VIDYADHAR MANDREKAR

Central limit problem and invariance principles on Banach spaces

Séminaire de probabilités (Strasbourg), tome 17 (1983), p. 425-497

<http://www.numdam.org/item?id=SPS_1983__17__425_0>

CENTRAL LIMIT PROBLEM AND INVARIANCE PRINCIPLES ON BANACH SPACES

V. MANDREKAR

0. INTRODUCTION. These notes are based on eight lectures given at the University of Strasbourg. The first three sections deal with the Central Limit Problem. The approach taken here is more along the methods developped by Joel Zinn and myself and distinct from the development in the recent book of Araujo and Giné (Wiley, New York, 1980). The first Section uses only the finite dimensional methods. In the second Section we use Le Cam's Theorem, combined with the ideas of Feller to derive an approximation theorem for a convergent triangular array. This includes the theorem of Pisier in CLT case. As the major interest here is to show the relation of the classical conditions to the geometry of Banach spaces (done in Section 3), we restrict ourselves to symmetric case. Also in this case, the techniques being simple, I feel that the material of the first three Sections should be accessible to graduate students.

In section 4, we present de Acosta's Invariance Principle with the recent proof by Dehling, Dobrowski, Philipp. In the last section we present Dudley and Dudley-Philipp work. I thank these authors for providing me the preprints. I thank Walter Philipp for enlightenning discussions on the subject.

As for the references the books by Parthasarathy and Billingsley are necessary references for understanding the main theme and the basic techniques. To understand the classical problem, one needs the books by Loéve and Feller, where Central Limit Problem is defined. Other needed references are embodied in the text. Remaining references are concerned with Sections 4 and 5 . For those interested in the complete bibliography, it can be foúnd in the book of Araujo-Giné.

I want to thank Professor X. Fernique for inviting me to present the course and the participants of the course for their patience and interest. Further, I want to thank M. Fernique and M. Heinkel for their hospitality and help during my stay, as well as discussions on the subject matter of the notes. I also would like to thank M. Ledoux for interesting discussions.

Finally, I express my gratitude to my wife Veena who patiently gave me a lot of time to devote to these notes.

1. PRELIMINARY RESULTS AND STOCHASTIC BOUNDEDNESS.

Let us denote by B a separable Banach space with $\| \ \|$ and (topological) dual B'. Let $(\Omega,\mathfrak{F},P)$ be a probability space and $\mathcal{B}(B)$ be the Borel sets of B. A measurable function on $(\Omega,\mathfrak{F}) \longrightarrow (B,\mathcal{B}(B))$ will be called a random variable (r.v.). We call its distribution $P \circ X^{-1}$ the law of X and denote it by $\mathcal{L}(X)$.

A sequence $\{\mu_n\}$ of finite measures on $(B,\mathcal{B}(B))$ is said to converge weakly to a finite measure μ on $(B,\mathcal{B}(B))$ if $\int f d\mu_n \longrightarrow \int f d\mu$ for all bounded continuous functions f on B. It is said to be relatively compact if the closure of $\{\mu_n\}$ is compact in the topology of weak convergence. By Prohorov Theorem, we get that a sequence $\{\mu_n\}$ of finite measures is relatively compact iff for $\varepsilon > 0$, there exists a compact subset K_ε of B such that $\mu_n(K_\varepsilon^c) < \varepsilon$, for all n and $\sup_n \mu_n(B) < \infty$. A sequence satisfying this condition will be called tight.

With every finite measure F on B we associate a probability measure $e(F)$ (the exponential of F) by

$$e(F) = \exp(-F(B)) \left\{ \sum_{n=0}^{\infty} \frac{F^{*n}}{n!} \right\}.$$

where F^{*n} denotes the n-fold convolution of F and $F^{*0} = \delta_0$, the probability measure degenerate at zero.

Remark : Note that the set of all finite (signed) measures form a Banach algebra under the total variation norm and multiplication given by the convolution. $F*G(A) = \int_B F(A-x)\, G(dx)$; thus the exponential is well-defined and the convergence of the series is in the total variation norm.

With every cylindrical (probability) measure we associate (uniquely) its characteristic function (c.f.) $\varphi_\mu(y) = \int \exp(i\langle y,x\rangle)\, d\mu$ for $y \in B'$. Here $\langle\ \rangle$ denotes the duality map on (B',B). We note that φ_μ determines μ uniquely on cylinder sets and hence, if μ is a probability measure, then

φ_μ determines μ uniquely on $\mathcal{B}(B)$, as B is separable. It is easy to check that for $y \in B'$.

$$\varphi_{e(F)}(y) = \exp[\int(\exp(i<y,x>)-1)dF]$$

for a finite measure. From this, one easily gets

1) $e(F_1+F_2) = e(F_1) * e(F_2)$ and in particular $e(F) = e(F/n)^{*n}$.

2) $e(F) = e(G)$ iff $F = G$ and $e(c\,\delta_0) = \delta_0$ for $c > 0$.

Furthemore, if $\{F_n\}$ is tight then $\{e(F_n)\}$ is tight, as

$$e(F_n) = \exp(-F_n(B))[\sum_{k=0}^{r} F_n^{*k}/k! + \sum_{k=r+1}^{\infty} F_n^{*k}/k!].$$

For $\varepsilon > 0$, choose r large to make the variation

$$\|e(F_n) - \exp(F_n)\sum_{k=0}^{r} F_n^{*k}/k\|_V < \varepsilon$$

and note that under the hypothesis $\{F_n^{*k}\}$ tight for each k. We also observe that F_n converges weakly to F implies $e(F_n)$ converges weakly to $e(F)$ for F_n and F finite measures. This we get as $\varphi_{e(F_n)}(y) \longrightarrow \varphi_{e(F)}(y)$ in view of the following theorem. (See for example, Parthasarathy, p. 153).

1.1. THEOREM. Let $\{\mu_n\}$ and μ be probability measures on B such that $\{\mu_n\}$ is tight and $\varphi_{\mu_n}(y) \longrightarrow \varphi_\mu(y)$ for $y \in B'$ then μ_n converges weakly to μ (in notation, $\mu_n \Rightarrow \mu$).

Let us consider how Poisson theorem results from this. Let $\{X_{n1},\ldots,X_{nn}\}$ be i.i.d. Bernoulli r.v.'s., $P\{X_{n1} = 1\} = 1-P\{X_{n1} = 0\} = p_n$. Then

$$e(\sum_{j=1}^{n}\mathcal{L}(X_{nj})) = e(np_n\,\delta_1 + n(1-p_n)\delta_0) = e(np_n\delta_1) * e(n(1-p_n)\delta_0)$$

$$= e(np_n\,\delta_1).$$

Hence as $np_n \to \lambda$, $e(np_n\,\delta_1) \Rightarrow e(\lambda\,\delta_1)$ = Poisson with parameter λ . As $p_n \to 0$, one can easily check that

$$\lim_n \left|\varphi_{\mathcal{L}(\sum_{j=1}^{n} X_{nj})}(y) - \varphi_{e(\sum_{j=1}^{n} \mathcal{L}(X_{nj}))}(y)\right| = 0 \quad \text{for } y \in R .$$

Thus associating $\lim_n \mathcal{L}(\sum_{j=1}^{n} X_{nj})$ the $\lim_n e(\sum_{j=1} \mathcal{L}(X_{nj}))$ is called the principle of Poissonization. Note that in this case the limit is $e(F)$, F finite.

We need some facts on weak convergence and convolution. We associate with every finite measure F a measure $\bar{F}(A) = F(-A)$, $A \in \mathcal{B}(B)$ and say that F is symmetric if $\bar{F} = F$.

1.2. THEOREM. (Parthasarathy, p. 58). Let G be a complete separable metric abelian group and $\{\lambda_n\}$, $\{\mu_n\}$, $\{\nu_n\}$ be sequences of probability measures such that $\lambda_n = \mu_n * \nu_n$ for each n .

a) If $\{\mu_n\}$ and $\{\nu_n\}$ are tight then so is $\{\lambda_n\}$.

b) If λ_n is tight then there exists $x_n \in G$ such that $\{\mu_n * \delta_{x_n}\}$ and $\{\nu_n * \delta_{-x_n}\}$ are tight. Further, if $\{\lambda_n\}$, $\{\mu_n\}$, $\{\nu_n\}$ are symmetric, then the tightness of $\{\lambda_n\}$ is equivalent to that of $\{\mu_n\}$ and $\{\nu_n\}$.

Let $q : B \to [0,\infty]$ be a measurable function satisfying $q(x+y) \leqslant q(x) +$ $q(x+y) \leqslant q(x) + q(y)$ and $q(\lambda x) = |\lambda|\, q(x)$. Then q is called a measurable seminorm. An example of such a measurable seminorm we shall use, is the Minkowski functional of a symmetric convex, compact set K in B defined by

$$q_K(x) = \inf\{\alpha \; ; \; \alpha > 0 \; , \; \alpha^{-1} x \in K\} .$$

1.3. THEOREM. (Lévy inequality). Let $\{X_j, \; j = 1,2,\ldots,n\}$ be independant, symmetric, random variables with values in B and $S_k = \sum_{j \leqslant k} X_j$ for $k = 1, 2,\ldots,n$, $S_0 = 0$. Then for each $t \geqslant 0$

$$P\{\sup_{k\leq n} q(S_K) > t\} \leq 2P(q(S_n) > t)$$

for any measurable seminorm q.

<u>Proof</u> : Let $E_k = \{q(S_j) \leq t,\ j = 1,2,\ldots,k-1, q(S_k) > t\}$ for $k = 1,2,\ldots,n$. Then with $E = \{\sup_{k\leq n} q(S_k) > t\}$ we have $E = \bigcup_k E_k$ and E_k are disjoint. Let $T_k = 2S_k - S_n$, then

$$\{q(S_n) \leq t\} \cap \{q(T_k) \leq t\} \subseteq \{q(S_k) \leq t\}$$

and hence using $E_k \subseteq \{q(S_k) > t\}$, we get

$$E_k = [E_k \cap \{q(S_n) > t\}] \cup [E_k \cap \{q(T_k) > t] .$$

Now set

$$Y_j = X_j \quad j \leq k \quad \text{and} \quad Y_j = -X_j \quad \text{for} \quad j > k ,$$

then by the symmetry and independence

$$\mathcal{L}(X_1,\ldots,X_n) = \mathcal{L}(Y_1,\ldots,Y_n)$$

giving $P(E_k \cap \{q(T_k) > t\}) = P(E_k \cap \{q(S_n) > t\})$ i.e. $P(E_k) \leq 2P(E_k \cap \{q(S_n) > t)$. Summing over k we get the result.

1.4. THEOREM. (Feller inequality). <u>Let</u> $\{X_j,\ j = 1,2,\ldots,n\}$ <u>be independent symmetric B-valued r.v.'s. with</u> $S_n = \sum_{j=1}^{n} X_j$, <u>then for</u> $t > 0$

$$1 - \exp(-\sum_{j=1}^{n} P(q(X_j) > t)) \leq P(q(S_n) > t/2) .$$

Further, for $t > 0$, such that $P(q(S_n) > t/2) < 1/2$

$$\sum_{j=1}^{n} P(q(X_j) > t) \leq -\log[1-2P(q(S_n) > 1/2)]$$

for a mesurable seminorm q on B.

<u>Proof</u> : Since $X_j = \sum_{k=1}^{j} X_k - \sum_{k=1}^{j-1} X_k$ we get $q(X_j) \leq q(\sum_{k=1}^{j} X_k) + q(\sum_{k=1}^{j-1} X_k)$

and hence

$$P(\max_{1\leq j\leq n} q(X_j) > t) \leq P(\max_{1\leq j\leq n} q(\sum_1^j X_k) > \frac{1}{2} t) .$$

But left hand side equals $1 - \prod_{j=1}^{n} (1 - P(q(X_j) > t)$ by independence.

As $1 - x \leq \exp(-x)$, $1 - P(q(X_j) > t) \leq \exp[-P(q(X_j) > t)]$ giving

$$1 - \exp(-\sum_{j=1}^{n} P(q(X_j) > t)) \leq 1 - \prod_{j=1}^{n} [1-P(q(X_j) > t)]$$

$$\leq P(\max_{1\leq j\leq n} q(\sum_1^j X_k) > t/2) .$$

Using theorem 1.3, we get the first inequality. The second follows immediately from the first.

1.5. LEMMA : (Truncation). Let $X_1, X_2, \ldots, X_n$ be independent symmetric r.v.'s. Let $a_j > 0$ for $j = 1,2,\ldots,n$ and define $X_j' = X_j\, 1(\|X_j\| \leq a_j)$. Let q be a measurable seminorm on B and set $S_n = \sum_{j=1}^{n} X_j$ and $S_n' = \sum_{j=1}^{n} X_j'$.

Then for $t > 0$, $P(q(S_n') > t) \leq 2P(q(S_n) > t)$.

Proof : Define $Y_j' = X_j - X_j'$ then $X_j' + Y_j'$ and $X_j' - Y_j'$ have the same distribution as X_j. Let

$$\tilde{S}_n = \sum_{j=1}^{n} Y_j' \quad \text{then} \quad \{q(S_n') > t\} = \{q(S_n' + \tilde{S}_n + S_n' - \tilde{S}_n) > 2t\}$$

$$\subseteq \{q(S_n' + \tilde{S}_n) > t\} \cup \{q(S_n' - \tilde{S}_n) > t\}$$

$$\mathcal{L}(S_n' + \tilde{S}_n) = \mathcal{L}(S_n' - \tilde{S}_n) = \mathcal{L}(S_n)$$

$$P(q(S_n') > t) \leq 2P(q(S_n) > t) .$$

We say that a sequence $\{Y_k\}$ of real valued r.v.'s. is stochastically bounded if for every $\varepsilon > 0$, there exists t finite so that $\sup_n P(\|Y_n\| > t) < \varepsilon$.

1.6. THEOREM. (Hoffman-Jørgensen). Let $\{X_i, i=1,2,\ldots\}$ be independent, symmetric, B-valued r.v.'s. with $q(X_i)$ in $Lp(\Omega,\mathfrak{F},P)$ for some p and a measurable seminorm q. Then $\{q(S_n)\}$ is stochastically bounded and $E\sup_j|q(X_j)|^p<\infty$ implies

$$\sup_n E\left|q\left(\sum_{j=1}^n X_j\right)\right|^p \leqslant 2.3.^p\, E\sup_i[q(X_i)]^p + 16.3^p\, t_o^p$$

where $t_o=\inf\{t>0\ ;\ \sup_n P(q(\sum_{j=1}^n X_j)^p>t)<\frac{1}{8.3}\,p\}$.

Proof : By theorem 1.4., (more precisely, its proof) we get that under the hypothesis, $\sup_n q(S_n)$ is finite a.e. and $\sup_i q(X_i)\leqslant 2\sup_n q(S_n)$. For $t,s>0$, we prove

$$(1.6.1)\quad (P(q(S_k)>2t+s)\leqslant P(\sup_n q(S_n)>t)+4[P(q(S_k)>t)]^2$$

$T=\inf\{n\geqslant 1\ ;\ q(S_n)\geqslant t\}$ where $T=\infty$ if the set is $\emptyset$. Now $q(S_k)\geqslant 2t+s$ implies $T\leqslant k$ giving $P(q(S_k)>2t+s)=\sum_{j=1}^k P(q(S_k)>2t+s,\ T=j)$. If $T=j$, then $q(S_{j-1})<t$ and hence for $T=j$ and

$$\begin{aligned}q(S_k)\geqslant 2t+s\ ,\ q(S_k-S_j)&\geqslant q(S_k)-q(S_{j-1})-q(X_j)\\&\geqslant 2t+s-t-\sup_j q(X_j)=t+s-N\end{aligned}$$

$$\begin{aligned}P(T=j,\ q(S_k)\geqslant 2t+s)&\leqslant P(T=j,\ q(S_k)\geqslant t+s-N)\\&\leqslant P(T=j,\ N\geqslant s)+P(T=j,\ q(S_k-S_j)\geqslant t).\end{aligned}$$

By independence of $T=j$ and S_k-S_j we get summing over $j\leqslant k$

$$P(q(S_k)>2t+s)\leqslant P(N\geqslant s)+\sum_{j=1}^k P(T=j)\,P(q(S_k-S_j)\geqslant t).$$

Now $Y_1=S_k-S_j$ and $Y_2=S_j$ then Y_1,Y_2 are symmetric independent and hence by Lévy inequality

$$P(q(Y_1)\geqslant t)\leqslant P(\max(q(Y_1),q(Y_1+Y_2))\geqslant t)\leqslant 2P(q(Y_1+Y_2)\geqslant t).$$

This proves (1.6.1). Since $\{q(S_k)\}$ is stochastically bounded

$$P(q(S_k) > t) \leq P(\max_j q(X_j) > t) \leq 2 \sup_k P(q(X_k) > t) .$$

Hence

$$P(\sup_k q(S_k) > 2t + s) \leq P(\max_j q(X_j) > s) + 8[P(\sup_k q(S_k) > t)]^2$$

i.e. $\quad R(2t + s) \leq Q(s) + 8R(t)^2 \quad$ (say).

Choose t_o as in the theorem and observe that for $a > 3t_0$

$$\begin{aligned}\int_0^a px^{p-1} R(x)dx &= 3^p p \int_0^{a/3} x^p R(3x)dx \leq 3^p p.2 \int_0^{a/3} x^p Q(x)dx \\ &\quad + 8p3^p \int_0^{a/3} x^{p-1} R^2(x)dx \\ &\leq 2.3.^p EN^p + 8.3.^p t_0^p p + 8p3^p \int_0^{a/3} x^{p-1} R(t_0)R(x)dx \\ &\leq C + \frac{1}{2}\int_0^a px^{p-1} R(x)dx .\end{aligned}$$

where $C = 2.3^p EN^p + 8.3^p t_0^p$. This gives the resoult.

Let $\{X_{nj}, j = 1,2,\ldots,k_n\}$ $n = 1,2,\ldots$ ($k_n \to \infty$ as $n \to \infty$) be a row independent triangular array of symmetric B-valued random variables. In these lectures, we shall consider only these triangular arrays and refer to them as triangular array, unless otherwise stated. For each $c > 0$, let

$$X_{njc} = X_{nj}\, 1(\|X_{nj}\| \leq c) \quad , \quad \tilde{X}_{njc} = X_{nj} - X_{njc} \quad ;$$

$$S_{nc} = \sum_{j=1}^{k_n} X_{njc} \quad , \quad S_n = \sum_{j=1}^{k_n} X_{nj} \quad , \quad \tilde{S}_{nc} = S_n - S_{nc} .$$

We shall denote by $F_n = \sum_{j=1}^{k_n} \mathcal{L}(X_{nj})$, $0_t = \{x \in B, \|x\| \leq t\}$.

The following is an extension of Feller's theorem.

1.7. THEOREM. Let $\{X_{nj}, j = 1,2,\ldots,k_n\}$ $n = 1,2,\ldots$ be a triangular array. Then $\{\|S_n\|\}$ is stochastically bounded iff

a) <u>For every</u> $\varepsilon > 0$, <u>there exists</u> t <u>large, so that</u> $\sup_n F_n(O_t^c) < \varepsilon$

b) <u>For every</u> $c > 0$, $\sup_n E\|S_n(c)\|^p < \infty$.

<u>Proof</u> : Put $q(x) = \|x\|$ in theorem 1.4., then we get condition a) .

By stochastic boundedness of $\|S_n\|$. Condition (b) follows from Lemma 1.5. and theorem 1.6. To prove the converse for $t > 0$

$$P(\|S_n\|) > 2t) \leqslant P(\|S_{nc}\| > t) + P(\|\widetilde{S}_{nc}\| > t) . \quad \text{Now}$$

$$\widetilde{S}_{nc} = \sum_{j=1}^{k_n} X_{nj}\, 1(\|X_{nj}\| > c) \quad \text{so} \quad \{\|\widetilde{S}_{nc}\| > t\} \subseteq \{\max_j \|X_{nj}\| > c\} .$$

Thus by Chebychev's inequality we get

$$P(\|S_n\| > 2t) \leqslant \frac{1}{t^p} E\|S_{nc}\|^p + \sum_{j=1}^{k_n} P(\|X_{nj}\| > c) .$$

Given $\varepsilon > 0$, choose c_o so that $F_n(O_{c_o}^c) < \varepsilon/2$ and then choose t_o so that $\frac{1}{t_o^p} \sup_n E\|S_{nc_o}\|^p < \varepsilon/2$.

We now derive some consequences of the above result in special cases.

1.8. <u>Special Examples.</u>

1.8.1. Example $B = L_p$, $p \geqslant 2$ and $X_{nj} = X_j/\sqrt{n}$, $\{X_j,\ j = 1,2,..\}$ i.i.d. sequence of L_p-valued r.v.'s. Before we study this example we need some general facts : We define $\Lambda(X) = \sup_{t>0} t^2 P(\|X\| > t)$.

<u>Rosenthal inequality</u>. Let $2 \leqslant p < \infty$, then there exists $c_p < \infty$ so that for any sequence $\{X_j,\ j = 1,2,\ldots,n\}$ of independent real-valued random variables with $E|X_j|^p < \infty$ and $EX_j = 0$ $(j = 1,2,\ldots,n)$ we have for all $n \geqslant 1$

$$\frac{1}{2} \max \{(\sum_{j=1}^{n} E|X_j|^p)^{1/p} , (\sum_{j=1}^{n} E|X_j|^2)^{1/2}\}$$

$$\leqslant (E|\sum_{j=1}^{n} X_j|^p)^{1/p} \leqslant C_p \max \{(\sum_{j=1}^{n} E|X_j|^p)^{1/p}, (\sum_{j=1}^{n} E|X_j|^2)^{1/2}\} .$$

We also observe that for a B-valued r.v. X $n \geqslant 1$, $\delta > 0$, $2 < p < \infty$

(*) $$n E \left\| \frac{X}{\sqrt{n}} 1(\|X\| \leqslant C\sqrt{n}) \right\|^p < \frac{p}{p-2} C^{p-2} \sup_{u>0} u^2 P(\|X\| > u)$$

To see this

$$E\|X\|^p 1(\|X\| \leqslant C\sqrt{n}) \leqslant \int_0^{(C\sqrt{n})^p} P(\|X\|) > u^{1/p}) du$$

$$\leqslant \int_0^{(C\sqrt{n})^p} \Lambda^2(X)/u^{2/p} du .$$

Evaluating the integral we get (*) . In this case, we observe that $F_n(O_t^c) = n P(\|X\| > \sqrt{n}\, t)$. Now if $\Lambda^2(X) < \infty$ then

$$n P(\|X\| > t\sqrt{n}) = \frac{t^2 n P(\|X\| > t\sqrt{n})}{t^2} \leqslant \frac{\Lambda^2(X)}{t^2} .$$

Given $\varepsilon > 0$, there exists t_o , so that

$$F_n(O_{t_o}^c) < \varepsilon \text{ for all } n .$$

Conservely, if such a t_o exists then $\sup_n t_o^2 n P(\|X\| > t_o\sqrt{n}) < M$ giving $\Lambda^2(X) < \infty$. Thus condition (b) of theoreme 1.7. is satisfied iff $\Lambda^2(X) < \infty$. Thus $\{\|X_1 + \dots + X_n/\sqrt{n}\|\}$ is stochastically bounded iff $\Lambda^2(X) < \infty$ and $\sup_n E \int | \sum_{j=1}^n X_j/\sqrt{n}\, 1(\|X_j\| \leqslant C\sqrt{n})(u)|^p d\mu < \infty$.

