–34.
[16] — *Entropy and approximation*, Bull. Amer. Math. Soc. 72 (1966), pp. 903–937.
[17] — *Metric entropy, widths, and superpositions of functions*, Amer. Math. Monthly 69 (1962), pp. 469–489.
[18] — *Approximation of Functions*, Holt, Rinehart, and Winston 1966.
[19] Y. Meyer, *Algebraic numbers and harmonic analysis*, North-Holland 1972.
[20] W. Rudin, *Real and Complex Analysis*, McGraw-Hill 1966.
[21] R. Salem, *On singular monotonic functions whose spectrum has a given Hausdorff dimension*, Ark. Mat. 1 (1951), pp. 353–365.
[22] A. C. Vosburg, *Metric entropy of certain classes of Lipschitz functions*, Proc. Amer. Math. Soc. 17 (1966), pp. 665–669.
[23] A. Zygmund, *Trigonometric Series, I and II*, Cambridge 1959 and 1968.

Received August 23, 1973 (715)

STUDIA MATHEMATICA, T. LII. (1974)

An inequality for the distribution of a sum of certain Banach space valued random variables

by

J. KUELBS* (Madison, Wi.)

Abstract. We prove an inequality for the distribution of a sum of independent Banach space valued random variables provided they take values in a space having a norm with a smooth second directional derivative and the random variables have $2+\delta$ moments. This inequality is applied to obtain the central limit theorem and the law of the iterated logarithm, and it is shown that these results apply to the L^p spaces, $2 \leqslant p < \infty$.

1. Introduction. Throughout the paper B is a real separable Banach space with norm $\|\cdot\|$, and all measures on B are assumed to be defined on the Borel subsets of B generated by the norm open sets. We denote the topological dual of B by B^*.

A measure μ on B is called a *mean zero Gaussian measure* if every continuous linear function f on B has a mean zero Gaussian distribution with variance $\int_B [f(x)]^2 \mu(dx)$. The bilinear function T defined on B^* by

$$T(f,g) = \int_B f(x)g(x)\mu(dx) \qquad (f,g \in B^*)$$

is called the *covariance function of* μ. It is well known that a mean zero Gaussian measure on B is uniquely determined by its covariance function. This is so because T uniquely determines μ on the Borel subsets of B generated by the weakly open sets, and since B is separable, the Borel sets generated by the weakly open sets are the same as those generated by the norm open sets.

However, a mean zero Gaussian measure μ on B is also determined by a unique subspace H_μ of B which has a Hilbert space structure. The norm on H_μ will be denoted by $\|\cdot\|_\mu$ and it is well known that the B norm $\|\cdot\|$ is weaker than $\|\cdot\|_\mu$ on H_μ. In fact, $\|\cdot\|$ is a measurable norm on H_μ in the sense of [7]. Since $\|\cdot\|$ is weaker than $\|\cdot\|_\mu$ it follows that B^* can be linearly embedded (by the restriction map) into the dual of H_μ, call it

* Supported in part by NSF Grant GP-18759.

H_μ^*, and identifying H_μ with H_μ^* in the usual way we have $B^* \subseteq H_\mu \subseteq B$. Then, by the main result in [7], the measure μ is the extension of the canonical normal distribution on H_μ to B. We describe this relationship by saying *μ is generated by H_μ*. For details on these matters as well as additional references see [7], [8].

Let $X_1, X_2, \ldots$ be independent B-valued random variables such that for some $\delta > 0$

$$\sup_k E\|X_k\|^{2+\delta} < \infty, \tag{1.1}$$

$$E(X_k) = 0 \quad (k = 1, 2, \ldots),$$

and having common covariance function

$$T(f, g) = E\big(f(X_k)g(X_k)\big) \quad (k = 1, 2, \ldots).$$

Then, if T is the covariance function of a mean zero Gaussian measure μ on B, and the norm on B has a second directional derivative with certain smoothness properties, we have for $t \geqslant 0$ and any $\beta > 0$ that

$$P\left(\left\|\frac{X_1 + \ldots + X_n}{\sqrt{n}}\right\| \geqslant t\right) \leqslant \mu(x\colon \|x\| \geqslant t - \beta) + O(n^{-\min(\alpha,\delta)/2}), \tag{1.2}$$

where the bounding constant is uniform in t for $t \geqslant 2\beta$ and $\alpha > 0$ determines a Lipschitz condition on the second directional derivative of the norm.

This result is proved in Theorem 2.1 using a method that is due to Trotter [13]. The application of Trotter's method in this setting hinges on the fact that the Gaussian measure μ determines a generating Hilbert space H_μ in B, and also on a number of nontrivial properties of Gaussian measures on B.

In Theorem 3.1 and Theorem 3.2 we apply the inequality (1.2) to obtain the central limit theorem and the law of the iterated logarithm. Finally, in Section 4 we show that the L^p spaces $2 \leqslant p < \infty$ satisfy the smoothness condition we impose on the second directional derivative of the norm on B.

Previous results and further references regarding the central limit theorem in the Banach space setting can be found in [4] and [6]. The law of the iterated logarithm for Hilbert space valued random variables was proved in [10], and for Gaussian random variables with values in an arbitrary Banach space in [11], [12].

2. The basic inequality.

The norm $\|\cdot\|$ on B is *twice directionally differentiable* on $B - \{0\}$ if for $x, y \in B$, $x + ty \neq 0$, we have

$$\frac{d}{dt}\|x + ty\| = D(x + ty)(y), \tag{2.1}$$

where $D\colon B - \{0\} \to B^*$ is measurable from the Borel subsets of B generated by the norm topology to the Borel subsets of B^* generated by the weak-star topology, and

$$\frac{d^2}{dt^2}\|x + ty\| = D^2_{x+ty}(y, y), \tag{2.2}$$

where D_x^2 is a bounded bilinear form on $B \times B$. We call D_x^2 the *second directional derivative* of the norm, and without loss of generality we can assume D_x^2 is a symmetric bilinear form. That is, if T_x is a bilinear form which satisfies (2.2) then $A_x(y, z) = [T_x(y, z) + T_x(z, y)]/2$ also satisfies (2.2) and A_x is symmetric. Hence in all that follows we assume D_x^2 is a symmetric bilinear form. Of course, if the norm is actually twice Fréchet differentiable on B with second derivative at x given by A_x, then it is well known that A_x is a symmetric bilinear form on $B \times B$, and in this case D_x^2 would be equal to A_x since symmetric bilinear forms are uniquely determined on the diagonal of $B \times B$.

If $D_x^2(y, y)$ is continuous in x ($x \neq 0$) and for all $r > 0$ and $x, h \in B$ such that $\|x\| \geqslant r$ and $\|h\| \leqslant r/2$, we have

$$|D^2_{x+h}(h, h) - D^2_x(h, h)| \leqslant C_r\|h\|^{2+\alpha} \tag{2.3}$$

for some fixed $\alpha > 0$ and some constant C_r, then we say *the second directional derivative is* Lip(α) *away from zero*.

We now mention briefly a number of properties which $D(x)$ and D_x^2 enjoy. For example, it is easily shown that $D(\lambda x) = (\operatorname{sgn}\lambda)D(x)$ for all $x, \lambda \neq 0$ and also that $\|D(x)\| = 1$ for $x \neq 0$. Somewhat more involved, but of no great difficulty, is the following lemma regarding the second derivative.

