EXISTENCE AND CONVERGENCE OF PROBABILITY MEASURES IN BANACH SPACES[1]

BY

ALEJANDRO D. DE ACOSTA

Abstract. Theorems of the Bochner-Sazonov type are proved for Banach spaces with a basis. These theorems give sufficient conditions of a topological nature under which a positive definite function is the characteristic functional of a probability measure. The conditions are, in a certain natural sense, best possible. Central limit theorems of the Lindeberg type for triangular systems of random variables taking values in a Banach space with a basis are obtained. Applications to l_p and $C[0, 1]$ are given.

0. Introduction. In this work we will be concerned with the following problems which we describe in general terms.

(1) *Generalization of Bochner's theorem.* Let X be a topological vector space X' its dual space. Let $g: X' \to C$ be positive definite, $g(0) = 1$. Under what conditions, or more specifically, under what topological conditions is g the characteristic functional of a probability measure on X?

(2) *Generalization of the central limit theorem.* Let $\{Z_j\}$, $j = 1, 2, \ldots$, be an independent sequence of X-valued random variables with a common distribution. Suppose for each $y \in X'$, $E\langle Z, y \rangle = 0$ and $E\langle Z, y \rangle^2 < \infty$. Under what conditions (on the common distribution) does $\mathscr{L}(n^{-1/2} \sum_{j=1}^n Z_j)$ converge weakly? More generally, when are limit theorems of the Lindeberg type valid?

For the first problem, the classical theorem of Bochner provides a complete answer when the space X is finite dimensional: g is a characteristic function if and only if g is continuous at 0. (See, for example, [11, p. 207]). A complete answer is also possible when the space is a locally compact abelian group. (See, for instance, [17].)

In the infinite-dimensional case, Kolmogorov showed that g is the Fourier-Stieltjes transform of a *finitely* additive measure defined on the field of cylinder sets induced by X' and σ-additive on each σ-field induced by a finite subset of X' if and only if the restriction of g to each finite-dimensional subspace of X' is

continuous. (See [16], [4].) Bochner showed that under the same conditions, g is the characteristic functional of a σ-additive measure on the σ-field induced by X' in its algebraic dual $(X')^*$, and conversely. (See, for instance, [4].)

The question is then: When is it possible to extend the finitely additive measure corresponding to g, called *weak distribution* (or *cylinder measure*), to a σ-additive (or even tight) measure on a σ-field of subsets of X? Or, when does the corresponding σ-additive measure on $(X')^*$ concentrate its mass in X?

It is easy to show that, for infinite-dimensional Banach spaces, continuity in the norm topology of X' is not enough to make g the characteristic functional of a σ-additive measure. For a discussion of the difficulties involved, as well as various results, we refer to LeCam [10], Prohorov [16], and the recent work of Dudley [4]. For the case of a (separable) Hilbert space, Sazonov [18] gave a solution to the problem: g is the characteristic functional of a σ-additive measure if and only if g is continuous in the topology generated by the trace class operators. (An exposition may be found in [14].) The same result was later obtained by Gross with a different approach [8].

For the case of nuclear spaces, a complete solution was obtained by Minlos [12] (see also [7]). A contribution in the case of l_p spaces was made by Vakhania ([22], [23])[2].

A solution to the second problem is known when X is a (separable) Hilbert space ([13], [15]). More generally, a theory of limit laws of infinitesimal systems of random variables closely paralleling the classical one has been developed in the case of Hilbert space ([25]; see also the exposition in [14]). Fortet and Mourier [6] have studied the problem for a special class of separable reflexive Banach spaces, the G-spaces. Results for the case of l_p spaces were obtained by Vakhania ([23]; a theorem for l_1 was also announced in [24]). In a recent work, Dudley and Strassen [5] give a sufficient condition in terms of ε-entropy for the validity of the central limit theorem when $X = C(S)$, S a compact metric space. These authors also present an example showing that the classical statement of the central limit theorem is false in $C[0, 1]$.

In the present work (§§4 and 6) sufficient conditions are given under which the two problems have solutions, for the class of Banach spaces with a Schauder basis. (It should be remarked here that no example of a separable Banach space without a Schauder basis is known. See [2, p. 67], or [19, p. 114].) It is also shown that the conditions are, in a certain natural sense, the best possible in the situation we study. Generalizations of Prohorov's compactness criterion (for a family of probability measures in Hilbert space) are obtained for spaces with a Schauder basis (§4).

---

[2] After the present work had been completed, the author learned about the existence (in manuscript form) of the paper "Harmonic analysis on certain vector spaces," by J. Kuelbs and V. Mandrekar. This work contains a detailed study of measures on l^p spaces, including versions of the Bochner theorem and of the Lévy continuity theorem.

Some applications to special spaces are presented, in particular to the study of the continuity of Gaussian processes (§§7 and 8). In the first section, we study the situation that arises when g is assumed w^*-continuous.

In the second section, certain convenient formulations of tightness and convergence criteria are given for Banach spaces.

Notation and terminology. For the standard notions concerning the weak topology of the space of probability measures on a topological space, we refer to [1], [14]. For us, *probability measure* will always mean a *σ-additive measure*, defined on the σ-field specified in each case. Given a topological space X, $\mathscr{F}$ will denote the Borel σ-field of X, that is, the σ-field generated by the open sets of X.

The basic facts about the characteristic functional of a measure,

$$\hat{\mu}(y) = \int \exp{(i\langle x, y \rangle)}\mu(dx), \quad (y \in X'),$$

defined on a topological vector space, may be found in [4], [10], [16].

For the linear space notions we use, we refer to [2], [19].

1. **Positive definite functions continuous in the w^*-topology.** Let X be a Hausdorff locally convex topological vector space, X' its dual space. Let $g: X' \to C$ be positive definite, $g(0) = 1$. In this section we show that continuity of g in the $\sigma(X', X)$ topology, while providing a sufficient condition for the validity of Bochner's theorem, greatly restricts the class of measures obtained. It should be remarked that *sequential* $\sigma(X', X)$-continuity is a necessary condition for g to be the characteristic functional of a probability measure, as can be proved easily using the Lebesgue dominated convergence theorem. The situation is, however, quite different for topological continuity.

We start by proving a proposition which describes the support of a probability measure in terms of its characteristic functional.

We recall that the support of a measure is defined as follows:

DEFINITION 1.1. *Let X be a Hausdorff topological space, $\mathscr{F}$ the Borel σ-field of X. Let μ be a nonnegative measure on $(X, \mathscr{F})$. The support of μ is the set*

$$C(\mu) = \{x \in X : \mu(U) > 0 \text{ for every open } U \text{ containing } x\}.$$

The following properties of $C(\mu)$ are easily proved:

(a) $C(\mu)$ is closed.

(b) If D is closed and $\mu(D^c) = 0$, then $C(\mu) \subset D$.

(c) If μ is a regular Borel measure, then $\mu[C(\mu)^c] = 0$.

(d) If $f: X \to R^1$ is continuous and nonnegative, then $\int f \, d\mu = 0$ implies $f(x) = 0$ for $x \in C(\mu)$.

PROPOSITION 1.1. *Let X be a Hausdorff topological vector space, $\mathscr{F}$ the Borel σ-field of X. Let μ be a regular Borel probability measure on $(X, \mathscr{F})$. Then*

$$[C(\mu)]^\perp = \{y \in X' : \hat{\mu}(ty) = 1 \text{ for all } t \in R^1\}.$$

Proof. Suppose $y \in X'$ and $\hat{\mu}(ty) = 1$ for all $t \in R^1$. Then $1 = \int e^{i\langle x, ty \rangle} \mu(dx)$, so $\int (1 - \cos \langle x, ty \rangle) \mu(dx) = 0$ for all $t \in R^1$. Using (d), one obtains: for each $t \in R^1$, $1 - \cos \langle x, ty \rangle = 0$ for all $x \in C(\mu)$. This implies: for all $x \in C(\mu)$, $\langle x, y \rangle = 0$. Thus $y \in [C(\mu)]^\perp$.

Conversely, let $y \in [C(\mu)]^\perp$. Then: for all $t \in R^1$, $\hat{\mu}(ty) = \int_{C(\mu)} e^{i\langle x, ty \rangle} \mu(dx) = \mu(C(\mu)) = 1$. $\square$

THEOREM 1.2. *Let X be a Hausdorff locally convex topological vector space, X' its dual space, $\mathscr{F}$ be the Borel σ-field on X. Let $g \colon X' \to C$ be positive definite, $g(0) = 1$.*

(a) g is $\sigma(X', X)$ continuous if and only if g is the characteristic functional of a probability measure μ on $(X, \mathscr{F})$ which satisfies the condition: for every $\varepsilon > 0$, there is a finite-dimensional subspace S of X, such that $\mu(S) \geqq 1 - \varepsilon$.

(b) If g is as in (a) and, further, for each $y \in X'$, $(\mu \circ y^{-1})(\{0\}) = 0$ or 1, then there is a finite-dimensional subspace S of X such that $\mu(S) = 1$.

Proof. (a) Suppose μ is a probability measure on $(X, \mathscr{F})$ satisfying the stated condition. Then, given $\varepsilon > 0$, one can find a finite-dimensional subspace S such that $\mu(S) \geqq 1 - (\varepsilon/6)$; then

$$|\hat{\mu}(y) - 1| \leqq \int_S |1 - e^{i\langle x, y \rangle}| \mu(dx) + \varepsilon/3, \qquad y \in X'.$$

Let $\{e_1, \ldots, e_k\}$ be a basis for S, $\{f_1, \ldots, f_k\}$ the dual basis in S', so that $x = \sum_{j=1}^k \langle x, f_j \rangle e_j$ when $x \in S$. Then if $A_r = \{x \in S : |\langle x, f_j \rangle| \leqq r, j = 1, \ldots, k\}$, one can choose r so that $\mu(S - A_r) \leqq \varepsilon/6$. Then

$$\int_S |1 - e^{i\langle x, y \rangle}| \mu(dx) \leqq \int_{A_r} |1 - e^{i\langle x, y \rangle}| \mu(dx) + \varepsilon/3$$

$$\leqq \int_{A_r} |\langle x, y \rangle| \mu(dx) + \varepsilon/3$$

$$\leqq \sum_{j=1}^k |\langle e_j, y \rangle| \int_{A_r} |\langle x, f_j \rangle| \mu(dx) + \varepsilon/3$$

$$\leqq r \sum_{j=1}^k |\langle e_j, y \rangle| + \varepsilon/3.$$

If we put now $V = \{y \in X' : |\langle e_j, y \rangle| \leqq \varepsilon/(3kr), j = 1, \ldots, k\}$, then $y \in V$ implies $|\hat{\mu}(y) - 1| \leqq \varepsilon$. This proves the $\sigma(X', X)$-continuity of $\hat{\mu}$.

Now let $g \colon X' \to C$ be positive definite, $g(0) = 1$, g continuous in the $\sigma(X', X)$ topology. Then there exists a weak distribution μ on the cylinder sets of X such that $\hat{\mu} = g$ (see [16, p. 410]).

We wish to show that μ extends to a tight probability measure on $(X, \mathscr{F})$. According to a theorem of Prohorov ([16, p. 411]; see also [9, p. 147] for a more general theorem on probability measures constructed from projective systems) it is enough to show: for every $\varepsilon > 0$, there is a compact K such that for any cylinder set

E which contains K, $\mu(E) \geq 1 - \varepsilon$. (The assumptions on the space make possible the use of this theorem.) Given $\varepsilon > 0$, there are $\delta > 0$, $x_1, \ldots, x_k \in X$ such that if $V = \{y \in X' : |\langle x_j, y \rangle| \leq \delta, j = 1, \ldots, k\}$, then $y \in V$ implies $|\hat{\mu}(y) - 1| \leq \varepsilon/2$. Let S be the subspace generated by $\{x_1, \ldots, x_k\}$. Then for any $y_1, \ldots, y_m \in S^\perp$ and any $t_1, \ldots, t_m \in R^1$, one has $|\hat{\mu}(\sum_{j=1}^m t_j y_j) - 1| \leq \varepsilon/2$. By integrating on the cube $[-n, n]^m$ of R^m with respect to normalized Lebesgue measure, changing the order of integration and finally letting $n \to \infty$, one obtains $\mu(\{x : \langle x, y_j \rangle = 0, j = 1, \ldots, m\}) \geq 1 - (\varepsilon/2)$. This says: for any closed subspace T of finite codimension which contains S, $\mu(T) \geq 1 - (\varepsilon/2)$.