By Rosenthal's inequality the second condition is equivalent to

$$\sup_n \sum_{j=1}^n E\int |X_j/\sqrt{n}\, 1(\|X_j\| \leqslant C\sqrt{n})(u)|^p d\mu < \infty \quad \text{and}$$

$$\sup_n \sum_{j=1}^n (\int (E(X_j 1(\|X_j\| \leqslant C\sqrt{n})/\sqrt{n})^2(u))^{p/2} d\mu < \infty$$

Here one chooses a jointly measurable version of $(X_j(u))$. The first term finite by the observation (*) and the second is finite by the monotone convergence iff

$\int (E(X_1(u))^2)^{p/2} \, d\mu < \infty$. Thus $\{\|X_1 + \dots + X_n/\sqrt{n}\|\}$ is stochastically bounded iff $\Lambda^2(X_1) < \infty$ and $\int (E\, X_1(u)^2)^{p/2} \, d\mu < \infty$.

1.8.2. Example : $B = H$ a separable Hilbert space. Let $\{e_k, k = 1,2,\dots\}$ be a a complete orthonormal basis in H. $X_{nj} = X_j/\sqrt{n}$, $\{X_j\}$ i.i.d. Then $\{X_1 + \dots$ $\dots + X_n/\sqrt{n}\}$ stochastically bounded, implies condition (b) of theorem 1.7. with $p = 2$ i.e.

$$\sup_n E\|\sum_{j=1}^{n} X_j/\sqrt{n}\; 1(\|X_j\| \leq C\sqrt{n})\|^2 < \infty . \quad \text{But this implies}$$

$$\sup_n E\|X_1 \; 1(\|X_1\| \leq C\sqrt{n})\|^2 = E\|X_1\|^2 < \infty .$$

From this (a) follows. Let π_k = Projection onto $\overline{sp}\{e_1,\dots,e_k\}$. Then by Chebychev inequality for $\varepsilon > 0$

$$P\{\|\frac{X_1 + \dots + X_n}{\sqrt{n}} - \pi_k(\frac{X_1 + \dots + X_n}{\sqrt{n}})\| > \varepsilon\}$$

$$\leq \frac{1}{\varepsilon^2} E\|X_1 - \pi_k(X_1)\|^2 < \varepsilon \quad \text{for } k \text{ large as } \; E\|X_1\|^2 < \infty .$$

Hence we get $\{X_1 + \dots + X_n/\sqrt{n}\}$ is flatly concentrated and, by one-dimensional central limit theorem, we get that $\mathcal{L}(X_1 + \dots + X_n/\sqrt{n}) \Rightarrow V$ where V is a Gaussian measure with covariance $E<y,X_1><y',X_1>$ for $y,y' \in H'$. We thus have the equivalence of :

i) Central Limit Theorem (CLT) holds in H for $\mathcal{L}(X_1)$.

ii) $E\|X_1\|^2 < \infty$ and (iii) $\{X_1 + \dots + X_n/\sqrt{n}\}$ is stochastically bounded.

1.8.3. Example : $(B = \mathbb{R}^k, k < \infty)$. Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ be row independent triangular array of (symmetric) R^k-valued r.v.'s satisfying for every $\varepsilon > 0$

$$(*) \qquad \max_{1 \leq j \leq k_n} P\{\|X_{nj}\| > \varepsilon\} \to 0$$

and assume that $\{S_n\}$ is stochastically bounded. Let for $y \in B'$, $\|y\|$ denote the strong norm on B' and $M < \infty$.

$$\sup_n \sup_{\|y\| \leq M} \sum_{j=1}^{k_n} |\varphi_{\mathcal{L}(X_{nj})}(y) - 1|$$

$$\leq \sup_n \sup_{\|y\| \leq M} \{\int_{\|x\| \leq c} (1-\cos \langle y,x\rangle) F_n(dx) + 2F_n(\|x\| > c) .$$

Now choose c_o so that the second term is $< \varepsilon/2$. Use on the first term inequalities,

$$(1-\cos \langle y,x\rangle) \leq \langle y,x\rangle^2 \leq \|y\|^2 \, \|x\|^2$$

to conclude that it does not exceed $M^2 \sup_n \int_{\|x\| \leq c_o} \|x\|^2 F_n(dx)$ which is finite by condition (b) of theorem 7.1. Hence for n large $\log \varphi_{nj}(y)$ exists where $\varphi_{nj}(y) = \varphi_{\mathcal{L}(X_{nj})}(y)$. Now

$$\sup_{\|y\| \leq M} |\log \prod_{j=1}^{k_n} \varphi_{nj}(y) - \log \varphi_{e(F_n)}(y)|$$

$$\leq \sup_{\|y\| \leq M} \sum_{j=1}^{k_n} |\log \varphi_{nj}(y) - \varphi_{nj}(y) + 1| \leq \underline{\text{Constant}} \ \sup \sum_{j=1}^{k_n} |\varphi_{nj}(y) - 1|^2$$

$$\leq \text{constant} \max_{1 \leq j \leq k_n} |\varphi_{nj}(y) - 1| \sup_{\|y\| \leq M} \sum_{j=1}^{k_n} (\varphi_{nj}(y) - 1) \to 0 \text{ by } (*) .$$

One can derive easily the following from above,

a) $\{S_n\}$ is stochastically bounded in $\mathbb{R}^k$ iff for some $c > 0$ (and hence for every) the finite measures defines by $\nu_n(A) = \int_A \min(c, \|x\|^2) F_n(dx)$, $A \in \mathcal{B}(\mathbb{R}^k)$ form a tight sequence.

b) For $B = \mathbb{R}^k$, the following are equivalent under $(*)$.

i) $\{S_n\}$ is otochastically bounded.

ii) $\{e(F_n)\}$ is otochastically bounded.

iii) For each $c > 0$, $\{\nu_n\}$ is tight.

c) Every limit law of $\{S_n\}$ satisfying (*) is infinitely divisible and conversely.

We note that condition 1.8.3. (*) is valid in general B. We define now infinitely divisible law.

1.9. DEFINITION. A probability measure μ on B is called infinitely divisible (i.d.) if for each integer n, there exists a probability measure μ_n on B such that $\mu = \mu_n^{*n}$.

We now prove converse part of 1.8.3.(c) in general. Let μ be i.d. and $\{X_{nj},\ j = 1,2,\dots,k_n\}$ be a row independent triangular array with $\mathcal{L}(X_{nj}) = \mu_n$ (this may not be symmetric unless μ is,in latter case, μ_n can be chosen so). Then $\mu = \lim_n \mathcal{L}(S_n)$. But

$$\varphi_{\mu_n}(y) = [\varphi_\mu(y)]^{\frac{1}{n}}. \quad \text{Hence} \quad \max_{1\leq j\leq n} |\varphi_{\mu_n}(y) - 1| \to 0 \quad \text{i.e.}$$

$\{X_{nj},\ j = 1,2,\dots,k_n\}$ satisfy 1.8.3.(*). We refer to this as the triangular array being uniformly infinitesimal (U.I.).

In view of theorem 1.2, symmetric i.d. laws are closed under weak limits. Hence we get $\lim_n e(F_n)$ is i.d. But under (*),

$\lim_n e(F_n) = \lim_n \mathcal{L}(\sum_{j=1}^{k_n} X_{nj})$, giving c) above for $B = \mathbb{R}^k$. This proof fails in general B. However 1.8.3. c) survives. To see this, denote for $T = \{y_1,\dots,y_k\} \subseteq B'$, $y_T(x) = (\langle y,x\rangle,\dots,\langle y_k,x\rangle)$ for $x \in B$.

1.10. LEMMA. Let μ be a symmetric probability measure on $\mathcal{B}(B)$. Then μ is i.d. iff $\mu \circ y_T^{-1}$ is i.d. for all finite subsets $T \subseteq B'$.

Proof : The "only if " part is obvious. For the other part, under the assumption, $\mu \circ y_T^{-1} = [\mu_n(T)]^{*n}$ for each n and T finite subset of B'.

Since $\varphi_{\mu o y_T^{-1}}(u) \neq 0$ for $u \in \mathbb{R}^k$, we get that $\{\mu_n(T), T$ finite subset of $B'\}$ is a cylinder measure μ_n satisfying for each y,

$$\varphi_\mu(y) = [\varphi_{\mu_n}(y)]^n \ .$$

Hence by theorem 1.2. (c), we get μ_n is a probability measure on $\mathcal{B}(B)$ i.e. μ is i.d.

Combining this with 1.8.3. c) we get

1.11. THEOREM. <u>The symmetric i.d. laws on</u> B <u>coincide with the limit laws of row sums of UI row-independent, symmetric triangular arrays.</u>

We note that by Lemma 1.5., $\{S_n\}$ is tight iff $\{S_{nc}\}$ and $\{\widetilde{S}_{nc}\}$ are tight. Hence for U.I. triangular arrays $\lim_n \mathcal{L}(<y,S_n>) = \lim_n e(F_n \circ y^{-1}) = \lim_n e(F_{nc} \circ y^{-1}) * e(\widetilde{F}_{nc} \circ y^{-1})$ with $F_{nc} = \sum_{j=1}^{k_n} \mathcal{L}(X_{njc})$ and $\widetilde{F}_{nc} = \sum_{j=1}^{k_n} \mathcal{L}(\widetilde{X}_{njc})$.

Thus $\lim_n \mathcal{L}(<y,S_n>) = \lim_n \mathcal{L}(S_{nc}) * \mathcal{L}(\widetilde{S}_{nc})$ at least for $B = \mathbb{R}^k$. In fact it is true in general.

1.12. THEOREM. <u>Let</u> $\{X_{nj}, j = 1,2,\dots,k_n\}$ <u>be U.I. triangular array such that</u> $\mathcal{L}(S_{nc}) \Rightarrow \mu$ <u>and</u> $\mathcal{L}(\widetilde{S}_{nc}) \Rightarrow \nu$. <u>Then</u> $(\mathcal{L}(S_{nc}), \mathcal{L}(\widetilde{S}_{nc})) \Rightarrow \mu \otimes \nu$.

<u>Proof</u> : is by the use of c.f.s and is left to the reader.

We can observe that all methods used so far are finite-dimensional. In the next chapter we bring out the methods particular to the infinite dimensional case.

2. CENTRAL LIMIT PROBLEM IN BANACH SPACES.

Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ be a (symmetric) row-independent triangular array of B-valued random variables as before for $n = 1,2,\dots$

$$S_n = \sum_{j=1}^{k_n} X_{nj} \quad \text{and} \quad F_n = \sum_{j=1}^{k_n} \mathcal{L}(X_{nj}) \quad .$$

2.1. THEOREM. (Le Cam). <u>Let</u> $\{\mathcal{L}(S_n)\}$ <u>be tight. Then for every</u> $t > 0$, <u>there exists a compact, convex symmetric set</u> $K_t \subseteq O_t$ <u>such that</u> $\{F_n | K_t^c\}$ <u>is tight.</u> <u>In particular</u> $F_n | O_t^c$ <u>is tight.</u>

<u>Proof</u> : Use theorem 1.4., with q the Minkowski functional of symmetric, compact, convex set $\widetilde{K}_\delta$, given from compactness of $\{\mathcal{L}(S_n)\}$, to get

$$(2.1.1) \quad \sup_n \sum_{j=1}^{k_n} P(X_{nj} \notin \widetilde{K}_\delta) < \delta \quad .$$

Let $K_t = \widetilde{K}_\delta \cap O_t$ (with δ fixed). We claim that

$$\sup_n \sum_j P(X_{nj} \in K_t) < M < \infty \quad .$$

As $\widetilde{K}_\delta \subseteq O_r$ and $P(X_{nj} \notin K_t) = P(X_{njr} \notin K_t) + P(\|X_{nj}\| > r)$ we assume that $\|X_{nj}\| \leqslant r$ a.s. Let

$$V_y = \{x \in B \; ; \; |\langle y,x\rangle| > t/2\} \quad .$$

Then $\{V_y, \|y\| \leqslant 1\}$ is a cover of $\overline{O_t^c} \cap \widetilde{K}_\delta$ and, hence by compactness there exists a finite cover $\{V_{y_1},\dots,V_{y_m}\}$. By theorem 1.7., $\sup_n E\langle y_j,S_n\rangle^2 < \infty$, $j = 1,2,\dots,m$. Hence,

$$\sum_j P(X_{nj} \notin K_t) \leqslant 2 \sum_j P(X_{nj} \notin \widetilde{K}_\delta) + \sum_j P(X_{nj} \in \overline{O_t^c} \cap \widetilde{K}_\delta) \quad .$$

The second term does not exceed $\Sigma_j \Sigma_i P(|\langle y_i,X_{nj}\rangle| > t/2)$. Using (2.1.1.) and Chebychev inequality we prove the claim. Now define $J_n = \{j \in (1,\dots,k_n) : P(X_{nj} \in K_t) < 3/4\}$ then by the claim $\sup_n \text{card}\,(J_n) \leqslant 4M$. As $\{X_{nj}, j = 1,2..k_n\}$

are tight for each j,n, we get using Lemma 1.5. and properties of K_t that $\{X_{nj}\,1(X_{nj} \notin K_t)\}$ is tight. Thus $\{\sum_{j\in J_n} P(X_{nj}\,1(X_{nj} \notin K_t)\}$ is tight. For $j \in J_n$, take $G = \widetilde{K}_\delta + K_t$, then $G^c \subseteq K_t^c$ since $\widetilde{K}_\delta$ is symmetric convex. For $j \in J_n$, $P(X_{nj} \in K_t) \geqslant 1/4$ and hence

$$\frac{1}{4} \sum_{j\in J_n} P(X_{nj} \notin G) \leqslant \sum_{j\in J_n} P(X_{nj} \notin K_\delta)\, P(X'_{nj} \in K_t)$$

where $\mathcal{L}(X_{nj}) = \mathcal{L}(X'_{nj})$ and they are independent. By (2.1.1.) we get the result.

We can derive the following corollaries :

2.2. COROLLARY. <u>For every</u> $c > 0$, $\{\mathcal{L}(\widetilde{S}_{nc})\}$ <u>tight implies</u> $\{e(\sum_{j=1}^{k_n} \mathcal{L}(\widetilde{X}_{njc})\}$ <u>tight, which gives</u> $\{e(\sum_{j=1}^{k_n} \mathcal{L}(\widetilde{X}_{njc}))\}$ <u>tight.</u>

2.3. COROLLARY. <u>Suppose</u> $\{\mathcal{L}(S_n)\}$ <u>is tight. Then there exists a</u> σ<u>-finite symmetric measure</u> F <u>such that for some subsequence</u> $\{n'\}$ <u>of integers</u> $F_{n'}^{(\varepsilon)} \Rightarrow F^{(\varepsilon)}$ <u>where</u> $F_n^{(\varepsilon)} = F_n|_{0_\varepsilon^c}$ <u>and</u> $F^{(\varepsilon)} = F|_{0_\varepsilon^c}$. <u>Furthermore,</u> $F^{(\varepsilon)}$ <u>is finite for each</u> $\varepsilon > 0$, $\int_{\|x\|\le\varepsilon} \langle y,x\rangle^2\, F(dx) < \infty$ <u>and</u> $F(\{0\}) = 0$.

<u>Proof</u> : By diagonalization procedure and Corollary 2.2., there exists a subsequence $\{n'\}$ such that $F_{n'}^{(\varepsilon_k)}$ converges for all k with $\varepsilon_k \downarrow 0$. Let $F_k = \lim_{n'} F_{n'}^{(\varepsilon_k)}$. Then $F_k(0_{\varepsilon_j}) = 0$ for $j \geqslant k$. Clearly, $F_k \uparrow$ and finite. If If we define $F = \lim_k F_k$; then F is σ-finite, $F^{(\varepsilon)}$ is finite and $F\{0\} = 0$. Since $\{\langle y, S_n\rangle\}$ is tight we get $\sup_n \int \langle y, S_{nr}\rangle^2\, dP < \infty$. This gives for $0 < \varepsilon_k < r$

$$\int_{0_r \cap 0_{\varepsilon_k}} \langle y,\cdot\rangle^2\, dF = \lim_n \int_{0_r \cap 0_{\varepsilon_k}} \langle y,\cdot\rangle^2 dF_n = \lim_n \sum_{j=1}^{k_n} E\langle y,(X_{njr})_{\varepsilon_k}\rangle^2$$

$$\leqslant \sup_n \Sigma_j\, E\langle y,X_{njr}\rangle^2 < \infty$$

Take limit over k to obtain the result.

2.4. COROLLARY. Let $\{\mathcal{L}(S_n)\}$ be tight, $\{X_{nj}\}$ be U.I. and $\lim_n \mathcal{L}(\tilde{S}_{n\varepsilon})$ exists for all $\varepsilon > 0$. Then $e(F^{(\varepsilon)}) = \lim_n \mathcal{L}(\tilde{S}_{n\varepsilon})$ and F is unique.

Proof : Using Corollary 2.2., Theorem 1.1. and arguments as in 1.8.3. we get for any other measure G $e(G^{(\varepsilon)}) \circ y^{-1} = \lim_n \mathcal{L}(\tilde{S}_{n\varepsilon}) \circ y^{-1} = e(F^{(\varepsilon)}) \circ y^{-1}$. Hence $G^{(\varepsilon)} \circ y^{-1} = F^{(\varepsilon)} \circ y^{-1}$ giving $G^{(\varepsilon)} = F^{(\varepsilon)}$ for all $\varepsilon > 0$ i.e., $F = G$.

We call F above as the Lévy measure associated with the i.d. law μ. We denote $\lim_k e(F^{(\varepsilon_k)})$ by $e(F)$ for F Lévy measure.

2.5. THEOREM. Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ be U.I. triangular array such that $\mathcal{L}(S_n) \Rightarrow \nu$. Then

a) There exists a Lévy measure F such that $F_n^{(c)} \Rightarrow F^{(c)}$ for each $c > 0$ and c continuity point of F. $(c \in C(F))$.

b) There exists a Gaussian measure γ with covariance $C_\gamma(y_1,y_2)$ such that for $y \in B'$,

$$(2.5.1)\quad \lim_{c\downarrow 0} \left\{\begin{matrix}\overline{\lim}\\ \underline{\lim}\end{matrix}\right\} \int_{\|x\|\leq c} \langle y,x\rangle^2 dF_n = \lim_{c\downarrow 0}\ \lim_{c\in C(F)} \int_{\|x\|\leq c} \langle y,x\rangle^2 dF_n = C_\gamma(y,y)$$

c) $\nu = e(F) * \gamma$ where F and γ are unique.

Proof : We have proved along a subsequence $\{n'\}$ of $\{n\}$, $F_{n'}^{(c)} \Rightarrow F^{(c)}$ for each $c \in C(F)$, where F is a Lévy measure since $\{\mathcal{L}(S_{n'})\}$ and $\{\mathcal{L}(S_{n'c})\}$ are tight, we can proceeding to the diagonal sequence get a probability measure ν_k such that for $c_k \downarrow 0$,

$$\mathcal{L}(S_{n''}) \Rightarrow \nu \text{ and } \mathcal{L}(S_{n''c_k}) \Rightarrow \nu_k .$$

By theorem 1.12, for each k,

$$\nu = \nu_k * e(F^{(c_k)}).$$

As $e(F^{(c_k)}) \Rightarrow e(F)$, $\{\nu_k\}$ is tight by Theorem 1.2. Since $\varphi_{e(F^{(c_k)})}(y) \neq 0$

for $y \in B'$, $\varphi_{\nu_k}(y) \to \varphi_{\nu_0}(y)$ for some cylinder measure ν_o. But $\nu = \nu_o * e(F)$ gives by Theorem 1.2. that ν_o is a probability measure γ. i.e. $\nu = \gamma * e(F)$. Let us assume that γ is Gaussian. (we shall prove it later). Thus every sequence has a convergent subsequence with limit $\nu = \gamma * e(F)$. We now prove that all limit points have the same Gaussian and non-Gaussian parts. Let $\gamma_1 * e(F_1) = \gamma_2 * e(F_2)$ then $\gamma_1 \circ y^{-1} * e(F_1 \circ y^{-1}) = \gamma_2 \circ y^{-1} * e(F_2 \circ y^{-1})$ giving by the one dimensional result,

$$\gamma_1 \circ y^{-1} = \gamma_2 \circ y^{-1} \quad \text{and} \quad F_1 \circ y^{-1} = F_2 \circ y^{-1} .$$

Thus a) and c) are proved. Let us now observe that $\mathcal{L}(S_{nc}) \Rightarrow \gamma * e(F|0_c)$ and $\{\langle y, S_{nc}\rangle^2\}$ is uniformly integrable in n by Theorem 1.7. Hence

$$\lim_n E\langle y, S_{nc}\rangle^2 = \int \langle y,x\rangle^2 d\gamma + \int_{\|x\| \leq c} \langle y,x\rangle^2 dF .$$

Take limit as $c \in C(F)$ goes to zero then $\int_{\|x\|\leq 1} \langle y,x\rangle^2 dF < \infty$ implies that the second term goes to zero, giving b). It remains to prove γ is Gaussian i.e. $\gamma \circ y^{-1}$ is Gaussian for $y \in B'$. For this we observe that there exists $n_k \uparrow$ such that $\mathcal{L}(S_{n_k c_k}) \Rightarrow \gamma (c_k \downarrow 0)$ by the proof. The following Lemma now completes the proof.

2.6. LEMMA. Let $\{X_{nj}, j = 1,2,\ldots,k_n\}$ $n = 1,2,\ldots$ be a triangular array such that

a) $\max_j \|X_{nj}\| \leq C_n$ a.s. and $C_n \downarrow 0$.

b) $\mathcal{L}(S_n) \Rightarrow \gamma$. Then γ is Gaussian.

Proof : Note as before, $\lim_n E\langle y, S_n\rangle^2 = C_\gamma(y,y)$ by Theorem 1.7. Hence it suffices to prove for $y \in B'$.

$$\Delta_n = E\left|\exp(i\langle y,S_n\rangle) - \exp(-\tfrac{1}{2}\langle y,S_n\rangle^2)\right| \to 0 .$$

But

$$\Delta_n \leq \sum_j |E \exp(i <y,X_{nj}>) - \exp - \frac{1}{2} E <y,X_{nj}>^2|$$

$E \exp i Y = 1 - \frac{1}{2} E Y^2 + E\{\exp i Y - 1 - i Y + \frac{1}{2} Y^2\}$ for Y symmetric and

$$\exp(- \frac{1}{2} EY^2) = 1 - \frac{1}{2} EY^2 + \{\exp(- \frac{1}{2} EY^2) - 1 + \frac{1}{2} EY^2\} .$$

Now use inequalities

$$|e^{it} - 1 - it + \frac{1}{2} t^2| \leq t^3 , \quad |e^x - 1 - x| < x^2 e^x (t,x \text{ real}) \quad \text{to get}$$

$$\Delta_n \leq \sum_j \{E| <y,X_{nj}>|^3 + (E <y,X_{nj}>^2)^2 \exp(\|y\| C_1)^2\}$$

$\to 0$ under the condition established.