LEMMA 2.1. *If the norm $\|\cdot\|$ is twice directionally differentiable on $B - \{0\}$ with second derivative D_x^2, then:*

(a) *If $\lambda \neq 0$, $x \neq 0$ then* $D^2_{\lambda x} = \dfrac{1}{|\lambda|}D^2_x$.

(b) $D^2_x(h, h) \geqslant 0$ *for all* $x, h \in B$.

Proof. For $\lambda, x \neq 0$ and $h \in B$

$$\begin{aligned} D^2_{\lambda x}(h, h) &= \lim_{t\to 0}\frac{D(\lambda x + th)(h) - D(\lambda x)(h)}{t} \\ &= \lim_{t\to 0}\frac{\lambda}{t}\big[D\big(\lambda(x + th/\lambda)\big)(h/\lambda) - D(\lambda x)(h/\lambda)\big] \\ &= \lim_{t\to 0}\frac{\lambda\operatorname{sgn}\lambda}{t}\,[D(x + th/\lambda)(h/\lambda) - D(x)(h/\lambda)] \\ &= |\lambda|\,D^2_x(h/\lambda, h/\lambda) = \frac{1}{|\lambda|}D^2_x(h, h). \end{aligned} \tag{2.4}$$

Hence $D^2_{\lambda x} = \frac{1}{|\lambda|} D^2_x$ since a symmetric bilinear form is determined by its values on the diagonal of $B \times B$. Now the non-negativity of D^2_x follows because the existence of the second derivative of $\|x+th\|$ implies

$$D^2_x(h, h) = \lim_{t\to 0} \frac{\|x+th\|+\|x-th\|-2\|x\|}{t^2}, \tag{2.5}$$

and since $\|2x\| \leqslant \|x+th\|+\|x-th\|$, we easily see that $D^2_x(h, h) \geqslant 0$.

THEOREM 2.1. *Let B denote a real separable Banach space with norm $\|\cdot\|$. Let $\|\cdot\|$ be twice directionally differentiable on B with the second directional derivative D^2_x being* Lip(a) *away from zero for some $a > 0$ and such that*

$$\sup_{\|x\|=1} \|D^2_x\| < \infty.$$

Let $X_1, X_2, \ldots$ be independent B-valued random variables such that for some $\delta > 0$

$$\sup_k E\|X_k\|^{2+\delta} < \infty, \tag{2.6}$$

$$EX_k = 0 \quad (k = 1, 2, \ldots)$$

and having common covariance function

$$T(f, g) = E\big(f(X_k)g(X_k)\big) \quad (f, g \in B^*).$$

Then, if T is the covariance function of a mean zero Gaussian measure μ on B, it follows for $t \geqslant 0$ and any $\beta > 0$ that

$$P\left(\left\|\frac{X_1+\ldots+X_n}{\sqrt{n}}\right\| \geqslant t\right) \leqslant \mu(x: \|x\| \geqslant t-\beta) + O(n^{-\min(a,\delta)/2}), \tag{2.7}$$

where the bounding constant is uniform in t for $t \geqslant 2\beta$.

Proof. If $0 \leqslant t \leqslant \beta$ then (2.7) is obvious, so fix $t > \beta$ and define a function $f: (-\infty, \infty) \to [0, 1]$ such that f is monotone increasing, $f(u) = 0$ for $u \leqslant t-\beta$, $f(u) = 1$ for $u \geqslant t$, and f'' is Lipschitz continuous (and hence in this case bounded) on $(-\infty, \infty)$. Let $g(x) = f(\|x\|)$, $W_n = (X_1+\ldots+X_n)/\sqrt{n}$, and assume $Y_1, Y_2, \ldots$ are independent random variables each with Gaussian distribution μ. To be specific we assume the sequences $\{X_k\}$ and $\{Y_k\}$ are defined on the probability space, $(\Omega, \mathfrak{F}, P)$. We also assume the Y_k's are independent of the X_k's and that $Z_n = (Y_1+\ldots+Y_n)/\sqrt{n}$. Then the distribution Z_n induced on B is μ and

$$\begin{aligned} P(\|W_n\| \geqslant t) &= \mu(x: \|x\| \geqslant t) + \{P(\|W_n\| \geqslant t) - \mu(x: \|x\| \geqslant t)\} \\ &\leqslant \mu(x: \|x\| \geqslant t-\beta) + E\{g(W_n)-g(Z_n)\}. \end{aligned} \tag{2.8}$$

Now

$$g(W_n)-g(Z_n) = \sum_{k=1}^{n} V_k,$$

where

$$\begin{aligned} V_k &= g\left(\frac{X_1+\ldots+X_k+Y_{k+1}+\ldots+Y_n}{\sqrt{n}}\right) - \\ &\quad - g\left(\frac{X_1+\ldots+X_{k-1}+Y_k+\ldots+Y_n}{\sqrt{n}}\right) \\ &= g(U_k+X_k/\sqrt{n}) - g(U_k+Y_k/\sqrt{n}) \end{aligned} \tag{2.9}$$

and

$$U_k = (X_1+\ldots+X_{k-1}+Y_{k+1}+\ldots+Y_n)/\sqrt{n}.$$

Let $h(\lambda) = g(U_k+\lambda X_k/\sqrt{n})$ for $-\infty < \lambda < \infty$. Since $g(x) = f(\|x\|)$ and f vanishes in a neighborhood of zero we have $h(\lambda)$ twice continuously differentiable on $(-\infty, \infty)$. Hence by Taylor's formula

$$\begin{aligned} g(U_k+X_k/\sqrt{n}) &= h(0)+h'(0)+\frac{h''(0)}{2}+\frac{[h''(\tau)-h''(0)]}{2} \\ &= g(U_k)+f'(\|U_k\|)D(U_k)(X_k/\sqrt{n}) + \\ &\quad + \tfrac{1}{2}f''(\|U_k\|)\{D(U_k)(X_k/\sqrt{n})\}^2 + \\ &\quad + \tfrac{1}{2}f'(\|U_k\|)D^2_{U_k}(X_k/\sqrt{n}, X_k/\sqrt{n}) + J_n(U_k, X_k) \\ &\qquad\qquad (0 < \tau < 1), \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} 2J_n(U_k, X_k) &= f''(\|U_k+\tau X_k/\sqrt{n}\|)\{D(U_k+\tau X_k/\sqrt{n})(X_k/\sqrt{n})\}^2 + \\ &\quad + f'(\|U_k+\tau X_k/\sqrt{n}\|)D^2_{U_k+\tau X_k/\sqrt{n}}(X_k/\sqrt{n}, X_k/\sqrt{n}) - \\ &\quad - f''(\|U_k\|)\{D(U_k)(X_k/\sqrt{n})\}^2 - \\ &\quad - f'(\|U_k\|)D^2_{U_k}(X_k/\sqrt{n}, X_k/\sqrt{n}) \end{aligned} \tag{2.11}$$

and τ is a non-negative random variable bounded by one.

A similar expression holds for $g(U_k+Y_k/\sqrt{n})$ except Y_k replaces X_k and τ is replaced by a random variable τ^* which is also non-negative and bounded by one.

