Chapter I : TYPES OF CONVERGENCE

I.1. Introduction

In this work, E will always stand for a real Banach space unless otherwise stated. Most of the results in this book are true however in complex Banach spaces without any real change. The norm on E is denoted by $\|\cdot\|$. When we say that F is a subspace of E we shall always mean a Banach subspace of E. The dual of E is denoted by E'. $(\Omega,\mathcal{F},P)$ will denote a complete probability space. This means that $\mathcal{F}$ is a σ-algebra of subsets of the set Ω and that P is a countably additive measure with $P(\Omega) = 1$ such that if $A \in \mathcal{F}$, $P(A) = 0$ and $B \subset A$ imply $B \in \mathcal{F}$.

I.1.1. Measurable functions

Let χ_A denote the characteristic function of $A \in \mathcal{F}$, i.e. the function

$$\chi_A(\omega) \begin{cases} = 0 & \text{if } \omega \notin A \\ = 1 & \text{if } \omega \in A \end{cases}$$

If $A_i \in \mathcal{F}$ for each $i = 1,\ldots,n$ and if $x_i \in E$ for each $i = 1,\ldots,n$, we say that the function

$$f = \sum_{i=1}^{n} x_i \chi_{A_i}$$

is a stepfunction. Let a function $f : \Omega \to E$ be the almost everywhere (a.e.) limit of f_n, i.e. : there exists $N \in \mathcal{F}$ such that $P(N) = 0$ and such that for each $\omega \in \Omega \setminus N$ and each $\varepsilon > 0$ there exists $n_o(\varepsilon) \in \mathbb{N}$ such that for each $n \geq n_o(\varepsilon)$:

$$\|f(\omega) - f_n(\omega)\| \leq \varepsilon$$

Then we say that f is strongly measurable or measurable, in case there cannot be any confusion. If x'(f) is measurable for each x' ∈ E', then we say that f is scalarly measurable.

Hence, in the definitions of scalar measurability, the nullset is dependent on the dual element x'. We obviously have that a strongly measurable function is scalarly measurable. The converse is not true, see Diestel and Uhl [1977], p.43. We shall often be dealing with scalar measurability as a working tool, but in this book only sequences of strongly measurable functions will be studied. A connection between strong and scalar measurability is given by Pettis' theorem.

Theorem I.1.1.1 (B.J. Pettis) : Let $f : \Omega \to E$ be a function. The following two assertions are equivalent :

(i) f is strongly measurable.

(ii) f is scalarly measurable and f is essentially separably valued. This means : there exists $A \in \mathcal{F}$, $P(A) = 0$ such that $f(\Omega \setminus A)$ is separable.

For a proof, see Diestel and Uhl [1977], p.42. Hence in a separable Banach space E the two notions coincide.

I.1.2. Integrable functions

Let $f : \Omega \to E$ be a function. We say that f is Bochner-integrable (w.r.t.P), or simply integrable, if there exists a sequence $(f_n)_{n \in \mathbb{N}}$ of stepfunctions such that

(i) $f_n \to f$, P-a.e., for $n \to \infty$

(ii) $\int_\Omega \|f_n - f_m\| dP \to 0$, for $m,n \to \infty$.

Instead of (ii) we might have required :

(ii)' $\int_\Omega \|f_n - f\| dP \to 0$, for $n \to \infty$, as is easily seen. Note that (i) implies that f is measurable. From (ii) (or (ii)') we can define the Bochner integral of f over $A \in \mathcal{F}$ (or simply integral of f over A) :

$$\int_A f\, dP = \lim_{n\to\infty} \int_A f_n\, dP$$

where, if f is a stepfunction, say

$$f = \sum_{i=1}^{n} x_i \chi_{A_i}$$

we define

$$\int_A f\, dP = \sum_{i=1}^{n} x_i\, P(A_i \cap A)$$

Theorem I.1.2.1 (S. Bochner) : Let $f : \Omega \to E$ be a function. Then f is integrable iff f is measurable and $\int_\Omega \|f\| dP < \infty$.

We also have the following easily proved result :

Theorem I.1.2.2 : Suppose that $f : \Omega \to E$ is integrable. Then $f = 0$, P-a.e. iff $\int_\Omega \|f\| dP = 0$.