2.7. COROLLARY. Every symmetric i.d. law has unique representation $\nu = \gamma * e(F)$ where γ is (centered) Gaussian and F is the Lévy measure.

2.8. COROLLARY. Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ $(n = 1,2,\dots)$ be a triangular array such that $\mathcal{L}(S_n) \Rightarrow \nu$. Then the following are equivalent

a) ν is Gaussian.

b) For every $y \in B'$ and $c > 0$, $\lim_n \sum_{j=1}^{k_n} P(| <y,X_{nj}>| > c) = 0$.

c) For every $c > 0$, $\lim_n F_n^{(c)} = 0$.

2.9. COROLLARY. Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ $(n = 1,2,\dots)$ be a U.I. triangular array such that $\mathcal{L}(S_n) \Rightarrow \nu * e(F)$. Then there exists $c_n \downarrow 0$ such that

$$\mathcal{L}(S_{nc_n}) \Rightarrow \gamma \quad \text{and} \quad \mathcal{L}(\widetilde{S}_{nc_n}) \Rightarrow e(F) .$$

Proof : Let π be the Prohorov metric then we know that $(\pi(\mathcal{L}(\widetilde{S}_{nc}), e(F^{(c)})) \to 0$. Hence there exists $c_n \downarrow 0$ such that $\pi(\mathcal{L}(\widetilde{S}_{nc_n}) , e(F^{(c_n)})) \to 0$. But $\pi(e(F^{(c_n)}),e(F)) \to 0$ giving the first conclusion.

Now $\lim_n \mathcal{L}(S_n) = \lim_n \mathcal{L}(S_{nc_n}) * \lim_n \mathcal{L}(\tilde{S}_{nc_n})$ i.e. $\gamma * e(F) = \lim_n \mathcal{L}(S_{nc_n}) * e(F)$.

Hence $\lim_n \mathcal{L}(S_{nc_n}) = \gamma$.

We note that although theorem 2.5. gives useful necessary conditions, they are far from satisfactory. In the case $X_{nj} = X_j/\sqrt{n}, \{X_j\}$ i.i.d., these conditions are $t^2 P(\|x\| > t) \to 0$ as $t \to \infty$ and X pregaussian. These are sufficient in ℓp , $p \geq 2$ but are not so even in $\ell_2(\ell p)$. Thus one needs to sharpen such a theorem. In the i.i.d. case such sharpening was done by Pisier. We present the following useful theorem in case the limit points are non-Gaussian.

2.10. THEOREM. <u>Let</u> $\{X_{nj}, j = 1,2,\ldots,k_n\}$ $n = 1,2,\ldots$ <u>be a U.I. triangular array.</u> <u>Then</u> $\{\mathcal{L}(S_n)\}$ <u>is tight with all limit points non-Gaussian (i.e.</u> $\nu = e(F)$) <u>iff</u>

a) <u>For each</u> $c > 0$, $\{F_n^{(c)}\}$ <u>is tight</u> ;

b) $\lim_{c\to 0} \sup_n E\|S_{nc}\|^p = 0$ <u>for all</u> p $(0 < p < \infty)$.

<u>Proof</u> : Necessity of a) is proved in theorem 2.1. and by Lemma 1.5., $\{S_{nc}\}_{nc}$ is tight. Further by one-dimensional result

$$\lim_{c\to 0} \sup_n \int_{\|x\|\leq c} \langle y,x\rangle^2 \, dF_n = 0 \quad .$$

Hence by Chebychev's inequality $\langle y, S_{nc}\rangle \xrightarrow[c\to 0]{P} 0$, for all $y \in B'$. Now $\{\mathcal{L}(S_{nc})\}$ is tight gives by theorem 1.1. that $\|S_{nc}\| \xrightarrow{P} 0$ uniformly in n as $c \to 0$. Given $\eta > 0$ choose c_o such that, for $c \leq c_o$,

$$\sup_n P\{\|S_{nc}\| > \tfrac{1}{3}\, \eta^{1/p}\, (16)^{-1/p}\} < \tfrac{1}{16}\, 3^{-p} \quad .$$

Then by theorem 1.6.,

$$\sup_n E\|S_{nc}\|^p \leq 4.3.^p\, C^p + \eta < \infty \quad \text{i.e. b)} .$$

To prove the converse. Given $\varepsilon > 0$, choose c so that $\sup_n E\|S_{nc}\|^p < \frac{1}{3}\, \varepsilon^{p+1}$.

and $K \subseteq 0_c^c$ symmetric compact so that for all n .

(2.10.1) $\quad F_n^{(c)}(K^c) < \frac{1}{3}\varepsilon$.

Choose a simple function $t : B \to B$ such that $\|x - t(x)\| < \eta$ on K and $t(x) = 0$ off K with $\eta < c$ and $\eta \sup_n F_n^{(c)}(B) < \frac{1}{3}\varepsilon^2$. Observe that

$$(2.10.2) \quad P\{\|S_n - \sum_{j=1}^{k_n} t(X_{nj})\| > 4\varepsilon\} \leqslant P\{\|\sum_{j=1}^{k_n} (X_{nj} - t(X_{nj}))_c\| > 2\varepsilon\}$$

$$+ P\{\|\sum_{j=1}^{k_n} \widetilde{(X_{nj} - t(X_{nj}))}_c\| > 2\varepsilon\} .$$

The second term on the RHS of the above inequality does not exceed

$$\sum_{j=1}^{k_n} P\{\|X_{nj} - t(X_{nj})\| > c\} = \sum_{j=1}^{k_n} P\{\|X_{nj} - t(X_{nj})\| > c,\ X_{nj} \notin K\}$$

as $\eta < c$. But for $X_{nj} \notin K$, $t(X_{nj}) = 0$ giving

$$(2.10.3) \quad P\{\|\sum_{j=1}^{k_n} \widetilde{(X_{nj} - t(X_{nj}))}_c\| > 2\varepsilon\} \leqslant F_n^{(c)}(K^c) .$$

The first term on the RHS of (2.10.2) does not exceed

$$(2.10.4) \quad P\{\|\sum_{j=1}^{k_n} (X_{nj} - t(X_{nj}))_c\, 1(X_{nj} \notin K)\| > \varepsilon\} +$$

$$+ P\{\|\sum_{j=1}^{k_n} (X_{nj} - t(X_{nj}))_c\, 1(X_{nj} \in K)\| > \varepsilon\} .$$

The first term above does not exceed

$$(2.10.5) \quad P\{\|\sum_{j=1}^{k_n} X_{njc}\| > \varepsilon\} \leqslant \frac{1}{\varepsilon^p} E\|S_{nc}\|^p \quad \text{as} \quad 0_c \subseteq K^c .$$

The second term does not exceed

$$\frac{1}{\varepsilon} \sum_{j=1}^{k_n} E\|(X_{nj} - t(X_{nj}))_c\, 1(X_{nj} \in K)\|$$

by Chebychev and triangle inequality. This in turn does not exceed $\frac{1}{\varepsilon}\eta F_n(K) \leqslant \frac{\eta}{\varepsilon} F_n^{(c)}(B)$. From this (2.10.1), (2.10.2), (2.10.3) and (2.10.5), we get

$\{\mathcal{L}(S_n)\}$ is flatly concentrated. Now for $y \in B'$, $c > 0$, $p > 1$ choose $\delta < c$ so that

$$(E|\langle y, S_{nc}\rangle|^p)^{1/p} \leq \|y\| \sup_n (E\|S_{n\delta}\|^p)^{1/p} + c[F_n^{(\delta)}(B)]^{1/p}$$

giving $\sup_n E|\langle y, S_{nc}\rangle|^p < \infty$. Clearly, there exists K, compact so that $\sup_n F_n^{(\delta)}(K^c) < \varepsilon$. Hence $\sup_n F_n(0_t^c) < \varepsilon$ choosing t so that $K \subset 0_t$ and $t > \delta$. Now $\{x : |\langle y,x\rangle| > t\} \subseteq 0^c_{t/\|y\|}$ giving by theorem 1.7. that $\{\langle y, S_n\rangle\}$ is stochastically bounded. Thus we get $\{\mathcal{L}(S_n)\}$ is tight by well-known theorem of de Acosta.

2.11. COROLLARY. Let $\{X_{nj},\ j = 1,2,\dots,k_n\}$ $n = 1,2,\dots$ be a U.I. triangular array such that $\{\mathcal{L}(S_n)\}$ is relatively compact with all limit points non-Gaussian then for every $\varepsilon > 0$, there exists a finite-dimensional subspace $\mathcal{M}$ and a triangular array $\{t(X_{nj})\}$ U.I. and uniformly bounded such that $\{\sum_{j=1}^{k_n} t(X_{nj})\}$ is tight

$$P(t(X_{nj}) \in \mathcal{M}) = 1 \quad \text{and} \quad P\{\|S_n - \sum_{j=1}^{k_n} t(X_{nj})\| > \varepsilon\} < \varepsilon \ .$$

2.12. COROLLARY. Let $\{X_{nj},\ j = 1,2,\dots,k_n\}$ be U.I. triangular array of uniformly bounded r.v.'s. with $\mathcal{L}(S_n) \Rightarrow \nu$. Then for each $p > 0$, $\varepsilon > 0$ there exists a symmetric U.I. triangular array $\{W_{nj}\}$ such that

i) $\{W_{nj}\}$ is a measurable function of $\{X_{nj}\}$ only for each n,j.

ii) There exists a finite-dimensional subspace $\mathcal{M}$ such that $P(W_{nj} \in \mathcal{M}) = 1$; $P(W_{nj} \in \mathcal{M}) = 1$.

iii) $\{\sum_{j=1}^{k_n} W_{nj})\}$ is tight in $\mathcal{M}$ and

iv) $\sup_n E\|\sum_{j \leq k_n} X_{nj} - \sum_{j \leq k_n} W_{nj}\|^p < \varepsilon$.

<u>Proof</u> : Choose $c_n \downarrow 0$ as in Corollary 2.9. Then $\{\tilde{S}_{nc_n}\}$ converges to a non-Gaussian limit. By the above corollary for $\varepsilon > 0$, $p > 0$ there exists $t : B \to B$ simple symmetric with finite dimensional rauge and $n_o \in \mathbb{N}$ such that for $n \geq n_o$

$$E\|\tilde{S}_{nc_n} - \sum_{j=1}^{k_n} t(X_{njc})\|^p < \varepsilon/4 \ .$$

As $\mathcal{L}(S_{nc_n}) \Rightarrow \gamma$ gaussian. Let $\mathcal{L}(Z) = \gamma$ and Z be written as a.s. convergent series

$$Z = \sum_{j=1}^{\infty} < y_j , Z > x_j$$

where $\{x_j\} \subseteq B$ and $y_j \in B'$. Since $\mathcal{L}(S_{nc_n}) \Rightarrow \mathcal{L}(Z)$, $\mathcal{L}(S_{nc_n} - \pi_k(S_{nc_n})) \Rightarrow$ $\mathcal{L}(Z - \pi_k(Z))$ with $\pi_k(x) = \sum_{j=1}^{k} <y_j , x> x_j$. By theorem 1.7., $\{\|S_{nc_n} - \pi_k(S_{nc_n})\|^p\}$ is uniformly integrable for $p > 0$. Hence

$$E\|S_{nc_n} - \pi_k(S_{nc_n})\|^p \to E\|Z - \pi_k(Z)\|^p \ .$$

Choose k_o so that $E\|Z - \pi_{k_o}(Z)\|^p < \delta$ and n_1 so that for $n \geq n_1$

$$E\|S_{nc_n} - \pi_{k_o}(S_{nc_n})\|^p < \varepsilon/4 \ .$$

Now $W_{nj} = t(X_{nj}) + \pi_{k_o}(X_{nj})$ for $n \geq n_o \vee n_1$. Then $\{W_{nj}\}$ satisfy the given conditions for $n \geq (n_o \vee n_1)$. For $n \leq n_o \vee n_1$, choose an appropriate simple function approximation.

We now look at this approximation in the case $X_{nj} = X_j/\sqrt{n}$ and $X_1 \dots$ $\dots X_n \dots$i.i.d. Let us observe that by the finite-dimensional result, the limit is Gaussian and by theorem 1.7., $\sup_n n\, P(\|X_1\| > \sqrt{n}t) < \infty$ giving $\Lambda^2(X_1) < \infty$. Hence $E\|X_1\|^p < \infty$, $p < 2$. Also $\frac{1}{\sqrt{nk}} \sum_{j=1}^{nk} X_j = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y^{(k)}$ where $Y^{(k)}$ are i.i,d with $\mathcal{L}(Y^{(k)}) = \mathcal{L}(X_1 + \dots + X_k/\sqrt{k})$. Again stochastic boundedness of

$\{\frac{1}{\sqrt{nk}} \sum_{j=1}^{nk} X_j\}$ implies $\Lambda^2(Y^{(k)}) < \infty$ and for $p < 2$,

$$E\|Y^{(k)}\|^p = \int_0^\infty P(\|Y^{(k)}\| > t)dt \leq 1 + \int_1^\infty M \frac{1}{t^2} dt = M + 1 \quad .$$

Hence $\sup_k E\|X_1 + \ldots + X_k/\sqrt{k}\|^p < \infty$ for $p < 2$. By Lemma 1.5., we get

$$E\|S_{nc}\|^p \leq 2\, E\|S_n\|^p \quad .$$

Now let π_k be approximating family so that $\sup_n E\|(I - \pi_k)S_{nc}\|^p < \varepsilon$. Choose $1 \leq p < 2$, then $E\|(I - \pi_k)(S_n - S_{nc})\|^p \leq 3 \sup_n E\|S_n\|^p$. This implies $\{\|(I - \pi_k)(S_n - S_{nc})\|\}$ is uniformly integrable in (n,c) . But $\|S_n - S_{nc}\| \to 0$ uniformly in n as $c \to \infty$ since

$$P(\|S_n - S_{nc}\| > \varepsilon) \leq n\, P(\|X_n\| > c\sqrt{n}) \leq \frac{1}{c^2} \Lambda^2(X_1) \quad .$$

Thus we get that $(I - \pi_k)(S_n - S_{nc}) \xrightarrow{P} 0$ uniformly in n as $c \to \infty$ and is uniformly integrable in (n,c) . Thus $E\|(I - \pi_k)\tilde{S}_{nc}\| \to 0$ as $c \to \infty$. In other words, uniformly in n ,

$$E\|(I - \pi_k)\, S_{nc}\| \longrightarrow E\|(I - \pi_k)\, S_n\| \quad \text{as} \quad c \to \infty \quad .$$

In particular,given $\varepsilon > 0$, there exists k_o such that

$$\sup_n E\|(I - \pi_k) \sum_{j=1}^{n} X_j/\sqrt{n}\| < \varepsilon \quad \text{for} \quad k \geq k_o \quad .$$

We thus have

2.14. PROPOSITION. <u>Let</u> X <u>be a symmetric B-valued random variable. Then</u> X <u>satisfies CLT iff for every</u> $\varepsilon > 0$ <u>there exists a simple random variable</u> Y <u>satisfying CLT so that</u> $\sup_n E\|X_1 + \ldots + X_n/\sqrt{n} - Y_1 + \ldots + Y_n/\sqrt{n}\| < \varepsilon$.

<u>Proof</u> : By the construction $\{\pi_k(X_1)\}$ satisfies CLT and hence is square integrable by example 1.8.2. Thus we can approximate $\pi_k(X_1)$ by Y_1 in $L_2(\pi_k(B))$ assuring Y_1 satisfy CLT. Converse is obvious by Corollary 2.12.

Remark : In order to obtain moment conditions we only use stochastic boundedness of $\{X_1 + \dots + X_n/\sqrt{n}\}$.

2.15. THEOREM. (Le Cam). Let $\{X_{nj}\}$ be a triangular array of B-valued random variables. Then $\{e(F_n)\}$ is tight implies $\{\mathcal{L}(S_n)\}$ is tight.

Proof : Note that $e(F_n) = \mathcal{L}(\sum_{j=1}^{k_n} \sum_{i=0}^{N_{nj}} X_{nji})$ where $\{N_{nj}\}$ are i.i.d. Poisson with parameter one, independent of $\{X_{nji}\}$ for all i,n,j and $\{X_{nji}\}$ $i = 0,1,\dots$ are i.i.d. with $\mathcal{L}(X_{nji}) = \mathcal{L}(X_{nj})$ for all i (always $S_o = 0$) . By theorem 1.2., $\{e(\lambda F_n)\}$ is tight for all λ iff $\{e(F_n)\}$ is tight. Hence $\{\mathcal{L}(\sum_{j=1}^{k_n} \sum_{i=0}^{N_{nj}} X_{nji})\}$ is tight with above assumptions except with $EN_{nj} = \lambda$. Choose λ so that $\exp(-\lambda) = \frac{1}{2}$ and let $T_n^* = S_n^* + \sum_1^{k_n} \sum_{\xi_{nj}<i\leqslant N_{nj}} X_{nji}$ with $T_n^* = \sum_{j=1}^{k_n} \sum_{i=0}^{N_{nj}} X_{nji}$ and $\xi_{nj} = \min(N_{nj},1)$. Then we have $\mathcal{L}(T_n^* - S_n^*) = \mathcal{L}(S_n^* - T_n^*)$. Use now an argument as in Lemma 1.5. with q , Minkowski functional of a convex, compact symmetric set K to obtain

$$P(T_n^* \in K^c) \geqslant \frac{1}{2} P(S_n^* \in K^c) \quad .$$

Thus $\{\mathcal{L}(S_n^*)\}$ is tight. But $\mathcal{L}(S_n^*) = \mathcal{L}(\sum_{j=1}^{k_n} \xi_{nj} X_{nj}) = \mathcal{L}(\sum_{j=1}^{k_n} (1-\xi_{nj})X_{nj})$ as ξ_{nj} is Bernoulli with $P(\xi_{nj} = 1) = \frac{1}{2}$. Hence $\mathcal{L}(\sum_{j=1}^{k_n} X_{nj})$ is tight.

The following theorem is now immediate from Corollary 2.12. and Theorem 2.15.

2.16. THEOREM. Let $\{X_{nj}\}$ $(j = 1,2,\dots,k_n$, $n = 1,2,\dots)$ be U.I. triangular array. Then $\mathcal{L}(S_n) \Rightarrow \nu = \gamma * e(F)$ iff for some c (and hence for all $c > 0$) we have

i) $F_n^{(\tau)} \Rightarrow F^{(\tau)}$ for all $\tau > 0$.

ii) For every $p > 0$, and $\varepsilon > 0$, there exists a symmetric U.I. triangular array $\{W_{nj}\}$ such that $\{W_{nj}\}$ is a measurable function of $\{X_{nj}\}$; a finite dimensional subspace $\mathcal{M}$ such that $P(W_{nj} \in \mathcal{M}) = 1$, $\{\mathcal{L}(\sum_{j=1}^{k_n} W_{nj})\}$ is tight in $\mathcal{M}$ and $\sup_n E\|\sum_{j=1}^{k_n} (X_{njc} - W_{nj})\|^p < \varepsilon$.

iii) Condition (2.5.1) holds.

We now consider some consequences of this theorem.

2.17. Applications :

2.17.1. Example : $B = H$ a Hilbert space. Then the above theorem implies for an H-valued triangular array,

$\mathcal{L}(S_n) \Rightarrow \gamma * e(F)$ iff

i) For each $c > 0$, $F_n^{(c)} \Rightarrow F^{(c)}$, $c \in C(F)$.

ii) For $\varepsilon > 0$ and for some complete orthonormal basis $\{e_i\}$

$$\lim_{N\to\infty} \sup_n \int_{\|x\|\leq 1} \|x - \pi_N(x)\|^2 F_n(dx) = 0 \quad \text{and} \quad \sup_n \int_{\|x\|\leq 1} \|\pi_N(x)\|^2 F_n(dx)$$

finite, with $\pi_N(x) = \sum_{j=1}^{N} (x,e_j)e_j$.

iii) $$\lim_{\varepsilon\downarrow 0} \left\{\begin{matrix}\overline{\lim}_n \\ \underline{\lim}_n\end{matrix}\right\} \int_{\|x\|\leq\varepsilon} \langle y,x\rangle^2 F_n(dx) = C_\gamma(y,y).$$

This can be seen by using theorem 1.7. and stochastic boundedness of $\{\pi_N(S_n)\}$. Let us now define T_n by

$$\langle T_n y,y\rangle = \int_{\|x\|\leq 1} \langle y,x\rangle^2 F_n(dX).$$

Then conditions (ii) and (iii) imply that $\{T_n\}$ has finite-trace and $\{T_n\}$ under the trace norm is compact i.e., for a complete orthonormal basis, $\sup_n$ trace $(T_n) < \infty$ and $\lim_N \sup_n \sum_N^{\infty} (T_n e_i, e_i) = 0$. Conversely if $\{T_n\}$ is compact then one can find a complete orthonormal basis satisfying (ii) and (iii). Thus we get the following : $\mathcal{L}(S_n) \Rightarrow \gamma * e(F)$ iff

i) $F_n^{(c)} \Rightarrow F^{(c)}$,

ii) $\{T_n\}$ is a compact sequence of trace-class operators,

iii) as above holds.

2.17.1. Example : $B = L_p$ $(p \geq 2)$, $X_{nj} = X_j/\sqrt{n}$ and $\{X_j\}$ i.i.d. Then $X_1 + \dots + X_n/\sqrt{n} \Rightarrow \gamma$ iff

i) $nP(\|X_1\| > \sqrt{n}) \to 0$,

ii) For $\varepsilon > 0$, $p > 0$ there exists π_k such that

$$E\|\sum_{j=1}^{n} X_j\ 1(\|X_j\| \leq c\sqrt{n})/\sqrt{n} - \pi_k(\sum_{j=1}^{n} (X_j\ 1(\|X_j\| \leq c\sqrt{n})/\sqrt{n}\| < \varepsilon$$

and $\{\mathcal{L}(\pi_k(S_{nc}))\}$ is tight.

iii) X_1 is Pre-Gaussian, i.e. X_1 has the same covariance as an L_p-valued gaussian r.v. $G(X_1)$.

We note that (i) $\Leftrightarrow t^2 P(\|X_1\| > t) \to 0$.