We will show below that

$$\begin{aligned} E\big(f'(\|U_k\|)D(U_k)(X_k)\big) &= E\big(f'(\|U_k\|)D(U_k)(Y_k)\big) = 0, \\ E\big(f''(\|U_k\|)\{D(U_k)(X_k)\}^2\big) &= E\big(f''(\|U_k\|)\{D(U_k)(Y_k)\}^2\big), \\ E\big(f'(\|U_k\|)D^2_{U_k}(X_k, X_k)\big) &= E\big(f'(\|U_k\|)D^2_{U_k}(Y_k, Y_k)\big) \end{aligned} \tag{2.12}$$

and hence by (2.9) and (2.10), we have

$$|E(V_k)| \leqslant E|J_n(U_k, X_k)| + E|J_n(U_k, Y_k)|. \tag{2.13}$$

Further, by showing both $E|J_n(U_k, X_k)|$ and $E|J_n(U_k, Y_k)|$ are $O(n^{-(1+\min(\alpha,\delta)/2)})$, we see from (2.13) that

$$|E\{g(W_n) - g(Z_n)\}| = O(n^{-\min(\alpha,\delta)/2}). \tag{2.14}$$

Combining (2.14) and (2.8) we get (2.7). Hence the theorem is proved provided (2.12) holds and the above estimates of $|E(J_n(U_k, X_k))|$ and $|E(J_n(U_k, Y_k))|$ are uniform in t for $t > 2\beta$.

We first establish the equalities in (2.12). Since $D: B - \{0\} \to B^*$ is measurable from the Borel subsets of B to the Borel subsets of B^* generated by the weak-star topology we have $f'(\|U_k\|)D(U_k)$ defined with probability one on Ω (recall f' vanishes in a neighborhood of zero) and it is measurable from $\mathfrak{F}_k$ (the minimal sigma algebra making U_k measurable from Ω to B) to the weak-star Borel sets of B^*. Furthermore, we actually have $D: B - \{0\} \to S = \{x^* \in B^*: \|x^*\|_{B^*} \leqslant 1\}$ and S is a compact metric space in the weak-star topology. Now f, f', and f'' are uniformly bounded on $(-\infty, \infty)$ so henceforth in the proof we let

$$C = \sup_{-\infty<u<\infty} \{|f(u)| + |f'(u)| + |f''(u)|\}. \tag{2.15}$$

(It is obvious that C can be taken uniform in t for $t > 2\beta$.) Thus by standard arguments there is a sequence of B^*-valued random variables $\{A_n: n \geqslant 1\}$ each of which takes on only finitely many values and is measurable from $\mathfrak{F}_k$ to the weak-star Borel subsets of B^* such that

$$\|A_n\|_{B^*} \leqslant C, \qquad \lim_n A_n = f'(\|U_k\|)D(U_k) \tag{2.16}$$

with probability one, where the convergence in (2.16) is in the weak-star sense. That is,

$$A_n(\cdot) = \sum_{j=1}^{j_n} \chi_{E_{j,n}}(\cdot) f_{j,n},$$

where $f_{j,n} \in C\cdot S \subseteq B^*$ and $E_{j,n} \in \mathfrak{F}_k$.

Since $E\|X_k\| < \infty$ we have by the dominated convergence theorem that

$$E(f'(\|U_k\|)D(U_k)(X_k)) = \lim_n E(A_n(X_k))$$
$$= \lim_n E\Big(\sum_{j=1}^{j_n} \chi_{E_{j,n}} f_{j,n}(X_k)\Big) = \lim_n \sum_{j=1}^{j_n} P(E_{j,n})\cdot 0 = 0$$

since $\chi_{E_{j,n}}$ is independent of $f_{j,n}(X_k)$. When X_k is replaced by Y_k we get the same result since Y_k is independent of U_k so the first equality in (2.12) holds.

Using the fact that the covariance functions for X_k and Y_k are the same, the above argument can be applied to show that the second equality in (2.12) also holds. Here, of course, we use the fact that (2.6) holds and $E\|Y_k\|^2 = E(Y_1\|^2 < \infty$ (see [5]) when we apply the dominated convergence theorem.

Now the bounded bilinear form D_x^2 is non-negative by Lemma 2.1 (b) and symmetric by assumption, and hence there is a non-negative operator $A_x: B \to B^*$ such that

$$D_x^2(y, z) = (A_x y, z) \qquad (y, z \in B).$$

If H_μ is the Hilbert space in B which generates μ on B (see [8] for details) then we know the identity map $i: H_\mu \to B$ is compact (see [7]). Identifying H_μ^* with H_μ, we have $B^* \subseteq H_\mu^* = H_\mu \subseteq B$ and hence A_x restricted to the Hilbert space H_μ is a non-negative compact symmetric operator. Thus the spectral theorem for compact symmetric operators on H_μ implies that for each $z \in H_\mu \subseteq B$

$$A_x(z) = \sum_j \lambda_j(x)\big(z, e_j(x)\big)e_j(x), \tag{2.17}$$

where $\{e_j(x): j \geqslant 1\}$ are orthonormal eigenvectors for A_x corresponding to the eigenvalues $\{\lambda_j(x): j \geqslant 1\}$ all of which are non-negative. Note that $e_j(x) \in B^* \subseteq H_\mu^* = H_\mu$ since $A_x: B \to B^*$.

Let $I_x(z) = (A_x z, z)^{1/2}$ for $z \in B$. Then I_x is a continuous semi-norm on B and for $z \in H_\mu \subseteq B$ we have

$$I_x^2(z) = \sum_j \lambda_j(x)\big(z, e_j(x)\big)^2. \tag{2.18}$$

In fact, we actually have (2.18) holding for all $z \in M = \bar{H}_\mu$ (the closure of H_μ in B). To see this note that for $z \in M$ and $\{y_n\} \subseteq H_\mu$ such that $y_n \to z$ we have by Fatou's lemma that

$$I_x^2(z) = \lim_j I_x^2(y_n) \geqslant \sum_j \lambda_j(x)\big(z, e_j(x)\big)^2. \tag{2.19}$$

To show equality in (2.19) for $z \in M$ note that if we define

$$\|z\|_1 = \Big(\sum_j \lambda_j(x)\big(z, e_j(x)\big)^2\Big)^{1/2}, \tag{2.20}$$

then $\|\cdot\|_1$ is finite on M by (2.19) and

$$\|z\|_1^2 \leqslant I_x^2(z) \leqslant \|A_x\|\,\|z\|^2, \tag{2.21}$$

so $\|\cdot\|_1$ is also a continuous semi-norm on M. Now $\|z\|_1 = I_x(z)$ for $z \in H_\mu$ so equality holds in (2.19) for $z \in M = \bar{H}_\mu$. Since it is well known that the support of μ is M (see, for example, [8]) and μ and the X_k's have common covariance operator T it follows easily that the support of each X_k is a subset of M. Hence with probability one we have

$$(2.22)\qquad \begin{aligned} D_x^2(X_k, X_k) &= \sum_j \lambda_j(x)(X_k, e_j(x))^2, \\ D_x^2(Y_k, Y_k) &= \sum \lambda_j(x)(Y_k, e_j(x)). \end{aligned}$$

Now we can choose a sequence of M-valued random variables $\{R_n: n \geq 1\}$ such that each R_n is finite-valued, $\mathfrak{F}_k$ measurable, and such that

$$\lim_n R_n = U_k,$$

where convergence is in the B norm.