Let $z_1, \ldots, z_k \in X'$ be such that $S^\perp \oplus H = X'$, where H is the subspace generated by $\{z_1, \ldots, z_k\}$. Then for any $r \geq 0$, the set $S \cap A_r$ is compact, where

$$A_r = \{x \in X : |\langle x, z_j \rangle| \leq r, j = 1, \ldots, k\}.$$

Choose r so that $\mu(A_r) \geq 1 - (\varepsilon/2)$.

Let E be any cylinder set such that $E \supset S \cap A_r$. We want to show $\mu(E) \geq 1 - \varepsilon$. Suppose, to the contrary, that $\mu(E^c) > \varepsilon$. Since $E^c \cap A_r$ is a cylinder set, there exist a continuous linear map Π of X into some R^p and a Borel set F in R^p such that $E^c \cap A_r = \Pi^{-1}(F)$. Now $E^c \cap A_r \cap \Pi^{-1}(\Pi(S)) \neq \varnothing$, for: $\mu(\Pi^{-1}(\Pi(S))) \geq 1 - (\varepsilon/2)$ because $\Pi^{-1}(\Pi(S))$ is a closed subspace of finite codimension which contains S, and hence

$$\mu(E^c \cap A_r \cap \Pi^{-1}(\Pi(S))) \geq 1 - (\mu(E) + \mu(A_r^c) + \mu(\Pi^{-1}(\Pi(S))^c))$$
$$> 1 - ((1 - \varepsilon) + \varepsilon/2 + \varepsilon/2) = 0.$$

But then $F \cap \Pi(S) \neq \varnothing$, so for some $x \in S$, $x \in \Pi^{-1}(F) = E^c \cap A_r$, and therefore $E^c \cap A_r \cap S \neq \varnothing$. This contradicts the assumption about E. Therefore $\mu(E) \geq 1 - \varepsilon$, and the compact $S \cap A_r$ satisfies the desired condition. The additional property of the extended measure is clear.

(b) Let $\delta > 0$ and $x_1, \ldots, x_k \in X$ be such that if

$$V = \{y \in X' : |\langle x_j, y \rangle| \leq \delta, j = 1, \ldots, k\},$$

then $y \in V$ implies $|1 - \hat{\mu}(y)| \leq 1/2$. If S is the subspace generated by $\{x_1, \ldots, x_k\}$, then for all $y \in S^\perp$, for all $t \in R^1$, $|1 - \hat{\mu}(ty)| \leq 1/2$, or $|1 - (\mu \circ y^{-1})\widehat{}(t)| \leq 1/2$. Reasoning as in the proof of (a), one can show that this implies $(\mu \circ y^{-1})(\{0\}) \geq 1/2$. Our assumption forces $\mu \circ y^{-1}$ to be point mass at 0. Therefore $\hat{\mu}(ty) = (\mu \circ y^{-1})\widehat{}(t)$ $= 1$ for all $t \in R^1$. This shows $S^\perp \subset \{y \in X' : \hat{\mu}(ty) = 1 \text{ for all } t \in R^1\}$. But, by Proposition 1.1, $[C(\mu)]^\perp = \{y : \hat{\mu}(ty) = 1 \text{ for all } t \in R^1\}$; therefore $C(\mu) \subset [C(\mu)]^{\perp\perp} \subset S^{\perp\perp} = S$, and $\mu(S) \geq \mu(C(\mu)) = 1$. $\quad\square$

REMARK. Theorem 1.2(a) is essentially equivalent to Theorem (3.1) of [4]. Our proof, however, is different from Dudley's. Part (b) of the theorem is a generalization of a remark of LeCam [10, p. 230] concerning the case when μ is Gaussian centered.

2. **Tightness and convergence.**   In this section we prove certain tightness and convergence criteria for measures in a Banach space. These propositions will be quite useful in our further work. Their convenience lies in the fact that the hypotheses are formulated "linearly"; thus, the application of these criteria to a particular Banach space does not require a description of the compacta of the space.

Let X be a normed linear space, $A \subset X$. Then we call

$$A^\varepsilon = \{x \in X : \inf\{\|x - z\| : z \in A\} \leq \varepsilon\}.$$

THEOREM 2.1. *Let X be a Banach space, $\mathscr{F}$ its Borel σ-field, $\{\mu_\alpha\}_{\alpha \in I}$ a family of probability measures on $(X, \mathscr{F})$. Then $\{\mu_\alpha\}_{\alpha \in I}$ is tight if and only if for every $\varepsilon > 0$, $\delta > 0$, there exist (a) a finite-dimensional subspace S of X such that $\mu_\alpha(S^\varepsilon) \geq 1 - \delta$ for all $\alpha \in I$, and (b) a finite subset $\{y_1, \ldots, y_k\} \subset X'$ and a number $r > 0$ such that*

$$\mu_\alpha(\{x : |\langle x, y_j \rangle| \leq r, j = 1, \ldots, k\}) \geq 1 - \delta,$$

and $S^\perp \oplus T = X'$, where T is the subspace generated by $\{y_1, \ldots, y_k\}$. Therefore, if the stated condition holds, then the family $\{\mu_\alpha\}_{\alpha \in I}$ is relatively compact. If μ_α is individually tight for all $\alpha \in I$ then the converse is also true.

Proof.   We prove sufficiency. Choose a finite-dimensional subspace S of X such that $\mu_\alpha(S^\varepsilon) \geq 1 - \varepsilon$ for all $\alpha \in I$, and a finite subset $\{y_1, \ldots, y_k\} \subset X'$ as described in condition (b). Consider now the set $A = S^\varepsilon \cap \{x : |\langle x, y_j \rangle| \leq r, j = 1, \ldots, k\}$. We will show that there is a compact K such that $A \subset K^\varepsilon$.

Let $\lambda_1, \ldots, \lambda_k$ be elements of X'' such that $y = z + \sum_{j=1}^k \langle y, \lambda_j \rangle y_j$ (with $z \in S^\perp$) for all $y \in X'$. Then $|\langle x, y \rangle| \leq |\langle x, z \rangle| + \sum_{j=1}^k |\langle y, \lambda_j \rangle| \, |\langle x, y_j \rangle|$.

Let $x \in A$. Then $|\langle x, z \rangle| \leq \varepsilon\|z\|$ and $\sum_{j=1}^k |\langle y, \lambda_j \rangle| \, |\langle x, y_j \rangle| \leq r\|y\|(\sum_{j=1}^k \|\lambda_j\|)$. Put

$$M_1 = \sum_{j=1}^k \|\lambda_j\|, \qquad M_2 = \sup\{\|y_j\| : j = 1, \ldots, k\}.$$

Then

$$\|z\| \leq \|y\| + \left\|\sum_{j=1}^k \langle y, \lambda_j \rangle y_j\right\| \leq \|y\|(1 + M_1 M_2).$$

Therefore

$$\|x\| = \sup\{|\langle x, y \rangle| : \|y\| \leq 1\} \leq \varepsilon[1 + M_1 M_2] + r M_1 = \rho.$$

This shows that $A \subset B_\rho$, where $B_\rho = \{x \in X : \|x\| \leq \rho\}$, for any $\rho \geq 0$.

It is easy to show that $B_\rho \cap S^\varepsilon \subset (S \cap B_{\rho + \varepsilon})^\varepsilon$. Since $A \subset B_\rho \cap S^\varepsilon$, we have $A \subset (S \cap B_{\rho + \varepsilon})^\varepsilon$. But $S \cap B_{\rho + \varepsilon}$ is compact.

We have then $\mu_\alpha(S \cap B_{\rho + \varepsilon})^\varepsilon \geq \mu_\alpha(A) \geq 1 - 2\varepsilon$ for all $\alpha \in I$. Therefore we have shown: for every $\varepsilon > 0$, there exists a compact K such that for all $\alpha \in I$, $\mu_\alpha(K^\varepsilon) \geq 1 - \varepsilon$. Now choose, for each $j \geq 1$, a compact K_j such that $\mu_\alpha(K_j^{\varepsilon/2^j}) \geq 1 - \varepsilon/2^j$ for all $\alpha \in I$. Then if $K = \bigcap_{j=1}^\infty K_j^{\varepsilon/2^j}$, K is compact and $\mu_\alpha(K) \geq 1 - \varepsilon$ for all $\alpha \in I$. This proves the tightness of $\{\mu_\alpha\}_{\alpha \in I}$. The necessity is clear. The rest of the statement

follows from Prohorov's theorem (see, for instance, [1], p. 35 and Appendix on theoretical complements). $\square$

Motivated by Theorem 2.1, we introduce

DEFINITION 2.1. *A family $\{\mu_\alpha\}_{\alpha \in I}$ of probability measures on $(X, \mathscr{F})$ is flatly concentrated if for every $\varepsilon > 0$, for every $\delta > 0$, there exists a finite-dimensional subspace S of X, such that $\mu_\alpha(S^\varepsilon) \geq 1 - \delta$ for all $\alpha \in I$.* It is clear that any tight family of probability measures is flatly concentrated.

For our next theorem, we need a geometric lemma:

LEMMA 2.2. *Let X be a topological vector space. Let T be a closed subspace of X, of finite codimension $k > 0$. Suppose A is a total subset of X. Then there exist $y_1, \ldots, y_k \in A$ such that:*

(1) *$\{y_1, \ldots, y_k\}$ is an independent set.*

(2) *If H is the subspace generated by $\{y_1, \ldots, y_k\}$, then $T \oplus H = X$.*

Proof. We show first that there exists $y_1 \in A$, $y_1 \notin T$. For suppose $A \subset T$. Taking closed linear hulls on both sides of the inclusion, we would obtain $X = T$, a contradiction.

Let $y_1 \in A \cap T^c$. Let T_1 be the subspace generated by T and y_1. It is not difficult to see that T_1 is closed and has codimension $(k - 1)$. Thus we can repeat the same argument as above, obtaining $y_2 \in A$, $y_2 \notin T_1$. Continuing in the same way we obtain $y_3, \ldots, y_k$. It is clear from the construction that (1) and (2) are satisfied. $\square$

THEOREM 2.3. *Let X be a Banach space, $\mathscr{F}$ its Borel σ-field, $\{\mu_\alpha\}_{\alpha \in I}$ a family of probability measures on $(X, \mathscr{F})$. Then $\{\mu_\alpha\}_{\alpha \in I}$ is tight if and only if:*

(a) *$\{\mu_\alpha\}_{\alpha \in I}$ is flatly concentrated.*

(b) *There exists a subset A of X' which is w^*-total such that $\{\mu_\alpha \circ y^{-1}\}_{\alpha \in I}$ is tight in R^1 for all $y \in A$. Therefore, if the stated conditions hold, then the family $\{\mu_\alpha\}_{\alpha \in I}$ is relatively compact. If μ_α is tight for each $\alpha \in I$, then the converse is also true.*

Proof. We only prove that (a) and (b) imply tightness. The rest of the statement is clear.

Choose a finite-dimensional subspace S of X such that $\mu_\alpha(S^\varepsilon) \geq 1 - \delta$ for all $\alpha \in I$. Then $S^\perp$ is w^*-closed and has finite codimension. By Lemma 2.2, one can choose $y_1, \ldots, y_k \in A$ such that $S^\perp \oplus H = X'$, where H is the subspace generated by $\{y_1, \ldots, y_k\}$. Since $\{\mu_\alpha \circ y_j^{-1}\}_{\alpha \in I}$ is tight, one can choose $r > 0$ such that

$$\mu_\alpha \circ y_j^{-1}([-r, r]) \geq 1 - \delta/k,$$

for $j = 1, \ldots, k$ and all $\alpha \in I$. Then

$$\mu_\alpha(\{x : |\langle x, y_j \rangle| \leq r, j = 1, \ldots, k\})^c \leq \sum_{j=1}^{n} \mu_\alpha(\{x : |\langle x, y_j \rangle| \leq r\})^c \leq \delta$$

for all $\alpha \in I$. The statement follows now from Theorem 2.1. $\square$

We obtain now a convergence theorem.