As $\pi_k(X_1)$ is pregaussian in $\pi_k(B)$ by (iii) it satisfies CLT in $\pi_k(B)$ by Cramer -Wold devise. Thus (iii) $\Rightarrow$ (ii), second part. We now show that (i) and (iii) imply the existance of π_k satisfying the first part of (ii) by Rosenthals inequality. With arguments as in 1.8.1. we get,

$$\sup_n n\, E\|X_1\ 1(\|X_1\| \leq c\sqrt{n})/\sqrt{n} - \pi_k(X_1\ 1(\|X_1\| \leq c\sqrt{n})/\sqrt{n})\|^p$$

$$\leq \text{Constant}\ \Lambda^2(X_1 - \pi_k(X_1)) \quad \text{and}$$

$$\sup_n \int [E(X_1\ 1(\|X_1\| \leq c\sqrt{n}) - \pi_k(X_1\ 1(\|X_1\| \leq c\sqrt{n}))(t)]^{p/2}\, d\mu$$

$$= \int [E(X_1 - \pi_k(X_1))(t)]^{p/2}\, d\mu\ .$$

Thus it suffices to show that in the norm $\Lambda(X_1) + (E\|G(X_1)\|^2)^{1/2}$ on $L_1(\Omega,\mathfrak{F},P)$, there is a finite-dimensional approximation. Let $\{\mathfrak{F}_k\}$ be an increasing subsequence of $\{\mathfrak{F}\}$. Define $\pi_k(X_1) = E(X|\mathfrak{F}_k)$, $X_o = 0$. Then one has by $\Lambda(X_1 - \pi_k(X_1)) \to 0$ as $k \to \infty$. Let $Y_k = \pi_{k+1}(X_1) - \pi_k(X_1)$, $\{Y_k\}$ are pregaussian and $\{G(Y_k)\}$ are independent Gaussian. Also,

$\mathcal{L}(\sum_k G(Y_k)) = \mathcal{L}(G(X_1))$. Using Fernique's theorem $\|G(X_1)\|^2$ is integrable giving $\lim_k (E\|G(X) - \sum_{j=1}^{k} G(Y_j)\|^2)^{1/2} = 0$. Thus we obtain π_k satisfying (ii).

We thus have the following theorem: X_1 satisfies CLT iff

i) $t^2 P(\|X_1\| > t) \to 0$ and

ii) X_1 is pregaussian.

2.17.3. Example : $X_{nj} = X_j/n$; X_j i.i.d., $Y = 0$, $F = 0$. Let X be a symmetric B-valued r.v. then we say that X satisfies WLLN iff for $X_1, X_2, \dots$ ... i.i.d. as X, $\mathcal{L}(\sum_{j=1}^{n} X_j/n) = \delta_o$ or equivalently $\sum_{j=1}^{n} X_j/n \xrightarrow{P} 0$.

We have X satisfies WLLN iff

i) $tP(\|X\| > t) \to 0$,

ii) $\lim_n n^{-1} E\|\sum_1^n X_1 1(\|X_i\| \leqslant n)\| = 0$.

By theorem 2.10., and theorem 2.5., X satisfies WLLN iff

1) $\forall\ c > 0$, $tP(\|X\| > t) \to 0$ and

2) For $\varepsilon > 0$, there exists δ_o such that $n^{-1}E\|\sum_{j=1}^{n} X_i 1(\|X_i\|) \leqslant \delta_o n)$ $< \varepsilon/2$ for all n. Now (1) $\Leftrightarrow$ (i) and (ii) $\Rightarrow$ 2) by writing expectation in terms of tails and using Lemma 1.4. Now choose δ_o by 2) and observe that

$$n^{-1} E\|\sum_{j=1}^{n} X_j (\delta_o n \leqslant \|X_j\| \leqslant n)\| \leqslant n^{-1} \sum_{j=1}^{n} E\|X_j\| 1(\delta_o n \leqslant \|X_j\| \leqslant n)$$

$$\leqslant n\, P(\|X_j\|) > \delta_o n) \to 0$$

as $n \to \infty$. Thus 2) $\Rightarrow$ ii).

3. CLASSICAL CLP AND GEOMETRY OF BANACH SPACES.

In this section we relate the validity of classical theorems with the associated geometry of Banach spaces. Our proofs will use freely the geometrical results. We shall not prove them but instead refer to the literature where they can be found.

3.1. Stochastic boundedness implies pregaussian : We first observe that stochastic boundedness of $\{X_1 + \dots + X_n/\sqrt{n}\}$, X_i i.i.d., does not imply X is pregaussian, as in c_o with $X = \{\varepsilon_n/\sqrt{\log n}\}$, ε_n i.i.d. symmetric Bernoulli, it is not true. We, in fact, have the following

3.1.1. THEOREM. The following are equivalent for any real separable Banach space B.

i) B does not contain an isomorphic copy of c_o.

ii) For every B-valued, integrable r.v. X, $\sup_n E\left\|\dfrac{X_1 + \dots + X_n}{\sqrt{n}}\right\| < \infty$

implies X is pregaussian.

Proof : As we have observed ii) ⇒ i), we consider now π_k as in example 2.17.2; and $X^k = \pi_k(X)$. Let $X_1^k \dots X_n^k$ be i.i.d. copies of X^k. Then

$$E\left\|\frac{X_1^k + \dots + X_n^k}{\sqrt{n}}\right\| \leqslant E\|X_1 + \dots + X_n/\sqrt{n}\| .$$

Thus X^k is pregaussian and $E\|G(X^k)\| \leqslant \lim_n E\left\|\dfrac{X_1^k + \dots + X_n^k}{\sqrt{n}}\right\|$ by CLT. Now

$G(X^k) = \sum_{i=1}^{k} G(Y^i)$ where $Y^i = X^i - X^{i-1}$. Now $\sum_{i=1}^{k} G(Y_i)$ is bounded in L_1 in

B and condition i) ⇒ by Kwapien theorem (Studia Math 52 (1974)) that

$\sum_{k=1}^{\infty} G(Y^k)$ converges. Clearly $G(X) = \sum_{k=1}^{\infty} G(Y^k)$.

3.2. Accompanying law theorem.

To start with we define

3.2.1. DEFINITION. A Banach space B contains ℓ_n^∞ uniformly [or c_o is finitely representable (f.r.) in B] if there exists $\tau \geq 1$ such that for each $n \in \mathbb{N}$ there are n vectors $x_{n_1}, \ldots, x_{n_n}$ in B satisfying

$$\max_{i \leq n} |t_i|/\tau \leq \left\| \sum_{i=1}^{n} t_i\, x_{n_i} \right\| \leq \tau \max_{i \leq n} |t_i| .$$

By a theorem of Maurey-Pisier (Studia Math 58 45-90) the following are equivalent for $q \geq 2$ and a sequence $\{\xi_i\}$ of i.i.d. centered real r.v.'s. such that $P(|\xi_1| > t) > 0$ for all t and $E|\xi_1|^q < \infty$.

(3.2.2)

(i) c_o is not f.r. in B

ii) There exists a constant $C = C(B,q,\{\xi_i\})$ finite s.t. for all sequences of points $\{x_i\} \subseteq B$,

$$E\left\|\sum_1^n x_i \xi_i\right\|^q \leq C\, E\left\| \sum_{i=1}^{n} x_i\, \varepsilon_i \right\|^q .$$

Thus we get that if c_o is f.r. B then there exists $\{x_i\} \subseteq B$ such that $\Sigma\, \varepsilon_i\, x_i$ converges but $\sum_j \overline{\xi}_j\, x_j$ diverges with $\overline{\xi}_j = e(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1})$.

There exist $k_n, \ell_n \to \infty$ such that

$$\Sigma_{\ell_n}^{k_n+\ell_n} x_j \varepsilon_j \longrightarrow 0 \quad \text{but} \quad \Sigma_{\ell_n}^{k_n+\ell_n} \overline{\xi}_j x_j \not\longrightarrow 0 .$$

Let us define $X_{nj} = \varepsilon_{j+\ell_n} x_{j+\ell_n}$. Then $\{X_{nj}\}$ is U.I. triangular array.

$\mathcal{L}(S_n) \Rightarrow \delta_o$ but $\{\mathcal{L}(\sum_{j=1}^{k_n} \overline{\xi}_j\, x_j)\}$ not tight. If it were tight by arguments as in Example 1.8.3. we get that $\sum_{j=1}^{k_n} \overline{\xi}_j\, x_j \longrightarrow 0$ as $\mathcal{L}(\sum_{j=1}^{k_n} \overline{\xi}_j\, x_j) = e(F_n)$, where $F_n = \sum_{j=1}^{k_n} \mathcal{L}(X_{nj})$.

Thus accompanying law theorem holds $\Rightarrow c_o$ is not f.r. in B. To prove the converse we need.

3.2.3. LEMMA. Let $\{X_j\}$ be i.i.d. and X_o be independent of $\{X_j\}$ with $E\|X_i\|^q < \infty$ $(i = 0,1)$. Then

$$E\|\sum_{i=0}^{n} X_i\|^q \leq E\|X_o + n X_1\|^q .$$

Proof : By Minkowski inequality,

$$(E\|X_o + \sum_1^n X_i\|^q)^{1/q} \leq (E\|\sum_{i=1}^{n} (\frac{X_o}{n} + X_i)\|^q)^{1/q}$$

$$\leq n\, E(\|\frac{X_o}{n} + X_1\|^q)^{1/q} \leq (E\|X_o + n X_1\|^q)^{1/q} .$$

3.2.4. LEMMA. The following are equivalent for $q \geq 2$.

i) c_o is not f.r. in B.

ii) There exists $L = L(B,q)$ such that for every finite sequence $X_1, X_2, \ldots, X_n$ of independent symmetric B-valued r.v.'s. with $E\|X_j\|^q < \infty$ $j = 1,2,\ldots,n$.

$$E\|\sum_{j=1}^{n} \sum_{i=1}^{N_j} X_{ji}\|^q \leq L\, E\|\sum_1^n X_j\|^q$$

where $\mathcal{L}(N_j) = e(\delta_1)$, $\{X_{ji}, i = 0,1,\ldots\}$ is i.i.d. with $\mathcal{L}(X_{ji}) = \mathcal{L}(X_j)$ and $\{X_{ji}\}$, $\{N_j\}$ are independent.

Proof : (ii) $\Rightarrow$ (i). Let $\{x_j\} \subseteq B$, $n \in \mathbb{N}$, $\{\varepsilon_j\}$ be i.i.d. symmetric Bernoulli, N with $E N = 1$, Poisson r.v. independent of $\{\varepsilon_j\},\{\xi_j\}$, i.i.d. Poisson, $E\,\xi_1 = 1$, and $\{\bar{\xi}_j\}$, independent symmetrization of $\{\xi_j\}$. Then

$$e(\mathcal{L}(x\varepsilon_i)) = \mathcal{L}(x \sum_{j=0}^{N} \varepsilon_j) \quad \text{and} \quad \mathcal{L}(\sum_{j=0}^{N} \varepsilon_j) = e(\tfrac{1}{2}\delta_{-1} + \tfrac{1}{2}\delta_{+1}) = \mathcal{L}(\bar{\xi}_1) .$$

From (3.2.2) and (ii) this gives ii) $\Rightarrow$ (i). To prove the converse, By (3.2.2) and Fubini theorem we get

$$E\|\sum_{j=1}^{n} X_j\,(N_j - 1)\|^q \leq L\, E\|X_j\varepsilon_j\|^q .$$

By Lemma 3.2.3., using E_2 for expectation on N_j and E_1 on X_j we get

$$E\Big\|\sum_{j=1}^{n}\sum_{i=0}^{N_j} X_{ji}\Big\|^q \leq E_2\, E_1\Big\|\sum_{j}\sum_{i=0}^{N_j} X_{ji}\Big\|^q$$

$$\leq E_2\, E_1\Big\|\sum_{j} N_j\, X_j\Big\|^q$$

and this in turn does not exceed

$$2^{q-1}\, E_2E_1\Big\|\sum_{j}(N_j-1)X_j\Big\|^q + 2^{q-1}\, E_2E_1\Big\|\sum_{j} X_j\Big\|^q$$

$$\leq 2^{q-1}(C+1)\, E\Big\|\sum_{j=1}^{n} X_j\Big\|^q \quad .$$

3.2.5. THEOREM. The following are equivalent for any real separable Banach space B.

i) c_o is not f.r. in B.

ii) For any symmetric U.I. triangular array $\{X_{nj}\}$, $\mathcal{L}(S_n)$ converges $\Rightarrow e(F_n)$ converges. In other words, accompanying law theorem holds.

Proof : As ii) $\Rightarrow$ i) is proved before we move to i) $\Rightarrow$ ii). Let $\delta > 0$, $\delta \in C(F)$ where F is the Lévy measure, associated with $\lim_n \mathcal{L}(S_n)$. Then by theorem 2.1. and 2.15. one can assume that $\{X_{nj}\}$ are uniformly bounded. Using Corollary 2.12. and Lemma 3.2.4. to $X_{nj} - W_{nj}$ where $\{X_{nji} - W_{nji}\}$ are i.i.d. as $X_{nj} - W_{nj}$ except $X_{njo} - W_{njo} = 0$. We get for every $\varepsilon > 0$,

$\sup_n E\Big\|\sum_{j=1}^{k_n}\sum_{i=0}^{N_j}(X_{nji} - W_{nji})\Big\|^q \leq L\varepsilon$. As $\{W_{nj}\}$ take values in a finite-dimensional space $\mathcal{L}(\sum_{j=1}^{k_n}\sum_{i=0}^{N_j} W_{nji})$ is tight. Thus by theorem 2.16. we get the result.

3.2.6. COROLLARY. The following are equivalent for a Banach space B.

i) c_o is not f.r. in B.

ii) For every B-valued symmetric U.I. triangular array $\{X_{nj}, j = 1,\ldots,k_n\}$ $n = 1,2,\ldots.$

$\{\mathcal{L}(S_n)\}$ tight implies $\{e(F_n)\}$ tight.

3.3. Lévy-Kinchine representation and type, cotype :

In the classical case the function (with F symmetric) $\varphi(y) = \exp\left(\int(\cos(y,x) - 1)\, F(dx)\right)$ is a c.f. of a (necessarily) i.d. law if F is a Lévy measure. One knows that, in general, such a functional is not a c.f. we want to examine conditions under which it is. If F has finite variation then such a function is a c.f. of $e(F)$. Hence without loss of generality, $F|_{O_1^c} = 0$. Let $F_n = F|_{O^c_{1/n}}$ and assume variations of F_n converge to ∞. Hence $F_n = k_n \mu_n$ with μ_n a probability measure. If c_o is not f.r. in B then by theorem 3.2.5., $\{e(F_n)\}$ converges iff $\mu_n^{*k_n}$ converges. Denote by $X_{nj} = \mathcal{L}(\mu_n)$ $j = 1,2,\ldots,k_n$. Then by theorem 2.16. we get with $S_n = \sum_{j=1}^{k_n} X_{nj}$. (Note that $\mu_n = F_n/\|F_n\|_V$).

Let c_o be not f.r. in B. Then φ is a c.f. of an i.d. law iff $\mathcal{L}(S_n)$ converges. For this to happen, the necessary and sufficient conditions are

i) For $\varepsilon > 0$ and $q > 0$ there exists a finite dimensional subspace $\mathcal{M}$ and a triangular array W_{nj}, $\mathcal{M}$-valued such that

$$\text{i)}\quad \sup_n E\Big\|\sum_{j=1}^{k_n} (X_{nj} - W_{nj})\Big\|^q < \varepsilon$$

$$\text{ii)}\quad \Big\{\mathcal{L}\Big(\sum_{j=1}^{k_n} W_{nj}\Big)\Big\} \text{ is tight .}$$

Of course, this is not a very good condition but in special cases we can reduce it to a simple condition.

We need for this the following.

3.3.1. DEFINITION.

a) Let B , $\mathfrak{X}$ be separable Banach spaces and $v : B \longrightarrow \mathfrak{X}$ be a linear map. Then $(v,B,\mathfrak{X})$ is said to be R-type p if there exists $\alpha > 0$, such that for $X_1, \dots, X_n$ symmetric independent B-valued, p-summable r.v.'s. ,

$$E\|v(S_n)\|_{\mathfrak{X}}^p \leq \alpha \sum_1^n E\|X_i\|^p \quad .$$

b) If $B = \mathfrak{X}$ and $v = I$, then B is called of R-type p .

If B is R-type p , then c_o is not f.r. in B by a result of Maurey-Pisier (referred earlier). Also, since $\lim_n \mathcal{L}(S_n)$ is non-Gaussian $W_{nj} = t(X_{nj})$ for a simple function t , $\|t(x)\| \leq \|x\|$. Thus a sufficient condition for i), ii) to happen is that for $\varepsilon > 0$, there exists a simple function t (theorem 1.7.), s.t.

$$\sup_n \int \|x - t(x)\|^p F_n(dx) = \int_{\|x\|\leq 1} \|x - t(x)\|^p F(dx)$$

does not exceed ε/α . Thus we have

3.3.2. PROPOSITION. The following are equivalent

i) B is of R-type p .

ii) For every Lévy measure F satisfying $\int \|x\|^p F(dx)$ finite, $\varphi(y)$ is a c.f. of a probability measure.

Proof : Under the condition we can choose a simple function t as above. Thus $\{e(F_n)\}$ is tight but $\varphi_{e(F_n)}(y) \longrightarrow \varphi(y)$. Hence $\varphi(y) = \varphi_\mu(y)$ for some probability measure μ and $e(F_n) \Rightarrow \mu$. Clearly F is the Lévy measure of μ . For the converse implication, suppose $\sum_j \|x_j\|^p < \infty$ and write

$$F = \lim_n \sum_{j=1}^n \left(\frac{1}{2}\delta_{x_j} + \frac{1}{2}\delta_{-x_j}\right) .$$ Then $\int \|x\|^p \, dF < \infty$. Hence $\varphi_1(y) = \varphi_\mu(y)$.

But $\varphi(y) = \lim_n \prod_{j=1}^n \varphi_{\xi_j x_j}(y)$ with $\{\xi_j\}$ i.i.d. symmetric Poisson real-valued r.v.'s. Hence $\mathcal{L}(\sum_{j=1}^{n} \xi_j x_j) \Rightarrow \mu$. Giving $\sum_{j=1}^{\infty} \xi_j x_j$ converges a.e. But this implies $\sum_{j=1}^{\infty} \varepsilon_j x_j$ converges a.e. by Contraction Principle.

3.3.3. DEFINITION. We say that B is of cotype q (Radmacher) $(q \geq 2)$ if there exists $\alpha > 0$, such that for $X_1, \ldots, X_n$ symmetric independent B-valued p-summable r.v.'s.

$$E\|S_n\|^q \geq \alpha \sum_1^n E\|X_i\|^q .$$

3.3.4. PROPOSITION. The following are equivalent

i) B is of cotype q.

ii) Every non-Gaussian i.d. law has Lévy measure satisfying $\int \|x\|^q \, dF$ finite.

Proof : We note that i) $\Rightarrow$ c_o is not f.r. in B. Hence by the necessary and sufficient conditions we get that

$$\sup_n E\| \sum_{j=1}^{k_n} X_{nj}\|^q < \infty .$$

Hence by cotype property of B, $\sup_n \sum_{j=1}^{k_n} E\|X_{nj}\|^q < \infty$.

But this gives $\int \|x\|^q F(dx) < \infty$ as $F_n \uparrow F$. To prove the converse assume $\Sigma x_i \xi_i$ converges then it follows by the assumption ii) that $\Sigma \|x_i\|^q$ converges. Thus by closed Graph theorem for every sequence $\{x_i\} \subseteq B$; $\sum_{i=1}^{n} \|x_i\|^q \leq$ constant $E\| \sum_1^n \xi_i x_i\|$. This implies that c_o is not f.r. in B. (Hamedani and Mandrekar Studia Math 66 (1978)). Hence by Section 3.2., $\Sigma \varepsilon_j x_j$ converges implies $\Sigma \|x_j\|^q < \infty$ giving cotype q property of B.

3.4. CLP and CLT in Banach spaces of type 2 :

We prove the following result.

3.4.1. THEOREM. The following are equivalent for a real separable Banach space of infinite dimension.

a) B is of type 2.

b) For any U.I. symmetric triangular array $\{X_{nj},\ j = 1,2,\dots,k_n\}$, $n = 1,2,\dots$ and F σ-finite measure,

i) $F_n^{(c)} \Rightarrow F^{(c)}$ for each $c \in C(F)$.

ii) For $\varepsilon > 0$, there exists a finite-dimensional subspace $\mathcal{M}$ valued r.v.'s. $\Phi(X_{nj})$ such that $\sup_n \sum_{j=1}^{k_n} E\|X_{njc} - \Phi(X_{njc})\|^2 \leq \varepsilon$.

iii) $\lim_{\varepsilon \downarrow 0} \overline{\lim}_n \langle y, S_{n\varepsilon}\rangle^2 = C_\gamma(y,y)$ for a cylindrical Gaussian γ imply $\mathcal{L}(S_n) \Rightarrow \gamma * e(F)$ with γ Gaussian.

c) For every U.I. symmetric triangular array $\{X_{nj},\ j = 1,2,\dots,k_n\}$ of B-valued random variables and a σ-finite measure F,

i) $F_n^{(c)} \Rightarrow F^{(c)}$ for $c \in C(F)$,

ii) $\lim_{c \to 0} \overline{\lim}_n \int_{\|x\| \leq c} \|x\|^2 dF_n = 0$ imply $\mathcal{L}(S_n) \Rightarrow e(F)$.

d) $E\|X\|^2 < \infty \Rightarrow$ CLT holds.

e) $E\|X\|^2 < \infty \Rightarrow X$ is pregaussian.

Proof : In view of theorem 2.16. and type 2 we get a) $\Rightarrow$ b). Condition ii) of c) $\Rightarrow C_\gamma(y,y) = 0$ and by Corollary 2.11. condition ii) of b). Hence b) $\Rightarrow$ c). We show c) $\Rightarrow$ a). Suppose $\sum_j \|x_j\|^2 < \infty$ but $\sum \varepsilon_j x_j$ does not converge for some $\{x_j\} \subseteq B$. Then there exist ℓ_n, k_n such that $(\ell_n \to \infty,\ k_n \to \infty)$

$$\sum_{j=\ell_n+1}^{\ell_n+k_n} \|x_j\|^2 \to 0 \quad \text{but} \quad \mathcal{L}\Big(\sum_{j=\ell_n+1}^{\ell_n+k_n} \varepsilon_j x_j\Big) \not\Rightarrow \delta_0 .$$

Define $X_{nj} = \varepsilon_{\ell_n+j}\, x_{\ell_n+j}$, $j = 1,2,\dots,k_n$. Then by c) $\mathcal{L}\big(\sum_{j=1}^{k_n} X_{nj}\big) \Rightarrow \delta_0$ reaching a contradiction. Thus $\sum \varepsilon_j x_j$ converges, giving a). To see

b) $\Rightarrow$ d) . Clearly, $E\|X\|^2 < \infty \Rightarrow t^2 P(\|X\| > t) \to 0$ as $t \to \infty$. Hence $F_n^{(c)} = n P(\|X\| > c\sqrt{n}) \to 0$ for each $c \to 0$. Condition b) (iii) is satisfied as $E<y,X>^2 < \infty$. Let $q(x) = \inf \{\|x-y\|, y \in \mathcal{M}\}$. The given condition b) (ii) is satisfied if for $\varepsilon > 0$ we can find $\mathcal{M}$ so that $\sup_n E\, q(\ 1(\|X\| \leqslant \sqrt{n}))^2 = E(q(X))^2 \leqslant \varepsilon$. Given $\varepsilon > 0$, choose simple function t , such that

$$E\|X - t(X)\|^2 < \varepsilon \quad .$$

Choose $\mathcal{M}$ such that $t(X) \in \mathcal{M}$ a.s. Obviously d) $\Rightarrow$ e) . For e) $\Rightarrow$ a) assume $\sum_j \|x_j\|^2 = 1$ and choose $\mathcal{L}(X) = \sum_{j=1}^{\infty} \frac{1}{2} \|x_j\|^2(\delta_{x_j} + \delta_{-x_j})$. Then $E\|X\|^2 < \infty$ and hence X is pregaussian i.e. $\exp(-\frac{1}{2} \sum_{j=1}^{\infty} <y,x_j>^2) = \varphi_\gamma(y)$ for $y \in B'$ and γ Gaussian measure. By Ito-Nisio theorem this implies that $\sum_{j=1}^{\infty} \gamma_j x_j$ converges a.s.