To be specific assume $R_n = \sum_{r=1}^{r_n} x_r(n)\chi_{E_{r,n}}$, where the $E_{r,n}$'s are disjoint and $\mathfrak{F}_k$ measurable. Then, by the dominated convergence theorem with dominating function $C\cdot \sup_{\|x\|=1} \|D_x^2\|/(t-\beta)\cdot[\|X_k\|^2+\|Y_k\|^2]$, we have (since $D_x^2(y, y)$ is continuous in x away from zero) that

$$(2.23)\qquad E\big(f'(\|U_k\|)D_{U_k}^2(X_k, X_k)\big) = E\big(\lim_n f'(\|R_n\|)D_{R_n}^2(X_k, X_k)\big)$$

$$= \lim_n \sum_{r=1}^{r_n} f'(\|x_r(n)\|)P(E_{r,n})E\big(D_{x_r(n)}^2(X_k, X_k)\big)$$

by independence of the $E_{r,n}$'s and X_k

$$= \lim_n \sum_{r=1}^{r_n} f'(\|x^r(n)\|)P(E_{r,n})E\big(D_{x_r(n)}^2)(Y_k, Y_k)\big)$$

by the common covariance of X_k and Y_k, and the representations (2.22)

$$= E\big(f'(\|U_k\|)D_{U_k}^2(Y_k, Y_k)\big)$$

by reversing the previous steps and U_k being independent of Y_k.

Hence the third equality in (2.12) holds.

We also point out that for $x \neq 0$

$$E\big(D_x^2(Y_k, Y_k)\big) < \infty$$

implies $\sum_j \lambda_j(x) < \infty$ so D_x^2 is actually trace class on H_μ. That is, by (2.22)

and that each $\lambda_j(x) \geq 0$ we have

$$\sum_j \lambda_j(x) = E\big(D_x^2(Y_k, Y_k)\big) < \infty.$$

We now turn to the proof that

$$(2.24)\qquad E\big(|J_n(U_k, X_k)|\big) = O\big(n^{-(1+\min(\alpha,\delta)/2)}\big)$$

and that the bounding constant is uniform in t for $t \geq 2\beta$.

Set $\gamma = \min\left(\frac{t-\beta}{4}, \beta\right)$ and let $E = \{x: \|x\| \leq \gamma\}$ and $E' = \{x: \|x\| > \gamma\}$ throughout the remainder of the proof.

First note that from (2.11) and (2.15) we have

$$(2.25)\qquad \begin{aligned} &\chi_{E'}(\|X_k\|/\sqrt{n})|2J_n(U_k, X_k)| \\ &\quad \leq \chi_{E'}(\|X_k\|/\sqrt{n})[2C\|X_k\|^2/n + 2C\sup_{\|x\|\geq\gamma}\|D_x^2\|\|X_k\|^2/n] \end{aligned}$$

since f' and f'' vanish on $(-\infty, t-\beta]$.

If $\|X_k\|/\sqrt{n} \leq \gamma$ we have two cases to consider. They are

$$\text{(a)}\qquad \|U_k + \tau X_k/\sqrt{n}\| \leq \frac{3}{4}(t-\beta),$$

$$\text{(b)}\qquad \|U_k + \tau X_k/\sqrt{n}\| > \frac{3}{4}(t-\beta).$$

Now case (a) is simple since $\|X_k\|/\sqrt{n} \leq \gamma \leq \frac{t-\beta}{4}$ and (a) implies $J_n(U_k, X_k) = 0$ since f' and f'' vanish on $(-\infty, t-\beta]$.

Now $\|X_k\|/\sqrt{n} \leq \gamma$ and (b) implies $\|U_k\| \geq \frac{t-\beta}{2}$. For $x, y \in B$ and $0 < \tau < 1$ we have

$$(2.26)\qquad \begin{aligned} 2J_n(x, \sqrt{n}y) = {} & [f''(\|x+\tau y\|) - f''(\|x\|)][D(x+\tau y)(y)]^2 + \\ & + f''(\|x\|)\big[(D(x+\tau y)(y))^2 - (D(x)(y))^2\big] + \\ & + [f'(\|x+\tau y\|) - f'(\|x\|)]D_{x+\tau y}^2(y, y) + \\ & + f'(\|x\|)[D_{x+\tau y}^2(y, y) - D_x^2(y, y)]. \end{aligned}$$

We now will estimate the right-hand side of (2.26) under the assumptions $\|y\| \leq \gamma \leq \frac{t-\beta}{4}$ and $\|x\| \geq \frac{t-\beta}{2}$.

Let C' denote a positive constant which dominates the Lipschitz constants of both f' and f'', and recall C from (2.15). Note that C' can be made uniform in t since $t > \beta$.

Then we have

$$\begin{aligned}(2.27)\quad &\chi_E(\|y\|)\,|[f''(\|x+\tau y\|)-f''(\|x\|)][D(x+\tau y)(y)]^2| \\ &\qquad\leqslant \min(2C, C'\|y\|)\,\chi_E(\|y\|)\,\|y\|^2, \\ &\chi_E(\|y\|)\,|f'(\|x+\tau y\|)-f'(\|x\|)|\,|D^2_{x+\tau y}(y, y)| \\ &\qquad\leqslant \min(2C, C'\|y\|)\,\chi_E(\|y\|)\cdot \sup_{\|z\|\geqslant\frac{3}{4}(t-\beta)} \|D_z^2\|\cdot\|y\|^2.\end{aligned}$$

Further, since $\|D(x)\| = 1$ for $x \neq 0$ we have

$$\begin{aligned}(2.28)\quad &\chi_E(\|y\|)\,|f''(\|x\|)|\,\big|\big(D(x+\tau y)(y)\big)^2-\big(D(x)(y)\big)^2\big| \\ &\qquad\leqslant 2C\|y\|\,|D(x+\tau y)(y)-D(x)(y)|\,\chi_E(\|y\|) \\ &\qquad= 2C\|y\|\,\Big|\int_0^\tau \frac{d^2}{dt^2}\|x+ty\|\,dt\Big|\,\chi_E(\|y\|)\end{aligned}$$

since $\|x\| \geqslant \dfrac{t-\beta}{2}$, $\|y\| \leqslant \gamma \leqslant \dfrac{t-\beta}{4}$, $0 < \tau < 1$

$$\leqslant 2C\|y\| \sup_{\|z\|\geqslant\frac{t-\beta}{4}} \|D_z^2\|\cdot\|y\|^2\cdot\chi_E(\|y\|).$$

Finally, since D_x^2 is Lip(α) away from zero we have for $\|x\| \geqslant \dfrac{t-\beta}{2}$, $\|y\| \leqslant \gamma \leqslant \dfrac{t-\beta}{4}$ that

$$(2.29)\quad \chi_E(\|y\|)\,|f'(\|x\|)|\,|D^2_{x+\tau y}(y, y)-D^2_x(y, y)| \leqslant C\cdot C_{\frac{t-\beta}{2}}\|y\|^{2+\alpha}\chi_E(\|y\|),$$

where C_r is defined as in (2.3).