Denote by $\mathcal{L}_E^1(\Omega,\mathcal{F},P)$, or $\mathcal{L}_E^1$ if there is no confusion, the set of all integrable functions. The natural operations (defined pointwise) make $\mathcal{L}_E^1$ into a vector space. It is semi-normed by the function $f \to \int_\Omega \|f\| dP$. Hence $L_E^1(\Omega,\mathcal{F},P) = \mathcal{L}_E^1(\Omega,\mathcal{F},P)|_{\mathcal{N}}$, or L_E^1 for short, is a normed space with

$$\mathcal{N} = \{f \in \mathcal{L}_E^1 | f = 0, \text{ P-a.e.}\} .$$

This follows trivially from theorem I.1.2.2. We denote the norm on L_E^1 by $\|\cdot\|_1$.

Furthermore, for $1 < p < \infty$, define L_E^p to be the space of those functions f in L_E^1 such that $\|f\|_p = (\int_\Omega \|f\|^p dP)^{1/p} < \infty$. We put on L_E^p the norm $\|\cdot\|_p$. Also, for $p = \infty$, we define L_E^∞ to be the space of those functions f in L_E^1 such that $\|f\|_\infty < \infty$. Here

$$\|f\|_\infty = \inf \{\lambda | P(\|f\| > \lambda) = 0\}$$

Also $L_E^\infty, \|\cdot\|_\infty$ is a normed space. If $1 \leq p \leq \infty$ and $E = \mathbb{R}$ we denote $L_{\mathbb{R}}^p = L^p$. We have, as is well known :

<u>Theorem I.1.2.3</u> : L^p_E is a Banach space ($1 \leq p \leq \infty$), for every Banach space E. As in the scalar case we recall that simple functions are dense in L^p_E ($1 \leq p \leq \infty$).

We note also the following result (see f.i. Dinculeanu [1967], p.187).

<u>Theorem I.1.2.4</u> : Suppose that $f \in L^1_E$. Then $f = 0$, a.e. iff $\int_A f\, dP = 0$, for each $A \in F$.

Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in L^1_E. We say that $(f_n)_{n \in \mathbb{N}}$ is <u>uniformly integrable</u> if

$$\lim_{\lambda \to \infty} \sup_{n \in \mathbb{N}} \int_{\{\|f_n\| > \lambda\}} \|f_n\| dP = 0$$

The following result is well known (see f.i. Chung [1974], p.96-97).

<u>Theorem I.1.2.5</u> : Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in L^1_E. $(f_n)_{n \in \mathbb{N}}$ is uniformly integrable iff the following two conditions are satisfied :

(i) $\sup_{n \in \mathbb{N}} \int_\Omega \|f_n\| dP < \infty$

(ii) $\lim_{P(A) \to 0} \sup_{n \in \mathbb{N}} \int_A \|f_n\| dP = 0$, i.e. : For each $\varepsilon > 0$, there is a $\delta(\varepsilon) > 0$ such that for any $A \in F$ with $P(A) \leq \delta(\varepsilon)$ we have :

$$\sup_{n \in \mathbb{N}} \int_A \|f_n\| dP \leq \varepsilon .$$

The following result is well known (see f.i. Breiman [1978], p.91-92).

<u>Theorem I.1.2.6</u> : Let $(f_n)_{n \in \mathbb{N}}$ be uniformly integrable. Suppose $f_n \to f$ a.e.. Then

$$\lim_{n \to \infty} \int_\Omega \|f_n - f\| dP = 0 .$$

From this result, the Lebesgue dominated convergence theorem is an obvious corollary.

On L_E^1 we consider a weaker norm, called the Pettis norm :

$$\|\cdot\|_{Pe} : L_E^1 \to \mathbb{R}^+$$

$$f \to \|f\|_{Pe} = \sup_{\substack{x' \in E' \\ \|x'\| \leq 1}} \int_\Omega |x'(f)| dP$$

We have $\|f\|_{Pe} \leq \|f\|_1$, but $\|\cdot\|_{Pe}$ and $\|\cdot\|_1$ are not equivalent norms, as we shall see later. That $\|\cdot\|_{Pe}$ is indeed a norm on L_E^1 is seen from the following result and theorem I.1.2.4 :