<u>Remark</u> : A reader is encouraged to state and prove equivalences of a),b),c),d),e), for a triplet $(v,B,\mathfrak{X})$ of R-type 2 . There is not much change in the proof. Also one can prove by the same proof equivalence of a), b) and c) for R-type p with 2 replaced by p .

3.5. <u>Domains of Attraction and Banach Spaces of Stable type</u> p $(p < 2)$:

We say that a Banach space B is of stable type p if for $\{x_j\} \subseteq B$, satisfying $\sum_j \|x_j\|^p < \infty$ we have $\sum_j x_j \eta_j$ converges a.s., where $\{\eta_j\}$ i.i.d. symmetric stable with $\varphi_{\mathcal{L}(\eta_1)}(t) = \exp(-|t|^p)$.

We say that a B-valued r.v. X is in the domain of attraction of a B-valued r. v. Y if there exist $b_n > 0$ and $x_n \in E$ $(n = 1,2,\ldots)$ such that

$$\mathcal{L}(X_1 + \ldots + X_n/b_n - x_n) \Rightarrow \mathcal{L}(Y)$$

(We write $X \in DA(Y)$) .

The domain of attraction problem is to characterize the $\mathcal{L}(X)$ so that $X \in DA(Y)$. We note that if $X \in DA(Y)$ then $aX + x \in DA(aY + x)$ for $a \in \mathbb{R}$, $x \in B$. Thus the domain of attraction problem is a problem of determination of type of $\mathcal{L}(X)$.

As in the classical case, one needs :

3.5.1. Convergence of Type Theorem : Let $\{X_n, n = 1,2,\ldots\}$ be B-valued r.v.'s. such that $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$ and there exist constants $\{a_n\} \subseteq \mathbb{R}$ such that $\mathcal{L}(a_n X_n + x_n) \Rightarrow \mathcal{L}(Y)$ then there exists $a \in \mathbb{R}$, such that $|a_n| \to |a|$ and $x_n \to x$ provided there exists $y \in B'$ such that $\alpha(<y,X>)$ and $\mathcal{L}(<y,Y>)$ are non-degenerate. In particular, $\mathcal{L}(aX + x) = \mathcal{L}(Y)$ if $a_n > 0$.

The proof is exactly as in the one dimensional case and hence is left to the reader.

Remark : For any $x \in B$ and for every sequence $\{\mathcal{L}(X_n)\}$ there exist x_n and $b_n \neq 0$ such that $\mathcal{L}(b_n X_n + x_n) \Rightarrow \delta_x$. To see this choose $\{c_n\}$ so that $P\{\|X_n\| > c_n\} < \frac{1}{n}$ to obtain $P(\|X_n/nc_n\| > \frac{1}{n}) < \frac{1}{n}$. Hence $\mathcal{L}(X_n/c_n n) \Rightarrow \delta_o$. Choose b $b_n = \frac{1}{nc_n}$ and $x_n = x$. Thus all laws are in the DA of degenerate law.

3.5.2. THEOREM. A r.v. $X \in DA(Y)$ with $<y,Y>$ non-degenerate for some $y \in B'$. Then

i) $b_n \to \infty$, $b_n/b_{n+1} \to 1$

and

ii) for all a,b real there exists a $c(a,b) \in B$ s.t.

$$\mathcal{L}(aY_1 + bY_2) = \mathcal{L}(c(a,b)Y + x(a,b)) \text{ with } Y_1, Y_2 \text{ i.i.d. as } Y.$$

In the one-dimensional case, such laws are called stable (as their type is stable under sums). As $\varphi_\eta(t) = \exp(-|t|^p)$ for some p in the one-dimensional case, we get, $c(a,b) = (|a|^p + |b|^p)^{1/p}$ and $x(a,b) = 0$ in the symmetric case. We say that a symmetric r.v. Y is stable r.v. of index p if Y satisfied Theorem 3.5.2. (b) with $c(a,b) = (|a|^p + |b|^p)^{1/p}$ and $x(a,b) = 0$. Note that $p \leq 2$. Using induction on the definition os stable r.v. with

$a_1 = a_2 = \dots = a_n = 1$ we get for $x_n \in B$

(3.5.3) $\mathcal{L}(n^{-1/p}(Y_1 + \dots + Y_n) - x_n) = \mathcal{L}(Y)$.

3.5.4. THEOREM. <u>A non-degenerate</u> Y <u>has non-empty domain of attraction iff</u> Y <u>is stable.</u>

Now (3.5.3) with $p = 2$ gives Y is Gaussian. As non-degenerate Gaussian laws do not satisfy (3.5.3) for $p < 2$, we call the laws with index $p < 2$ as non-Gaussian stable laws. Also (3.5.3) implies Y is i.d. and in the symmetric case $x_n = 0$. Let F be Lévy measure associated with $\mathcal{L}(Y)$. Let $F_n(\cdot) = F(n^{-1/p}\cdot)$, then by (3.5.3), for Y symmetric,

$$\mathcal{L}(n^{1/p}\, Y) = \mathcal{L}(Y_1 + \dots + Y_n)$$

and hence by uniqueness of Lévy measure, $F_n = nF$. Let A be Borel subset of $\{x|\ \|x\| = 1\}$, and $M(r,A) = F\{x \in B\ ;\ \|x\| > r\ ,\ \frac{x}{\|x\|} \in A\}$ $r > 0$. Then

$$nM(1,A) = M(n^{-1/p},A) = k\ M((k/n)^{\frac{1}{p}},A) \quad .$$

By monotonicity of M we get for $r > 0$

$$M(r,A) = r^{-p}\ M(1,A) = r^{-p}\ \sigma(A) \quad \text{(say)} .$$

3.5.5. COROLLARY. $\varphi_{\mathcal{L}(Y)}(y) = \exp\{\int_S |\langle y,s\rangle|^p\ \sigma(ds)\}$ <u>for a symmetric stable r.v.</u> Y <u>of index</u> p . <u>Here</u> σ <u>is the unique measure on the unit sphere</u> S <u>of</u> B .

By using (3.5.3) and Theorem 1.7. we have $\sup_c c^p\ P(\|Y\| > c) < \infty$ for Y symmetric stable. Hence $E\|Y\|^\beta < \infty$ for $\beta < p$. From Theorem 2.10 we get that a symmetric B-valued r.v. $X \in DA(Y)$ iff

$$(3.5.6)\quad \begin{array}{ll} \text{(a)} & nP(\|X\| > rb_n\ ,\ \frac{X}{\|X\|} \in A) \to r^{-p}\,\sigma(A) \quad \text{for} \quad r > 0 \quad \text{and} \\ & \sigma(\partial A) = 0\ . \\ \text{(b)} & \lim_{\varepsilon\to 0}\ \overline{\lim}_n\ b_n^{-q}\, E\|Z_1 + \dots + Z_n\|^q = 0 \quad \text{for some} \quad q > 0 \\ & \text{with} \quad Z_i = X_i\ 1(\|X_i\|) \leqslant \varepsilon b_n)\ . \end{array}$$

By elementary calculations, using $b_n \to \infty$ and $b_n/b_{n+1} \to 1$ and (3.5.6) (a) we get

$$(3.5.7)\quad \frac{P(\|X\| > rt)}{P(\|X\| > t)} \to r^{-p}\ , \quad \text{as} \quad t \to \infty\ .$$

i.e., $P(\|X\| > .)$ is regulary varying of index $(-p)$. Also for A with $\sigma(\partial A) = 0$, as $t \to \infty$

$$(3.5.8)\quad P(\|X\| > t,\ \frac{X}{\|X\|} \in A)/P(\|X\| > t) \to \sigma(A)/\sigma(S)\ .$$

In particular $X \in DA(Y)$ implies $E\|X\|^q < \infty$ for $q < p$. To obtain sufficiency we observe using regular variation

$$\frac{t^p\ P(\|X\| > t)}{E\|X\|^q 1(\|X\| \leqslant t)} \to \frac{q-p}{p} \quad \text{as} \quad t \to \infty\ .$$

Put $t = b_n\varepsilon$ and multiply the dominator and numerator by n to obtain from (3.5.7)

$$(3.5.9)\qquad \lim_n\ nb_n^{-q}\, E\|Z\|^q = \frac{p}{q-p}\ \varepsilon^{p-q}\ .$$

It is known that if B is of stable type p then for any family $(W_1,\dots,W_n)$ of symmetric independent B-valued r.v.'s. with $E\|W_i\|^q < \infty$ $(i = 1,\dots,n\ ;\ q < p)$ there exists C such that

$$E\|\sum_1^n W_i\|^q \leqslant C \sum_{i=1}^n E\|W_i\|^q \quad . \quad \text{(see e.g. Maurey-Pisier)}\ .$$

From this, (3.5.9.),(3.5.6),(3.5.7) and (3.5.8)

3.5.10. THEOREM. Let B be of stable type $p < 2$. Then $X \in DA(Y)$ iff Y is stable and X satisfies (3.5.7) and (3.5.8).

In the "if" part one produces b_n using (3.5.7).

3.5.11. THEOREM. The following are equivalent for $p < 2$.

a) B is of stable type p.

b) Conditions (3.5.7) and (3.5.8) for some σ are necessary and sufficient for $X \in DA(Y)$ with σ being the measure associated with Lévy measure of Y.

c) $t^p P(\|X\| > t) \to 0$ iff $\dfrac{X_1 + \ldots + X_n}{n^{1/p}} \xrightarrow{P} 0$

Proof : We have proved i) $\Rightarrow$ ii). To prove ii))implies iii), choose θ, symmetric, stable, real-valued r.v. independent of X and $e \in B$ s.t. $\|e\| = 1$. Then it is easy to check that $P(\|X + \theta e\| > .)$ is regularly varying of index $(-p)$. Note that $nP(n^{-1/p} \theta \varepsilon .) \Rightarrow d\Gamma \times r^{-(1+p)}dr$ with $\Gamma(+1) = \Gamma(-1) > 0$ and $\operatorname{supp} \Gamma = \{+1,-1\}$. Hence for $\lambda > 0$, there exists a closed symmetric interval J with interior of $J \supseteq [-\lambda,\lambda]$ and $\delta > 0$ such that $(J^c)^\delta \subseteq [-\lambda,\lambda]$ and $nP(\theta/n^{-1/p} \in (J^c)^\delta) < \varepsilon$. Here $(J^c)^\delta$ denotes δ neighbourhood of J^c. Now choose δ_o s.t. $[(Je)^c]^{\delta_o} \cap \mathbb{R}e \subseteq (J^c)^\delta e$. Then since $t^p P(\|X\| > t) \to 0$, there exists $n_o(\varepsilon,\delta_o) = n_o$ such that for $n \geqslant n_o$ $nP(\|X\| > \delta_o n^{1/p}) < \varepsilon$. Thus

$$\begin{aligned} nP(n^{-1/p} Y \notin Je) &\leqslant nP(n^{-1/p} Y \notin Je, \|X\| \leqslant \delta_o n^{1/p}) + nP(\|X\| > \delta_o n^{1/p}) \\ &\leqslant nP(n^{-1/p} \theta e \in [(Je)^c]^{\delta_o}) + \varepsilon \\ &\leqslant nP(n^{-1/p} \theta \in (J^c)^\delta) + \varepsilon = 2\varepsilon . \end{aligned}$$

Thus $\{nP(n^{-1/p} Y \varepsilon .)\}$ is tight outside every neighbourhood of zero. By one-dimensional result $\mathcal{L}[<y,\sum_1^n \frac{(X_i + \theta_i e)}{n^{1/p}}>] \Rightarrow \mathcal{L}(<y,\theta e>)$ for all $y \in B'$.

Here $\{\theta_i, i = 1,2,\dots,n\}$ are i.i.d., with $\mathcal{L}(\theta)$. This implies

$$n P(<y,Y>/n^{1/p} \in .) \Rightarrow F \circ y^{-1}(.) .$$

Here $dF = d\hat{\Gamma} \times r^{-(1+p)}dr$, $\operatorname{supp} \hat{\Gamma} = \{-e,e\}$, $\hat{\Gamma}(e) = \hat{\Gamma}(-e)$ equals $\Gamma(1)$. Hence

$$n P(n^{-1/p} Y \in .) \Rightarrow F .$$

This gives (3.5.7) and (3.5.8) for Y. Also by (ii) we get $b_n/n^{\frac{1}{p}} \to$ constant and

$$\sum_{j=1}^n X_j + \theta_j e/n^{1/p} \Rightarrow \theta e .$$

This gives the result. For (iii) $\Rightarrow$ (i) observe that exactly as in the proof for Proposition 2.14. we get $\sup_n E\|X_1 + \dots + X_n/n^{1/p}\|^r < \infty$ for $r < p$. Let $CL(X) = \sup_n E\|n^{-1/p} (\tilde{X}_1 + \dots + \tilde{X}_n)\|^r$ where $X_1 \dots X_n$ are i.i.d. B-valued r.v.s with $E\|X_1\|^r < \infty$ and $(\tilde{X}_1,\dots,\tilde{X}_n)$ is independent symmetrization of $(X_1,\dots ,X_n)$. Let $CL(p,r) = \{X ; X$ B-valued r.v. and $CL(X) < \infty\}$ and

$$L_0^{p,\infty} = \{X ; X \text{ B-valued r.v. and } C^p P(\|X\| > C) \to 0, C \to \infty\} .$$

On $L_0^{p,\infty}$ define $\Lambda p(X) = \sup_C C^p P(\|X\| > C)$ for $p \leq 1$ or $[\sup_C C^p P(\|X\| > C)]^{1/p}$ for $p > 1$. Under (iii), we can define T on $L_0^{p,\infty} \to CL(p,r)$. T is defined everywhere and closed. Thus by closed graph theorem $CL \leq$ Constant $\Lambda p(X)$. Let $K =$ constant. As in example 2.17.2. we can approximate $X \in L_0^{p,\infty}$ by simple functions in Λp-norm. Now if Y is a simple function then finite-dimensional CLT, $\lim_n E\|n^{-1/p} \sum_1^n Y_j\|^r = 0$ since $p < 2$. Hence range of T is included in the X is satisfying $\lim_n E\|n^{-1/p} \sum_{j=1}^n X_j\|^r = 0$, giving (iii) is a super property of B. By Maurey-Pisier-Krivine result (see Maurey-Pisier cited earlier) one has to show ℓ_p is not f.r. in B to get (i).

It suffices to show that (iii) fails in ℓ_p. Let $\{\varepsilon_j\},\{N_j\}$ be i.i.d. sequence with $\{\varepsilon_j\}$ i.i.d. symmetric Bernoulli and $P(N_j \geq n) = \dfrac{1}{n \log \log n}$

for $n \geq 27$ and 1 otherwise, $\{N_j\}$ $\mathbb{N}^+$-valued. Define

$$X_j = \varepsilon_j \sum_{N_j^2 - N_j < k < N_j^2 + N_j} e_k$$

$\{e_k\}$ natural basis of ℓ_p. One can check that $nP(\|X\|_p > (2n)^{1/p}) \to 0$ and $\{X_j\}$ i.i.d. but $\{\tilde{X}_1 + \dots + \tilde{X}_n/n^{1/p}\}$ is not stochastically bounded.

3.5.11. COROLLARY. Let B be of stable type one (B-convex). Then X satisfies WLLN iff $tP(\|X\| > t) \to 0$.

3.6. Results in the space of continuous functions : These results are special case of results in type 2 spaces. Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ be a symmetric triangular array of B-valued r.v.'s. Then $\{F_n^{(1)}\}$ is tight iff $\mathcal{L}(\sum_{j=1}^{k_n} \tilde{X}_{nj1})$ tight. Thus one wants to consider $\sum_{j=1}^{k_n} X_{nj1}$; i.e., without loss of generality, $\|X_{nj}\| \leq 1$. If we assume that $B = \bigcup_{n=1}^{\infty} nK$ with $\|x\|_K = \inf\{\lambda : x \in \lambda K\}$ for K compact and the injection $i : B \to B_K$ is continuous, i.e., if B is compactly generated, and R-type 2, then

$$E\|i(\sum_{j=1}^{k_n} X_{nj}\, 1(\|X_{nj}\| \leq 1)\|_K^2 \leq \sum_1^n E\|X_{nj1}\|^2 .$$

Since $P(\sum_{j=1}^{k_n} X_{nj1} \in (\lambda K)^c) = P(\|i(\sum_{j=1}^{k_n} X_{nj1})\|_K > \lambda)$,

by Chebychev's inequality, we get

3.6.1. THEOREM. Let $\{X_{nj}, j = 1,2,\dots,k_n\}$ $n = 1,2,\dots$ be a triangular array of B-valued r.v.'s. with B compactly generated and R-type 2. If $\{F_n^{(1)}\}$

is tight and $\sup_n \int_{\|X\|\leq 1} \|X\|^2 F_n(dX)$ is finite. Then $\{\mathcal{L}(S_n)\}$ is tight.

Remark :

1) A similar proof shows that $e(F_n)$ is tight as on type 2 space, $\int \|x\|^2 e(F_n)(dx) \leq \int \|x\|^2 F_n(dx)$.

2) By one-dimensional result $\{(1\wedge\|X\|^2)F_n(dx)\}$ tight $\Leftrightarrow$ $\{\mathcal{L}(\sum_{j=1}^{k_n}\|X_{nj}\|^2)\}$ is tight.

3) We note that the above result holds for triplet $(v,B,\mathfrak{X})$ of R-type 2 if $v(B)$ is compactly generated. In this case, $\{\mathcal{L}(\sum_{j=1}^{k_n} \|X_{nj}\|^2)\}$ tight implies $\mathcal{L}(\sum_{j=1}^{k_n} v(X_{nj}))$ tight.

We shall use the last fact to obtain results on the space of continuous functions.

Let (S,d) be a compact metric space and ρ a continuous metric on S . Define

$$|||f|||_\rho = \|f\|_\infty + \sup_{t\neq s} |f(t) - f(s)|/\rho(t,s) \quad .$$

On $C(S)$, the space of continuous functions with respect to d . Let

$$C^\rho(S) = \{f \in C(S) \ , \ |||f|||_\rho < \infty\}$$

$$C_o^\rho(S) = \{f \in C^\rho(S) \ ; \ \lim_{(t,s)\to(a,a)} |f(t)-f(s)|/f(t,s) = 0 \ , \ \forall \ a\} \quad .$$

3.6.2. LEMMA. $(C^\rho(S), |||\cdot|||_\rho)$ <u>is a Banach space and</u> $C_o^\rho(S)$ <u>is a(closed) separable subspace of</u> $C^\rho(S)$.

<u>Proof</u> : As other parts are standard, only proof needed is to show $C_o^\rho(S)$ is closed. Define T on $C_o^\rho(S)$ by

$$(Tf)(t,s) = \begin{cases} f(t)-f(s)/\rho(t,s) & \text{if } t \neq s \\ 0 & \text{if } t = s \end{cases}$$

Then T is continuous linear operator on $C_o^\rho(S)$ to $C(S \times S)$ and $Sf = (Tf,f)$ is an isometry on $C_o^\rho(S)$ into $C(S \times S) \times C(S)$ with $\|(f,g)\|_{C(S\times S)\times C(S)} = \|f\|_\infty + \|g\|_\infty$. Hence $C_o^\rho(S)$ is separable.

A continuous metric ρ is called pregaussian if for a centered Gaussian process $\{X_t , t \in S\}$

$E|X(t)-X(s)|^2 \leqslant C\rho(t,s) \Rightarrow X$ has continuous sample paths .

If on (S,ρ) there exists a probability measure λ satisfying

$$(3.6.3) \qquad \lim_{\varepsilon \to 0} \ \sup_{s\in S} \int_0^\varepsilon [\log(1 + 1/\lambda\{t \in S : e(s,t) \leqslant u\}]^{\frac{1}{2}} du = 0 \ .$$

or for metric entropy $H(S,\rho,x)$ of (S,ρ) , and some $\alpha > 0$

$$(3.6.3') \qquad \int_0^\alpha H^{1/2}(S,\rho,x)\, dx < \infty \ .$$

Then it is known (Fernique : Lecture notes in Math 480 or Dudley : J. Functional Anal. 1 (1967)) that ρ is pre-Gaussian.

3.6.4. LEMMA. *Let B be a Banach space and v a continuous operator on B into $C(S)$. If $v(B) \subseteq C^\rho(S)$ for some pregaussian metric ρ , then $(B, C(S),v)$ is of R-type 2 .*

Proof : Let $v : B \to (C^\rho(S),\||\ \||_\rho)$ is continuous by the closed graph theorem. Let $\Sigma \|x_j\|^2 < \infty$ for $\{x_j\} \subseteq B$. Then with $v(x_j) = f_j$, we have

$$\sum_{j=1}^\infty |f_j(t)-f_j(s)|^2 \leqslant \sum_{j=1}^\infty \rho^2(t,s)\||f_j|\|^2 \leqslant \text{constant } \rho^2(t,s) \sum_{j=1}^\infty \|x_j\|^2 \ .$$

By ρ being pregaussian we get $\Sigma \gamma_j f_j$ converges a.s. in $C(S)$ iff $\sum_{j=1}^\infty |f_j(t) - f_j(s)|^2 \leqslant C \rho^2(t,s)$. Hence we get $\Sigma \gamma_j f_j$ converges a.s. in $C(S)$ completing the proof.

We now recall some facts. Under (3.6.3) (or (3.6.3')), there exists ρ' satisfying (3.6.3) (or (3.6.3')) and $\rho(t,s) \leq a\,\rho'(t,s)$ with $\lim_{(t,s)\to(a,a)} \rho(t,s)/\rho'(t,s) = 0$ i.e., if a r.v. lies in $C^{\rho}(s)$, it lies in $C_o^{\rho'}(s)$. Also,

$$C_o^{\rho'}(S) = \bigcup_n nK \quad \text{with } K = \{x \, ; \, \|x\|_\rho \leq 1\} \text{ compact .}$$

Thus $C_o^{\rho'}(S)$ is compactly generated. We can thus use the remark following Theorem 3.6.1. to get

3.6.5. THEOREM. Let (S,ρ) be a compact pseudo-metric space satisfying (3.6.3) (or (3.6.3')). Let $\{X_{nj}\}$ be a $C(S)$-valued triangular array of row independent r.v.'s. Assume

i) $\mathcal{L}(S_n(t_1),\ldots,S_n(t_k))$ converges in $(C(S),\rho)$ weakly for each finite subset $(t_1,\ldots,t_k) \subseteq S$.

ii) $\|X_{nj}\|_\rho < \infty$ a.s. for j,n and $\mathcal{L}(\sum_{j=1}^{k_n} \|X_{nj}\|_\rho^2)$ is tight. Then

a) $\{e(F_n)\}$ converges and $\{\mathcal{L}(S_n)\}$ converges.