THEOREM 2.4. *Let X be a Banach space, $\{\mu_n\}_{n\in N}$ a sequence of probability measures on $(X, \mathscr{F})$. If*

    (a) *$\{\mu_n\}_{n\in N}$ is flatly concentrated, and*

    (b) *there exists a sequentially w^*-dense subspace W of X' and a function g: $W \to C$, whose restriction to each one-dimensional subspace of W is continuous, such that $\hat{\mu}_n$ converges (pointwise) to g on W,*

then there exists a probability measure μ on $(X, \mathscr{F})$ such that μ_n converges weakly to μ and $\hat{\mu}(y) = g(y)$ for $y \in W$.

Proof. We show first that $\{\mu_n\}_{n\in N}$ is relatively compact. Fix $y \in W$ and define $h(t) = g(ty)$ for $t \in R^1$. Then h is continuous at 0 and $(\mu_n \circ y^{-1})\hat{}(t) = \hat{\mu}_n(ty) \to g(ty) = h(t)$. By the one-dimensional P. Lévy continuity theorem, the sequence $\{\mu_n \circ y^{-1}\}_{n\in N}$ is tight. Theorem 2.3 implies now the desired conclusion. Clearly any two subsequential limits of $\{\mu_n\}_{n\in N}$, say μ and ν, satisfy $\hat{\mu}(y) = g(y) = \hat{\nu}(y)$ for $y \in W$. Let $y \in X'$, $y_n \in W$, $y_n \to y$ in the w^*-topology. Since characteristic functionals are always sequentially w^*-continuous, we have $\hat{\mu}(y) = \lim_{n\to\infty} \hat{\mu}(y_n) = \lim_{n\to\infty} \hat{\nu}(y_n) = \hat{\nu}(y)$. But μ and ν are tight measures; therefore the equality $\hat{\mu} = \hat{\nu}$ implies that $\mu = \nu$ [16, p. 414]. Hence there exists a (unique) probability measure μ on $(X, \mathscr{F})$ such that μ_n converges weakly to μ. $\square$

REMARK. It is well known that the classical form of P. Lévy's continuity theorem fails drastically to be true in the infinite-dimensional case. In a future work, we will study conditions, in a sense weaker than those in Theorem 2.4, under which P. Lévy's theorem holds. These conditions, first conjectured by L. LeCam, allow the finite-dimensional subspaces near which mass is concentrated to depend on each measure in the sequence.

3. A "flat concentration" inequality.

In this section and the following ones, we work with Banach spaces possessing a Schauder basis. First we establish certain terminology and recall basic facts about such spaces [2].

Let X be a Banach space with a Schauder basis $(e_i)_{i\in N}$. We assume $\|e_i\| = 1$, $i \in N$. We call S_n the subspace generated by $\{e_i, \ldots, e_n\}$. Let $(f_j)_{j\in N}$ be the coordinate functions for this basis. Then $f_j \in X'$ $(j \in N)$ and $\langle e_i, f_j \rangle = \delta_{ij}$. We call T_n the subspace of X' generated by $\{f_1, \ldots, f_n\}$, and let $T = \bigcup_{n=1}^{\infty} T_n$.

Let $\Pi_n: X \to X$ be defined by $\Pi_n(x) = \sum_{j=1}^{n} \langle x, f_j \rangle e_j$. Then Π_n is a projection with range S_n, and $\|\Pi_n x - x\| \to 0$ as $n \to \infty$ for all $x \in X$. If $\Pi_n^*: X' \to X'$ is the operator adjoint to Π_n, then one can write $\Pi_n^*(y) = \sum_{j=1}^{n} \langle e_j, y \rangle f_j$, and $\Pi_n^*(y) \to y$ in the w^*-topology. This implies that T is a sequentially w^*-dense subspace of X'. One also has $T_n \oplus S_n^{\perp} = X'$. In fact, $y - \Pi_n^*(y) \in S_n^{\perp}$.

For each closed subspace S of X, we define $p_S(x) = \inf\{\|x - z\| : z \in S\}$ for $x \in X$. Then one can show that p_S is a continuous seminorm. If $S = S_n$, we put $p_{S_n} = p_n$. Then $p_n(\Pi_m(x)) \to p_n(x)$ as $m \to \infty$.

We introduce now the concept of "space of type p."

DEFINITION 3.1. *Let X be a Banach space with a Schauder basis $(e_i)_{i \in N}$ and coordinate functions $(f_j)_{j \in N}$. Let $p > 0$. We say that X is of type p, with respect to $(e_i)_{i \in N}$ if $[p_n(\Pi_m(x))]^p \leq \sum_{j=n+1}^m |\langle x, f_j \rangle|^p$ for all $x \in X$ and $n, m \in N$ $(n < m)$.*

In the future, when dealing with a space of type p, we will omit reference to the basis.

As an example, we mention the fact that l_p is of type p $(1 \leq p < \infty)$, as one can easily verify.

LEMMA 3.1. (a) *All Banach spaces with a Schauder basis are of type 1.*
(b) *If X is of type p and $q < p$, then X is of type q.*

Proof. (a) Since p_n is a seminorm, we have

$$p_n(\Pi_m(x)) = p_n\left(\Pi_n(x) + \sum_{j=n+1}^m \langle x, f_j \rangle e_j\right)$$

$$\leq p_n(\Pi_n(x)) + \sum_{j=n+1}^m |\langle x, f_j \rangle| p_n(e_j).$$

But $p_n(\Pi_n(x)) = 0$ and $p_n(e_j) \leq 1$, $j = n+1, \ldots, m$. Therefore

$$p_n(\Pi_m(x)) \leq \sum_{j=n+1}^m |\langle x, f_j \rangle|.$$

(b) Using the inequality $\sum_{i=1}^k a_i^r \leq (\sum_{i=1}^k a_i)^r$, valid for $a_i \geq 0$, $r \geq 1$, we have

$$[p_n(\Pi_m(x))]^p \leq \sum_{j=n+1}^m |\langle x, f_j \rangle|^p \leq \left[\sum_{j=n+1}^m |\langle x, f_j \rangle|^q\right]^{p/q},$$

and then $[p_n(\Pi_m(x))]^q \leq \sum_{j=n+1}^m |\langle x, f_j \rangle|^q$. $\square$

We prove now an inequality which gives an estimate for the mass a probability measure puts in an ε-neighborhood of S_n in terms of the characteristic functional of the measure. It is stated so as to allow some flexibility in applications. We denote by γ the standard normal distribution in R^1.

THEOREM 3.2. *Let X be a Banach space with Schauder basis $(e_i)_{i \in N}$ and coordinate functions $(f_j)_{j \in N}$, of type p $(1 \leq p < \infty)$. Let μ be a probability measure on $(X, \mathcal{F})$.*
(a) *If $1 \leq p < 2$ and $(b_j)_{j \in N}$ is a sequence of positive real numbers, then*

$$\mu[(S_n^\varepsilon)^c] \leq K_\varepsilon \liminf_{m \to \infty} \int_{R^{m-n}} \left|1 - \operatorname{Re} \hat{\mu}\left(s(n, m) \sum_{j=n+1}^m t_j b_j^{1/p} f_j\right)\right| \gamma(dt_{n+1}) \cdots \gamma(dt_m)$$

where $K_\varepsilon = [1 - \exp(-\varepsilon^2/2)]^{-1}$ and $s(n, m) = [\sum_{j=n+1}^m b_j^{-2/(2-p)}]^{(2-p)/2p}$.
(b) *If $2 \leq p < \infty$, then*

$$\mu[(S_n^\varepsilon)^c] \leq K_\varepsilon \liminf_{m \to \infty} \int_{R^{m-n}} \left|1 - \operatorname{Re} \hat{\mu}\left(\sum_{j=n+1}^m t_j f_j\right)\right| \gamma(dt_{n+1}) \cdots \gamma(dt_m).$$

Proof. (a) First observe that $(S_n^\varepsilon)^c = \{x : p_n(x) > \varepsilon\}$. By Markov's inequality, we have

$$\mu(\{x : p_n(x) > \varepsilon\}) = \mu(\{x : \tfrac{1}{2}[p_n(x)]^2 > \varepsilon^2/2\}) \leq K_\varepsilon \int [1 - \exp(-\tfrac{1}{2}[p_n(x)]^2)]\mu(dx).$$

Then, by Lebesgue's dominated convergence theorem,

(1) $\mu(\{x : p_n(x) > \varepsilon\}) \leq K_\varepsilon \lim\limits_{m \to \infty} \int [1 - \exp(-\tfrac{1}{2}[p_n(\Pi_m(x))]^2)]\mu(dx).$

By assumption, $[p_n(\Pi_m(x))]^p \leq \sum_{j=n+1}^m |\langle x, f_j\rangle|^p$. By Hölder we have

$$\sum_{j=n+1}^m |\langle x, f_j\rangle|^p = \sum_{j=n+1}^m b_j^{-1} b_j |\langle x, f_j\rangle|^p$$

$$\leq \left[\sum_{j=n+1}^m b_j^{-q}\right]^{1/q}\left[\sum_{j=n+1}^m b_j^{2/p}|\langle x, f_j\rangle|^2\right]^{p/2},$$

where $q = 2/2 - p$ is the exponent conjugate to $2/p$. Then

(2)
$$[p_n(\Pi_m(x))]^2 \leq \left[\sum_{j=n+1}^m b_j^{-q}\right]^{2/pq}\left[\sum_{j=n+1}^m b_j^{2/p}|\langle x, f_j\rangle|^2\right]$$

$$= \sum_{j=n+1}^m |\langle x, s(n,m)b_j^{1/p}f_j\rangle|^2.$$

Since $\hat{\gamma}(t) = \exp(-t^2/2)$, one can write

$$1 - \exp\left(-\frac{1}{2}\sum_{j=n+1}^m |\langle x, s(n,m)b_j^{1/p}f_j\rangle|^2\right)$$

$$= \int_{R^{m-n}}\left(1 - \cos\sum_{j=n+1}^m t_j\langle x, s(n,m)b_j^{1/p}f_j\rangle\right)\gamma(dt_{n+1})\cdots\gamma(dt_m)$$

$$= \int_{R^{m-n}}\left(1 - \cos\left\langle x, \sum_{j=n+1}^m s(n,m)t_j b_j^{1/p}f_j\right\rangle\right)\gamma(dt_{n+1})\cdots\gamma(dt_m).$$

Integrating with respect to μ, using Fubini's theorem and inequality (2), we obtain

$$\int [1 - \exp(-\tfrac{1}{2}[p_n(\Pi_m(x))]^2)]\mu(dx)$$

$$\leq \int_{R^{m-n}}\left|1 - \operatorname{Re}\hat{\mu}\left(s(n,m)\sum_{j=n+1}^m t_j b_j^{1/p}f_j\right)\right|\gamma(dt_{n+1})\cdots\gamma(dt_m).$$

Taking lim inf on both sides and using (1), we obtain (a).

(b) If X is of type p, with $p \geq 2$, then it is of type 2, by Lemma 3.1. Thus

$$[p_n(\Pi_m(x))]^2 \leq \sum_{j=n+1}^m |\langle x, f_j\rangle|^2.$$

The proof proceeds in a way similar to the proof of (a); the only difference is that the sequence $(b_j)_{j \in N}$ is not introduced. $\square$

REMARK. Theorem 3.2 may be regarded as a generalization of an inequality used implicitly by Sazonov [18] (see also [14, Chapter 6]). Our formulation is,

however, somewhat different. A generalization of Sazonov's method along different lines has been obtained by L. Schwartz [20]. It should be observed that the use of the standard normal distribution in Theorem 3.2 is just a matter of convenience. For example, if $p \in [1, 2]$ and $p < r \leqq 2$, then a corresponding inequality could be obtained using the standard symmetric stable distribution with index r. However, the fact that a stable distribution lacks a second moment if $r < 2$ would create difficulties in future application.