Theorem I.1.2.7 : $\|\cdot\|_{Pe}$ is equivalent to the seminorm on $\mathcal{L}_E^1$ (or norm on L_E^1)

$$f \to \sup_{A \in \mathcal{F}} \|\int_A f\, dP\|$$

Indeed :

$$\sup_{A \in \mathcal{F}} \|\int_A f\, dP\| \leq \|f\|_{Pe} \leq 2 \sup_{\substack{x' \in E' \\ \|x'\| \leq 1}} \sup_{A \in \mathcal{F}} |\int_A x'(f) dP|$$

$$= 2 \sup_{A \in \mathcal{F}} \|\int_A f\, dP\|$$

The second inequality is a standard inequality for scalar measures; see f.i. Dunford and Schwartz [1957], p.97. We have said already that $\|\cdot\|_{Pe}$ and $\|\cdot\|_1$ are not equivalent norms. In fact more is true :

Theorem I.1.2.8 : The following assertions are equivalent :

(i) $\|\cdot\|_{Pe}$ on L_E^1 is complete.

(ii) $\|\cdot\|_{Pe}$ and $\|\cdot\|_1$ are equivalent.

(iii) $\dim E < \infty$.

Proof : (iii) ⇒ (ii) ⇒ (i) is obvious.

(i) ⇒ (ii) follows from the closed graph theorem.

(ii) ⇒ (iii). Suppose $\dim E = \infty$. From the well-known Dvoretzky-Rogers theorem there exists a sequence $(x_n)_{n \in \mathbb{N}}$ in E such that Σx_n is unconditionally convergent and such that $\Sigma\|x_n\| = \infty$; see e.g. Lindenstrauss and Tzafriri [1979], p.16 or Dvoretzky and Rogers [1950]. From the unconditional convergence, it follows that :

$$\sup_{\substack{x' \in E' \\ \|x'\| \leq 1}} \Sigma |x'(x_n)| < \infty$$

Take now $\Omega = (0,1]$ with Lebesgue measure, and take for $n = 1,2,\ldots$

$$A_n = (\frac{1}{2^n}, \frac{1}{2^{n-1}}]$$

Put, for every $n \in \mathbb{N}$

$$f_n = \sum_{k=1}^{n} 2^k x_k \chi_{A_k}$$

We have

$$\sup_{n \in \mathbb{N}} \|f_n\|_{Pe} = \sup_{n \in \mathbb{N}} \sup_{\substack{\|x'\| \leq 1 \\ x' \in E'}} \int_0^1 |x'(f_n)| =$$

$$= \sup_{n \in \mathbb{N}} \sup_{\substack{\|x'\| \leq 1 \\ x' \in E'}} \sum_{k=1}^{n} |x'(x_k)| < \infty$$

while

$$\sup_{n \in \mathbb{N}} \|f_n\|_1 = \sum_{n=1}^{\infty} \|x_n\| = \infty ,$$

contradicting (ii). □

This method of proof will also be used in a more intricate way later on in this book.

I.2. Adapted sequences

I.2.1. Definitions

Suppose that $(F_n)_{n \in \mathbb{N}}$ is an increasing sequence of sub-σ-algebras of the σ-algebra F. $(F_n)_{n \in \mathbb{N}}$ is called a stochastic basis. Suppose that $(X_n)_{n \in \mathbb{N}}$ is a sequence of integrable functions. Hence each X_n belongs to $L^1_E(\Omega, F, P)$. If every X_n is F_n-measurable, i.e., is the a.e. limit in E of a sequence of stepfunctions with steps in F_n, we say that

$(X_n, F_n)_{n \in \mathbb{N}}$ is an adapted sequence or a stochastic process. We remark that any sequence $(X_n)_{n \in \mathbb{N}}$ in L^1_E can be considered as an adapted sequence w.r.t. $(F_n)_{n \in \mathbb{N}}$, where

$$F_n = \sigma(X_1, \ldots, X_n)$$

for every $n \in \mathbb{N}$. Here $\sigma(X_1, \ldots, X_n)$ denotes the smallest σ-algebra keeping every $X_1, \ldots, X_n$ measurable.