If in addition $\{X_{nj},\ j = 0,\ldots,k_n\}$ are U.I. then $\lim_n e(F_n) = \lim_n \mathcal{L}(S_n)$.

b) As c_o is f.r. in $C(S)$, we can find a triangular array, U.I. such that the above conditions are not necessary.

3.6.6. COROLLARY. Let (S,ρ) be a compact pseudo-metric space satisfying (3.6.3) (or (3.6.3')). If $E\|X\|_\rho^2 < \infty$ and X symmetric, then X satisfies CLT.

Proof : $X_{nj} = X_j/\sqrt{n}$, $\sum_{j=1}^{n} \|X_{nj}\|_\rho^2 = \frac{1}{n} \sum_{j=1}^{n} \|X_j\|^2$.

Hence by WLLN in $\mathbb{R}$ we get the result.

One can, of course, study CLP and CLT in cotype 2 spaces. Analogue of theorem 3.4.1. holds for cotype 2 spaces (involving necessary conditions). It therefore suffices to study CLT only in cotype 2 spaces. We refer the reader for this to (Chobanian and Tarieladze (1977) J. Mult. Analysis 7) .

One should note that original motivation (from the probabilistic point of view !) for probability on Banach spaces was to study Donsker's invariance principle. However theorem 3.6.5. does not include this because in this case, with

$$X_{nj}(t) = \begin{cases} 0 & 0 \leq t \leq j-1/n \\ 1 & j/n \leq t \leq 1 \\ \text{linear between } j-1/n \text{ and } j/n \end{cases} .$$

and ξ_1 satisfying CLT, one needs to show $\mathcal{L}(\sum_{j=1}^{n} X_{nj}\, \xi_j/\sqrt{n}) \Rightarrow \mathcal{L}(W)$, W being the Brownian motion on $[0,1]$. Take $\rho(t,s) = |t-s|$. Then $\frac{1}{n}\|\xi_j X_{nj}\| \geq |\xi_i|^2$. Hence $\mathcal{L}(\sum_{j=1}^{n} \|X_{nj}\|_\rho^2)$ is not tight with $X_{nj} = \xi_j X_{nj}/\sqrt{n}$. Thus, what is the influence of such CLT on classical probability theory ?

J. Kuelbs observed that CLT holds in B iff the invariance principle holds in B , for B separable. However such invariance principles are of interest in non-separable case (empirical processes). Recently, Dudley-Phillips circumvented the theory on Banach space except for the finite-dimensional approximation to construct Invariance Principle in probability (to be defined !). In the meantine, de Acosta extended Kuelbs result and obtained an a.s. Invariance Principle for non-Gaussian limit. We shall present it next for row i.i.d. triangular array. The theorem is due to de Acosta and the proof is due to Dehling-Dobrowski-Philipp.

4. INVARIANCE PRINCIPLES IN SEPARABLE BANACH SPACES.

Given an i.d. law μ on B, we can write it as $\gamma * e(F)$ if it is symmetric, with γ symmetric Gaussian, F a Lévy measure. In general, if μ is not symmetric, one can write for $\tau > 0$, $\mu = \gamma * S_\tau\, e(F) * \delta_{x_\tau}$ for $x_o \in B$ and $S_\tau\, e(F)$ denotes the probability measure whose c.f. is of the form

$$\int \exp(i < y,x >) - 1 - i < y,x\, 1(\|x\| \leqslant 1) > F(dx) .$$

Let $\mu_t = \gamma_t * S_\tau\, e(tF) * \delta_{tx_\tau}$ where $\gamma_t = \gamma(t^{-\frac{1}{2}}(\cdot))$. Then $\{\mu_t,\ t \geqslant 0\}$ is well defined (and is in fact a convolution semigroup). Here $\mu_o = \delta_o$. If $\{X_{nj}\}_{j=1}^{k_n}$ is a triangular array of row-independent B-valued r.v.'s. with $\lim\limits_n \mathcal{L}(S_n) = \mu$ (i.d.) then we get the following :

4.1. LEMMA. $\mathcal{L}\Big(\sum\limits_{\frac{k}{2^r} < j/k_n \leqslant \frac{k+1}{2^r}}\Big) \Rightarrow \mu_{1/2^r}$ $\quad k = 0,1,\ldots,2^r$

<u>Proof</u> : The proof is by induction on r. If $r = 0$, $k = 0$ then the Lemma reduces to $\mathcal{L}(S_n) \Rightarrow \mu_1 = \mu$, which is given. Assume the conclusion holds for $r - 1$ and k be fixed $= 0,1,\ldots,2^r-1$. Then k or $k+1$ is divisible by 2. First assume $k = 2i$, $i = 0,1,\ldots,2^{r-1}-1$. Then by induction hypothesis

$$\mathcal{L}\Big(\sum_{i/2^{r-1} < j/k_n \leqslant \frac{i+1}{2^{r-1}}} X_{nj}\Big) \Rightarrow \mu_{2^{r-1}} \quad \text{as } n \to \infty .$$

Let

$$\lambda_n = \mathcal{L}\Big(\sum_{k/2^r < j/k_n \leqslant \frac{k+1}{2^r}} X_{nj}\Big) \text{ and } \nu_n = \mathcal{L}\Big(\sum_{\frac{k+1}{2^r} < j/k_n \leqslant \frac{k+2}{2^r}} X_{nj}\Big) .$$

Then

$$(4.1.1)\quad \lambda_n * \nu_n = \mathcal{L}\Big(\sum_{i/2^{r-1} < j/k_n \leqslant \frac{i+1}{2^{r-1}}} X_{nj}\Big) \underset{n\to\infty}{\Rightarrow} \mu_{1/2^{r-1}} .$$

Hence there exists a sequence $\{x_n\}$ such that $\{\lambda_n * \delta_{x_n}\}$ and $\{\nu_n * \delta_{x_{-n}}\}$

is tight. But $\lambda_n = \nu_n$, $\nu_n = \lambda_n * \mathcal{L}(X_{n1})$ or $\lambda_n = \nu_n * \mathcal{L}(X_{n1})$ and $\mathcal{L}(X_{n1}) \underset{n\to\infty}{\Rightarrow} \delta_o$. $\lim_n \lambda_n * \delta_{x_n} = \lim_n \nu_n * \delta_{x_n}$ exists over a subsequence. But $\lim_n \lambda_n * \delta_{x_n} = \lim_n(\nu_n * \delta_{-x_n})* \delta_{2x_n}$. Hence $\lim_n \delta_{x_n}$ exists and is equal to δ_{x_o} . Hence $\lim_n \lambda_n = \lim_n \nu_n$, all this over the same subsequence. Using (4.1.1) , we get using linear functionals that

$$\lim \lambda_n = \mu_{1/2^r}$$

i.e.
$$\mathcal{L}\Big(\sum_{k/2^r \leq j/k_n \leq k/2^r} X_{nj}\Big) \Rightarrow \mu_{1/2^r} .$$

Let us now denote by (for r to chosen)

$$H_{nk} = \{j \; ; \; k/2^r < j/k_n \leq k+1/2^r\} \qquad (0 \leq k < 2^r)$$

and by $t_{nk} = \min H_{nk}$, $p_{nk} = \text{card. } H_{nk}$. Then we have proved that with $\mu_n = \mathcal{L}(X_{nj})$, $j = 1,2,\dots,k_n$.

4.2. COROLLARY. $\mu_n^{*pnk} \Rightarrow \mu_{1/2^r}$.

Let us denote b π the Prohorov distance and $S_{nk} = \sum_{j\leq k} X_{nj}$, we have the continuity of μ_t at zero.

4.3. LEMMA. $\lim_{c\to 0} \overline{\lim_{n\to\infty}} \sup_{k\leq ck_n} \pi(\mathcal{L}(S_{nk}),\delta_o) = 0$.

<u>Proof</u> : If the Lemma were not true we can find a sequence $\{j_n, n \geq 1\}$ of integers such that $j_n/k_n \to 0$ but $S_{nj_n} \not\to 0$ in probability. Let $\alpha_n = \mu_n^{*j_n}$ and $\beta_n = \mu_n^{*(k_n-j_n)}$, then $\alpha_n * \beta_n \Rightarrow \mu$. Hence there exists an $\{x_n\} \subseteq B$ such that $\{\alpha_n * \delta_{x_n}\}$ is tight. Now $[\varphi_{\mu_n}(y)]^{k_n} \longrightarrow \varphi_\mu(y)$ uniformly for $\|y\| \leq M$ $(M < \infty)$.

Hence $[\varphi_{\mu_n}(y)]^{j_n} \longrightarrow 1$ uniformly on $\|y\| \leq M$, noting that log of all c.f. involved exist. Hence $\mu_n^{*j_n} \Rightarrow \delta_0$ contradition.

We also note the

4.4. COROLLARY. If $j_n/k_n \to 0$, then $\mu^{j_n/k_n} \Rightarrow \delta_0$. As $\left|\frac{p_{nk}}{k_n} - \frac{1}{2^r}\right| \to 0$ we get that

$$\pi(\mu^{p_{nk}/k_n}, \mu^{1/2^r}) \to 0 \quad \text{as} \quad n \to \infty \qquad 0 \leq k < 2^r . \tag{4.5}$$

Thus we get for $n \geq n_0$,

$$\pi(\mu_n^{*p_{nk}}, \mu^{p_{nk}/k_n}) < \varepsilon\, 2^{-r} \qquad 0 \leq k \leq 2^r . \tag{4.6}$$

Re : In view of Strassen's theorem, this would say that on each block the points on the process given by $\{\mu_t, t \geq 0\}$ are close to the partial sums. But the process may have jumps.

To take care of this we need the following.

4.7. LEMMA. Let X and Y be independent B-valued random variables, with Y Gaussian. Then

$$P(\|X + Y\| \leq t)$$

is continuous.

Proof : Since X can be approximated arbitrary closely in norm by discrete r.v. we can assume X discrete. It is enough to show

$$\sum_{i=1}^{\infty} P(X = x_i)\, P(t-\varepsilon \leq \|x_i + Y\| \leq t+\varepsilon) \quad 0 \quad \text{as} \quad \varepsilon \to 0 .$$

It suffice to prove $P(t-\varepsilon \leq \|x_i + Y\| \leq t+\varepsilon) \to 0$. But $\|x_i + Y\| = \sup \{ \langle y_j, x_i + Y\rangle ; \|y_j\| \leq 1 , y_j \in B\}$. Hence this is known.

4.8. LEMMA. Let $\{Z_i , i = 1,2,\dots,n\}$ be a finite sequence of independent identically distributed r.v.'s. and with distribution of $\|Z_i\|$ continuous. Then L, defined by $\|Z_L\| = \max_{1 \leq j \leq n} \|Z_i\|$ is a well defined r.v. a.e., uniform on $\{1,2,\dots,n\}$ and independent of $S_n = \sum_{j=1}^{n} Z_j$.

Proof : $P(S_n \in A) = \sum_{j=1}^{n} P(S_n \in A, L = j) = nP(S_n \in A, L = 1)$ as the distribution of S_n is permutation invariant. Now

$$P(L = j) = P(\omega : \|Z_j\| \geq \|Z_\ell\| , \forall j \neq \ell) .$$

Hence the $P(L = j)$ is independent of j ; i.e. $P(L = j) = \frac{1}{n}$ giving the result.

Let τ_{nk} be probability measure on integers so that to each integer in H_{nk}, it assigns mass $1/p_{nk}$ and zero otherwise then τ_{nk} $(0 \leq k < 2^r ;$ $n = 1,\dots)$ is the distribution of L_{nk} such that $X_{L_{nk}} = \max_{j \in H_{nk}} \|X_{nj}\|$ (if $\|X_{nk}\|$ has continuous distribution). Now we observe that $\forall\, n \geq n_o$

$$(4.9) \qquad \pi(\mu_n^{*p_{nk}} \times \tau_{nk},\ \mu^{p_{nk}/k_n} \times \tau_{nk}) < \varepsilon/2^r \qquad\qquad 0 \leq k < 2^r .$$

Using Strassen's Theorem, we obtain, for each n, triangular arrays $\{x_{nj}, j = 1,2,\dots,k_n\}$ and $\{y_{nj}, j = 1,2,\dots,k_n\}$ $n = 1,2,\dots$ of row-wise i.i.d. r.v.'s. and triangular arrays $\{L_{nj}, 0 \leq j < 2^r\}$ and $\{M_{nj}, 0 \leq j < 2^r\}$ $n = 1,2,\dots$ with $\mathcal{L}(x_{nj}) = \mu_n$, $\mathcal{L}(y_{nj}) = \mu^{j/k_n}$ $j = 1,2,\dots,k_n$.

$$S_{nk} = \sum_{j \in H_{nk}} x_{nj} \qquad T_{nk} = \sum_{j \in H_{nk}} y_{nj} \qquad\qquad 0 \leq k < 2^r \text{ for } n \geq n_o .$$

$$(4.10) \qquad P\|S_{nk} - T_{nk}\| > \varepsilon 2^{-r} , \ \underline{\text{or}}\ \ L_{nk} \neq M_{nk}\} < \varepsilon/2^r \qquad\qquad 0 \leq k < 2^r .$$

We have shown that the sums over the block are close. The assumption of continuity of the distribution of the norm is removed by convolution μ_n and μ with a Gaussian measure of small variance (Lemma 4.7) .

Theorem we want to prove is the following

4.11. THEOREM. Let $\{\mu_n\}$ be a sequence of prob. measures such that $\mu_n^{*k_n} \Rightarrow \mu$ ($k_n \to \infty$ as $n \to \infty$). There exists a probability space and two row-wise independent triangular arrays of B-valued random variables $\{x_{nj}, 1 \leq j \leq k_n\}$ and $\{y_{nj}, 1 \leq j \leq k_n\}$ such that

(4.11.1) $\quad \mathcal{L}(x_{nj}) = \mu_n , \quad \mathcal{L}(y_{nj}) = \mu^{1/k_n} \qquad (1 \leq j \leq k_n)$

and

(4.11.2) $\quad \max_{k \leq k_n} \| \sum_{j \leq k} x_{nj} - \sum_{j \leq k} y_{nj} \| \longrightarrow 0 \quad$ a.s.

Let $\quad S_k^{(n)} = \sum_{j \leq k} x_{nj} \qquad T_k^{(n)} = \sum_{j \leq k} y_{nj}$.

Define $X_n(t) = S_k^{(n)}$ $\quad t = k/k_n \quad 0 \leq k \leq k_n$ and linear in between and $Y_n(t) = T_k^{(n)}$ $\quad t = k/k_n \quad (0 \leq k \leq k_n)$ and linear in between.

Then $Z_n = X_n - Y_n$ are $C([0,1],B)$ and $Z_n \longrightarrow 0$ in distribution. Therefore by Skorokhod's theorem $\exists\, Z_n' \ni \mathcal{L}(Z_n') = \mathcal{L}(Z_n)$ and $Z_n' \longrightarrow 0$ a.s. Thus it suffices to prove (4.11.2) with $\xrightarrow{P} 0$.

To do this on the same probability space one needs the following lemma.

4.12. LEMMA. Let $S, S_1, S_2, \dots$ be Polish spaces with distribution λ_n on $S \times S_n$ such that marginals of λ_n on S are identical. Then there exists a sequence of random variables $X, X_1, X_2, \dots$ taking values in $S \times S_1 \times \dots$ such that $\mathcal{L}((X, X_n)) = \lambda_n$.

Proof : Let $\Phi_m = S \times S_1 \times \dots \times S_m$. First we observe that for $m = 2$ we have the measure

$$\nu_2(A_1 \times A_2 \times A_3) = \int_{A_1} \lambda_1(A_2 | x)\, \lambda_2(A_3 | x)\, \lambda(dx)$$

where λ is the marginal on S and $\lambda_1(\cdot|x)$ and $\lambda(\cdot|x)$ are conditional distributions (which exist). Suppose that the lemma is proved for Φ $(j \le m)$. Apply the case $m = 2$ to Φ_m and λ_{m+1} on $S \times S_{m+1}$ to get the result.

Reduction of the theorem : It suffices to prove that given $\epsilon > 0$ $\exists$ two triangular array's $\{x_{nj}\}$ and $\{y_{nj}\}$ satisfying (4.11.1) such that

(4.13) $$\limsup_{n \to \infty} P(\max_{k \le k_n} \|S_k^{(n)} - T_k^{(n)}\| > \epsilon) < \epsilon .$$

Suppose for each m, we can find two triangular arrays $\{x_{nj}^{(m)}, j \le k_n\}$, $(\{y_{nj}^{(m)}, j \le k_n\}$ such that for $n \ge n_m$

$$P(\max_{k \le n} \|S_k^{(n)}(m) - T_k^{(n)}(m)\| > \frac{1}{m}) < \frac{1}{m} .$$

We can and do assume that that for different m's $\{(x_{nj}^{(m)}, y_{nj}^{(m)}), 1 \le j \le k_n\}$ are independent. The arrays defined by $x_{nj} = x_{nj}^{(m)}$, $y_{nj} = y_{nj}^{(m)}$ $n_m \le n \le n_{m+1}$ satisfy (4.11.1) and (4.11.2) with $\xrightarrow{P} 0$. Thus the problem is to prove (4.13). This is what we have essentially shown except the maximum is within blocks. To get maximum otherwise we need Shorokhod's inequality.

Let $D([0,1];B)$ be the space of "cadlag" functions on $[0,1]$ into B and ξ be a process with independent increments which with probability one is $D([0,1];B)$-valued

$$\Delta^P(c,\delta) = \sup \min(P(\|\xi(t)-\xi(t_1)\| > \delta;\ P(\|\xi(t_2)-\xi(t)\| > \delta)$$

and

$$\Delta(c) = \sup \min(\|\xi(t) - \xi(t_1)\| ;\ \|\xi(t_2) - \xi(t)\|)$$

the supremum is taken over all (t,t_1,t_2) $(0 \le t \le 1$, $t-c \le t_1 < t < t_2 \le t+c$.

The following lemma can be found in (Theory of Prob. Appl. 1956).

SKOROHOD LEMMA. Let $0 < c \le 1$ be such that $\Delta^P(c,\delta/20) < \frac{1}{4}$. Then for any positive integer $\ell \ge 3/c$

$$P(\Delta(1/\ell) \le 10^3\, \Delta^P(3/\ell,\delta/12)/c .$$

$$P(\Delta(1/\ell) > \delta) \le 10^3 \ \Delta^P(3/\ell,\delta/12)/c \ .$$

Let $\xi_n(t) = \sum_{i\le tk_n} x_{ni}$ and $\Delta^P(c,\delta,n)$ and $\Delta(c,n)$ be defined as above for ξ_n. Lemma 4.3.

4.14. COROLLARY. Let $\epsilon > 0$. Then

$$\lim_{c\to 0} \ \limsup_n \ \Delta^P(c,\epsilon,n) = 0 \ .$$

Now using Skorohod Lemma we get for $\epsilon > 0 \ \exists \ r = r(\epsilon) \ge 1$ such that for $n \ge n_1(\epsilon)$

(4.15) $\quad P\{\Delta(2^{-r},n) > \epsilon\} \le \epsilon \ .$

Using this r we can define H_{nk} and from (4.10) and (4.15) we get get with $S(m) = \sum_{i=m} x_{ni}$, $T(m) = \sum_{i\le m} y_{ni}$ and $n \ge \max(n_0,n_1)$,

$$\max_{k<2^r, t_{nk}<m\le t_{n,k+1}} \min(\|S(m) - S(t_{n,k})\|, \|S(m)-S(t_{n,k+1}))<\epsilon$$

$$\max_{k<2^r, t_{n,k}<m\le t_{n,k+1}} \min(\|T(m) - T(t_{n,k})\|, \|T(m)-T(t_{n,k+1})\|)<\epsilon$$

and

$$\sum_{k<2^r} \|S_{nk} - T_{nk}\| < \epsilon \ , \quad L_{nk} = M_{nk} \qquad 0 \le k < 2^r$$

except on a set E of probability $< 3\epsilon$.

Let $\omega \in E^c$ and $m \le k_n$ be given choose k sotthat $t_{nk} < m \le t_{n,k+1}$ we want to show that

$$\|S(m) - T(m)\| < 8\epsilon \ .$$

Suppose first that $\|S_{nk}(\omega)\| \le 5\epsilon$. If for all $(m\|T(m) - T(t_{nk})\| < \epsilon$, then

$$\|S(m) - T(m)\| \le \sum_{k<2^r} \|S_{nk} - T_{nk}\| + \|S(m) - S(t_{nk})\| + \|T(m) - T(t_{nk})\| \ .$$

Note : $\|S(m)-S(t_{nk}) \le \|S(m)-S(t_{nk+1})\| + \|_{nk}\|$ such similary for $T(m)$ but $\|T_{nk}\| < \epsilon$ as $\|T(m) - T(t_{nk})\| < \epsilon) \ . \le \epsilon + \epsilon + \|S_{nk}\| + \epsilon \le 8\epsilon$.

If jump in the 1st process at is $t_{n,k+1}$ and the second at t_{nk} then there is problem ! .

If $\|T(m) - T(t_{n,k+1})\| < \epsilon$, then we can write

$$\|S(m)-T(m)\| \le \|S(m)-S(t_{n,k+1})\| + \|T^{(m)}-T(t_{n,k+1})\|$$

$$+ \|T(t_{n,k+1}) - S(t_{n,k+1})\| < 8\epsilon$$

as above. (Here S has jump at $t_{n,k}$) .

It remains to prove the above if $\|X_{nk}\| > 5\epsilon$. By (4.15) and $\sup_n k_n \mu_n(\|X\| > \epsilon) = c(\epsilon) < \infty$, which follows from Theorem 2.16., we get

$$\sum_{0\le k\le 2^r} \sum_{i,j\in H_{nk}} P(\min(\|x_{ni}\|,\|x_{nk}\|) > \epsilon) \le c(\epsilon)2^r (p_{nj}/k_n)^2 \le c(\epsilon)$$

$$\le c(\epsilon)\ 2^r\ (p_{nj}/k_n)^2 \le c(\epsilon)\ 2^{-r+1} < \epsilon$$

(by choosing r in (4.15) large) .

Thus we can discard the set E_1 on which at least two $\|x_{ni}\|$ or $\|y_{ni}\|$ with in H_{nk} exceed 2ϵ . Thus if $\omega \in E_1^c$ then in each block exactly one of $\|x_{ni}\|$ exceeds 2ϵ and this happens at $i = L_{nk}$ and similary for $\|y_{ni}\|$ at $i = M_{nk}$. Hence on $E^c \cap E_1^c$ we have for all k , $0 \le k < 2^r$.

$$\|S(t_{nk}) - S(t_{nk} + h)\| < \epsilon \qquad 1 \le h \le L_{nk}$$

and

$$\|S(t_{n,k+1} - S(t_{nk} + h)\| < \epsilon \quad \text{if} \quad L_{nk} \le h < p_{nk} \ .$$

Analogously for $T(t_{nk} + h)$. Hence $1 \le m \le k_n$, $\exists\, k$ such that

$$\|S(m) - T(m)\| = \|S(t_{nk} + h) - T(t_{nk} + h)\| < 3\epsilon$$

using $\|S(t_{nk}) - T(t_{nk})\| < \epsilon$ on $\omega \in E^c \cap E_1^c$ for $n \ge \max(n_0, n_1)$. Hence we have proved (4.13) . By the reduction of the problem we get (4.11.2) holds with $\xrightarrow{P} 0$.