Consequently, by combining (2.27), (2.28), and (2.29), we have for $\|x\| \geqslant \dfrac{t-\beta}{2}$, $\|y\| \leqslant \gamma \leqslant \dfrac{t-\beta}{4}$, $0 < \delta \leqslant 1$, and $0 < \alpha \leqslant 1$ a positive constant C'' (which is uniform in t for $t \geqslant 2\beta$) such that

$$(2.30)\quad \chi_E(\|y\|)\cdot|2J_n(x, \sqrt{n}y)| \leqslant C''\|y\|^{2+\min(\alpha,\delta)}\chi_E(\|y\|).$$

Combining (2.25) and (2.30) we get from (2.6) that

$$E|2J_n(U_k, X_k)| = O\big(E(\|X_k/\sqrt{n}\|^{2+\min(\alpha,\delta)})\big) = O\big(n^{-(1+\min(\alpha,\delta)/2)}\big),$$

where the bounding constant is uniform in t for $t > 2\beta$. A similar estimate is valid for $|E(2J_n(U_k, Y_k))|$, so the theorem is proved.

3. Applications of the basic inequality. We will apply Theorem 2.1 to obtain the central limit theorem and the law of the iterated logarithm for a sequence of B-valued random variables $\{X_k\}$.

THEOREM 3.1. *Let B and $\{X_k\}$ satisfy the conditions in Theorem 2.1, and assume μ is a Gaussian measure on B with covariance function T. Then, if μ_n denotes the measure induced on B by $(X_1+\ldots+X_n)/\sqrt{n}$, we have $\lim_n \mu_n = \mu$ in the sense of weak convergence.*

Proof. For each $f \in B^*$ let $\mu^f(\mu_n^f)$ denote the distribution of f on $(-\infty, \infty)$ with respect to $\mu(\mu_n)$. For any nonempty subset A of B let $A^\varepsilon = \{x: \|x-A\| < \varepsilon\}$, where

$$\|x-A\| = \inf_{y\in A}\|x-y\|.$$

Fix $\varepsilon > 0$. Now by standard finite-dimensional arguments, we have for each $f \in B^*$ that $\mu_n^f \xrightarrow[n\to\infty]{} \mu^f$, where the convergence is in the sense of weak convergence. Hence by [1, Theorem 2.3] it suffices to prove there exists a finite-dimensional subspace E of B such that

$$(3.1)\quad \mu_n(E^\varepsilon) > 1-\varepsilon.$$

Let H_μ denote the Hilbert space in B which generates μ on B. Let $\{a_k\}$ be a complete orthonormal sequence in H_μ which lies in B^* under the usual embedding of B^* in H_μ. Such a sequence exists since B^* is dense in H_μ (see [9] for details). We define

$$(3.2)\quad \Pi_N(x) = \sum_{k=1}^N a_k(x)a_k \quad (N = 1, 2, \ldots),$$

$$Q_N(x) = x-\Pi_N x = (I-\Pi_N)(x),$$

where $a_k(x)$ denotes the value of the linear functional corresponding to a_k at the point x.

Now it is well known (see, for example, [9]) that $Q_N x \to 0$ with μ-probability one on B. Further, if μ^{Q_N} denotes the measure μ induced on B under the mapping Q_N, then it is easy to see that μ^{Q_N} is a mean-zero Gaussian measure on B with generating Hilbert space $Q_N H_\mu = \{x \in H_\mu: x = \sum_{k\geqslant N+1} a_k(x)a_k\}$. Applying Theorem 2.1 to the random variables $\{Q_N X_k, k \geqslant 1\}$, we have by (2.7) with $t = \varepsilon$, $\beta = \varepsilon/2$ that

$$(3.3)\quad P\left(\left\|\frac{Q_N X_1+\ldots+Q_N X_n}{\sqrt{n}}\right\| \geqslant \varepsilon\right) \leqslant \mu^{Q_N}(x: \|x\| \geqslant \varepsilon/2) + O\big(n^{-\min(\alpha,\delta)/2}\big).$$

Since $Q_N x \to 0$ with μ-probability one and $\mu^{Q_N}(A) = \mu(x: Q_N x \in A)$ we have an N_0 such that for $N \geqslant N_0$

$$(3.4)\quad \mu^{Q_N}(x: \|x\| \geqslant \varepsilon/2) \leqslant \varepsilon/4.$$

Thus we can choose n_0 such that $n \geqslant n_0$ implies

$$(3.5)\qquad P\left(\left\|\frac{Q_{N_0}X_1+\ldots+Q_{N_0}X_n}{\sqrt{n}}\right\| \geqslant \varepsilon\right) < \varepsilon/2+\varepsilon/2 = \varepsilon.$$

Now for $n < n_0$ we can choose N_1 such that

$$(3.6)\qquad P\left(\left\|\frac{Q_{N_1}X_1+\ldots+Q_{N_1}X_n}{\sqrt{n}}\right\| \geqslant \varepsilon\right) < \varepsilon.$$

Let $E = \Pi_{N_2}B$, where $N_2 = \max(N_0, N_1)$. Then by (3.5) and (3.6) for every $n \geqslant 1$

$$\begin{aligned}\mu_n(E^\varepsilon) &= \mu_n(x\colon \|x-E\| < \varepsilon) = 1-\mu_n(x\colon \|x-E\| \geqslant \varepsilon)\\ &\geqslant 1-\mu_n(x\colon \|Q_{N_0}x\| \geqslant \varepsilon)\,\chi_{\{k: k\geqslant n_0\}}(n)-\\ &\quad -\mu_n\{x\colon \|Q_{N_1}x\| \geqslant \varepsilon)\,\chi_{\{k: k< n_0\}}(n) \geqslant 1-\varepsilon,\end{aligned}$$

and the theorem is proved.

We now turn to the law of the iterated logarithm in this setting. $\mathrm{LL}n$ denotes $\log\log n$ if $n \geqslant 3$ and 1 for $n = 1, 2$.

THEOREM 3.2. *Let B and $\{X_k\}$ satisfy the conditions in Theorem 2.1, and assume μ is a Gaussian measure on B with covariance function T. If K is the unit ball of the Hilbert space H_μ which generates μ then*

$$(3.7)\qquad P\left(\lim_n\left\|\frac{X_1+\ldots+X_n}{\sqrt{2n\,\mathrm{LL}n}}-K\right\| = 0\right) = 1,$$

$$(3.8)\qquad P\left(C\left(\left\{\frac{X_1+\ldots+X_n}{\sqrt{2n\,\mathrm{LL}n}}\right\}\right) = K\right) = 1,$$

where $C(\{a_n\})$ denotes the cluster set of the sequence $\{a_n\}$.

Remark. It is known (see, for example, [11]) that K is a compact subset of B thus (3.7) implies that with probability one the sequence $\left\{\frac{X_1+\ldots+X_n}{\sqrt{2n\,\mathrm{LL}n}}\right\}$ is conditionally compact in B.

Proof. Let $S_n = X_1+\ldots+X_n$ for $n \geqslant 1$. To prove (3.7) it suffices to show that for each rational $\varepsilon > 0$ we have with probability one that $S_n/\sqrt{2n\,\mathrm{LL}n} \notin K^\varepsilon$ only finitely often. Fix $\varepsilon > 0$ and let $A_n = \{S_n/\sqrt{2n\,\mathrm{LL}n} \notin K^\varepsilon\}$. Let $n_r = [\beta^r]$, where $[\cdot]$ denotes the greatest integer function and $\beta > 1$. Let

$$B_r = \left\{\frac{S_n}{\sqrt{2n_r\,\mathrm{LL}n_r}} \notin K^\varepsilon \quad \text{for some } n\colon n_r \leqslant n < n_{r+1}\right\}.$$

Then $\limsup_n A_n \subseteq \limsup_r B_r$.