A function $\tau : \Omega \to \mathbb{N} \cup \{+\infty\}$ is called a stopping time w.r.t. $(F_n)_{n \in \mathbb{N}}$ if for each $n \in \mathbb{N}$, $\{\tau = n\} \in F_n$. Here we denote $\{\tau = n\} = \{\omega \in \Omega \mid \tau(\omega) = n\}$. The set of all stopping times is denoted by T^*. We order T^* as follows : for $\sigma, \tau \in T^*$ we denote $\sigma \leq \tau$ if $\sigma(\omega) \leq \tau(\omega)$ for each $\omega \in \Omega$. The subset of T^* consisting of all bounded stopping times is denoted by T. So $\tau \in T$ if $\tau \in T^*$ and $\tau(\Omega)$ is a finite subset of $\mathbb{N}$. The order induced by T^* on T has the property that $\mathbb{N}$ is cofinal in T. For $\tau \in T$, denote

$$F_\tau = \{A \in F \mid A \cap \{\tau = n\} \in F_n, \text{ for each } n \in \mathbb{N}\}$$

and

$$X_\tau = \sum_{n=\min \tau}^{\max \tau} X_n \chi_{\{\tau = n\}} \quad .$$

This means, for each $\omega \in \Omega$: $(X_\tau)(\omega) = X_{\tau(\omega)}(\omega)$. It is obvious that X_τ is F_τ-measurable and that X_τ is integrable. If $\sup_{\tau \in T} \int_\Omega \|X_\tau\| < \infty$ then $(X_n, F_n)_{n \in \mathbb{N}}$ is said to be of class (B).

Stopping times satisfy a "localization property" : Let $\{\sigma_1, \ldots, \sigma_n\} \subset T$ and $A_i \in F_{\sigma_i}$, for each $i = 1, \ldots, n$ such that $(A_i)_{i=1,\ldots,n}$ is a partition of Ω. Then $\tau \in T$ if $\tau =: \sigma_i$ on A_i. Indeed, for each $k \in \mathbb{N}$:

$$\{\tau = k\} = \bigcup_{i=1}^{n} [A_i \cap \{\sigma_i = k\}]$$

$$\in F_k$$

For the same reason we have that, if σ and $\tau \in T$, then $\{\sigma = \tau\} \in F_\sigma \cap F_\tau$. We denote $F_\infty = \sigma(\bigcup_n F_n)$, the σ-algebra generated by $(F_n)_{n \in \mathbb{N}}$. If $\sigma \in S$, a subset of T, denote $S(\sigma) = \{\tau \in S \| \tau \geq \sigma\}$.

All these notions can be generalized to adapted nets $(X_i, F_i)_{i \in I}$, where I is a directed index set, $F_i \subset F_j \subset F$ if $i \leq j$, $i,j \in I$ and X_i is F_i-measurable; for this, see chapter VI.

I.2.2. Conditional expectations

Let G be a sub-σ-algebra of the σ-algebra F. We state and prove the following existence theorem for conditional expectations of functions in L^1_E.

Theorem I.2.2.1 : There exists a unique map :

$$E^G : L^1_E(\Omega,F,P) \to L^1_E(\Omega,G,P/G)$$

such that for each $A \in G$ and each $f \in L^1_E(\Omega,F,P)$ we have :

$$\int_A f \, dP = \int_A E^G(f) \, dP$$

Furthermore for each $p \in [1,+\infty)$, $E^G \in \mathcal{L}(L^p_E(\Omega,F,P), L^p_E(\Omega,G,P/G))$ and $\|E^G\|_p \leq 1$, where $\mathcal{L}(\cdot,\cdot)$ denotes the space of the continuous linear operators between two Banach spaces and P/G denotes the restriction of P to the sub-σ-algebra G.

Proof : Fix $A \in F$. Define

$$\mu : G \to [0,1]$$

by $\mu(B) = P(A \cap B)$ for $B \in G$. Obviously $\mu \ll P$ on G. So, by the Radon-Nikodym theorem, there exists a function $\varphi_A \in L^1(\Omega,G,P/G)$ such that :

$$\mu(B) = P(A \cap B) = \int_B \varphi_A \, dP$$

for each $B \in G$. Let f be a stepfunction of the form

$$f = \sum_{i=1}^{n} x_i \chi_{A_i}$$

where $x_i \in E$ and $A_i \in F$. Define

$$E^G(f) = \sum_{i=1}^{n} x_i \varphi_{A_i} .$$

Then we see that the conditional expectation property is fullfilled on stepfunctions. Extension to $L^1_E(\Omega,F,P)$ is possible since stepfunctions are dense in this space. The uniqueness of E^G is obvious from theorem I.1.2.4. Linearity is obvious too. So we only have to show that