A distribution which can provide a successful substitute for the normal is the bilateral exponential.

4. A theorem of the Bochner type. In this section we prove a compactness theorem for a family of probability measures and establish a sufficient condition of topological character for the validity of Bochner's theorem, for Banach space with a Schauder basis. We also show that the condition is the best possible of its kind. We use the notation of §3.

Let us recall that a real-valued function Φ defined on a vector space E is called a *quadratic form* if there exists a symmetric bilinear form ψ on $E \times E$ such that $\Phi(x) = \psi(x, x)$ for all $x \in E$. Φ is called *nonnegative definite* if $\Phi(x) \geqq 0$, for all $x \in E$.

We introduce now the

DEFINITION 4.1. *Let E be a vector space, $\{z_j\}_{j \in N}$ a sequence of elements of E. Let $p > 0$. A nonnegative definite quadratic form Φ defined on E is said to be of trace class p, relative to $\{z_j\}_{j \in N}$, if $\sum_{j=1}^{\infty} [\Phi(z_j)]^p < \infty$.*

We will be concerned with quadratic forms defined on spaces $E = X'$, where X is a Banach space with Schauder basis $(e_i)_{i \in N}$ and coordinate functions $(f_j)_{j \in N}$. *The property of being of trace class p will always be taken relative to $(f_j)_{j \in N}$.* In general, the property being satisfied or not depends on the particular basis chosen.

REMARK. This terminology is consistent with standard Hilbert space usage in the following way: Let X be a separable Hilbert space with inner product $(\cdot, \cdot)$, $(f_j)_{j \in N}$ an orthonormal basis. If Φ is continuous and of trace class 1 and A is an operator on X such that $\Phi(x) = (Ax, x)$, then A is of trace class in the usual sense. In fact, in this case the condition on A (or Φ) does not depend on the particular orthonormal basis (see, for instance, [7, Chapter 1]).

As in §3, let T be the linear subspace of X', spanned by $(f_j)_{j \in N}$.

THEOREM 4.1. *Let X be a Banach space with a Schauder basis of type p ($1 \leqq p < \infty$). Let $\{\mu_\alpha\}_{\alpha \in I}$ be a set of probability measures on $(X, \mathcal{F})$.*

(a) *Let $p \in [1, 2)$. Suppose for every $\varepsilon > 0$, there exists a family $\{\Phi_{\alpha, \varepsilon}\}_{\alpha \in I}$ of quadratic nonnegative definite forms of trace class $p/2$, defined on T and satisfying $\sup_{\alpha \in I} \sum_{j=1}^{\infty} [\Phi_{\alpha, \varepsilon}(f_j)]^{p/2} < \infty$, and $\sup_{\alpha \in I} \sum_{j=r}^{\infty} [\Phi_{\alpha, \varepsilon}(f_j)]^{p/2} \to 0$ as $r \to \infty$, such that $|1 - \operatorname{Re} \hat{\mu}_\alpha(y)| \leqq \Phi_{\alpha, \varepsilon}(y) + \varepsilon$ for all $y \in T$. Then $\{\mu_\alpha\}_{\alpha \in I}$ is relatively compact.*

(b) *Let $p \in [2, \infty)$. Suppose for every $\varepsilon > 0$, there exists a family $\{\Phi_{\alpha, \varepsilon}\}_{\alpha \in I}$ of quadratic nonnegative definite forms of trace class 1 defined on T and satisfying*

$\sup_{\alpha\in I} \sum_{j=1}^{\infty} [\Phi_{\alpha,\varepsilon}(f_j)] < \infty$, and $\sup_{\alpha\in I} \sum_{j=r}^{\infty} [\Phi_{\alpha,\varepsilon}(f_j)] \to 0$ as $r \to \infty$, such that $|1 - \operatorname{Re} \hat{\mu}_\alpha(y)| \leq \Phi_{\alpha,\varepsilon}(y) + \varepsilon$, for all $y \in T$. Then $\{\mu_\alpha\}_{\alpha\in I}$ is relatively compact.

Proof. (a) By Theorem 3.2, one has

$$\mu_\alpha(S_n^\delta)^c \leq K_\delta \liminf_{m\to\infty} \int_{R^{m-n}} \left|1 - \operatorname{Re}\hat{\mu}_\alpha\left(s(n,m)\sum_{j=n+1}^{m} t_j b_j^{1/p} f_j\right)\right| \gamma(dt_{n+1})\cdots\gamma(dt_m).$$

Given $\varepsilon > 0$, choose a family $\{\Phi_{\alpha,\varepsilon}\}_{\alpha\in I}$ as described in the statement of (a). Then

$$\left|1 - \operatorname{Re}\hat{\mu}_\alpha\left(s(n,m)\sum_{j=n+1}^{m} t_j b_j^{1/p} f_j\right)\right| \leq \Phi_{\alpha,\varepsilon}\left(s(n,m)\sum_{j=n+1}^{m} t_j b_j^{1/p} f_j\right) + \varepsilon$$

$$\leq \sum_{i=n+1}^{m}\sum_{j=n+1}^{m} [s(n,m)]^2 t_i t_j b_i^{1/p} b_j^{1/p}\psi_{\alpha,\varepsilon}(f_i,f_j) + \varepsilon,$$

where $\psi_{\alpha,\varepsilon}$ is the bilinear form associated with $\Phi_{\alpha,\varepsilon}$. Then

$$\int_{R^{m-n}} \left|1 - \operatorname{Re}\hat{\mu}_\alpha\left(s(n,m)\sum_{j=n+1}^{m} t_j b_j^{1/p} f_j\right)\right| \gamma(dt_{n+1})\cdots\gamma(dt_m)$$

$$\leq [s(n,m)]^2 \sum_{j=n+1}^{m} b_j^{2/p}\Phi_{\alpha,\varepsilon}(f_j),$$

because $\int t\gamma(dt) = 0$ and $\int t^2\gamma(dt) = 1$. Thus we have: for each $\alpha \in I$,

$$\mu_\alpha(S_n^\delta)^c \leq K_\delta\left(\liminf_{m\to\infty} [s(n,m)]^2 \sum_{j=n+1}^{m} b_j^{2/p}\Phi_{\alpha,\varepsilon}(f_j) + \varepsilon\right).$$

The sequence $(b_j)_{j\in N}$ is still an arbitrary positive sequence. Now we make a proper choice of $(b_j)_{j\in N}$ for each $\alpha \in I$.

Put $b_{\alpha,j} = [\Phi_{\alpha,\varepsilon}(f_j)]^{p(p-2)/4}$. Then, on the one hand,

$$s(n,m) = \sum_{j=n+1}^{m} ([\Phi_{\alpha,\varepsilon}(f_j)]^{p(p-2)/4})^{-2/(2-p)} = \sum_{j=n+1}^{m} [\Phi_{\alpha,\varepsilon}(f_j)]^{p/2};$$

on the other hand

$$\sum_{j=n+1}^{m} b_j^{2/p}\Phi_{\alpha,\varepsilon}(f_j) = \sum_{j=n+1}^{m} [\Phi_{\alpha,\varepsilon}(f_j)^{p(p-2)/4}]^{2/p}\Phi_{\alpha,\varepsilon}(f_j)$$

$$= \sum_{j=n+1}^{m} [\Phi_{\alpha,\varepsilon}(f_j)]^{p/2}.$$

Let $M = \sup_{\alpha\in I} \sum_{j=1}^{\infty} [\Phi_{\alpha,\varepsilon}(f_j)]^{p/2}$. Then, for all $\alpha \in I$,

$$\mu_\alpha(S_n^\delta)^c \leq K_\delta\left[M^2 \sup_{\alpha\in I} \sum_{j=n+1}^{\infty} [\Phi_{\alpha,\varepsilon}(f_j)]^{p/2} + \varepsilon\right].$$

It is clear now that for a prescribed δ, one can choose first ε and then n so that the right-hand side of the inequality is smaller than any given number.

Next we observe that $|1 - (\mu_\alpha \circ f_j^{-1})\hat{\ }(t)| = |1 - \hat{\mu}_\alpha(tf_j)| \leq t^2\Phi_{\alpha,\varepsilon}(f_j) + \varepsilon \leq t^2 M^{2/p}$

$+\varepsilon$. This implies that for each $j \in N$, the family $\{\mu_\alpha \circ f_j^{-1}\}_{\alpha \in I}$ is relatively compact in R^1. Since $T_n \oplus S_n^\perp = X'$, an application of Theorem 2.1 concludes the proof of (a).

(b) The proof of this assertion is similar to that of (a), with the difference that now we apply Theorem 3.2(b) and no choice of b_j's is necessary. □

Obviously in Theorem 4.1, as well as in future propositions, one can take $p=1$ and obtain a statement valid for all Banach spaces with a Schauder basis, by Lemma 3.1.

PROPOSITION 4.2. *Let X be a Banach space with a Schauder basis. Let $g: T \to C$ be positive definite, $g(0)=1$. If the restriction of g to each finite-dimensional subspace of T is continuous, then for each $n \in N$ there exists a probability measure μ_n on $(X, \mathscr{F})$ such that $\mu_n(S_n)=1$ and $\hat{\mu}_n(y)=g(\sum_{j=1}^n \langle e_j, y \rangle f_j)$.*

Proof. The function g induces a weak distribution μ on the field generated by T on X. Let $\mathscr{F}_n = \{\Pi_n^{-1}(B) : B \text{ Borel in } S_n\}$. Then $\mathscr{F}_n$ is precisely the σ-field induced by $\{f_1, \ldots, f_n\}$. Therefore μ is σ-additive on $\mathscr{F}_n$. It follows that the measure ν_n, defined on the Borel sets of S_n by $\nu_n(B)=\mu(\Pi_n^{-1}(B))$, is σ-additive. We extend it to a probability measure μ_n on $(X, \mathscr{F})$ by putting $\mu_n(A)=\nu_n(A \cap S_n)$ for $A \in \mathscr{F}$. Then clearly $\mu_n(S_n)=1$ and $\mu_n = \mu \circ \Pi_n^{-1}$. Also, $\hat{\mu}_n(y)=(\mu \circ \Pi_n^{-1})^\wedge(y)=\hat{\mu}(\Pi_n^*(y))=g(\sum_{j=1}^n \langle e_j, y \rangle f_j)$.

Having fixed in X a Schauder basis $(e_i)_{i \in N}$ and the corresponding coordinate functions $(f_j)_{j \in N}$, we state the

DEFINITION 4.2. *We call τ_p the (locally convex) topology defined on X' (or T, as the case may be) by the system of seminorms $\{\Phi^{1/2} : \Phi \text{ is a quadratic nonnegative definite form on } X' \text{ (or } T), \text{ of trace class } p\}$.*

Obviously, $\tau_p \subset \tau_q$ if $p < q$.

THEOREM 4.3. *Let X be a Banach space with a Schauder basis of type p ($1 \le p < \infty$). Let $g: T \to C$ be positive definite, $g(0)=1$. If either (a) $p \in [1, 2)$ and g is $\tau_{p/2}$-continuous, or (b) $p \in [2, \infty)$ and g is τ_1-continuous, then there exists a unique probability measure μ on $(X, \mathscr{F})$ such that $\hat{\mu}(y)=g(y)$ for each $y \in T$.*

Proof. (a) Let $\eta > 0$ be given. By assumption, there exists Φ of trace class $p/2$ such that $\Phi(y) \le 2$ implies $|1 - \operatorname{Re} g(y)| \le \eta$. Then for any $y \in T$, one has

$$(1) \qquad\qquad |1 - \operatorname{Re} g(y)| \le \Phi(y) + \eta.$$

It follows that the restriction of g to each finite-dimensional subspace of T is continuous. Therefore, one can construct a sequence $\{\mu_n\}_{n \in N}$ of probability measures as described in Proposition 4.2. Also, it is clear that $\hat{\mu}_n \to g$ on T; in fact, $\hat{\mu}_m(y) = g(y)$ for all $m \ge n$ whenever $y \in T_n$. Since T is sequentially w^*-dense in X', in order to prove (a) we only have to show that $\{\mu_n\}_{n \in N}$ is relatively compact, by Theorem 2.4.