Now if $Y_1, \ldots, Y_N$ are successive sums of independent B-valued random variables such that

$$\sup_{1\leqslant j\leqslant N} P(\|Y_N-Y_j\| > \varepsilon/2) = c < 1,$$

then by a standard argument we can show

$$P\Big(\bigcup_{j=1}^{N}\{Y_j \notin K^\varepsilon\}\Big) \leqslant \frac{1}{1-c}P(Y_N \notin K^{\varepsilon/2}).$$

Hence

$$(3.9)\qquad P(B_r) \leqslant \frac{1}{1-d}P\left(\frac{S_{n_{r+1}-1}}{\sqrt{n_{r+1}-1}} \notin \sqrt{\frac{n_r}{n_{r+1}-1}}\sqrt{2\,\mathrm{LL}n_r}\,K^{\varepsilon/2}\right),$$

where

$$(3.10)\qquad d = \sup_{n_r\leqslant n<n_{r+1}} P\left(\|S_{n_{r+1}-1}-S_n\| > \tfrac{1}{2}\varepsilon\sqrt{2n_r\,\mathrm{LL}n_r}\right) \leqslant \tfrac{1}{2}$$

provided r is sufficiently large and $\beta > 1$. That is, if $n_{r+1}-1-n \leqslant N$ then the supremum of the corresponding probability over these n is $\leqslant 1/2$ provided r is sufficiently large. On the other hand, by Theorem 2.1, there is an N such that $(n_{r+1}-1-n) > N$ implies the supremum of the corresponding probability over these n is also $\leqslant 1/2$ provided r is sufficiently large. Hence (3.10) is valid.

Now K compact in B (and hence bounded) implies that for any fixed $\varepsilon > 0$ there exists a $\beta > 1$ sufficiently close to one and a $\delta > 0$ such that for all r sufficiently large

$$(3.11)\qquad \sqrt{\frac{n_r}{n_{r+1}-1}}K^{\varepsilon/2} \supseteq K^\delta.$$

Choosing $\delta > 0$ and $\beta > 1$ so that (3.11) holds, we obtain from (3.9) that for all r sufficiently large

$$(3.12)\qquad P(B_r) \leqslant 2P\left(\frac{S_{n_{r+1}-1}}{\sqrt{n_{r+1}-1}} \notin \sqrt{2\,\mathrm{LL}n_r}\,K^\delta\right).$$

Let $\{\Pi_N\colon N \geqslant 1\}$ and $\{Q_N\colon N \geqslant 1\}$ be defined as in (3.2). Then for each $N = 1, 2, \ldots$ we have

$$\begin{aligned}(3.13)\qquad & P\left(\frac{S_{n_{r+1}-1}}{\sqrt{n_{r+1}-1}} \notin \sqrt{2\,\mathrm{LL}n_r}\,K^\delta\right)\\ &\leqslant P\left(\Pi_N\left(\frac{S_{n_{r+1}-1}}{\sqrt{n_{r+1}-1}}\right) \notin \sqrt{2\,\mathrm{LL}n_r}\,\Pi_N(K_\mu^\gamma)\right)+P\left(\left\|Q_N\left(\frac{S_{n_{r+1}-1}}{\sqrt{n_{r+1}-1}}\right)\right\| \geqslant \frac{\delta}{2}\cdot\sqrt{2\,\mathrm{LL}n_r}\right),\end{aligned}$$

where $\gamma > 0$ is such that if $K_\mu^\gamma = \{x \in H_\mu\colon \|x-K\|_\mu < \gamma\}$ then $\Pi_N K_\mu^\gamma \subseteq \Pi_N K^{\delta/2}$. The existence of such a $\gamma > 0$ is obvious since $\Pi_N B$ is finite-dimensional and hence the norms $\|\cdot\|$ and $\|\cdot\|_\mu$ are equivalent on $\Pi_N B$

(γ may depend on N but this will be no problem). Further,

$$\Pi_N K_\mu^\gamma = \{x \in \Pi_N H_\mu \colon \|x\|_\mu \leqslant 1+\gamma\}. \tag{3.14}$$

Fix $c > 1$ and choose λ so that $2\lambda(\delta/2)^2 \geqslant c$. Since $\|Q_N x\| \to 0$ with μ-probability one, we have by [5] that there exists an N_0 such that $N \geqslant N_0$ implies

$$\int_B \exp\{\lambda\|Q_N x\|^2\} d\mu(x) < \infty. \tag{3.15}$$

For fixed $N \geqslant N_0$ and λ we then obtain

$$\mu(\|Q_N x\| \geqslant \tfrac{1}{2}\delta\sqrt{2\mathrm{LL}n_r}) \leqslant \exp\{-c\,\mathrm{LL}n_r\} \tag{3.16}$$

for all sufficiently large r.

Hence fix $N \geqslant N_0$, $C > 1$, and choose $\gamma > 0$ as in (3.13). Combining (3.16), (3.14), and Theorem 2.1 to the random variables $\{\Pi_N X_k \colon k \geqslant 1\}$ and $\{Q_N X_k \colon k \geqslant 1\}$, we have by (3.13) that

$$P\left(\frac{S_{n_{r+1}-1}}{\sqrt{n_{r+1}-1}} \notin \sqrt{2\mathrm{LL}n_r}K^\delta\right) \leqslant \mu(x\colon \|\Pi_N x\|_\mu \geqslant (1+\gamma)\sqrt{2\mathrm{LL}n_r}-1) + \mu(x\colon \|Q_N x\| \geqslant \tfrac{1}{2}\delta\sqrt{2\mathrm{LL}n_r}-1) + O(n_r^{-\min(\alpha,\delta)/2}). \tag{3.17}$$

Now it follows easily that there exists a $d > 1$ such that

$$\mu(x\colon \|\Pi_N x\|_\mu \geqslant (1+\gamma)\sqrt{2\mathrm{LL}n_r}-1) \leqslant \exp\{-d\,\mathrm{LL}n_r\} \tag{3.18}$$

for all large r.

Combining (3.18), (3.17), and (3.16) we see

$$\sum_r P(B_r) < \infty.$$

Then $P(\limsup_n A_n) = 0$ and (3.7) holds.

To prove (3.8) one can proceed exactly as in the second part of Theorem 3.1 in [10] except that for (3.16) in [10] we need estimates of the type used under the assumption of $2+\delta$ moments and not 3 moments. Since the range space of the random variables in this inequality is finite dimensional we can prove rather easily using the techniques of Theorem 2.1 of [10] that the error is $O(n_r^{-\gamma/8})$ where $\gamma > 0$ is $\gamma = \min(\alpha, \delta)$. This estimate allows us to complete the proof.

We also mention that Strassen's functional form of the law of the iterated logarithm for B-valued variables can also be proved in this setting using (2.7) and the techniques developed in [10] when B was a real separable Hilbert space, but it will not be included here.