4.16. COROLLARY. Let $\{X_{nj}, j = 1,\ldots,k_n\}$ be triangular array of row wise i.i.d. r.v.'s. Let $X_n(t) = \sum_{j\le k} X_{nj}$ $t = k/k_n$ $0 \le k \le k_n$ $(X_n(0) = 0)$ with linearly interpolated in between. Then $\{X_n(t)\}$ converges to a process $\{Y(t)\}$ of stationary independent increments associated with the semigroup $\{\mu_t\}$ on $D([0,1],B)$ iff $\mathcal{L}(S_n) \Rightarrow \mu$.

In particular, CLT holds in B iff invariance Principle holds.

We note that necessary and sufficient condition for CLT to hold is that X be approximated by a simple function in CL(X) norm (Proposition 2.14). Hence if one assumes that in non-separable case one has finite-dimensional approximation in outer measure P^*, then does the CLT hold ? We shall answer this in the next section, but we first want to show that in separable case CLT holds in outer measure implies measurability of X (at least under completion). Thus the problem studied next is a proper generalization of the work on separable case and reduces to it under such hypothesis.

Let us first explain the set up in the non-separable case. Let $(A,\mathcal{G},Q)$ be a probability space and $(A^\infty,\mathcal{G}^\infty,Q^\infty)$, the countable product of $(A,\mathcal{G},Q)$ with elements $\{x_j\}$ and denote by

$$(\Omega,\mathcal{F},P) = ([0,1],\mathcal{B}[0,1],\text{Leb.}) \times (A^\infty,\mathcal{G}^\infty,Q^\infty) .$$

Here $\mathcal{G}$ is assumed to be included in the completion under Q of a countably generated σ-algebra of $\mathcal{G}$. Let

$$P^*(A) = \inf\{P(C),\ C \supseteq A,\ C \in \mathcal{F}\}$$

$$P_*(A) = \sup\{P(C),\ C \subseteq A,\ C \in \mathcal{F}\} .$$

Let $\mathcal{L}_0(\Omega,\mathcal{F},P) = \{f : f : \Omega \to [-\infty,+\infty],\ f \text{ measurable}\}$. For any $f : \Omega \to [-\infty,+\infty]$, define

$$f^* = \text{ess inf}\,\{j \in \mathcal{L}^0(\Omega,\mathcal{F},P),\ j \ge f\}$$

$$f_* = -((-f)^*) = \text{ess sup}\{g\ ;\ g \le f,\ g \in \mathcal{L}^0(\Omega,\mathcal{F},P)\} .$$

4.17. LEMMA. The function f^* exists ans is $\mathfrak{F}$-measurable. Moreover, we can take $f^* \geq f$ everywhere, for all $g : \Omega \to [-\infty, +\infty]$, $(f+g)^* \leq f^* + g^*$ a.s. and $(f-g)^* \geq f^* - g^*$ a.s. if both sides are defined a.s.

Proof : Define $L_o(\Omega, \mathfrak{F}, P)$, the equivalence classes in $\mathcal{L}_o(\Omega, \mathfrak{F}, P)$ with metric

$$d(f,g) = \inf\{\varepsilon > 0 \ ; \ P(|\tan^{-1} f - \tan^{-1} g| > \varepsilon) < \varepsilon\} \ .$$

Then $(L_o(\Omega, \mathfrak{F}, P), d)$ is a separable metric space and hence ess inf $(\mathcal{J})$ for $\mathcal{J} \subseteq L_o(\Omega, \mathfrak{F}, P)$ can be written as $\min_{k \leq n} j_k \downarrow$ ess inf $(\mathcal{J})$ with $\{j_k\}$ dense subset of $L_o(\Omega, \mathfrak{F}, P)$. Thus f^* is measurable and by construction, the other properties follow.

Let $(S, \| \ \|)$ be a Banach space and h be a map (not necessarly measurable) of $(A^\infty, \mathfrak{A}^\infty, Q^\infty)$ into S. We call $X_j = h(x_j)$ a sequence of independent identically formed (i.i.f.) elements.

4.18. THEOREM. Let $X_n = h(x_n)$ $n = 1,2,\dots$ be i.i.f. elements. Let

$$\lim_{n \to \infty} P^*(X_1 + \dots + X_n/\sqrt{n} \leq t) = \lim_{n \to \infty} P_*(X_1 + \dots + X_n/\sqrt{n} \leq t)$$

$$= \gamma(-\infty, t] \quad \forall \ t \in \mathbb{R} \ .$$

where γ is $N(0,1)$ r.v. Then h is measurable for the completion of $\mathfrak{A}$ under $\mathcal{L}(x_1)$. So X_i are measurable and $EX_i = 0$, $EX_i^2 = 1$.

For this we need the following lemma. Its proof is presented in the appendix.

4.19. LEMMA. Let $(A_j, \mathfrak{A}_j, P_j)$ be probability spaces such that $\mathfrak{A}_j$ is the completion of a contably generated σ-algebra. Let $f_j : A_j \to [0,\infty]$ be any functions $j = 1,2,\dots,n$. Then on $\prod_{j=1}^n (A_j, \mathfrak{A}_j, P_j)$ with co-ordinate functions (x_j)

$$\left(\prod_{j=1}^n f_j(x_j)\right)^* = \prod_{j=1}^n f_j^*(x_j) \quad \text{a.s.}$$

where $0.\infty = 0$. If $n = 2$, $f_1 = 1$ then the same holds for f_2.

<u>Proof of theorem 4.18.</u> As X_n are non-measurable we consider

$$X_{n*} \leq X_n \leq X_n^* \quad .$$

Let $D = \{X_1^* = \infty\}$ then D is measurable. If $P(D) > 0$ then $P|_D$ is non-atomic, as

$$\{X_1^* + \cdots + X_n^*/\sqrt{n} \leq t\} \subseteq \{X_1 + \cdots + X_n/\sqrt{n} \leq t\} \; .$$

Define on D, $Y_1 \geq 0$ (finite-valued) such that $P(Y_1 \geq nM_n + 2n) \geq n^{-\frac{1}{2}}$, where $M_n \uparrow \infty$ are chosen so that $P(X_1^* \leq -M_n) \leq n^{-3}$. This is possible as $X_1^* > -\infty$. Since $P(\min_{j\leq n} X_j^* \leq -M_n) \leq n^{-2}$, by Borel-Cantelli we get for n large, $X_j^* \geq -M_n$ for all $j \leq n$. Thus $\sum_{1\leq j\leq n} X_j^* \geq -nM_n$. We define off D, $Y_1 = X_1^* - 1$. Repeatedly , we can define Y_j from X_j, then they are independent.

$$P\{\max_{1\leq j\leq n} Y_j \geq nM_n + 2n\} \geq 1 - (1 - n^{-\frac{1}{2}})^n \longrightarrow 1 \quad .$$

Hence for n large, there exists a j with $Y_j \geq nM_n + 2n$. Thus on D (by non-negatively) and off D (as $\sum_{j=1}^n X_j^* \geq -nM_n$) we get

$$\sum_{j=1}^n Y_j \geq n \quad .$$

But $Y_j < X_j^*$ and hence by Lemma 4.19. and independence

$$P^*(X_j \geq Y_j \,,\; j = 1,2,\ldots) = 1$$

Hence $P^*\left(\frac{X_1 + \cdots + X_n}{n^{\frac{1}{2}}} \geq n^{\frac{1}{2}}\right) = 1$, contraditing the assumption unless $P(D) = 0$. Let $D(j) = \{X_j \geq X_j^* - 2^{-j}\}$. Then $P^*(D(j)) = 1$. Apply Lemma 4.19. with $P_j = \mathcal{L}(x_j)$, $f_j = 1_{D(j)}$. Then $P^*(\bigcap_{j=1}^n D(j)) = 1$. On $\bigcap_{j=1}^n D(j)$

$$X_1 + \cdots + X_n/\sqrt{n} \leq X_1^* + \cdots + X_n^*/\sqrt{n} \leq X_1 + \cdots + X_n/\sqrt{n} + 1/\sqrt{n} \quad .$$

Hence X_1^* satisfies CLT giving $EX_1^* = 0$. Similar arguments give $EX_{1*} = 0$. Now $X_1^* - X_{1*} \geq 0$ gives $X_1^* = X_{1*} = X_1$ a.e. completing the proof.

4.20 COROLLARY. If $S = B$ is a separable Banach space and $X_j = h(x_j)$ satisfy CLT as above then $\{X_j\}$ are completion measurable for Borel subsets of B.

Proof : Since $\langle y, X_j \rangle$ satisfies CLT with $\gamma = N(0, \sigma_y^2)$, $(\sigma_y^2 > 0)$ for $y \in B'$ we get that $\langle y, \cdot \rangle$ are measurable with respect to

$$\mathfrak{F}_o = \{C \; ; \; h^{-1}(C) \text{ is measurable for } \mathcal{L}(x_1) \text{ completion of } \mathcal{C}\} .$$

But B is separable, $\mathcal{B}(B) = \sigma\{\langle y, \cdot \rangle \; ; \; y \in B'\}$, giving the conclusion.

APPENDIX

Proof of Lemma 4.19. Clearly, $(\prod_{j=1}^{n} f_j)^* \le \prod_{j=1}^{n} f_j^*$. For the converse, take $n = 2$ and suppose g is measurable on $A_1 \times A_2$ and for $\epsilon > 0$

$$C(\epsilon) = \{(x,y) \; ; \; g(x,y) + \epsilon < f_1^*(x) \, f_2^*(y)\} .$$

Suppose $(P_1 \times P_2)(C(0)) > 0$. Then for some $\epsilon > 0$ $(P_1 \times P_2)(C(\epsilon)) > 0$. Fix such ϵ. For $m = 1,2,\ldots$, let $B_m = \{y : m < f_2^*(y) < \infty\}$. Then for some m, $P_1 \times P_2(C(\epsilon) \setminus A_1 \times B_m) > 0$. Fix such m and let $D = C(\epsilon) \setminus (A_1 \times B_m)$, $D_x = \{y \; ; \; (x,y) \in D\}$ and $H = \{x \; ; \; P_2(D_x) > 0\}$. Suppose $f_1(x) f_2(y) \le g(x,y)$ everywhere. Let $x \in H$, if $f(x) = +\infty$, then $f_2 \ge 0$ and P_2-almost all $y \in D_x$, $f_1(x) \, f_2(y) < f_1^*(x) \, f_2^*(y)$ so $f_2(y) = 0 = f_2^*(y)$, a contradiction. If $0 < f_1(x) < \infty$, then for P_2-almost all $y \in D_x$, $f_2^*(y) \le g(x,y)/f_1(x)$ so

$$f_2^*(y) < (f_1^*(x) \, f_2^*(y) - \epsilon)/f_1(x)).$$

Then $f_2^*(y) < +\infty$, so $f_2^*(y) \le m$. If $f_2^*(y) \le 0$, we get a contradiction since $f_1^*(x) \ge f_1(x) > 0$. So for any such y, $0 < f_2^*(y) \le m$ and $f_1(x) < f_1^*(x) - \epsilon/m$. If $f_1 = 1$ this is a contradiction and finishes proof for this case. In case $f_j \ge 0$, $j = 1,2,\ldots$, we have

$$f_1(x) \le \max(0, f_1^*(x) - \epsilon/m)$$

for all $x \in H$. If $f_1^* > 0$ on some subset J of H with $P_1(J) > 0$, this allows f_1^* to be chosen smaller, a contradiction. So $f_1 = f_1^* = 0$ a.e. on H, but then $0 \le g < 0$ on D again a contradiction. For $n \ge 3$, use indution.

5. CLT AND INVARIANCE PRINCIPLES FOR SUMS OF BANACH SPACE VALUED RANDOM ELEMENTS AND EMPIRICAL PROCESSES.

Throughout the section we shall use the notations f^*, f_*, P^*, P_* as in the last Section. In order to induce the reader to familiarise with these, we state the following Lemma whis is immediate from Lemma 4.17.

5.1. LEMMA. Let $(S, \| \ \|)$ be a vector space with norm $\|\cdot\|$. Then for, $X, Y : \Omega \longrightarrow S$,

$$\|X + Y\|^* \le (\|X\| + \|Y\|)^* \le \|X\|^* + \|Y\|^* \quad \text{a.s.}$$

and

$$\|c\, X\|^* = |c|\ \|X\|^* \quad \text{a.s. for all } c \in \mathbb{R}\ .$$

Also we state the following consequence of Lemma 4.19.

5.2. LEMMA. Let $(\Omega, \mathfrak{F}, P) = (\Omega_1 \times \Omega_2 \times \Omega_3, \mathfrak{F}_1 \times \mathfrak{F}_2 \times \mathfrak{F}_3, P_1 \times P_2 \times P_3)$ and denote the projections $\pi_i : \Omega \longrightarrow \Omega_i (i = 1,2,3)$. Then for any bounded non-negative function f,

$$E\{f^*(\omega_1,\omega_3) | (\pi_1,\pi_2)^{-1}(\mathfrak{F}_1 \times \mathfrak{F}_2)\} = E\{f^*(\omega_1,\omega_3) | \ \pi_1^{-1}(\mathfrak{F}_1)\}$$

a.s. P

Proof : By Lemma 4.19 (with $f_2(\omega_2) = 1$), $f^*(\omega_1,\omega_3)$ equals P-a.e. a measurable function not depending on ω_2 and thus is independent of $\pi_2^{-1}(\mathfrak{F}_2)$.

For not necessarily measurable real-valued functions g_n on Ω, we say that $g_2 \xrightarrow{P} 0$ if $\lim\limits_{n\to\infty} P^*(|g_n| > \epsilon) = 0$, $\forall\, \epsilon > 0$ and $g_n \longrightarrow 0$ in L_p if there exists $\{f_n, n \ge 1\}$, f_n measurable $f_n \ge |g_n|$ and $f_n \longrightarrow 0$ in L_p.

5.3. LEMMA. Let $X : \Omega \longrightarrow \mathbb{R}$. Then for all $t \in \mathbb{R}$ and $\epsilon > 0$.

$$P^*(X \ge t) \le P(X^* \ge t) \le P^*(X \ge t-\epsilon)\ .$$

In particular, for any $X_n : \Omega \longrightarrow \mathbb{R}$, $X_n \xrightarrow{P} 0$ or in L_p iff $|X_n|^* \xrightarrow{P} 0$ or in L_p, respectivelly.

Proof : Since $\{X \geq t\} \subseteq \{X^* \geq t\}$, it remains to prove the last inequality. Let for $j \in Z$

$$C_j = \{\omega : X \geq j\epsilon\} \quad \text{and} \quad D_j \supseteq C_j$$

be measurable such that $P^*(C_j) = P(D_j)$. W log, D_j in non-increasing. Since $X(\omega) > -\infty$ we get $\bigcup_j D_j = \bigcup_j C_j = \Omega$. Let

$$Y(\omega) = (j+1)\epsilon \quad \text{on} \quad D_j \setminus D_{j+1} \quad \text{for } j \in Z .$$

$$= +\infty \quad \text{on} \quad \bigcap_j D_j .$$

We claim that $X^*(\omega) \leq Y(\omega)$. To prove the claim, we observe that the result is true for $\{\omega : Y(\omega) = +\infty\}$. If $\omega \in D_j \setminus D_{j+1}$ for some j, then $\omega \notin C_{j+1}$. Hence $Y(\omega) = (j+1)\epsilon$ exceeds $X(\omega) < (j+1)\epsilon$. Thus $X(\omega) \leq Y(\omega)$ and Y measurable giving $X^*(\omega) \leq Y(\omega)$. Given $t \in \mathbb{R}$, there exists unique $j \in Z$ such that

$$j\,\epsilon \leq t < (j+1)\,\epsilon .$$

Thus

$$P(X^* \geq t) \leq P(X^* \geq j\epsilon) \leq P(Y \geq j\epsilon) .$$

But $\{Y \geq j\epsilon\} = D_{j-1}$. Thus

$$P(D_{j-1}) = P^*(D_{j-1}) = P^*(X > (j-1)\epsilon)$$

$$\leq P^*(X \geq t - 2\epsilon) .$$

The following lemma is an immediate extension of the classical theorem. Hence we indicate only the changes needed in the classical proof as is given for example in Breiman.

5.4. LEMMA. (Ottaviani Inequality) Let $\{X_j, 1 \leq j \leq n\}$ be an independent sequence of random elements where X_j takes values in a normed-vector space $(S, \| \ \|)$. Write $S_n = \sum_{j \leq n} X_j$ and suppose that $\max_{j \leq n} P(\|S_n - S_j\|^* > \alpha) = c < 1$. Then $P(\max_{j \leq n} \|S_j\|^* > 2\alpha) \leq (1-c)^{-1} P(\|S_n\|^* > \alpha)$.

Proof : In the classical proof, replace $|\ |$ by $\|\ \|^*$ using Lemmas 4.17. and 5.1. One really needs $\|S_j\|^* \le \|S_n\|^* + \|S_n - S_j\|^*$. To complete the argument involving independence we argue as follows.

Let $\omega_1 = (x_{j+1},\dots,x_n)$ and $\omega_2 = (x_1,\dots,x_j)$. Then $F(\omega_1,\omega_2) = S_n - S_j$ depends only on ω_1 and by Lemma 4.21., $\|S_n - S_j\|^*$ depends only on ω_1 and is thus independent of $\{j^* - j\}$ (j^* stopping time in the usual way). The remaining parts are as before.

The following lemma is also technical and hence we defer the proof to the appendix.

5.5. LEMMA. Let S and T be Polish spaces and λ be a law on $S \times T$ with marginal μ on S. Let $(\Omega,\mathcal{B},P)$ be a probability space and X a r.v. on Ω with values in S and $\mathcal{L}(X) = \mu$. Assume that $\exists$ a r.v. U on Ω independent of X with values in a separable metric space R and $\mathcal{L}(U)$ on R being atomless. Then there exists $Y : \Omega \longrightarrow T$ a r.v. such that $\mathcal{L}(X,Y) = \lambda$.

5.6. THEOREM. Let $\{X_j,\ j \ge 1\}$ be a sequence of independent identically formed S-valued random elements $X_j = h(x_j)$, $(j \ge 1)$. Suppose that for each $m \ge 1$ there is a mapping $\Lambda_m : S \longrightarrow S$ with the following properties

(5.6.1) The linear span $L_m S$ of $\Lambda_m S$ is finite-dimensional

(5.6.2) For each $m \ge 1$, $\exists$ $n_o = n_o(m)$ so that for all $n \ge n_o$

$$P^*\{n^{-\frac{1}{2}} \| \sum_{j\le n} (X_j - \Lambda_m X_j)\| \ge \frac{1}{m} \} \le \frac{1}{m}$$

(5.6.3) For each $m \ge 1$, the mapping $\Lambda_m \circ h$ is measurable from $(A,\mathcal{A})$ into $L_m S$.

(5.6.4) $E\,\Lambda_m X_1 = 0$, $E\|\Lambda_m(X_1)\|^2 < \infty$, $\forall\ m \ge 1$.

Let T be the completion of the linear span of $\bigcup_{m\ge 1} \Lambda_m(S)$, so that T is a separable Banach space. Then there exists a sequence $\{Y_j,\ j \ge 1\}$ of i.i.d. T-valued Gaussian r.v.'s. defined on $(\Omega,\mathfrak{F},P)$ such that

(5.6.5) $\quad EY_1 = 0$

(5.6.6) $\quad E<y,Y_1><y',Y_2> = \lim_{m\to\infty} E<y,\Lambda_m(X_1)><y',\Lambda_m(X_1)>\ , \quad \forall\ y,y'\in T^*$

as $n\to\infty$.

(5.6.7) $\quad n^{-\frac{1}{2}} \max_{k\le n} \|\sum_{j\le k}(X_j - Y_j)\| \to 0$ in Probability and in L^p for $p < 2$.

Proof : We first show the desired Gaussian limit. Let $k,m,r \ge 1$. Consider i.i.d. vectors $\{(\Lambda_k X_j, \Lambda_m X_j, \Lambda_r X_j),\ j\ge 1\}$. Let $0<\epsilon<\frac{1}{2}$ fixed by (5.6.2) we get for $k,m \ge 6/\epsilon$ and $\forall\ n \ge n_o(k) \vee n_o(m)$

(5.6.8) $\quad P\{n^{-\frac{1}{2}} \|\sum_{j\le n}(\Lambda_k X_j - \Lambda_m X_j)\| > \epsilon/2\} < \epsilon/2$.

Let

$$U_{n,kmr} = n^{-\frac{1}{2}} \sum_{j\le n}((\Lambda_k X_j, \Lambda_m X_j, \Lambda_r X_j)\ .$$

and for $(u,v,w) \in L_k S \times L_m S \times L_r S$,

$$\|(u,v,w)\| = \|u\| + \|v\| + \|w\|\ .$$

By CLT there exists μ_{kmr} on $L_k S \times L_m S \times L_r S$ centered Gaussian so that

(5.6.9) $\quad \pi(\mathcal{L}(U_{nkmr}), \mu_{kmr}) < \epsilon/2\ , \quad n \ge n_1(\epsilon,k,m,r)$.

Let $\mu_{km}, \mu_{kr}, \mu_{mr}, \mu_k, \mu_m, \mu_r$ be the marginals of μ_{kmr} . Now $\mu_k, \mu_{km}, \mu_{kmr}$ can be regarded as Borel probability measures on $T, T\times T$ and $T\times T\times T$. Now (5.6.8) for m,r implies

(5.6.10) $\quad \mu_{mr}\{(v,w)\in T\times T\ ;\ \|v-w\| > \epsilon\} < \epsilon\ , \quad m,r > \frac{6}{\epsilon}$.

On $T\times T$ we take $\|(u,v)\| = \|u\| + \|v\|$. We rewrite the above as

$$\mu_{kmr}\{(u,v,w) : \|(u,v) - (u,w)\| > \epsilon\} < \epsilon, m,r \ge 6/\epsilon\ ,\ k \ge 1$$

and obtain that

$$\pi(\mu_{jm},\mu_{kr}) \le \epsilon \;, \quad m,r \ge 6/\epsilon \quad k \ge 1 \;.$$

Hence $\{\mu_{km}\}_{m\ge 1}$ and each $k \ge 1$ is a Cauchy seuqence for the Prohorov metric. Hence $\exists\ \mu_{k\infty}$ on $T \times T$ such that

$$\mu_{km} \Rightarrow \mu_{k\infty} \quad \text{as} \quad m \to \infty \;.$$

By (5.6.10),

(5.6.11) $\quad \mu_{k\infty}\{(u,v);\ \|u-v\| > \epsilon\} \le \epsilon \;, \quad \forall\ k \ge 6/\epsilon \;.$

As marginal of μ_{km} is μ_m, we get that there exists μ_∞ on T such that

$$\mu_m \Rightarrow \mu_\infty \quad \text{as} \quad m \to \infty \;.$$

Further, μ_{km} has marginals μ_k and μ_m we conclude that $\mu_{k\infty}$ is Gaussian with marginals μ_k and μ_∞.