But (1) implies: for all $n \in N$, $|1 - \operatorname{Re} \hat{\mu}_n(y)| = |1 - \operatorname{Re} g(\Pi_n^*(y))| \leq \Phi(\Pi_n^*(y)) + \eta$. Since clearly the family of quadratic forms $\{\Phi \circ \Pi_n^*\}_{n \in N}$ satisfies the hypotheses of Theorem 4.1(a), the relative compactness of $\{\mu_n\}_{n \in N}$ follows.

(b) Similar to (a); Theorem 4.1(b) is used now. $\square$

Next, we construct two examples, for cases (a) and (b) respectively, showing that as far as continuity in a τ_α topology is concerned, our conditions cannot be relaxed.

EXAMPLE 4.1. *We show that for every $p \in [1, 2)$ and for every $\alpha > p/2$, there exists a Banach space with a Schauder basis, of type p, and a function $g: T \to C$ such that (1) g is positive definite and $g(0) = 1$, (2) g is τ_α-continuous but g is not the characteristic functional of a probability measure on $(X, \mathscr{F})$.*

Let $X = l_p$, $1 \leq p < 2$; then $X' = l_q$, where $q = p/p - 1$ if $p > 1$, $q = \infty$ if $p = 1$. Let $p/2 < \alpha < 1$. Putting $r = p/\alpha$, we have $p < r < 2$.

Let $(a_n)_{n \in N}$ be a sequence of positive numbers (to be completely specified later). For $y = (y_n)_{n \in N} \in T \subset l_q$, we define $g(y) = \exp(-\sum_{n=1}^\infty a_n |y_n|^r)$ (g is well defined, because $y_n = 0$ except for a finite number of n's). Clearly $g(0) = 1$. The fact that g is positive definite follows from the positive definiteness of $\exp(-|t|^r)$ in R^1. We will show that it is possible to choose $(a_n)_{n \in N}$ so that g is τ_α-continuous but not a characteristic functional.

Introducing a second sequence $(b_n)_{n \in N}$ of positive numbers and applying Hölder's inequality, we obtain

$$|1 - g(y)| \leq \sum_{n=1}^\infty a_n |y_n|^r = \sum_{n=1}^\infty (a_n b_n^{-1})(b_n |y_n|^r)$$

$$\leq \left[\sum_{n=1}^\infty (a_n b_n^{-1})^s \right]^{1/s} \left[\sum_{n=1}^\infty b_n^{2/r} y_n^2 \right]^{r/2},$$

where $s = 2/2 - r$ is the exponent conjugate to $2/r$. Thus

$$(1) \qquad |1 - g(y)| \leq \left[\sum_{n=1}^\infty (a_n b_n^{-1})^s \right]^{1/s} [\Phi(y)]^{r/2},$$

where $\Phi(y) = \sum_{n=1}^\infty b_n^{2/r} y_n^2$. Observe that $\Phi(f_j) = b_j^{2/r}$. Suppose $g = \hat{\mu}$ (on T) for some probability measure μ on $(X, \mathscr{F})$. Then, under μ, $\{f_1, \ldots, f_n, \ldots\}$ is an independent sequence such that $\mathscr{L}(a_j^{-1/r} f_j) = \sigma_r$, where σ_r is the standard symmetric stable distribution of exponent r.

Since for all $x \in X = l_p$, $\sum_{n=1}^\infty |\langle x, f_n \rangle|^p = \|x\|^p$, the Kolmogorov three series theorem [11, p. 237] implies that

$$(2) \qquad \sum_{n=1}^\infty E_\mu(|f_n|^p I_{[|f_n| \leq 1]}) < \infty.$$

Now

$$E_\mu(|f_n|^p I_{[|f_n|\leq 1]}) = a_n^{p/r} E_\mu(|a_n^{-1/r}f_n|^p I_{[|a_n^{-1/r}f_n|\leq a_n^{-1/r}]})$$

$$= a_n^{p/r} \cdot 2 \int_0^{a_n^{-1/r}} t^p \sigma_r(dt)$$

$$\geq a_n^{p/r} \int_1^{a_n^{-1/r}} t^p \sigma_r(dt)$$

$$\geq a_n^{p/r} \int_1^{a_n^{-1/r}} M \frac{t^p \, dt}{t^{r+1}}.$$

(We have used here the known fact that if ϕ_r is the density of σ_r, then there is a constant M such that $\phi_r(t) \geq M/t^{r+1}$ for $t \geq 1$.)

But $\int_1^{a_n^{-1/r}} t^{p-r-1} \, dt = (1/(p-r))[a_n^{-p/r+1} - 1]$. Therefore

$$E_\mu(|f_n|^p I_{[|f_n|\leq 1]}) \geq \frac{M}{p-r} [a_n - a_n^{p/r}] = \frac{M}{r-p} [a_n^\alpha - a_n].$$

Then we have

(3) $$\sum_{n=1}^\infty E_\mu(|f_n|^p I_{[|f_n|\leq 1]}) \geq \frac{M}{r-p} \sum_{n=1}^\infty [a_n^\alpha - a_n].$$

We show next that $(a_n)_{n\in N}$ and $(b_n)_{n\in N}$ can be chosen so that $\sum a_n < \infty$, $\sum b_n^{2\alpha/r} < \infty$, $\sum (a_n b_n^{-1})^{2/(2-r)} < \infty$, and $\sum a_n^\alpha = \infty$. It will follow then from (1) that g is τ_α-continuous; but the contradiction between (2) and (3) proves that μ cannot exist.

Put $a_n = n^{-1/\alpha}$ and $b_n = a_n^v = n^{-v/\alpha}$, where v is any number such that $p/2\alpha < v < 1 - (2\alpha-p)/2$ (this choice is possible because $p/2\alpha < 1 - (2\alpha-p)/2$ if and only if $p/2 < \alpha < 1$). Then we have $\sum a_n = \sum n^{-1/\alpha} < \infty$ because $1/\alpha > 1$, $\sum b_n^{2\alpha/r} = \sum n^{-2\alpha v/p} < \infty$ because $2\alpha v/p > 1$, $\sum (a_n b_n^{-1})^{2/(2-r)} = \sum n^{-(1/\alpha)(1-v)(2/(2-r))} < \infty$ because

$$(1/\alpha)(1-v)\frac{2}{2-r} > (1/\alpha)\left(\frac{2\alpha-p}{2}\right)\frac{2\alpha}{2\alpha-p} = 1,$$

and finally $\sum a_n^\alpha = \sum n^{-1} = \infty$.

REMARK. Since $(a_n)_{n\in N}$ is chosen so that $\sum a_n < \infty$, g can be taken to be defined on X'.

EXAMPLE 4.2. *We show that for every $p \in [2, \infty)$ and for every $\alpha > 1$, there exists a Banach space with a Schauder basis, of type p, and a function $g : T \to C$ such that the same properties stated in Example 4.1 hold.*

Let $X = l_p$, $p \leq 2 < \infty$. Let $\alpha > 1$. Choose $r \in (1, 2)$. As in Example 4.1, define for $y = (y_n)_{n\in N} \in T \subset X' : g(y) = \exp(-\sum_{n=1}^\infty a_n |y_n|^r)$. Proceeding as before, we obtain $|1 - g(y)| \leq [\sum_{n=1}^\infty (a_n b_n^{-1})^s]^{1/s} [\Phi(y)]^{r/2}$, where $s = 2/(2-r)$ and $\Phi(y) = \sum_{n=1}^\infty b_n^{2/r} y_n^2$.

Suppose $g = \hat{\mu}$ (on T) for some probability measure μ on $(X, \mathscr{F})$. Reasoning as in Example 4.1, we have the two inequalities:

$$\sum_{n=1}^{\infty} E_\mu(|f_n|^p I_{[|f_n| \leq 1]}) < \infty,$$

$$\sum_{n=1}^{\infty} E_\mu(|f_n|^p I_{[|f_n| \leq 1]}) \geq \frac{M}{p-r} \sum_{n=1}^{\infty} (a_n - a_n^{p/r}).$$

We show next that $(a_n)_{n \in N}$ and $(b_n)_{n \in N}$ can be chosen so that $\sum a_n^{p/r} < \infty$, $\sum b_n^{2\alpha/r} < \infty$, $\sum (a_n b_n^{-1})^{2/(2-r)} < \infty$ and $\sum a_n = \infty$.

As in Example 4.1, this will show that μ cannot exist. Put $a_n = n^{-1}$, $b_n = a_n^v = n^{-v}$, where v is any number such that $r/2\alpha < v < r/2$ (this choice is possible because $\alpha > 1$). Then we have $\sum a_n^{p/r} < \infty$ because $p/r > 1$, $\sum b_n^{2\alpha/r} = \sum n^{-2\alpha v/r} < \infty$ because $2\alpha v/r > 1$, $\sum (a_n b_n^{-1})^{2/(2-r)} = \sum n^{-(1-v)2/(2-r)} < \infty$ because $(1-v)2/(2-r) > (1 - r/2)(2/(2-r)) = 1$ and $\sum a_n = \infty$.

COROLLARY. *Let $p > 2$. Then there exists no topology σ on l_p' such that if g is positive definite on l_p' and $g(0) = 1$, then $g = \hat{\mu}$ for some probability measure μ on l_p if and only if g is σ-continuous.*

Proof. In the terminology of Example 4.2, let $\alpha = p/2$. Then $\sum_{j=1}^{\infty} [\Phi(f_j)]^{p/2} = \sum_{j=1}^{\infty} b_j^{2\alpha/r} < \infty$. According to Theorem 5.2 (§5 of this paper), $\exp(-\Phi)$ is the characteristic functional of a Gaussian probability measure on l_p. Now suppose σ exists. Then Φ would be σ-continuous, and consequently the function g of Example 4.2 would be σ-continuous, hence the characteristic functional of a probability measure on l_p. But Example 4.2 shows that this is impossible. □

REMARK. Professor LeCam communicated to us an example of R. M. Dudley concerning the question of "decomposability" of the composition of a linear process and a nuclear operator. This example has recently appeared in [4, Proposition (8.1)]. Although Dudley uses the Cauchy distribution and deals with a different (though closely related) problem, we found his example useful for the construction of Examples 4.1 and 4.2.

5. Gaussian measures.   We recall several definitions (see [10, p. 229]).

DEFINITION 5.1. *Let X be a locally convex topological vector space, μ a weak distribution on X. μ is said to have a covariance if $\int |\langle x, y \rangle|^2 \mu(dx) < \infty$ for each $y \in X'$. The quadratic form $y \to \int |\langle x, y \rangle|^2 \mu(dx)$ will be called the covariance form of μ.*

DEFINITION 5.2. *Let X and μ be as in Definition 5.1. If $\int |\langle x, y \rangle| \mu(dx) < \infty$ for each $y \in X'$, and there exists in X an element x_0 (necessarily unique), such that $\langle x_0, y \rangle = \int \langle x, y \rangle \mu(dx)$, then x_0 is called the expectation of μ.*

DEFINITION 5.3. *Let X and μ be as in Definition 5.1. Then μ is said to be Gaussian if for every $y \in X'$, the probability measure $\mu \circ y^{-1}$ in R^1 is Gaussian (possibly degenerate).*

Let us remark that a Gaussian weak distribution μ always has a covariance, for $\int |\langle x, y \rangle|^2 \mu(dx) = \int t^2 (\mu \circ y^{-1})(dt) < \infty$. Let $\rho(y) = \int \langle x, y \rangle \mu(dx)$,

$$\Phi(y) = \int |\langle x, y \rangle|^2 \mu(dx).$$

Then $\hat{\mu}(y) = \exp [i\rho(y) - \tfrac{1}{2}(\Phi(y) - \rho^2(y))]$.