$$\|E^G\|_p \leq 1$$

for each $p \in [1,+\infty)$. Suppose first that $f \in L^p$ and let $at+b$ be a support line of $\Phi(t) = t^p$. Then

$$aE^G(f) + b = E^G(af+b) \leq E^G(|f|^p)$$

Hence $|E^G f|^p \leq E^G(|f|^p)$, by taking suprema over the support lines. This inequality is called Jensen's inequality and is valid for any convex function Φ. From this :

$$\int_\Omega |E^G(f)|^p \leq \int_\Omega |f|^p \qquad (1)$$

Now let f be an E-valued stepfunction $f = \sum_{i=1}^n x_i \chi_{A_i}$. Then :

$$\|E^G(f)\|_p \leq (\int_\Omega (\sum_{i=1}^n \|x_i\| E^G(\chi_{A_i}))^p \, dP)^{1/p}$$

$$= (\int_\Omega (E^G(\sum_{i=1}^n \|x_i\| \chi_{A_i}))^p \, dP)^{1/p}$$

$$\leq (\int_\Omega \|f\|^p \, dP)^{1/p} = \|f\|_p ,$$

the last inequality by (1). This finishes the proof since stepfunctions are dense in L^p_E. □

$E^G(f)$ is called the conditional expectation of f w.r.t. G.

Example I.2.2.2 : If we start with $(F_n)_{n\in\mathbb{N}}$ as in I.2.1 and if we take $f \in L^1_E(\Omega,F,P)$, the sequence $(E^{F_n}(f),F_n)_{n\in\mathbb{N}}$ is an adapted sequence. In fact it is a martingale, converging to f a.e. and in the $\|\cdot\|_1$ sense as we shall see later in chapter II.

Properties I.2.2.3 : We have the following properties of conditional expectations E^G :

(i) $E^G E^G = E^G$

(ii) $E^G(x) = x$ where $x \in E$

(iii) E^G is linear.

(iv) E^G is a contraction.

It is interesting to note that these properties characterize conditional expectation operators in $\mathcal{L}(L^1_E(\Omega,F,P),\ L^1_E(\Omega,F,P))$ if E has a strictly convex norm, i.e. : the unit sphere $\{x \in E \mid \|x\| = 1\}$ contains no straight lines. For the proof, see Landers and Rogge [1981]. For properties of strict convexity, see Diestel [1975b], from p.23 on.

I.3. Convergence

In this section we shall discuss some types of convergence which will be studied later on in the book. $(X_n,F_n)_{n\in\mathbb{N}}$ denotes an adapted sequence with values in the Banach space E. One might as well take any sequence $(X_n)_{n\in\mathbb{N}}$ in L^1_E, but we study only adapted sequences in this book. On the other hand $(X_n)_{n\in\mathbb{N}}$ is always adapted w.r.t. $(F_n)_{n\in\mathbb{N}}$, where $F_n = \sigma(X_1,\ldots,X_n)$. Relations between different types of convergence will only be stated if they are well known. For comparatively new results we shall provide a proof.

I.3.1. Pointwise convergence

$(X_n,F_n)_{n\in\mathbb{N}}$ is said to be strongly convergent a.e. to a (measurable) function X_∞ if there is a nullset N such that for each $\omega \in \Omega \setminus N$ and each $\varepsilon > 0$, there exists an $n_o \in \mathbb{R}$ such that for each $n \in \mathbb{N}(n_o)$, $\|X_\infty(\omega) - X_n(\omega)\| \leq \varepsilon$. $(X_n,F_n)_{n\in\mathbb{N}}$ is said to be weakly convergent a.e. to a (measurable) function X_∞ if there is a nullset N such that for each $\omega \in \Omega \setminus N$ and each weak-zero-neighbourhood V, there exists an $n_o \in N$ such that for each $n \in \mathbb{N}(n_o)$, $X_\infty(\omega) - X_n(\omega) \in V$. $(X_n,F_n)_{n\in\mathbb{N}}$ is said to be scalarly convergent a.e. to a (not necessarily measurable) function X_∞ if for every $x' \in E'$, $(x'(X_n))_{n\in\mathbb{N}}$ converges a.e. to $x'(X_\infty)$. Note that X_∞ is always scalarly measurable.