4. Some spaces with smooth norm. L^p denotes the real vector space $L^p(S, \Sigma, m)$, where m is a sigma-finite positive measure on S. Then L^p is a real separable Banach space provided $p \geqslant 1$ and if $p \geqslant 2$ we will show the usual norm on L^p is twice directionally differentiable. We also show that the second directional derivative is $\mathrm{Lip}(\alpha)$ away from zero. These results are suggested by those in [3], but do not seem to be immediate corollaries of [3].

THEOREM 4.1. *If $p \geqslant 2$ and if for $x \in L^p(S, \Sigma, m)$ we define*

$$\|x\| = \left\{\int_S |x(s)|^p\, dm(s)\right\}^{1/p}$$

then the norm $\|\cdot\|$ has two directional derivatives and the second derivative is $\mathrm{Lip}(\alpha)$ away from zero with $\alpha = 1$ if $p = 2$ or $p \geqslant 3$ and $\alpha = p-2$ for $2 < p < 3$. Furthermore,

$$\sup_{\|x\|=1} \|D_x^2\| \leqslant 2(p-1).$$

Proof. It is easy to prove that the first and second directional derivative of the norm are given for $\|x+ty\| \neq 0$ by

$$\frac{d}{dt}\|x+ty\| = \frac{\int |x+ty|^{p-1}\operatorname{sgn}(x+ty)\,y\,dm}{\|x+ty\|^{p-1}} \tag{4.1}$$

and

$$\frac{d^2}{dt^2}\|x+ty\| = (p-1)\left[\frac{\int |x+ty|^{p-2}y^2\,dm}{\|x+ty\|^{p-1}} - \frac{\{\int |x+ty|^{p-1}\operatorname{sgn}(x+ty)\,y\,dm\}^2}{\|x+ty\|^{2p-1}}\right], \tag{4.2}$$

where $\operatorname{sgn}(x)$ equals $+1$ if $x > 0$, -1 if $x < 0$, 0 if $x = 0$, and is interpreted as $+1$ in (4.1) and (4.2) if $p = 2, 4, 6, \ldots$ We do not include the details here as (4.1) and (4.2) are easily obtained directly, or by applying the results in [3] which yield first and second Fréchet derivatives for $\|\cdot\|$.

In the notation of (2.1) and (2.2) we have for $\|x\| \neq 0$ that

$$D(x)(y) = \frac{\int |x|^{p-1}\operatorname{sgn}(x)\,y\,dm}{\|x\|^{p-1}}, \tag{4.3}$$

$$D_x^2(y, y) = (p-1)\left[\frac{\int |x|^{p-2}y^2\,dm}{\|x\|^{p-1}} - \frac{(\int |x|^{p-1}\operatorname{sgn}(x)\,y\,dm)^2}{\|x\|^{2p-1}}\right], \tag{4.4}$$

where $\operatorname{sgn}(x)$ is interpreted as above.

Using the fact that the dual of L^p $(2 \leqslant p < \infty)$ is $L^{p'}$ for p' such that $1/p+1/p' = 1$ it follows that $D\colon L^p - \{0\} \to L^{p'}$ is continuous so clearly D is measurable in the sense required. Furthermore, the computations involved in establishing the continuity of D are similar (but simpler) than those showing D_x^2 is $\mathrm{Lip}(\alpha)$ away from zero so we only demonstrate the $\mathrm{Lip}(\alpha)$ property.

From (4.4) $D_x^2(y, y)$ is easily seen to be continuous in x $(x \neq 0)$. Now for $\|x\| \geqslant r$ and $\|h\| \leqslant r/2$, where $r > 0$ let

$$(4.5)\qquad \begin{aligned} I &= \left| \frac{\int |x+h|^{p-2} h^2 dm}{\|x+h\|^{p-1}} - \frac{\int |x|^{p-2} h^2 dm}{\|x\|^{p-1}} \right|, \\ J &= \left| \frac{\left(\int |x+h|^{p-1} \operatorname{sgn}(x+h)\, h\, dm\right)^2}{\|x+h\|^{2p-1}} - \frac{\left(\int |x|^{p-1} \operatorname{sgn}(x)\, h\, dm\right)^2}{\|x\|^{2p-1}} \right|. \end{aligned}$$

To prove the second directional derivative of the norm has the Lip(a) property it is easily seen from (4.2) that it suffices to prove that there exists a constant C_r such that

$$(4.6)\qquad |I| \leqslant \frac{C_r}{2} \|h\|^{2+a}$$

and

$$(4.7)\qquad |J| \leqslant \frac{C_r}{2} \|h\|^{2+a},$$

where $a = 1$ if $p = 2$ or $p \geqslant 3$ and $a = p-2$ for $2 < p < 3$.

Now

$$\begin{aligned} (4.8)\qquad |I| &\leqslant \frac{1}{\|x+h\|^{p-1}} \left| \int [|x+h|^{p-2} h^2 - |x|^{p-2} h^2]\, dm \right| + \\ &\quad + \left| \int |x|^{p-2} h^2 dm \right| \left| \frac{1}{\|x+h\|^{p-1}} - \frac{1}{\|x\|^{p-1}} \right| \\ &\leqslant \frac{1}{\|x+h\|^{p-1}} \int \left| |x+h|^{p-2} - |x|^{p-2} \right| h^2 dm + \\ &\quad + \frac{\|x\|^{p-2} \|h\|^2}{\|x+h\|^{p-1} \|x\|^{p-1}} \left| \|x+h\|^{p-1} - \|x\|^{p-1} \right|. \end{aligned}$$

To estimate (4.8) we use a lemma due to Banach and Saks [2] which asserts that if a, b are any two real numbers and $1 \leqslant p < \infty$, then there exists a constant M independent of a, b such that

$$(4.9)\qquad \left| |a+b|^p - |a|^p \right| \leqslant M |b|^p + \sum_{j=1}^{[p]} \binom{p}{j} |a|^{p-j} |b|^j,$$

where $[t]$ = greatest integer in t.

Using (4.9) with $a = \|x+h\|$, $b = \|x\| - \|x+h\|$, we have

$$\begin{aligned} (4.10)\qquad \left| \|x\|^{p-1} - \|x+h\|^{p-1} \right| &= \left| (\|x+h\| + (\|x\| - \|x+h\|))^{p-1} - \|x+h\|^{p-1} \right| \\ &\leqslant M \|h\|^{p-1} + \sum_{j=1}^{[p-1]} \binom{p-1}{j} \|x+h\|^{p-1-j} \|h\|^j. \end{aligned}$$

Hence for $\|x\| \geqslant r$, $\|h\| \leqslant r/2$ we have a constant M_r such that

$$(4.11)\qquad \frac{1}{\|x+h\|^{p-1} \|x\|} \left| \|x\|^{p-1} - \|x+h\|^{p-1} \right| \leqslant M_r \|h\|.$$

Using (4.9) again we see that if $p-2 \geqslant 1$ there exists a constant M' such that

$$\begin{aligned} (4.12)\qquad \int \left| |x+h|^{p-2} - |x|^{p-2} \right| h^2 dm &\leqslant M' \int \Big\{ |h|^p + \sum_{j=1}^{[p-2]} |x|^{p-2-j} |h|^{2+j} \Big\} dm \\ &\leqslant M' \Big\{ \|h\|^p + \sum_{j=1}^{[p-2]} \|x\|^{p-2-j} \|h\|^{2+j} \Big\}. \end{aligned}$$

Thus if $p-2 \geqslant 1$, $\|x\| \geqslant r$, $\|h\| \leqslant r/2$ we have a constant M_r' such that

$$(4.13)\qquad \frac{1}{\|x+h\|^{p-1}} \int \left| |x+h|^{p-2} - |x|^{p-2} \right| h^2 dm \leqslant M_r' \|h\|^3.$$

If $0 < p-2 < 1$ then

$$(4.14)\qquad \int \left| |x+h|^{p-2} - |x|^{p-2} \right| h^2 dm \leqslant \int |h|^p dm = \|h\|^p = \|h\|^{2+a},$$

where $a = p-2$, and if $p = 2$ (4.14) is zero. Combining (4.14), (4.13), (4.11), and (4.8), we get a constant C_r such that for $\|x\| \geqslant r$, $\|h\| \leqslant r/2$ we have (4.6) holding.