THEOREM 5.1. *Let X be as in Definition 5.1. Suppose μ is a Gaussian probability measure on $(X, \mathcal{F})$. Then ρ and Φ are sequentially $\sigma(X', X)$ continuous.*

Proof. Suppose $y_n \to y$ in the $\sigma(X', X)$ topology. Then $\hat{\mu}(ty_n) \to \hat{\mu}(ty)$, for each $t \in R^1$. But it is not difficult to see from the form of $\hat{\mu}$ that this implies $\rho(y_n) \to \rho(y)$ and $\Phi(y_n) \to \Phi(y)$.

REMARK. Questions concerning the continuity of ρ and Φ are studied in [10].

We construct now Gaussian measures in Banach spaces with a Schauder basis. Although Theorem 4.3 could be applied, it turns out that a direct approach yields a stronger result.

THEOREM 5.2. *Let X be a Banach space with a Schauder basis of type p ($1 \le p < \infty$). Let Φ be a nonnegative definite quadratic form of trace class $p/2$, defined on T. Let $x_0 \in X$. Then there exists a Gaussian probability measure μ on $(X, \mathcal{F})$ such that $\hat{\mu}(y) = \exp [i\langle x_0, y \rangle - \Phi(y)/2]$ on T.*

Proof. Let $g(y) = \exp(-\Phi(y)/2)$ for $y \in T$. Then it is not difficult to prove that g is positive definite (for example by using the fact that the Fourier transform of a measure in R^n is always positive definite).

Let $\{\mu_k\}$ be the sequence of measures constructed from g in Proposition 4.2. Since $\hat{\mu}_k \to g$ on T, to prove the existence of a probability measure ν such that $\hat{\nu} = g$ on T, it is enough to show that $\{\mu_k\}$ is flatly concentrated.

Now

$$\mu_k(S_n^\varepsilon)^c = \mu_k(\{x : [p_n(x)]^p > \varepsilon^p\})$$

$$\le \varepsilon^{-p} \int [p_n(x)]^p \mu_k(dx)$$

(1)
$$\le \liminf_{m \to \infty} \varepsilon^{-p} \int [p_n(\Pi_m(x))]^p \mu_k(dx)$$

$$\le \liminf_{m \to \infty} \varepsilon^{-p} \int \sum_{j=n+1}^{m} |\langle x, f_j \rangle|^p \mu_k(dx)$$

$$\le \varepsilon^{-p} \sum_{j=n+1}^{\infty} \int |\langle x, f_j \rangle|^p \mu_k(dx).$$

We use now the following fact, which can be easily verified by calculation: For every $r > 0$ and $s > 0$, there exists a constant $C(r, s)$, such that if γ is a normal distribution with mean 0 and arbitrary variance, then

$$\int |t|^r \gamma(dt) = C(r, s) \left[\int |t|^s \gamma(dt) \right]^{r/s}.$$

We have then $\int |\langle x, f_j \rangle|^p \mu_k(dx) = C(p, 2)[\int |\langle x, f_j \rangle|^2 \mu_k(dx)]^{p/2}$. But for any k, $\int |\langle x, f_j \rangle|^2 \mu_k(dx) = \Phi(f_j)$ or 0. From (1) we obtain now: for any k,

$$\mu_k(S_n^\varepsilon)^c \leqq \varepsilon^{-p} C(p, 2) \sum_{j=n+1}^\infty [\Phi(f_j)]^{p/2}.$$

Since Φ is of trace class $p/2$, this proves that $\{\mu_k\}$ is flatly concentrated.

Now if δ_{x_0} is point mass at x_0, then $\mu = \delta_{x_0} * \nu$. $\square$

REMARK. The property of the normal distribution mentioned in this theorem is also used by Vakhania [22].

DEFINITION 5.4. *We call $\gamma_{(\Phi, x_0)}$ the Gaussian measure constructed in Theorem 5.2. If $x_0 = 0$, we write $\gamma_{(\Phi, 0)} = \gamma_\Phi$.*

6. A Lindeberg-type central limit theorem. For the sake of convenience, we state the classical theorem of Lindeberg (see, for instance, [11, p. 280]), in the form:

CLASSICAL LINDEBERG THEOREM. *Let $(\Omega, \mathscr{A}, P)$ be a probability space. Let $Z_{n,j}: \Omega \to R^1$, $n \in N$, $j = 1, \ldots, k_n$, be a triangular system of random variables such that $E(Z_{nj}^2) < \infty$ and $E(Z_{nj}) = 0$ for $n \in N$, $j = 1, \ldots, k_n$, and $\{Z_{nj}\}$, $j = 1, \ldots, k_n$, is an independent set for each n. Suppose*:

(1) *For every $\varepsilon > 0$, $\sum_{j=1}^{k_n} E(Z_{nj}^2 I_{[|Z_{nj}| > \varepsilon]}) \to 0$ as $n \to \infty$.*

(2) *$\sum_{j=1}^{k_n} E(Z_{nj}^2) \to a$ as $n \to \infty$, for some $a > 0$.*

Then for each $t \in R^1$, $E[\exp(it \sum_{j=1}^{k_n} Z_{nj})] \to \exp(-at^2/2)$ as $n \to \infty$, and therefore $\mathscr{L}(\sum_{j=1}^{k_n} Z_{nj})$ converges weakly to the Gaussian distribution with mean 0 and variance a.

We prove next a generalization of this theorem to the case of a triangular system of random variables taking values in a Banach space with a Schauder basis. Let us remark that, since we deal with separable Banach spaces, measurability questions pose no difficulty. In particular, the sum of two random variables is again a random variable.

THEOREM 6.1. *Let X be a Banach space with a Schauder basis $(e_i)_{i \in N}$ and coordinate functions $(f_j)_{j \in N}$, of type p with respect to $(e_i)_{i \in N}$, where $p \in [1, 2)$. Let T_n be the subspace of X' generated by $\{f_1, \ldots, f_n\}$, and let $T = \bigcup_{n=1}^\infty T_n$. Let $(\Omega, \mathscr{A}, P)$ be a probability space. Let $Z_{nj}: \Omega \to X$, $n \in N$, $j = 1, \ldots, k_n$, be a triangular system of X-valued random variables, such that for each $y \in T$, $E|\langle Z_{nj}, y \rangle|^2 < \infty$ and $E\langle Z_{n,j}, y \rangle = 0$, $n \in N$, $j = 1, \ldots, k_n$, and for each n $\{\langle Z_{nj}, y \rangle\}_{j=1,\ldots,k_n}$ is an independent set of real-valued random variables. Let us call $\Phi_{nj}(y) = E|\langle Z_{nj}, y \rangle|^2$. Suppose*:

(1) *For every $y \in T$, for every $\varepsilon > 0$, $\sum_{j=1}^{k_n} E(|\langle Z_{nj}, y \rangle|^2 I_{[|\langle Z_{nj}, y \rangle| > \varepsilon]}) \to 0$ as $n \to \infty$.*

(2) *There exists a function Φ defined on T such that for each $y \in T$, $\sum_{j=1}^{k_n} \Phi_{nj}(y) \to \Phi(y)$.*

(3) *If $\psi_n(y) = \sum_{j=1}^{k_n} \Phi_{nj}(y)$, then $\{\psi_n\}_{n \in N}$ satisfies $\sup_{n \in N} \sum_{i=1}^\infty [\psi_n(f_i)]^{p/2} < \infty$, $\sup_{n \in N} \sum_{i=r}^\infty [\psi_n(f_i)]^{p/2} \to 0$ as $r \to \infty$.*

Then Φ is a quadratic nonnegative definite form of trace class $(p/2)$ and $\mathscr{L}(\sum_{j=1}^{k_n} Z_{nj})$ converges weakly to γ_Φ.

Proof. Since $\Phi(y) = \lim_{n\to\infty} \psi_n(y)$, Fatou's lemma yields

$$\sum_{i=1}^{\infty} [\Phi(f_i)]^{p/2} \leq \liminf_{n\to\infty} \sum_{i=1}^{\infty} [\psi_n(f_i)]^{p/2}$$

$$\leq \sup_{n\in N} \sum_{i=1}^{\infty} [\psi_n(f_i)]^{p/2} < \infty,$$

and hence Φ has the desired properties.

Let $\mu_n = \mathscr{L}(\sum_{j=1}^{k_n} Z_{nj})$. Then for each $y \in T$

$$|1 - \operatorname{Re} \hat{\mu}_n(y)| = \left| \int (1 - \cos \langle x, y \rangle) \mu_n(dx) \right|$$

$$\leq \tfrac{1}{2} \int |\langle x, y \rangle|^2 \mu_n(dx).$$

But

$$\int \langle x, y \rangle^2 \mu_n(dx) = E \left\langle \sum_{j=1}^{k_n} Z_{nj}, y \right\rangle^2$$

$$= \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} E(\langle Z_{ni}, y \rangle \langle Z_{nj}, y \rangle) = \sum_{j=1}^{k_n} E\langle Z_{nj}, y \rangle^2$$

$$= \sum_{j=1}^{k_n} \Phi_{nj}(y) = \psi_n(y),$$

using the independence of $\{Z_{nj}\}_{j=1,\ldots,k_n}$ and $E\langle Z_{nj}, y \rangle = 0$. Thus, for each $y \in T$, $|1 - \operatorname{Re} \hat{\mu}_n(y)| \leq \tfrac{1}{2}\psi_n(y)$. Applying Theorem 4.1 and taking into account the properties (3) of $\{\psi_n\}_{n\in N}$, it follows that $\{\mu_n\}_{n\in N}$ is relatively compact. Assumptions (1) and (2) and the Classical Lindeberg Theorem imply that

$$\hat{\mu}_n(y) = E \left[\exp \left(i \left\langle \sum_{j=1}^{k_n} Z_{nj}, y \right\rangle \right) \right] \to \exp(-\Phi(y)/2)$$

for each $y \in T$. But, as we know, T is sequentially w^*-dense in X'; therefore, by Theorem 2.4, μ_n converges weakly to γ_Φ. $\square$

Of course, by taking $p = 1$, this theorem includes a statement valid for all Banach spaces with a Schauder basis.

For spaces of type $p \geq 2$, we have the proposition:

THEOREM 6.2. *Suppose the conditions of Theorem 6.1 are verified, with the only changes: $p \geq 2$ instead of $p \in [1, 2)$, and (3) is replaced by $\sup_{n\in N} \sum_{i=1}^{\infty} \psi_n(f_i) < \infty$, $\sup_{n\in N} \sum_{i=r}^{\infty} \psi_n(f_i) \to 0$ as $r \to \infty$.*

Then Φ is a quadratic nonnegative definite form of trace class 1 and $\mathscr{L}(\sum_{j=1}^{k_n} Z_{nj})$ converges weakly to γ_Φ.

Proof. Similar to that of Theorem 6.1, only that instead of Theorem 4.1(a) we apply Theorem 4.1(b). □

From Theorems 6.1 and 6.2 we obtain now

THEOREM 6.3. *Suppose the space X is as in Theorem 6.1., only that we assume $p \in [1, \infty)$.*

Let μ be a probability measure on $(X, \mathscr{F})$ such that for each $y \in T$, $\int |\langle x, y \rangle|^2 \mu(dx) < \infty$ and $\int \langle x, y \rangle \mu(dx) = 0$. Let $(\Omega, \mathscr{A}, P)$ be a probability space and $Z_j : \Omega \to X$, $j \in N$, independent X-valued random variables with $\mathscr{L}(Z_j) = \mu$ for all $j \in N$. Let $\Phi(y) = \int |\langle x, y \rangle|^2 \mu(dx)$, $y \in T$. If either:

(1) *$p \in [1, 2)$ and Φ is of trace class $(p/2)$ or*

(2) *$p \in [2, \infty)$ and Φ is of trace class 1,*

then $\mathscr{L}(n^{-1/2}(\sum_{j=1}^n Z_j))$ converges weakly to γ_Φ.