For $k \ge 1$, fixed, let $\{(Z_{kj},Z_j)\ j \ge 1\}$ be a sequence of i.i.d. random vectors on Ω' with values in $T \times T$

$$\mathcal{L}(Z_{kj},Z_j) = \mu_{k\infty} \quad j \ge 1 \quad (\text{Note } \{Z_j\} \text{ depends on } \epsilon) \;.$$

Now $\mu_{k\infty}$ is centered Gaussian gives by (5.6.11)

$$P\{n^{-\frac{1}{2}} \|\sum_{j\le n} (Z_{kj} - Z_j)\| > \epsilon\} \le \epsilon \qquad k \ge 6/\epsilon \;.$$

By Lévy inequality $n \ge 1$

$$P\{n^{-\frac{1}{2}} \max_{m\le n} \|\sum_{j\le m} (Z_{kj} - Z_j)\| > \epsilon\} \le 2\epsilon \;.$$

Let $k > 6/\epsilon$, then $\{\Lambda_k X_j,\ j \ge 1\}$ satisfies CLT with limit μ_k. Hence by Section 4, there exists Ω'' and a sequence $\{V_{kj},\ j \ge 1\}$ of independent r.v.'s., having the same distribution as $\{\Lambda_k X_j\ ,\ j \ge 1\}$ and a sequence $\{W_{kj}\}$ of i.i.d. r.v.'s. with common distribution μ_k such that

$$n^{-\frac{1}{2}} \max_{m\le n} \left\| \sum_{j\le m} (V_{kj} - W_{kj})\right\| \xrightarrow{P} 0 \quad .$$

By Lemma 4.12 $(m = 2)$, we can assume $\Omega' = \Omega''$ and $Z_{kj} = W_{kj}$ for all j . Hence we get for some $n_2(\epsilon,k) \ge n_o(k)$ and $n \ge n_2(\epsilon,k)$,

$$P\{n^{-\frac{1}{2}} \max_{m\le n} \left\| \sum_{j\le m} (V_{kj} - Z_j)\right\| > 3\epsilon\} < 3\epsilon \quad .$$

(Note that Z_j depends on $k \ge 6/\epsilon$, i.e. on ϵ).

Let us overcome this problem. Choose $\epsilon = \epsilon_p = 2^{-p-3}$ $p = 1,2,\ldots$ and $k = k(p) = 2^{p+6} > 6/\epsilon_p + 1$. By what has been proved we obtain two sequences

$$\{V_j^{(p)}, \ j \ge 1\} \quad \text{and} \quad \{Z_j^{(p)}, \ j \ge 1\}$$

with the following properties

$V_j^{(p)} = V_{k(p)j}$ $j \ge 1$, $\mathcal{L}(\{Z_j^{(p)}, \ j \ge 1\}) = \mathcal{L}(\{Z_j \ , \ j \ge 1\})$ and for some $n_3(p) \ge n_2(2^{-p-6}, k(p))$ and $n \ge n_3$

(5.6.12) $$P\{n^{-\frac{1}{2}} \max_{m\le n} \left\| \sum_{j\le m} (V_j^{(p)} - Z_j^{(p)})\right\| > 2^{-p}\} < 2^{-p} \quad .$$

We can assume V-sequences are independent of each others and Z-sequences. Put $r(p) = \sum_{q\le p} n_3(q)$.

Define

(5.6.13) $$V_j = V_j^{(p)} \quad \text{and} \quad Z_j' = Z_j^{(p)} \quad \text{if} \quad r(p) < j \le r(p+1) \ .$$

Then $\{V_j \ , \ j \ge 1\}$, $\{Z_j' \ , \ j \ge 1\}$ are sequences of independent r.v.'s. Moreover, for $\epsilon > 0$, there exists $n_4(\epsilon)$ such that

(5.6.14) $$P(n^{-\frac{1}{2}} \max_{m\le n} \left\| \sum_{j\le m} (V_j - Z_j')\right\| > 4\epsilon) < 4\epsilon \quad .$$

We now prove (5.6.14) to get rid of dependence of Z_j on ϵ .

Let s be such that $2^{-s} < \epsilon$ and $N_o = N_o(\epsilon)$ be so large that for all $n \ge N_o$, (as s is fixed)

$$P\{n^{-\frac{1}{2}} \max_{m\le r(s)} \|\sum_{j\ge m} V_j\| > \epsilon\} < \epsilon$$

and

$$P\{n^{-\frac{1}{2}} \max_{m\le r(s)} \|\sum_{j\le m} Z_j'\| > \epsilon\} < \epsilon \quad .$$

Let $n \ge \max(N_o, n_3(s)) = n_4(\epsilon)$. Choose M so that $r(M) < n \le r(M+1)$. Then $n \ge n_3(p)$, $p \le M$ by définition of $r(M)$. By (5.6.12) and (5.6.13) , we get

$$\max_{m\le n} \|\sum_{j\le m} (V_j - Z_j')\| \le \max_{m\le r(s)} \|\sum_{j\le m} V_j\|$$

$$+ \max_{m\le r(s)} \|\sum_{j\le m} Z_j'\|$$

$$+ \sum_{p=s}^{m-1} \max_{r(p)<m\le r(p+1)} \|\sum_{j=r(p+1)}^{m} (V_j - Z_j')\|$$

$$+ \max_{r(M)<m\le n} \|\sum_{j=r(M+1)}^{m} (V_j - Z_j')\|$$

$$\le 2\epsilon\, n^{\frac{1}{2}} + \sum_{p=s}^{M} 2^{-p}\, n^{\frac{1}{2}} = 4\epsilon\, n^{\frac{1}{2}}$$

by (5.6.12) . This holds except on a set of measure $< 4\epsilon$ giving (5.6.14) .

Now we want to show that $\{X_j \, , \, j \ge 1\}$ and $\{Z'_j, \, j \ge 1\}$ are defined on on the same probability space. For this we need Lemma 5.5. For $j \ge 1$, define $p(j)$ such that $j \in (r(p), r(p+1)]$ and $\rho(j) = 2^{p(j)+6}$. Then

$$\mathcal{L}(\{\Lambda_{\rho(j)} X_j \, , \, j \ge 1\}) = \mathcal{L}(\{V_j \, , \, j \ge 1\})$$

by construction. In Lemma 5.5.

$$\lambda = \mathcal{L}(\{V_j, \, j \ge 1\}, \{Z_j' \, , \, j \ge 1\}), \; X = \{\Lambda_{\rho(j)} X_j \, , \, j \ge 1\}$$

and U uniform. Then by the above equality of the law and independence of uniform $[0,1]$ and X, we get existence of $\{Y_j \, , \, j \ge 1\}$ defined on Ω such that

$$\lambda = \mathcal{L}(\{\Lambda_{\rho(j)} X_j, \, j \ge 1\} \, , \{Y_j \, , \, j \ge 1\}) \, .$$

Thus by (5.6.14) we get as $n \to \infty$

$$(5.6.15) \quad n^{-\frac{1}{2}} \max_{m \le n} \left\| \sum_{j \le m} (\Lambda_{\rho(j)} X_j - Y_j) \right\| \xrightarrow{P} 0 .$$

Since $n_3(p) \ge n_2(\epsilon_p , k(p)) \ge n_o(2^{p+6})$ for $p \ge 1$. By (5.6.2) we have

$$P^*\{n^{-\frac{1}{2}} \left\| \sum_{j \le n} X_j - \Lambda_{k(p)} X_j \right\| \ge 2^{-p-6}\} \le 2^{-p-6} .$$

By Ottavani Inequality and Lemma 5.3., for $n \ge n_3(p)$

$$P\{n^{-\frac{1}{2}} \max_{k \le n} \left\| \sum_{j \le k} (X_j - \Lambda_{k(p)}) \right\| > 2^{-p}\} < 2^{-p} .$$

This is analogue of (5.6.12). Following proof as for (5.6.14), we get for $\epsilon > 0$ and some $n_5(\epsilon)$ and $n \ge n_5(\epsilon)$

$$n^{-\frac{1}{2}} \max_{m \le n} \left\| \sum_{j \le m} (X_j - \Lambda_{\rho(j)} X_j) \right\| \xrightarrow{P} 0 , \quad \text{as } n \to \infty .$$

Combining with (5.6.15) we get the result in terms of convergence in probability. As in the proof of Proposition 2.14., $\sup_\lambda \lambda^2 P\{n^{-\frac{1}{2}} \|S_n\|^* > \lambda\} < \infty$.

Since

$$P\{\|n^{-\frac{1}{2}} S_n\|^* > \lambda\} \ge P\{\max_{k \le n} n^{-\frac{1}{2}} \|S_k\|^* > \lambda\}$$

with $S_k = \sum_{j \le k} X_j$, we get for $p < 2$. Using Fernique's theorem

$$n^{-\frac{1}{2}} \max_{k \le n} \left\| \sum_{j \le k} (X_j - Y_j) \right\|^{*p}$$

is uniformly integrable. Hence convergence in L_p follows for $p < 2$.

Also $E\{\langle s, \Lambda_k X_j \rangle^2\} = E\{\langle s, Z_{k1} \rangle^2\}$, $s \in T'$ as $\Lambda_k X_j$ satisfies CLT with limit μ_k. As $\mu_k \Rightarrow \mu_\infty$ Gaussian, we have $E\langle s, Z_{k1} \rangle^2 \to E\langle s, Z_1 \rangle^2$ as $k \to \infty$. But $E\langle s, Z_1 \rangle^2 = E\langle s, Y_1 \rangle^2$ proving (5.6.5) and (5.6.6).

Let us now apply the theorem to empirical processes. Let $\{x_j\}$ be a sequence of i.i.d. uniform r.v.'s. and h be a map on $[0,1] \to (D([0,1], \|\cdot\|)$

given by $1_{[0,s]}(\cdot) - s$ for $0 \le s \le 1$. Then $X_j(s) = 1(X_j \le s) - s$ and

$F_n(s) = n^{-1} \sum_{j=1}^{n} 1(X_j \le s)$ is called empirical distribution function. We get

$n^{-\frac{1}{2}} \sum_{j=1}^{n} X_j(s) = n^{\frac{1}{2}} (F_n(s) - s)$.

The classical result says that $\mathcal{L}(n^{\frac{1}{2}}(F_n(\cdot) - \cdot)) \Rightarrow \mathcal{L}(W_o)$ in the supremum norm on $D[0,1]$ where $W_o(s) = W(s) - s\,W(1)$, the Brownian Bridge, W being Wiener process.

In general, if $\{x_j\}$ are i.i.d. r.v. and $B \in \mathcal{Q}$, we can define empirical measure by

$$Q_n(B) = n^{-1} \sum_{j=1}^{n} 1(x_j \in B)$$

and the following gives analogue of the above result.

5.7. THEOREM. *Let* $\mathcal{G} \subseteq \mathcal{L}_2(A,\mathcal{Q},Q)$ *be a class of functions so that*

(5.7.1) $\mathcal{G}$ *is totally bounded in* $\mathcal{L}_2$.

For every $\epsilon > 0$, *there exists* $\delta > 0$ *such that for all* $n \ge n_o$,

(5.7.2) $P^*(\sup\{|\int (f-g)\,d\nu_n| : f,g \in \mathcal{G},\ \int (f-g)^2 dP < \delta^2\} > \epsilon) < \epsilon$.

Then there exists a sequence $\{Y_j, j \ge 1\}$ *of i.i.d. Gaussian processes defined on* Ω *indexed by* $f \in \mathcal{G}$ *and sample functions of* Y_1 *are a.s. uniformly continuous on* $\mathcal{G}$ *in* $\mathcal{L}_2$*-norm such that*

a) $E\,Y_1(f) = 0$ *for all* $f \in \mathcal{G}$.

b) $EE\,Y_2(f)\,Y_1(g) = \int fg\,dQ - \int f\,dQ \int g\,dQ$ *for all* $f,g \in \mathcal{G}$.

and as $n \to \infty$.

c) $n^{-\frac{1}{2}} \max_{k\le n} \sup_{f\in\mathcal{G}} |\sum_{j\le k} [f(x_j) - \int f\,dQ - Y_j(f)]| \xrightarrow{P} 0$

as well as in L_p, $p < 2$.

We observe now how Theorem 5.7. can be put in the form of Theorem 5.6. Let $m \ge 1$ and $\epsilon = \frac{1}{m}$. Choose δ and n_o according to (5.7.2). Let

$$\|f-g\|_{2,Q} = [\int (f-g)^2 dQ]^{\frac{1}{2}} \quad , \; f,g \in \mathcal{G} \quad .$$

Since $\mathcal{G}$ is totally bounded in $\| \ \|_{2,Q}$ there exist $f_k = f_{km} \in \mathcal{G}$, $1 \le k \le N(\delta)$ such that for $f \in \mathcal{G}$, there exists a $k(f)$, with $\|f-f_k\|_{2,Q} < \delta$. Choose $k = k(f)$ minimal. Hence by (5.7.2) and definition of empirical measure, we get

$$P^*\{n^{-\frac{1}{2}} \sup_{f \in \mathcal{G}} |\sum_{j \le n} (f-f_k)(x_j) - \int (f-f_k)\, dQ| > 1/m\} < \frac{1}{m} \quad .$$

Now set S as the space of all bounded real-valued functions on $\mathcal{G}$. Define for $\Psi \in S$

$$\|\Psi\| = \{|\Psi(f)| \; ; f \in \mathcal{G}\} \quad .$$

Then $(S,\|\cdot\|)$ is a Banach space (not necessarily separable).

Define $h : A \longrightarrow S$ by $h(x)(f) = f(x) - \int f dQ$ for $x \in A$ and $\wedge_m : S \longrightarrow S$ by setting

$$\wedge_m \Psi(f) = \Psi(f_k) \; .$$

Let $X_j = h(x_j)$. Then

$$(\wedge_m X_j)(f) = (\wedge_m h(x_j)f) = f_k(x_j) - \int f_k dQ \quad , \quad f \in \mathcal{G} \quad .$$

Now $\dim L_m(S) = N(\delta) < \infty$ and WLOG assume $\delta(\epsilon) \downarrow$ as $\epsilon \downarrow$. Clearly assumptions of Theorem 5.6. are satisfied. Now $(T,\|\cdot\|)$ be as in that theorem. Then there exist i.i.d. Gaussian T-valued Y_j satisfying a),b),c), of Theorem 5.7. by Theorem 5.6., if we show Y_1 has uniformly continuous sample paths on $\mathcal{G}$ for $\| \ \|_{2,p}$ (for a),b)) .

Let $Z_n = n^{-\frac{1}{2}} (Y_1 + \dots + Y_n)$, then $\mathcal{L}(Z_n) = \mathcal{L}(Y_1)$ on T and $\|Z_n - \nu_n\| \xrightarrow{P} 0$. Given $\epsilon > 0$, take $\delta(\epsilon) > 0$ and n_o from (5.7.2) s.t. for $n \ge n_o$

$$P^*(\|Z_n - \nu_n\| > \epsilon) < \epsilon \quad .$$

For $\Psi \in S$, let

$$p_\delta(\Psi) = \sup\{|\Psi(f) - \Psi(g)|, \; f, \, g \in \mathcal{G} \; , \; \|f-g\|_{2,Q} < \delta\}$$

Then p_δ is a seminorm on S with $p_\delta(\Psi) \le 2\|\Psi\|$ for all $\Psi \in S$ and by (5.7.2).

$$P^*\{p_\delta(\nu_n) > \epsilon\} < \epsilon \quad \text{for } n \ge n_o .$$

Thus

$$P^*\{p_\delta(Z_n) > 3\epsilon\} < 2\epsilon .$$

But p_δ is continuous and hence measurable on T. As $\mathcal{L}(Z_n) = \mathcal{L}(Y_1)$, $P(p_\delta(Y_1) > 3\epsilon) < 2\epsilon$. Let $a_k = \delta(2^{-k})$ and $W_k = \{\Psi \in S ; p_{a_k}(\Psi) < 3.2^{-k}\}$. Then

$$P(Y_1 \notin W_k) < 2^{1-k} (\epsilon = 2^{-k}) .$$

Let $W = \bigcup_{j\ge1} \bigcap_{k\ge j} W_k$. Then W is a Borel set in T, consisting of functions uniformly continuous on $\mathcal{G}$ and $P(Y_1 \in W) = 1$ by Borel-Cantelli lemma.

A class $\mathcal{G}$ of functions satisfying (5.7.1) and (5.7.2) is called a Donsker Class of sets for Q. In case $\mathcal{G} = \{1_C , C \in \mathcal{C}\}$, we call $\mathcal{C}$ a Donsker Class of sets. Our purpose now is to give conditions on $\mathcal{C}$ and Q in order that $\mathcal{C}$ is a Donsker Class.

For $\delta > 0$ and $\mathcal{C} \subseteq \mathfrak{a}$, a class of sets, we define, $N_I(\delta) = N_I(\delta, \mathcal{C}, Q)$ to be the smallest number d of sets $A_1 \dots A_d \in \mathfrak{a}$ satisfying.

For each $C \in \mathcal{C}$, there exist A_r and A_s $(1 \le r,s \le d)$ such that $A_r \subset C \subset A_s$ and $P(A_s \setminus A_r) < \delta$. We call $\log(N_I(\delta)$ a metric entropy with inclusion. It is shown by Dudley (Ann. Prob. 6 (1978)) that

$$(5.8) \qquad \int_0^1 (\log N_I(x^2))^{\frac{1}{2}}\, dx < \infty$$

implies (5.7.1) and (5.7.2). Hence we get

5.9. THEOREM. *Let $\mathcal{C}$ be a class of sets for which (5.8) holds. Then there exists a sequence $\{Y_j , j \ge 1\}$ of i.i.d. Gaussian processes defined on the same probability space indexed by $C \in \mathcal{C}$ with sample functions of Y_1 a.s.*

uniformly continuous on $\mathcal{C}$ in the $d_Q(C,D) = Q(C \Delta D)$ on $\mathcal{G}$. The processes Y_j have following properties.

a) $EY_1(C) = 0$ for all $C \in \mathcal{C}$.

b) $EY_1(C)\, Y_1(D) = P(C \cap D) - P(C)P(D)$ for all $C,D \in \mathcal{C}$ and as $n \to \infty$,

c) $n^{-\frac{1}{2}} \max_{k \leq n} \sup_{C \in \mathcal{C}} \left| \sum_{j \leq k} 1(x_j \in C) - Q(C) - Y_j(C) \right| \to 0$

in probability as well as L_2.

Note $1_C \leq 1$, one gets uniform integrability $\|\ \|$ in the proof of Theorem 5.6.

A collection $\mathcal{C}$ is called Vapnik-Cervonerkis class (VCC) if for some $n < \infty$, no set D with n elements has all its subsets of the form $C \cap D$. The Vapnik-Cervonenkis number $V(\mathcal{C})$ denotes smallest such n.

5.10. DEFINITION.

a) If $(A,\mathcal{G})$ and $(\mathcal{C},\mathcal{S})$ are measurable spaces with $\mathcal{C} \subseteq \mathcal{G}$, we call $(A,\mathcal{G}\ ;\ \mathcal{C},\mathcal{S})$ a chair.

b) A chair is called admissible iff $\{(x,C) : x \in C\} \in \mathcal{G} \otimes \mathcal{S}$ for all $C \in \mathcal{C}$.

c) A chair is called a-Suslin iff it is admissible and $(A,\mathcal{G})$, $(\mathcal{C},\mathcal{S})$ are Suslin spaces.

d) A chair is called Qa-Suslin iff it is a-Suslin and d_Q-open subsets of $\mathcal{C}$ belong to $\mathcal{S}$.

If $\mathcal{C}$ is a VCC and for some σ-algebra $\mathcal{G}' \supseteq \mathcal{C}$ and σ-algebra $\mathcal{S}$ of $\mathcal{C}$ s.t. $(A,\mathcal{G}'\ ;\ \mathcal{C},\mathcal{S})$ is Qa-Suslin then $\mathcal{C}$ satisfies (5.7.1) and (5.7.2).

For proof see Dudley (cited before).

Thus one can produce large class of examples for which approximation condition (5.6.2) holds and also Theorem 5.9. holds.

Appendix : Proof of Lemma 5.5. :

Proof : We may assume R complete, hence Polish. Any uncountable Polish space is Borel isomorphic to $[0,1]$ (Parthasarathy, p. 14). Every Polish space is Borel-isomorphic to some compact subset of $[0,1]$. Thus there is no loss of generality in assuming $S = T = R = [0,1]$ with the usual topology, metric and Borel structure. Next, we take disintegration of λ on $[0,1] \times [0,1]$ (N. Bourbaki, VI, Integration p. 58-59). There exists a map λ_s from s into the set of all probability measures on T s.t. $\int f(s,t)d\lambda = \int_0^1\int_0^1 f(s,t)d\lambda_s d\mu$ for all bounded, Borel measure functions f on $[0,1] \times [0,1]$. For each s, let F_s be the distribution function of λ_s. $F_s^{-1}(t) = \inf\{z \,;\, F_s(z) \geq t\}$ for $0 \leq t \leq 1$. We may assume U has uniform distribution over $[0,1]$. For each t, the map $s \to F_s^{-1}(t)$ is measurable. Since $F_s^{-1}(1)$ is non-decreasing and left-continuous.

$$F_s^{-1}(t) = \lim_{n\to\infty} \sum_{j=0}^{n} F_s^{-1}(j/n)\ 1\{j/n \leq t \leq j+1/n\} .$$

Hence $F_s^{-1}(t)$ is jointly measurable in (s,t). Let $Y(\omega) = F_{X(\omega)}^{-1}(U(\omega))$, then Y is a r.v. Moreover, for any bounded Borel function g on $[0,1]\times[0,1]$ using Fubini Theorem and the fact $\text{leb.}O(F_s^{-1})^{-1} = \lambda_s$

$$\int g d\lambda = \int_0^1\int_0^1 g(s,t)d\lambda_s d\mu = \int_0^1\int_0^1 g(s,F_s^{-1}(t))\ dt\ d\mu$$

$$= \int_0^1\int_0^1\int_0 g(s,F_s^{-1}(t))\ d(\mu \otimes \text{leb.})$$

$$= E\ g(X,F_X^{-1}(U)) = Eg(X,Y) .$$

REFERENCES / BOOKS

[1] BILLINGSLEY P. (1968) : Convergences of Probability measures, Wiley, New York.

[2] FELLER W. (1971) : Introduction to Probability Theory and its applications, Vol. 2, Wiley, New York.

[3] LOEVE M. (1968) : Probability Theory, Van Neustrand Princeton.

[4] PARTHASARATHY K.R. (1967) : Probability measures on metric spaces, Academic Press, N. Y.

PAPERS :

[1] DE ACOSTA A. (1982) : An invariance principle in probability for triangular arrays of B-valued random vectors, Annals of Probability, 10 .

[2] DABROWSKI A., DEHLING H., PHILIPP W. (1981) : An almost sure invariance principle for triangular arrays of Banach space valued random variables (preprint).

[3] DUDLEY R.M. and PHILIPP W. (1982) : Invariance principles for sums of Banach space valued random elements and empirical processes. Preprint.

V. MANDREKAR

MICHIGAN STATE UNIVERSITY

and

UNIVERSITE DE STRASBOURG

This work was partially supported on NSF-MCS-78-02878 and AFOSR 80-0080 .