Now

$$\begin{aligned} (4.15)\qquad |J| &\leqslant \frac{1}{\|x+h\|^{2p-1}} \left| \left(\int |x+h|^{p-1} \operatorname{sgn}(x+h)\, h\, dm \right)^2 - \right. \\ &\quad \left. - \left(\int |x|^{p-1} \operatorname{sgn}(x)\, h\, dm \right)^2 \right| + \\ &\quad + \left(\int |x|^{p-1} |h|\, dm \right)^2 \left(\frac{1}{\|x+h\|^{2p-1}} - \frac{1}{\|x\|^{2p-1}} \right) \\ &\leqslant \frac{\|x+h\|^{p-1} \|h\| + \|x\|^{p-1} \|h\|}{\|x+h\|^{2p-1}} \times \\ &\quad \times \left| \int [|x+h|^{p-1} \operatorname{sgn}(x+h) - |x|^{p-1} \operatorname{sgn}(x)]\, h\, dm \right| + \\ &\quad + \frac{\|x\|^{2(p-1)} \|h\|^2}{\|x+h\|^{2p-1} \|x\|^{2p-1}} \left| \|x\|^{2p-1} - \|x+h\|^{2p-1} \right|. \end{aligned}$$

Proceeding as in (4.10), (4.11) we have for $\|x\| \geqslant r$, $\|h\| \leqslant r/2$ a constant N_r such that

$$(4.16)\qquad \frac{\|h\|^2}{\|x\| \, \|x+h\|^{2p-1}} \left| \|x\|^{2p-1} - \|x+h\|^{2p-1} \right| \leqslant N_r \|h\|^3.$$

If $p-1=1$ $(p=2)$ then

$$(4.17)\quad \int [|x+h|^{p-1}\operatorname{sgn}(x+h)-|x|^{p-1}\operatorname{sgn}(x)]\,h\,dm = \int h^2\,dm = \|h\|^2,$$

and if $p-1>1$ then by the mean-value theorem applied to the function $f(x) = |x|^{p-1}\operatorname{sgn} x$ we have

$$(4.18)\quad \big||a+b|^{p-1}\operatorname{sgn}(a+b)-|a|^{p-1}\operatorname{sgn}(a)\big| \leqslant (p-1)\sup_{0\leqslant\tau<1}|a+\tau b|^{p-2}|b| \leqslant (p-1)[|a+b|^{p-2}+|a-b|^{p-2}]|b|.$$

Hence from (4.18) we obtain for $p-1>1$ that

$$(4.19)\quad \Big|\int [|x+h|^{p-1}\operatorname{sgn}(x+h)-|x|^{p-1}\operatorname{sgn}(x)]\,h\,dm\Big| \leqslant (p-1)\Big[\int |x+h|^{p-2}h^2\,dm+\int |x-h|^{p-2}\,h^2\,dm\Big] \leqslant (p-1)\|h\|^2[\|x+h\|^{p-2}+\|x-h\|^{p-2}].$$

Combining (4.19), (4.17), (4.16), and (4.15) we see that for $\|x\|\geqslant r$, $\|h\|\leqslant r/2$ we have a constant C_r such that (4.6) and (4.7) hold.

Finally, the estimate $\sup_{\|x\|=1}\|D_x^2\|\leqslant 2(p-1)$ follows since

$$\sup_{\|x\|=1}\|D_x^2\| = \sup_{\|x\|=1}\sup_{\substack{\|y\|\leqslant 1\\ \|z\|\leqslant 1}}|D_x^2(y,z)| = \sup_{\|x\|=1}\sup_{\substack{\|y\|\leqslant 1\\ \|z\|\leqslant 1}}\left|D_x^2\Big(\frac{y+z}{2},\frac{y+z}{2}\Big)-D_x^2\Big(\frac{y-z}{2},\frac{y-z}{2}\Big)\right| \leqslant \sup_{\|x\|=1}\big(2\sup_{\|w\|\leqslant 1}|D_x^2(w,w)|\big)\leqslant 2(p-1).$$

Hence the theorem is proved.

Using Theorem 4.1 we see the L^p spaces $(2\leqslant p<\infty)$ satisfy the conditions used in sections 2 and 3. Thus the central limit theorem and the law of the iterated logarithm are valid in these spaces for sequences of random variables of the type employed in Theorems 2.1, 3.1, and 3.2. A central limit theorem in this setting was known previously due to [6], but the law of the iterated logarithm for non-Gaussian random variables is new when $p>2$.

Bibliography

[1] A. D. de Acosta, *Existence and convergence of probability measures in Banach spaces*, Trans. Amer. Math. Soc. 152 (1970), pp. 273–298.

[2] S. Banach and S. Saks, *Sur la convergence forte dans les champs L^p*, Studia Math. 2 (1930), pp. 52–57.

[3] R. Bonic and J. Frampton, *Smooth functions on Banach manifolds*, Journal of Math. and Mechanics 15 (1966), pp. 877–898.

[4] R. M. Dudley, *Metric entropy and the central limit theorem in $C(S)$* (to appear in the Annales Institute Fourier Grenoble).

[5] X. Fernique, *Integrabilite des vecteurs Gaussiens*, C. R. Acad. Sci. Paris 270 (1970), pp. 1698–1699.

[6] R. Fortet and E. Mourier, *Les fonctions aléatoires comme éléments aléatoires dans les espace de Banach*, Studia Math. 19 (1955), pp. 62–79.

[7] L. Gross, *Lectures in modern analysis and applications II*, Vol. 140, Lecture notes in mathematics, Springer-Verlag, New York.

[8] J. Kuelbs, *Some results for probability measures on linear topological vector spaces with an application to Strassen's log log law*, Journal of Functional Analysis 14 (1973), pp. 28–43.

[9] — *Expansions of vectors in a Banach space related to Gaussian measures*, Proc. Amer. Math. Soc. 27 (1971), pp. 364–370.

[10] — and T. Kurtz, *Berry–Essen estimates in Hilbert space and an application to the law of the iterated logarithm*, Annals of Probability 2 (1974), pp.

[11] — and R. LePage, *The law of the iterated logarithm for Brownian motion in a Banach space*, Trans. Amer. Math. Soc. 185 (1973), pp. 253–264.

[12] R. LePage, *The* Log Log *law for Gaussian processes* (to appear in Z. Wahrscheinlichkeitstheorie).

[13] H. F. Trotter, *An elementary proof of the central limit theorem*, Arch. Math. 9 (1959), pp. 226–234.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN, MADISON

Received October 10, 1973 (744)