Proof. If we put $Z_{nj} = n^{-1/2} Z_j$, $j = 1, \ldots, n$, then $\sum_{j=1}^n \Phi_{nj}(y) = \Phi(y)$ and

$$\sum_{j=1}^n E(|\langle Z_{nj}, y \rangle|^2 I_{[|\langle Z_{nj}, y \rangle| > \varepsilon]}) = \int_{[|\langle x, y \rangle| > \varepsilon n^{1/2}]} |\langle x, y \rangle|^2 \mu(dx).$$

The statement follows at once. □

REMARK. Theorem 6.3 can be proved directly as follows:

$$P\left[n^{-1/2} \sum_{j=1}^n Z_j \in (S_r^\varepsilon)^c \right] = P\left[p_r\left(n^{-1/2} \sum_{j=1}^n Z_j \right) > \varepsilon \right]$$

$$\leqq \varepsilon^{-p} \sum_{i=r+1}^\infty E\left| \left\langle n^{-1/2} \sum_{j=1}^n Z_j, f_i \right\rangle \right|^p.$$

Let $1 \leqq p < 2$. But

$$E\left| \left\langle n^{-1/2} \sum_{j=1}^n Z_j, f_i \right\rangle \right|^p \leqq \left[E\left| \left\langle n^{-1/2} \sum_{j=1}^n Z_j, f_i \right\rangle \right|^2 \right]^{p/2} = [\Phi(f_i)]^{p/2}.$$

Therefore $P[n^{-1/2} \sum_{j=1}^n Z_j \in (S_r^\varepsilon)^c] \leqq \varepsilon^{-p} \sum_{i=r+1}^\infty [\Phi(f_i)]^{p/2}$. It follows that

$$\left\{ \mathscr{L}\left(n^{-1/2} \sum_{j=1}^n Z_j \right) \right\}$$

is flatly concentrated. Since the characteristic functionals converge to

$$\exp(-\Phi(y)/2),$$

the statement follows.

If $p \geqq 2$, then X is of type 2 and one can proceed similarly. □

7. Application to the spaces l_p. In this section we show that certain results we have proved can be sharpened in the case of l_p spaces.

We take in l_p, $1 \leqq p < \infty$, the usual Schauder basis. Let Q_r ($r > 0$) be the class of all quadratic nonnegative definite forms of trace class r, defined on T.

THEOREM 7.1. *The map* $(\Phi, z) \to \gamma_{(\Phi,z)}$ *is a bijection of* $Q_{p/2} \times l_p$ *onto the class of all Gaussian probability measures on* l_p.

Proof. The fact that to each pair $\Phi \in Q_{p/2}$, $z \in l_p$ there corresponds a Gaussian measure $\gamma_{(\Phi,z)}$ has been proved in Theorem 5.2. Now suppose μ is a Gaussian measure on l_p. Let $\rho(y) = \int \langle x, y \rangle \mu(dx)$, $\Phi_1(y) = \int |\langle x, y \rangle|^2 \mu(dx)$ for $y \in X'$. Then $\hat{\mu}(y) = \exp[i\rho(y) - \frac{1}{2}(\Phi_1(y) - \rho^2(y))]$. We want to show that $(\Phi_1 - \rho^2)$ is of trace class $p/2$. Call $\Phi = \Phi_1 - \rho^2$. Define $\bar{\mu}$ on $(X, \mathscr{F})$ by $\bar{\mu}(B) = \mu(-B)$ for $B \in \mathscr{F}$. Then if $\nu = \mu * \bar{\mu}$, we have $\hat{\nu}(y) = \exp[-\Phi(y)]$. Let $\varepsilon < 1/2$. If we choose a ball

$$B = \{x \in X : \|x\| \leqq r\}$$

such that $\nu(B^c) < \varepsilon/2$, we have:

$$1 - \exp(-\Phi(y)) = \int [1 - \exp(i\langle x, y \rangle)]\nu(dx)$$

$$\leqq \int_B |1 - \exp(i\langle x, y \rangle)|\nu(dx) + \varepsilon$$

$$\leqq \int_B |\langle x, y \rangle|\nu(dx) + \varepsilon$$

$$\leqq \left[\int_B |\langle x, y \rangle|^p \nu(dx)\right]^{1/p} + \varepsilon.$$

Let $\psi(y) = \int_B |\langle x, y \rangle|^p \nu(dx)$. Then if $\psi(y) \leqq \varepsilon^p$, one must have $\Phi(y) \leqq M_1 < \infty$. Let $y \in X'$. If $\psi(y) \neq 0$, then $\psi[\varepsilon[\psi(y)]^{-1/p}y] = \varepsilon^p$; therefore $\Phi(\varepsilon[\psi(y)]^{-1/p}y) \leqq M_1$, so $\Phi(y) \leqq M_1 \varepsilon^{-2}[\psi(y)]^{2/p}$, and finally $[\Phi(y)]^{p/2} \leqq M_2 \psi(y)$. If $\psi(y) = 0$, then

$$\nu\{x \in B : \langle x, y \rangle = 0\} = \nu(B) > 0.$$

Then $\nu \circ y^{-1}(\{0\}) > 0$, and the fact that ν is Gaussian implies that $\nu \circ y^{-1}(\{0\}) = 1$. Therefore $\Phi(y) = \int_{\{x:\langle x,y \rangle = 0\}} \langle x, y \rangle^2 \nu(dx) = 0$. Hence $[\Phi(y)]^{p/2} \leqq M_2 \psi(y)$ holds for all $y \in X'$.

Then we have

$$\sum_{j=1}^{\infty} [\Phi(f_j)]^{p/2} \leqq M_2 \sum_{j=1}^{\infty} \int_B |\langle x, f_j \rangle|^p \nu(dx)$$

$$= M_2 \int_B \left(\sum_{j=1}^{\infty} |\langle x, f_j \rangle|^p\right)\nu(dx)$$

$$= M_2 \int_B \|x\|^p \nu(dx) \leqq M_2 r^p < \infty.$$

This proves $\Phi \in Q_{p/2}$.

A similar argument using homogeneity properties shows that

$$\sum_{j=1}^{\infty} \left|\int \langle x, f_j \rangle \mu(dx)\right|^p = \sum_{j=1}^{\infty} |\rho(f_j)|^p < \infty.$$

Define now $z = \sum_{j=1}^{\infty} [\int \langle x, f_j \rangle \mu(dx)] e_j$. Then if $y \in T$, $y = \sum_{i=1}^{n} a_i f_i$, we have

$$\langle z, y \rangle = \sum_{i=1}^{n} a_i \int \langle x, f_i \rangle \mu(dx)$$

$$= \int \left\langle x, \sum_{i=1}^{n} a_i f_i \right\rangle \mu(dx)$$

$$= \rho(y).$$

Thus $\langle z, y \rangle = \rho(y)$ for all $y \in T$. But T is sequentially w^*-dense in X' and both functions are sequentially w^*-continuous. It follows that $\langle z, y \rangle = \rho(y)$ for all $y \in X'$. $\square$

REMARK. If $p > 2$ and μ is a probability measure on l_p, then $\hat{\mu}$ is $\tau_{p/2}$-continuous. For:

$$|1 - \operatorname{Re} \hat{\mu}(y)| = \left| \int (1 - \cos \langle x, y \rangle) \mu(dx) \right|$$

$$\leq \tfrac{1}{2} \int_B |\langle x, y \rangle|^2 \mu(dx) + \varepsilon,$$

where $B = \{x : \|x\| \leq r\}$ and $\mu(B) > 1 - \varepsilon/2$. Put $\Phi(y) = \int_B |\langle x, y \rangle|^2 \mu(dx)$. Then $[\Phi(y)]^{p/2} \leq \int_B |\langle x, y \rangle|^p \mu(dx)$. Therefore

$$\sum_{j=1}^{\infty} [\Phi(f_j)]^{p/2} \leq \sum_{j=1}^{\infty} \int_B |\langle x, f_j \rangle|^p \mu(dx)$$

$$\leq \int_B \|x\|^p \mu(dx) \leq r^p < \infty.$$

Therefore Φ is of trace class $p/2$, and the statement follows. $\square$

THEOREM 7.2. *Let $X = l_p$. Let μ be a probability measure on $(X, \mathscr{F})$ such that for each $y \in T$, $\int |\langle x, y \rangle|^2 \mu(dx) < \infty$ and $\int \langle x, y \rangle \mu(dx) = 0$. Let $(\Omega, \mathscr{A}, P)$ be a probability space and $Z_j : \Omega \to X$, $j \in N$, independent X-valued random variables with $\mathscr{L}(Z_j) = \mu$ for all j. Let $\Phi(y) = \int |\langle x, y \rangle|^2 \mu(dx)$. Then:*

(1) *If $\mathscr{L}(n^{-1/2}(\sum_{j=1}^{n} Z_j))$ converges weakly, then Φ is of trace class $p/2$.*

(2) *If $1 \leq p \leq 2$, then $\mathscr{L}(n^{-1/2}(\sum_{j=1}^{n} Z_j))$ converges weakly if and only if Φ is of trace class $p/2$.*

Proof. It is clear that only (1) requires proof. Suppose $\mathscr{L}(n^{-1/2}(\sum_{j=1}^{n} Z_j))$ converges weakly to ν. Then ν is Gaussian and has covariance form Φ. By Theorem 7.1, Φ is of trace class $p/2$. $\square$

REMARK. Vakhania [22], [23] has obtained with different methods Theorem 7.1 and part of Theorem 7.2. He proves also that $\int \|x\|^r \mu(dx) < \infty$ for any $r > 0$ when μ is a Gaussian measure on l_p ($1 \leq p < \infty$).

8. Applications to the space $C[0, 1]$ and to Gaussian processes. From Theorems 2.3 and 2.4 we obtain now compactness and convergence criteria for measures on $C[0, 1]$. Although these criteria are close to standard ones (see [1, Chapter 2]),

the replacement of a direct tightness assumption by the assumption of "flat concentration" seems to be of some interest. Let $M[0, 1]$ be the space of (finite) signed measures on $[0, 1]$. Then $(C[0, 1])' = M[0, 1]$ by the Riesz representation theorem. For $t \in [0, 1]$, δ_t will denote unit point mass at t. We denote by W the subspace of $M[0, 1]$ consisting of all measures of the form $\sum_{j=1}^{n} \lambda_j \delta_{t_j}$, for $\lambda_1, \ldots, \lambda_n \in R^1$, $t_1, \ldots, t_n \in [0, 1]$. Finally, $(X, \mathscr{F})$ will denote the space $C[0, 1]$ with its Borel σ-field.

THEOREM 8.1. *Let $\{\mu_\alpha\}_{\alpha \in I}$ be a set of probability measures on $(X, \mathscr{F})$. Then $\{\mu_\alpha\}_{\alpha \in I}$ is relatively compact if and only if*
   (a) *$\{\mu_\alpha\}_{\alpha \in I}$ is flatly concentrated,*
   (b) *for every $t \in [0, 1]$, the set of measures $\{\mu_\alpha \circ \delta_t^{-1}\}_{\alpha \in I}$ is tight (in R^1).*

Proof. The subspace W is (sequentially) w^*-dense in $M[0, 1] = X'$. Therefore $\{\delta_t : t \in [0, 1]\}$ is w^*-total. The statement follows then by Theorem 2.3. $\square$

THEOREM 8.2. *Let $(\Omega, \mathscr{A}, P)$ be a probability space. Let $Z_m: \Omega \to X$, $m \in N$, be a sequence of X-valued random variables. Suppose:*
   (a) *$\{\mathscr{L}(Z_m)\}_{m \in N}$ is flatly concentrated.*
   (b) *There exists a function $g: W \to C$, continuous in $(\lambda_1, \ldots, \lambda_n)$ for each fixed $(t_1, \ldots, t_n)$ such that for each $(\lambda_1, \ldots, \lambda_n)$ and $(t_1, \ldots, t_n)$, as $m \to \infty$,*

$$E\left[\exp\left(i \sum_{j=1}^{n} \lambda_j Z_m(t_j)\right)\right] \to g\left(\sum_{j=1}^{n} \lambda_j \delta_{t_j}\right).$$

Then there exists a probability measure μ on $(X, \mathscr{F})$ such that $\mathscr{L}(Z_m)$ converges weakly to μ and $\hat{\mu}(y) = g(y)$ for $y \in W$.

Proof. Our assumption on g means that the restriction of g to each finite-dimensional subspace of W is continuous. The statement follows from Theorem 2.4 and the already used fact that W is sequentially w^*-dense in $M[0, 1]$. $\square$

Next we determine the form of the characteristic functionals of Gaussian measures on $C[0, 1]$. We prove first a lemma relating the usual notion of continuous Gaussian process to the notion of Gaussian measure.

LEMMA 8.3. *Let μ be a probability measure on $(X, \mathscr{F})$ such that for all finite sets $\{t_1, \ldots, t_n\} \subset [0, 1]$, the measure $\mu \circ (\delta_{t_1}, \ldots, \delta_{t_n})^{-1}$ is Gaussian (possibly degenerate) in R^n. Then μ is Gaussian. The reciprocal also holds.*

Proof. The hypothesis implies that for any $(\lambda_1, \ldots, \lambda_n)$, $(t_1, \ldots, t_n)$, the measure $\mu \circ (\sum_{j=1}^{n} \lambda_j \delta_{t_j})^{-1}$ is Gaussian in R^1. If $y \in X'$ and $\{y_n\}$ is a sequence in W which converges w^* to y, then $\mu \circ y_n^{-1}$ converges weakly to $\mu \circ y^{-1}$ (by the dominated convergence theorem). But this implies that $\mu \circ y^{-1}$ is Gaussian. The reciprocal is obvious. $\square$

THEOREM 8.4. *Let P be a Gaussian measure on $(X, \mathscr{F})$. Let $m(t) = \int x(t) P(dx)$,*

$K(s, t) = \int [x(s) - m(s)][x(t) - m(t)]P(dx)$, $s, t \in [0, 1]$. *Then m and K are continuous functions, and* $\hat{P}(v) = \exp [i \int m(t)v(dt) - \frac{1}{2} \int K(s, t)v(ds)v(dt)]$, $v \in X'$.

Proof. Since the map $t \to \delta_t$ of $[0, 1]$ into $M[0, 1]$ is continuous when $M[0, 1]$ is given the w^*-topology, it follows from Theorem 5.1 and the fact that $x(t) = \langle x, \delta_t \rangle$ that m and K are sequentially continuous, hence continuous.

If $\rho(v) = \int \langle x, v \rangle P(dx)$ and $\Phi(v) = \int \langle x, v \rangle^2 P(dx)$, then we know from §5 that $\hat{P}(v) = \exp [i\rho(v) - \frac{1}{2}[\Phi(v) - \rho^2(v)]]$. But

$$\rho(v) = \int \left[\int x(t)v(dt) \right] P(dx) = \int v(dt) \int x(t)P(dx) = \int m(t)v(dt).$$

(The use of Fubini's theorem is not difficult to justify.) Also, by a similar computation, $\Phi(v) - \rho^2(v) = \int K(s, t)v(ds)v(dt)$. $\square$

REMARK. For the particular case in which P is Wiener measure (see, for instance, [1, Chapter 2]), we have $m(t) \equiv 0$ and $K(s, t) = \min (s, t)$. Hence:

COROLLARY 8.5. *Let P be Wiener measure on $(X, \mathcal{F})$. Then*

$$\hat{P}(v) = \exp \left[-\frac{1}{2} \int \min (s, t)v(ds)v(dt) \right], \qquad v \in M[0, 1].$$

We prove next a criterion for the continuity of Gaussian processes.

We use in this proposition a basis for $C[0, 1]$. One such basis was constructed by Schauder. This basis is described in [2, p. 69]; we will just display the coordinate functions. For convenience, we write $\delta(t)$ instead of δ_t. For $n \geq 1$, we denote by (p_n, q_n) the unique pair of nonnegative integers such that $n = 2^{q_n - 1} + p_n$. The system of coordinate functions is

$$f_1 = \delta(1), \qquad f_2 = \delta(0),$$

$$f_{n+2} = \delta\left(\frac{2p_n + 1}{2^{q_n}}\right) - \frac{1}{2}\left[\delta\left(\frac{p_n}{2^{q_n - 1}}\right) + \delta\left(\frac{p_n + 1}{2^{q_n - 1}}\right)\right] \quad \text{for } n \geq 1.$$

We formulate the theorem in terms of this basis; it is clear that another basis could be used instead (with the appropriate modifications).

THEOREM 8.6. *Let $\{Z_t\}_{t \in [0, 1]}$ be a real-valued stochastic process such that for each finite set $\{t_1, \ldots, t_n\} \subset [0, 1]$, $\mathcal{L}(Z_{t_1}, \ldots, Z_{t_n})$ is centered Gaussian. Let $K(s, t) = E(Z_s Z_t)$. Let $\Phi(v) = \int K(s, t)v(ds)v(dt) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(t_i, t_j)$ for $v = \sum_{j=1}^n \lambda_j \delta_{t_j} \in W$.*

Suppose $\sum_{j=1}^\infty [\Phi(f_j)]^{1/2} < \infty$. Then there exists a probability measure P on $(X, \mathcal{F})$ such that for each $\{t_1, \ldots, t_n\} \subset [0, 1]$, $P \circ (\delta_{t_1}, \ldots, \delta_{t_n})^{-1} = \mathcal{L}(Z_{t_1}, \ldots, Z_{t_n})$.

Proof. If T_n is the subspace of X' generated by $\{f_1, \ldots, f_n\}$, and $T = \bigcup_{n=1}^\infty T_n$, then $T \subset W$ and hence Φ is defined on T. Clearly Φ is a quadratic nonnegative definite form on T, of trace class 1/2 by assumption. The statement follows now from Theorem 5.2 and the fact that

$$E\left[\exp \left[i \sum_{j=1}^n \lambda_j Z_{t_j} \right] \right] = \exp \left[-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j K(t_i, t_j) \right]. \qquad \square$$

REMARK. Computation shows that in the case of Wiener measure, that is, when $K(s, t) = \min (s, t)$, one has (for our special basis) $\sum_{j=1}^{\infty} [\Phi(f_j)]^{1/2} = \infty$. Of course, Wiener measure can be constructed in several ways, exploiting the special properties of K in this case (such as the independent increment property of a process with covariance kernel K).

Thus Theorem 8.6 provides a comparatively simple sufficient condition for a Gaussian process to have continuous sample paths; however, this condition is by no means necessary. In fact, no necessary and sufficient conditions for the continuity of the paths of a Gaussian process seem to be known.

Sufficient conditions of a rather different type have been obtained by Fernique, Delporte and Dudley. (See [3] for these results.)

In connection with Theorem 8.6, a natural question arises: given a (centered) Gaussian process $\{Z_t\}_{t\in[0,1]}$ which has continuous paths, is it possible to find a basis of $C[0, 1]$ such that $\sum_{j=1}^{\infty} [\Phi(f_j)]^{1/2} < \infty$, where $(f_j)_{j\in N}$ are the coordinate functions for the basis and Φ is the quadratic form on $M[0, 1]$ associated to the covariance kernel of the process?

REFERENCES

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

2. Mahlon M. Day, *Normed linear spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Heft 21, Springer-Verlag, Berlin, 1962. MR **26** #2847.

3. R. M. Dudley, *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes*, J. Functional Analysis **1** (1967), 290–330. MR **36** #3405.

4. ――――, *Random linear functionals*, Trans. Amer. Math. Soc. **136** (1969), 1–24.

5. R. M. Dudley and V. Strassen, "The central limit theorem and ε-entropy," in *Probability and information theory*, Lecture Notes in Math., no. 89, Springer-Verlag, Berlin, 1969.

6. R. Fortet and E. Mourier, *Les fonctions aléatoires comme éléments aléatoires dans les espaces de Banach*, Studia Math. **15** (1955), 62–79. MR **19**, 1202.

7. I. M. Gel'fand and N. Ja. Vilenkin, *Generalized functions*. Vol. 4: *Some applications of harmonic analysis*, Fizmatgiz, Moscow, 1961; English transl., Academic Press, New York, 1964. MR **26** #4173; MR **30** #4152.

8. L. Gross, *Harmonic analysis on Hilbert space*, Mem. Amer. Math. Soc. No. 46 (1963). MR **28** #4304.

9. J. Kisyński, *On the generation of tight measures*, Studia Math. **30** (1968), 141–151. MR **38** #290.

10. L. LeCam, *Convergence in distribution of stochastic processes*, Univ. Calif. Publ. Statist. **2** (1957), 207–236. MR **19**, 128.

11. M. Loève, *Probability theory. Foundations. Random sequences*, 2nd ed., University Series in Higher Math., Van Nostrand, Princeton, N. J., 1960. MR **23** #A670.

12. R. A. Minlos, *Generalized random processes and their extension to a measure*, Trudy Moskov. Mat. Obšč. **8** (1959), 497–518; English transl., Selected Transl. Math. Stat. and Prob., vol. 3, Amer. Math. Soc., Providence, R. I., 1963, pp. 291–313. MR **21** #7563; MR **27** #4266.

13. E. Mourier, *Eléments aléatoires dans un espace de Banach*, Ann. Inst. H. Poincaré **13** (1953), 161–244. MR **16**, 268.

14. K. R. Parthasarathy, *Probability measures on metric spaces*, Probability and Mathematical Statistics, no. 3, Academic Press, New York, 1967. MR **37** #2271.

15. Ju. V. Prohorov, *Convergence of random processes and limit theorems in the theory of probability*, Teor. Verojatnost. i Primenen. **1** (1956), 177–238 = Theor. Probability Appl. (1956), 106–134. MR **18**, 943.

16. ———, *The method of characteristic functionals*, Proc. Fourth Berkeley Sympos. Math. Statist. and Prob. (Berkeley, Calif., 1960), vol. 2, Univ. of California Press, Berkeley, 1961, pp. 403–419. MR **24** #A3670.

17. W. Rudin, *Fourier analysis on groups*, Interscience Tracts in Pure and Appl. Math., no. 12, Interscience, New York, 1962. MR **27** #2808.

18. V. V. Sazonov, *On characteristic functionals*, Teor. Verojatnost. i Primenen. **3** (1958), 201–205. (Russian) MR **20** #4882.

19. H. H. Schaefer, *Topological vector spaces*, Macmillan, New York, 1966. MR **33** #1689.

20. Laurent Schwartz, *Extension du théorème de Sazonov-Minlos à des cas non hilbertiens*, C.R. Acad. Sci. Paris Sér. A–B **265** (1967), A832–A834. MR **37** #2275.

21. ———, *Réciproque du théorème de Sazonov-Minlos dans des cas non hilbertiens*, C. R. Acad. Sci. Paris Sér. A–B **266** (1968), A7–A9. MR **37** #2276.

22. N. Vakhania, *Sur une propriété des répartitions normales de probabilités dans les espaces l_p ($1 \leq p < \infty$) et H*, C. R. Acad. Sci. Paris **260** (1965), 1334–1336. MR **30** #4282.

23. ———, *Sur les répartitions de probabilités dans les espaces de suites numériques*, C. R. Acad. Sci. Paris **260** (1965), 1560–1562. MR **30** #4283.

24. V. S. Varadarajan, *Convergence of stochastic processes*, Bull. Amer. Math. Soc. **67** (1961), 276–280. MR **23** #A2924.

25. S. R. S. Varadhan, *Limit theorems for sums of independent random variables with values in a Hilbert space*, Sankhya **24** (1962), 213–238.

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA 94720